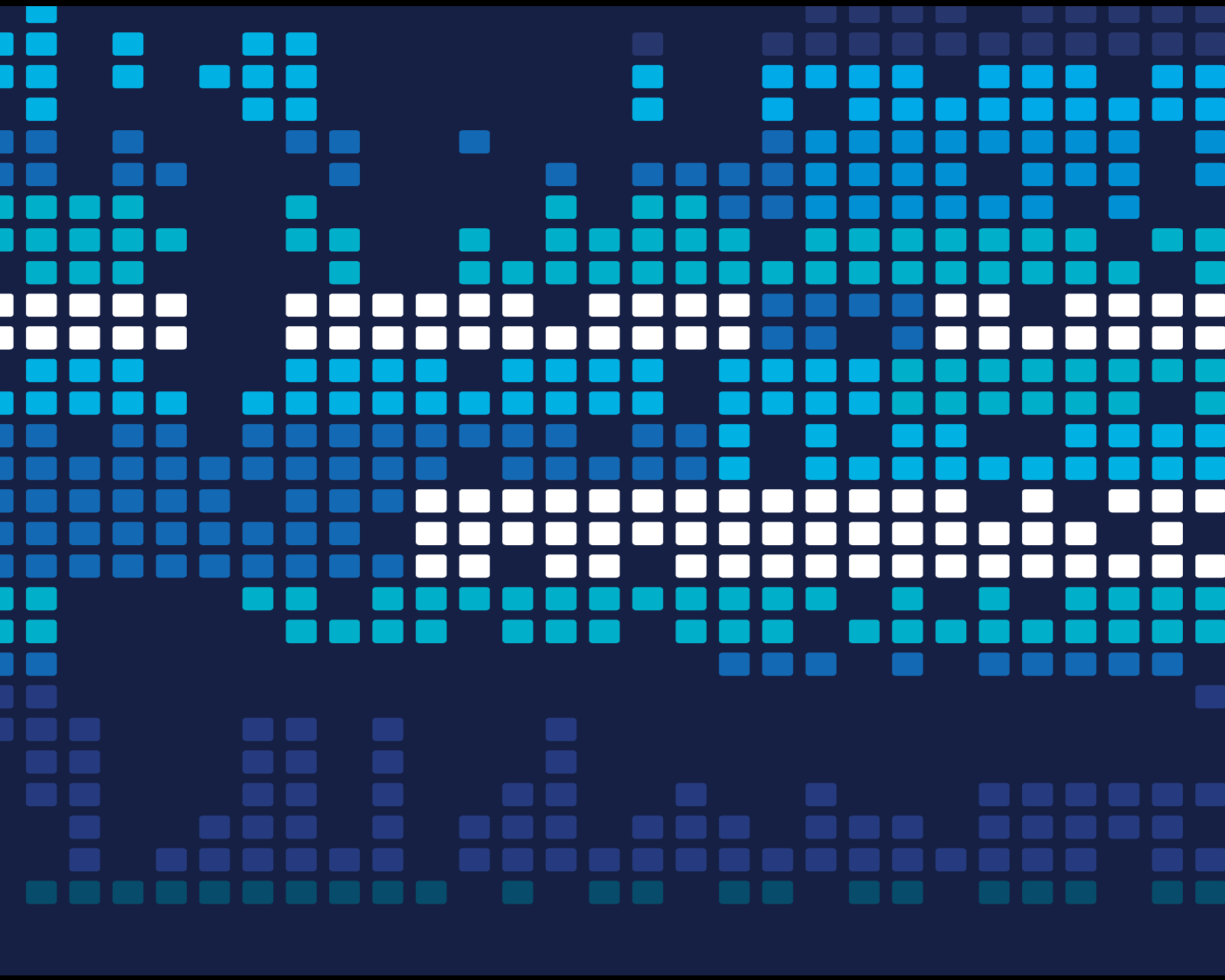
Scientific Programming

Machine Learning and Scientific Programing in Multi-Sensor Data Processing

Lead Guest Editor: Baiyuan Ding
Guest Editors: Gohar Ali and Guobin Chen



Machine Learning and Scientific Programing in Multi-Sensor Data Processing

Scientific Programming

Machine Learning and Scientific Programing in Multi-Sensor Data Processing

Lead Guest Editor: Baiyuan Ding
Guest Editors: Gohar Ali and Guobin Chen

Contents

Contents

Contents

Contents

Hindawi
Scientific Programming
Volume 2023, Article ID 9803492, 1 page
https://doi.org/10.1155/2023/9803492

Retraction

Retracted: Dynamic Early-Warning Model of College Students' Psychological Crisis Based on Characteristic Attribute

Scientific Programming

Received 26 September 2023; Accepted 26 September 2023; Published 27 September 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] X. Chen, "Dynamic Early-Warning Model of College Students' Psychological Crisis Based on Characteristic Attribute," *Scientific Programming*, vol. 2022, Article ID 6026248, 11 pages, 2022.

Hindawi
Scientific Programming
Volume 2023, Article ID 9893506, 1 page
https://doi.org/10.1155/2023/9893506

Retraction

Retracted: LSTM-Based Attentional Embedding for English Machine Translation

Scientific Programming

Received 29 August 2023; Accepted 29 August 2023; Published 30 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] L. Jian, H. Xiang, and G. Le, "LSTM-Based Attentional Embedding for English Machine Translation," *Scientific Programming*, vol. 2022, Article ID 3909726, 8 pages, 2022.

Hindawi
Scientific Programming
Volume 2023, Article ID 9825727, 1 page
https://doi.org/10.1155/2023/9825727

Retraction

Retracted: Construction of the 3D Reconstruction System of Building Construction Scene Based on Deep Learning

Scientific Programming

Received 15 August 2023; Accepted 15 August 2023; Published 16 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Z. Lu, "Construction of the 3D Reconstruction System of Building Construction Scene Based on Deep Learning," *Scientific Programming*, vol. 2021, Article ID 5839391, 9 pages, 2021.

Hindawi
Scientific Programming
Volume 2023, Article ID 9829285, 1 page
https://doi.org/10.1155/2023/9829285

Retraction

Retracted: Classroom Visualization Based on the Relationship between User Roles and Scene Content

Scientific Programming

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Y. Li, D. Yu, K. Yu, and X. Chen, "Classroom Visualization Based on the Relationship between User Roles and Scene Content," *Scientific Programming*, vol. 2022, Article ID 5300840, 10 pages, 2022.

Hindawi
Scientific Programming
Volume 2023, Article ID 9791531, 1 page
https://doi.org/10.1155/2023/9791531

Retraction

Retracted: Design of Packaging Design Evaluation Architecture Based on Deep Learning

Scientific Programming

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] L. Shi, "Design of Packaging Design Evaluation Architecture Based on Deep Learning," *Scientific Programming*, vol. 2022, Article ID 4469495, 8 pages, 2022.

Hindawi
Scientific Programming
Volume 2023, Article ID 9782074, 1 page
https://doi.org/10.1155/2023/9782074

Retraction

Retracted: Remote English Teaching Resource Sharing Based on Internet O2O Model

Scientific Programming

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Z. Hou, "Remote English Teaching Resource Sharing Based on Internet O2O Model," *Scientific Programming*, vol. 2022, Article ID 1217807, 10 pages, 2022.

Hindawi
Scientific Programming
Volume 2022, Article ID 6367755, 8 pages
https://doi.org/10.1155/2022/6367755

Research Article

Design of Literary Theory Teaching System in the Visual Field of Language Graphic Relations

Yuan Zhang[1] **and Xiaolei Zhang**[2]

[1]*College of Humanities and International Education, Xi'an Peihua University, Xi'an, Shaanxi 710065, China*
[2]*Service and Software R&D Management Department, Xi'an Huawei Technologies Co., Ltd., Xi'an, Shaanxi 710065, China*

Correspondence should be addressed to Yuan Zhang; zhangyuan@peihua.edu.cn

Received 31 March 2022; Revised 23 April 2022; Accepted 1 May 2022; Published 23 May 2022

Academic Editor: Baiyuan Ding

Literary theory has developed with the development of the times. In traditional literary theory teaching, we always adhere to the "literature is the art of language" and "literature is a reflection of social life." The essence of literature, the writer, the text, and the interpretation of the content, for example, are established on the basis of the understanding of language, in order to understand the world of literature reflect and express, and the image idea is basically not involved. This paper discusses the teaching of literary theory based on the relationship between language and picture, close to the literary reality to optimize the teaching content, carry out research-based teaching, and cultivate students' thinking ability and practical ability, so as to explore and innovate the teaching of literary theory. For the objective questions or questions close to the objective questions, the machine is used for marking, while the subjective questions need teachers to mark. In addition, in terms of evaluation, qualitative and quantitative methods are used to comprehensively evaluate learners, and a preliminary evaluation table for students is designed.

1. Introduction

In the long history of human aesthetic ideography, the relationship between the language symbol system represented by the image/word and the image symbol system represented by the image/image is very complicated, and their respective functions in literary ideography show different trends due to the difference of their symbolic attributes. Therefore, in the scientific link of literary ideography, the study of the relationship between language and image may analyze the attributes of the two symbols and thus take it as the starting point to systematically reveal the relationship between language and image in literature and art and the theoretical basis of criticism [1, 2].

In recent years, with the deepening and development of the study of literary imagology, the study of the relationship between language and image has been included in the important stage of the current literary research. In fact, the context of this era includes not only literary studies, but also a variety of humanities and social sciences such as philosophy and sociology, all of which are responding to the study of the relationship between language and picture [3, 4]. In the study of this issue, the depth of the art history is relatively wide, based on which the corresponding new art history is established. The depth and breadth of the discussion on the relationship between language and image are very deep, thus providing different research paths for the development of the current literature imagology [5, 6].

As a symbol to express certain ideas, language signifier is a sound image inseparable from ideas. While the sound image is the psychological imprinting of the physical pronunciation of language, the character fixes the sound image into the visual image with sound. Referring to the process of language symbols is more dependent on social historical and cultural impact of the external factors, but most of these factors depend on the human population characteristics and unique temperament. This kind of expression mode also marked the humans use language symbols of group communication and aesthetic expression way using symbols to communicate with other nature species essential differences

[7, 8]. In contrast, the image is not; the image is first of all a physical phenomenon of physiological-psychological refraction. The generation of images is the result of the interaction between subject and object, while the generation of language is the result of pure socialization. In terms of structure and form, image is quite different from language [9]. Image is a symbol system of "mutual appearance of form and meaning." Its core element is "image" and its "simulation" function, and its meaning expression also relies on the intuitive examination of the symbol itself and the perceptual connection between subject and object. The image simulates the form and structure of the natural image through shape, color space, and other elements. At the same time, it also presents the external representation of the concept and image [10, 11].

When literature encounters "the age of reading pictures," the problem that literature faces is not only what language to write and how to write, but the relationship between language and image, which is an urgent realistic problem. Faced with the impact and siege of images, if we still stick to the teaching concept of literature theory like ivory tower, we will not be able to face and answer the fresh and realistic problems faced by literature, which will make the teaching content appear old and empty and fail to stimulate students' interest in learning [12, 13].

Based on this, in the teaching of literary theory, we should introduce the realistic problems faced by literary theory to students as a core concept, guide students to master the basic knowledge and basic laws of literary theory, broaden their horizons, and constantly stimulate their aesthetic interest [14, 15]. In the author's opinion, this core concept is a literary theory view based on "the relationship between literature and image," that is, through exploring and elucidating the complex relationship between literature and image, the coexistence of homology, and the unity of opposites. It focuses on cultivating the critical thinking ability and practical innovation ability of the students of the department of Chinese literature who focus on the reality of literature, inherit the tradition of literary theory, and face the future of literature [16, 17].

Research-based teaching is a teaching paradigm under the guidance of innovative education theory. It makes full use of the advantages of the previous teaching model and makes students master relevant knowledge and improve their ability to solve practical problems by studying major practical problems. Taking the relationship between language and picture as the main line, research-based teaching aims to guide students to use knowledge and ability creatively through classroom teaching process and to accumulate knowledge, cultivate ability, and exercise thinking in the discussion of problems, so as to achieve creative design of teaching objectives. Problem is oriented to achieve the diversity of teaching objectives [18, 19]. Question is the starting point of all research activities. In the past, classroom teaching often took questions as a means to check whether students have mastered the knowledge and played the role of imparting knowledge. Even guiding students to discover and solve problems is based on theoretical presupposition. For example, the previous lectures on "reflection theory," "aesthetic reflection theory," "the relationship between literature and ideology," and other contents take the teaching idea of essence first, around the original Canon or some abstract concepts to sketch the literary prospect, but ignore the fresh literary reality and the dilemma faced by the development. Research-based teaching is based on realistic literary problems, and teachers guide students to choose topics according to their interests. For example, when explaining literary images, we can choose the TV adaptation of a Dream of Red Mansions as a topic for discussion [20, 21]. In reading and watching classic pieces, with supplementary reading on the basis of some representative academic papers, the teacher carries out special background knowledge counselling, around the treasure jade, dai jade, treasure chai image in the novel text, and TV performance in the text of the similarities and differences, as well as the interaction of the language-figure, and inspires the student to think of two different media and have more in-depth understanding of image. In this way, students master important knowledge of literary images in the process of discussion [22, 23]. In the process of carrying out the practice of thematic research thinking, oral expression, analysis, synthesis, and other aspects of the ability to improve, the comprehensive quality has been further improved. Establish a dialogue relationship between teachers and students with two-way interaction as the mechanism. In research-based teaching, the teacher is the guide and the guide to introduce knowledge, while the students are learners and researchers. As the learning interest and knowledge experience of contemporary college students are largely based on network technology and information dissemination platform, they are sensitive and curious about visual culture and fashion culture and can obtain a large amount of information and resources through the Internet to find the problems and solutions. Therefore, students have full power of speech and expression in some teaching content. For example, in narrative thematic discussion, students' creative spirit and research ability can be stimulated by teaching content such as text narration and image narration [24, 25]. Hence, this paper designs a literature theory teaching system that integrates self-service learning and testing mode, curriculum learning and testing mode, and selection learning and testing mode. In self-service learning and testing mode, learners can choose questions in the system to study and test independently according to their own needs.

2. Teaching System Design

In the end of 2019, COVID-19 pandemic broke out around the world, and people's social environments were greatly affected. Due to the epidemic, Chinese primary and secondary schools and institutions of higher learning are unable to provide normal classroom teaching activities as scheduled, making online teaching as the main teaching mode of educational institutions during the epidemic. Therefore, reducing social activities through online teaching is still one

of the main methods to effectively prevent and control the epidemic. In this context, people pay more and more attention to online teaching system, such as Figure 1.

From the point of view of hardware requirements, the online teaching system can be completed only by personal PC during development, and the server requirements are not high after implementation. In terms of software requirements, many of the technologies required for development are open source and free, so the cost of use is not high. From the perspective of the economic value realized after the successful design of the system, this system can greatly save the cost of online teaching, so it is cost-effective; therefore, the system research and development is completely feasible economically. In addition, the framework of the teaching system is based on the currently popular Java development language, which has outstanding advantages of open source, high efficiency, and simplicity. Besides, the technology is relatively mature, and there are many related literatures and cases, which can play a good reference and reference role. Before the project development, relevant technical standards and literature cases have been collected and sorted out, which can basically meet the development and use of the project, so the technical feasibility can be basically guaranteed. In this paper, considering the design and implementation of online teaching system, users are mainly teachers and students with a certain cultural basis, in the operation of the software system with certain ability. At the same time, in the system interface design, this paper will try to consider the user's use habits, on the basis of not affecting the system function and performance, to ensure the simplicity of the system interface and functional convenience, so the system is basically feasible in user operation. After the completion of course learning and testing and selective learning and testing, students' scores shall be statistically analyzed, and the parameters of statistical analysis of students' scores are as follows: calculate students' test scores, calculate class average scores, divide score segments, count the number of students in each score segment and calculate percentage, count the number of excellent students and excellent rate, and calculate the pass number, pass rate, median, mode, region, standard deviation, etc. These indicators can be used to reflect the current learning status of learners more vividly in the form of charts.

Here are some important indicators:

(1) Average value reflects the overall level achieved by students of this class in the exam, and this value can be used as a reference for comparison of learning levels of different classes.

The calculation formula is to let N test scores be

$$x_1, x_2, \ldots, x_N. \tag{1}$$

The average score is

$$\overline{x} = \frac{1}{N}\left(x_1 + x_2 + \ldots + x_N\right). \tag{2}$$

The above equation can also be written as

Figure 1: Online teaching system.

$$\overline{x} = \frac{1}{N}\sum_{i=1}^{N} x_i. \tag{3}$$

(2) Median refers to the score in the middle of the test score ranking, which can be used as the representative value of the test score, reflecting the general level.

(3) Mode is the score that appears most frequently in the test results. It reflects a central tendency in test scores.

(4) Region: the difference between the highest and lowest scores reflects the biggest gap in learning outcomes of all the current test takers.

(5) Standard deviation: standard deviation reflects the overall distribution of test scores, or the degree of dispersion of test scores relative to the mean. The larger the standard deviation is, the greater the difference between good and bad grades is, that is, the more imbalanced the distribution among students is; thus it can be seen that the distance between learners is larger. In addition, standard deviation can also be used to compare students' two test scores. Since students have different test papers, standard deviation can be used as the basis to convert the two test scores into standard scores to judge whether students have made progress in the learning process.

The standard deviation is calculated by the following formula:

$$S_x = \sqrt{1/N \sum \left(x_i - \overline{x}\right)^2}. \tag{4}$$

The difficulty of the questions can be expressed by the correct answer rate of the students to the questions. The calculation formula is as follows:

$$p_i = \frac{x_i}{S_i}, \tag{5}$$

where p_i is the difficulty of problem i; x_i is the average of the correct answers to question i; S_i is the total score for problem

i. We consider a parameter that measures the ability of students at different levels to grasp knowledge. Students with good grades get high marks, and students with poor grades get low marks. The calculation formula is

$$D_i = \frac{H_i - L_i}{S_i}, \tag{6}$$

where H_i is the average score of the correct answer to question i in the high group; L_i is the average score of the correct answer to question i in the lower group; S_i is the total score of question i; D_i is the distinction on problem i. Reliability indicates the reliability of the test. The higher the accuracy of the test, the higher the reliability of the test. It reflects the stability and reliability of the test. The calculation formula is

$$T = \frac{N}{N-1}\left[1 - \frac{\sum_i^N \partial_i^2}{\partial^2}\right], \tag{7}$$

where N is the total number of problems. Validity is a parameter used to measure the validity and correctness of a test. The calculation formula is

$$E = \frac{1}{N}\sum_i^N D_i, \tag{8}$$

where N is the total number of questions. D_i is the distinction of problem i; E is the validity of the test paper. The difficulty of the paper can be reflected by the average difficulty of the paper. The calculation formula is

$$\begin{aligned} P &= \frac{1}{N}\sum_i^N P_i, \\ \sum_{i=1}^K A[i] * B[i] &= G, \\ C[i,j] - \partial &\le D \le C[i,j] + \partial. \end{aligned} \tag{9}$$

The classical measurement theory is based on the true fraction theory and is built on the linear relationship between real score, true score, and error score. Its mathematical model is as follows:

$$X = T + E, \tag{10}$$

where X is the real score, T is the true score, and E is the error score.

3. System Framework

In order to ensure the upload and play of course media under the condition of high concurrency of big data, the teaching management module also covers the Hadoop framework based on cloud computing distributed system and file reading, as well as the optimization of load balancing algorithm. The module consists of three submodules: online teaching, video uploading, and video maintenance. Online teaching, including course information, is carried out by students through the online teaching interface, so that the online teaching function depends on the entity of course information. Video uploading by teacher through the video uploading interface also includes course information, so that the video uploading function also depends on the course information entity. Video maintenance performed by administrators on the video maintenance interface also includes course information, so that the video maintenance function also depends on course information entities. According to the dependence of the teaching management module function on the data entity, the attributes of the course information entity can be analyzed based on the design of the course information in the database.

The functional module of data management is composed of three submodules: student's homework, teacher's homework, and job maintenance. The distributed system of cloud computing based on Hadoop framework and file reading and writing method proposed in this paper are also introduced, as well as optimized load balancing algorithm. Student user can operate student's homework through student's homework interface, including homework information, so student's homework function depends on homework information entity. The operation of teacher's homework by teacher user through teacher's homework interface also includes homework information, so the function of teacher's homework also depends on homework information entity. Job maintenance operations performed by administrators on the job maintenance interface also include job information, so the job management function also relies on job information entities. According to the dependency of the job management module on the data entity, the attributes of the job information entity can be analyzed as the basis for the design of the job information data table of the database.

The system adopts the autonomous learning and test module, and the so-called self-help is in the specific system environment; according to the demand of oneself independently accomplishing a task, in this module the learner can choose according to their own needs the corresponding questions to study and test, and the conditions enable the learners to choose a course content condition, question topic condition, and item difficulty. That is, these three groups of paper conditions form the test questions that learners need to learn and test. It is not enough just to provide the reference answers for learners, but more importantly, to help learners understand the real problems of the test questions.

In this module, to help learners learn and test, we provide the following functions. (1) Extract test questions: according to the value that the learner fills in the selection condition, the corresponding test questions are extracted from the question extraction algorithm. (2) Test timing: display the test time and remaining time used by learners, and specify the test within the specified time range. If the time exceeds the specified time range, the system will automatically save the results of students' answers. (3) Test submission: the results of individual answers are stored in the system for the convenience of marking and evaluation. Answer the idea. Students in the process of test learning will inevitably encounter some problems without ideas to solve the problem and then provide students with some ideas to

solve the problem, which is conducive to students to save time and improve efficiency. (4) Reference answer: provide reference answers to test questions for students to check their answers.

When learners do not have solutions to the questions, it provides solutions for learners and provides motivation for learners to further answer the questions. To analyze the reference answer, eliminate the doubts in the hearts of students, so that students have a more comprehensive understanding of the connection between knowledge, knowledge mastery, and mastery of knowledge.

The system also contains course learning and testing modules. Curriculum learning and testing is a set of learning and testing mode, which is the network transfer of traditional testing, and based on the paper formation strategy it is the traditional test paper formation strategy (a two-way detail table); this module is mainly to help learners to check the stage of learning results and deepen the understanding and mastery of knowledge. In this module, learners can not only study and test the papers composed by teachers, but also study and test the real questions of midterm technical exams in the past years.

In this module, we provide the following subfunctions: (1) Extract test questions: this function is generally taken charge of by teachers, who can correct questions that do not meet the test requirements by extracting questions from the test questions resource base according to the two-way list. (2) Test timing displays the time questions used by learners for tests and submits student tests for responses within the prescribed time. (3) Test submission: learners can submit their own test answers within the specified time, or the system can automatically save learners' test answers when they arrive at the specified time, for later marking and evaluation. Answer the idea: in the study of test questions, when learners have no clue in the face of the topic, appropriately giving learners hints can stimulate students' interest in learning, expand their learning thinking, and improve their learning efficiency. (4) Reference answer: provide reference answers for students to judge their own answers. (5) Personal answer: show the results of the answers so that learners can compare their personal answers with the reference answers.

Help students understand the misunderstandings in the answer, enrich the connection between knowledge points, and integrate mastery and application practice. Provide reference answers for the test questions done by learners, so that learners can check their answers with reference answers to find out their own shortcomings and deepen their understanding of knowledge.

The system also includes optional learning and testing modules. Selected learning and testing is a kind of comprehensive learning and test mode; the model is based on the selected test and can be fixed mode test (graduate) selected learning and testing, in a fixed mode of learning and testing, we provide two ways of learning and test papers: one aspect is the fixed model of examination paper composed by teachers according to selective learning and test paper group strategy. Another aspect is the study and test of selective real questions over the years. The purpose of this module is to help learners deepen their understanding of knowledge and then get good results in the selection test and have an advantage in the selection competition.

4. Simulation Experiment Analysis

"Literary theory guidance" is the basic theory of Chinese language and literature professional courses; through the study of this course, learners can gain a systematic knowledge of literary theory, to deepen students' understanding and the understanding of literature and to help students establish a scientific concept of literature, and its importance is beyond doubt. In order to ensure the quality of learning, it is necessary to develop a learning and evaluation system platform for learners. In view of this, we seek to enable the platform to provide multilevel networked testing, learning and evaluation services for learners, to meet the needs of learners in all stages of learning and evaluation.

Because this system is based on "literature theory guidance," this course is an example of design and development, so in the survey object, we mainly include university Chinese language and literature students; in the research level, for both undergraduate students and also graduate students, the research object has a certain universality and representativeness. The survey results can reflect some of the needs of learners for the system.

Figure 2 shows the basic situation of the online class users: male students account for 32.4% and female students account for 67.6% (see Figure 2). Han nationality accounted for 87.2%, and ethnic minorities accounted for 12.8% (see Figure 2). Among the students surveyed, 62.5% were undergraduates and 37.5% were postgraduates (see Figure 3).

At the same time, this teaching system has done an online survey on the distribution of the test questions of the course "Literature theory Guidance" to the question "Have you taken an online adaptive test?" The net answer of the question: 18.3% of the students chose to participate in the adaptive examination, while 81.7% of the students chose not to participate in the adaptive examination. It can be seen that the vast majority of students did not participate in the white adaptive examination.

For the students who have participated in the above, according to the student experience, and the existing adaptive examination system of the deficiencies of the investigation: 34.6% of the students think that most of the examination system does not provide learning function, 42.3% of the students think that the existing adaptive examination system needs to be improved in terms of adaptation, and 23% of the students think that the interface of the existing examination system is too monotonous. 35.8 percent of students believe that the existing adaptive testing system does not provide good communication skills. Show learners' answers in the test, so that learners can judge and understand the answers according to the reference answers.

To the question "What questions would you prefer to have in the test bank?," the answers to questions were the following: 35.2 percent of the students chose multiple-choice, 29.4 percent blank, 39.7 percent noun explanation, 45.3 percent discrimination, 43.6 percent analysis, 41.7

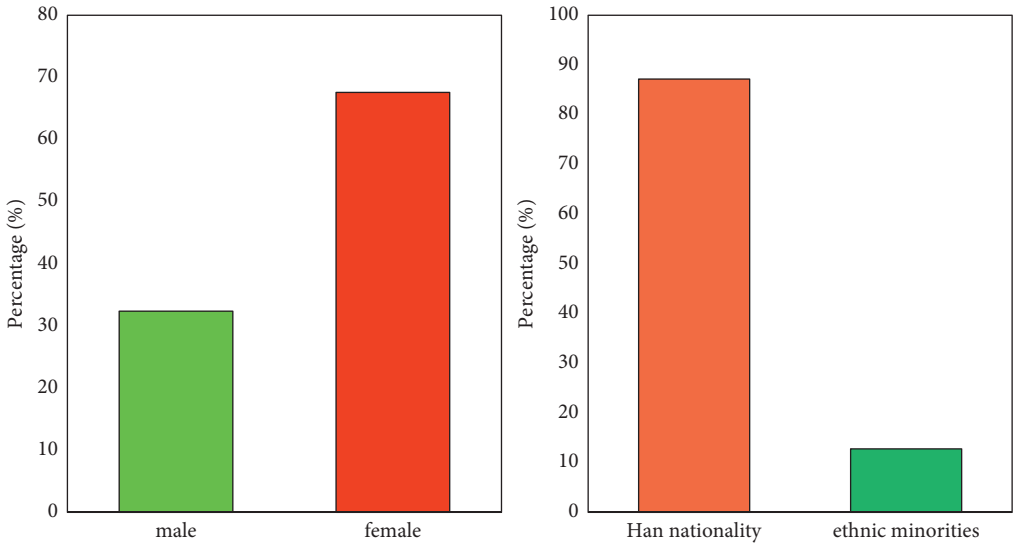


Figure 2: Gender and nationality comparison.

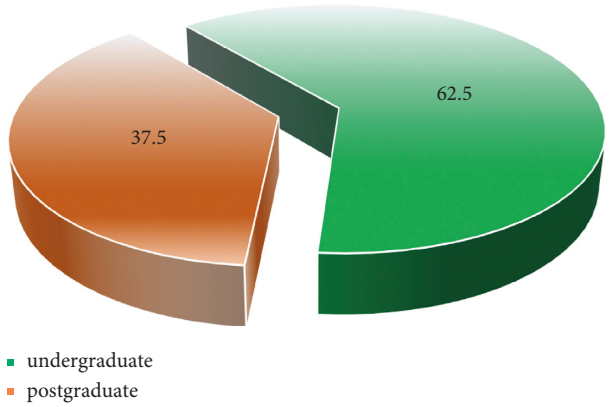


Figure 3: Grade comparison.

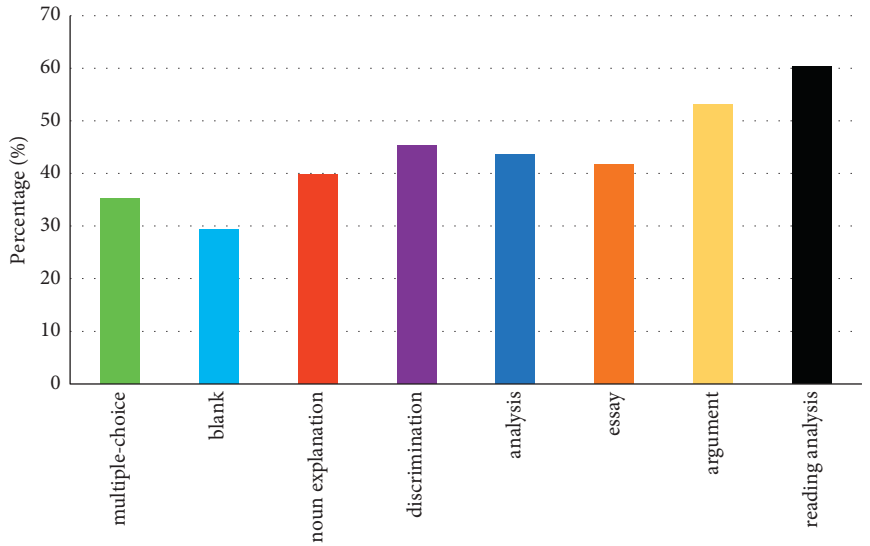


Figure 4: Testing types.

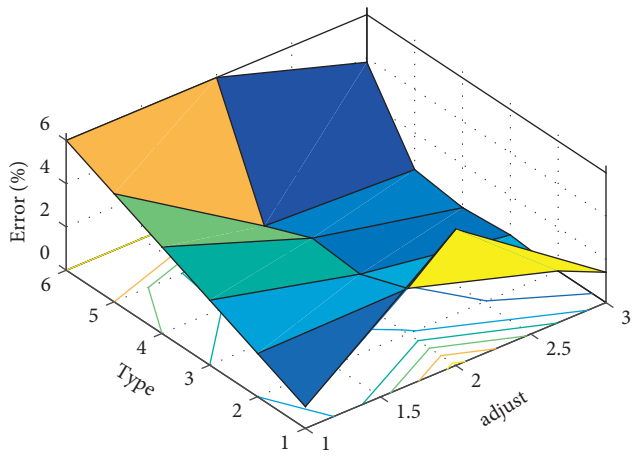


Figure 5: Error comparison.

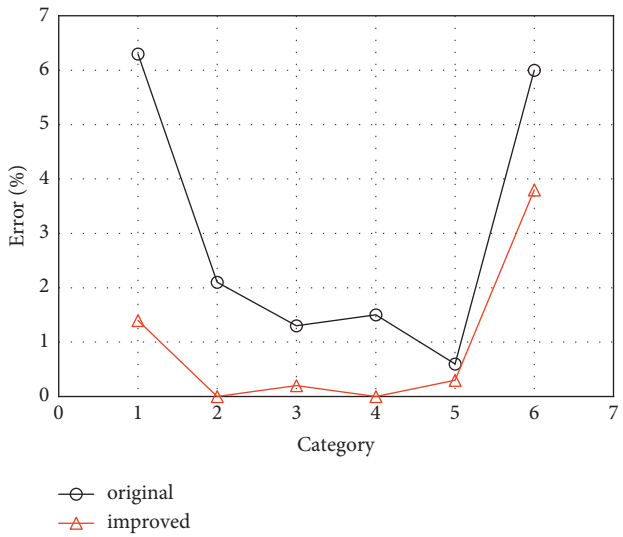


Figure 6: Error of the original and improved model.

percent essay, 41.7 percent essay, and 53.1 percent argument; reading analysis accounted for 60.4%; and other types accounted for 24% (see Figure 4).

In addition, the system studied was optimized in this paper. As can be seen from Figure 5, compared with the model before adjustment, the error rate of type 1, type 2, type 3, type 4, type 5, and type 6 in the training dataset decreased by 0.049, 0.021, 0.011, 0.015, 0.003, and 0.022, respectively. The error rate of category 1, category 2, category 4, and category 6 in the test dataset decreased by 0.025, 0.025, 0.025, and 0.033, respectively, thus achieving the purpose of model optimization. The scale of data processed in this paper is large, and the evaluation of the model is carried out from the following aspects.

Stability is mainly reflected by the degree of difference between the results of training samples and test samples. The error rate difference between training samples and test samples is small, so it is relatively stable. The greater the difference in error rate between training samples and test samples is, the more unstable the system will be.

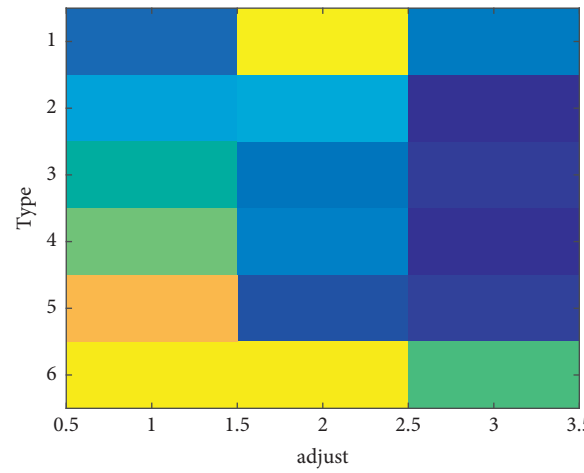


Figure 7: Predicted results of the model.

Accuracy is mainly reflected by the total error rate. The higher the total error rate, the lower the accuracy. Conversely, the lower the total error rate, the higher the accuracy. Provide reference answers for the test questions done by learners, so that learners can check their answers with reference answers to find out their own shortcomings and deepen their understanding of knowledge.

For the model in this paper, in terms of stability, the total error rate of the training sample is 175, while that of the test sample is 238, which is within the controllable range. The model is generally stable. In terms of accuracy, the error rate of training samples and test samples is less than 10%, indicating relatively high accuracy. This can also be seen from Figures 6 and 7.

5. Conclusion

At present, the teaching of literary theory must face the relationship between image and language and actively solve this important realistic problem, so as to explore the innovative way of literary theory teaching. However, this does not mean that some basic problems of literary theory are cancelled and the advantages of traditional teaching mode are ignored. On the contrary, we should combine the advantages of traditional literature theory teaching, so that the literature theory teaching based on the relationship between language and picture and the main line should return to the realistic problems faced by literature, inspire and encourage students to think about the realistic problems, and cultivate their innovative consciousness and ability.

However, due to the limitation of time and development experience in the system functional design, this paper only develops and realizes the basic functions of online teaching, without considering short-term commercial application. Therefore, in the future research and work, we can optimize and expand the functionality and practicability of the system according to the actual application requirements of the industry and fully combine the current cutting-edge technologies such as Internet of Things teaching and artificial

intelligence, so as to realize a more intelligent online teaching system.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

[1] R. Duschinsky, D. Jacobvitz, L. Peake, and S. Messina, ""An extraordinarily pernicious influence": the discursive figure of the spoiling grandmother before 1937," *Journal of Family History*, vol. 45, no. 2, pp. 158–171, 2020.

[2] B. Spicer, ""Translating women: breaking borders and building bridges in the English-language book industry" IMLR, 31 october-1 november 2019," *Journal of Romance Studies*, vol. 20, no. 1, pp. 189–191, 2020.

[3] C. Shan, "A study on discourse analysis and the teaching of literature," *International Journal of Education and Teaching Research*, vol. 2, no. 1, pp. 21–28, 2021.

[4] J. Li, "The cultivation of cultural confidence of college students in the new era in Chinese American literature teaching," in *Proceedings of the 2nd International Conference on Education and Education, Sport and Psychological (ESPS 2021)*, pp. 257–260, Xi'an, China, March 2021.

[5] T. Cheng and M. Deng, "Research on introduction to famous works and cultivation of humanistic quality in the teaching of British and American literature based on network information," *Journal of Physics: Conference Series*, vol. 1744, no. 3, Article ID 032240, 2021.

[6] Ľ. Simanová, A. Sujová, and P. Gejdoš, "Improving the performance and quality of processes by applying and implementing six sigma methodology in furniture manufacturing process," *Drvna Industrija: Znanstveni časopis za pitanja drvne tehnologije*, vol. 1, pp. 876–889, 2019.

[7] P. Vendy Eko, B. Benoit, and O. Barbara, "A proposed method and its development for wood recovery assessment in the furniture manufacturing process," *Bioresources*, vol. 13, no. 2, pp. 214–223, 2018.

[8] S. Ruday and K. Caprino, *Inquiry-Based Literature Instruction in the 6-12 Classroom: A Hands-On Guide for Deeper Learning*, Taylor & Francis, Milton Park, UK, 2020.

[9] X. Huang, "A study on the strategies of introducing film resources into foreign literature teaching," *International Journal of Education and Teaching Research*, vol. 1, no. 4, pp. 213–222, 2020.

[10] B. Cheng and Y. Yuan, "A study on corpus stylistics assisted English and American literature teaching," in *Proceedings of the 2020 Northeast Asia International Symposium on Linguistics, Literature and Teaching(2020 NALLTS)*, pp. 727–731, Dalian, China, October 2020.

[11] S. Kumar, A. K. Kar, and P. V. Ilavarasan, "Applications of text mining in services management: a systematic literature review," *International Journal of Information Management Data Insights*, vol. 1, no. 1, Article ID 100008, 2021.

[12] L. Yu, "A brief analysis of ways to cultivate students' language ability and comprehensive quality in Spanish literature teaching," *International Journal of Educational Technology*, vol. 1, no. 3, pp. 1–7, 2020.

[13] Y. Sun, "The effective application of situational teaching method in English language and literature teaching," *International Journal of Social Sciences in Universities*, vol. 3, no. 3, pp. 54–59, 2020.

[14] A. L. Alfeo, G. C. A. C. Mario, and G. Vaglini, "Measuring physical activity of older adults via smartwatch and stigmergic receptive fields," *CoRR*, vol. 1, pp. 1–9, 2019.

[15] N. A. Callistus, Q. Christopher, S. Jamie, M. Stephen, and R. Payne Terry, "Wearable smartwatch technology to monitor symptoms in advanced illness," *BMJ Supportive & Palliative Care*, vol. 8, no. 2, p. 237, 2018.

[16] X. Li, "Problems and countermeasures of Chinese language and literature teaching in higher vocational education," *Advances in Higher Education*, vol. 4, no. 7, pp. 1–10, 2020.

[17] M. Fuad, A. Efendi, and U. A. Muhammad, "The use of pepaccur local wisdom for Indonesian literary teaching materials," *JPI (Jurnal Pendidikan Indonesia)*, vol. 9, no. 2, p. 213, 2020.

[18] S. Asaithambi, S. Venkatraman, and R. Venkatraman, "Big data and personalisation for non-intrusive smart home automation," *Big Data and Cognitive Computing*, vol. 5, no. 1, p. 6, 2021.

[19] T. Yun, "Exploration of multimodal English and American literature teaching based on computer network," *Journal of Physics: Conference Series*, vol. 1533, no. 2, pp. 543–552, 2020.

[20] X. Li, "English and American literature teaching in colleges and universities in the new media age," *Higher Education of Social Science*, vol. 18, no. 1, pp. 41–44, 2020.

[21] S. Morgan, "Why are we reading this? Hermeneutic inquiry into the practice of teaching (with) literature," *Educational Studies*, vol. 56, no. 2, pp. 145–159, 2020.

[22] D. Radhika and D. Aruna Kumari, "The smart triad: big data analytics, cloud computing and Internet of Things to shape the smart home, smart city, smart business & smart country," *International Journal of Recent Technology and Engineering*, vol. 8, pp. 3594–3600, 2019.

[23] S. Rintoul, "Kevin binfield and william J. Christmas, eds. Teaching laboring-class British literature of the eighteenth and nineteenth centuries. Options for teaching 43. New york: modern language association of America, 2018. Pp. 348. $45.00 (cloth)," *Journal of British Studies*, vol. 59, no. 1, pp. 193–195, 2020.

[24] N. Ma and J. Zhang, "English and American children's literature teaching and the cultivation of English normal students' core accomplishment," in *Proceedings of the 2020 International Conference on Education, Sport and Psychological Studies(ESPS 2020)*, pp. 143–146, Xi'an, China, January 2020.

[25] C. Mai, "A study on the current situation and countermeasures of college English translation teaching," *International Journal of New Developments in Education*, vol. 2, no. 4, pp. 1–7, 2020.

Hindawi
Scientific Programming
Volume 2022, Article ID 5090023, 9 pages
https://doi.org/10.1155/2022/5090023

Research Article

Analysis on the Three-Dimensional Intervention Mode of Public Art in Rural Culture from the Perspective of 3D Video

Xiang Li and Hang Liu

Art College of Southwest, Minzu University, Chengdu, Sichuan 610225, China

Correspondence should be addressed to Hang Liu; 21300175@swun.edu.cn

Received 1 April 2022; Accepted 25 April 2022; Published 16 May 2022

Academic Editor: Baiyuan Ding

With the development of 3D technology, based on the perspective of 3D video, this paper analyzes the principle of 3D video, establishes the perspective model of 3D video, and discusses the modeling process in detail. Then, from the perspective of 3D video, this paper studies the three-dimensional intervention model of public art in Beijing, Shanghai, Chongqing, and Tianjin. Through the combination of public art and rural culture, the countryside can obtain vitality through art and popularize and promote rural culture and image. In general, this paper provides some ideas and experience for the research of public art in rural culture from the perspective of 3D video.

1. Introduction

Public art is also known as social art. It does not belong to a certain kind of art genre or art style, nor does it only refer to a certain kind of art form. It refers to the art created by any media in public space, which is usually located outdoors and open to the public and meets the three characteristics of publicity, popularity, and artistry at the same time [1]. In terms of form, public art includes sculptures, murals, gardening, decoration, landscape, and public spaces (squares, buildings, roads, urban roads, etc.). Meanwhile, the public art also includes some artistic activities such as celebrations and public performances. These public arts existing in public spaces serve the public and embody the spirit and value of cultural openness, sharing, and exchange in public spaces [2]. Public art refers to works or designs created by artists for a given special public space. Based on this, we understand that public art is a very broad artistic abstract concept, and there is no specific definition of art types. But what is certain is that public art exists in society or public space and has an impact on the broad public and society, which is also an artificially created behavior or art [3, 4].

Generally speaking, in the 1930s, public art emerged in the United States during the great depression. At that time, with the help of the new deal implemented by the government, public art entered the social public space and people's daily life space one after another [5]. It was intended to alleviate people's pain caused by the economic crisis and cultivate the pride of American culture through public art. After that, the function of public art has gradually developed from simplification to diversification. Public artworks in cities no longer only have aesthetic functions but also have new functions such as activating space, promoting interaction, and harmonious society [6].

In China, the public's awareness of public art is not enough and even mistakenly believes that it is an independent art type or style. Of course, there are not many behaviors with clear public art characteristics in the real sense in China [7]. In terms of regions, public art appeared earlier in economically developed areas, earlier and more in cities than in villages, and later and less in villages because of the large development gap between them and cities.

The development of rural public art is mainly concentrated in rural residential spaces, individual scenic spots, scenic spots, and other spatial places after entering the twenty-first century [8, 9]. Another way of presentation is in pure artistic forms, such as the current various artist villages and rural art seasons. All kinds of artists concentrate on rural

areas for creation, and most of the works of art are displayed outdoors. Of course, the scale is usually larger than indoor, is more intuitive, and has strong interaction with culture, nature, and environment [10].

In terms of space design, since the twenty-first century, a series of architectural art villages have emerged in China [11]. Most of these activities are organized by social forces, and some well-known architect teams at home and abroad go to a fixed rural area for free design, creation, and construction, with the nature of the architectural experiment. These architectural experiments in the countryside have gone beyond the scope of architecture and have formed a strong touch and guidance to the rural architectural culture and traditional culture, especially enlightening the generally chaotic and low architectural cognition in the countryside [12]. The most direct effect is that the surrounding area of the experimental area benefits directly. Due to the social effect of architects, the society pays attention to the area, which has greatly improved its popularity and tourism attraction. The sudden increase in the flow of people has directly increased the local development opportunities.

These experimental behaviors of space design in rural areas are more of the nature of public art. They are more professional in space design, interact more with the public and society, and use professional exploration to greatly affect society and their professional fields. Since they are experimental, they have more reliance on art so that more people pay attention to architectural art and feel the charm of architectural art [13, 14].

With the promotion of the National Rural Revitalization Strategy, new rural construction, and rural poverty alleviation, rural public art practice projects have developed rapidly in China, but there are also many problems with the booming development trend [15, 16].

1.1. Ignoring the Subjective Status of Villagers, It Is Difficult to Realize the Effective Upgrading of Industry. Many artists lack research on the actual situation in rural areas [17]. They ignore the subjectivity of villagers and the living state and spiritual world of local residents. They imagine the rural environment as an ideal Utopia or simply take it as another display space and simply move the works originally placed in the art museum to the rural environment so that the works are incompatible with the rural environment and cannot resonate between villagers and viewers. Therefore, it is unable to achieve industrial upgrading.

1.2. Without Long-Term Planning, It Is Difficult to Develop Sustainably. Some art practice projects only take the villages where the art practice projects are carried out as a one-time location for art practice [18]. They only pay attention to their immediate interests without long-term planning, so it is difficult to develop sustainably. When the art activities ended, everything in the countryside returned to its original state and even suffered some damage. For example, since 2014, Phoenix Ancient Village and the China Academy of Art have established Phoenix Art Exhibition Center and sculpture, painting, and various literary and artistic activities that have gradually been involved in the cultural construction of the village. Although these public art projects actively mobilize villagers to participate, because most of the residents of the village are immigrants with low cultural identity and low participation, many practical projects take artists as the main body. There is no diversified linkage mechanism for sustainability with the local natural environment and residents' lives, and the follow-up is weak.

1.3. Lack of Regional Characteristics and the Loss of Rural Traditional Cultural Spirit. Using public art to carry forward rural local culture, restore traditional folk customs, and rebuild the declining public space in rural areas can effectively arouse the villagers' sense of pride and cultural identity [19]. However, in many contemporary art rural construction activities in China, the loss of rural traditional culture is serious. The reasons are as follows: first, the historical resources such as rural traditional buildings and intangible cultural heritage have been continuously destroyed, and the traditional regional cultural resources and spirit have gradually weakened and even disappeared, which increases the difficulty of art rural construction; second, the construction mode of urban assimilation also affects the development of rural areas, and the phenomenon of rural homogenization is obvious; and third, many artists have forcibly implanted some concepts that have nothing to do with the local cultural environment into the construction of rural culture. To some extent, such public art is not only the rescue and revival of rural culture but also the invasion of rural resources.

1.4. Lack of Linkage and Cooperation Mechanism, and the Short Output Income of Rural Cultural Industry. The involvement of public art in the rural cultural industry can produce different paths according to different rural types [20]. For example, villages with original ecological culture and ancient buildings can develop cultural tourism, tourist souvenirs, experience halls of intangible cultural heritage, and other industries. Rural areas with superior natural conditions and beautiful environment can develop homestay, ecological agriculture experiences, and other industries. Suburban villages use public art projects to package and develop the leisure tourism industry. However, all kinds of villages still face many problems in industrial output. First, the income of public art combined with rural cultural industries is short, and these industries are idle after art festivals and related art activities. Second, excessive industrial packaging has affected the lives of indigenous people. Many migrant people have become the vested interests of the rural cultural industry, while indigenous people have become the marginal groups of rural industrial economy. For example, the overcommercialized packaging of some rural ancient buildings is incompatible with local culture and breaks the daily life of aborigines. Third, the development mode of rural cultural industry is single and lacks the linkage and cooperation mechanism of multiresources, so it is difficult to achieve long-term benefits.

Therefore, based on 3D technology, Kharroubi et al. [21] studied the impact of public art on rural culture, promoted the development of rural areas while carrying forward and disseminating national excellent traditional culture, and laid a practical foundation for the development of rural public art. Florian et al. [22] studied the relationship between public art and rural culture from the perspective of time-space correlation depth video, revived the countryside by means of public art intervening in rural culture, obtained certain value, and provided some ideas for the later practice of "art intervening in the countryside." Zhou et al. [23] proposed a perspective algorithm for converting 2D video into 3D video, studied the impact of public art on rural development caused by the intervention of rural culture, met the mutual communication methods among villagers, promoted the linkage development of public art and rural cultural industry, and realized the revival and independent and sustainable development of new rural culture. Nico et al. [24] focused on the algorithm structure of a 3D video convolution network, established the relationship between public art and rural culture, and predicted the enforceability of public art in rural revitalization according to China's current national conditions.

In general, based on the perspective of 3D video, this paper first studies the principle of 3D video, then discusses the pattern analysis algorithm of 3D video in detail, and establishes the corresponding framework. Finally, from the perspective of 3D video, this paper analyzes the development of public art in rural culture in various villages of Beijing, Shanghai, Chongqing, and Tianjin; makes a comparative analysis; and puts forward corresponding countermeasures for the development of rural culture, so as to provide a new idea for public art to intervene in rural culture.

2. Principle of 3D Video

Human behavior recognition in video has important application prospects in intelligent monitoring, smart home, and other environments. However, due to the influence of camera perspective, the research of human behavior recognition based on multiview video is highly challenging. In the field of behavior recognition, according to different video types, the mainstream methods can be divided into two categories: behavior recognition based on 3D video and behavior recognition based on depth video.

Although there are many hardware devices with 3D functions at present, there has always been the problem of insufficient film sources for 3D materials that can be used for display, which greatly restricts the promotion of 3D technology. The existing 2D video resources have a wide variety of simple acquisition methods, and 2D technology is relatively mature. Converting these existing 2D videos into 3D videos with three-dimensional effects is a good way to solve the problem of insufficient film sources. The process from ordinary 2D video to 3D video mainly includes two key steps [25, 26]: (1) recover the lost depth information from the input 2D video and generate the corresponding depth map and (2) synthesize virtual stereo image pairs.

At present, there are two mainstream 3D content generation methods: one is that researchers use computer imaging tools and synchronous cameras to generate 3D images or videos, but this method is not only time-consuming but also expensive; The second is to generate 3D content by using the depth map. This method can be divided into the manual method, semiautomatic method, and fully automatic method. The manual method is to manually allocate the depth value for the image/video [4]. The 3D content generated by this method is of the highest quality, but it consumes much more energy and financial resources than other methods. The semiautomatic method is to conduct 2D-3D conversion by manually intervening with the computer. This method is more efficient than the manual method, but it also requires additional labor and time overhead. The fully automatic method can generate 3D content with little or no manual intervention, and the whole process is in the charge of the computer, which not only improves the speed but also saves a lot of resources.

In order to better apply 2D-3D conversion technology to daily life, researchers have invested a lot of energy in the fully automatic conversion method based on the depth map and finally use significance detection for 2D-3D conversion [27]. The task of saliency detection is to detect the most attractive target/region in an image. This mechanism can provide great help for the processing task based on visual information. Therefore, we can use the properties and characteristics of saliency detection to convert the saliency map into another depth map for 2D-3D conversion so that the objects with high saliency are closer to the observer and the objects that are not interesting are far away from the observer. When converting 2D video to 3D video, we should standardize the image and then improve the video signal through the significance principle, so as to obtain a 3D video image through different electromagnetic waves. The 2D-3D conversion process is shown in Figure 1. The main focus is to generate the saliency map conducive to 3D video generation and then calculate the best 3D simulation value through parallax.

3. Model Building from 3D Video Perspective

The model from the perspective of 3D video is mainly based on a convolution neural network, which is composed of a set of convolution and pooling layers with nonlinear activation functions. Although the pooling layer can effectively increase the receptive field, for pixel-level prediction tasks, using the pooling layer for downsampling will cause the loss of detail information and even lead to the failure of reconstruction of objects with small targets. Recently, researchers use hole convolution as a substitute for feature extraction. This method is not easy to cause loss of detail when calculating features at any scale. The reason is that when using hole convolution for calculation, the receptive field can be increased without a pooling layer so that each convolution output contains a large range of information. Figure 2 is a schematic diagram of target tracking from a 3D video perspective. It can be seen that in the perspective of a 3D video, target tracking is divided into keyframe collection, preprocessing, extracting the moving area, obtaining the

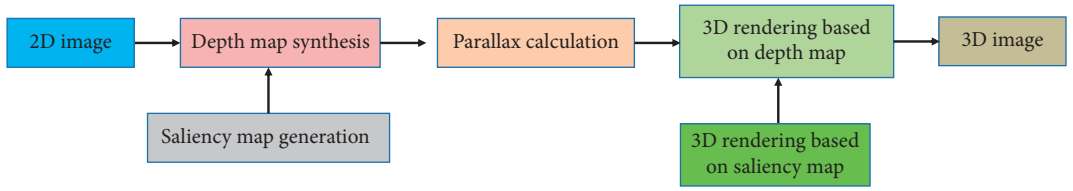


Figure 1: Process of 2D-3D conversion.

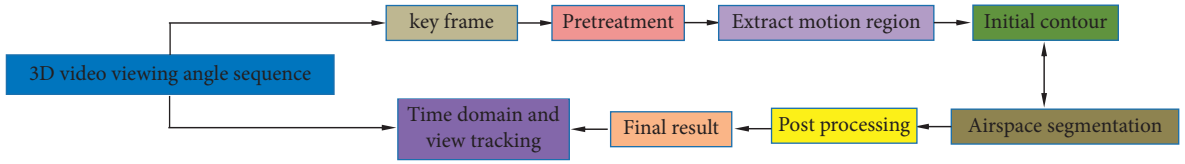


Figure 2: Schematic diagram of target tracking from a 3D video perspective.

initial contour, and then spatial segmentation. After post-processing, the final result can be obtained, so as to realize the comprehensive tracking analysis of the time domain and the field of view.

Multiscale feature extraction is a common method to improve the accuracy of 3D video models. This method extracts spatial features of different scales by using hole convolution kernels with different hole rates; the hole convolution kernels reflect the animation definition in 3D video. Through the hole convolution kernels, the size of each resolution of 3D video can be obtained, and then the 3D video can be clarified. Then, these feature maps and outputs of the final spatial feature map can be connected. Let the dimension of the input three-dimensional tensor $\boldsymbol{T}$ be $H \times W \times D$, and H and W are the length and width of the image, and D is the number of channels of the tensor. L is a set of convolution kernels $\{\mathbf{C_l}\}_{l=1}^{L}$, and the size of the convolution kernel is $n \times n$. The void ratio is $\{r_l\}_{l=1}^{L}$, and the convolution kernels with different void rates have different receptive fields (α). The receptive field represents the spatial structure of the convolution kernel, which can reflect the number of frames and quality resolution of the image. The size of the receptive field is shown in (1). In the calculation process, each layer will output a feature map $\{\mathbf{F_l}\}_{l=1}^{L}$, as shown in (2).

$$\alpha = \left[(n-1)r_l + 1\right]^2, \tag{1}$$

$$\mathbf{F}_l = \mathbf{C}_l \times \mathbf{T}. \tag{2}$$

It can be seen from (2) that when $\boldsymbol{T}$ remains unchanged, the output characteristic graph is only related to the size and void rate of the void convolution kernel. Although the size of each output feature image is the same, due to the different void rates, the size of the receptive field is also different. Sometimes, the size of the receptive field will exceed the input image, which is just like observing the image at different distances. If the distance is too far or too close, you will not see all the information about the object. Only from the appropriate distance and background can you get the effective salient object/region.

After obtaining the feature map $\{\mathbf{F_l}\}_{l=1}^{L}$, they need to be connected together and sent to the improved bidirectional ConvLSTM as input. Bidirectional ConvLSTM is a direct input algorithm, which can greatly improve the accuracy of the algorithm in the input process and make the feature maps more closely connected. Therefore, the first part of the model can learn features from different scales through the data set. The formula of characteristic diagram connection operation is shown in the following formula:

$$\mathbf{P} = \left[\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_\mathbf{n}\right], \tag{3}$$

where [,] represents the connection operation and $\mathbf{P}$ is the characteristic diagram obtained after the connection operation.

In addition, since the depth residual framework is also applicable to the model in this paper, the source input characteristic graph $\mathbf{T}$ is also added to the connection operation, that is,

$$\mathbf{P} = \left[\mathbf{T}, \mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_n\right]. \tag{4}$$

As we all know, ConvLSTM is mainly composed of four parts: memory cell $\mathbf{c}_t$, input gate $\mathbf{i}_t$, output gate $\mathbf{o}_t$, and forgetting gate $\mathbf{f}_t$. $\mathbf{c}_t$ is a kind of state accumulator, which is controlled by the other three parts $\mathbf{i}_t$, $\mathbf{o}_t$, and $\mathbf{f}_t$. When the input door is open, the input new data will be added to $\mathbf{c}_t$. Similarly, when the forgetting gate is activated, the past cell state $\mathbf{c}_{t-1}$ will be discarded. Whether the final hidden layer state $\mathbf{H}_{t-1}$ will be affected by the current cell state depends on the output gate $\mathbf{o}_t$. As mentioned above, ConvLSTM can be expressed as follows:

$$\left\{\begin{array}{c} \mathbf{i}_t = \sigma\left(\mathbf{W}_i^P \times P_t + \mathbf{W}_i^H \times \mathbf{H}_{t-1}\right) \\ \mathbf{f}_t = \sigma\left(\mathbf{W}_f^P \times P_t + \mathbf{W}_f^H \times \mathbf{H}_{t-1}\right) \\ \mathbf{o}_t = \sigma\left(\mathbf{W}_o^P \times P_t + \mathbf{W}_o^H \times \mathbf{H}_{t-1}\right) \\ \mathbf{c}_t = \mathbf{f}_t \times \mathbf{c}_{t-1} + \mathbf{i}_t \tanh\left(\mathbf{W}_c^P \times \mathbf{P}_t + \mathbf{W}_c^H \times \mathbf{H}_{t-1}\right) \\ \mathbf{H}_{t-1} = \mathbf{o}_t \times \tanh \mathbf{c}_t \end{array}\right\}. \tag{5}$$

According to (5), ConvLSTM can simply "remember" the spatiotemporal information in the sequence. However, the spatiotemporal information in the front and back directions of video plays a very important role and significance

in significance detection. Therefore, use bidirectional ConvLSTM to $\{\mathbf{P_m}\}_{m=1}^{M}$ for bidirectional spatiotemporal feature extraction. In the process of extraction, attention should be paid to maintaining the space-time information in the left, right, front, and back directions of 3D video and tracking and acquisition at any time to maintain the timeliness of the extraction process.

$$\mathbf{Y}_t = \tanh\left(\mathbf{W}_y^{Hf} \times \mathbf{H}_t^f + \mathbf{W}_y^{Hb} \times \mathbf{H}_{t=1}^b\right), \tag{6}$$

where $\mathbf{H}^f$ and $\mathbf{H}^b$ represent the hidden layer states of forward and backward ConvLSTM, respectively; $\mathbf{Y}_t$ represents the final saliency map output after fusing the two-way spatiotemporal information; tanh (•) indicates the tanh activation function.

When using a saliency map for 2D-3D conversion, the calculation steps of parallax calculation and image rendering based on a saliency map are basically the same as those based on the depth map. This is because, although saliency detection and depth map prediction are two different fields, the prediction maps (saliency map and depth map) generated by the two methods are the same in image attributes, that is, gray-scale images with the brightness of 0–255. It should be noted that the 2D-3D conversion method based on a depth map generates 3D content according to the distance between the object/area and the camera lens, while the method based on a saliency map performs 2D-3D conversion according to whether the human eye is interested in the target/area.

In the calculation of parallax, there is little difference between this method and the traditional method. Parallax calculation is to ensure the same number of frames in the conversion of 2D and 3D videos, so as to make the data information of saliency map and depth map consistent. This is to make the 3D video closer to reality and restore the original ecological scene. If the parallax value of the pixel in the figure at (x, y) is $R(x.\ y)$, then

$$\begin{aligned} R(x, y) &= Z\left[1 - \frac{Y(x, y)}{128}\right], \\ Z &= \frac{GK}{G + D}, \end{aligned} \tag{7}$$

where Z represents the maximum parallax and G represents the maximum depth. K is the distance between human eyes, and D represents the distance between human eyes and screen. $Y(x, y)$ represents the significant value in the image (x, y).

When evaluating the performance of the model from the perspective of 3D video, there are three widely used metrics, namely precision-recall curve, F-measure, and mean absolute error (MAE). In the precision-recall curve, precision represents the percentage of significant pixels correctly predicted, and recall represents the proportion of significant pixels correctly predicted in the truth map. F-measure is used to measure the overall performance of the model and reflects the accuracy of the model, which can accurately characterize the use performance of 3D video model and better represent the accuracy of 3D video. The higher the F-measure of the model, the better the performance.

$$\omega_\beta = \frac{\left(1 + \beta^2\right) A_{\text{precision}} A_{\text{recall}}}{\beta^2 A_{\text{precision}} + A_{\text{recall}}}, \tag{8}$$

where ω_β for β is the F-measure of the weight parameter and $A_{\text{precision}}$ and A_{recall} represent accuracy and recall, respectively. The value of β is 0.3. MAE represents the absolute error between the prediction map and the truth map, which represents the difference between the two lines in the precision-recall curve, and it is a microcharacterization method. The lower the MAE value of the system, the better the performance.

$$\delta = \frac{1}{HW} \sum_{x=1}^{H} \sum_{y=1}^{W} |Y - G|, \tag{9}$$

where δ indicates the MAE value and H and W represent the length and width of the image, respectively. Y and G represent the binary graph generated after binarization of prediction graph Y and truth graph G, respectively, and the interval is [0, 1].

4. Comparison and Analysis of Experimental Results

4.1. Analysis of the Application of Public Art to Rural Culture from the Perspective of 3D Video. With the advancement of China's economy, various public art activities are booming. This is not only the need for the development of art itself but also the awakening of people's public consciousness. However, we must understand that whether these rural art activities we have sprung up can be carried out for a long time is not just a hothead. It can be called art changing the countryside by drilling into the village and painting a few wall paintings and placing a few sculptures. The entry of art into the countryside does not mean that artists enter the village. The entry of new types of public art into the countryside should be a continuous, in-depth, all-round, and three-dimensional intervention.

This paper selects the rural culture in four municipalities directly under the central government to study. Figure 3 shows the application rate of public art to rural culture from the perspective of 3D video. It can be seen that with the increased time, the application rate of public art to rural culture in the four municipalities directly under the central government shows a downward trend. However, Beijing has the highest application rate, followed by Shanghai, while Tianjin has the lowest application rate. The difference in rural culture between different cities is mainly the difference in local rural traditional culture. The per capita GDP of rural people in developed areas is high. They have time and energy to accept public art, which improves the application rate of rural culture to a certain extent. Meanwhile, the main reason may be that Beijing and Shanghai have a long history, and the rural people have a high degree of acceptance of public art. Therefore, in the process of implementing public art in the rural culture, we should pay attention to the advanced nature and practicality of public art, so as to ensure the public art's application rate in rural culture.

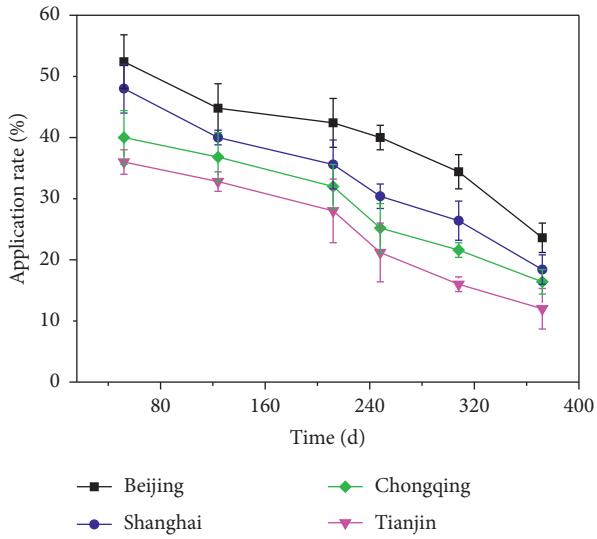


Figure 3: Application rate of public art to rural culture from the perspective of 3D video.

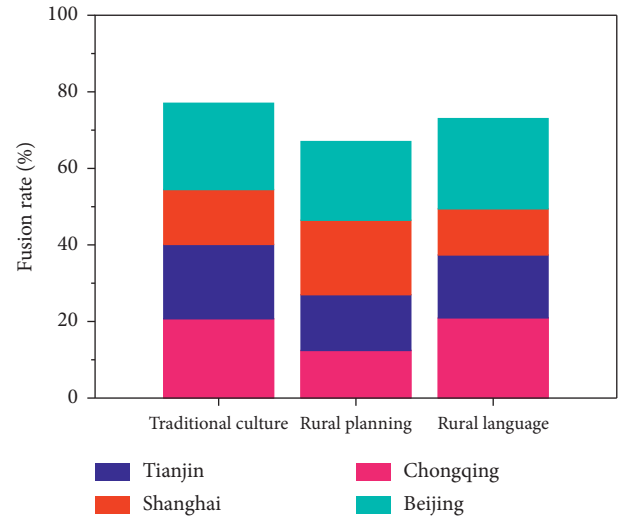


Figure 4: The influence of public art on rural cultural diversity from the perspective of 3D video.

4.2. Statistics of Public Art on Rural Cultural Diversity from the Perspective of 3D Video. Today, the development of new types of public art should be integrated into all aspects of our life and create an artistic atmosphere that affects the public. It can be sculpture, mural, architecture, lighting, new media, dialogue, performance, festival design, and so on, which occur in space and affect people's visual, auditory, tactile, and other feelings. The way for new types of public art to intervene in rural culture must be cross-border, multidisciplinary, all-round, and three-dimensional. In addition to pure art, architecture, sociology, psychology, history, statistics, and even politics can be used as a way to realize public art.

Figure 4 shows the influence of public art on rural cultural diversity from the perspective of 3D video. It can be seen that the integration rate between rural traditional culture, rural planning, and rural language is different in different cities. Among them, Beijing has more in these three aspects, while Shanghai has more in rural planning and less in rural language. Chongqing has a low integration rate in rural planning and a high integration rate in traditional culture and rural language. For Tianjin, due to its rich history and culture, the integration rate of traditional culture and rural language is high, but there is a lack of rural planning. In short, in order to better reflect the integration of public art into rural cultural diversity, the promotion of public art should be carried out according to local conditions according to the culture, financial resources, and rural style of local cities, so as to contribute to the integration of public art into rural cultural diversity and promote the promotion and application of public art.

4.3. Analysis of the Development of Public Art on Rural Culture from the Perspective of 3D Video. Low productivity, dilapidated and backward public infrastructure, massive labor loss, and lack of core competitiveness are common problems in many rural areas in China. It is an ideal direction to change the backward appearance and productivity through the intervention of "art" in the countryside. Encourage people to participate, persuade the government to support, and obtain funds through enterprises to help promote. In the process of transforming art and cultural resources into economic capital for sustainable development in the countryside, whether it is led by the government or the spontaneous behavior of non-governmental organizations, it can form its own relatively perfect hematopoietic function and improve the growth mechanism from production, consumption, and promotion, which is the core issue that needs to be developed urgently for the current "art intervention in the countryside." Therefore, this paper studies the development prospect of public art for rural culture between different cities and villages.

The development rate of public art in a rural culture based on 3D video between villages in the same city is shown in Figure 5. It can be seen that with the increase of time, the development rate of public art for rural culture shows a trend of increasing first and then decreasing. The development rate is the highest in about 100 days and then slowly decreases to a stable state. Among them, the development prospect of Tianjin is the best, while that of Shanghai and Beijing is poor. The main reason may be that Tianjin is located in the coastal area, and the rural people are more open-minded and easier to accept public art. However, due to the deep-rooted local traditional culture, the acceptance of public art in Beijing and Shanghai has been reduced, resulting in a reduction in the development rate. Therefore, in order to improve the development rate of public art for rural culture, we should improve the innovation of public art, form art in line with local cultural characteristics, stimulate the endogenous vitality of art, and promote the dissemination and development of public art in rural culture.

4.4. Analysis of the Participation of Public Art in Rural Culture from the Perspective of 3D Video. The countryside is not a canvas that can be freely sprinkled. Even if you are a famous

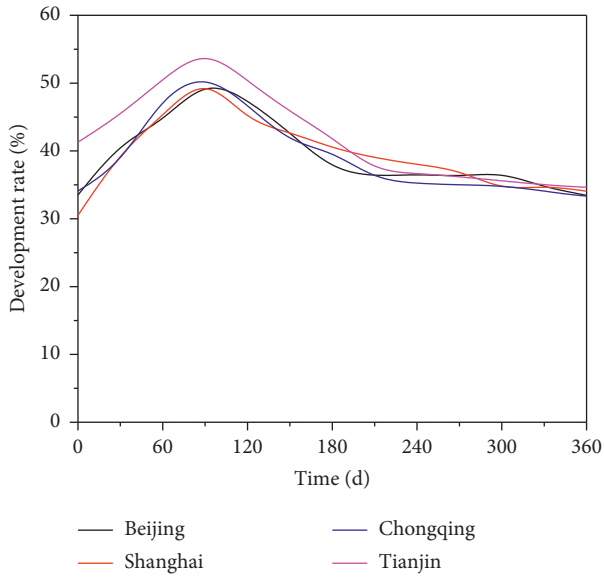


Figure 5: Development rate of public art to a rural culture based on 3D video between villages in the same city.

artist, you cannot forcibly implant your own ideas and works. Forcibly implanting works of art that are divorced from the local environment will not only destroy the life comfort of the local people but also make the work itself extremely embarrassing and even destroyed. The process of art intervening in the countryside must be completed imperceptibly. Art practitioners entering the countryside should strive to be part of the local people, lay down their bodies, enjoy with the people, integrate themselves in the way of emotional transfer, and feel their own feelings in the process of practice, so as to achieve growth. Public art is definitely not an artist's personal stage. Artists leave after finishing work, while local residents have to stay with the work all their life. Therefore, equal participation and mutual understanding are very important. Let them slowly understand why they do this, whether it is meaningful to them, and whether it can trigger emotion, rather than being imposed as a guide.

Figure 6 shows the participation of public art in rural culture from the perspective of 3D video. It can be seen that Beijing's public art has the highest participation in rural culture, followed by Shanghai and Chongqing, while Tianjin has the lowest. The possible reason is that Beijing and Shanghai have a high comprehensive economic capacity and higher education culture, which affects the rural people's pursuit and appreciation of public art, so their participation is high. Due to the low comprehensive economic capacity of Tianjin, rural people pay attention to improving their quality of life and have no time and energy to consider public art, which leads to the reduction of participation. Therefore, if we want to improve the participation of public art in rural culture, we should focus on solving the income level of local rural people, increase the publicity of local traditional culture, and make the villagers accept the new public art.

When evaluating the participation of public art in rural culture in different cities, we should consider the proportion

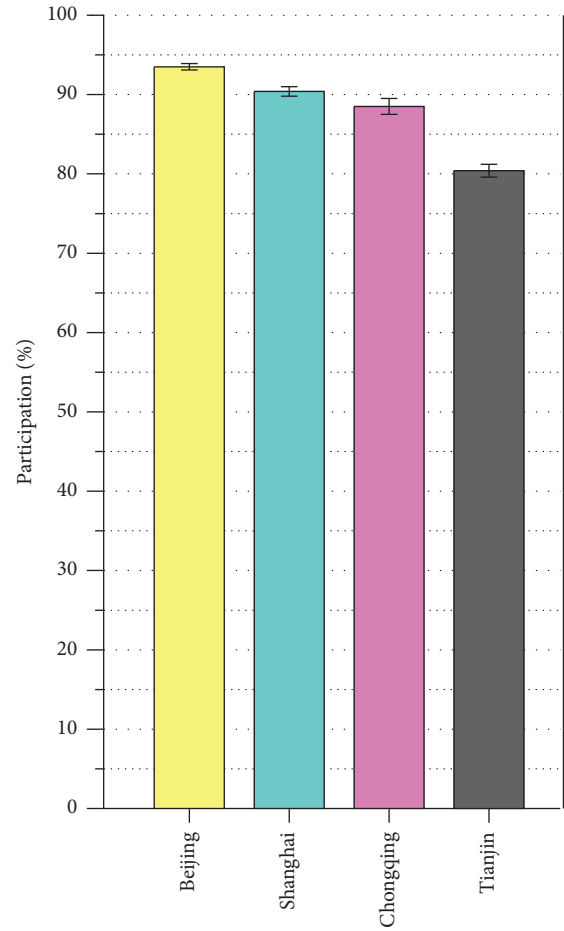


Figure 6: Participation of public art in rural culture from the perspective of 3D video.

of traditional culture, rural planning, and rural language. Figure 7 shows the proportion of rural culture among different cities. It can be seen that the proportion of traditional culture and rural planning in Beijing is relatively high, that of rural language in Shanghai is the highest, that of Chongqing is the third, and that of Tianjin is the lowest, which is consistent with the participation of public art in rural culture. Therefore, if we want to improve the participation of urban public art in rural culture, we should start from the local rural traditional culture. Planning and language should be arranged to mobilize the labor enthusiasm of local people and make them have the courage to accept advanced public art, so as to promote the prosperity and development of rural culture.

4.5. Statistics of Rural People's Sense of Achievement of Public Art from the Perspective of 3D Video. The sense of achievement of the rural people represents the yearning and expectation of the local people for a better life. If the public art can be integrated with the local rural culture, the sense of achievement of the people will be higher. Investigate and analyze the rural people in four cities to analyze the people's sense of achievement. The fulfillment of public art from the perspective of 3D video among people in different cities and

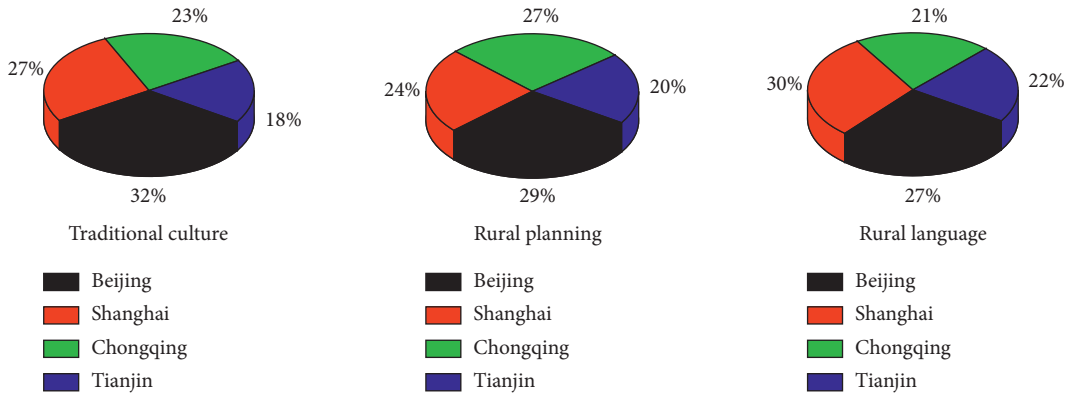


Figure 7: Proportion of rural culture among different cities.

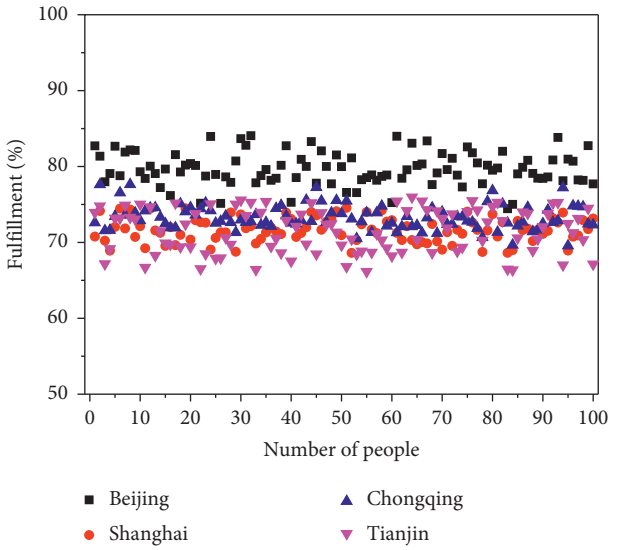


Figure 8: The fulfillment of public art from the perspective of 3D video among people in different cities and villages.

villages is shown in Figure 8. It can be seen that the rural people in Beijing have the highest sense of achievement in public art, and the fluctuation of everyone's sense of achievement is small, followed by Chongqing and Shanghai, while the people in Tianjin have the lowest sense of achievement, and the fluctuation is also large. The main reason may be related to the local economy. The citizens of Tianjin are busy making a living and lack appreciation of public art, while the per capita income of Beijing and Chongqing is high; rural people have time to enjoy the beauty of public art, so their sense of achievement is higher.

5. Conclusion

Public art has a great impact on the development of rural culture. From the perspective of 3D video, this paper studies the application principle of 3D video and establishes the corresponding model. Then, the model is applied to four municipalities directly under the central government to study the three-dimensional intervention mode of public art for rural culture, analyze the development prospect of public art for rural culture in each city, and put forward countermeasures according to the corresponding problems, so as to return public art to public life and promote the prosperity and development of rural culture.

Data Availability

The data set can be accessed upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

The authors thank the Fundamental Research Funds for the Central Universities (no. 2020SQN09) for the support.

References

[1] E. Samiei, A. Seyfoori, B. Toyota, and Ghavami, "Investigating programmed cell death and tumor invasion in a three-dimensional (3D) microfluidic model of glioblastoma," *International Journal of Molecular Sciences*, vol. 21, no. 9, pp. 23–35, 2020.

[2] S. Kolovos, A. P. Finch, H. Ploeg et al., "Five-year cost-effectiveness analysis of the European Fans in Training (EuroFIT) physical activity intervention for men versus no intervention," *International Journal of Behavioral Nutrition and Physical Activity*, vol. 17, no. 1, pp. 91–101, 2020.

[3] M. A. Sarge, H. S. Kim, and J. A. Velez, "An as intervention: the role of perspective taking in combating public stigma with virtual simulations," *Cyberpsychology, Behavior, and Social Networking*, vol. 23, no. 1, pp. 41–51, 2020.

[4] L. A. Adutwum, J. K. Kwao, and J. J. Harynuk, "Unique ion filter a data reduction tool for chemometric analysis of raw comprehensive two dimensional gas chromatography mass spectrometry data," *Journal of Separation Science*, vol. 44, no. 14, pp. 2773–2784, 2021.

[5] G. P. Esteves, P. Swinton, C. Sale, R M. James, G G. Artioli, and Roschel, "Individual participant data meta-analysis provides No evidence of intervention response variation in individuals supplementing with beta-alanine," *International Journal of Sport Nutrition and Exercise Metabolism*, vol. 23, no. 2, pp. 5–16, 2021.

[6] P. Fernández-González, A. Koutsou, A. Cuesta-Gómez, C. Carratalá-Tejada, and Molina-Rueda, "Reliability of kinovea software and agreement with a three-dimensional motion system for gait analysis in healthy subjects," *Sensors*, vol. 20, no. 11, pp. 35–44, 2020.

[7] M. D. Murphy, D. Pinheiro, R. Iyengar, and Lim, "A data-driven social network intervention for improving organ donation awareness among minorities: analysis and optimization of a cross-sectional study," *Journal of Medical Internet Research*, vol. 22, no. 1, pp. 13–25, 2020.

[8] D. T. Robinson, A. Schertenleib, B. M. Kunwar, and J. Shrestha, "Environmental research and public health assessing the impact of a risk-based intervention on piped water quality in rural communities: the case of mid-western Nepal," *The Journal*, vol. 3, no. 2, pp. 45–56, 2019.

[9] E. L. Andrade, J. B. Bingenheimer, M. C. Edberg, K. L. Zoerhoff, and E. M. Putzer, "Evaluating the effectiveness of a community-based hygiene promotion program in a rural Salvadoran setting," *Global health promotion*, vol. 26, no. 1, pp. 69–80, 2019.

[10] J. Xia, H. H. Ip, N. Samman, H. T. Wong, and Gateno, "Three-dimensional virtual-reality surgical planning and soft-tissue prediction for orthognathic surgery," *IEEE Transactions on Information Technology in Biomedicine*, vol. 34, no. 1, pp. 121–128, 2019.

[11] J. J. Hong, T. S. Nam, D. Park, and I. H. Jeon, "Three-dimensional morphometric analysis of penetrative depth and size of nonarthritic and degenerative arthritic glenoids: implications for glenoid replacement in shoulder arthroplasty," *Clinical Orthopaedic Surgery*, vol. 12, no. 2, pp. 136–147, 2020.

[12] E. Kentaro, M. Toru, K. Kunihisa, M. Nakahama, and Doi, "Protective effect of remote ischemic preconditioning on myocardial damage after percutaneous coronary intervention in stable Angina patients with complex coronary lesions - subanalysis of a randomized controlled trial," *Circulation Journal: Official Journal of the Japanese Circulation Society*, vol. 82, no. 7, pp. 1788–1796, 2019.

[13] B. Janet, S. Colin, C. G. Julio, D. Carpenter, T M. Link, and S. Majumdar, "Three-dimensional image registration of MR proximal femur images for the analysis of trabecular bone parameters," *Medical physics*, vol. 35, no. 10, pp. 4630–4639, 2019.

[14] A. K. Alina, F. J. Gates, V. M. Mays, K. W. Chang, and S D. Cochran, "Aggression, escalation, and other latent themes in legal intervention deaths of non-hispanic black and white men: results from the 20032017 national violent death reporting system," *American Journal of Public Health*, vol. 11, no. 2, pp. 107–115, 2021.

[15] J. P. Y. Ho, A. M. Merican, M. S. Hashim, A. A. Abbas, C. K. Chan, and J. A. Mohamad, "Three-dimensional computed tomography analysis of the posterior tibial slope in 100 knees," *The Journal of Arthroplasty*, vol. 32, no. 10, pp. 3176–3183, 2017.

[16] G. V. Beznoussenko, A. Ragnini-Wilson, C. Wilson, and A. A. Mironov, "Three-dimensional and immune electron microscopic analysis of the secretory pathway in *Saccharomyces cerevisiae*," *Histochemistry and Cell Biology*, vol. 146, no. 5, pp. 515–527, 2016.

[17] J. L. Fowler, L. S. Seung-Young, Z. C. Wesner, S. K. Olehnik, S. J. Kron, and M. Hara, "Three-dimensional analysis of the human pancreas," *Endocrinology*, vol. 3, no. 3, pp. 13–20, 2018.

[18] O. Benjamin, L. Christopher, A. Neil et al., "Computational hemodynamics of abdominal aortic aneurysms: three-dimensional ultrasound versus computed tomography," *Proceedings of the Institution of Mechanical Engineers - Part H: Journal of Engineering in Medicine*, vol. 230, no. 3, pp. 201–210, 2017.

[19] M. Royen, E. I. Verhoef, C. F. Kweldam et al., "Three-dimensional microscopic analysis of clinical prostate specimens," *Histopathology*, vol. 69, no. 6, pp. 985–992, 2017.

[20] R. Tiara, V. R. Claudia, C. R. Pischke, S. Muellmann, M. Peters, and S. Lippke, "Health-related lifestyle and dropout from a web-based physical activity intervention trial in older adults: a latent profile analysis," *Health psychology Official Journal of the Division of Health Psychology, American Psychological Association*, vol. 40, no. 8, pp. 481–490, 2021.

[21] S. A. Kharroubi, Y. Beyh, E. Abdul Fattah, and T. Young, "The importance of accounting for parameter uncertainty in SF-6D value sets and its impact on studies that use the SF-6D to measure health utility," *International Journal of Environmental Research and Public Health*, vol. 17, no. 11, Article ID 3949, 2020.

[22] S. Florian, P. Barbara, K. Andreas, A. Stauffer, and S. Farr, "Outcomes of patients with single-bone-forearm surgery: a clinical assessment and three-dimensional motion analysis," *The Journal of hand surgery, European volume*, vol. 44, no. 8, pp. 838–844, 2020.

[23] X. Zhou, A. Crippa, A. K. Danielsson, M. R. Galanti, and N. Orsini, "Effect of tobacco control policies on the Swedish smoking quitline using intervention time-series analysis," *BMJ Open*, vol. 9, no. 12, pp. e033650–12, 2019.

[24] B. Nico, V. Stefan, K. Michiel et al., "Three-dimensional and quantitative analysis of atherosclerotic plaque composition by automated differential echogenicity Catheterization and cardiovascular interventions," *Official Journal of the Society for Cardiac Angiography & Interventions*, vol. 70, no. 7, pp. 968–978, 2019.

[25] A. Kotalik, A. Eaton, Q. Lian, C. Serrano, J. Connett, and J. D. Neaton, "A win ratio approach to the re-analysis of multiple risk factor intervention trial," *Clinical Trials*, vol. 16, no. 6, pp. 626–634, 2019.

[26] Z. Robert and M. S. O'Toole, "The effect of expressive writing intervention on psychological and physical health outcomes in cancer patients--a systematic review and meta-analysis," *Psycho-Oncology*, vol. 24, no. 11, pp. 1349–1359, 2021.

[27] E. Kazuhiro, S. Naohiro, S. Kunihiko, and Y. Yasuda, "Pancreatic arteriovenous malformation: a case report of hemodynamic and three-dimensional morphological analysis using multi-detector row computed tomography and post-processing methods," *JOP Journal of the pancreas*, vol. 10, no. 1, pp. 59–63, 2019.

Hindawi
Scientific Programming
Volume 2022, Article ID 5726718, 9 pages
https://doi.org/10.1155/2022/5726718

Research Article

Analysis on Transmission Characteristics of Stimulated Raman Scattering Based on the Multi-Sensor Signal Enhancement Technique

Wentao He and Zhiwei Men

College of Physics Jilin University, Changchun, Jilin 130012, China

Correspondence should be addressed to Wentao He; hewt19@mails.jlu.edu.cn

Received 28 March 2022; Revised 20 April 2022; Accepted 22 April 2022; Published 11 May 2022

Academic Editor: Baiyuan Ding

In recent 20 years, fibre laser system has been developed rapidly and widely used for its high quality, high efficiency, high robustness, and compactness. However, there are still many factors (such as non-linear effect, thermal effect, and mode instability) that limit the further increase of power of fibre laser system. Stimulated Raman scattering (SRS) is one of the main limitations in the transmission process of fibre lasers. It not only reduces the output efficiency of fibre lasers, but also increases the damage risk of reverse Stokes light to the system. Recent studies have shown that SRS in low-mode fibres can lead to quasi-static mode degradation in addition to mode instability. With the introduction of multi-sensor enhancement technology in the fibre field, it becomes an effective means to popularise high-power and high-beam quality fibre lasers. Based on the multi-sensor signal enhancement technology, this paper explores the influence of this technology on the output efficiency of SRS in the mode-reducing fibre laser, which provides a new idea and method for the output efficiency and transmission analysis method of fibre laser.

1. Introduction

In 1961, Snitzer of America Optical developed the world's first fibre laser [1, 2]. With the unique advantages of high conversion efficiency, good beam quality, and compactness, fibre lasers have shown promise for a wide range of applications in earth sciences, strong field physics, atomic and molecular physics, frequency conversion, beam synthesis, and industrial processing. Led by pioneers such as Dr. Charles Kao, Dr. E. Snitzer, Dr. Oleg G. Okhotnikov, and Dr. Valentin Gapontsev, fibre laser technology is constantly evolving and single-mode laser power has reached the 10 kW threshold [3]. However, the special waveguide structure of optical fibres allows the laser energy to be confined mainly within the μm level of the fibre core, and as the laser power of the fibre increases, very high energy densities will develop within the core, causing various harmful non-linear effects [4, 5]. Lawrence Livermore Laboratory (USA) and the National University of Defense Technology (China) have conducted theoretical analyses of the output power limit of laser systems under semiconductor laser (LD) pumping structure and the same band pumping structure, respectively, and the results show that the SRS effect is one of the main factors limiting the power increase, and the same conclusion has been obtained experimentally. In order to achieve high-power laser output, few-mode fibres are widely used in high-power fibre laser systems [6]. Recent results have shown that in few-mode fibres, SRS induces a quasi-static mode degradation in conjunction with the excitation of mode instabilities (TMI), and that the corresponding threshold power is well below the threshold defined in the previous model for limiting the near-diffraction output power [7]. Therefore, effective strategies must be adopted to suppress SRS and improve the output of existing high-power, high-beam quality laser systems.

With the joint efforts of researchers, the understanding of SRS has been deepened, and related inhibition techniques have been developed and summarised. Early reviews on

technologies to improve SRS transmission performance mainly focused on fibre structure and design of long-period Raman suppression gratings [8, 9]. In addition to fibre design, the results also show that the system parameters have a great influence on SRS. Therefore, it is necessary to summarise the influence of each parameter on SRS strength at a general level and study its optimisation direction. At the same time, with the development of technology, new fibre structure design and Raman suppression grating technology emerge and show excellent performance, one of which is multi-sensor enhancement technology [10]. In this paper, the influence of multi-sensor enhancement technology on transmission efficiency characteristics of SRS is analysed, and a fibre structure design and transmission analysis method based on multi-sensor signal enhancement technology are proposed. On this basis, the mechanism of multi-sensor enhancement technology on improving transmission efficiency of fibre laser is further clarified.

2. New Physical Characterisation of the SRS Effect Based on the Multi-Sensor Signal

The semi-classical physical formulation of SRS is that an incident photon of frequency ω_p interacts with a medium of intrinsic frequency ω_v to form a Stokes photon of frequency ω_s and an anti-Stokes photon of frequency ω_a [11–13]. On the one hand, this shifts the energy from the signal light to the Raman scattered light, causing a decrease in the power and energy conversion efficiency of the output beam; on the other hand, as the Stokes light generated by the SRS is bi-directional Stokes light, when the backward Stokes light power reaches a certain value, it increases the risk of damage to the optical components in the laser system, limiting the further increase in the power of the laser system [14, 15]. As a light intensity-dependent non-linear effect, increasing the mode field area with a few-mode fibre can effectively suppress SRS; however, the increased mode field area allows the fibre to support multiple modes. Recent studies have shown that the SRS effect in few-mode fibres leads to mode degradation of the output beam, which can be divided into dynamic mode degradation (i.e., transverse mode instability) and quasi-static mode degradation.

As one of the main limiting factors for the power enhancement of current fibre laser systems, TMI is physically characterised by a sharp degradation of the output laser beam quality after the average power of the output laser reaches a threshold value, with dynamic energy transfer between fundamental modes (FMs) and higher-order modes (HOMs) on the order of kHz, which is manifested by transverse mode instability of the output beam [16]. In Raman fibre amplifiers (RFAs), the quantum loss heat generation during the SRS process can also form thermally refractive index gratings, resulting in TMI in Raman light and leading to dynamic energy coupling between the fundamental and higher-order modes [17, 18]. This phenomenon has also been recently observed in ytterbium-ion-Raman hybrid gain fibre laser systems. However, the experimental results also show that when SRS induces TMI, the distribution of the signal light (Raman-pumped light) shows a lateral temporal instability, while the Raman Stokes light only varies in intensity and the lateral distribution remains unchanged, corresponding to the temporal evolution of the spot shape and the mode decomposition of the signal light as shown in Figure 1. This means that the SRS-induced TMI occurs on signal light rather than Raman light.

Since the formation of TMI requires a certain intensity of thermal loading, SRS-induced TMIs usually occur with a Raman power share greater than 50% [19]. Recently, another type of quasi-static mode degradation has been observed in cases where the Raman power share is relatively low (see Figure 2). Experimentally, when the Raman power share reaches 3%, the output laser beam quality degrades rapidly and the spot appears jittered. This phenomenon can be effectively suppressed by suppressing the SRS [20]. The time-domain signal results show corresponding eigenfrequencies in the order of Hz. There are currently two explanations for the physical mechanism of quasi-static mode degradation caused by SRS: one is the coupling of energy from the fundamental mode of the signal light to the higher-order mode of the Raman light due to the inter-mode mixing (IM-WM) effect; the other is the mode degradation of the signal light due to the core-pumped Raman effect.

3. Analysis Method on Transmission Characteristics of SRS Based on the Multi-Sensor Signal Enhancement Technique

On the one hand, the new physical characterisation of SRS in few-mode fibres further limits the power enhancement of high-power, high-beam quality fibre laser systems; on the other hand, as a new phenomenon, there is still a lack of a clear and uniform understanding of the mechanisms underlying the transmission properties of SRS in fibres based on multi-sensor enhancement techniques, and experimental results do not fully match the theoretical explanation [21]. The active suppression strategy in SRS can dynamically regulate the Raman intensity, which can help to further investigate the physical mechanism behind this phenomenon. The use of sensors to dynamically modulate the Raman intensity in SRS transmission properties can help to further investigate the physical mechanisms behind this phenomenon. Thus, in view of the fibre optic transmission efficiency bottleneck, it is necessary to systematically review the beneficial effects of multi-sensor signal enhancement techniques on SRS transmission characteristics and further develop relevant transmission analysis ideas and new methods. Based on the structural design of the multi-sensor signal enhancement technology, this paper presents an overview of optical fibre transmission characterisation methods from the perspective of fibre design and system optimisation.

3.1. Optical Sensors. The coupling equation describing the SRS effect in continuous wave operation, neglecting the Raman optical gain of rare earth ions, dispersion, and non-

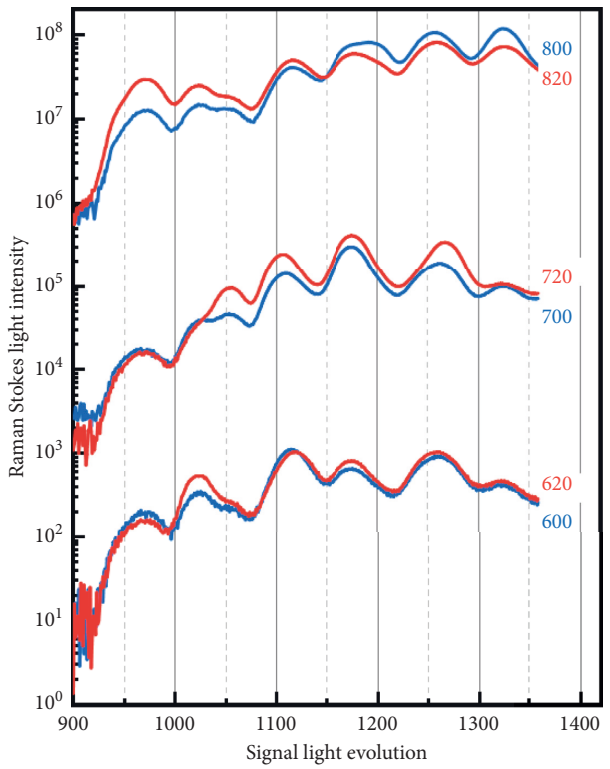


Figure 1: Raman Stokes light intensity corresponding to the temporal evolution of the signal light.

linear effects such as self-phase modulation (SPM) and cross-phase modulation (XPM) on the SRS, can be formulated as

$$\begin{aligned}\frac{dI_R}{dz} &= i_g - i_\alpha,\\ i_g &= g_R I_s I_R,\\ i_\alpha &= \alpha_R I_R,\\ \frac{dI_R}{dz} &= g_R I_s I_R - \alpha_R I_R,\\ \frac{dI_s}{dz} &= -\frac{\omega_s}{\omega_R} g_R I_s I_R - \alpha_s I_s,\end{aligned} \tag{1}$$

where I_s and I_R, are the signal and Raman light intensities; g_R is the Raman gain coefficient; ω_s and ω_R are the signal and Raman light frequencies; and α_s and α_R correspond to the loss factors of the signal and Raman light.

The loss term of Raman light can be expressed as αRIRLeff and the gain term as g RIsIRLeff. Therefore, in order to improve the efficiency of SRS transmission, it is necessary to "increase the Raman loss and decrease the Raman gain, while ensuring the optical power and beam quality of the output signal as much as possible." Compared to increasing the Raman loss, reducing the Raman gain fundamentally reduces the production of Raman light. On the one hand, this improves the transmission efficiency and stability of the system; on the other hand, reducing the gain

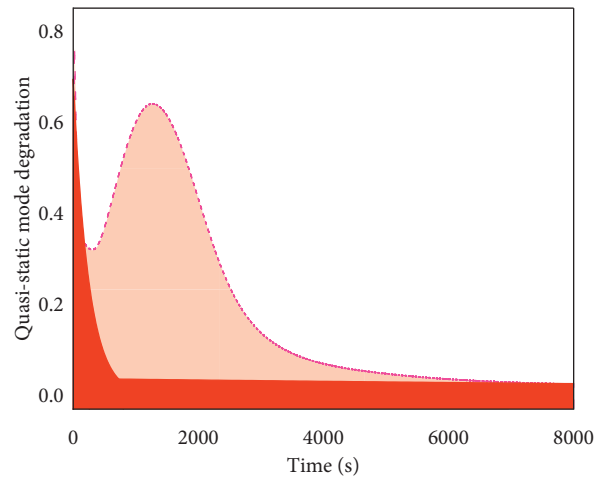


Figure 2: Quasi-static mode degradation with the Raman power share.

also reduces the heat generation due to quantum losses caused by SRS. This is why "Raman light gain reduction" is a common strategy in current optical sensor signalling technology. The latter can be achieved by optical sensors in addition to increasing the core area using large mode field fibres. In addition, the total gain can be reduced by decreasing the effective length of action.

3.1.1. Material Components. Hu et al. [22] suggested that the SRS effect could be suppressed by using Yb: YAG-derived fibres instead of conventional SiO_2 quartz fibres to reduce the Stokes optical gain coefficient. It was shown that an increase in the $Y_2O_3 + Al_2O_3$ content in the fibre leads to a decrease in the Raman gain coefficient, as shown in Figure 3.

In 2021, Sciortino et al. [23] summarised and collated the effect of fibre material on the non-linear effect, stating that the SRS power can be expressed as

$$P_s^R \propto V_m \Lambda^2, \tag{2}$$

where V_m is the molar volume and Λ is the bond compression factor when fibre material is under non-linear effect.

In order to achieve low Raman gain, the fibre material needs to be able to meet the following conditions:

(1) The material is highly disordered, thereby broadening the Raman gain spectrum and reducing the peak.

(2) A material with a high concentration and low gain coefficient g_R is used.

(3) The Raman spectrum of the laser system has minimal overlap with the Raman gain spectrum corresponding to the material components.

3.1.2. Large Mode Field Sensors. As a light intensity-dependent non-linear effect, increasing the mode field area is one of the most effective means of improving the SRS transmission efficiency, which on the one hand leads to a

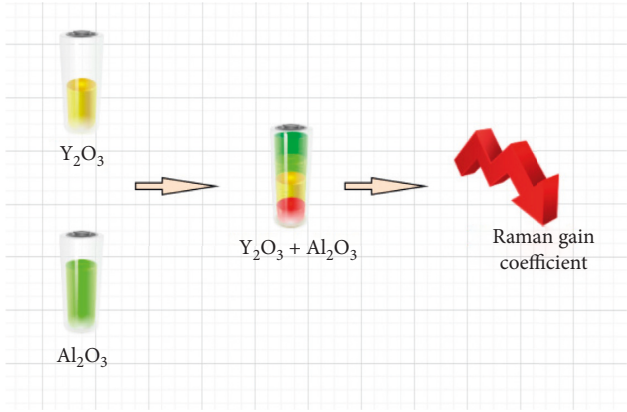


Figure 3: The diagram of the $Y_2O_3 + Al_2O_3$ content in the fibre with the Raman gain coefficient.

reduction in the Raman-pumped light intensity and on the other hand implies a reduction in the effective action distance by increasing the mode field area. In general, the fibre material has a high refractive index, which leads to a further increase in the mode field area of the few-mode fibre, thus causing a decrease in the TMI threshold reduction, which is related to the material used in the fibre. The core of the design of large mode field fibres is therefore to suppress TMI by embedding a large mode field sensor, thus improving the efficiency of the fibre transmission. The most commonly used large mode field sensor fibres are partially doped fibres, tapered fibres, and spindle fibres (SSC). Partially doped fibres are partially doped in the centre to enhance the gain capability of the fundamental mode and achieve higher-order mode suppression through large mode field sensors; tapered fibres have a non-uniform core size distribution along the fibre axis, which can effectively strip higher-order modes and increase the TMI threshold; spindle fibres can be equated to a pair of tapered fibres fused to a section of large mode field fibre. The reduced mode field area at the end of the fibre further improves the output laser beam quality compared to tapered fibres. However, as the output power is further increased to 8.5 kW, the SNR rapidly decreases to 13 dB, limiting the power increase of the laser system.

3.1.3. Off-Domain Sensors. In 2021, Yang et al. [24] designed an Yb-doped ring-distributed filter fibre with the refractive index distribution shown in Figure 4, where the Yb-doped core is enveloped by a high refractive index germanium-doped ring and the preform is ground into a "star" shape. The Yb-doped core is enveloped by a high refractive index germanium-doped ring, and the preform is ground into a "star shape" and coated with a low refractive index polymer to achieve cladding pumping. Figure 5 shows the mode distribution in the fibre, where the unique transverse distribution of the refractive index allows for a different mode distribution of signal light and "noise" such as Raman scattered light, which can be effectively attenuated by reducing the overlap factor between Raman light and signal light. In 2021, Pei et al. [25] introduced a bend-

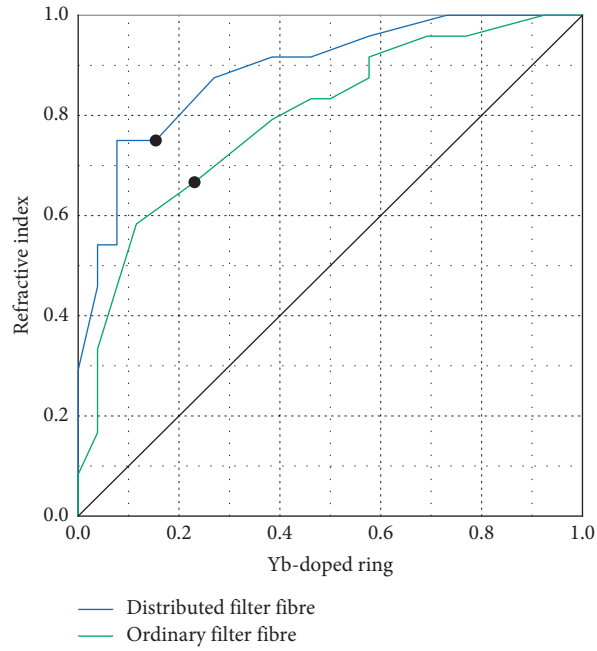


Figure 4: Yb-doped ring-distributed filter fibre with the refractive index distribution.

compensating cladding that allows the fibre to maintain a low gain over a long length, making it possible to use it for long-distance (>20 m) kW-class laser transmission.

3.1.4. Geometric Sensors. The reduction in the perceived length of the geometric sensor is based on the fact that reducing the active length by increasing the absorption coefficient of the gain fibre is an effective strategy for transmission efficiency such as SRS, which can usually be achieved by both increasing the doping concentration of rare earth ions and increasing the area ratio of the inner cladding of the core. However, increasing the doping concentration leads to an enhanced photon darkening effect in the gain fibre which affects the performance of the laser. Increasing the area ratio of the core cladding increases the core area on the one hand, i.e., large mode field fibre technology, and reduces the diameter of the cladding on the other. In 2018, Choi et al. [26] improved the cladding pump absorption by a factor of 1.5 by implanting a fluorine-doped low refractive index quartz rod (LCA) in the inner cladding of a circular double cladding fibre (the fibre structure is shown in Figure 6). In the optical fibre oscillator (OSC), the higher pump absorbance results in a 33% reduction in the effective action distance length compared to a conventional 20/400 μm double-clad fibre (DCF), while the SRS threshold is increased from 1.6 kW to 2.4 kW.

Compared to other transmission efficiency strategies, increasing the mode field area reduces the optical intensity on the one hand and shortens the fibre length on the other, thus providing excellent Raman suppression. High-beam quality laser outputs of 8 kW have been achieved with centrally doped fibres, while high-beam quality laser outputs of 5 kW have also been achieved with spindle fibres. The fact

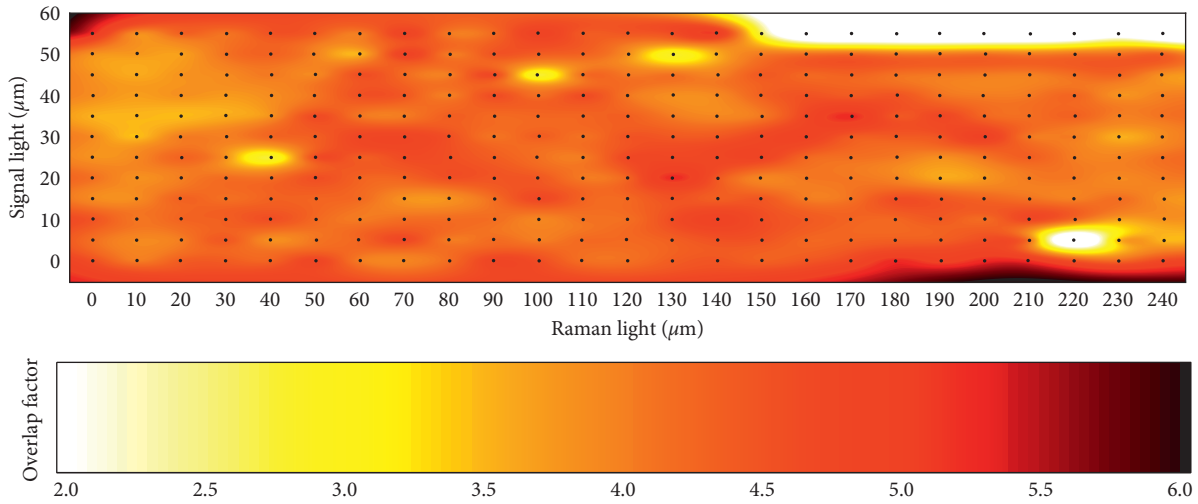


Figure 5: Mode distribution in the fibre with transverse distribution of the refractive index.

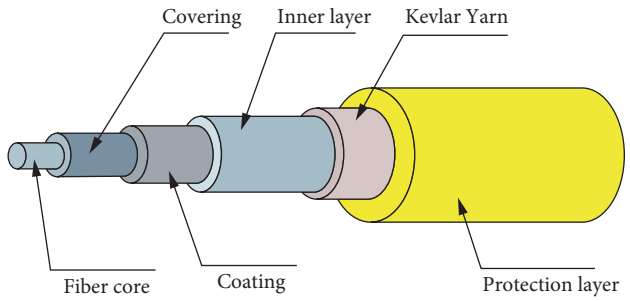


Figure 6: The fibre structure based on the geometric sensor signal enhancement technique.

that the various fibre designs are not opposed to each other means that a combination of design strategies is possible. By combining multiple strategies, such as introducing central partial doping into the spindle fibre design and improving large mode field fibre materials, SRS can be further suppressed and the output power of high-beam quality fibres can be increased.

3.2. SRS Transmission Characterisation Method. On the one hand, special optical fibre technology requires special fibre designs, which are relatively complex and expensive to prepare. On the other hand, fibres are an integral part of the laser system and the output laser performance is determined by all system components. Therefore, in addition to the use of special fibres, improving SRS transmission performance can also be achieved by optimising other system parameters (e.g., overall structure, pumping mode, and seed characteristics.) when building a high-power fibre laser system. As a common optical component, fibre gratings also demonstrate excellent Raman filtering properties.

3.2.1. Overall Structure. In order to suppress the SRS effect in kW-class multi-sensor signals, it is often necessary to use large mode field fibres (core diameter >20 μm) to reduce the fibre length and decrease the effective action distance. In 2017, Berto et al. [27] suggested that the effective action distance in dual-range amplification is twice as high as in single-range amplification, despite the fact that the fibre length is below 3 m. The intensity of SRS produced in a dual-range fibre amplifier with bidirectional pumping is comparable to that in a single-range amplifier with reverse pumping. In 2017, Wang et al. [28] investigated the SRS effect in a kilowatt cascade-pumped ytterbium-doped fibre amplifier. The model pointed out that although the Raman effect can be suppressed by lower doping concentrations and shorter fibre lengths, its effect on pump absorption and amplification efficiency should be considered. To avoid excessive pump light residuals, the fibre length should be greater than 50 m for pumping conditions with a central wavelength of 1018 nm.

3.2.2. Embedded with Multi-Sensor Enhancement Technology. The single oscillation cavity fibre laser system has the advantages of simple structural system and low production cost, and largely eliminates the damage to the laser system caused by material reflection. In order to increase the SRS threshold in a single oscillating cavity, the effect of the fibre Bragg grating (FBG), which is the core component of the oscillating cavity, on the SRS has been investigated. Lin showed that although the SRS can be suppressed by increasing the bandwidth of the OC-FBG, the increase in bandwidth will also increase the backward leakage power of the HR-FBG. In addition, the reflectivity of the OC-FBG in a single oscillating cavity structure also has an effect on the Raman light. The results show that the power of the Raman light decreases as the reflectivity of the OC-FBG decreases, but a low OC-FBG reflectivity leads to fluctuations in the output power due to the material reflecting back light into the resonant cavity. Analysis of transmission is based on filtered fibre optic gratings.

When Sun et al. [29] first used multiple long-period gratings (LPGs) in series to form a lumped filter, Raman fibre grating filters have gained widespread attention and developed rapidly due to their excellent Raman rejection ratio and relatively low insertion loss.

The principle of operation of a long-period grating is shown where Raman light of a specific wavelength is coupled and leaked to the cladding after the grating, and the corresponding transmission spectrum is shown in Figure 7. Earlier, Cheng et al. [30] wrote LPGs in a 10/125 μm fibre using a CO_2 laser and reduced the insertion loss of the 1030 nm signal light to 0.01 dB by a toe-cutting process, which made it possible to connect several gratings in series at the same time. In the experiment, the tandem of three long-period gratings doubled the SRS threshold and further additions to the number of gratings in tandem could be made to achieve better suppression. The resonance loss of the Raman light reached 26 dB, and the inscription of the long-period grating did not significantly affect the transmission loss of the signal light. The optimised 14/250 μm long-period grating was inserted into a MOPA laser system and achieved an SNR of 24 dB at 805 W output power.

In 2016, De et al. [31] first proposed the use of chirped tilted Bragg gratings to suppress SRS, as shown in Figure 8, where the chirped tilted Bragg grating suppresses Raman light by reflecting it backwards at an angle into the cladding compared to a long-period grating. Compared to long-period gratings, the spectrum of a chirped tilted Bragg grating is continuous and can be tuned by varying the tilt angle with relatively low sensitivity to the environment. However, the suppression principle of chirped tilted Bragg gratings allows (a) some of the reflected light to propagate backwards along the core, lowering the TMI threshold and causing non-linear effects such as four-wave mixing (FWM); (b) the light leaking from the tilted reflections to the cladding heats up the coating of the CTFBG, necessitating the addition of a cladding photo stripper (CPS) on the grating surface at high power, making the grating more difficult to fabricate, increasing the process difficulty, and reducing the lifetime of the grating. All-fibre single oscillators, kW-level LD-pumped fibre amplifiers, and kW-level cascaded fibre amplifiers have been implemented using chirped tilted Bragg gratings. The output spectra of the cascaded fibre amplifier with 0, 1, and 2 CTFBGs based on the multi-sensor signal are shown in Figure 8, and it can be seen that the SRS is significantly suppressed as the number of inserted chirped tilted Bragg gratings increases.

3.2.3. Analysis of Optical Fibre Transmission Based on Pumping Methods. The common pumping methods used in high-power multi-sensor signal systems are end-face pumping and distributed lateral pumping (DSCCP). In the end-face pumping method, the effective propagation distance of the high-power signal light is shorter (at the end of the Yb-doped fibre) in the backward pumping strategy, thus significantly reducing the distance between the signal light and Raman light and suppressing the SRS effect. Although the SRS threshold is lower for distributed lateral pumping compared to end-face pumping for the same fibre size, this pumping method results in a more uniform thermal load in the fibre and therefore significantly mitigates the thermal effects in the fibre. High-beam quality kW-class fibre oscillators and fibre amplifiers using distributed lateral pumping have also been implemented.

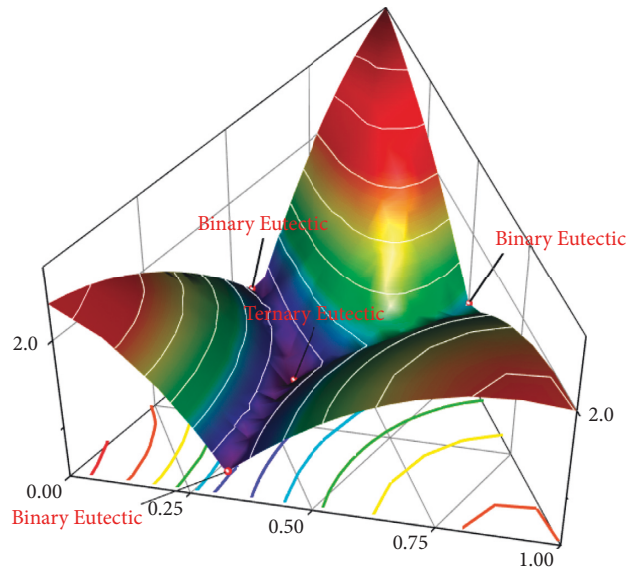


Figure 7: Raman light of a specific wavelength relationship of the corresponding transmission spectrum.

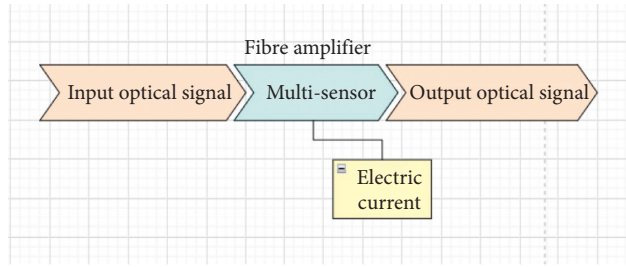


Figure 8: The cascaded fibre amplifier with CTFBGs based on the multi-sensor signal.

3.2.4. Transmission Analysis Based on Seed Characteristics and Self-Pulsation. In 2019, Koushki et al. [32] investigated the evolution of SRS with seed optical power in an Yb-doped fibre amplifier. The results showed that the SRS threshold was inversely proportional to the injected seed optical power. However, due to the limited pumping power, the reduction of the seed optical power decreases the output optical power. At the same time, the stability of the fibre optic amplifier is reduced by spontaneous radiation amplification (ASE) when it is operated at high power. In addition, studies have shown that ASE can also act as Raman seed light to cause Raman amplification, leading to a reduction in the SRS threshold. Lang et al. [33] investigated the effect of Raman scattering noise in the seed light on the SRS and gave a formula for calculating the Raman threshold of a high-power amplifier when considering Raman scattering noise in the seed light, showing that when the Raman scattering noise in the seed light is greater than 10–8 W, it plays an important role in the Raman threshold of the high-power amplifier.

Shi et al. [34] showed that in high-power continuous fibre lasers, relaxation oscillations can cause self-pulses with high peaks, which may trigger a series of non-linear effects such as SRS when their peak power reaches the non-linear threshold. According to Shi, the generation of self-pulses significantly

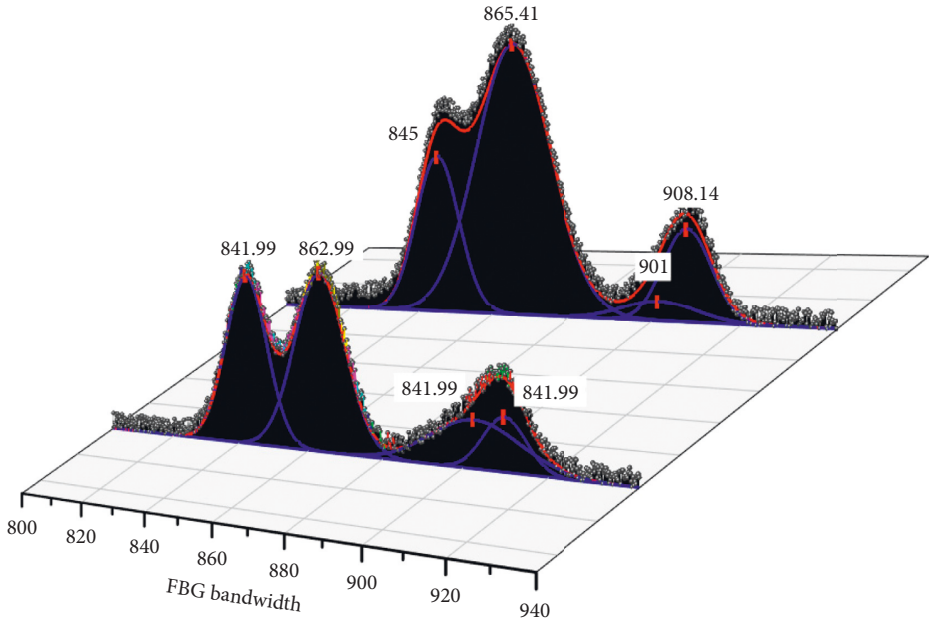


Figure 9: Corresponding simulation with combination of FBG bandwidth.

lowers the SRS threshold and mainly increases the forward Stokes optical power, so effective strategies are needed to suppress them. The commonly used improving SRS transmission performance includes using a single-frequency seed source with phase modulation, increasing the FBG bandwidth of the seed oscillation stage, and increasing the length of the energy transfer fibre between the seed and amplification stages. Figure 9 shows the corresponding simulation results, which show that the combination of FBG bandwidth increase and energy transfer fibre lengthening strategy can further suppress SRS. The combination of the two strategies results in a high SNR of ~50 dB at ~1 kW output power.

3.2.5. Analysis of Transmission Based on Four-Wave Mixing. In high-power multi-sensor signal systems, when the phase-matching conditions are met, photons at frequencies $\omega 1$ and $\omega 2$ pass through the parametric process to produce Stokes and anti-Stokes photons at frequencies $\omega 3$ and $\omega 4$. This process is called four-wave mixing (FWM) when the four frequencies are in the same fibre transverse mode in a few-mode fibre, or inter-mode four-wave mixing (IM-FWM) if they correspond to different fibre transverse modes. Shatarah et al. [35] showed that the normalised phase mismatch ΔB changes the effective Raman gain coefficient when FWM occurs

$$\begin{aligned} \Delta B &= \frac{(2\beta_L - \beta_S - \beta_a)}{g_0}, \\ \frac{g_{\text{eff}}}{g_0} &= \text{Re}\left[\sqrt{2(1 - f + f\chi_0) - \Delta B}\sqrt{\Delta B}\right], \\ g_{\text{eff}} &= \text{Re}\left[\sqrt{2(1 - f + f\chi_0) - \Delta B}\sqrt{\Delta B}\right] g_0, \\ X_0 &= 1.38i, \end{aligned} \tag{3}$$

where β_L, β_S, and β_a correspond to the propagation constants of the signal, Stokes, and Raman light; g_0 is the small-signal Raman gain coefficient; the value of the parameter f is 0.18; and $\chi_0 = 1.38i$ corresponds to the peak of the imaginary part of the Raman non-linear polarisation rate. The simulations suggest that the coupling between FMW and SRS may lead to a significant enhancement of SRS (~1.8 times), and therefore, a fibre with a suitable dispersion value needs to be selected to suppress SRS by optimising the parameter ΔB.

Gorelik et al. [36] theoretically analysed the effect of FWM and IM-FWM on RFL. Through the control variables method, the model states that FWM leads to a second-order Stokes optical power enhancement, which reduces the Raman threshold by 50%. Meanwhile, the effect of IM-FWM on RFL is explored by varying the power share of the LP11 mode in the seed light while ignoring FWM, and the results show that the Raman threshold decreases significantly as the power share of the LP11 mode increases, and IM-FWM has a stronger effect on the Raman threshold than FWM when the share is higher than 10^{-3}. The FWM can be effectively suppressed by adopting a time-domain stable pump source, while the IM-FWM can be suppressed by introducing higher-order mode losses.

The theoretical optimisation of the system parameters affecting the SRS threshold is summarised. It should be noted that, on the one hand, multi-sensor enhancement techniques cause changes in many important parameters that are often coupled with each other in affecting the SRS, and they in turn determine other characteristics of the laser system. Therefore, when designing the system, all factors should be taken into account in the context of the requirements and improving SRS transmission performance should not be pursued blindly.

4. Conclusion

As one of the main limiting factors for the single-link power increase of fibre laser systems, SRS increases the risk of system damage in addition to the reduction of system power and efficiency by backward Stokes light. In addition, a new phenomenon of mode degradation due to SRS has recently been observed in few-mode fibres. As a new physical characterisation of SRS in few-mode fibres, the available experimental results do not fully match the theoretical explanation and need further investigation. The study of strategies to enhance the transmission efficiency associated with combing SRS can, on the one hand, improve the power output of existing high-power, high-beam quality fibre laser systems; on the other hand, the dynamic adjustment of SRS intensity through multi-sensor means can help to deepen the understanding of the analytical methods for SRS transmission characteristics. As a result of continuous efforts by researchers, various Raman-enhanced transmission techniques have been developed and are now divided into two main categories: special fibre design techniques and overall system structure optimisation. The introduction of multi-sensor enhancement techniques in special fibre design techniques has shown excellent capabilities for improving transmission characteristics and high-power applications. In addition to the use of fibre gratings as Raman filtering components, the optimisation of the system structure is based on the influence of various sensor enhancement techniques (e.g., ASE, self-pulsing, and FWM) on the Raman light and the optimisation of the relevant parameters to achieve the upgrading of the transmission characteristics. In the next step, we will continue to explore the physical mechanism of SRS transmission characteristics optimisation and the corresponding analysis methods in conjunction with SRS transmission techniques based on multi-sensor enhancement techniques and mode decomposition techniques. The feasibility study of introducing new multi-sensor enhancement techniques will also be carried out to further enhance the power level of existing high-power, high-beam quality fibre laser systems to provide a reference.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] F. Yang, Y. Zhao, Y. Qi, and Y. Z. H. L. W. Tan, "Towards label-free distributed fiber hydrogen sensor with stimulated Raman spectroscopy," *Optics Express*, vol. 27, no. 9, pp. 12869–12882, 2019.

[2] M. Lawrence and J. A. Dionne, "Nanoscale nonreciprocity via photon-spin-polarized stimulated Raman scattering[J]," *Nature Communications*, vol. 10, no. 1, pp. 1–8, 2019.

[3] K. Jiao, J. Shu, H. Shen, and Z. F. R. Guan, "Fabrication of kW-level chirped and tilted fiber Bragg gratings and filtering of stimulated Raman scattering in high-power CW oscillators," *High Power Laser Science and Engineering*, vol. 7, p. e31, 2019.

[4] R. Ranjan, M. Indolfi, M. A. Ferrara, and L. Sirleto, "Implementation of a nonlinear microscope based on stimulated Raman scattering," *Journal of Visualized Experiments*, no. 149, Article ID e59614, 2019.

[5] M. Wang, Y. Zhang, Z. Wang, and J. J. J. X. X. Sun, "Fabrication of chirped and tilted fiber Bragg gratings and suppression of stimulated Raman scattering in fiber amplifiers," *Optics Express*, vol. 25, no. 2, pp. 1529–1534, 2017.

[6] K. Jiao, H. Shen, Z. Guan, and F. R. Yang, "Suppressing stimulated Raman scattering in kW-level continuous-wave MOPA fiber laser based on long-period fiber gratings," *Optics Express*, vol. 28, no. 5, pp. 6048–6063, 2020.

[7] M. Heck, V. Bock, R. G. Krämer et al., "Mitigation of stimulated Raman scattering in high powermulti-sensor signals using transmission gratings[C]//Fiber Lasers XV: technology and Systems," *International Society for Optics and Photonics*, vol. 10512, Article ID 105121I, 2018.

[8] D. Lioe, K. Mars, S. Kawahito, and K. K. T. M. Yasutomi, "A stimulated Raman scattering CMOS pixel using a high-speed charge modulator and lock-in amplifier," *Sensors*, vol. 16, no. 4, p. 532, 2016.

[9] K. Mars, D. X. Lioe, S. Kawahito, and K. K. T. M. Yasutomi, "Label-Free biomedical imaging using high-speed lock-in pixel sensor for stimulated Raman scattering," *Sensors*, vol. 17, no. 11, p. 2581, 2017.

[10] F. Yang and W. Jin, "All-fiber hydrogen sensor based on stimulated Raman gain spectroscopy with a 1550 nm hollow-core fiber[C]," in *Proceedings of the 2017 25th Optical Fiber Sensors Conference (OFS)*, pp. 1–4, IEEE, Jeju, South Korea, 24 April 2017.

[11] H. J. Lee, K.-C. Huang, G. Mei, and C. N. W. J. K. J. J.-X. Zong, "Electronic preresonance stimulated Raman scattering imaging of red-shifted proteorhodopsins: toward quantitation of the membrane potential," *The Journal of Physical Chemistry Letters*, vol. 10, no. 15, pp. 4374–4381, 2019.

[12] Y. Li, B. Shen, S. Li, Y. Zhao, J. Qu, and L. Liu, "Review of stimulated Raman scattering microscopy techniques and applications in the biosciences," *Advanced Biology*, vol. 5, no. 1, Article ID 2000184, 2021.

[13] Z. Amira, B. Mohamed, and E. Tahar, "Monitoring of temperature in distributed optical sensor: Raman and Brillouin spectrum," *Optik*, vol. 127, no. 8, pp. 4162–4166, 2016.

[14] M. Mehrabi, H. Beyranvand, and M. J. Emadi, "Multi-band elastic optical networks: inter-channel stimulated Raman scattering-aware routing, modulation level and spectrum assignment," *Journal of Lightwave Technology*, vol. 39, no. 11, pp. 3360–3370, 2021.

[15] Z. Wang, W. Yu, J. Tian, and T. D. Q. P. M. Qi, "5.1 kW tandem-pumped fiber amplifier seeded by random fiber laser with high suppression of stimulated Raman scattering," *IEEE Journal of Quantum Electronics*, vol. 57, no. 2, pp. 1–9, 2021.

[16] F. Yang, Y. Zhao, Y. Qi et al., "Label-free distributed hydrogen sensing with stimulated Raman scattering in hollow-core fibers[C]," in *Proceedings of the Optical Fiber Sensors*, Optical Society of America, Lausanne Switzerland, 24 September 2018.

[17] B. Manifold, E. Thomas, A. T. Francis, and A. H. D. Hill, "Denoising of stimulated Raman scattering microscopy images via deep learning," *Biomedical Optics Express*, vol. 10, no. 8, pp. 3860–3874, 2019.

[18] Y. Ozeki, T. Asai, J. Shou, and H. Yoshimi, "Multicolor stimulated Raman scattering microscopy with fast

wavelength-tunable Yb fiber laser," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 25, no. 1, pp. 1–11, 2019.

[19] V. S. Gorelik, P. P. Sverbil, V. V. Filatov, and D. G. T. S. H. Bi, "Transmission spectra of one-dimensional porous alumina photonic crystals," *Photonics and Nanostructures - Fundamentals and Applications*, vol. 32, pp. 6–10, 2018.

[20] F. Hu, L. Shi, and W. Min, "Biological imaging of chemical bonds by stimulated Raman scattering microscopy," *Nature Methods*, vol. 16, no. 9, pp. 830–842, 2019.

[21] S. Yan, S. Cui, K. Ke, and B. X. S. P. Zhao, "Hyperspectral stimulated Raman scattering microscopy unravels aberrant accumulation of saturated fat in human liver cancer," *Analytical Chemistry*, vol. 90, no. 11, pp. 6362–6366, 2018.

[22] X. Hu, W. Chen, M. Chen, and Z. Meng, "Experimental observation of the competition between stimulated Brillouin scattering, modulation instability and stimulated Raman scattering in long single mode fiber," *Journal of Optics*, vol. 18, no. 8, p. 085501, Article ID 085501, 2016.

[23] G. Sciortino, A. Ragni, A. De la Cadena, and M. G. D. G. Sampietro, "Four-channel differential lock-in amplifiers with autobalancing network for stimulated Raman spectroscopy," *IEEE Journal of Solid-State Circuits*, vol. 56, no. 6, pp. 1859–1870, 2021.

[24] Y. Yang, Y. Yang, Z. Liu, and L. S. X. Z. M. Guo, "Microcalcification-based tumor malignancy evaluation in fresh breast biopsies with hyperspectral stimulated Raman scattering," *Analytical Chemistry*, vol. 93, no. 15, pp. 6223–6231, 2021.

[25] W. Pei, H. Li, W. Huang, and M. Z. Wang, "All-fiber tunable pulsed 1.7 μm fiber lasers based on stimulated Raman scattering of hydrogen molecules in hollow-core fibers," *Molecules*, vol. 26, no. 15, p. 4561, 2021.

[26] D. S. Choi, B. J. Rao, D. Kim, and S.-H. H. M. Shim, "Selective suppression of CARS signal with three-beam competing stimulated Raman scattering processes," *Physical Chemistry Chemical Physics*, vol. 20, no. 25, pp. 17156–17170, 2018.

[27] P. Berto, C. Scotté, F. Galland, and H. H. B. Rigneault, "Programmable single-pixel-based broadband stimulated Raman scattering," *Optics letters*, vol. 42, no. 9, pp. 1696–1699, 2017.

[28] X. Wang, H. Qi, Y. Li, and F. H. F. Y. Z. X. X. Yu, "Synthesis and characterization of new Sr3(BO3)2 crystal for stimulated Raman scattering applications," *Crystals*, vol. 7, no. 5, p. 125, 2017.

[29] X. Sun, J. Li, and M. Hines, "SNR improvement in a Raman based distributed temperature sensing system using a stimulated Raman scattering filter[C]//Fiber Optic Sensors and Applications XIV," *SPIEL*, vol. 10208, pp. 69–74, 2017.

[30] T. Cheng, W. Gao, X. Xue et al., "Experimental observation of stimulated Raman scattering in a fluoride fiber[C]//Optical Components and Materials XIV," *SPIEL*, vol. 10100, pp. 363–368, 2017.

[31] L. De Xing, "A study on CMOS image sensors for stimulated Raman scattering using high-speed lateral electric field charge modulators[J]," *Dr. Diss*, 2016.

[32] E. Koushki and B. Maleki, "Induced photoacoustic gratings due to Raman scattering in organic components," *Dyes and Pigments*, vol. 164, pp. 82–86, 2019.

[33] X. Lang and K. Welsher, "Mapping solvation heterogeneity in live cells by hyperspectral stimulated Raman scattering microscopy," *The Journal of Chemical Physics*, vol. 152, no. 17, p. 174201, Article ID 174201, 2020.

[34] L. Shi, A. A. Fung, and A. Zhou, "Advances in stimulated Raman scattering imaging for tissues and animals," *Quantitative Imaging in Medicine and Surgery*, vol. 11, no. 3, pp. 1078–1101, 2020.

[35] I. S. M. Shatarah and R. Olbrycht, "Distributed temperature sensing in optical fibers based on Raman scattering: theory and applications[J]," *Measurement Automation Monitoring*, vol. 63, 2017.

[36] V. S. Gorelik, D. Bi, Y. P. Voinov, and A. I. V. A. A. I. Vodchits, "Spontaneous and stimulated Raman scattering in protium and deuterium water," *Optics and Spectroscopy*, vol. 126, no. 6, pp. 687–692, 2019.

Hindawi
Scientific Programming
Volume 2022, Article ID 9662292, 10 pages
https://doi.org/10.1155/2022/9662292

Research Article

Application of 3D Visualisation Video Technology for the Modern Business Management

Ying Li[1] and Yangxiang Ye[2]

[1]*Jinneng Holding Group Party School, Datong, Shanxi 037000, China*
[2]*Zhejiang Industry & Trade Vocational College, Wenzhou, Zhejiang 325000, China*

Correspondence should be addressed to Yangxiang Ye; yeyangxiang@zjitc.edu.cn

Received 17 March 2022; Revised 11 April 2022; Accepted 15 April 2022; Published 28 April 2022

Academic Editor: Baiyuan Ding

Nowadays, identifying problems is often more difficult and important than solving them in the complex business environment. Especially in modern enterprises, with more staff and complex working environment, it is more difficult to find the problems buried under the "iceberg." Thus the task of managers becomes how to find problems and create an environment conducive to management, so that problems naturally emerge. Based on the modern enterprise management as the research object, the information technology and 3D visualisation video technology in the application of modern enterprise management were carried out. Then it explains the concept and theory of 3D visualisation video technology, analyses the feasibility of applying visualisation video management in modern enterprises and, in particular, explores the existing problems of modern enterprise management, and comes up with methods and specific measures for implementing visualisation video management in large modern enterprises. It provides an innovative approach to the management of modern enterprises in China when facing new challenges and new turning points in the context of the financial crisis.

1. Introduction

The development of computer network technology and changes in the way information is transmitted have had a huge impact on the production, operation, and management models of modern enterprises [1, 2]. With the increasing competition in the market, the old mode of company operation (cost + profit = price) has been gradually replaced by a new mode of company operation (market-driven price – controllable costs for the producer = profit) [3]. This means that it is now difficult for companies to control the market price to achieve the desired profit. This challenging change has led companies to reduce costs, eliminate waste, and maximise profits by improving management [4]. In other words, all activities must be focused on how to achieve the same value for the customer at the lowest possible cost [5, 6].

But in order to eliminate these unnecessary wastes, increase efficiency and add value to customers. The chief problem is to identify these wastes. In today's complex business environment, identifying problems is often more difficult and more important than solving them [7, 8]. This is especially true in modern companies, where the number of staff and the complexity of the working environment make it more difficult to find the problems buried underneath the "iceberg" [9]. The task of the manager then becomes how to identify problems and how to create an environment, which is conducive to management and allows problems to come to the fore naturally.

As a result, the manager's task now becomes how to make anomalies visible automatically through effective management methods [10, 11]. But this is not easy to do, as there are often many complicated processes involved in the complex activities of a production site. The challenge for managers and academics alike is how to make problems automatically visible by using a number of methods [12].

This paper attempts to investigate the feasibility of using visualisation video in modern enterprises, by adopting the theories related to visualisation video technology to identify potential problems, find out solutions, and improve economic performance in modern enterprises.

2. Basic Theories Related to Visualisation Technology

The visual technique approach was first systematically introduced by Professor Yoshijiro Sawada of Kurume University in his book "Visual Management of Workshop Management." The method was first used at the Japanese company Toyota Motor Corporation) [13, 14]. After its introduction at Toyota, the company's efficiency was greatly improved, waste was significantly reduced, and products were delivered on time and with excellent quality. Many Japanese companies have since adopted the visual technology approach [15, 16]. For example, KODAK has set up a dedicated visualisation video team to promote visualisation video technology, which has yielded significant benefits. HONEYWELL has also achieved good results in the field of aviation maintenance using visual technology, which has made a great contribution to aviation safety. Many other major companies such as DELL, GE, and HP are also implementing visualisation video [17]. In view of the success of these large companies in visualisation video management, many small- and medium-sized enterprises are also trying to promote visualisation video management [18]. The image below shows a 3D scan of DELL's face visualisation technology, as illustrated as Figure 1.

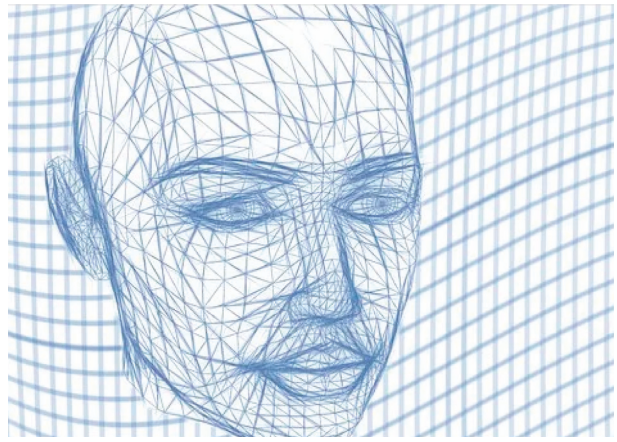


Figure 1: Technical 3D scan of DELL's face visualisation.

2.1. The Concept and Content of Visualisation Video. Visual management is a systematic way of enabling all members of a work area to see, understand, and appreciate the situation in the area (both physical and virtual) through simple visualisation video and to make continuous improvements [19]. This is also called visual communication, where the work area is clear to the human eye. It guarantees self-explanation, self-regulation, and continuous improvement of the work area. Visual management can also be interpreted as the use of IT systems to enable managers to effectively grasp corporate information and to achieve transparency and visualisation video in management, so that the effects of management can permeate all aspects of the company's human resources, supply chain, and customer management. Using visual and colour information to organise on-site production activities is one way to increase productivity. Visual management uses visual signals as a basic tool and the basic principle of openness to make the requirements and intentions of the manager as visible as possible, thereby promoting self-management and self-control [20]. Visual management is therefore a form of management characterised by openness and visual display.

Visualisation video is already well established in business management and is used at four levels: in internal processes, in learning and growth, in customer management, and in finance. Focused on modern enterprises with numerous personnel and chaotic production environment, visual video management is introduced in the internal management process of modern enterprises, and on-site management through visual video can improve management efficiency. To show the efficiency advantages of visualisation video technology in modern business, it is compared with traditional technology. Figure 2 shows the comparison between the use of visualisation video technology and traditional technology in processing efficiency changes with the development of modern companies. As can be seen from the figure, with the continuous development of modern enterprises, the amount of information that needs to be processed increases dramatically, resulting in traditional technologies first increasing and then gradually decreasing, while visualisation video screen technology performs strongly in the context of large information.

2.2. Classification of Visualisation Video. Visualisation video can be divided into three categories according to its scope of application: scientific computing visualisation video, digital visualisation video, and information visualisation video. Figure 3 illustrates the relationship curves among scientific computing visualisation video, digital visualisation video, and information visualisation video in terms of processing data flow and 3D processing intensity. From the figure, it can be seen that there is a pattern of change in the peaks of the three in increasing order. Scientific computing visualisation video refers to the theory, method, and technology of using computer graphics and image processing technology to convert data and calculation results generated in the process of scientific computing into graphics or images and then display them on the screen and to process them interactively. Information visualisation video includes operations research and related subjects, focusing more on the visual video of abstract information. The main areas are visualisation video of hierarchical information structures, visualisation video of multidimensional data structures, visualisation video of network structures, operational status, browsing history, and network users [21, 22].

3. Enterprise Management Objects and the Advantages of Applying Visual Management

In a modern company, the focus of management is on managing exceptions. In modern enterprises, exceptions can be divided into two categories: firstly, what should be done is not done and what should not happen [23]; secondly,

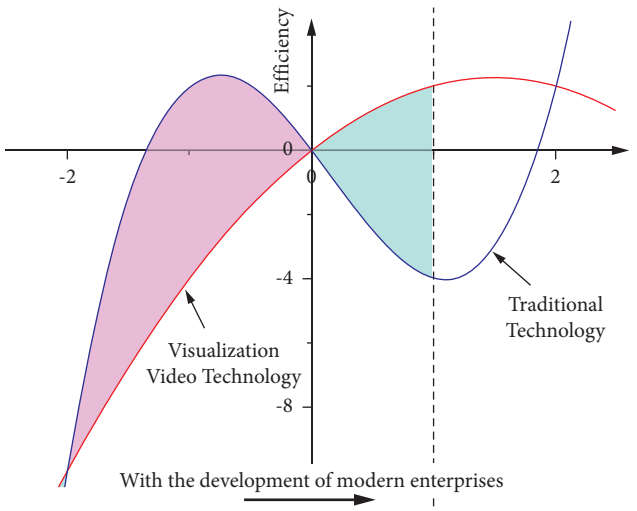


Figure 2: Comparison between the use of visualisation video technology and traditional technology.

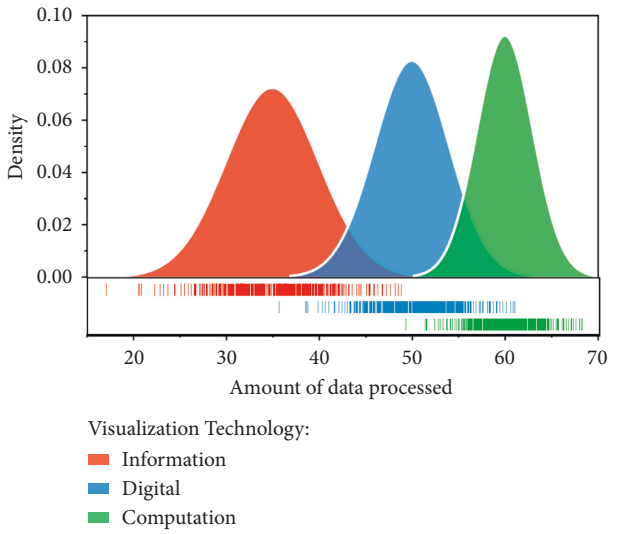


Figure 3: The relationship curves among scientific computing visualisation video, digital visualisation video, and information visualisation video in terms of processing data flow and 3D processing intensity.

although what should be done is done and what should happen happens, it does not occur according to the process or the prescribed way of working, even without processes to follow, and if a replacement of the staff, the scenario, or the customer occurs, problems will arise. For example, there are no work instructions for a particular process and workers operate according to their own experience, which results in the following: if a skilled worker operates the process, the product is guaranteed to be satisfactory; if that worker leaves and is replaced by a new worker, a large number of defective products will be produced [24, 25]. Both of these things lead to abnormal production waste. Therefore, as a manager, you need to be able to detect anomalies in the operation of the company, and when you find anomalies, find solutions and organise implementation, which can be expressed in the following equation:

$$E_{cs} = \frac{(A_P E_P + A_S E_S)}{A}, \quad E_{cs} = mE_P + (1-m)E_S. \tag{1}$$

In this equation, E_{cs} and E_s represent total and partial anomalies in the course of a company's operations, respectively; A_p and A_s represent total and partial abnormal production waste respectively; m represents degree of visualisation.

In the production industry, large companies can afford to spend a lot of money on a range of management software or carry out a range of consultancy activities to visualise their management. Modern companies, however, are mainly small- and medium-sized enterprises that do not have the financial resources and energy to create the conditions for costly, scientifically calculated visualisation video management [26]. In fact, visualisation video itself is an art that solves the problem of how to express the unique perspective of visualisation video to build modern enterprises. Management processes will help enterprises to identify and solve internal problems in a timely manner, reduce costs, inventory, and waste, improve performance, and increase customer and employee satisfaction through image information visualisation video management. Figure 4 illustrates the schematic diagram of image information visualisation video technology to improve customer and satisfaction.

The benefits of visualisation video for modern companies at the production management level can be seen in the following areas.

3.1. Visual Management can Improve Efficiency. Visual management is intuitive and conducive to improving efficiency. On-site managers organise and direct production, in essence releasing all kinds of information. The process of receiving information and then acting on it is the process by which operators carry out their production work in an orderly manner. The production system runs at high speed, which requires the information to be transmitted and processed quickly and accurately under the condition of machine production; otherwise it will cause huge direct economic losses. If information relating to each worker had to be communicated directly by the manager, there would be a relatively large number of managers on the production site of a modern company with hundreds or thousands of workers. Visualisation video offers a shortcut to this problem [27]. By far the most common sensory organs used by operators and other managers to receive information are the eyes, ears, and nerve endings, where visual sense is used most frequently. Visual signals, sent out by instruments, television, signals, signs, diagrams, etc., are easy to be read and recognised. Where possible, visual signals can be used to transmit information quickly and accurately, so that production can be organised effectively without the need for on-site management. This relationship can be expressed by the following mathematical equation:

$$\alpha = K - \frac{2}{3}G. \tag{2}$$

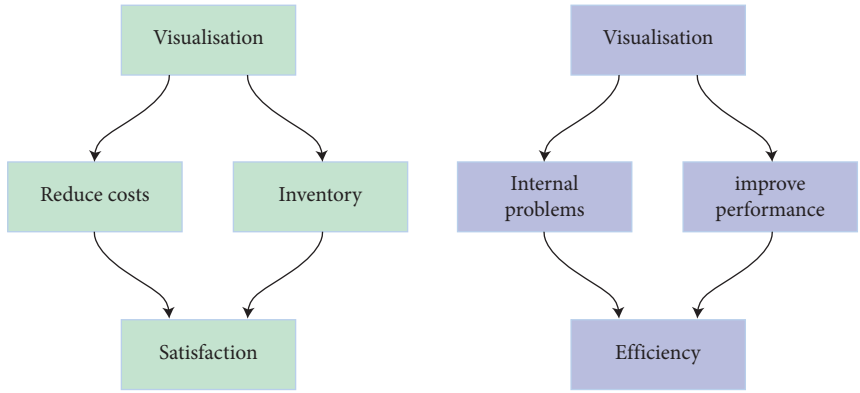


Figure 4: Schematic diagram of image information visualisation video technology to improve customer and employee satisfaction.

In this equation, α represents coefficient of on-site management; K and G represent intensity and frequency of visual signals, respectively.

Visual management is highly transparent and facilitates mutual supervision and motivation of the personnel on site. It is easy to see what to do, how to do it, how much to do it, when to do it, and where to do it, which is conducive to tacit cooperation and mutual supervision and makes it less easy to hide any breaches of labour discipline [28, 29]. For example, if workers are required to wear different uniforms and caps according to the characteristics of different workshops and types of work, it is easy to put those who leave their jobs without permission or chat on the job in full view of the public, so that they can exercise self-restraint and gradually develop good habits. For example, some countries (data from China, India, the United States, and Japan) have implemented a plaque system for enterprises, in which units are assessed and put up different coloured signs according to four grades: excellent, good, poor, and bad; individuals are assessed and wear different coloured armbands for those who are orderly and qualified. Figure 5 shows the rose chart comparing the popularity of visualisation video screen technology in the production and sales sectors in China, India, the United States, and Japan. As can be seen from the figure, sales departments in the four countries are greater than traditional production departments in the application of visualisation technology. In this way, visual management can play the role of encouraging the advanced and spurring the backward. In short, the production of modern enterprises requires both strict management and the development of people's habits and abilities of self-management and self-control. Above all, visualisation video screen technology provides an effective and concrete way of doing this. The mathematical expression can be given by the following equations:

$$S_1 = S_p + \Delta,$$
$$S_p = \int_0^L \frac{p_p(z)}{E_p} \mathrm{d}z, \tag{3}$$

where S_1 and S_p represent excellent and good plaque system, respectively; Δ represents the degree of self-restraint.

3.2. Visualisation Video Is Good for Staff Management: Visual Management Facilitates. Visual video is conducive to employee management. Visual management helps to produce good physiological and psychological effects. The physical and technical aspects of improving production conditions and environment are often ignored at the expense of the physical, psychological, and social characteristics of on-site personnel. For example, instruments and meters used to control machinery and production processes must be available, which is an essential material condition to strengthen on-site management.

3.3. Visualisation Video Facilitates Performance Management. Visualisation video allows managers to understand at a glance the overall sales, new market development, new product promotion, etc., so that they can take timely action to correct problems or increase market share, etc.

4. Approach to Visual Management Implementation

Modern enterprises generally lack production knowledge training for their employees, who are not aware of the need to streamline production and are unable to identify problems in a timely manner [30]. Visualisation video can be a very important method of production management, as it allows for the standardisation of visual displays. It includes mark line, mark plate, and sign colours.

4.1. Signage. Signage as a means of production control is intuitive and easy to use. The signage is an effective means of controlling production operations so that each production step and each process is carried out in strict accordance with the quantity standards and so that overproduction and overstocking are eliminated. The use of signage needs to be adapted to the working conditions of the site and be simple and practical. Signage includes some of the following elements.

(1) Kanban. If a process breaks down or stops for other reasons, the operator can see the signage and stop production when the previous process does not need to supply work-in-progress. Kanban is therefore a means of transmitting information that serves as a reminder.

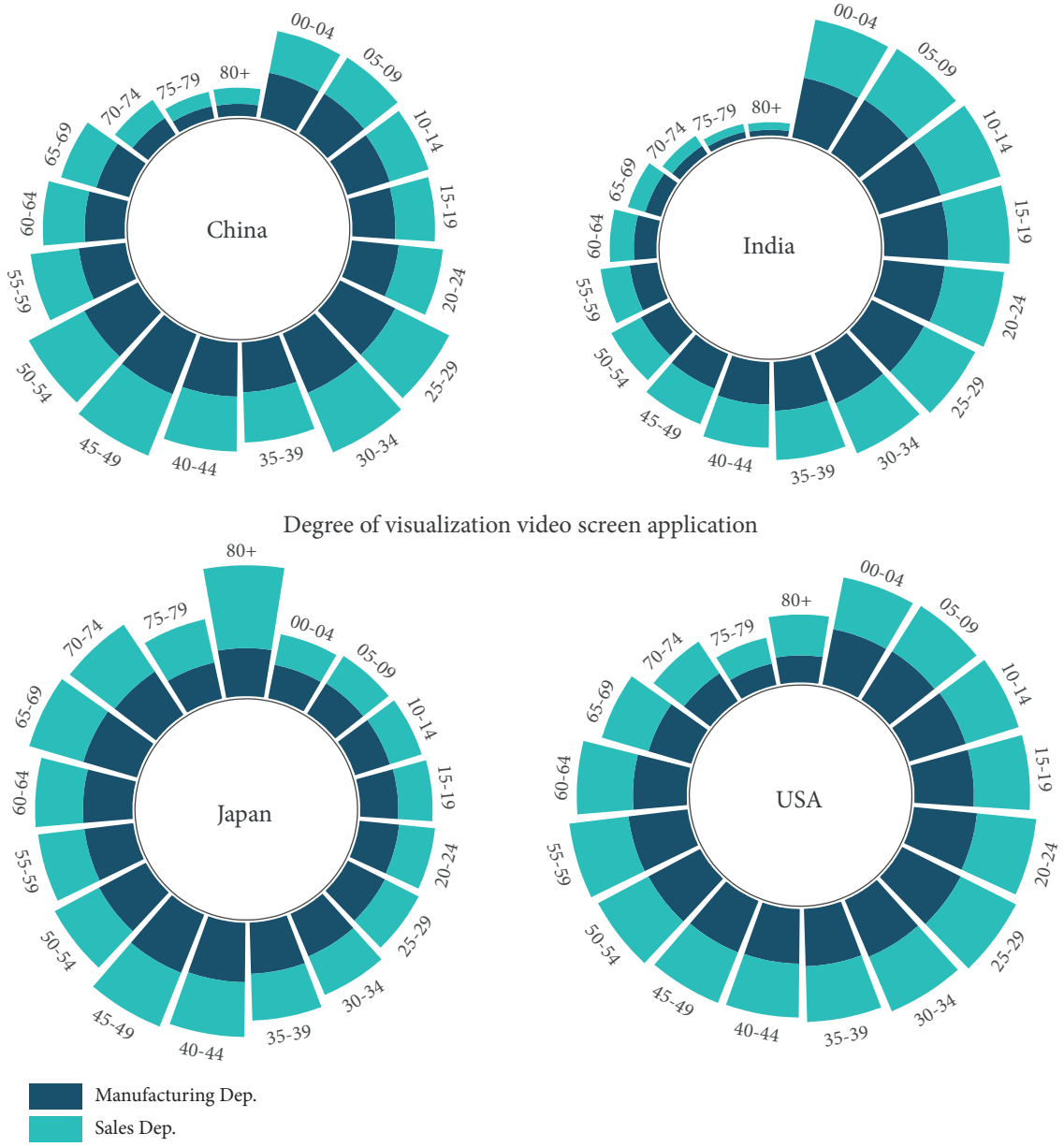


Figure 5: Rose chart comparing the popularity of visualisation video screen technology in the production and sales sectors in China, India, the United States, and Japan.

(2) Contact signs for each process. To minimise the loss of working hours and to improve the continuity of production, it is important to set up convenient and practical signals to communicate between the various production stages and types of work. For example, red lights can be installed on machinery and equipment, and workstation fault displays can be configured on assembly lines so that if a stoppage occurs, a signal can be sent out and a roving maintenance worker can see it and come to repair it in time.

(3) Standard signage. First, all kinds of work station apparatus, including boxes, boxes, trays, carts, etc., should be loaded according to the standard number of regulations, so that the operation, transport, and inspection personnel points are convenient and accurate. Secondly, the uniformity of the site personnel dress and the implementation of the tagging system site personnel dress can play a role not only in labour protection, in machine production conditions, but also in formalization and standardisation of one of the contents. Standard signage can not only reflect the excellent quality of the workforce, with access to identify different units, professions, and positions within the enterprise, but also bring about a certain psychological effect, such as a sense of belonging, a sense of honour, and a sense of responsibility. The signage system includes unit signage and

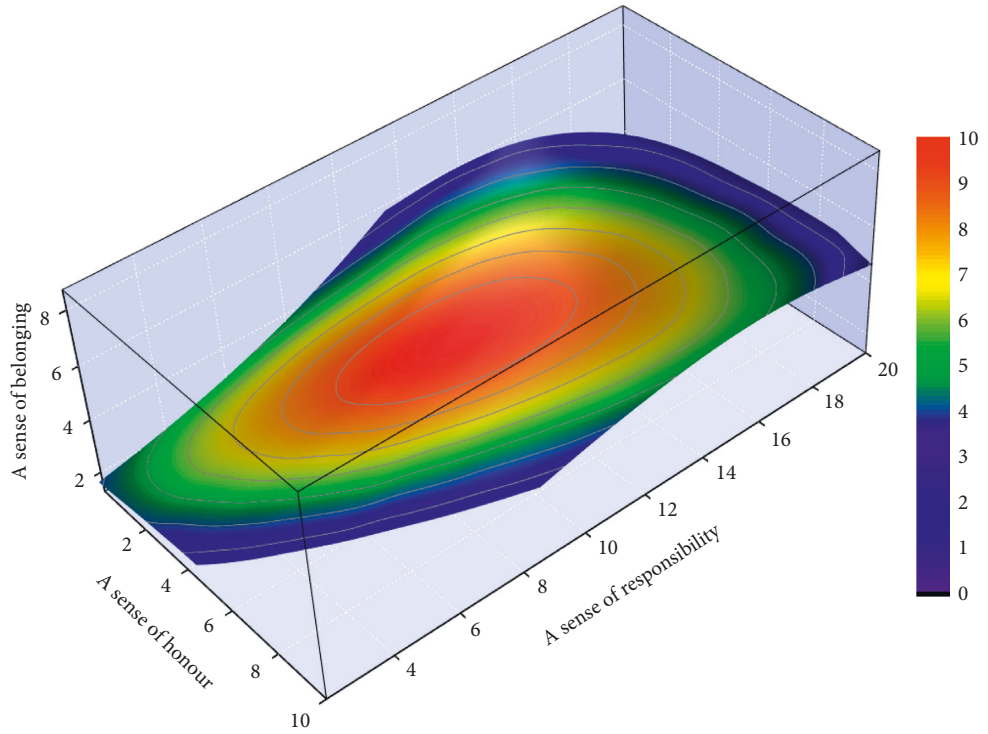


Figure 6: A visual three-dimensional representation of the sense of belonging, honour, and responsibility in modern enterprises.

individual signage. To visually describe the interrelationship among a sense of belonging, a sense of honour, and a sense of responsibility, Figure 6 shows a visual three-dimensional representation of the sense of belonging, honour, and responsibility in modern enterprises. The diagram indicates that there is a mutually reinforcing relationship between the three. The mathematical expression can be given by the following equations:

$$S = \frac{1}{M},$$

$$S = \frac{1}{M} + \frac{n}{\rho_w g L_p},$$

$$c = \frac{k}{S = 1/M + n/\rho_w g L_p}, \tag{4}$$

$$S = \frac{1}{M} + \frac{\alpha^2}{K + 4G/3},$$

where S represents plaque system; M and G represent coefficients of visual management, whose value depends on the degree of visualisation applied.

(4) Evaluation signage. In accordance with the internal inspection and evaluation system, the results of those assessment projects with a significant relationship with the achievement of the enterprise's strategic tasks and objectives are attached to the apartment in a visual and intuitive way, which can motivate the advanced units to go to the higher level and spur the backward units to catch up. Wearing signs for individuals, such as badges, badges, armbands, etc., serve a similar purpose with the dress code. In addition, it can also be combined with the appraisal to give people pressure and motivation to achieve the purpose of motivating people and promoting work.

4.2. Marker Lines. In addition to quantity control, there is also quality and cost control, which is also managed visually. For example, for quality control, at each quality control point (control), there should be a quality control chart, with clear critical lines drawn on the chart, so that quality fluctuations can be clearly displayed and abnormalities found and dealt with in a timely manner [31]. The workshop should make use of a board to publish the "Daily Statistics of Defective Products" and display the rejects of the day on a showcase for consultation and analysis by the relevant personnel to determine improvement measures to prevent recurrence.

4.3. Standard Colours. For colour standardisation management, colour is a visual signal commonly used in site management, and visualisation video management should be scientific, reasonable, and clever use of colour and achieving uniform standardisation management and not allowing arbitrary painting.

Danger signals are mostly red, while high-temperature workshops should be painted in light blue, blue-green, white, and other cool colours, resulting in a sense of refreshment and comfort; on the contrary, low-temperature workshops

Colour standardisation management			
	Phase 1	Phase 2	Phase 3
High temperature workshop	Light blue	Blue-green	White
Cryogenic workshop	Red	Orange	Yellow
Heat treatment equipment	Lead grey	Grey ink	Grey Purple

Figure 7: Illustration of the practical application of visualisation technology in the distribution of standard colours in different work scenarios.

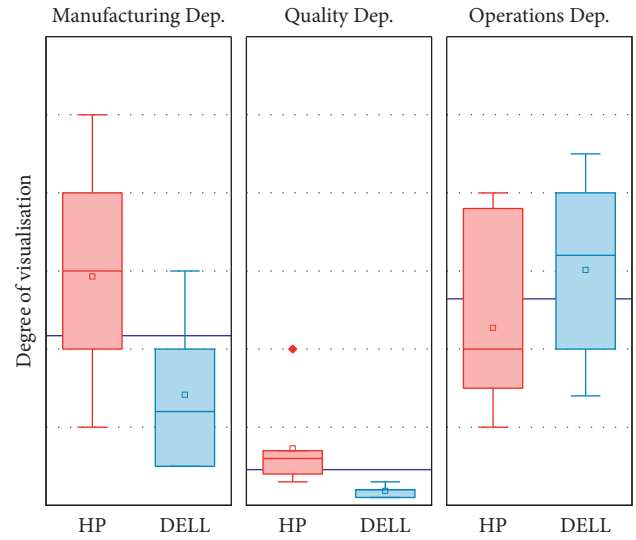


Figure 8: Groups box chart of degree of visualisation technology in different sectors in DELL and HP.

are suitable for red, orange, yellow, and other warm colours, bringing about warm feeling. Heat treatment equipment mostly adopts cold colours, such as lead grey, which can play a role in reducing the "psychological temperature." Since employees of furniture factories see warm wood colours all day long, wood processing equipment should be painted light green, relieving the irritation of the operator caused by the warm colour surroundings. The mathematical derivation for the relationship between the colour and sensation is given in equations (5)–(7). And Figure 7 shows illustration of the practical application of visualisation technology in the distribution of standard colours in different work scenarios.

$$p = \frac{P}{A}, \tag{5}$$

$$p_s = \mu_s p, \tag{6}$$

$$p_s = \frac{p}{1 + m(n-1)}, \tag{7}$$

where p and p_s represent psychological temperature; μ_s represents heat coefficient; m and n represent high- and low-temperature value.

5. Application Examples of Visualisation Management

A typical representative of modern enterprises, a modern company, was selected for this paper. During the survey, it was found that the enterprise suffered from rising costs. In recent years, with the impact of international oil price fluctuations, chemical fibre raw material prices rose, cotton prices rose faster, and since the second half of 2009, the average price of third-grade cotton market rose by more than 20%. Foreign trade market development difficulties increased pressure. The financial crisis has affected orders from overseas markets, with orders received by the company falling by as much as 70% year-on-year compared to the period before the financial crisis, leaving less and less room for profit. The backend yarn and fabric products only rose by about 5 percent [32]. In addition to the impact of adverse external factors, the company also suffers from low production efficiency, frequent delays in product delivery, and a large number of reworking due to quality problems. During a visit to the company, it was found that the production workshops were cluttered with machines and there were no obvious work zones; the responsibilities of the production process staff were unclear and problems were not detected and reported in a timely manner. The products were not placed in a standard way and the commissioner to sort out, check, and accept was required. Here are two examples of companies that illustrate the importance of the level of visualisation, illustrated as Figure 8. The figure shows groups box chart of degree of visualisation technology in different sectors in DELL and HP. As can be seen from the graph, HP has a greater degree of adoption of visualisation video screen technology than HP. The study therefore focuses on how to reduce production waste in the company, identify problems in all stages of production, and improve the value of production in four areas: visual area management, product management, equipment management, and production process management.

5.1. Visual Area Management

(1) Overall layout. Based on the TPS production model of lean production and the production process of the modern company, the layout of the entire factory is planned. The machines in each line should be closely linked and compactly arranged to reduce wasted space and to integrate the overall platform, moving

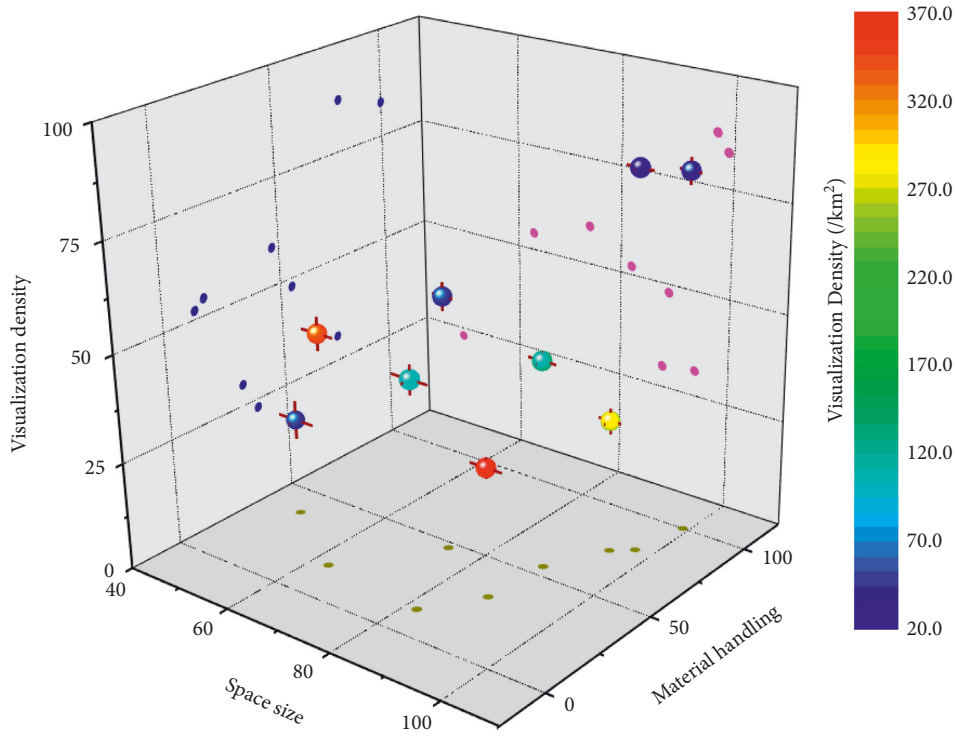


Figure 9: Three-dimensional scatter plot of interrelationship among space size, material handling, and visualisation density.

away from the previous independent and scattered state. This will result in continuous and compact processes, less space, less material handling, visualisation density, etc. In order to clarify the interrelationship among space size, material handling, and visualisation density, Figure 9 illustrates three-dimensional scatter plot of interrelationship among space size, material handling, and visualisation density. From the figure, it can be seen that the three show a positive correlation.

(2) Workshop layout. Similarly, for each workshop, it is also necessary to design a compact and easy-to-communicate humanised layout of the production process based on the theory of lean production in the TPS production model combined with the workshop production operation process.

(3) Posting. After the planning is complete, the workshop layout should be posted in each workshop, so that staff, management, new members, and customers can clearly understand it during the workshop production, so that every employee understands the layout of the process and can quickly find their own tools and items, reduce the time to find tools and items, reduce the time for training, let everyone participate in the process, and put forward ideas for improvement to make the area planning more rational.

(4) Colour coding. The marking of the areas is then done in two ways: firstly, the areas are marked with a specific text, such as a sign. The other requires visual colour management of the areas, with the following specifications.

5.2. Visual Product Management. There are two types of product identification: one is attribute identification and the other is status identification. Attribute labelling uses clear labels to indicate the type, date of manufacture, production batch, specification, quantity, and attribution of the product. This enables staff to know at a glance what the product is, so that they can easily do the right thing at once when shipping, collecting materials, etc., and also facilitates future traceability, etc. [33, 34]. In this modern company, red marks indicate defective products, yellow marks indicate products to be judged or detained, and green marks indicate qualified products. This enables any employee to know at a glance (unnecessary being a professional inspector) which product is qualified and which one is not. If an employee in the next process sees a red indicator attached to the product, he will not take it to the next process for further processing [35, 36]. This prevents further waste caused by defective products continuing to be produced. In the case of finished goods, the dispatcher can tell at a glance which products are non-conforming, thus preventing incorrect products from being delivered to the customer. This prevents the delivery of incorrect products to the customer.

5.3. Visual Equipment Management. The various states of the equipment are identified by the tagging of the equipment. A red tag means that the equipment is new or has undergone a special change and that it has not undergone "Q-TEST" or "F-test" since it was bought or after the change. The equipment cannot be used to officially mass-produce products. Even if the product is produced, it cannot be shipped, and if it is to be shipped, it must go through a strict

approval process; the yellow sign represents equipment for training practice, the product produced by this equipment cannot be shipped either, and the shipment must go through a full inspection: the green sign represents qualified equipment for normal use; the product produced by this equipment can be shipped according to the normal process; the "under repair" sign identifies that the equipment is under maintenance; the equipment is prohibited to be operated by employees, in order to avoid staff misuse resulting in casualties or economic losses.

6. Conclusion

This study focuses on the visualisation video and innovation management of modern enterprises in China, providing a theoretical basis and practical guidance for the implementation of visualisation video management, which is of great significance for the application of low-cost visualisation video management in modern enterprises. However, many aspects of this paper are worthy of further study, such as the specific visualisation video measures and methods for each production step and the evaluation of the effects of visualisation video implementation from a qualitative perspective.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by Planning Project of Philosophy and Social Sciences in Shanxi Province; Research on the integration of deepening the reform of state assets and promoting the transformation and development of Shanxi energy state-owned enterprises no. 2019B466.

References

[1] A. Chinnaswamy, A. Papa, L. Dezi, and A. Mattiacci, "Big data visualisation video, geographic information systems and decision making in healthcare management," *Management Decision*, vol. 57, no. 8, pp. 1937–1959, 2019.

[2] Y. Jiang, B. W. Ritchie, and P. Benckendorff, "Bibliometric visualisation: An application in tourism crisis and disaster management research," *Current Issues in Tourism*, vol. 22, no. 16, pp. 1925–1957, 2019.

[3] M. Weck, I. Humala, P. Tamminen, and F. A. F. Ferreira, "Knowledge management visualisation video in regional innovation system collaborative decision-making," *Management Decision*, vol. 60, no. 4, pp. 1017–1038, 2021.

[4] R. Roberts and R. Laramee, "Visualising business data: A survey," *Information*, vol. 9, no. 11, p. 285, 2018.

[5] T. Gerrish, K. Ruikar, M. Cook, M. Johnson, M. Phillip, and C. Lowry, "BIM application to building energy performance visualisation and management: Challenges and potential," *Energy and Buildings*, vol. 144, pp. 218–228, 2017.

[6] M. Abubakre, A. Fayoumi, and I. Eleburuike, "Implementing process improvement initiative: The role of visualisation video and standardisation methods," *Business Process Management Journal*, vol. 27, no. 3, pp. 965–986, 2020.

[7] S. Bachhofner, I. Kis, C. Di Ciccio, and J. Mendling, "Towards a multi-parametric visualisation approach for business process analytics," in *Proceedings of the International Conference on Advanced Information Systems Engineering*, pp. 85–91, Essen, Germany, June 2017.

[8] R. Venkatraman and S. Venkatraman, "Big data infrastructure, data visualisation video and challenges," in *Proceedings of the 3rd International Conference on Big Data and Internet of Things*, pp. 13–17, VIC, Melbourn, Australia, August 2019.

[9] S. L. P. D. Guerreiro, "Decision-making in partially known business process environments using Markov theory and policy graph visualisation," *International Journal of Business Information Systems*, vol. 36, no. 3, pp. 355–392, 2021.

[10] R. D. Raut, S. K. Mangla, V. S. Narwane, B. B. Gardas, P. Priyadarshinee, and B. E. Narkhede, "Linking big data analytics and operational sustainability practices for sustainable business management," *Journal of Cleaner Production*, vol. 224, pp. 10–24, 2019.

[11] G. H. Tang and H. Zeng, "Visualisation technology in digital intelligent warehouse management system," *International Journal of Grid and Utility Computing*, vol. 12, no. 4, pp. 406–414, 2021.

[12] P. A. Castillo, A. M. Mora, H. Faris et al., "Applying computational intelligence methods for predicting the sales of newly published books in a real editorial business management environment," *Knowledge-Based Systems*, vol. 115, pp. 133–151, 2017.

[13] M. T. Wynn, E. Poppe, J. Xu et al., "ProcessProfiler3D: A visualisation framework for log-based process performance comparison," *Decision Support Systems*, vol. 100, pp. 93–108, 2017.

[14] C. Tsan-Ming, Y. C. Chang, and C. Jui-Kun, "A case study of visual management in TOYOTA manufacturing enterprise," *Journal of Quality*, vol. 16, no. 1, pp. 73–86, 2009.

[15] A. Ishizaka, S. A. Khan, S. Kusi-Sarpong, and I. Naim, "Sustainable warehouse evaluation with AHPSort traffic light visualisation video and post-optimal analysis method," *Journal of the Operational Research Society*, vol. 73, no. 3, pp. 558–575, 2020.

[16] L. C. George, Y. Guo, D. Stepanov, V. K. Reddy Peri, R. L. Elvitigala, and M. Spichkova, "Usage visualisation for the AWS services," *Procedia Computer Science*, vol. 176, pp. 3710–3717, 2020.

[17] M. Katuščáková, E. Capková, and J. Grečnár, "Capturing and sharing intangible cultural heritage through knowledge visualisation video and knowledge modelling tools," in *Proceedings of the European Conference on Knowledge Management*, pp. 612–617, Lisbon, Portugal, July 2019.

[18] M. Daradkeh, "A preliminary study of user acceptance and adoption of data visualisation tools for decision support in business organisations," *International Journal of Business Information Systems*, vol. 26, no. 3, pp. 297–317, 2017.

[19] J. L. K. Nußholz, "A circular business model mapping tool for creating value from prolonged product lifetime and closed material loops," *Journal of Cleaner Production*, vol. 197, pp. 185–194, 2018.

[20] T. Wolfenstetter, M. R. Basirati, M. Böhm, and H. Krcmar, "Introducing TRAILS: A tool supporting traceability, integration and visualisation of engineering knowledge for

product service systems development," *Journal of Systems and Software*, vol. 144, pp. 342–355, 2018.

[21] B. Taneja and K. Bharti, "Mapping unified theory of acceptance and use of technology (UTAUT) 2: A taxonomical study using bibliometric visualisation video," *Foresight*, vol. 24, no. 2, pp. 210–247, 2021.

[22] S. Alfadhel, S. Liu, and F. O. Oderanti, "Business process modelling and visualisation to support e-government decision making: Business/IS alignment," in *Proceedings of the International Conference on Decision Support System Technology*, pp. 45–57, Namur, Belgium, May 2017.

[23] M. Karlsson, S. Haraldson, M. Lind, E. Olsson, T. Andersen, and M. Tichavska, "Data visualisation tools for enhanced situational awareness in maritime operations," *Progress in IS, Maritime Informatics*, Springer, Cham, pp. 355–372, 2021.

[24] J. Závadský, Z. Závadská, C. Stępniak, and A. Stępniak, "Spatial visualisation video of stakeholders based on cartographic methodology in management systems," *Economic Annals-XXI*, vol. 179, no. 9-10, pp. 91–104, 2019.

[25] P. Quattrone, "Embracing ambiguity in management controls and decision-making processes: On how to design data visualisations to prompt wise judgement," *Accounting and Business Research*, vol. 47, no. 5, pp. 588–612, 2017.

[26] S. Joel-Edgar and J. Gopsill, "Understanding user requirements in context: A case study of developing a visualisation video tool to map skills in an engineering organisation," in *Proceedings of the 2018 International Conference on Information Management and Processing (ICIMP)*, pp. 6–10, London, UK, January 2018.

[27] P. Levontin, P. Baranowski, A. W. Leach et al., "On the role of visualisation in fisheries management," *Marine Policy*, vol. 78, pp. 114–121, 2017.

[28] A. E. Ripoll-Zarraga, F. Portillo, and C. Mar-Molinero, "The impact of the economic crisis on the efficiency of Spanish airports: A DEA visualisation analysis," *Research in Transportation Business & Management*, vol. 41, In press, Article ID 100689, 2021.

[29] W. Z. Low, W. M. P. V. D. Aalst, A. H. M. Hofstede, M. T. Wynn, and J. D. Weerdt, "Change visualisation: A," *Information Systems*, vol. 65, pp. 106–123, 2017.

[30] S. Pozniak, "Business value from visualisation video technologies," Dissertation, University of Stavanger, Norway, 2017.

[31] S. D. N. Wessels and R. Dixon, "Geotechnical data aggregation and visualisation video supporting informed risk management: The one-stop geotech shop," in *Proceedings of the Slope Stability 2020: Proceedings of the 2020 International Symposium on Slope Stability in Open Pit Mining and Civil Engineering*, pp. 703–712, Perth, Australia, January 2020.

[32] D. Lingerfelt and J. Dockins, "Medical practice office location analysis and operational metric assessment using spatial data visualisation video tools," *Management in Healthcare*, vol. 4, no. 3, pp. 221–230, 2020.

[33] S. Meier, B. Gebel-Sauer, and P. Schubert, "Knowledge graph for the visualisation of CRM objects in a social network of business objects (SoNBO): Development of the SoNBO visualiser," *Procedia Computer Science*, vol. 181, pp. 448–456, 2021.

[34] R. Adey, C. Peratta, and J. Baynham, "Corrosion data management using 3D visualisation and a digital twin," in *Proceedings of the CORROSION 2020*, Lviv, Ukrine, October 2020.

[35] N. Gerber, S. Hofer, and C. Beck, "Scientific visualisation of complex interdependencies in hospitals," *Transfer*, vol. 2019, no. 1, p. 8, 2019.

[36] B. Zhou, C. Maines, S. Tang, and Q. Shi, "A framework for the visualisation of cyber security requirements and its application in BPMN," *Guide to Vulnerability Analysis for Computer Networks and Systems*, Springer, Cham, pp. 339–366, 2018.

Hindawi
Scientific Programming
Volume 2022, Article ID 9993271, 8 pages
https://doi.org/10.1155/2022/9993271

Research Article

Personalized Learning Behavior Evaluation Method Based on Deep Neural Network

Hengyao Tang, Guosong Jiang, and Qingdong Wang

Computer School of Huanggang Normal University, Huanggang, Hubei 43800, China

Correspondence should be addressed to Qingdong Wang; wangqingdong@hgnu.edu.cn

Received 11 February 2022; Revised 4 March 2022; Accepted 10 March 2022; Published 22 April 2022

Academic Editor: Baiyuan Ding

In recent years, the research on personalized learning under the background of "Internet +" mainly focuses on the theory, design, and application and there is less research on learning evaluation. As an important means to measure the learning process and results, learning assessment plays an important role in supporting the effectiveness of personalized learning. From the perspective of educational services, how to realize learning evaluation that meets the needs of personalized learning is an important issue to be studied in the field of personalized learning. In this paper, the big data generated by learners on the online learning platform are used as the research target, and according to the level of learners' learning ability, a deep neural network is established to cluster and group them according to the cognitive thinking method. In order to reduce data redundancy and improve processing efficiency, a deep neural network with five hidden layers is used to extract typical features, so as to obtain more accurate evaluation results. Finally, the neural network model is used to obtain the clustering results of different groups of learning behaviors and the evaluation curves of the five-course knowledge points of learners at different levels. From the experimental results, the proposed personalized evaluation method can effectively analyze the learning differences between learners with different ability levels, and it is basically consistent with the evaluation standards of artificial experts.

1. Introduction

Personalized learning in the context of Internet+ aims at the development of students' individuality, respects individual differences of students, and emphasizes the individual support of information technology. This way of learning can promote students' individual potential to be maximized, which is very in line with the current needs of talent training in colleges and universities [1–3]. The development of information technology, especially the application of the Internet in education, enables learners to obtain abundant resources according to their own interests and learning needs, receive personalized services and guidance, and control the learning process autonomously, making learning more personalized change [4].

The current society is entering the information age and the "Internet +" age from the industrial age. With the improvement of social and economic level and living standards, people's educational needs are developing from standardized teaching to personalized learning and lifelong learning, and the supply of educational services will also change from "standardized supply" to "personalized service." Relevant statistics show that in recent years, the research on personalized learning under the background of "Internet +" mainly focuses on the theory, design, and application, and there is less research on learning evaluation [5–7]. As an important means to measure the learning process and results, learning assessment plays an important role in supporting the effectiveness of personalized learning. Although personalized learning can simplify the cost and operation mode of learners, there is

currently a lack of suitable means for evaluating the teaching effect of personalized learning; especially, some unique characteristics displayed by personalized learning, such as learners before starting to learn. The cognition of the knowledge structure and scope that one has, the knowledge blind spots that should be made up after a period of study, the improvement of learning ability and preferences after the study is over. If the characteristics of this personalized learning are not accurately grasped, it will be difficult to achieve the true sense of teaching students in accordance with their aptitude. In addition, if the traditional artificial force is used to analyze the activities of each learner, it is impossible to face the vast ocean of data generated every moment on the learning platform. Using information technology, the machine analyzes the big data generated by the user's learning behavior, retains key and effective features, and generates a complete and accurate personalized evaluation. From the perspective of educational services, how to realize learning evaluation that meets the needs of personalized learning is an important issue to be studied in the field of personalized learning [8–12].

In response to this core issue, many researchers have conducted in-depth research from various aspects. For example, for the massive big data generated on the learning platform, Hadoop technology is used to aggregate, store, and obtain massive learning data; the literature forms a multidimensional three-dimensional household data model by recording learners' online learning behaviors, records, habits and preferences, and other activities to quantify the behavior patterns of learners in the process of personalized learning, so as to provide targeted personalized services, use the gradient descent method to synthesize the data, obtain relevant data models for learning evaluation, and put forward an intelligent learning guidance environment [13, 14, 15]. No matter which method is used, the core of the evaluation is how to effectively classify the data generated by the learner in the learning process and to effectively reduce the dimension of the multidimensional data features that describe the learner's behavior in an appropriate way, which can ensure the uniqueness of the data features [16–19]. It is representative and can ensure that the system platform can realize the calculation with the minimum calculation cost. And how to improve the accuracy of the classification and effectively use the classification results to provide help for subsequent processing, the performance of the deep neural network in this aspect is very significant. At present, neural networks have good performance in almost all problem areas related to classification [20].

In view of the advantages of deep neural network in feature training, a method based on deep neural network learning is proposed in this paper. First, the feature vectors of the original learning data generated by the learners are automatically clustered, and then, the multidimensional data vector features are dimensionally reduced and cleaned by using the deep learning network DNN to ensure the real validity and real-time performance of the evaluation data. Finally, through the relevant experimental data, the validity of this personalized evaluation behavior is verified [21–24].

2. Analysis of the Existing Problems in Personalized Learning Evaluation

In the "Internet +" environment, more and more students' personalized learning evaluations use process evaluation methods. Process evaluation is an evaluation activity aimed at optimizing the learning process, improving the learning effect, and promoting the development of individual life. Process-oriented learning evaluation can be used as an effective means to focus, record, guide, motivate, and promote learners' learning experience and growth and is a key link in the quality assurance of online education. Through literature analysis and teaching practice, the author found that there are some problems in the process of implementing personalized learning evaluation.

2.1. Analysis of Evaluation Subjects. The subject of evaluation refers to the implementer of evaluation activities, that is, the evaluator. The evaluation concept of the evaluator should be considered from two aspects, one is the current knowledge level of the evaluator, and the other is its possible development potential. The main problems of evaluation subjects are as follows. (1) There is over-emphasis on the diversification of evaluation subjects, such as students' self-evaluation and mutual evaluation, intragroup and intergroup evaluation, teachers' evaluation, and even the evaluation of students' parents, but such a multievaluation method is not suitable to evaluate personalized learning, because personalized learning focuses on changes in the learning process, not just learning results, and self-evaluation and mutual evaluation lack the necessary process data support. (2) There is lack of personalized evaluation standards, even if the same students in a teaching class vary in their knowledge level, learning ability, personal goals, academic mood, etc. It is difficult to use unified evaluation standards to provide students with personalized and accurate evaluations, and it is difficult to help students achieve their own learning goals. (3) The evaluation is based on a single source of data, and the learning effect evaluation of learners is often only measured by academic performance, lacking data support for learning behaviors, such as independent study, participation in problem discussions, completion of homework, display and sharing of works, etc.

2.2. Evaluation Object Analysis. The object of evaluation refers to the recipient of the evaluation activity, that is, the person being evaluated. Contemporary college students have flexible thinking, strong ability to accept new things, high enthusiasm for professional learning and ability improvement, and advocating competition and have a positive and enterprising spirit. However, it is undeniable that with the continuous changes in the international and domestic situation, under the combined effect of ambivalence such as enjoying life and employment pressure, some students are anxious and confused in their hearts. Some problems

include not knowing oneself correctly, failing to find a suitable learning method for oneself, and lack of self-confidence to become an excellent student.

2.3. Analysis of the Evaluation Process. Process evaluation should be able to effectively guide students' learning process, adjust learning behavior, and timely feedback of learning results, which has a positive impact on improving students' learning motivation and guiding students' future development direction. At present, the traditional procedural evaluation process has the following problems: (1) lack of individualized assessment of learners; (2) lack of data tracking of learning behavior in the learning process; (3) lack of timely feedback of learning results.

3. Construction of the Personalized Learning Evaluation Model from the Perspective of Education Service

In the "Internet +" environment, the process evaluation of personalized learning requires appropriate service processes to solve the actual problems and solve the problems that lack the evaluation of learners' personalized characteristics, the lack of data tracking of learning behavior in the learning process, and the lack of learning results timely feedback and other issues. This research builds a personalized learning evaluation model from the perspective of educational services, which mainly provides services such as learner personalized feature testing, personalized learning evaluation plan matching, personalized learning process data tracking, and timely feedback of learning results. The specific evaluation service process is described as follows.

3.1. Personalized Trait Testing for Learners. The evaluation service first evaluates the evaluation object, that is, the knowledge level of the learner, and then accurately selects a personalized evaluation plan according to the learning goal selected by the learner: (1) to test the knowledge level, it is necessary to understand the learner before the learning evaluation service. Match the appropriate individualized assessment plan through the knowledge level; (2) after the learners complete the knowledge level test, the selected learning objectives are divided into mastery and understanding; (3) test the academic mood; Academic Mood Self-assessment is a subjective test for learners. The purpose of testing learners' academic emotions is to study the impact of personalized learning evaluation services on students' personalized learning.

3.2. Automatic Evaluation Scheme Matching for Learners. The results of the learner's knowledge level test and the choice of learning objectives determine the type of learning. According to the learning type, the learning evaluation plan is automatically matched from the preset personalized evaluation plan library. In different evaluation schemes, for learners with high knowledge levels and learning goals, the evaluation standard will be higher; for learners with low knowledge levels and low learning goals, the evaluation standard will be appropriately lowered, requiring them to complete basic learning. The content can be improved on an original basis.

3.3. Personalized Learning Process Tracking for Learners. The purpose of tracking the learning process of individualized learning is to grasp the first-hand information of the learner's learning status. Based on the concept of big data, the learner's learning activities and learning results are recorded in detail. Through data analysis, problems are identified, and timely reminders and help are given. Intervention and adjustment of learners' learning behaviors are carried out through learning process tracking and data analysis methods.

3.4. Giving Learners Feedback on the Results of Learning Activities. The learning activities and learning results of the personalized learning process must have accurate and timely feedback. Quantitative feedback should be given to the learning activities that learners participate in, such as the degree of interaction, participation in discussions, resource utilization, etc., and the learning points will be recorded according to the set rules; in the correction of homework, the objective questions will be feedback immediately, and the subjective questions will be given within a limited time limit. Feedback is given inside; the test results of each stage are displayed in visual graphics so that learners can keep abreast of the trend of their academic performance; excellent practical works are displayed and exchanged, and their advantages and disadvantages are pointed out. This service link not only allows learners to understand their own test scores in time but also understands the results of their own learning behavior investment, so as to clarify the next step forward.

4. Deep Neural Network Evaluation Model

4.1. Deep Neural Network. DNN is developed based on artificial neural network, and the main difference between the two is that a deep neural network contains multiple hidden layers and the number of network nodes. DNN hidden layers discover the inherent properties of the data, thereby improving the modeling ability of neural networks to learn multiple layers of abstract data. Multiple neurons in DNN can obtain common core features of the dataset from a small amount of training data and have powerful modeling capabilities for complex problems. The specific process of DNN is as follows.

After the data is preprocessed, the initialization data are passed from the input layer to the first hidden layer. The input-output relationship of the first hidden layer is

$$r_1 = f(w_1 \cdot x + b_1). \tag{1}$$

Suppose all output values in r_1 are the original column vector x transformed by the activation function f:

$$r_{1,m} = f\left\{\sum_{i=1}^{n} w_{1,m,i} \cdot x_i + b_{1,m}\right\}. \tag{2}$$

According to the principle of DNN, the output r_p of the Pth hidden layer in the DNN model can be obtained.

$$r_p = f\left(w_p \cdot r_{p-1} + b_p\right). \tag{3}$$

The value of all elements in the output r_p of the hidden layer of $r_{p,m}$ layer P is

$$r_{p,m} = f\left\{\sum_{i=1}^{q} w_{p,m,i} \cdot r_{p-1,i} + b_{p,m}\right\}. \tag{4}$$

After the input vector X is processed by the input layer and all hidden layers, it will be transmitted to the output layer. The result is as follows:

$$y = g\left(w_{n+1} \cdot r_n\right) + b_{n+1}. \tag{5}$$

For a training set $\{(x_1,y_1),\ldots,(x_m,y_m)\}$ containing m labeled samples, during the neural network training process, the cost function of each sample (x, y) is

$$G(W,b) = \frac{1}{m}\sum_{i=1}^{m}\frac{1}{2}\left\|h_{W,b}\left(x^i\right) - y^i\right\|^2. \tag{6}$$

Using the gradient descent method can obtain a good convergence effect and reach the local optimal value, so the parameters W and b are set, and the update formula is as follows:

$$\begin{aligned} W_{ij}^l &= W_{ij}^l - \alpha\frac{\partial}{\partial W_{ij}^l}G(W,b), \\ b_i^l &= b_i^l - \alpha\frac{\partial}{\partial b_i^l}G(W,b). \end{aligned} \tag{7}$$

4.2. An Efficient Clustering Method Based on Multidimensional Data Vectors. According to Bloom's theory, human cognitive thinking is divided into six levels: memory, understanding, application, analysis, evaluation, and creation. Therefore, two levels are adopted for the abstraction of online learning behavior data features. The first level is low-level features, which mainly includes login time, learning time, number of learning times, selected knowledge points, number of discussion, number of questions asked, number of questions answered, number of questions solved, time to complete the test, success rate of completing the test, homework score, etc. $U_{\text{low}} = \{x_1, x_2, \ldots, x_n\}$. The second layer is the high-level feature, which mainly includes the degree of homework completion, the accuracy rate of homework completion, the quality of learning questions, the quality of answering questions, the quality of problem-solving solutions, etc. $U_{high} = \{y_1, y_2, \ldots, y_n\}$. Each sample can be further divided into multidimensional features, such as $x_i = (x_{i1}, x_{i2}, \ldots, x_{in})$ and $y_i = (y_{i1}, y_{i2}, \ldots, y_{in})$, where each component represents a characteristic of learning activities, such as logging in times and study time. After obtaining multidimensional data, a clustering algorithm can be used to divide the required data features.

4.3. Feature Extraction Methods for Clustered Data. The feature extraction of clustered data adopts the hidden Markov model based on DNN, which is a forward neural network with multiple hidden layers. The input layer represents the underlying features of the clustered data, and the output layer represents the typical features after dimensionality reduction. The nonlinear activation function of each node in the hidden layer adopts SIGMOD, and the nonlinear output value of each node is

$$\begin{aligned} y_j^h &= \text{Sig}\left(x_j\right) = \frac{1}{1+e^{-x_j}}, \\ x_j &= b_j + \sum_i y_j^{h-1} w_{ij}. \end{aligned} \tag{8}$$

Among them, y_j^h is the nonlinear output value of the jth node in the hth layer; x_j is the node input value; b_j is the bias; w_{ij} is the connection weight between node j and i. The DNN training parameters are obtained by iterative training of the BP network propagation algorithm.

$$J = (w_1, b_1, w_2, b_2) = \frac{1}{N}\sum_{i=1}^{N}\left(x_i' - x_i\right)^2. \tag{9}$$

The initial network parameters are initialized by the RBM restricted Boltzmann machine.

$$\begin{cases} \dfrac{\partial \log p(v|\theta)}{\partial w_{i,j}} \approx \dfrac{1}{N}\displaystyle\sum_{n=1}^{N}\left[v_i^n h_j^{(n)} - v_i^n h_j^{(n)}\right], \\ \dfrac{\partial \log p(v|\theta)}{\partial b_i} \approx \dfrac{1}{N}\displaystyle\sum_{n=1}^{N}\left[v_i^n - v_i^n\right], \\ \dfrac{\partial \log p(v|\theta)}{\partial b_j} \approx \dfrac{1}{N}\displaystyle\sum_{N=1}^{N}\left[h_j^{(n)} - h_j^{(n)}\right]. \end{cases} \tag{10}$$

In this paper, the deep neural network modeling unit is used as the basic unit. As shown in Figure 1, each multidimensional feature is divided into three HMM states, and all multidimensional feature HMM states correspond one-to-one with each node of the DNN output layer. In the experiment, 8-dimensional features are used as input, and 5 hidden layers are used, and each hidden layer has 1024 nodes.

As shown in Figure 1, the input of the deep neural network is the learning activity data, and the output corresponds to the key typical features after dimensionality reduction and cleaning. After the layered processing of the neural network, the distinguishing degree of the features is enhanced. Some distinguishing degrees are relatively poor and are not used. The feature quantities that highlight the characteristics of the learner are processed by the hidden layer, which can ensure that the final extracted features can

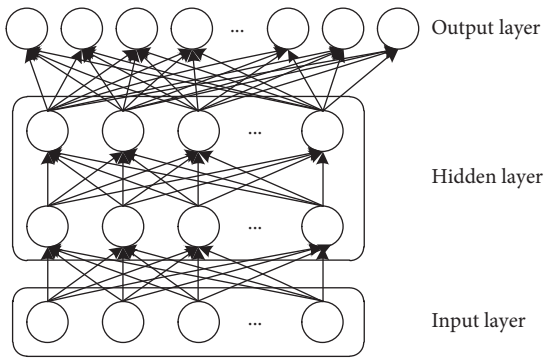


Figure 1: Deep neural network structure.

greatly reduce the dimension while maintaining the maximum degree of discrimination. The activation of each hidden layer node of the deep neural network is mainly determined by the output value of the node. In order to make the feature distribution more similar to the Gaussian distribution, the output value of the hidden layer is averaged to make the feature output close to the Gaussian distribution.

The obtained clustering feature is denoted as $\cup\sigma_i$, where σ_i corresponds to the i-th learning feature, and the mean value of this feature in the hidden layer is obtained through the neural network.

$$H_{i,t} = \frac{1}{L}\left(h_{i,1} + h_{i,2} + \cdots + h_{i,L}\right). \tag{11}$$

Among them, $h_{i,L}$ is the nonlinear output vector of the i-th feature on the L-th layer, and the average value of each hidden layer feature is taken as the network feature of the feature.

$$F = \frac{1}{T}\sum_{i=1}^{N}\sum_{i=1}^{N} H_{i,t}. \tag{12}$$

In order to obtain the effective feature components in the mean feature of the hidden layer, the effective feature after the final dimension reduction is obtained by using

$$E = H_{i,t} - F. \tag{13}$$

5. Analysis of Results

In the experiment, a well-known video teaching website in China was used as the test platform, and the records of 900 students' learning activities on the platform were compared and analyzed. First, the experimental samples are divided into three groups: elementary, intermediate, and advanced according to the level of learning ability, and each group includes 300 students' records of learning activities of a certain knowledge point in 5 online courses within one month.

The knowledge points of different courses basically show a normal distribution pattern for learners at different levels. Regardless of the level of the learner's ability, the mastery of the knowledge points of the course basically conforms to the same principle. Since the difficulty coefficient of each course is different, different courses also show different degrees of distinction. For example, it can be found that the distribution state and style of the intermediate group and the advanced group are basically the same, while the distribution of the elementary group is different from the other two. There are clear differences between groups. This shows that for the five courses currently tested, the level of entry level still has a certain influence on the establishment of learners' later learning behavior.

Figure 2 shows the performance curves of learning each course in different groups, mainly showing the clustering results of learning knowledge points of different courses. As can be seen from Figure 2(a), since the "Web Technology Fundamentals" course does not require very high prerequisite knowledge for learners and the knowledge points are simple, the primary group surpasses the intermediate and advanced groups, which may be related to the learners' learning experience mentality related. Figures 2(b) and 2(c) two courses have higher requirements for preknowledge and need to have a certain computer and mathematics foundation, so it can be found that the students in the advanced group have a better grasp of knowledge points than the other two groups. As Figures 2(d) and 2(e) are both programming courses, the knowledge points to be learned are similar and the difficulty is equal, so the performance curves are basically similar, but it is obvious that the primary group's mastery of knowledge points is similar to other. There is still a small difference between the two groups. It can be seen that the learning ability and behavior of the advanced group samples are significantly higher than those of the intermediate group and the primary group, which are related to the learners' previous learning experience, knowledge accumulation, learning habits, and understanding ability, which also shows that the learning ability high and low is a step-by-step cultivation process.

In order to further measure the accuracy of the evaluation method in this experiment, the method of manual expert review was introduced to compare with the method of machine learning evaluation. Experienced teachers are selected in this major to evaluate the knowledge points of each course according to different levels of groups. The main basis of manual evaluation methods still uses traditional methods, such as class attendance rate, homework completion rate, completion quality, test scores, project capability, and other easily quantifiable indicators, and in order to ensure statistical fairness, all human experts are set to have the same weight. The results of the statistical analysis are shown in Figure 3. It can be seen from the figure that due to the advantages of learning background and ability, the curve of the advanced group is slightly different from that of the other level groups and the manual group, while the curve obtained by the primary group and the intermediate group is highly consistent with the curve obtained by the manual review, indicating that the machine learning evaluation is not effective. The method can truly reflect the human evaluation criteria for the learning effect.

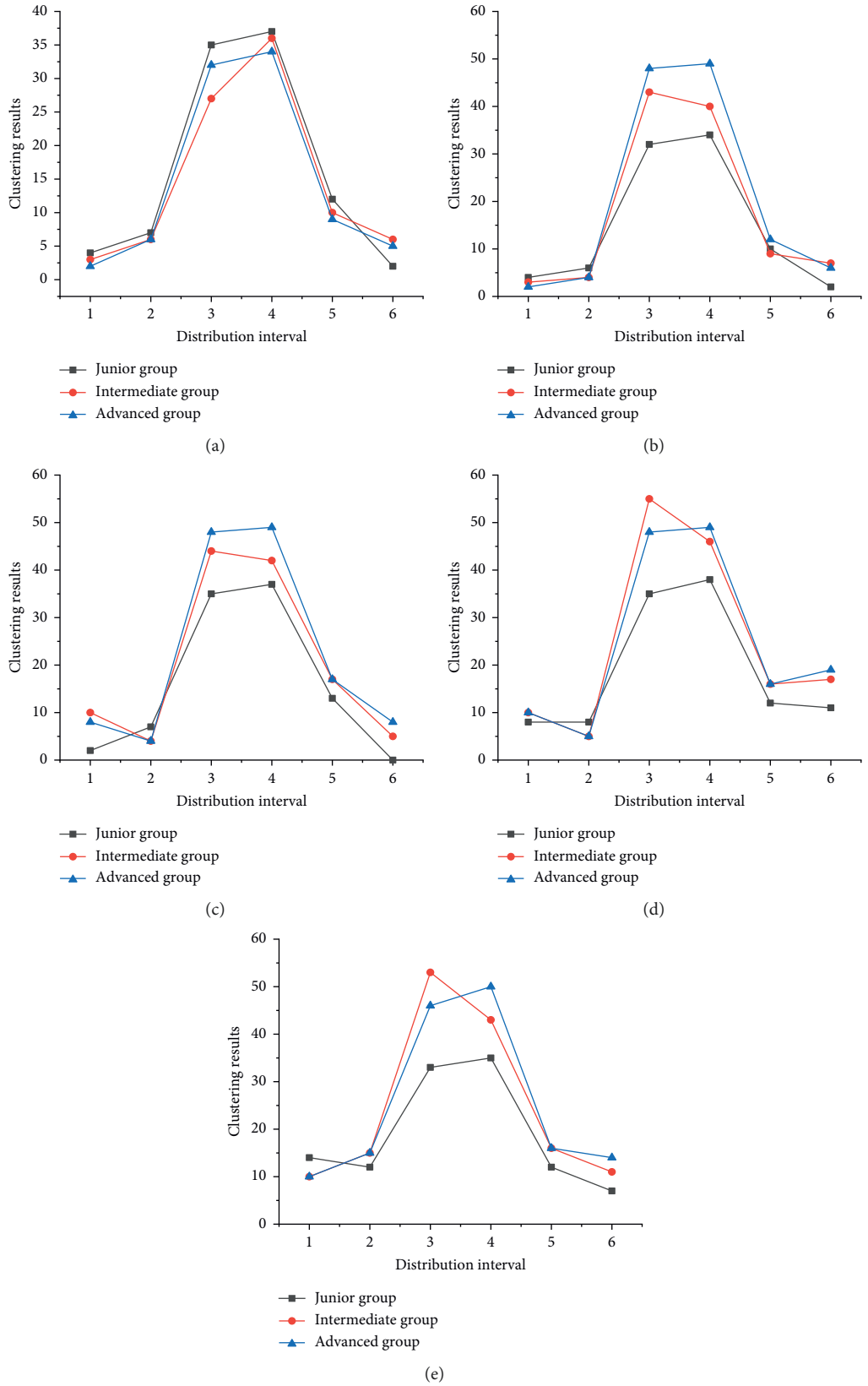


Figure 2: Evaluation curves of different courses for learners at different levels. (a) Assessment of knowledge points in the basic course of web technology. (b) Assessment of knowledge points in artificial intelligence courses. (c) Evaluation of knowledge points in the data structure course. (d) Evaluation of knowledge points in the python degree design courses. (e) Evaluation of knowledge points in a java programming course.

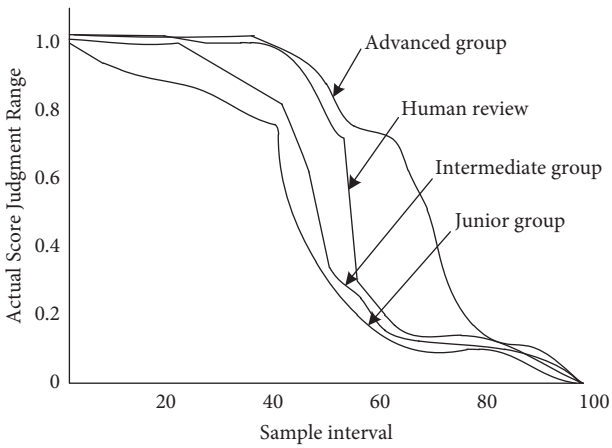


Figure 3: Comparison of results between human expert evaluation and machine statistical evaluation.

6. Conclusion

The development of information technology will inevitably bring about fundamental changes in traditional learning methods. How generate personalized and accurate learning behavior evaluation for learning users, so that learners can better understand their learning situation. The state and behavior state truly realize the "targeted" and "individualized teaching" of the learning method. This paper uses machine learning to classify the user data generated by the network platform and then uses the deep neural network to reduce the dimension of the feature data to extract typical features, which can not only reduce the computational cost of training but also ensure that the training features are the best. It can reflect the characteristics of learners' individual learning. Judging from the experimental results, the evaluation results obtained by machine learning are basically consistent with the evaluation results obtained by manual expert review, which shows that under the premise of ensuring the amount of data, machine evaluation can replace manual evaluation, and machine evaluation is more objective and comparative. However, according to the latest research results, the gradient of the deep neural network will be unstable during training, and the existence of this phenomenon will affect the performance of the neural network. The next research can consider introducing different learning methods to analyze the learning data, in order to obtain a more general personalized learning evaluation method.

Data Availability

The dataset can be accessed upon request to the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] J. Shin and O. Bulut, "Building an intelligent recommendation system for personalized test scheduling in computerized assessments: a reinforcement learning approach," *Behavior Research Methods*, vol. 54, no. 2, 2021.

[2] X. Tang, Y. Chen, X. Li, J. Liu, and Z. Ying, "A reinforcement learning approach to personalized learning recommendation systems," *The British journal of Mathematical and Statistical Psychology*, vol. 72, no. 1, 2018.

[3] F. Zhang and S.-X. Feng, "The process of personalized learning based on flipped classroom," *Sino-US English Teaching*, vol. 14, no. 4, 2017.

[4] D. H. C. Nguyen, N. Arch-Int, and S. Arch-Int, "FUSE: a fuzzy-semantic framework for personalizing learning recommendations," *International Journal of Information Technology and Decision Making*, vol. 17, no. 4, 2018.

[5] Y. Chen, "Research on resource personalization recommendation algorithm and model based on computer deep learning," *Journal of Physics: Conference Series*, vol. 1915, no. 2, 2021.

[6] A. Shemshack, Kinshuk, and J. M. Spector, "A comprehensive analysis of personalized learning components," *Journal of Computers in Education*, vol. 8, no. 4, 2021.

[7] H. Almed, "Determinant factors and adaptation features of mobile personalized learning system," *Journal of Physics: Conference Series*, vol. 1858, no. 1, 2021.

[8] S. Chaichumpa, S. Wicha, and P. Temdee, "Personalized learning in a virtual learning environment using modification of objective distance," *Wireless Personal Communications*, vol. 118, pp. 2055–2072, 2021.

[9] L. Fulton, D. L. Hoffman, and S. Paek, "Language learning in an era of datafication and personalized learning," *Educational Technology Research and Development*, vol. 69, no. 5, 2020.

[10] D. Zou and H. Xie, "Personalized word-learning based on technique feature analysis and learning analytics," *Journal of Educational Technology & Society*, vol. 21, no. 2, 2018.

[11] L. Zhang, J. D. Basham, and S. Yang, "Understanding the implementation of personalized learning: a research synthesis," *Educational Research Review*, vol. 31, 2020.

[12] W. Intayoad, T. Becker, and P. Temdee, "Social context-aware recommendation for personalized online learning," *Wireless Personal Communications*, vol. 97, no. 1, 2017.

[13] C. Troussas, K. Chrysafiadi, and M. Virvou, "An intelligent adaptive fuzzy-based inference system for computer-assisted language learning," *Expert Systems with Applications*, vol. 127, 2019.

[14] A. T. Tadesse and P. . I. Davidsen, "Framework to support personalized learning in complex systems," *Journal of Applied Research in Higher Education*, vol. 12, no. 1, 2020.

[15] X. Su, "Personalized learning resource recommendation methods from the perspective of deep learning," *Organic Chemistry: An Indian Journal*, vol. 14, no. 3, 2017.

[16] H. Wang and W. Fu, "Personalized learning resource recommendation method based on dynamic collaborative filtering," *Mobile Networks and Applications*, vol. 26, pp. 473–487, 2020.

[17] "Personalized learning environment in higher education through big data and blended learning analytics," *International Journal of Recent Technology and Engineering*, vol. 8, no. 3, 2019.

[18] M. Alberto, "Optimizing student-driven learning (SdL) through a framework designed for tailoring personal student paths," *Education Sciences*, vol. 10, no. 9, 2020.

[19] H. Xie, Di Zou, R. Zhang, M. Wang, and R. Kwan, "Personalized word learning for university students: a profile-based method for e-learning systems," *Journal of Computing in Higher Education*, vol. 31, no. 2, 2019.

[20] H. A. Alamri, S. Watson, and W. Watson, "Learning technology models that support personalization within blended learning environments in higher education," *TechTrends*, vol. 65, no. 3, 2020.

[21] M. Somasundaram, K. A. Mohamed Junaid, and M. Srinivasan, "Artificial intelligence (AI) enabled intelligent quality management system (IQMS) for personalized learning path," *Procedia Computer Science*, vol. 172, no. C, 2020.

[22] F. M. Andi Besse, Y. Norazah, and B. Ahmad Hoirul, "Personalized learning model based on deep learning algorithm for student behaviour analytic," *Procedia Computer Science*, vol. 163, no. C, 2019.

[23] X. Duan, "Automatic generation and evolution of personalized curriculum based on genetic algorithm," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 14, no. 12, 2019.

[24] M. Zhang, J. Zhu, Z. Wang, and Y. Chen, "Providing personalized learning guidance in MOOCs by multi-source data analysis," *World Wide Web*, vol. 22, no. 3, 2019.

Hindawi
Scientific Programming
Volume 2022, Article ID 6298248, 10 pages
https://doi.org/10.1155/2022/6298248

Research Article

A Deep Convolutional Neural Network Based Risk Identification Method for E-Commerce Supply Chain Finance

Qian Tang,[1,2] Yan Lu,[2] Bin Wang,[1] and Zhen Li [3,4]

[1]*Faculty of Science and Technology, University of Macau, Taipa, Macau 999078, China*
[2]*College of Business, Beijing Institute of Technology, Guangdong, Zhuhai 519088, China*
[3]*Centre for Efficiency and Performance Engineering University of Huddersfield, Huddersfield, HD1, UK*
[4]*College of Industrial Automation, Beijing Institute of Technology, Guangdong, Zhuhai 519088, China*

Correspondence should be addressed to Zhen Li; mb75437@connect.um.edu.mo

Received 28 February 2022; Revised 20 March 2022; Accepted 23 March 2022; Published 14 April 2022

Academic Editor: Baiyuan Ding

With the popularity of the Internet, the rise of e-commerce platforms has led to the rapid development of supply chain (SC) financial services in China, and the competitiveness of commercial banks and core enterprises in the supply chain is now gradually increasing, rapidly expanding into an important area of competition between the two. As an emerging force rebounding from the economic downturn, e-commerce platform transactions, with their unique characteristics of informatization, diversification, and convenience, have provided a broad space for Internet SC finance. The article mainly analyzes the risk identification method of e-commerce SC finance, analyzes its risk from the financing process, gives corresponding data support for the matters or processes that may cause financing risk based on DCNN model, and takes Jingdong SC finance as an example and analyzes its main financing methods and risk identification process; based on different experimental comparisons, a multigroup experimental study shows that the accuracy of supply chain finance risk identification using deep convolutional neural network models can reach 95.36%, which demonstrates the effectiveness of the proposed method by providing better performance compared to traditional BP and SVM networks.

1. Introduction

With the popularization of the Internet, more and more e-commerce enterprises rely on the Internet platform to complete multichannel financing for their projects. The steadily increasing number of online platform institutions and e-commerce enterprises has led to increasing diversity. With the encouragement of economic policies and the wave of "Internet+," people are trying to use the Internet to build a modern intelligent SC system that integrates big data [1], shared Internet, artificial intelligence, and blockchain. In response to the domestic economic transformation and upgrading and the impact of the Internet culture, traditional enterprises are beginning to change and open up new markets based on the concept of following the past. Banking and finance as intermediaries are actively building their own data platforms, forming the advantages of numerous branches, gathering financial talents, and extensive reach to make customer data more three-dimensional while retaining the original customers [2].

In the supply chain, apart from the core enterprises, the upstream often has difficulties in financing and expensive financing due to their own scale, which has become a bottleneck in the development of the SC, which integrates logistics, information flow, commercial flow, and capital flow in the SC, provides a new way to break the bottleneck of SC development, and is developing rapidly. In 2014, Shanghai Chunyu SC Management Co [3], Ltd. was exposed by the media that its capital chain was in trouble and it owed a huge amount of money, and the "Chengxing International Incident" in 2019 involved a huge amount of money. In 2019, the CBIRC issued the "Guidance on Promoting SC Financial Services," stating that it is necessary to adhere to precise financial services, the authenticity of transaction

backgrounds, the availability of transaction information [4]; from offline SC finance to online SC finance transformation and upgrading, due to the rapid development, the research of SC finance has long existed in a situation where theory seriously lags behind practice [5].

Foreign scholars mostly start from the perspective of enterprise financing. Timme [6] believes that SC finance is a new type of financing business in which all the participants in the SC chain cooperate with financial institutions outside the chain to provide convenient SC financing. Hoffmann [7], from the perspective of the flow of funds between industrial organizations, proposes that SC finance is a number of organizations inside and outside the SC, through planning and controlling; therefore, Berger and Udeel [8] proposed a "government policy financial structure-lending technology" financing approach. Michael Lamoureux [9] made a new definition of the concept of SC finance on the basis of a synthesis of previous research done by scholars, and he believes that SC finance is a process of capital optimization, in which the flow of capital, logistics, information, and commerce is integrated, and the various means of financing and costs are optimized. Lamoureux [10] believes that SC finance is a process of optimizing the financing of the system, through the integration and use of transaction information, the integration of capital flow, logistics, information flow and commercial flow, embedded cost management analysis, and the optimization of the availability of funds and costs by various financing instruments. Li et al. [11] start from the perspective of enterprise capital flow and point out that SC finance is a financial service activity that provides credit and settlement for enterprises in the SC through capital optimization and monitoring in order to reduce the costs of suppliers and retailers. It can be seen that foreign scholars understood SC finance from the earliest as only a means of financing, to later focus on the flow of funds in the SC system and the capital optimization process [12], and gradually considered SC finance as a comprehensive financial service to improve the efficiency of the SC and financial value added.

Compared with foreign scholars, domestic scholars' research on SC finance started later and with a different entry point. At the earliest, Khan et al. [13] extracted the concept of "rongtong warehouse" from logistics finance, and "rongtong warehouse" is divided into two definitions: narrow and broad, but nowadays, the narrow definition is mostly taken: rongtong warehouse is a comprehensive service platform that integrates credit for SMEs, covers logistics enterprises, and connects traditional commercial platforms with e-commerce platforms. It is a comprehensive service platform that integrates credit for SMEs [14, 15], covers logistics enterprises, and connects traditional commercial platforms with e-commerce platforms and is believed to break the "bottleneck" of external financing for SMEs. Some other scholars define SC financial services in terms of the business model of commercial banks.

As the research progressed, it gradually shifted from the research on logistics finance and financing to the analysis of SC structure [16]. This phase defines it as follows: SC finance is based on the analysis of its internal transaction structure, introducing variables such as core enterprises, logistics supervisory companies, and capital flow instruments, and using a self-repaying trade credit model to provide closed credit, settlement, financial management, and other financial services to enterprises in the SC nodes [17]. Apostolopoulos and Mpesiana [18] have since made a new interpretation and expansion of the definition, arguing that SC finance is the provision of targeted financial management solutions by specific financial institutions or organizers in the SC (core enterprises, third-party logistics enterprises) for specific links in the SC or for the whole process of the whole node.

The concept of online SC finance (OSCF) was first proposed by Wang and Wong [19]. Internet SC finance refers to a new comprehensive financial service model in which an e-commerce platform, which is qualified to operate and provide funds at the same time, or a commercial bank, has a credit system with a large amount of customer data accumulated by the platform, and uses the credit model of self-reimbursement trade finance to provide commercial credit, issue loans, and make payments and settlements for SMEs and individual users on the platform. According to Goodwill et al. [20], Internet SC finance is an industrial ecosystem and financial, ecological platform that builds cross-sectoral and cross-regional through the Internet, the Internet of Things, and other advanced technological means on an online platform and is widely integrated with enterprises, industries, and governments. This paper will analyze the main risks faced by SC finance and the corresponding risk control methods from both theoretical and practical aspects, DCNN model, and take Jingdong SC finance as an example to analyze its main financing methods and risk identification process. Through experimental comparison, the results show that the accuracy of the proposed deep convolutional neural network model for supply chain finance risk recognition is 95.36%, which is an improvement compared with BP, SVM, and LeNet-5 models, showing the superiority of the model, but because of a large number of parameters in the operation of the model, improving the operation speed is also a direction worthy of research.

2. SC Financing Model for E-Commerce Enterprises

The traditional SC finance model takes commercial banks as the center of capital financing, with commercial banks playing the role of credit assessment and credit decision-making. There are three main modes of operation: accounts receivable financing, order financing, and warehouse financing, among which accounts receivable financing involves accounts receivable and commercial paper pledge mode; order pledge financing has spot financing and order pledge [21–23], warehouse financing is divided into inventory pledge and ticket first and then goods mode.

The Internet SC finance model is a way to use the Internet platform for transaction financing. At present, Internet SC finance contains three financing models, namely intermediary-type e-commerce transactions, P2P, and self-operated e-commerce platform-led. The intermediary-type e-commerce trading platform-led SC finance model is to put

forward loan applications to the platform when the e-commerce enterprise has capital needs, and the merchant repays the principal and interest to the intermediary platform after completing capital turnover; the P2P platform-led is a third-party network platform that connects individual funds borrowing and lending parties [24], and it is difficult for the P2P platform to act as a financial institution alone for independent risk control; the self-operated e-commerce platform-led model participates in all aspects of its self-operated The self-operated e-commerce platform-led model is involved in all aspects of its own products and can provide credit services to its own suppliers, so the self-operated model has access to the most comprehensive data information of the three models.

2.1. The Financing Process of E-Commerce Enterprises under SC Finance. In the context of SC finance, e-commerce enterprises use data directly or indirectly to direct resources into play through emerging fields such as data analysis, network cloud computing, and blockchain economy, forming the initial authorized credit and risk control system of the e-commerce platform [25]. At the same time, e-commerce enterprises rely on national policy support and the vast resources of the Internet platform to accumulate financial data related to the operation and sales of suppliers involved in the financing chain, So it makes the financing much more time-efficient. This allows e-commerce businesses to have more goods supplied by their suppliers, even in the case of e-commerce shopping festivals or promotions under the new media, where there may be a shortage of supply, so the solution of pledging goods to third-party storage enterprises is proposed, and the resulting receivables can be applied for financing from factoring financing products, and the suppliers will receive a certain percentage of the loan for the goods. The process is illustrated in Figure 1.

When suppliers of e-commerce enterprises generate a large number of orders, they may also choose to use the orders as pledges for financing. After the supplier platform applies for the SC factoring financing product, the commercial bank analyzes and compares the data transmitted by the e-commerce company to determine the credit rating of the supplier and issues a certain percentage of the loan. When the purchaser pays the final payment for the goods to the e-commerce company. The remaining funds are transferred back to the supplier. The process is illustrated in Figure 2.

The financing warehouse refers to the inventory pledge loan model. Considering the instability and variability of the goods, it usually requires the participation of a third-party storage enterprise to cooperate with the e-commerce company to carry out value estimation and goods management of the enterprise's inventory, so the e-commerce company is both a distributor and a logistics enterprise at the same time. The process is illustrated in Figure 3. The commercial bank and the e-commerce company receive an application from an enterprise with a financing need for finance through warehouse financing and agree that the enterprise will store its inventory in the warehouse of the captive logistics, which takes over the management inspection and assessment work; subsequently, the commercial bank determines the limits of the loan based on the feedback assessment report; finally, the financing enterprise withdraws the inventory in batches after repaying the loan in batches [26].

Self-supply is a more specific management model, while self-managed e-commerce is involved in the whole process of "production—purchase—transaction—transport" of its own products, and the self-managed e-commerce platform can provide credit to its own suppliers. The role in the supply chain is more like that of a core business, so access to data is more comprehensive. It plays a more central role in the SC and therefore has access to more comprehensive data. At present, Jingdong Group's SC financing model is more complete than that of its peers, and this paper will explore its financing risks in conjunction with an analysis of Jingbao Bei's financing model.

3. Principles of DCNN Models

3.1. Time-Frequency Transformation. The time-frequency diagram obtained by the time-frequency analysis method of transformation contains a wealth of information about the state of the equipment and intuitively reflects the transformation. The Fourier transform uses a fixed window function and its resolution is also fixed [27, 28], making the short-time Fourier transform somewhat limited. The wavelet basis of the continuous wavelet transform is scalable and solves the time resolution and frequency. The one-dimensional continuous S-transform of the signal $x(t)$ is defined as follows:

$$S(\tau, f) = \int_{-\infty}^{+\infty} x(t) \frac{|f|}{\sqrt{2\pi}} e^{-(t-\tau)\cdot r/2} e^{-j2\pi f t} \mathrm{d}t, \tag{1}$$

where $w(t, f) = (|f|/\sqrt{2\pi})\exp(-t^2 f^2/2)$ is the Gaussian window function. Its corresponding S-inverse transform is as follows:

$$x(t) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} S(\tau, f) e^{-j2\pi f t} \mathrm{d}\tau \mathrm{d}f, \tag{2}$$

where τ means time shift factor, f means frequency, t means time, Gaussian window function $\omega(t, f)$ is both a function of time and frequency, the window width at low frequencies is large with good frequency resolution, and the window width at high frequencies is small with good time resolution.

3.2. Deep Learning Convolutional Neural Networks. Different from the traditional neural network, the deep learning convolutional neural network model has the characteristics of weight sharing. This structure can greatly reduce the complexity of the model and also has the ability to extract features automatically. Compared with the traditional neural network, the deep convolution The performance of the neural network in recognition of two-dimensional images is even better. The functions of each part of the DCNN [29] are shown in Figure 3.

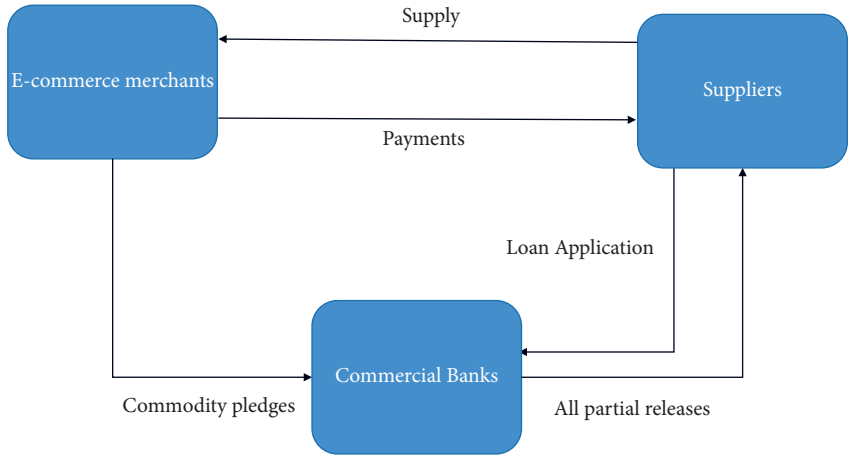


Figure 1: Flow chart of accounts receivable financing.

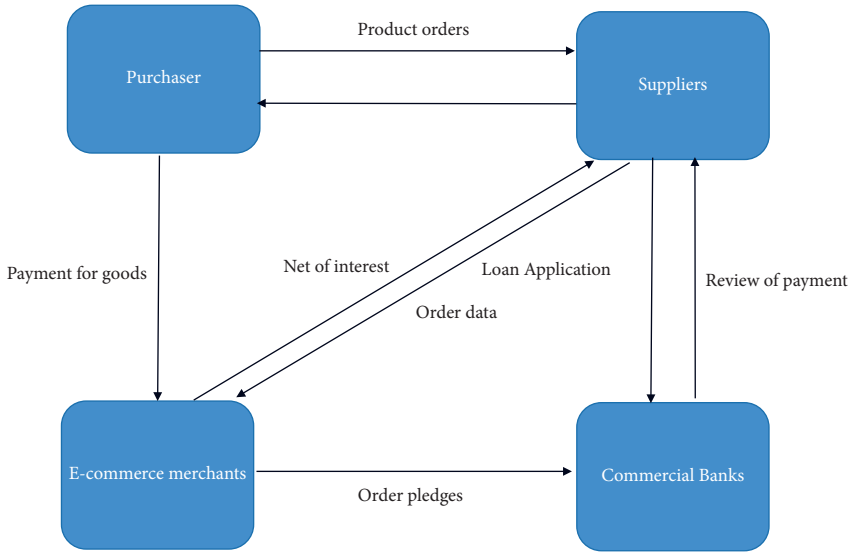


Figure 2: Flow chart of order financing.

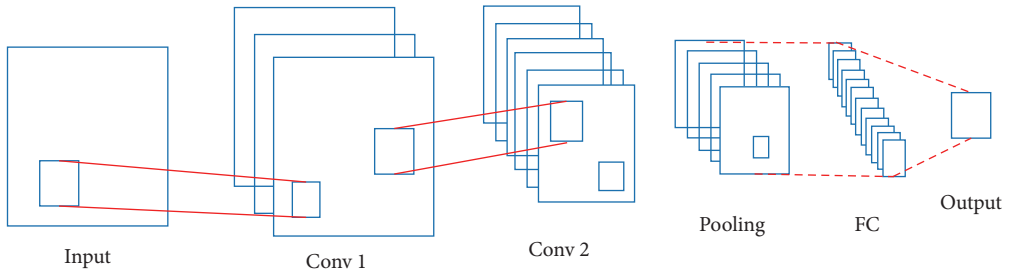


Figure 3: Convolutional neural network structure diagram.

(1) The input layer, which is the spectral image obtained by the time-frequency transform.

(2) The convolution layer, where the input image matrix is convolved with a convolution kernel from the original image. One of the features of the convolution operation is the sharing of weights, which reduces the noise of the original signal through the convolution operation.

(3) The pooling layer, also known as the downsampling layer, is usually referred to as the convolution layer and the pooling layer as the convolution layer. we use maximum pooling to build the pooling layer, through convolutional operations, retaining the main information of the output features of the convolutional layer on the basis of reducing the

number of feature dimensions to improve the network training network.

(4) Fully-connected layers.

(5) The number of nodes in the output layer corresponds to the category of classification, and the output value represents the probability of the corresponding category. The Softmax classifier was chosen to implement the classification task.

3.3. Forward Propagation. The convolutional neural network with sample sets (X, Y), X as inputs and Y as ideal output vectors is forward propagated as follows [30]:

(1) Random initialization of network parameters. Before training, the parameters of the convolutional neural network are initialized: the weights of the convolutional kernel, the bias, and the parameters of the tail perceptron. A small random number close to zero is chosen to ensure that the training is carried out properly, but also to ensure that the weights do not become saturated and lead to training failure.

(2) Calculating the network output.

3.4. Backpropagation. The operation of the deep convolutional neural network model is forward-transmitted, but in the process of adjusting the model parameters, it is from the back to the front. The model adjusts the parameters of the corresponding model layer according to its own recognition accuracy, that is, minimizing the error cost function is used to fine-tune the network parameters to gradually improve the model performance. The error between the ideal output vector Y and the actual output vector O is calculated.

$$e = \frac{1}{2}\sum_{r}^{k=1}\left(Y_k - O_k\right)^2 = \frac{1}{2}\mathbf{Y}_k - \mathbf{O}_k^2, \tag{3}$$

where r represents the total number of categories, Y_k represents the k th dimension of the label corresponding to input X, and O_k represents the k th dimension of the output corresponding to X the input network.

The convolutional neural network error back propagation is trained using a backpropagation algorithm.

(1) Calculate the partial derivative of the error with respect to the weights, backpropagating along the gradient direction with a partial derivative of the following:

$$\Delta\omega_{ij}(n) = -\rho\frac{\partial e(n)}{\partial\omega_{ij}(n)}. \tag{4}$$

Weights updated to the following:

$$\omega_{ij}(n+1) = \Delta\omega_{ij}(n) + \omega_{ij}(n). \tag{5}$$

The gradient can be solved by the partial derivative:

$$\frac{\partial e(n)}{\partial\omega_{ij}(n)} = \frac{\partial e(n)}{\partial e_j(n)}\cdot\frac{\partial e_j(n)}{\partial v_L^j(n)}\cdot\frac{\partial v_L^j(n)}{\partial\mu_L^j(n)}\cdot\frac{\partial\mu_L^j(n)}{\partial\omega_{ij}(n)}. \tag{6}$$

Of which,

$$\begin{aligned} e_j(n) &= \frac{\partial e(n)}{\partial e_j(n)}, \\ \frac{\partial v_L^j(n)}{\partial\mu_L^j(n)} &= g\prime\mu_L^j(n), \\ \frac{\partial\mu_L^j(n)}{\partial\omega_{ij}(n)} &= v_K^j(n). \end{aligned} \tag{7}$$

Forward propagation of error signal: Adjustment of the input layer and implied layer weights ω_{mi}.

$$\Delta\omega_{mi}(n) = \rho\delta_K^j v_M^m(n), \tag{8}$$

where $v_M^m(n)$ is the output of the input neuron:

$$\delta_K^j = f'\left(\mu_K^j(n)\right)\sum_{L}^{j=1}\delta_L^j\omega_{ij}. \tag{9}$$

3.5. Risk Identification Model Based on Deep Learning. Softmax is used as the classification layer, and the classification is performed by the probability $p(y^{(i)} = j/f^{(i)})$ that the sample vector X belongs to the j nd classification (j indicates the number of classes). The output of the classification layer is a k-dimensional vector whose unit values sum to 1, with the following equation:

$$h_{\gamma^r}\left(f^i\right) = \begin{bmatrix} p\left(y^{(i)} = 1|f^i;\gamma_1^T\right) \\ p\left(y^{(i)} = 2|f^i;\gamma_2^T\right) \\ \vdots \\ p\left(y^{(i)} = k|f^i;\gamma_k^T\right) \end{bmatrix} = \frac{1}{\sum_k^{j=1} e^{\gamma^\gamma f^i}}\begin{bmatrix} e^{\gamma^\top f^{(i)}} \\ e^{\gamma_2^T f^{(i)}} \\ \vdots \\ e^{\gamma_i^T f^{(i)}} \end{bmatrix}, \tag{10}$$

where $\gamma_1^T, \gamma_2^T, \ldots, \gamma_k^T$ is the parameter of the iterative regression model, $(1/\sum_k^{j=1} e^{\gamma^\gamma f^i})$ is used to normalize the output, k represents the dimensionality, $e^{\gamma^\gamma f^i}$ represents the exponential function of the current i element, $f^{(i)}$ represents the current i th element, $y^{(i)}$ represents the classification probability of the i th element and $y^{(i)}$ represents the relative probability output of Softmax.

The output of Softmax is logged and the higher the output value, the higher the relative probability of the correct category. The training samples are fed into the DL network and the network parameters are trained to build the deep learning network model. The test samples are fed into the deep learning network to test the classification performance. The actual vibration signal features are fed into the trained deep learning model to obtain the risk identification results of SC finance.

4. Experimental Analysis

4.1. Data Sources. After the selection of risk factors for Jingdong SC finance, that is, after constructing the target layer, subtarget layer, and criterion layer, this paper scores these risk factors through the expert review scoring method. In the selection of risk factors for Jingdong SC finance, this paper is based on the research on SC-related financial risks at home and abroad. Through the advice of experts within Jingdong Finance, six primary indicators and 19 secondary indicators were selected. Questionnaire data from practitioners related to online SC finance and academic researchers in related fields of finance were then used.

4.2. Selection of Risk Factors for SC Finance in Jingdong. Through the results of the analysis of the risk factors of Jingdong SC finance, we selected these risk factors as the object of our study. For Jingdong SC Finance, first of all, our target layer is Jingdong SC Finance Risk. In terms of the macro aspect, the risk of Jingdong SC finance is mainly influenced by the macroeconomic and legal system. Therefore, the macroeconomic environment and the improvement of the legal system are selected as secondary indicators: For the microaspect, the risk of Jingdong SC finance mainly includes of the relationship between the e-commerce platform and bank. Overall credit risk, supply chain relationships, pledge risk, and operational risk all play an important role in the selection of risk factors for Jingdong's supply chain finance. The risk of Jingdong's relationship with banks mainly includes the degree of close cooperation between Jingdong and commercial banks, the current status of commercial banks, and the current status of Jingdong's e-commerce platform. We can represent the risk hierarchy model of Jingdong's SC finance in Table 1.

4.3. Constructing Pairwise Comparison Matrices. The "1–9 scale" is used when comparing two risk factors, i.e., on a scale of 1 to 9. Then, n secondary indicators are compared to obtain a pairwise comparison matrix, which gives the following:

$$A_{ij} = \frac{1}{A_{ji}},$$

$$\mathrm{A} = \left(A_{ij}\right)_{n\times n} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix}. \tag{11}$$

Next, we need to carry out the relative weight calculation, let the relative weight of the first level indicator A correspond to the second level indicator $A_{11}, A_{12}, \ldots, A_{1n}$ of Jingdong SC financial risk is $W_{11}, W_{12}, \ldots, W_{1n}$, and its vector form is $W1 = (W_{11}, W_{12}, \ldots, W_{1n})^T$.

A graph of the linear relationship of these factors to Tier 1 indicator A_2 is shown in Figure 4.

Credit risk A_2 is used here as an example. The results of the two-by-two comparison of the impact of the factors of the Level 2 indicator $A_{21}, A_{22}, A_{23}, A_{24}$ on Level 1 indicator A_2 are shown in Table 2 below.

The comparative effect of the relationship between the secondary indicators is shown in Figure 5.

The weights of the Tier 1 indicators were then collated as shown in Table 3 before proceeding to risk identification.

The comparative weighting relationships for these first-level indicators are shown in Figure 6, which gives a pie chart of the weights twice.

Through the weight values of each level of indicators given in Table 3, it is easy to see that among the level 1 indicators, corporate credit risk has the largest influence on the risk of Jingdong's SC finance: the next major risks are economic environment and legal and regulatory risks. The subsequent risk weights are, in descending order, the relationship risk between Jingdong e-commerce and commercial banks, platform operational risk, and the relationship risk between SC enterprises. Relatively speaking, enterprise pledge risk has the least impact weight on Jingdong's SC finance risk. Meanwhile, it can be seen from Table 3 that, in terms of enterprise credit risk, the secondary indicators that have the greatest weight on the risk of Jingdong's SC finance mainly focus on the credit status of enterprise counterparties; in economic environment and legal regulatory risk, the indicators that have the greatest impact on it are mainly legal regulatory mechanisms. In the risk of the relationship between Jingdong's e-commerce and banks is mainly influenced by the cooperation between Jingdong's e-commerce and banks and the support of commercial banks together.

The data of the above first-level indicators were randomly selected using a DCNN model, in which a total of 15,000 sample data were obtained, and the training and test sets were randomly selected in the ratio of 8 : 2. The experiments were conducted twice, and the specific composition of the data of the two experiments is shown in Table 4.

A comparison of the data set distributions for the two experiments is shown in Figure 7.

On the basis of the two data sets, five experiments were conducted on each data set in order to improve the accuracy of risk identification and to avoid the influence of chance on the experimental results, and the specific identification results are given in Table 5.

A comparison of the recognition results for the two sets of data is shown in Figure 8.

In this paper, the average value of five experiments was finally selected as the identification result of SC finance risk. The comparison results of this paper's method with other methods are given in Table 6.

A visual comparison of the recognition accuracy of the proposed method with other methods is shown in Figure 9.

As can be seen from Figure 8, compared to other research methods, the risk identification accuracy of the

Table 1: Jingdong SC finance risk evaluation system table.

Macroeconomics and legal regulation (A1)	Macroeconomic environment (A11) Legal regulation (A12)
Corporate credit risk (A2)	Corporate credit profile (A21) Corporate solvency (A22) Business development prospects (A23) Corporate credit profile (A24)
Relationship risk between SCs (A3)	Enterprise SC cooperation (A31) SC status (A32)

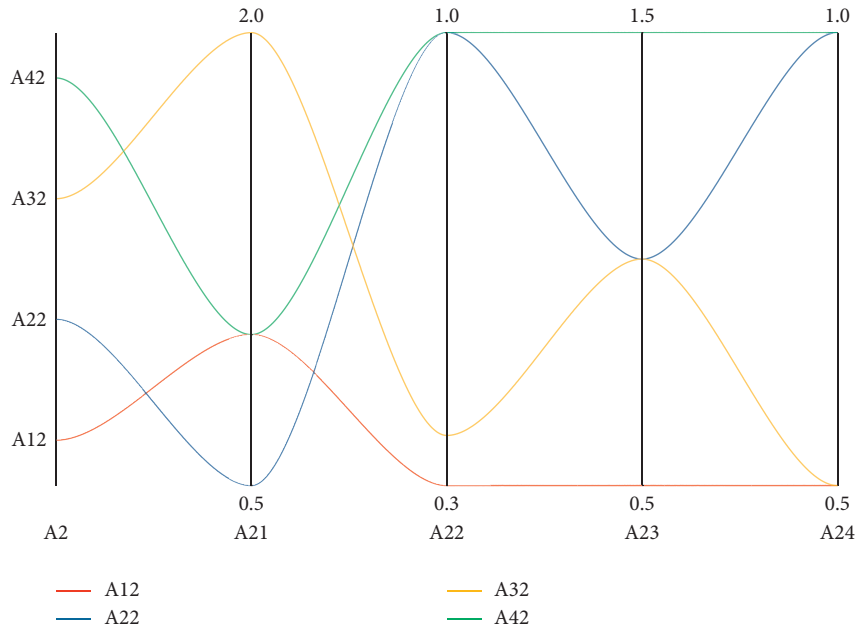


Figure 4: Linear relationship between secondary and primary indicators.

Table 2: Comparison chart of the impact of secondary indicators on primary indicators.

A_2	A_{21}	A_{22}	A_{23}	A_{24}
A_{12}	1	0.25	0.5	0.5
A_{22}	0.5	1	1	1
A_{32}	2	0.333	1	0.5
A_{42}	1	1	1.5	1

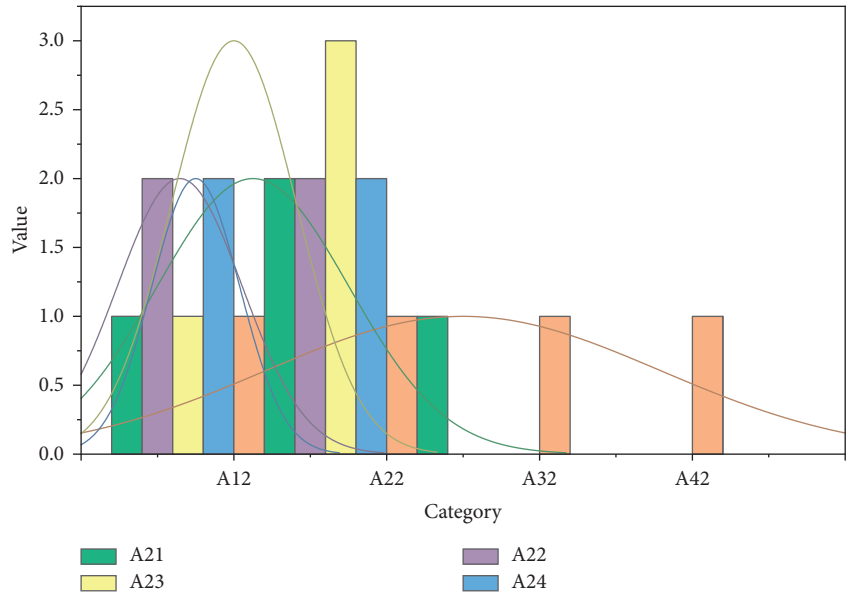


Figure 5: Comparative effect of the relationship between secondary indicators.

Table 3: Distribution of weight values for level 1 indicators.

Tier 1 indicators	Tier 1 indicator weighting values
EELR	0.222
CCR	0.300
ICRC	0.102
CPR	0.058
POR	0.150
RBRB	0.168

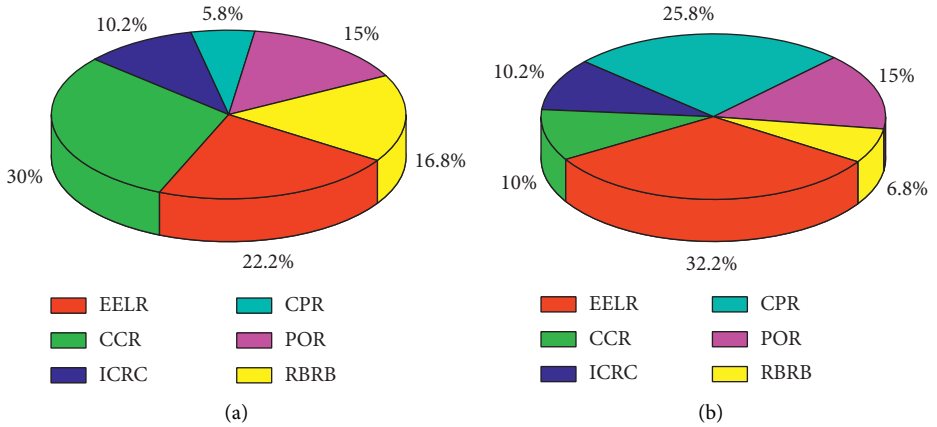


Figure 6: Comparison of the two weighted pie charts. (a) Experiment 1. (b) Experiment 2.

Table 4: Specific distribution of SC finance risk identification datasets.

Category	Experiment 1	Experiment 2
EELR	2000	2500
CCR	3000	2000
ICRC	2500	2000
CPR	3500	3000
POR	1500	2500
RBRB	2500	3000

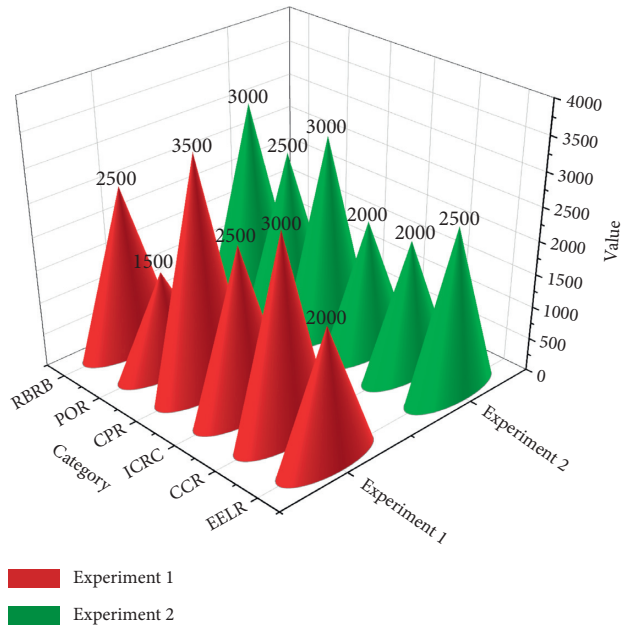


Figure 7: Comparison of the data set distribution between the two experiments.

Table 5: Five experimental identification accuracies for both data sets showing.

Category	Experiment 1 (%)	Experiment 2 (%)	Experiment 3 (%)	Experiment 4 (%)	Experiment 5 (%)
Data1	95.62	94.56	97.14	96.21	93.28
Data2	91.28	93.65	92.54	95.86	94.25

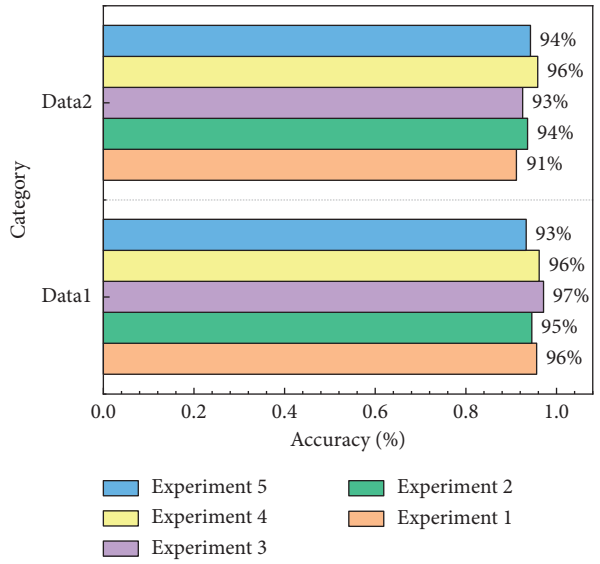


Figure 8: Comparison of the recognition results of the two data sets.

Table 6: Comparison of the results of the proposed method with other methods.

Category	Accuracy (%)
BP	89.26
LetNet-5	90.48
SVM	93.65
CNN	91.43
Proposed model	95.36

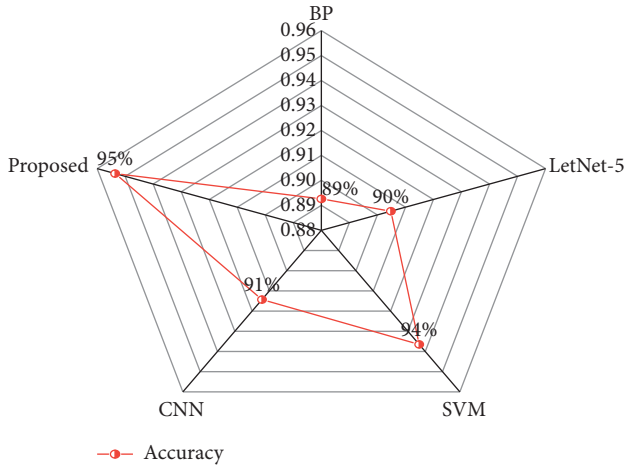


Figure 9: Comparison of the recognition accuracy of the proposed method with other methods.

proposed method has improved significantly, mainly due to the unique model composition produced, which, combined with the optimal weight selection on the data, has led to a further improvement in the identification results and validated the effectiveness of the proposed model.

5. Conclusion

Deep learning provides a way for the risk identification technology of e-commerce SC finance. The risk identification model based on DL proposed in this paper overcomes the problem of cross-interference between the components of the time-frequency analysis method and improves the effectiveness and stability of the time-frequency analysis method. The deep learning model directly processes the time-frequency image, extracts the time-frequency features of the signal, and establishes a two-layer CNN model applied to risk analysis and recognition. The results show that the model can achieve a higher accuracy rate compared with previous algorithms, providing a new way for risk recognition. Although the deep convolutional neural network model used in this paper has shown excellent performance in the process of financial risk identification, it has a longer processing time than other models and is difficult to apply in reality. Therefore, future research work should focus on improving the processing speed of the model, reducing the number of parameters in the model run, and further speeding up the processing speed while improving the recognition accuracy.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

[1] H. Li, S. Zhuang, and Y. Ma, "Benign and malignant classification of mammogram images based on deep learning," *Biomedical Signal Processing and Control*, vol. 51, pp. 347–354, 2019.

[2] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, https://arxiv.org/abs/1409.1556.

[3] G. Huang, Z. Liu, and L. Van Der Maaten, "Densely connected convolutional networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4700–4708, Minneapolis, MN, USA, June 2017.

[4] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le, "Learning Transferable Architectures for Scalable Image Recognition," in *Proceedings of the IEEE Conference on Computer Vision and*

Pattern Recognition, pp. 8697–8710, Saly Lake, UT, USA, June 2018.

[5] C. Szegedy, W. Liu, Y. Jia, and P. Sermanet, "Going deeper with convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9, Boston, MA, USA, June 2015.

[6] J. Deng, W. Dong, and R. Socher, "Imagenet: a large-scale hierarchical image database," in *Proceedings of the 2009 IEEE conference on computer vision and pattern recognition*, no. 5, pp. 248–255, IEEE, Miami, FL, USA, June 2019.

[7] F. Perez, S. Avila, and E. Valle, "Choosing a cnn architecture for melanoma classification," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, Long Beach, CA, USA, June 2019.

[8] Z. Wu, S. Li, C. Chen, A. Hao, and H. Qin, "A deeper look at image salient object detection: Bi-stream network with a small training dataset," *IEEE Transactions on Multimedia*, vol. 24, 2020.

[9] X. Wang, S. Li, C. Chen, Y. Fang, A. Hao, and H. Qin, "Data-level recombination and lightweight fusion scheme for rgb-d salient object detection," *IEEE Transactions on Image Processing*, vol. 30, pp. 458–471, 2020.

[10] G. Ma, S. Li, C. Chen, A. Hao, and H. Qin, "Stage-wise salient object detection in 360 omnidirectional image via object-level semantical saliency ranking," *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 12, pp. 3535–3545, 2020.

[11] Y. Li, S. Li, C. Chen, A. Hao, and H. Qin, "A plug-and-play scheme to adapt image saliency deep model for video data," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, no. 6, pp. 2315–2327, 2020.

[12] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with DCNNspp," in *Proceedings of the Advances in Neural Information Processing Systems*, no. 3, pp. 1097–1105, london, 2018.

[13] S. Khan, N. Islam, Z. Jan, I. U. Din, and J. J. C. Rodrigues, "A novel deep learning based framework for the detection and classification of breast cancer using transfer learning," *Pattern Recognition Letters*, vol. 125, pp. 1–6, 2019.

[14] K. He, X. Zhang, S. Ren, and J. Sun, "Proceedings of the IEEE conference on computer vision and pattern recognition," *Deep residual learning for image recognition*, vol. 50, pp. 770–778, 2016.

[15] S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, "Aggregated residual transformations for deep neural networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1492–1500, honolulu, June 2017.

[16] A. Narin, C. Kaya, and Z. Pamuk, "Automatic detection of coronavirus disease (Covid-19) using X-ray images and DCNNs," *Pattern Analysis and Applications*, vol. 24, Article ID 10849, 2013.

[17] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. R. Acharya, "Automated detection of covid-19 cases using deep neural networks with X-ray Images," *Computers in Biology and Medicine*, vol. 121, Article ID 103792, 2020.

[18] I. D. Apostolopoulos and T. A. Mpesiana, "Automatic detection from X-ray images utilizing transfer learning with convolutional neural networks," *Physical and Engineering Sciences in Medicine. Covid-19*, vol. 43, pp. 1–8, 2020.

[19] L. Wang and A. Wong, "Covid-net: a tailored DCNN design for detection of Covid-19 cases from chest X-ray images," *Scientific Reports*, Article ID 19549, 2020.

[20] B. D. Goodwin, C. Jaskolski, C. Zhong, and H. Asmani, "Intra-model variability in Covid-19 classification using chest X-ray images," p. 02167, 2020, https://arxiv.org/abs/2005.02167.

[21] S. D. Deb and R. K. Jha, "Covid-19 detection from chest x-ray images using ensemble of cnn models," in *Proceedings of the 2020 International Conference on Power, Instrumentation, Control and Computing (PICC)*, pp. 1–5, IEEE, Thrissur, India, March 2020.

[22] M. L. Duca, A. Koban, M. Basten, and E. Bengtsson, "A new database for financial crises in european countries," *European Central Bank Occasional Paper Series*, vol. 194, 2017.

[23] T. Duprey, B. Klaus, and T. Peltonen, "Dating systemic financial stress episodes in the EU countries," *Journal of Financial Stability*, vol. 32, pp. 30–56, 2017.

[24] C. M. Reinhart and K. S. Rogoff, "This time is different: a panoramic view of eight centuries of financial crises," *Ann. Econom. FInance Soc. AEF*, vol. 15, pp. 1065–1188, 2014.

[25] C. M. Reinhart and M. B. Sbrancia, "The liquidation of government debt. Econ," *Polity*, vol. 30, no. 82, pp. 291–333, 2015.

[26] C. M. Reinhart and C. Trebesch, "Sovereign debt relief and its aftermath," *Journal of the European Economic Association*, vol. 14, pp. 215–251, 2016.

[27] C. M. Reinhart and C. Trebesch, "The international monetary fund: 70 years of reinvention," *The Journal of Economic Perspectives*, vol. 30, pp. 3–28, 2016.

[28] L. Ureche-Rangau and A. Burietz, "One crisis, two crises the subprime crisis and the European sovereign debt problems," *Economic Modelling*, vol. 35, pp. 35–44, 2018.

[29] T. D. Willett and C. Wihlborg, "Varieties of European Crises," *Handbook of Safeguarding Global Financial Stability*, Elsevier, no. 3, , pp. 309–322, London, 2019.

[30] Y. Zhao, J. de Haan, B. Scholtens, and H. Yang, "Sudden stops and currency crashes," *Review of International Economics*, vol. 22, no. 4, pp. 660–685, 2018.

Hindawi
Scientific Programming
Volume 2022, Article ID 6775667, 10 pages
https://doi.org/10.1155/2022/6775667

Research Article

A Method Design of English Teaching System Based on Video Feedback Method

Yizhi Li[1] **and Yalan Gou**[2]

[1]*School of Foreign Language and International Business, Guilin University of Aerospace Technology, Guilin, Guangxi 541004, China*
[2]*Language and Literature Department, Guilin University, Guilin, Guangxi 541006, China*

Correspondence should be addressed to Yalan Gou; 857081945@qq.com

Received 20 January 2022; Revised 12 February 2022; Accepted 23 February 2022; Published 11 April 2022

Academic Editor: Baiyuan Ding

The traditional English teaching mode is that the teacher simply imparts textbook knowledge and students understand and absorb it. However, this method has obvious problems; that is, it is difficult to ensure that students can quickly understand the content taught and cannot get fast feedback. The video feedback method is a method of English teaching combined with audiovisual technology. Teachers can use video technology to record key English knowledge through video and then give feedback and explain the learning through video. At the same time, the intelligent video feedback system teaching method will greatly improve the teaching quality of English classroom and make students' learning more fun. This paper mainly designs the English teaching system based on the video feedback method and finally realizes the intelligent feedback scheme of the English teaching system. Firstly, the neural network method, classification method, and video shooting technology are used to extract and predict the characteristics of students' classroom expressions, speech, and so on, and analyze through the video feedback system of English classroom. The research results show that the classification method proposed in this study can better complete the body movements such as student expressions and speech collected by the English teaching system, and the neural network method can more accurately predict and feed back the teaching content through the video feedback method with the largest error. It was only 2.98%, and the linear correlation also reached more than 0.98. The minimum error of the video feedback information is only 0.95%, and the prediction errors of the other two kinds of English classroom information are also 2%. The classification and prediction of intelligent video feedback information have achieved good results.

1. Introduction

The traditional English teaching method is that teachers teach students through blackboard writing through textbooks and teaching plans, and students receive them through listening and recording. Due to the difficulty of English teaching and teaching of other subjects, it is easy for students to lose interest in English subjects [1]. This method of teaching is also difficult to ensure high efficiency, and it cannot timely reflect the effectiveness of teachers' teaching methods. In recent years, with the development of computer hardware [2] technology, computer-aided systems have improved the way of English teaching, and the teaching efficiency has been improved to a certain extent [3]. Compared with the traditional English teaching mode, this mode only changes the medium of teaching. It also has certain defects. It cannot provide timely feedback according to the actual situation in the English classroom, nor can it provide real-time feedback according to the actual situation in the English classroom [4]. Actions shown are recorded and given feedback. Both of these two methods limit the further improvement of English teaching efficiency, which also limits the improvement of students' interest in learning English. The video feedback method is a new English teaching mode, which will change the disadvantages of the traditional teaching mode [5].

The video feedback method has been widely used in many applications, such as tennis teaching classrooms,

traffic teaching classrooms, and other fields, which has shown certain feasibility and advantages in other fields [6]. The video feedback method is a teaching feedback method based on audiovisual technology. It can record the teaching content to form the classroom content used in teaching and can realize functions such as playback, zooming, and slow playback [7]. For the application of the video feedback method in the English teaching classroom, it can record a course, and the teacher analyzes and slows down the English video through computer-aided calculation. For difficult content and knowledge points, teachers can teach by means of pause, slow release, and circulation, which can greatly improve students' memory ability and learning interest [8]. At the same time, machine learning methods have developed rapidly in recent years, and many intelligent classification and prediction methods have been derived. The English teaching system can combine video feedback methods and integrate elements of intelligent classification and prediction to realize intelligent English feedback method teaching [9]. In this way, the responses of students and teachers in the English teaching classroom can be recorded in real time, and the responses can be predicted and analyzed through the terminal of the intelligent computer-aided system.

In recent years, the English teaching mode has undergone great changes. It is committed to finding a teaching mode that is more efficient and more suitable for students. A lot of research has also been carried out on the research of English teaching methods [10]. Xu and Tsai [11] took the difficulties in college students English vocabulary learning as the research breakthrough point; they analyzed the relationship between college students and electronic media and the impact of interactive teaching mode on college students' vocabulary memory. At the same time, they studied the theory of multimedia-based interactive English teaching mode and intervention model for English learning adaptation. Zhao [12] believes that oral English teaching is the worst part of multimedia teaching. In response to this problem, he proposed a data mining teaching model. This model first uses the DBN network to send information to the DBN-DELM network, which significantly improves the multimedia performance. The efficiency of the English teaching mode also improves the learning interest in pronunciation and accent in oral English teaching. In view of the problems existing in computer English teaching, Liu [13] used SPOC method to research and analyze the problems in IT English teaching, from the perspectives of collecting student data, uploading relevant English resources, and doing a good job in teaching design. The conclusion shows that the SPOC flipped English teaching method proposed by him improves the effect of daily English teaching and also improves students' satisfaction with English classroom and daily English learning time. Li [14] believes that the teaching level of oral English represents the level of English teaching, and good oral ability is a value for evaluating English teaching. He has used multi-interactive multimedia technology to take the environment, language sense, emotion, and other factors into consideration in English teaching. The two interactive modes of classroom interaction and vitality improve the interaction of English teaching compared with the traditional oral English teaching mode. Zhang [15] has used image recognition technology to design a new way for English teaching mode. Image superresolution technology GAN model is used by him to reconstruct superresolution images from low-resolution images in English teaching. At the same time, he established a mixed-sample spine regression model to estimate the behavioral characteristics of learning in English teaching, and the validation test showed the feasibility and accuracy of this method. Han [16] studied the feasibility of online English teaching mode by combining deep learning methods and remote supervision, and he simulated and analyzed the application of supervised learning algorithms in English teaching. Two perspectives, student evaluation and teacher evaluation, were used to verify the feasibility of the model. The conclusion shows that this model is suitable for the online teaching task of English. Xie and Ma [17] improved the traditional MOOC English teaching model based on cloud computing technology and artificial technology and corrected and verified the MOOC model according to the situation in English cross-cultural communication and the needs of the online teaching model. The results show that the improved MOOC model can improve English. The efficiency of teaching cross-cultural communication. Zhao [18] found that the demand for English in all walks of life is relatively high. In response to this problem, he proposed a SPOC English teaching model and carried out experimental verification. This model has good applicability in vocational English teaching. Gao [19] introduced the data mining software Clementine into the English teaching work, extracted the information hidden in the English teaching evaluation system, and determined the minimum support degree and the minimum confidence degree. This method has high performance in English teaching evaluation management. The applicability can better reflect the actual value of teachers.

Through the above review of the design and research of the English teaching method system, it can be found that the current English teaching research mainly focuses on the research of computer-aided systems, and a small amount of research involves the research on the English teaching method of the video feedback method, the English teaching mode based on the intelligent video feedback method will change the shortcomings of the traditional teaching model of single teaching, and it can form a feedback learning system. For students, they can discover the defects and problems of learning English through the video feedback system in a timely manner, thereby making corrections in a timely manner [20]. There are few English teaching modes. Aiming at the above-mentioned research status and the shortcomings of current English teaching, this paper designs an intelligent video feedback English teaching mode [21]. The English teaching system based on the video feedback method will use the video recording technology to record the speech, body movements, and teaching courseware content of the students and teachers in a course to form a video [22]. Then, these videos will be classified and intelligently predicted and analyzed through computer-aided systems and intelligent Internet technology, and the recorded videos will be displayed to students in the form of videos [23]. Teachers

can slow down and play back the important and difficult knowledge of the videos to achieve English. Teaching feedback technology and intelligent Internet technology can provide intelligent algorithms and knowledge from the Internet to students [24].

This paper mainly designs and predicts the English teaching system based on the video feedback method and intelligent algorithm. The first part mainly introduces the defects of the English teaching mode and the advantages of the video feedback method for English teaching [25]. The second part introduces the significance of intelligent video feedback technology for English teaching. The third part explains the method and process of the English teaching mode based on the video feedback method. The fourth part analyzes the feasibility of intelligent video feedback technology in English teaching from the perspective of accuracy and error, and the last part is the summary of the article.

2. The Significance and Necessity of Intelligent Video Feedback Method for English Teaching and Data Sources

2.1. The Necessity of Intelligent Video Feedback Method to Improve English Teaching. Video feedback is a new technology that integrates audiovisual technology. Its development benefits from the rapid development of computer hardware equipment and streaming media technology [26]. The video feedback method has been applied in many fields, and it shows good results. English teaching is a relatively cumbersome subject among many subjects. It is difficult to stimulate students' interests and hobbies. If we only rely on traditional teaching methods, because it is a language subject, it has a lot of cumbersome grammatical information and expression habits. Information is a discipline that requires long-term persistence [27]. At the same time, in English teaching, there are many boring grammar sentences, and so on, which require students to memorize and review constantly. If a new English teaching mode is produced, it can not only impart English knowledge to students, but also stimulate students' interest in learning, which is a suitable method for English teaching. The computer-aided system teaching method is a relatively new teaching mode [28]. It can transmit English [29] knowledge to students through video, but it has certain defects. It cannot be carried out according to the difficult points of each student and most students. Targeted teaching cannot show the performance of students in real time [30]. The video feedback method can record students' performance and classroom content in the form of video, which will be analyzed and displayed to students in a targeted manner. At the same time, the feedback mechanism of the video feedback method can well integrate the algorithm of intelligent prediction and classification, which can better find solutions to similar problems according to the students' response [31]. This method can not only effectively analyze students' performance through video, but also match students' difficult points in a similar context, which greatly stimulates students' interest in learning English and their understanding of English. And it also can improve the memory ability of trivial knowledge [32]. The traditional English teaching mode has been going on for many years, but the current social English teaching mode is different from the previous teaching of grammar, sentence patterns, and so on, and now the teaching of spoken English is more demanding. The expression of spoken English has a greater correlation with students' emotions and mouth shapes. The video feedback method can better show students' performance to students for error correction, which cannot be achieved by traditional teaching models.

2.2. Data Sources of English Video Feedback Teaching Analysis. The purpose of this paper is to use video technology to collect audiovisual information of students and teachers in English classroom and courseware content, process it into video and save it. At the same time, this video information needs to be classified and predicted in combination with intelligent algorithms, and finally the information is displayed to students and teachers through the computer-aided system terminal for them to learn and summarize feedback knowledge. Video technology can collect students' language and image information, such as speech and body movements. Intelligent algorithms need to extract features from these images and language information and perform intelligent prediction and classification. For the English teaching classroom, the content of the courseware and the verbal responses and expressions of the students are more critical. Therefore, this paper needs to use the content of the courseware, the expressions of the students, the physical movements of the students, and the words of the teachers saved by the video technology as the data sources of the intelligent classification prediction algorithm. First of all, intelligence needs to effectively classify the four types of information obtained from video technology and use the decision tree method, and then combine the Internet technology to predict the classified information, and then analyze the difficulties faced by students in English learning, and then feed back through the video. Technology will effectively learn these difficult points. This method can not only extract effective information according to the content of students and courseware, but also analyze difficult points in combination with the Internet. At the same time, it can also combine the computer-aided system to realize the English teaching mode of video feedback method, which will be more pertinent.

3. Algorithms and Technologies Used in Intelligent Video Feedback English Teaching Method

3.1. Introduction to the Video Feedback Method. Feedback teaching method is a teaching method combining system statistics theory, information theory, and cybernetics, which can guide teachers and students to be in a relatively harmonious teaching environment, rather than the traditional teaching-listening mode. In this teaching mode, students need to spend time to obtain certain learning outcomes and feed back these learning outcomes to learners, a way for

learners to give feedback and improve learning based on the results. The video feedback method is a special method in the feedback teaching method. It uses audiovisual technology to record the behavior and posture of the students during the learning process. Teacher will show these videos to the students through the computer-aided system and guides and corrects them. Students make adjustments to their learning styles based on this feedback. After students make adjustments, teachers will further video record their learning behaviors, which can supervise students' bad habits and poor learning outcomes in the process of learning, and correct them in time. The video feedback English teaching method adopted in this study not only uses audiovisual technology to record and feedback students' learning process, but also provides a feedback method of intelligent analysis. This model will automatically classify images captured by video technology. And make intelligent predictions on the classification results. Figure 1 shows the working mode of intelligent video feedback English teaching. First, teachers will use video to record students' speech, body movements, courseware content, and so on, in the process of learning as the data source of intelligent analysis algorithm. The tree and intelligent prediction methods give feedback on the learning effect of learning, and these feedback results will be displayed to students for viewing and analysis through the computer-aided system. This method will not only realize the feedback technology in the English teaching process, but also realize the intelligent interaction effect in the English teaching.

3.2. Classification Method Decision Tree for Intelligent Video Feedback. The video feedback method is to record students' learning behaviors through videos, which will contain many types of data, such as students' expressions, speech, and courseware. If it is not classified effectively, inputting these data directly into the intelligent prediction model will produce poor results, which is not good for intelligent feedback in English teaching. Therefore, it is necessary to classify the student behavior data collected by video technology. There are many classification methods with good performance, such as clustering, decision tree, support vector machine, and other methods. Because the data in this study was collected through video technology, and the types of features that need to be classified are relatively obvious, such as student speech, courseware teaching content, student expressions and other behavioral information. Therefore, this study will adopt decision tree as an effective classification of video information in English teaching in this paper. The purpose of the classification method in this paper is to classify the English classroom information collected by video through decision tree, and then input these classification features into the neural network to predict it, and finally achieve the purpose of intelligently assisting video English teaching, rather than simply using video feedback method. Through the intelligent video feedback method, relevant information from the Internet can be collected, which will make students more interested in receiving relevant information about English teaching. Figure 2 shows a schematic diagram of classifying the information collected by video technology through the decision tree method. It can be seen that the English video information is effectively divided into student speech, student expression, courseware content, and other information. At present, in the field of English teaching, the most common classification algorithms are decision trees and clustering methods. The clustering methods are mainly based on distance and density-based classification methods. However, this paper uses the intelligent video feedback system to collect student behavior information. There are obvious differences in the characteristics, and the decision number method is more suitable.

Entropy represents the measure of uncertainty. The smaller the entropy, the better the classification effect of the decision tree. This is an important evaluation index of the decision tree. The following equation shows the expression of entropy:

$$H(D) = -\sum_{l=1}^{L} \frac{|C_l|}{|D|} \log_2 \frac{|C_l|}{|D|}. \tag{1}$$

Equation (2) illustrates the expression for conditional entropy, which is a measure of uncertainty that represents D in the case of condition A. Conditional entropy is also an important evaluation index in decision trees.

$$H\left(\frac{D}{A}\right) = \sum_{J=1}^{n} \frac{|C_j|}{|D|} H(D)_j. \tag{2}$$

Information gain describes the degree to which the uncertainty of feature A for classification dataset D is reduced, which is also a conditional probability event. It is also an important indicator for the classification of English video information, especially for video feedback. The following equation shows the expression for the information gain:

$$g_r(D, A) = \frac{g(D, A)}{H_A(D)}. \tag{3}$$

For the classification task of the English video feedback problem, the index of the probability distribution of the sample points belonging to the Lth class is defined as the Gini index, and the expression is shown as follows:

$$\text{Gini}(p) = \sum_{l=1}^{L} p_k(1 - p_k) = 1 - \sum_{l=1}^{L} p_k^2. \tag{4}$$

The samples of video information in English teaching work are often displayed in the form of sets, and the expression of the Gini index of the sample sets is shown as follows:

$$\text{Gini}(D) = 1 - \sum_{l=1}^{L} \left(\frac{|C_k|}{|D|}\right)^2. \tag{5}$$

The following equation reflects the probability Gini index of a conditional distribution. For classification tasks with conditions, this Gini index is often used to express as

$$\text{Gini}(D, A) = \frac{|D_1|}{|D|} \text{Gini}(D_1) + \frac{|D_2|}{|D|} \text{Gini}(D_2). \tag{6}$$

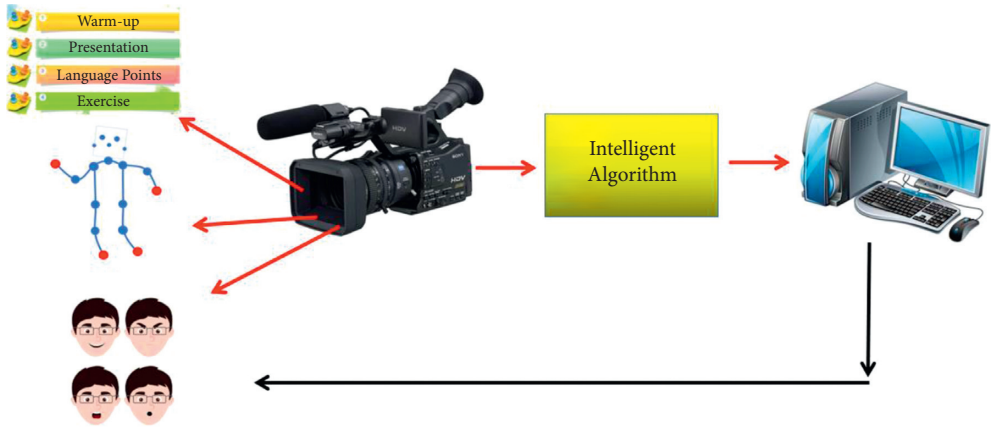


Figure 1: The workflow of intelligent video feedback English teaching method.

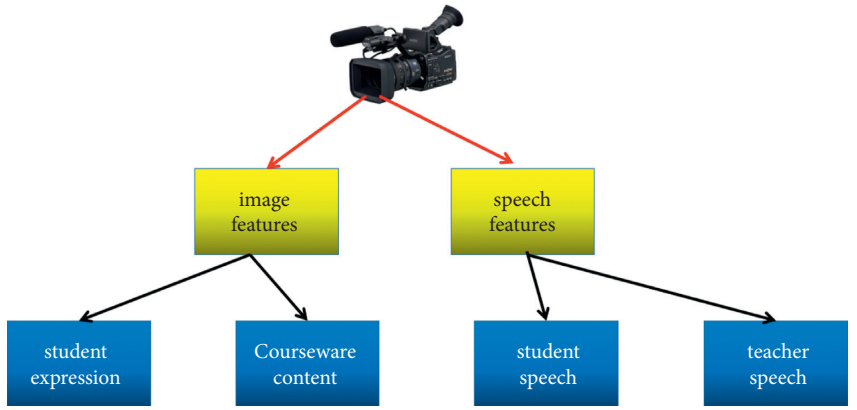


Figure 2: The classification process of English video information through decision tree.

3.3. The Prediction and Extraction Method of English Teaching Video Feedback Information. This research is not only to realize the English teaching mode of video feedback method, but also to realize the intelligentization of video feedback by combining intelligent algorithms. Teachers use audiovisual technology to classify the collected student behavior information and courseware content in the English classroom and extract these behavior features through intelligent algorithms and then feed the information back to students. The intelligent prediction method adopted in this paper adopts the convolutional neural network method. The convolutional neural network has been proved to have obvious advantages in extracting features, and it allows deeper networks to extract video features more accurately. The input of the convolutional neural network is the classification data of English classroom video information obtained by the decision tree method in the previous section, and the output is the analysis of the students' effective behavioral characteristics, which will be used as the final result of the video feedback method, and then used for students and teachers. And then students and teacher will use it for reference. Figure 3 shows the prediction process of video feedback information for English teaching through a convolutional neural network.

Convolutional neural network is also a special algorithm of perceptron, which is also a nonlinear operation in accordance with weight and bias. Equation (7) shows the operation rules between layers of convolutional neural network. The product of the input data and the weight plus the bias is summed, and the output is passed through an activation function. $x_i^{\varsigma-1}$ is the input datasets, and k_{ij}^{ς} is the weights of the model. f is the mapping relations.

$$x_j = f\left(\sum_{i\in M_j} x_i^{\zeta-1} * k_{ij}^{\zeta} + b_j^{\zeta}\right). \tag{7}$$

In a convolutional neural network, there is a pooling layer whose purpose is to extract features with strong correlation to reduce the amount of computation and reduce the risk of overfitting. The pooling layer generally has two ways of upsampling and downsampling. Equations (8) and (9) illustrate the process of upsampling and downsampling, respectively.

$$\delta_j^{\zeta} = \beta_j^{\zeta+1}\left(f'(u)_l^{\zeta} \circ \mathrm{up}\left(\delta_j^{\zeta+1}\right)\right), \tag{8}$$

$$x_j = f\left(\sum_{u,v} \beta_j^{\zeta} \mathrm{down}\left(x_i^{\zeta-1}\right) + b_j^{\zeta}\right). \tag{9}$$

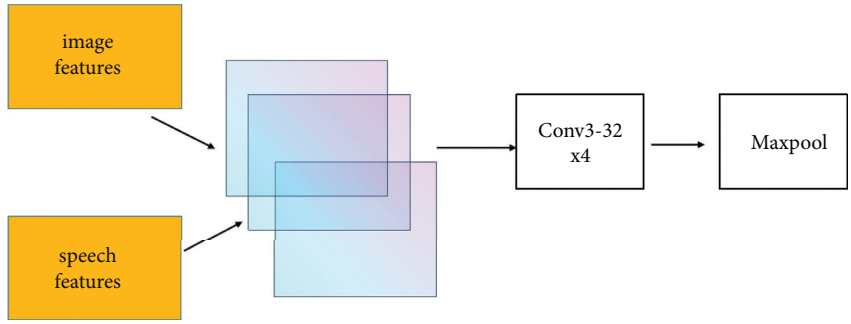


Figure 3: The application process of convolutional neural network in English teaching video information extraction.

The following equation illustrates the automatic derivative operation function during the convolution operation:

$$\frac{\partial W}{\partial b_j} = \sum_{u,v} \delta_j^{\zeta} uv. \tag{10}$$

For student speech, a common information feature in English classrooms, it has obvious temporal characteristics. Equations (11) and (12), respectively, illustrate the processing flow for the temporal feature of student speech, which can memorize historical information. W is the weights of the model, and C_t is the state data at the moment. And h_{t-1} is the historical status information data.

$$o_t = \sigma\left(W_{xo} * x_t + W_{ho} * h_{t-1} + W_{co} \circ C_t + b_o\right), \tag{11}$$

$$h_t = o_t \circ ELU\left(C_t\right). \tag{12}$$

3.4. Data Processing of English Teaching Video Feedback System. From the description of the above three sections, it can be seen that the video feedback English teaching method designed in this study is not only a single video feedback idea, but also an intelligent video feedback system. Two intelligent algorithms are involved in this process, so it is necessary to perform normalization processing and feature normalization processing on the information data collected by these video feedback systems. The video feedback system will collect information such as words, expressions, and courseware content learned in the English classroom. It can be clearly seen that these characteristics are not within an order of magnitude range. In order to more accurately feed back students' English classroom behavior information, it is necessary to process the video data in the process of intelligent classification and prediction so that they keep the same distribution and the same magnitude range. In this study, the standard normal distribution method is used to preprocess the information data of the English video feedback system in order to better distribute the weight distribution.

4. The Feasibility and Accuracy Analysis of Intelligent Video Feedback English Teaching Method

The English teaching mode based on the video feedback method can be classified as the image features and language features by the decision tree method, and the students' expressions, speech, and courseware content can be classified well. This is the first step to realize the intelligent video feedback system. Figure 4 shows a schematic diagram of the classification of the four features collected by the video feedback system. It can be seen that the classification errors of all English classroom information features are within an acceptable range, and the maximum error is only 2.98%. This error comes from students since the speech characteristics of students are constantly changing with time and courseware content and the emotional characteristics of students; this characteristic information is more difficult to classify. The classification errors for the other three types of features are all within 2%. These three types of features are mainly courseware content, student expressions and actions, and so on. It can be seen that the correlation between these features is relatively large, and the variability of these features over time is also relatively large. Therefore, the classification error of these three features is relatively small, and the smallest error is only 0.95%. Compared with the behavioral information of students, the content of the courseware used in teachers' teaching has little change over time, and the error is 1.79%, and the error of another kind is only 1.48%. In order to improve the classification error of students' speech features, the weight of this part can be appropriately increased or more ratios can be collected. It can require to increase the number of sample features. Figure 5 shows the distribution of hotspots after classification of four different English classroom features. It can be clearly seen that the distribution of hotspots in different locations and time points is relatively consistent. In general, the feature information of the English video feedback system can be better classified by means of decision tree.

After the feature information of the English teaching video feedback system is classified by the decision tree method, this feature information needs to be input into the intelligent prediction system to match the information on

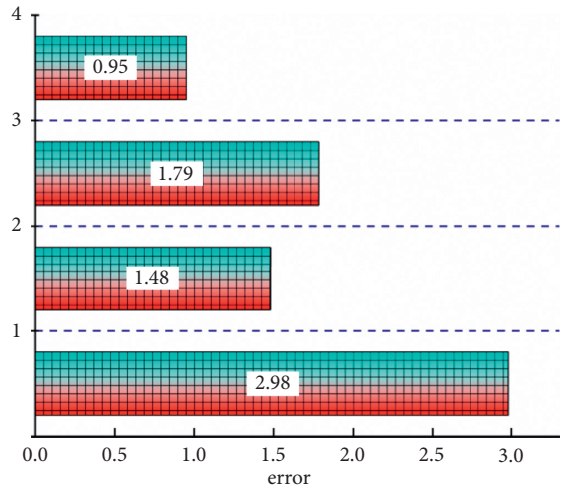


Figure 4: The classification error of English video feedback system features.

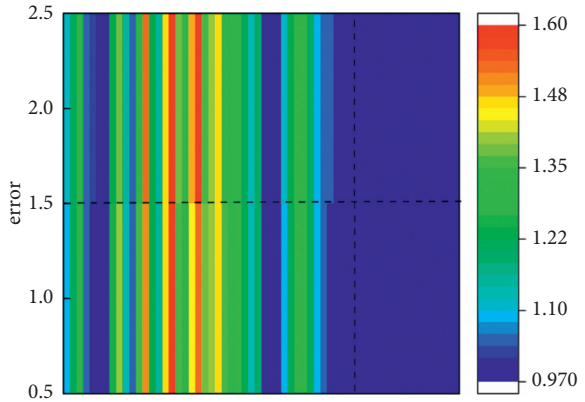


Figure 5: The classification hotspot distribution map of English video feedback system features.

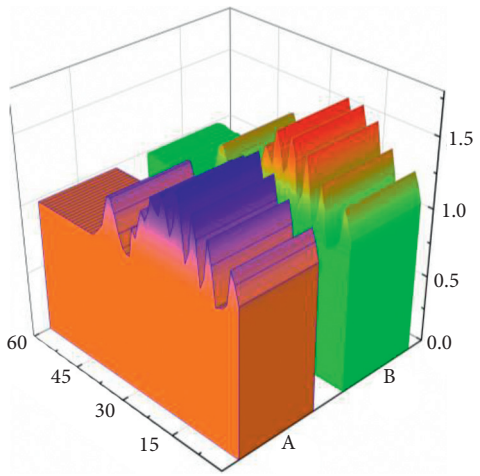


Figure 6: The prediction value of characteristics of intelligent video feedback system in English teaching.

the Internet. This part mainly shows the prediction accuracy of the English teaching video feedback features. Figure 6 shows the change trend between the predicted value and the

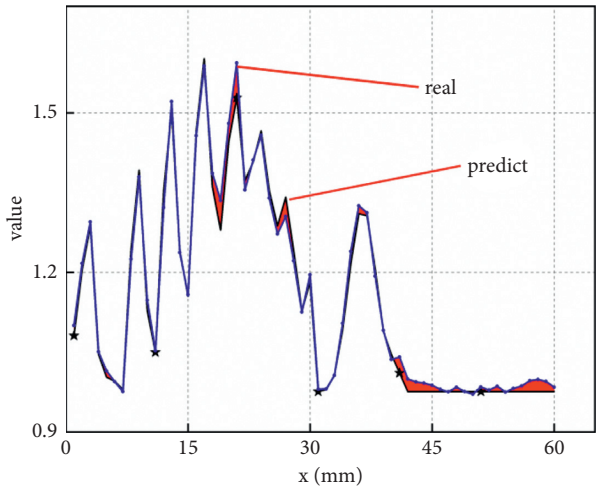


Figure 7: The prediction curve and real curve of characteristics of intelligent video feedback system in English teaching.

actual value of the students' speech features of the English teaching video feedback system. The reason for choosing this student speech feature as a prediction is that this feature is less efficient in classification. This feature is a feature that changes with time. It can be clearly seen from Figure 6 that the predicted student speech is in good agreement with the actual student speech feature, whether it is the change trend of the feature or the peak and trough of the speech feature. Overall, this intelligent prediction algorithm can well match the characteristics of students' speech information collected by the English teaching video feedback system. Figure 7 shows the predicted curve and the actual change curve trend of the feature of student speech. From Figure 7, it can be seen more intuitively that the error between these two values is relatively small; that is, the area of the red area is relatively small. The main error of students' speech is mainly concentrated in the latter part of the video system information, which is mainly due to the increasing error caused by the accumulation of time. However, in general, this error is

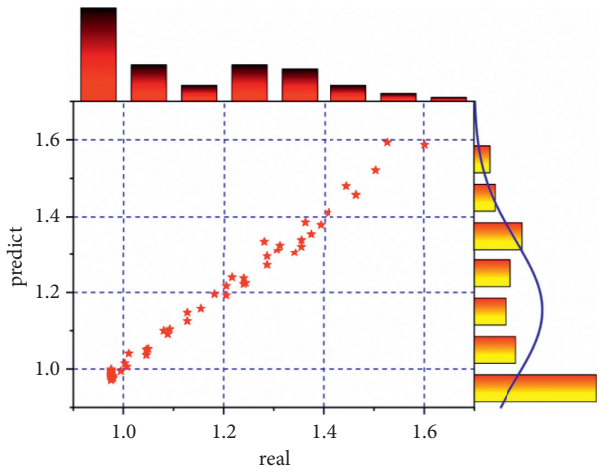


Figure 8: The linear correlation for intelligent video feedback system for English teaching.

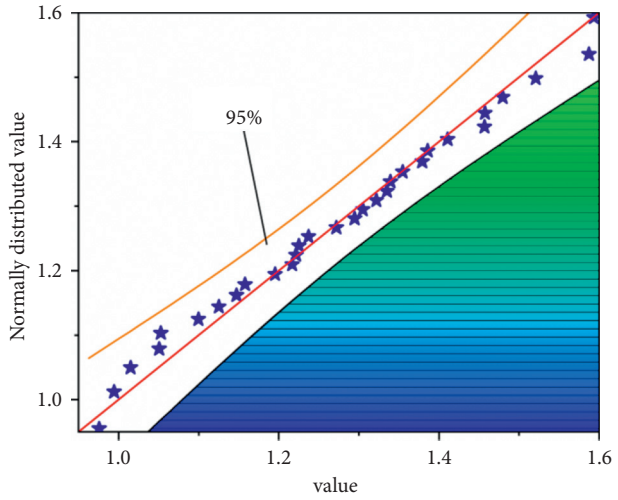


Figure 9: The normal distribution of feedback characteristics of English teaching videos.

acceptable for an intelligent video feedback system for English teaching. From Figure 7, it can be clearly seen that the error is relatively large at the inflection point of the curve, which may be due to the existence of variables that change greatly over time in the English teaching classroom, such as students' behavior information, which leads to a large error.

Linear correlation is an important indicator to measure the prediction accuracy of the video feedback system. It can more intuitively reflect the difference between the student's speech and the real student's speech, which can ensure that the video feedback system will more realistically respond to students' real learning of English situation. Figure 8 shows a plot of the linear correlation coefficient between the predicted and actual speech of the students. From Figure 8, it can be clearly seen that the two data maintain a good linear correlation, and the data points are well distributed on both sides of the linear function, which proves that this intelligent algorithm is suitable for intelligent video feedback systematic. It can also be clearly seen from Figure 8 that the correlation between the prediction of the video feedback information of the students' speech and the actual data has exceeded 0.98, which proves that this model has good accuracy in the English teaching video intelligent feedback system. Figure 9 shows the normal distribution of the feature information prediction of the English teaching video feedback system. It can be clearly seen that the predicted values are well distributed in the normal interval, and the confidence level has reached 95%. This further proves the feasibility of this intelligent prediction algorithm in the English teaching video feedback system.

5. Summary of Intelligent Video Feedback English Teaching Plan

The knowledge will be imparted by a peer-to-peer format for the traditional English teaching model, which is difficult to ensure that each student can more efficiently learn English knowledge, especially in such a tedious subject as English. Computer-aided systems only change the mode of teaching, and it cannot reflect the actual learning situation of students. The video feedback method is a specific audiovisual technology learning feedback method. Students and teachers can analyze and timely feed back according to their own learning status and achievements and then correct the shortcomings in the English learning process.

Aiming at the characteristics of the video feedback method and the characteristics of English subjects, this research designs an intelligent English teaching video feedback scheme and classifies and predicts the characteristics of video information. Teachers can use cameras to record students' speech, students' facial expressions, and courseware content in English class, and then the information obtained from the video recording can be well matched to the actual situation of students through decision tree and intelligent prediction method, and finally the computer-aided system can be used to record the students' actual situation. Feedback information is output, and students can conduct targeted analysis and corrections. Both decision tree and intelligent prediction method are more suitable for classification and prediction of English teaching video feedback system information. The maximum error of decision tree classification is only 2.98%, which appears in the behavioral feature of students' speech. For the prediction of video information, the intelligent prediction algorithm also matches the changing trend of the actual students' speech characteristics and the peak and valley values, and the correlation coefficient also achieves a good effect.

Data Availability

The dataset can be accessed upon request to the corresponding author.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This study was supported by the 2021 Undergraduate Education Reform Project of Guangxi, Research and Practice of Ideological and Political Education of College English Courses in Application-Oriented Colleges and Universities in Guangxi (no. 2021JGA346).

References

[1] L. H. Li, "Design of college English process evaluation systembased on data mining technology and internet of things," *International Journal of Data Warehousing and Mining*, vol. 16, no. 2, pp. 8–13, 2020.

[2] S. Zou, "Designing and practice of a college English teaching platform based on artificial intelligence," *Journal of Computational and Theoretical Nanoscience*, vol. 14, no. 1, pp. 104–108, 2017.

[3] L. Diao and P. Hu, "Deep learning and multimodal target recognition of complex and ambiguous words in automated English learning system," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 4, pp. 7147–7158, 2021.

[4] M. Zhang, "Application of brain neuroscience in the discussion of multimedia English teaching mode," *Neuro-Quantology*, vol. 16, no. 5, 2018.

[5] M. S, K. Mujeeb, R. Madhavi, and S. Praveen, "An empiricalstudy of the big data classification methodologies," *International Journal of Bioinformatics Research and Applications*, vol. 16, no. 2, pp. 195–215, 2020.

[6] D. Qiu, "A study on English teaching model under thebackground of big data," *Journal of Physics: Conference Series*, vol. 1744, no. 3, 2021.

[7] Z. Turan. and B. Akdag-Cimen, "Flipped classroom in englishlanguage teaching: a systematic review," *Computer Assisted Language Learning*, vol. 33, no. 6, pp. 590–606, 2020.

[8] S. Franceschini, P. Trevisa, L. Ronconi, and S. Bertoni, "Action videogames improve reading abilities and visual-to-auditory attentional shifting in English-speaking children with dyslexia," *Scientific Reports*, vol. 7, no. 1, pp. 1–12, 2017.

[9] T. Hu, "Design of English teaching resource ManagementSystem based on collaborative recommendation," *Techniques of Automation and Applications*, vol. 38, no. 09, pp. 162–165, 2019.

[10] M. X. Wang, J. Li, and Q. Y. Jia, "Design andimplementation of an assistant teaching system for englishcreative writing based on dynamic reference recommendation," *Computer Applications and Software*, vol. 36, no. 02, pp. 104–108, 2019.

[11] D. Y. Xu and S. B. Tsai, "A study on the application of interactive English-teaching mode under complex data analysis," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 2675786, 12 pages, 2021.

[12] H. X. Zhao, "Construction of multimedia-assisted English teaching mode in big data network environment," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 1609187, 10 pages, 2021.

[13] L. L. Liu, "Research on IT English flipped classroom teaching model based on SPOC," *Scientific Programming*, vol. 2021, Article ID 7273981, 9 pages, 2021.

[14] N. Li, "Construction and application of the multi-intermediate multi-media English oral teaching mode," *Intelligent Automation & Soft Computing*, vol. 26, no. 4, pp. 807–815, 2020.

[15] F. Zhang, "Innovation of English teaching model based on machine learning neural network and image super resolution," *Journal of Intelligent and Fuzzy Systems*, vol. 39, no. 2, pp. 1805–1816, 2020.

[16] Y. Han, "Evaluation of English online teaching based on remote supervision algorithms and deep learning," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 4, pp. 7097–7108, 2021.

[17] H.Y. Xie and Q. Ma, "College English cross-cultural teaching based on cloud computing MOOC platform and artificial intelligence," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 4, pp. 7335–7345, 2021.

[18] M. M. Zhao, "Construction and research design of vocational English blended teaching model based on SPOC," in *Proceedings of the 2019 7Th International Conference On Information And Education Technology (Iciet 2019)*, pp. 238–243, Aizu-Wakamatsu, Japan, March 2019.

[19] F. Gao, "Establishment of college English teacher's teaching ability evaluation based on Clementine data mining," *Journal of Intelligent & Fuzzy Systems*, vol. 38, no. 6, pp. 6833–6841, 2020.

[20] F. Zhang, L. L. Dong, C. X. Wang, N. Ye, and Q. He, "Researchon project driven it English teaching mode for software engineering majors," *Education Teaching Forum*, vol. 43, no. 5, pp. 192–194, 2019.

[21] F. Yu, "Research on the innovation of ESP English blendedteaching under the background of informatization-taking itEnglish as an example," *Campus English*, vol. 29, no. 1, pp. 36-37, 2019.

[22] L. Shao and X. Liu, "E application of ESP in college Englishteaching-taking it English as an example," *Science & Education*, vol. 56, no. 4, pp. 192-193, 2019.

[23] J. G. Liu, "A study on the teaching mode of college Englishunder the background of modern science and technologydevelopment," *Food Research and Development*, vol. 41, no. 21, pp. 258-259, 2020.

[24] S. N. Chu, J. X. Chen, and Y. H. Liu, "Research on the practiceof science and trade English translation teaching under thetranslation studio mode," *Educational Research*, vol. 3, no. 11, pp. 108-109, 2020.

[25] M. Nophawu, C. P. Madoda, T. s. P. Baba, and L. Nhlanhla, "Analysing English first additional language teachersunderstanding and implementation of reading strategies," *Reading and Writing*, vol. 9, no. 1, pp. 1–10, 2018.

[26] Z. Wen and L. Li, "A study on SPOC-based blended English teaching mode," *Journal of Chengdu Normal University*, vol. 35, no. 10, pp. 59–64, 2019.

[27] K. Nasradin, M. V Michos, and V. B. Markovic, "Traditional language teaching versusICT oriented classroom," *In Sinteza 2019-international scientific conference on information technology and data relatedresearch*, vol. 1, pp. 627–632, 2019.

[28] M. K. Ahmed, "Multimedia aided language teaching: an ideal pedagogy in the English language teaching of Bangladesh," *American International Journal of Social Science Research*, vol. 3, no. 1, pp. 39–47, 2018.

[29] R. A. Rashid, S. Basree, A. Rahman, and K. Yunus, "Reforms inthe policy of English language teaching in Malaysia," *Policy Futures in Education*, vol. 15, no. 1, pp. 100–112, 2017.

[30] J. S. Zhang, "Evaluation of English teaching quality based onGA optimized RBF neural network," *Computer Systems &Applications*, vol. 29, no. 03, pp. 171–176, 2020.
[31] L. Geng, "Evaluation model of college English multi-mediateaching effect based on deep convolutional neural networks," *Mobile Information Systems*, vol. 8, Article ID 1874584, 2021.
[32] C. Zhen, "Using big data fuzzy K-means clustering and information fusion," *Algorithm in English Teaching Ability Evaluation*, vol. 1, Article ID 5554444, 2021.

Hindawi
Scientific Programming
Volume 2022, Article ID 5203022, 10 pages
https://doi.org/10.1155/2022/5203022

Research Article

Expression Recognition of Classroom Children's Game Video Based on Improved Convolutional Neural Network

Xiaohong Li

Henan Institute of Economics and Trade, Zhengzhou, Henan 450000, China

Correspondence should be addressed to Xiaohong Li; xiaoxiao518@henetc.edu.cn

Received 12 February 2022; Revised 8 March 2022; Accepted 11 March 2022; Published 8 April 2022

Academic Editor: Baiyuan Ding

Humans express emotions in many ways, such as gestures, limbs, and expressions. Among them, facial expressions are the most intuitive way to express human inner emotional activities in human-to-human communication. With the rapid development of computer vision, facial expression recognition is an important research topic in the field of computer vision. It plays a key role in nonverbal communication and can be applied to human-computer interaction, social robotics, video games, and other fields. Traditional expression recognition algorithms require complex manual feature extraction, which takes a long time, and the accuracy of expression recognition in complex scenes is not high. However, with the development of deep learning, especially the convolutional neural network, facial expression recognition technology has also developed rapidly, and the recognition accuracy has been greatly improved. This paper studies the facial expression recognition method of classroom children's game video based on convolutional neural network and proposes a convolutional neural network with deeper layers. The full connection is modified to 4 layers of convolution, 4 layers of pooling, and 2 layers of full connection. Firstly, the facial expression image is preprocessed by, for example, key point location, face cropping, and image normalization; then, the convolutional layer is used to extract the low-dimensional and high-dimensional feature information of the face image; and the pooling layer is used to extract the face image. The feature information is dimensionally reduced. Finally, the softmax classifier is used to classify and recognize the expressions of the training sample images. In order to improve the accuracy of expression recognition, a self-made set of labeled pictures was added to the expression training set. Simulation and comparison experiments show that the improved model has higher accuracy and smoother loss curve, which verifies the effectiveness of the improved network.

1. Introduction

Expression recognition refers to the separation of a specific expression state from a given static image or dynamic video sequence, thereby determining the psychological emotion of the recognized object. With the rapid development of computer technology, artificial intelligence technology, and related disciplines, the automation degree of the whole society continues to increase, and people's demand for human-computer interaction similar to the way people communicate with each other is growing. If computers and robots can understand and express emotions like humans, this will fundamentally change the relationship between humans and computers, enabling computers to serve humans better. Expression recognition is the basis of emotion understanding, the premise for computers to understand people's emotions, and an effective way for people to explore and understand intelligence [1–5].

Facial expression recognition (as shown in Figure 1) refers to separating a specific expression state from a given static image or dynamic video sequence, so as to determine the psychological emotion of the recognized object, realize the understanding and recognition of facial expression by computer, and fundamentally change the relationship between humans and computers, so as to achieve better human-computer interaction. Therefore, facial expression recognition has great potential application value in the field of education. In particular, the evaluation of classroom teaching efficiency can have a great application, but there is currently a lack of an effective expression recognition

method for classroom teaching. In order to solve the above shortcomings, it is urgent to provide a solution [6–9].

Facial expression recognition technology has a wide range of application scenarios and is mainly used in the following real-world scenarios:

(1) In the field of human-computer interaction, the traditional mouse and keyboard and human input commands are abandoned, and expressions, actions, and voices are used to control and operate the machine, so that the machine can understand human emotions and make corresponding responses, thereby saving time and improving machine performance operating efficiency.

(2) In the field of safe driving, the facial expression recognition technology can be used to monitor the driver's facial expression status at all times, so as to determine whether driver fatigue occurs.

(3) In the field of short video, with the launch and continuous development of 5G networks and smartphones, mobile phones have gradually become one of the main ways for people to understand current affairs and for leisure and entertainment. People can watch short videos anytime and anywhere through their mobile phones. The time of each video is only a few dozen seconds. In this short period of time, users can watch what the publisher wants to express or learn about current news and current affairs. However, the current short videos can only be recommended according to the type of videos that viewers usually like. If it can be assisted by facial expression recognition, the camera captures the user's facial features and then analyzes and judges the category of their emotions to recommend short videos that match their current emotions type.

(4) In terms of case detection, when examining suspects, the machine can automatically identify and learn complex psychological changes based on changes in the suspect's facial expressions, so as to figure out the other party's behavioral motives and provide certain help for the police to solve the case [10–16].

Facial expression recognition is a complex learning process, and how to improve the recognition rate is a problem that researchers need to solve. With the rapid development of artificial intelligence, many scholars have devoted themselves to the field of expression recognition, applied some related algorithms such as image processing and feature extraction, and achieved good results. However, how to further improve the recognition rate still needs to be studied today.

In 1862, French researcher Duchenne studied electrical stimulation of the individual facial muscles responsible for the production of facial expressions. Later, in 1872, in Darwin's work "The Expression of the Emotions in Man and Animals," he described the changing process of facial expressions in the process of human-to-human communication, indicating that one of the important ways for humans to express emotions is the facial expressions. A comprehensive and in-depth study of facial expression recognition began with the MIT Media Lab led by Professor Picard, which applied facial emotion recognition technology to the analysis of social behaviors of autistic teenagers. MIT has also developed a robot that can recognize the facial expressions of the other person in the communication with the interlocutor, then analyze the facial expression of the interlocutor, and make different responses according to the results of the analysis. Since then, with the rapid development of computer vision, the research on facial expression recognition by domestic and foreign researchers has progressed rapidly, and different methods for facial expression recognition research have emerged and have achieved remarkable results. For example, the optical flow method proposed by Mase et al. achieves an 80% facial expression recognition rate. With the rise of deep learning, neural networks are favored by researchers due to their high recognition rate. Therefore, the facial expression recognition algorithm based on deep learning has become a research hotspot. For example, Liu et al. proposed an AUDN (AU-inspired Deep Network), which divides facial expressions into different facial expression units, encodes them, and uses deep neural networks for deeper feature extraction, so that the network model can achieve better facial expression recognition effect. Lopes et al. used preprocessing to extract specific features of facial expression. They established a five-layer convolutional neural network to extract facial expression features and input the extracted facial expression features into a classifier for facial expression analysis. The proposed method achieved a facial expression recognition rate of 97.75% on the facial expression dataset CK+. However, the traditional facial expression recognition method requires step-by-step processing. First, the facial image features are manually extracted, and then the corresponding classifiers are selected for classification. This process is relatively complicated. The images collected in real scenes are mainly affected by the illumination angle and posture. The influence of factors such as difference and occlusion greatly reduces the robustness of traditional methods [17–23].

Figure 1: Facial expression recognition.

With the introduction and continuous development of deep learning, deep learning methods have gradually shown good results in computer vision tasks, among which convolutional network and recurrent network algorithms have

been used in feature extraction, image recognition, and classification tasks. CNN is a deep neural network composed of convolutional layers, pooling layers, nonlinear activation functions, and fully connected layers. The local features of the input data itself are used for autonomous learning, the global features in the image, and the data enhancement methods such as translation, scaling, and rotation of the image make it robust. The CNN algorithm does not need to manually extract features but performs end-to-end learning and training by directly inputting the pixel values of the image sequence, reducing the degree of dependence on facial image samples and data preprocessing methods. Therefore, CNN has made breakthroughs in tasks such as image processing, face recognition, automatic detection, and scene analysis [24–30].

2. Convolutional Neural Network

At present, the research method using deep learning is the most effective method in solving the problem of facial expression recognition, and using artificial neural network models of different depths in the convolutional neural network has different effects on the lighting environment, different angles, posture changes, whether there is occlusion, and other factors. Feature extraction comes from facial images. In most computer vision tasks, the artificial neural network of deep learning method can avoid the tedious process of manually extracting facial expression image features in traditional methods and extract facial expression features through autonomous learning, so that the obtained image features have strong discrimination. At the same time, the CNN model has high accuracy and better robustness. In addition, the convolutional network model in the deep learning method integrates each link of the traditional method into an end-to-end network model for learning and training, which effectively reduces the complexity of the target task.

Neural network is an abstract mathematical model proposed and developed on the basis of modern neuroscience, which aims to reflect the structure and function of the human brain. A neural network is composed of one or several neurons; that is to say, a neuron is the basic unit of a neural network, as shown in Figure 2, which is a neuron structure. From the neuron structure diagram, the output h can be obtained, and its formula is as follows:

$$h = f\left(\sum_{i=1}^{n} w_i x_i + b\right), \tag{1}$$

where x_i is the ith input, w_i is the weight of the ith input data, b is the bias value, f is the activation function, and h is the output.

2.1. Convolutional Layer. Convolution calculation is the core operation of convolutional neural network, and it is also a special linear operation. The convolution layer is calculated by multiple convolution kernels to form multiple feature maps. Figure 3 is an example of a convolution calculation,

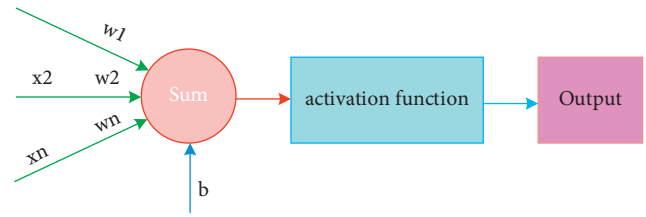


Figure 2: The basic unit of a neural network.

where the size of the input image is $3 * 3$, the red square indicates a size of the convolution kernel of $2 * 2$, the sliding step size is 1, the size of the generated feature map is $2 * 2$, and the convolution calculation process is represented in the dotted box.

Suppose the size of the input image is $W_0 * W_0$, the size of the convolution kernel is $K * K$, the sliding step size is S, the padding is P, and the size of the output feature map is $W_1 * H_1$; then, the calculation formulas of W_1 and H_1 are as follows:

$$W_1 = \frac{(W_0 - K + 2P)}{S} + 1,$$
$$H_1 = \frac{(H_0 - K + 2P)}{S} + 1. \tag{2}$$

The number of channels of the output feature map is equal to the number of convolution kernels. The specific operation of padding is to add 0 to the periphery of the input image and add a layer, and padding is recorded as 1. The role of padding is to prevent the loss of image edge information.

2.2. Pooling Layer. In the convolutional neural network, the function of the pooling layer is to compress the image and reduce the dimension of the feature map, so it is also called the downsampling layer. In the convolutional neural network, usually after a series of convolution operations, a pooling layer is used to halve the width and height of the feature map extracted by the convolutional neural network, and through the compression of the features, the main image of the image is achieved. The purpose of efficient extraction of feature information is to simplify the computational complexity and improve the computational speed of the network. There are many types of pooling, such as max pooling, average pooling, overlap pooling, and spatial pyramid pooling. The most commonly used pooling methods in image recognition and classification tasks are max pooling and average pooling. The maximum pooling refers to taking the largest feature point in the neighborhood, which means that the maximum value of the feature is saved; the average pooling refers to averaging the feature points in the neighborhood, which means that the average value of the feature is saved. The calculation process of pooling is similar to the calculation process of convolution. During pooling, the convolution kernel goes through the feature map from top to bottom and from left to right according to a certain step size, and the corresponding window area will be pooled. The calculation process of pooling is

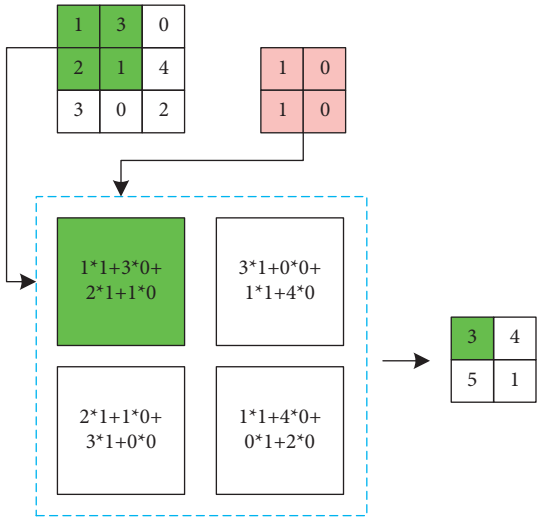


Figure 3: An example of a convolution calculation.

$$y = \text{pooling}(a), \tag{3}$$

where pooling ($*$) represents the pooling function and a is the result after convolution.

2.3. Activation Function. In order to ensure that the convolutional network can better fit the data model, an activation function is usually added after the feature information is extracted by the convolution module for nonlinear calculation. The four activation functions often used in current convolutional network models are the Sigmoid function, the tanh function, the ReLU function, and the Leaky ReLU function.

(a) Sigmoid function: it is also called S-model function. By mapping variables to (0, 1), it has the characteristics of monotonic continuity, limited output range, and easy derivation. The function formula is expressed as follows:

$$\sigma(x) = \frac{1}{1 + e^{-x}}. \tag{4}$$

(b) Tanh function: it is the hyperbolic tangent function. Its shape is similar to the Sigmoid function, but the whole function takes the zero point as the center of symmetry, and the transformation range is (–1, 1), which can solve the mean shift phenomenon of the Sigmoid function. The function expression is shown as follows:

$$\tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}. \tag{5}$$

(c) ReLU function: ReLU function is the most commonly used nonlinear activation function in CNN model, which has the characteristics of simple expression and fast operation speed. The most important thing is to solve the problem of gradient dispersion during model backpropagation. In addition, the ReLU function's feature of setting the negative semiaxis input to zero can make the connections between the network's convolutional layers more sparse during training, but this operation can also lead to neuron death. The function expression is shown as follows:

$$\text{relu}(x) = \begin{cases} x, & \text{if } x > 0, \\ 0, & \text{if } x \le 0. \end{cases} \tag{6}$$

(d) Leaky ReLU function: this function is the same as the ReLU function on the positive half axis and performs a simple weighting operation on the negative half axis. It resolved the issue of neuron death caused by the ReLU function in training. Its function expression is shown in (7), where a is a very small fixed value. In this case, the function becomes PReLU.

$$\text{leaky relu}(x) = \begin{cases} x_i, & \text{if } x > 0, \\ ax_i, & \text{if } x \le 0. \end{cases} \tag{7}$$

2.4. Dropout Layer. Convolutional neural networks are prone to overfitting problems. The so-called overfitting means that the training set shows a high accuracy rate, but the accuracy rate is poor in the test set. The proposal of dropout technology has effectively improved this problem. The idea is to randomly discard some neurons with a certain probability during the training process to reduce the network's dependence on certain neurons. The network structure after each dropout is different. The results can be regarded as the average of multiple models, which can improve the generalization ability of the network.

Suppose the output of the ith hidden unit in the $l+1$th layer of the standard neural network is y_i, y represents the output of the lth layer, w represents the weight, b represents the bias, and $\sigma()$ represents the activation function; then, the calculation formula of y_i is as follows:

$$y_i = \sigma(w_i^{*} y + b_i). \tag{8}$$

If dropout is added, the calculation formula of y_i is as follows:

$$\begin{aligned} r &\sim \text{Bernoulli}(p), \\ \tilde{y} &= r^{*} y, \\ y_i &= \sigma(w_i^{*} \tilde{y} + b_i). \end{aligned} \tag{9}$$

In the formula, r is the generated probability vector, which is multiplied by y to get a reduced version of the output vector y, and y is applied to each hidden unit of the next layer to obtain y_i, which is equivalent to sampling a large network.

3. Improving Convolutional Neural Networks

The solution of expression recognition mainly consists of 4 steps: face detection, data preprocessing, feature extraction, and expresssion classification. The most used method at

present is facial expression recognition through deep learning, which combines the two steps of feature extraction and classification to achieve an end-to-end training mode. However, the traditional method is to extract various features in the image and then select the corresponding classifier for identification.

For expression recognition tasks, the selection of expression data samples is equally important. Therefore, this paper will describe in detail different expression datasets and related theoretical and practical methods used in expression recognition tasks. The facial expression recognition process based on deep learning is shown in Figure 4.

In the study of the FER problem, many databases have been subjected to comparative experiments by many researchers. Traditional methods use two-dimensional still images or two-dimensional dynamic image sequences for the research of expression recognition. In recent years, the spontaneous facial expression task has formed a new research hotspot in the process of research development. The application of 3D face image and expression analysis is of great help to the understanding of the internal subtle structural changes of spontaneous expressions. This section will briefly introduce relevant datasets related to the research of expression recognition problem, among which there are various common 2D and 3D dynamic image sequences as well as still images.

3.1. CK+ Facial Expression Database. The CK+ (Extended Cohn–Kanade) facial expression database contains 593 image sequences. The last frame of each image sequence contains action unit (AU) markers. Among all image sequences, 327 have expression labels, including spontaneous expressions and posed expressions. The 123 participants were between the ages of 18 and 30, and most were women. The resolutions of the images are the precision of the grayscale values, which is 8 bits.

3.2. JAFFE Expression Database. There are 213 gray images in the facial expression database, including seven facial expressions, six of which are basic facial expressions and one neutral facial expression. The original unit of each image is described as pixel.

3.3. FER2013 Expression Database. The FER2013 expression database is a face recognition contest database provided by the Kaggle website in 2013. There are 35,887 grayscale images in the database, including a total of seven facial expressions. Each expression corresponds to a numerical label, where 0 = anger, 1 = disgust, 2 = fear, 3 = happy, 4 = sad, 5 = surprised, and 6 = neutral. The original size of each image is that all images were downloaded from the Internet.

3.4. MMI Expression Database. The MMI Facial Expression Database consists of over 2900 video sequences and high-resolution still images of 75 participants. It fully annotates the presence of AU in video sequences and partially encodes it at the frame level, indicating whether each frame is in the neutral, onset, vertex, or offset phase. There were 75 participants, both male and female, ranging in age from 19 to 62. The original size of each face image is pixels.

3.5. BP4D Spontaneous Expression Database. BP4D Spontaneous (Binghamton-Pittsburgh 4D Dynamic Spontaneous) Expression Database is a 3D video database including spontaneous expressions of 41 young adults (23 females, 18 males). Participants were 18–29 years old, 11 were Asian, 6 were African American, 4 were Hispanic, and 20 were European American. This database facilitates the exploration of 3D spatiotemporal features in fine facial expressions, leading to a better understanding of the relationship between posture and motion dynamics in facial AU, and a deeper understanding of naturally occurring facial movements. The original size of each face image is pixels.

3.6. B+ Expression Database. The B+ (Extended Yale B face expression) database consists of 16,128 facial images of 28 subjects. The subjects are photographed from 9 different poses as needed, and 64 shooting parameters are used to shoot under a single set of light sources. The original size of each face image is pixels.

3.7. KDEF Expression Database. KDEF (The Karolinska Directed Emotional Face) expression database consists of 4,900 facial expression images of 70 subjects, each photographed from five different angles, and each angle takes multiple facial expressions, including seven different expressions. The original size of each face image is pixels.

Convolutional layer, downsampling layer, fully connected layer, and output layer are the general structure of CNN. It mainly uses the two basic ideas of local perception and weight sharing. This paper uses the CNN after transforming the number of layers to do the task of expression classification. We improve the convolutional neural network LeNet, from the original 2-layer convolution, 2-layer pooling, and 1-layer full connection to 4-layer convolution (C1, C2, C3, C4), 4-layer pooling (S1, S2), S3, S4), and 2-layer full connection. After the convolutional layer, a Rectified Linear Unit (ReLU) activation function is added, and Batch Normalization (BN) is added before the activation function for normalization to prevent the disappearance of gradients. Finally, dropout technology is used to prevent overfitting.

In the convolution layer, features are extracted by using convolution kernels, and the number of convolution kernels is the same as the feature map. Generally, the number of convolution kernels increases with the depth of the convolutional neural network, in order to better extract the high-level features of the input image. The corresponding operation is shown in Figure 5(a). The dark area in the left image of Figure 5(a) indicates that the convolution kernel acts on the image pixel. The convolution kernel is multiplied by the corresponding image pixel and then added to obtain the value of the dark area in the right image; the rest of the

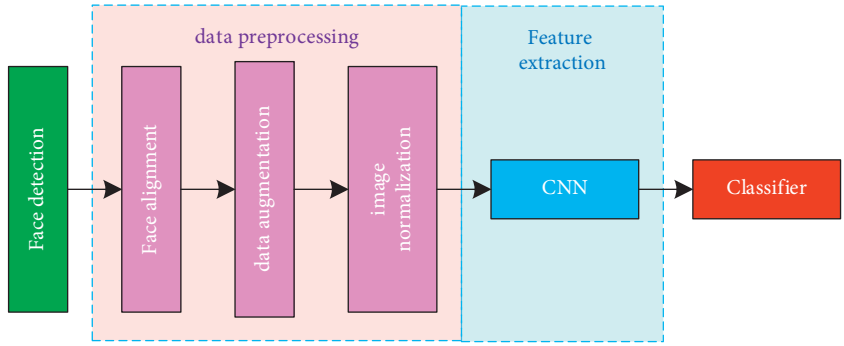


Figure 4: The facial expression recognition process based on deep learning.

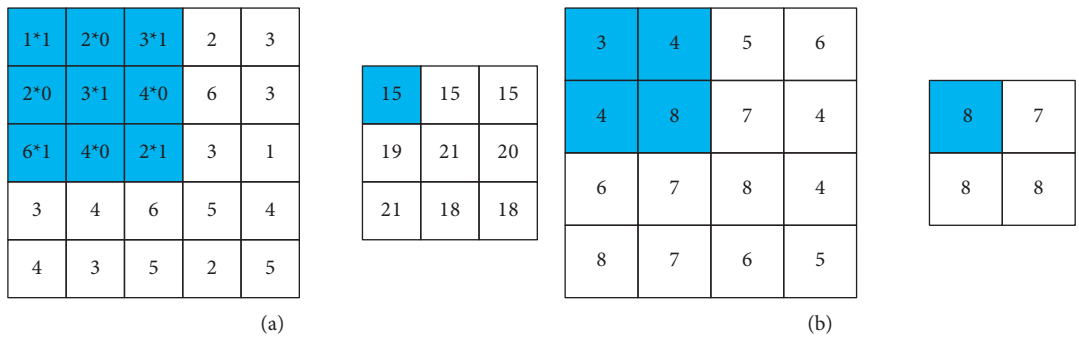


Figure 5: Operation. (a) Convolution operation. (b) Max pooling operation.

values in the right image of Figure 5(a) are calculated by moving the position of the convolution kernel in the image.

In the convolutional neural network, sampling operations are divided into two types, namely, upsampling and downsampling. There are two commonly used downsampling methods: maximum downsampling and average downsampling. Maximum downsampling is also called max pooling, and average downsampling is also called average pooling. In order to avoid the problem that the feature dimension of the feature map extracted by the convolutional layer is too high, a pooling layer is often connected after the convolutional layer for dimensionality reduction. In this paper, the maximum pooling layer is used for feature dimensionality reduction. If the input image is large, it is also possible to connect continuous 2-layer pooling and perform 2 dimensionality reduction operations. The feature training classifier learned by this method will not have the problem of excessive dimensionality. At the same time, the downsampling operation can reduce the sensitivity of the feature map output to rotation, scaling, translation, etc. The size of the feature map after downsampling becomes the original $2n/s$, where n is the size of the sampling window. This paper uses max pooling, and its operation is shown in Figure 5(b). The dark area in the left picture shows that the sampling window acts on the image pixels, and the maximum value of the image pixel in the sampling window is taken out as the final sampling result, which corresponds to the value of the dark area in the right picture of Figure 5(b). The rest of the values on the right are calculated by shifting the position of the sampling window in the image.

The fully connected layer is connected before the output layer of the CNN, and there is generally 1 or 2 layers at the back of the CNN. Its connection method is special; each neuron in the fully connected layer must be connected with all the neurons in the previous layer to integrate the local information in the convolutional layer and the pooling layer. A ReLU activation function is added after each neuron in the fully connected layer. The input of the fully connected layer must be an array, and it must be one-dimensional. Therefore, the two-dimensional array output by the pooling layer S4 is converted into a one-dimensional array, and then all the converted one-dimensional arrays are connected, and finally becomes 1 A 4 608-dimensional ($3 \times 3 \times 512 = 4\,608$) feature vector, which is used as the input of the fully connected layer. The fitting ability and training speed of the network are closely related to the number of neurons in the fully connected layer, so it is necessary to select a suitable number of neurons. Through the experimental test, the network learning effect is better when the number is 800.

4. Simulation Experiments

The experimental environment of the algorithm in this paper is as follows: Ubuntu 16.04 system, Intel Core i5-7200U CPU, and 8 GB of memory. Because most of the images in the public dataset have complex backgrounds and different light intensities, the static images in the public dataset CK+ are selected for the experiments in this paper, and some self-collected images are added to expand the dataset. We select 7 kinds of expressions, with a total of 648 face image samples,

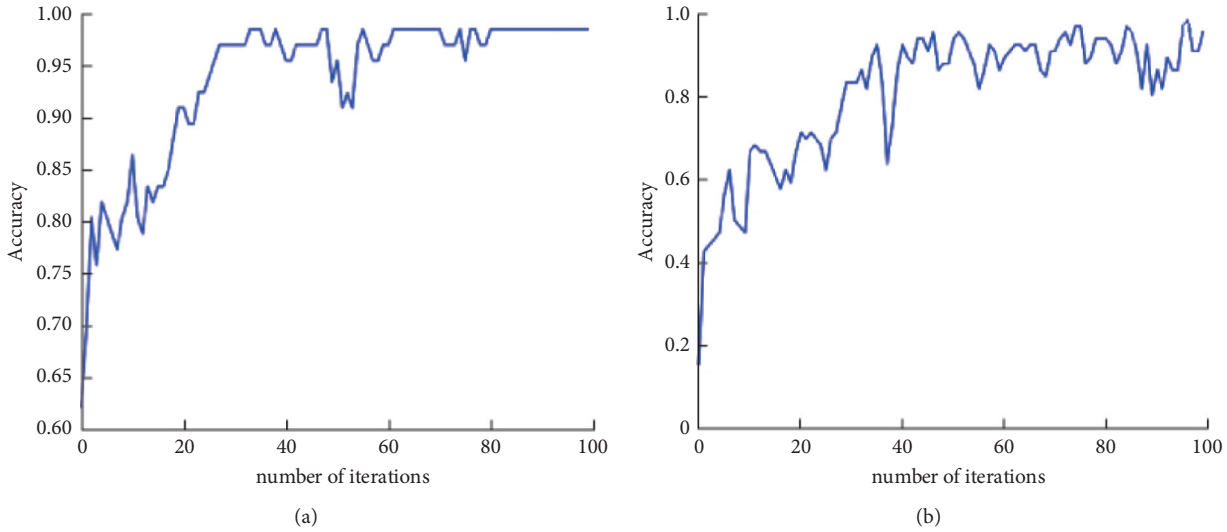


Figure 6: Accuracy. (a) Proposed method. (b) LeNet.

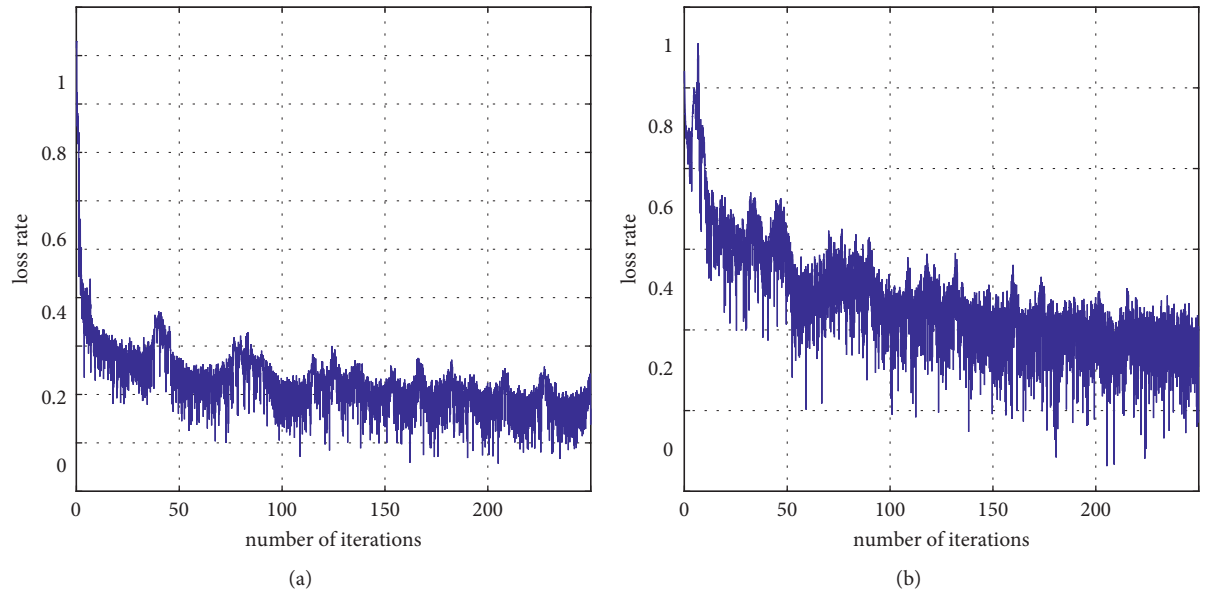


Figure 7: Loss rate. (a) Proposed method. (b) LeNet.

of which there are 90, 50, 138, 36, 50, 166, and 118 samples of angry, scared, happy, neutral, sad, surprised, and disgusted expressions, respectively. These 648 samples are all used to train the training set of the model. The test set used in the experiment includes some remaining static images and self-collected images of the CK+ dataset, including 128 samples of 7 kinds of expression labels.

We convert the image format of the training set to CSV format, then input it into the improved convolutional neural network in this paper for training, and also input it into the LeNet network before the improvement for training. The network accuracy image in this paper is shown in Figure 6(a), the LeNet accuracy image is shown in Figure 6(b), the image of the network loss function in this paper is shown in Figure 7(a), and the image of the LeNet loss function is shown in Figure 7(b).

From the accuracy curve, it can be seen that the accuracy curve of the improved network in this paper is faster and smoother than the accuracy curve of the LeNet network. According to the loss function curve, it can also be seen that the loss function of the improved network and the loss of the LeNet network are faster and smoother. By evaluating the performance of the model based on these two indicators, it can be seen that the improved network is more robust than the LeNet network.

In order to further illustrate the performance advantages of the improved network in this paper compared with the original network LeNet, the test set is input into the improved network and the LeNet network, and the confusion matrix is used as the performance evaluation index. The confusion matrices of the two models are shown in Figure 8. It can be seen that the average recognition rate of the seven

Figure 8: Confusion matrix on test set. (a) Proposed method. (b) LeNet.

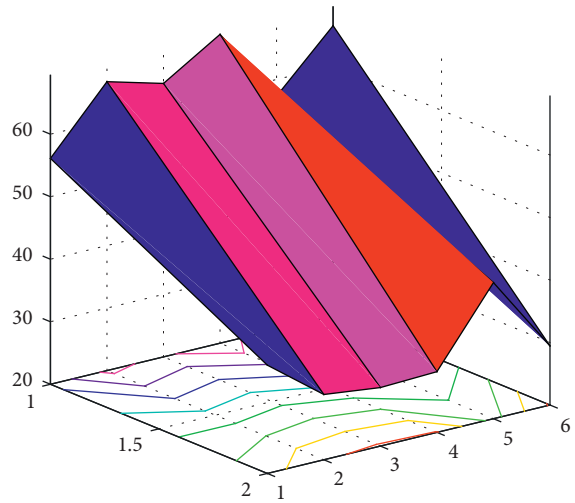


Figure 9: Prediction.

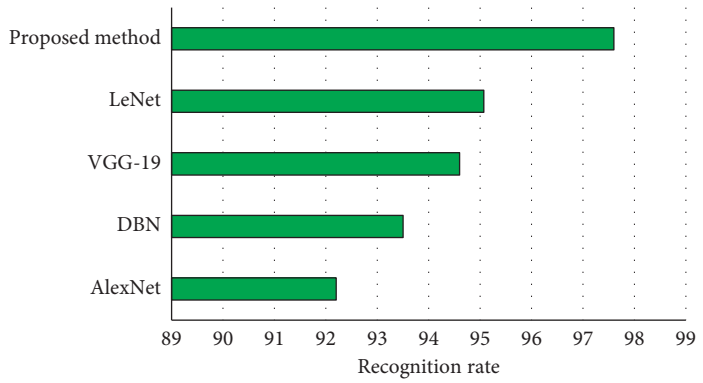


Figure 10: Recognition rate.

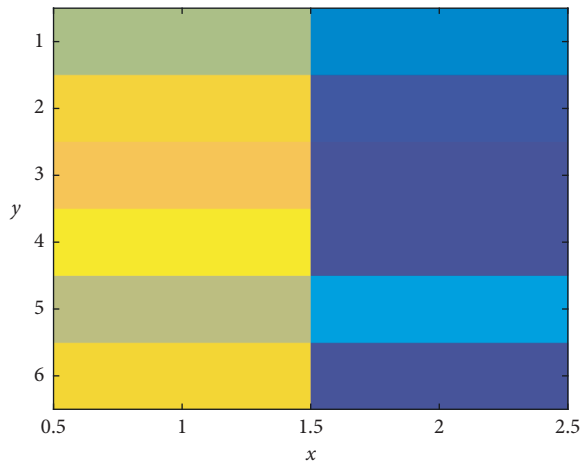


Figure 11: x and y variation.

expressions on the test set of the improved network is 93.7%, and the average recognition rate of the seven expressions of LeNet on the test set is 90.5%, indicating that the recognition accuracy of the improved network has been improved. It is also proved that the robustness of the improved network is better compared with the original network. The prediction is shown in Figure 9.

In order to further verify the recognition rate of the CNN model proposed in this paper, under the same hardware environment, the same dataset was input into the following four networks: DBN, AlexNet, VGG-19, and LeNet for experiments, and they were obtained in training. It can be seen from the results in Figure 10 that the average recognition rate of the improved CNN in this paper is increased by 5.4% compared with the classic AlexNet algorithm. Compared with the LeNet algorithm before the improvement, the recognition rate is also increased by 2.53%. The x and y variation are shown in Figure 11.

5. Conclusion

Humans express emotions in many ways, such as gestures, limbs, and expressions. Among them, facial expressions are the most intuitive way to express human inner emotional activities in human-to-human communication. With the rapid development of computer vision, facial expression recognition is an important research topic in the field of computer vision. It plays a key role in nonverbal communication and can be applied to human-computer interaction, social robotics, video games, and other fields. Traditional expression recognition algorithms require complex manual feature extraction, which takes a long time, and the accuracy of expression recognition in complex scenes is not high. However, with the development of deep learning, especially the convolutional neural network, facial expression recognition technology has also developed rapidly, and the recognition accuracy has been greatly improved. This paper studies the facial expression recognition method of classroom children's game video based on convolutional neural network and proposes a convolutional neural network with deeper layers. The full connection is modified to 4 layers of convolution, 4 layers of pooling, and 2 layers of full connection. Firstly, the facial expression image is preprocessed by, for example, key point location, face cropping, and image normalization; then, the convolutional layer is used to extract the low-dimensional and high-dimensional feature information of the face image; and the pooling layer is used to extract the face image. The feature information is dimensionally reduced. Finally, the softmax classifier is used to classify and recognize the expressions of the training sample images. In order to improve the accuracy of expression recognition, a self-made set of labeled pictures was added to the expression training set. Simulation and comparison experiments show that the improved model has higher accuracy and smoother loss curve, which verifies the effectiveness of the improved network.

Although the recognition effect has been greatly improved, the training speed of the improved network is lower than the training speed of the original classic network. Therefore, how to optimize the network to improve the recognition rate and also ensure the speed of network training is also the future problem to be improved.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

The author thanks 2022 Henan Province Key Research and Development and Promotion Special Project (Science and Technology Research, Soft Science Research Support Project) and Research on the Cultivation Path of "Artisan Spirit" for Students in Henan Vocational Colleges under the Background of Quality Improvement and Excellence (no. 149).

References

[1] H. Hamada, S. Miki, and R. Nakatsu, "Automatic evaluation of English pronunciation based on speech recognition techniques," *IEICE—Transactions on Info and Systems*, vol. E76-D, no. 3, pp. 352–359, 1993.

[2] K. Truong, *Automatic Pronunciation Error Detection in Dutch as a Second Language: An Acoustic-Phonetic Approach*, Utrecht University, Utrecht, Netherlands, 2014.

[3] B. Dong, Q. Zhao, and Y. Yan, "Automatic scoring of flat tongue and raised tongue in computer-assisted Mandarin learning," in *Proceedings of the Proceedings of ISCSLP*, Tianjin, China, October 2016.

[4] S. M. Witt and J. Young, "Phone-level pronunciation scoring and assessment for interactive language learning," *Speech Communication*, vol. 30, no. 4, pp. 5–108, 2000.

[5] H. Chao, Z. Feng, F. K. Soong, M. Chu, and R. Wang, "Automatic mispronunciation detection for Mandarin," in *Proceedings of the IEEE International Conference on Acoustics*, Las Vegas, NV, USA, April 2008.

[6] Y. B. Wang and L. S. Lee, "Improved approaches of modeling and detecting error patterns with empirical analysis for computer-aided pronunciation training," in *Proceedings of the Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, Kyoto, Japan, March 2012.

[7] A. Neri, C. Cucchiarini, and H. Strik, "ASR-based corrective feedback on pronunciation: does it really work?" in *Proceedings of the International Conference on Interspeechicslp*, San Francisco, CA, USA, September 2016.

[8] Y. Ishida and S. Hashimoto, "Asymmetric characterization of diversity in symmetric stable marriage problems: an example of agent evacuation," *Procedia Computer Science*, vol. 60, no. 1, pp. 1472–1481, 2015.

[9] P. Zoha and R. Kaushik, "Image edge detection based on swarm intelligence using memristive networks," *IEEE Trans. on CAD of Integrated Circuits and Systems*, vol. 37, no. 9, pp. 1774–1787, 2018.

[10] W. Li, S. M. Siniscalchi, N. F. Chen, and C.-H. Lee, "Improving non-native mispronunciation detection and enriching diagnostic feedback with DNN-based speech attribute modeling," in *Proceedings of the 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Shanghai, China, March 2016.

[11] X. Qian, H. Meng, and F. Soong, "The use of DBNHMMs for mispronunciation detection and diagnosis in L2 English to support computer-aided pronunciation training," in *Proceedings of the proc interspeech*, Brno, Czechia, September 2021.

[12] K. Li, X. Qian, and H. Meng, "Mispronunciation detection and diagnosis in L2 English speech using multidistribution deep neural networks," *IEEE ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, pp. 193–207, 2016.

[13] A. Lee, Y. Zhang, and J. Glass, "Mispronunciation detection via dynamic time warping on deep belief network-based posteriorgrams," in *Proceedings of the IEEE International Conference on Acoustics*, Vancouver, BC, Canada, May 2013.

[14] Y. Hua, J. Zhao, and L. Jia, "Improve mispronunciation detection with Tandem feature," in *Proceedings of the International Symposium on Chinese Spoken Language Processing*, Hongkong, China, January 2020.

[15] J. Pais, "Random matching in the college admissions problem," *Economic Theory*, vol. 35, no. 1, pp. 99–116, 2018.

[16] J. J. Jung and G. S. Jo, "Brokerage between buyer and seller agents using constraint satisfaction problem models," *Decision Support Systems*, vol. 28, no. 4, pp. 291–384, 2020.

[17] Y. Liu and K. W. Li, "A two-sided matching decision method for supply and demand of technological knowledge," *Journal of Knowledge Management*, vol. 21, no. 3, p. 0183, 2017.

[18] J. Byun and S. Jang, "Effective destination advertising: matching effect between advertising language and destination type," *Tourism Management*, vol. 50, no. 10, pp. 31–40, 2015.

[19] A. N. Nagamani, S. N. Anuktha, N. Nanditha, and V. K. Agrawal, "A genetic algorithm-based heuristic method for test set generation in reversible circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 37, no. 2, pp. 324–336, 2018.

[20] C. Koch and S. P. Penczynski, "The winner's curse: conditional reasoning and belief formation," *Journal of Economic Theory*, vol. 174, pp. 57–102, 2018.

[21] C. K. Karl, "Investigating the winner's curse based on decision making in an auction environment," *Simulation & Gaming*, vol. 47, no. 3, pp. 324–345, 2016.

[22] D. Ettinger and F. Michelucci, "Creating a winner's curse via jump bids," *Review of Economic Design*, vol. 20, no. 3, pp. 173–186, 2016.

[23] J. A. Brander and E. J. Egan, "The winner's curse in acquisitions of privately-held firms," *The Quarterly Review of Economics and Finance*, vol. 65, pp. 249–262, 2017.

[24] Z. Palmowski, "A note on var for the winner's curse," *Economics/Ekonomia*, vol. 15, no. 3, pp. 124–134, 2017.

[25] B. R. Routledge and S. E. Zin, "Model uncertainty and liquidity," *Review of Economic Dynamics*, vol. 12, no. 4, pp. 543–566, 2009.

[26] D. Easley and M. O'Hara, "Ambiguity and nonparticipation: the role of regulation," *Review of Financial Studies*, vol. 22, no. 5, pp. 1817–1843, 2019.

[27] P. Klibano, M. Marinacci, and S. Mukerji, "A smooth model of decision making under ambiguity," *Econometrica*, vol. 73, no. 6, pp. 1849–1892, 2005.

[28] Y. Halevy, "Ellsberg revisited: an experimental study," *Econometrica*, vol. 75, no. 2, pp. 503–536, 2017.

[29] D. Ahn, S. Choi, and D. Gale, "Estimating ambiguity aversion in a portfolio choice experiment," *Working paper*, vol. 5, no. 2, pp. 195–223, 2019.

[30] T. Hayashi and R. Wada, "Choice with imprecise information: an experimental approach," *Theory and Decision*, vol. 69, no. 3, pp. 355–373, 2010.

Hindawi
Scientific Programming
Volume 2022, Article ID 2306177, 10 pages
https://doi.org/10.1155/2022/2306177

Research Article

BP Neural Network-Based Big Data Intelligent Travel Algorithm and Its Application

Yan Huang

Chongqing Chemical Industry Vocational College, Chongqing 401220, China

Correspondence should be addressed to Yan Huang; tsxy20160127@cqcivc.edu.cn

Received 31 December 2021; Accepted 17 February 2022; Published 6 April 2022

Academic Editor: Baiyuan Ding

With the increase of urbanization rate, a large number of people flood into cities, increasing pressure of urban traffic, and problems accumulated in the taxi industry are gradually prominent. The phenomenon of crowded queue for taxi is frequent in peak hours, and vehicles patrol and sweep streets during peak hours. The key to solve these problems lies in mastering the rules and patterns of taxi travel and finding the factors affecting the relationship between taxi supply and demand. It is difficult to effectively understand taxi travel as a whole due to the large number of taxis and their large scale and strong mobility. Comprehensive application of trajectory data mining method can extract the spatiotemporal characteristics of massive taxi trajectory data and reveal the nature of its occurrence. This paper mainly focuses on the problems faced by urban traffic governance, such as the mismatch between data sources and demand systems, the uncoordinated operation of comprehensive transportation system, and the difficulty in sharing big data resources between government and enterprises. With massive data resources across departments, this paper designed a networked intelligent computing platform for big data of urban transportation integrated with various modes of transportation to sense the operation situation of urban comprehensive transportation system in real time, accurately grasp the space-time distribution of urban transportation supply and demand, and significantly improve the ability of coordinated operation, organization, and management of transportation in large cities. It also effectively enhances the quality of transportation information sharing and integration services and comprehensively improves the efficiency and overall carrying capacity of the urban comprehensive transportation system.

1. Introduction

Modern urban comprehensive transportation is an important carrier to support the normal operation of a city and guarantee residents' life [1]. Data integration mechanism for urban traffic management in our country and innovative applications platform as a whole are still in the primary stage of data integration, facing the construction unit partition, system, information island, fragmentation data management, artificial intelligence, and low level of big data applications challenge that cannot meet the urban traffic system management efficiently and orderly, wisdom resource scientific configuration. The need for data openness and integration. Based on time and space distribution of taxi data research, most can reflect the transportation personnel in the city of dynamic distribution of geographical space for the dynamic space-time distribution directly reflects the population, employment, travel, road and living space and a series of with the spatial layout of land use governance is directly related to spatial planning such as influencing factors [2, 3].

Through the acquisition, analysis, and fusion of urban traffic big data, as well as the analysis and prediction of operation state, the effective monitoring and management of urban traffic can be realized, the capacity and service level of the whole road network can be improved, and a refined data set for modern urban traffic governance can be formed. At present, the city space layout research mostly focus on the seeking of the static relationship between urban spatial elements, such as traffic network population and land use, ignoring the study of urban spatial dynamic indexes such as active degree, and dynamic index such as

urban space activity can intuitively reflect the aggregation degree of city space quickly [4]. It is of great significance to analyze and guide the distribution of urban economic and social land and transportation, especially for the layout and optimization of urban commercial land. In the past, there have been abundant researches on indicators of urban spatial activity degree, and the research on acquisition methods of urban Hot Spots and other similar indicators is also relatively mature. Based on previous studies, it can be seen that traffic travel is an important cause of urban space activity. Therefore, this paper assumes that the active area of traffic can reflect the active degree of urban space in this region, that is, the active area of urban space. Through big data analysis of smart travel, active urban space is found, and its spatial correlation with urban commercial land is verified through correlation analysis, so as to guide the evaluation of urban commercial land [5–8].

Since the twenty-first century, with the method of cloud platform, based on the cloud computing architectures high compatibility, set up the urban traffic data intelligent computing cloud platform, through the integration of data mining, deep learning, edge, heterogeneous computing technology, such as building for refinement, intelligent, real-time traffic monitoring scheduling, and operation management of urban intelligent traffic management paradigm [9]. Traffic congestion governance is a governance problem faced by all cities in the world. Comprehensive, objective, accurate, and timely grasp of city-level large-scale traffic operation rules is of vital significance to traffic congestion governance. Used in traffic data-driven governance mainly face the following four aspects of the problem: one is the cognitive modeling ability is insufficient, the existing traffic data collection facilities generally single function isolated deployment, multisource heterogeneous data in time and space fragmentation distribution, and the lack of data correlation analysis. The second is that the reasoning ability is insufficient, and the traffic situation analysis focuses on the traffic flow. It is urgent to systematically sort out and apply the knowledge and experience of traffic governance [10]. Third, the capacity of large-scale computing is insufficient, and the elastic expansion capacity and large-scale instantaneous computing capability of the existing centralized computing platform of urban traffic are facing challenges. Fourth, the practical ability of governance is insufficient. Cities at different stages of transport development face different governance problems and scenarios and lack professional and effective platform and tool support, so as to enable multiparty cooperation to comprehensively improve the operation efficiency and service level of transport system [12, 13].

In terms of the intelligent governance of urban traffic, major cities in China have initially established various traffic big data collection systems. Developing the evaluation system. But how to effectively integrate the various types of data, forming a powerful traffic computing platform, to control traffic real-time running state to state for a long time, is the key to the city, city transportation management at home and abroad in the video image structured processing multisource data fusion traffic data modeling chart database system state nowcasting cloud computing traffic management carried out in such aspects as the related research and application of exploration. But scene data sensing technology based on the traffic control is not yet mature, cross-media multisource heterogeneous traffic data of data fusion and knowledge mining technology bottleneck, urban transport complex adaptive system both short-term and long-term mechanisms of cognitive are not clear, there is no intelligent computing platform for urban complex transportation system governance. The data accumulated in the process of traffic operation have the characteristics of multitypes, multisources, and heterogeneity [14]. The data provided by different information sources are all in their own reference frame, resulting in the disunity of the spatiotemporal datum and scale of multisource and heterogeneous traffic big data and the incomplete image view in the spatial coverage area of traffic big data. Based on the algorithm of feature representation feature fusion and multistage segment fusion of multisource heterogeneous traffic big data, the multidimensional fusion of traffic big data from different sources and granularity is carried out by combining the self-encoder transfer learning, multitask learning, and multiview learning [15–17]. Based on the results of multisource heterogeneous fusion representation and analysis of traffic big data, heterogeneous fusion analysis applications such as individual behavior trajectory prediction, local traffic flow real-time estimation, overall traffic trend analysis, and OD time estimation are carried out from three aspects of coverage time granularity accuracy [18–25].

2. Related Works

In recent years, many experts and scholars at home and abroad have mined the spatial-temporal characteristics of taxi track data from different disciplinary perspectives and proposed many research methods, including descriptive statistics, density analysis, and cluster analysis. These methods can intuitively describe the overall distribution characteristics of data sets and are not limited by research assumptions and analysis models. They can be used as an independent method to mine trajectory data and can also be used for data preprocessing to provide basic support for subsequent research. Nesmachnow et al. [26] divided the study area of Lisbon into several grids, counted and visualized the trips in the grids, respectively, and represented the trips by the depth of color. Wang et al. [27] calculated the number of taxi arrivals within a certain range of service facilities such as restaurants and shopping based on taxi trajectory data, quantified the attractiveness level of various service facilities, and analyzed the spatial distribution pattern of the attractiveness through global and local spatial autocorrelation. Xiong et al. [28] extracted the pick-up and drop-off points from the taxi track data of Wuhan in a week and used the bar chart to show the daily taxi travel volume and the line chart to show the taxi travel volume in different periods of time. They counted the number of pick-up and drop-off points of Beijing in a month and calculated the working days and rest days. The average number of pick-up

and drop-off points is visualized in the form of broken line graph to reveal the time series pattern of taxi trips in Beijing. Zheng et al. [29] studied the visualization of a large number of floating car data from the perspectives of global view and local view. The global view displays the distribution of regions selected by users in the form of focal graph, and the local view has 2D and 3D traffic parameter visualization symbols and numerous forms of charts.

Hou et al. [30] defined the external transport hub (airport railway station bus and passenger station) of Beijing as the research area and used the kernel density estimation method to compare and analyze the up-and-down passenger cluster area arriving at the transport hub and the disembarking passenger cluster area leaving the transport hub. Kibria et al. [31] taking Nanjing city as the research area, the bandwidth of kernel density estimation was determined by incremental spatial autocorrelation, and then the kernel density was used to detect the hot spots of taxi picking up and getting off, and the temporal and spatial law of taxi travel was analyzed. In order to effectively mine taxi trajectory data, Zhang [32] proposed a two-layer framework to automatically identify each taxi journey destination and estimate the return journey mode and destination using geographic points of interest data and taxi trajectory data, combined with three methods of spatiotemporal clustering Bayesian inference and Monte Carlo simulation. Massive information rich taxi trajectory data, using descriptive statistics method can understand the distribution of total body data, statistical travel frequency, travel distance trip length variables, such as, in turn, through a variety of straight view image visualization method to show the taxi travel mode, but has a certain generality, can accurately express travel fine feature kernel density analysis can extract regions with high point density, and the neighborhood size can affect the estimation results to a certain extent, which may lead to ring phenomenon. Chang [33] collected 222 traffic trips, with a total of nearly 580 000 GPS records, and each record includes speed, acceleration, and driving direction change characteristics. On this basis, the multilayer perceptron neural network, Bayesian network, and decision tree models are constructed, respectively, by using 75 quantiles of velocity, mean signal quality of deviation vector acceleration of velocity as the input of the model. The results show that the multilayer perceptron neural network and decision tree have good recognition accuracy.

Moreno et al. [34] used SVM model to identify travel modes, and the results showed that SVM had good application effect on travel mode recognition, but it did not optimize SVM parameters. As a result, although SVM model was adopted, the final identification accuracy was not high due to the nonoptimal combination of SVM parameters. Fonzone et al. [35] proposed a traffic mode identification method based on AGPS mobile phone. GPS data are collected by mobile phone software, such as instantaneous speed and acceleration, as the characteristics of traffic pattern recognition. The BP neural network (back propagation neural network, referred to as BP neural network) is used for traffic pattern recognition. Mohammadi and Al-Fuqaha [36] proposed transportation mode selection model based on neural network is established, to choose the gender, age, income, occupation, purpose of travel, the place of departure and arrival location, departure time, arrival time of the nine variables as the input of the model, the use of Xuzhou city, Jiangsu province transport in the large-scale urban population trip survey sampling investigation as an example, the measured data Identify walking, cycling and bus travel patterns. But the study identified a lack of taxis, subways and other modes of transportation, which are an important part of transportation.

From the above analysis, we know that the above methods have studied the big data intelligent travel to some extent. However, some problem still exists. For example, no scholar has applied the BP neural networks to this field till now, so the research here is still a blank, which has great theoretical research and practical application value for optimal management of financial assets.

This paper consists of five parts. Section 1 and Section 2 give the research status and background. Section 3 is BP neural network-based big data intelligent travel algorithm. Section 4 shows the experimental results and analysis. The experimental results of this paper are introduced and compared and analyzed with relevant comparison algorithms followed. Finally, Section 5 concludes the full paper.

3. BP Neural Network-Based Big Data Intelligent Travel

3.1. BP Neural Network. Backpropagation neural network (BP neural network) is a supervised learning algorithm, which is often used to train multilayer perceptron because BP neural network successfully solves the problem. In order to solve the weight adjustment problem of multilayer feedforward neural network of nonlinear continuous function, nearly 90% of neural network models in the real application of artificial neural network are BP neural network and its variation form. BP network has a structure of three or more layers, namely input layer, one or more layers of hide layer, and output layer. The neurons between each layer are fully connected, and all neurons in the layer are not connected. The nodes of each hidden layer are generally used. The brief structure of Sigmoid excitation function BP neural network is shown in Figure 1, where the threshold value is not drawn. The number of the nodes of the hidden layer is fix in this paper, and also the number of the output layer.

There are n neurons in the input layer. The input vector is

$$X = (x_1, x_2, \cdots, x_n)^T \tag{1}$$

The hidden layer has L spirit meridian elements, the output layer has M neurons, and the output vector is

$$Y = (y_1, y_2, \cdots, y_m)^T \tag{2}$$

Before the prediction of BP neural network, network training should be carried out first. By network training, the weight threshold can be adjusted to make the network have associative memory and prediction ability. The training process of BP network includes the following steps:

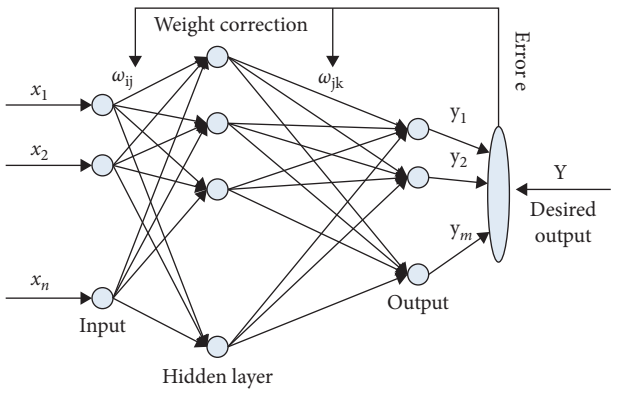


Figure 1: Flow chart of PSO algorithm.

The hidden layer output is calculated according to the input vector X, the connection weights between the input layer and the hidden layer I and J, and the hidden layer threshold, compute the hidden layer output H:

$$H_j = f\left(\sum_{i=1}^{n} \omega_{ij} x_i - a_j\right), j = 1, 2, \cdots, l. \tag{3}$$

Here, L is the number of nodes in the hidden layer and F is the excitation function of the hidden layer:

$$s.t. \sum_{i=1}^{n} w_i u_i = r, \sum_{i=1}^{n} w_i = 1. \tag{4}$$

The output calculation of the output layer is based on the output H of the hidden layer, and the connection weight between the hidden layer and the output layer is J k and the threshold of each neuron of the output layer:

$$\begin{aligned} &\text{Max} \sum_{i=1}^{n} w_i u_i \\ &O_k = \sum_{j=1}^{l} H_j \omega_{jk} - b_k, \quad k = 1, 2, \ldots, m. \end{aligned} \tag{5}$$

The network prediction error is calculated according to the network predicted output O and expected output Y:

$$e_k = Y_k - O_k, k = 1, 2, \cdots, m. \tag{6}$$

Network connection weights are updated according to network prediction error:

$$\begin{aligned} &\omega_{ij} = \omega_{ij} + \eta H_j\left(1 - H_j\right) x_i \sum_{k=1}^{m} \omega_{jk} e_k, \quad i = 1, 2, \ldots, n; j = 1, 2, \ldots, l, \\ &\omega_{jk} = \omega_{jk} + \eta H_j e_k, \quad j = 1, 2, \ldots, l; k = 1, 2, \ldots, m. \end{aligned} \tag{7}$$

Update the network node threshold according to network prediction error:

$$\begin{aligned} &a_j = a_j + \eta H_j\left(1 - H_j\right) \sum_{k=1}^{m} \omega_{jk} e_k, \quad j = 1, 2, \ldots, l, \\ &b_k = b_k + e_k, \quad k = 1, 2, \ldots, m. \end{aligned} \tag{8}$$

Determine whether the algorithm iteration is finished, if not:

$$E\left(r_k\right) = r_f + \beta\left(E\left(r_M\right) - r_f\right). \tag{9}$$

3.2. Particle Swarm Optimization. The research puts forward a multiauxiliary information fusion space-time model, which can better deal with the dynamic and complexity of traffic flow (Figure 2). Second, it breaks through the traffic knowledge mining technology based on reasoning of temporal knowledge association rules and realizes the completion of incomplete knowledge base and mining of tacit knowledge by using spatiotemporal matching and periodic rules.

Aiming at the problem that the efficiency of association retrieval of complex rules of mass knowledge is low due to the high requirement of effectiveness in traffic governance scenes, the knowledge efficient retrieval technology based on graph database is firstly broken through, and the community attribute search algorithm is used to realize the rapid retrieval of mass data and multidimensional complex association relations. Secondly, it breaks through the traffic knowledge mining technology in typical governance scenarios, and develops knowledge mining models of 7 individual activity travel rules, people affected by traffic events, and travel characteristics of large-scale activities, so as to realize in-depth mining and knowledge retrieval of complex rules in typical governance scenarios.

The overall process of constructing the influencing factor model of taxi travel is shown in Figure 3. First, a variety of tools such as fishing nets and grid calculator model builder are created in ArcGIS Pro to quantify track points, road network, and POI data of population density, and the dependent variables of the model are obtained. Independent variables, OLS GWR MGWR regression model was constructed, respectively. During model construction, experimental parameters were repeatedly adjusted and model results were comprehensively compared for many times to obtain the optimal solution of each model.

4. Experimental Results and Analysis

4.1. Introduction to Experimental Environment and Data Set. In the overall process of constructing the influencing factor model of taxi travel, first, a variety of tools such as fishing nets

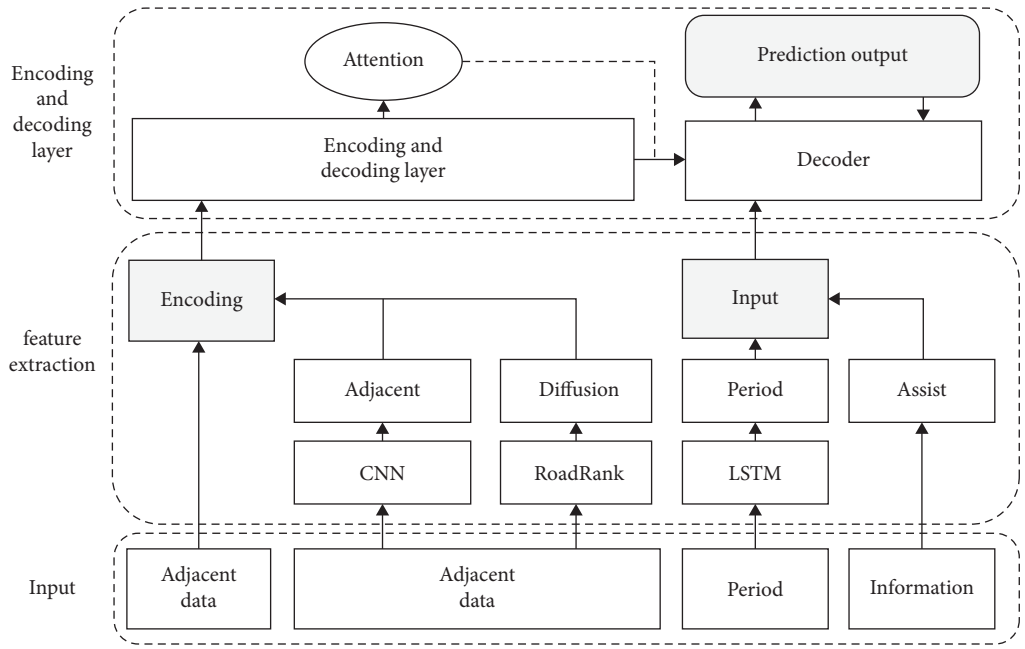


Figure 2: Multiauxiliary information fusion BP model considering traffic flow diffusion.

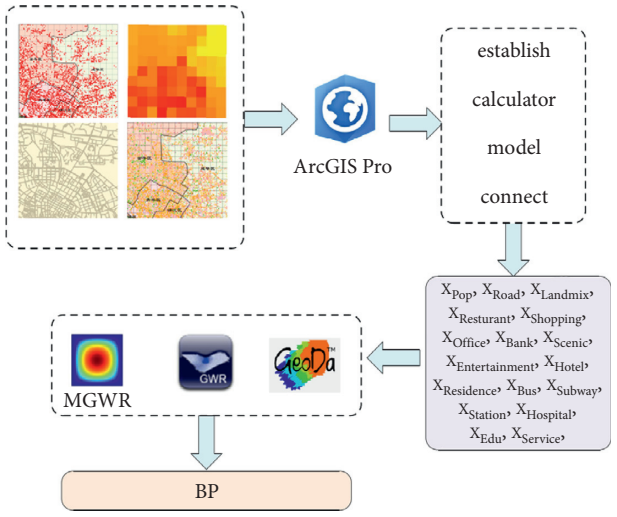


Figure 3: Comparison of optimization rate of three models.

and grid calculator model builder are created in ArcGIS Pro to quantify track points, road network, and POI data of population density, and the dependent variables of the model are obtained. For independent variables, OLS GWR MGWR regression model was constructed. During model construction, experimental parameters were repeatedly adjusted, and model results were comprehensively compared for many times to obtain the optimal solution of each model.

In this paper, the spatial data of Wuhan taxi provided by Sky Smart Travel Big data platform are used as the main data source, and other traditional statistics and related planning background data are used as auxiliary data for analysis and research.

4.2. Experimental Results Analysis. In order to verify the correlation of urban space active area and population distribution, the author extracted urban population density distribution and the urban space superposition found in the active region, and the urban space activity and has close relation with the city's population distribution, extracted the basic urban space active area that is located in the high-density urban core area, perfectly matching with the conventional judgment. It indicates that the urban spatial active area obtained by the above method is relatively accurate, as shown in Figure 4.

As shown in Figure 5, on a week-by-week basis, the highest number of trips were made on Friday, with about 3,000 fewer trips made on Saturday and the number of trips made on Sunday declined sharply, basically the lowest point or lower point in each cycle. On November 4, 11, 18, and 25, the four Fridays exceeded the usual number of trips. The time period when a large number of taxis took place at 18 : 00 in the evening is between 5 : 00 and 24 : 00, an average of 500 fewer trips per hour.

The law of arrival point is similar to that of the starting point, but the difference lies in the obvious fluctuation of arrival point, the peak value and valley value are more prominent, and the trend is not as smooth as that of the starting point. The peak time is relatively short, and 9 : 00, 13 : 00, and 18 : 00 constitute the peak of working days and 12 : 00 16 : 00 constitutes the low point. The low point of arrival on the rest day is at 16 : 00. Compared with the starting point, the trend change of arrival points and the increase of the number of arrival points have a certain backward delay, and the departure time is scattered and the arrival time is concentrated. In the morning of working days, the quantity of arrival points and starting points in the study

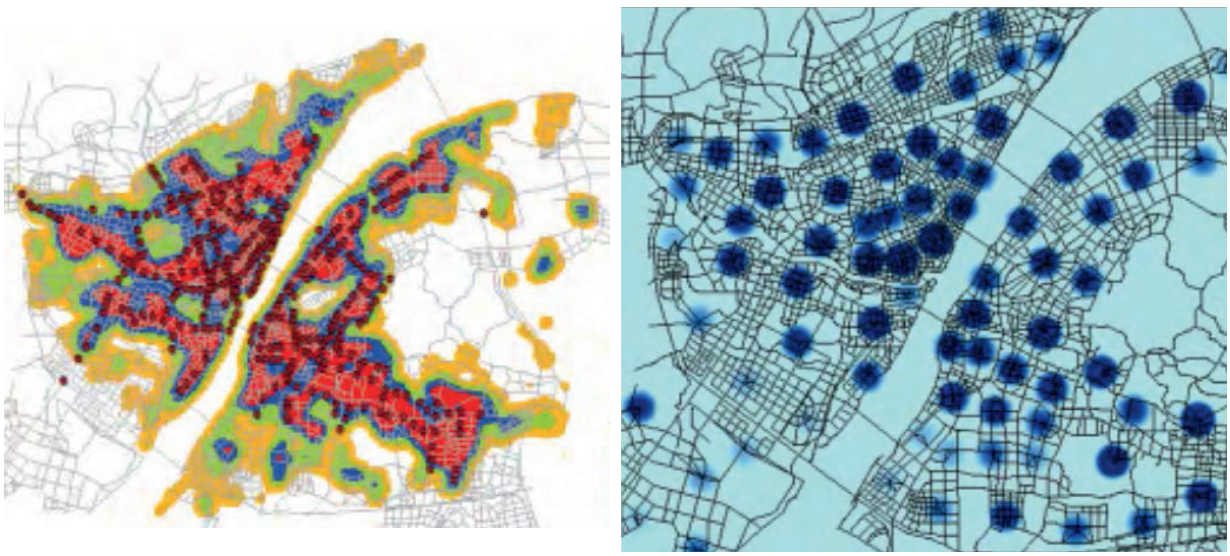


Figure 4: Matching map of population density and spatial active area.

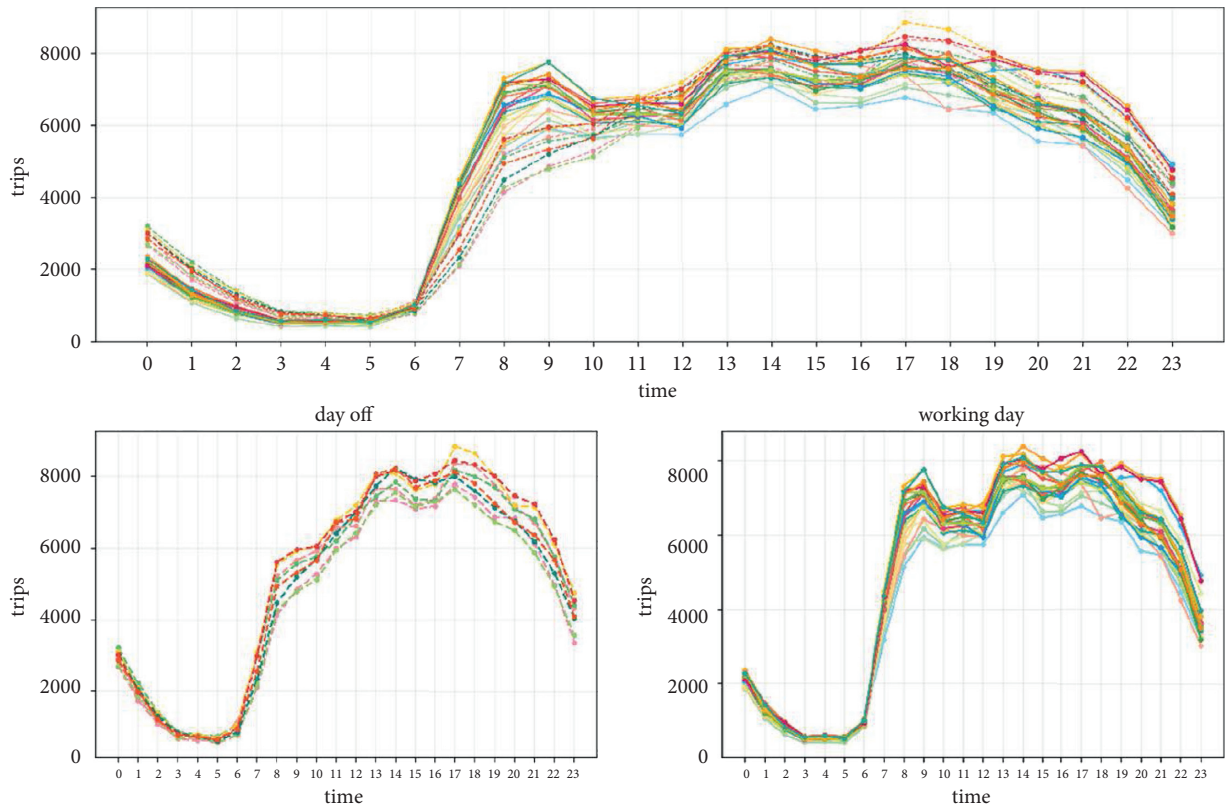


Figure 5: Distribution of arrivals on rest weekdays.

area differs greatly from 8 : 00 to 10 : 00, indicating that the regional inflow is greater than outflow in the morning.

The core points connected with density and the boundary points in the core point neighborhood together form a class. The red core points and yellow boundary points shown in Figure 6 belong to the same class, Density-Based Spatial Clustering of Applications with Noise (DBSCAN). The tightness of sample point distribution is measured by judging whether the sample points are connected in density. The closely connected samples are divided into one category to obtain a cluster category, and the samples connected in density are divided into different categories to obtain all cluster categories.

Advantages of hierarchical clustering: the definition of distance and similarity is simple, and there are few restrictions; there is no need to determine the number of classes before clustering; it can flexibly realize multilayer similarity clustering.

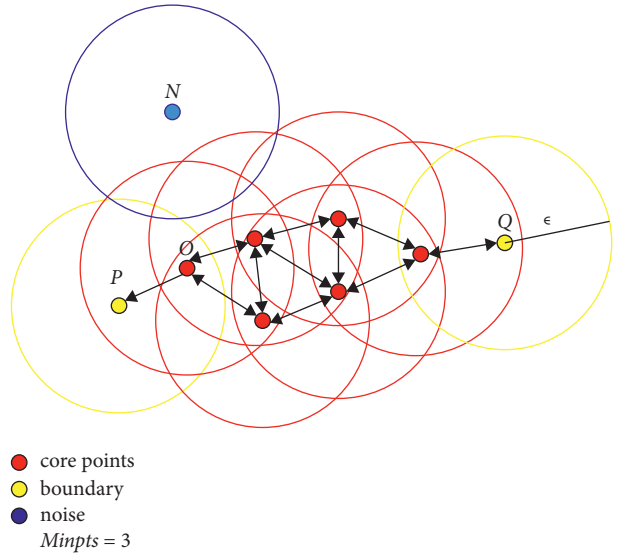


Figure 6: Cluster results of DBSCAN.

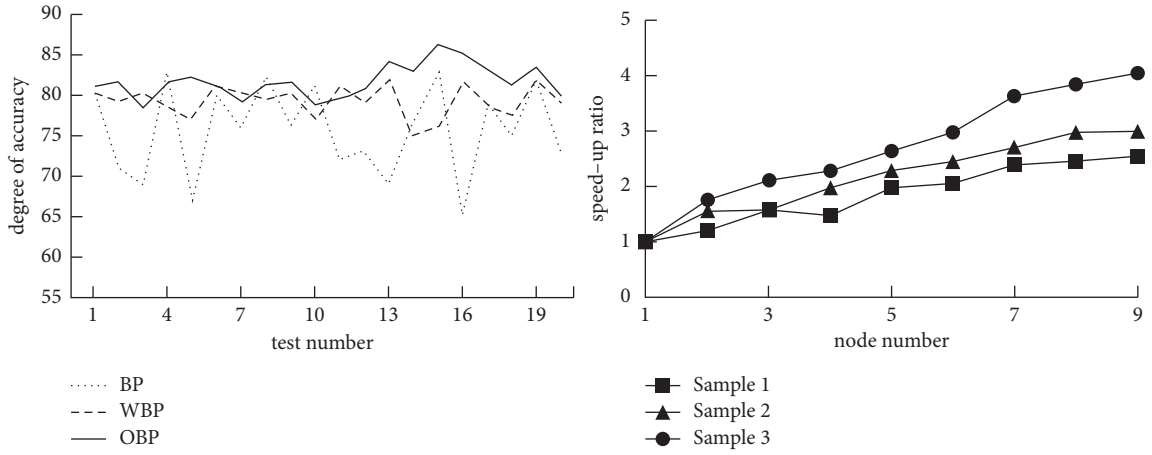
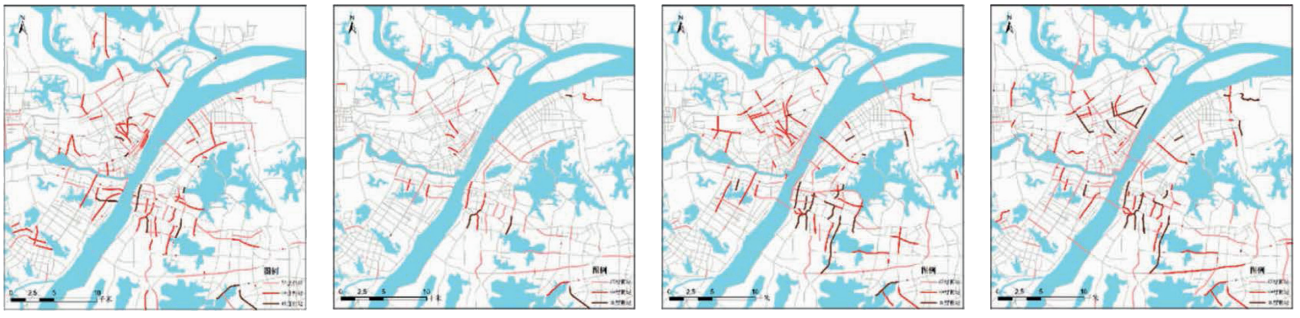


Figure 7: Comparison of accuracy and acceleration ratio test results.

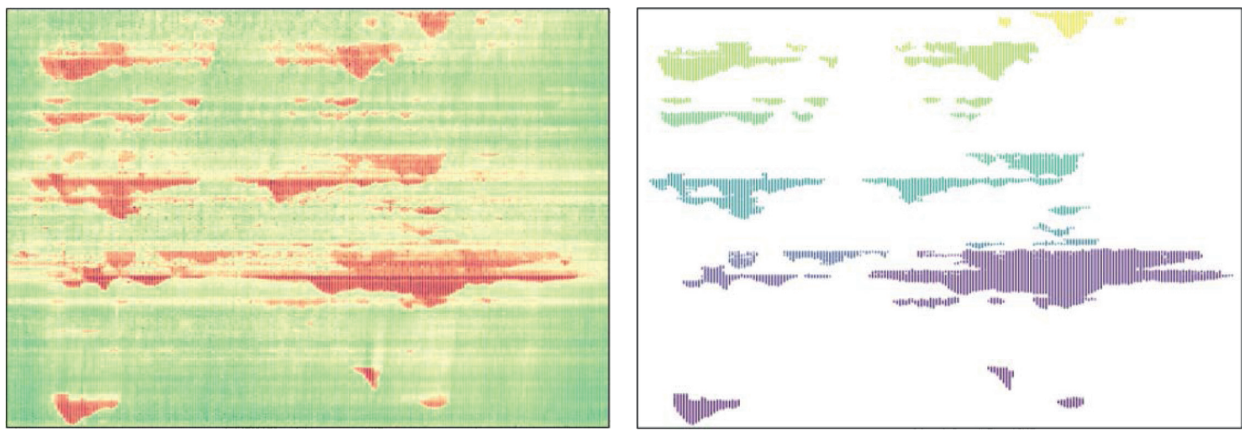


(a) (b) (c) (d)

Figure 8: Sketch of global marginal ranking.

Disadvantages: distance matrix needs to be calculated, which has high time and space complexity. The extreme distribution of sample points affects the clustering accuracy. There may be chain-like clusters, and in the case of uneven distribution of data points and large difference in density, a single hierarchical threshold can only extract fixed cluster.

Automatic identification of travel modes with smart phone data can not only provide a large amount of data basis for traffic planning but also provide a basis for real-time



(a) (b)

Figure 9: Travel congestion identification based on BP neural network.

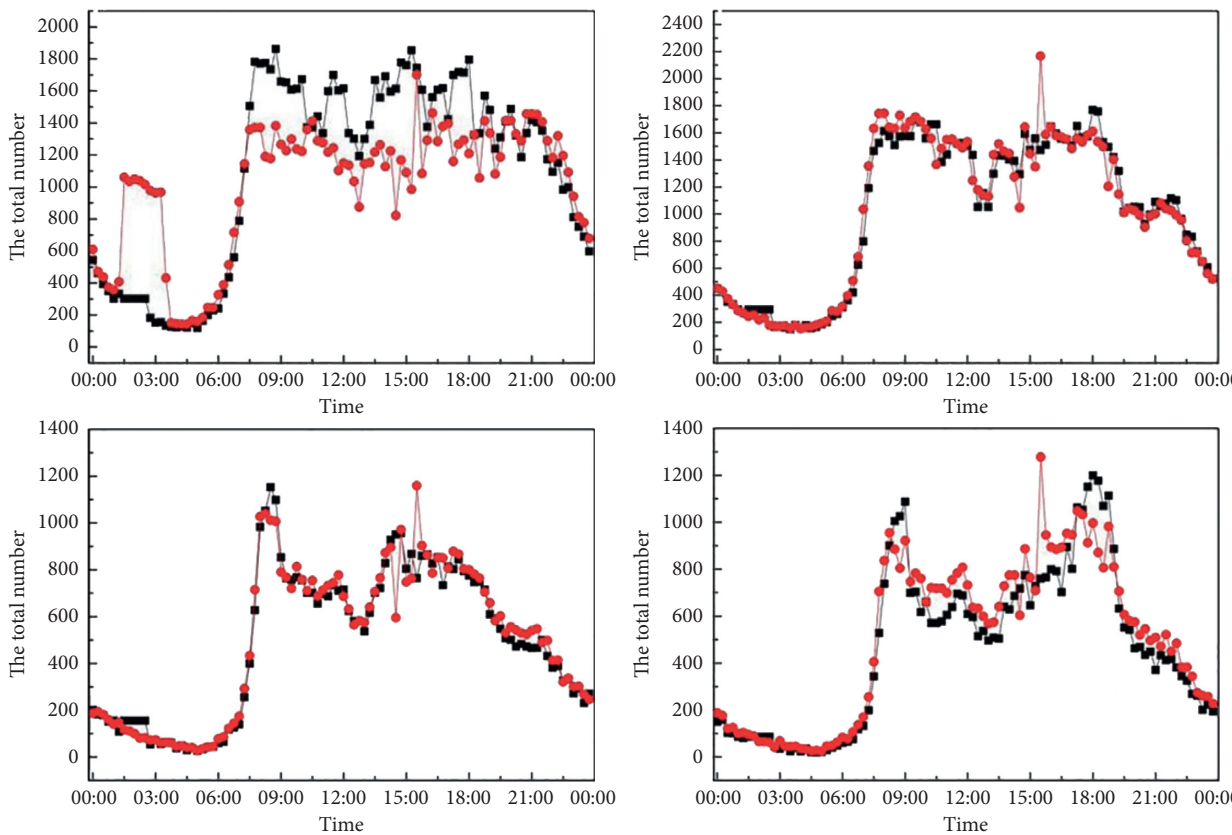


Figure 10: The prediction based on BP is compared with the measured data.

traffic control and guidance. However, traffic travel behavior itself is very complex, and travel background distance and other attributes are important factors affecting travel mode recognition. There is still a lot of work to be done in this direction in the future.

The number of nodes was successively increased to run the parallel FCM clustering integration algorithm, and the number of nodes was increased from 1 to 9. Data sample 1, sample 2, and sample 3 were clustered, respectively, and the time of each running was recorded. Each group of experiments was run for 10 times, respectively, and the average running situation was calculated, as shown in Figure 7. The acceleration ratio of all the three sample sizes shows an increasing trend, and the increasing trend is more obvious with the increase of the sample data volume, indicating that the algorithm has a strong ability to process large-scale data, and the acceleration ratio performance is good.

Cluster integration operation is carried out on 3 nodes, 6 nodes, and 9 nodes, respectively, for sample data of three sizes.

Record between the run time of each experiment, the experiment in each group is repeated 10 times, and the average operation is calculated, as shown in Figure 7. With the increase of the number of nodes, three samples of the algorithm running time are now to reduce the trend, the greater the amount of data, the decreasing trend, the more obvious can be seen from the diagram, sample 3 reduce the trend of the most obvious. The running time of the nine nodes is basically only one-third of that of the three nodes, while the other two samples are only reduced by about one half, so the algorithm has good scalability when processing large-scale data.

Before and after the epidemic, the morning rush hour congested roads on weekdays were visualized by levels, as shown in Figure 8. From the perspective of spatial distribution, the heavily congested roads before the epidemic (November 2019) mainly included Jiefang Avenue, Zhuyeshan interchange, Jiefang Road, Nanhu Road, Lhoshi Road, Fu Road, Guanggu Avenue, Yangguang Avenue, etc. In the initial stage of the resumption of work and production (March 2020), the distribution of congested sections was significantly reduced, including Baishazhou Avenue and Lhoshi Road and Auxiliary Road of Optics Valley Avenue. After the full resumption of work and production (June 2020), the urban traffic gradually recovered, and the traffic congestion gradually increased; After the recovery of the epidemic (August 2020), the number of severely congested roads increased significantly, mainly including Changqing Road, Zhuyeshan overpass, Sanyanqiao Road, Jiefang Road, Jiefang Road, Baishazhou Road, Zhongshan Road, and Luoshi Road, indicating that the urban traffic was further restored and the traffic pressure gradually increased.

Before and after the epidemic, the congested roads in the evening rush hours on weekdays were classified. Before the epidemic (November 2019), the heavily congested roads mainly included The Second Ring Road (Yangxin Road), Jiefang Road, Zhuye Mountain Interchange, Wuluo Road, Luoyu Road, Shengli Street, etc. In the early stage of the resumption of work and production (March 2020), the traffic congestion situation was significantly reduced, and all congested roads were significantly reduced. After the full resumption of work and production (June 2020), as in the morning rush hour, the congested sections increased significantly. The urban traffic gradually returned to normal, and the heavily congested sections mainly include Development Avenue, Jiefang Avenue, Sanyan Bridge Road, Joy Avenue, Jiefang Road, Bayi Road, Zhongbei Road, etc. On the one hand, it indicates that urban traffic has fully recovered, and on the other hand, traffic congestion has slightly intensified, which may be related to the increase in the proportion of self-driving trips after the recovery of the epidemic.

Aiming at the problem that it is difficult to accurately identify the influencing factors of traffic governance with multiple governance targets, the short-term congestion symptom recognition technology based on the trans-media traffic knowledge graph is broken through, and two symptom recognition models of road network congestion state and bus service level are constructed, with the success rate of congestion recognition reaching 92.4% (Figure 9). Based on the study of structural change points of bus passenger flow, the influence of track on bus is analyzed.

Aiming at the requirements of agile early warning of congestion in the whole urban road network, the technology method combining meso dynamic traffic simulation and deep learning was broken through. Taking Futian Central District of Shenzhen as the test object, the urban road network online deduced real traffic for 2 hours, and the test result of running time was 3.728 S, the prediction accuracy of key roads exceeds 85%, and the specific results are shown in Figure 10.

5. Conclusions

To sum up, this paper conducts theoretical and application research on the cognitive law of urban traffic complex system and the big data intelligent travel algorithm based on the BP neural network. The continued work has laid a solid foundation for the future to promote the integration of key technologies and more application demonstration scenarios and rely on the landing scenario to improve the data fusion perception traffic knowledge map traffic situation inference and cloud edge system platform indicators.

In terms of data fusion perception, it will rely on more application scenarios to further break through the theoretical and technical bottlenecks of traffic data processing and analysis, accurately extract and efficiently store effective information in traffic monitoring, and realize the restoration of the traffic panorama and in-depth analysis and cross analysis of its features.

Data Availability

The data set can be accessed upon request to the author.

Conflicts of Interest

The author declares no conflicts of interest.

References

[1] L. Zhu, F. R. Yu, Y. Wang, B. Ning, and T. Tang, "Big data analytics in intelligent transportation systems: A survey[J]," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 1, pp. 383–398, 2018.

[2] S. Kaffash, A. T. Nguyen, and J. Zhu, "Big data algorithms and applications in intelligent transportation system: A review and bibliometric analysis," *International Journal of Production Economics*, vol. 231, Article ID 107868, 2021.

[3] G. Zeng, "Application of big data in intelligent traffic system," *IOSR Journal of Computer Engineering*, vol. 17, no. 1, pp. 01–04, 2015.

[4] X. Hu and Z. Deng, "Research on perception bias of implementation benefits of urban intelligent transportation system based on big data[J]," *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no. 1, pp. 1–7, 2019.

[5] L. M. M. Visuwasam and D. P. Raj, "A distributed intelligent mobile application for analyzing travel big data analytics," *Peer-to-Peer Networking and Applications*, vol. 13, no. 6, pp. 2036–2052, 2020.

[6] R. Mau, "What's in a name? Purchases and sales of financial assets as a monetary policy instrument," *Purchases and Sales*

of Financial Assets As a Monetary Policy Instrument, vol. 2020, 2020.

[7] Q. Cui, Y. Wang, K. C. Chen et al., "Big data analytics and network calculus enabling intelligent management of autonomous vehicles in a smart city," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2021–2034, 2018.

[8] G. Fusco, A. Bracci, T. Caligiuri, C. Colombaroni, and N. Isaenko, "Experimental analyses and clustering of travel choice behaviours by floating car big data in a large urban area," *IET Intelligent Transport Systems*, vol. 12, no. 4, pp. 270–278, 2018.

[9] C. J. Wang, C. Y. Ng, and R. H. Brook, "Response to COVID-19 in Taiwan: Big data analytics, new technology, and proactive testing," *JAMA*, vol. 323, no. 14, pp. 1341-1342, 2020.

[10] A. I. Torre-Bastida, J. Del Ser, I. Laña, M. Ilardia, M. N. Bilbao, and S. Campos-Cordobés, "Big Data for transportation and mobility: Recent advances, trends and challenges," *IET Intelligent Transport Systems*, vol. 12, no. 8, pp. 742–755, 2018.

[11] H. Lu, Z. Sun, and W. Qu, "Big data and its applications in urban intelligent transportation system[J]," *Journal of Transportation Systems Engineering and Information Technology*, vol. 15, no. 5, pp. 45–52, 2015.

[12] V. Konecny, C. Barnett, and M. Poliak, "Sensing and computing technologies, intelligent vehicular networks, and big data-driven algorithmic decision-making in smart sustainable urbanism," *Contemporary Readings in Law and Social Justice*, vol. 13, no. 1, pp. 30–39, 2021.

[13] A. Farouk and D. Zhen, "Big data analysis techniques for intelligent systems," *Journal of Intelligent and Fuzzy Systems*, vol. 37, no. 3, pp. 3067–3071, 2019.

[14] N. Hoseinzadeh, R. Arvin, A. J. Khattak, and L. D. Han, "Integrating safety and mobility for pathfinding using big data generated by connected vehicles," *Journal of Intelligent Transportation Systems*, vol. 24, no. 4, pp. 404–420, 2020.

[15] M. Shengdong, X. Zhengxian, and T. Yixiang, "Intelligent traffic control system based on cloud computing and big data mining," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 12, pp. 6583–6592, 2019.

[16] Y. Chen, "Intelligent algorithms for cold chain logistics distribution optimization based on big data cloud computing analysis[J]," *Journal of Cloud Computing*, vol. 9, no. 1, pp. l–12, 2020.

[17] C. Anda, A. Erath, and P. J. Fourie, "Transport modelling in the age of big data," *International Journal of Urban Sciences*, vol. 21, no. sup1, pp. 19–42, 2017.

[18] W. Li, J. Zhu, Y. Zhang, and S. Zhang, "Design and implementation of intelligent traffic and big data mining system based on internet of things," *Journal of Intelligent and Fuzzy Systems*, vol. 38, no. 2, pp. 1967–1975, 2020.

[19] Y. Lv, Y. Duan, W. Kang, Z. Li, and F.-Y. Wang, "Traffic flow prediction with big data: A deep learning approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 2, pp. 865–873, 2014.

[20] J. Fiosina and J. Ã Maxims Fiosins, "Big data processing and mining for next generation intelligent transportation systems," *Jurnal Teknologi*, vol. 63, no. 3, 2013.

[21] H. Dadashi, "Optimal investment strategy post retirement without ruin possibility: A numerical algorithm," *Journal of Computational and Applied Mathematics*, vol. 363, pp. 325–336, 2020.

[22] W. Serrano, "A big data intelligent search assistant based on the random neural network," in *Proceedings of the INNS Conference on Big Data*, pp. 254–261, Springer, Cham, 2016.

[23] Y. Han, Y. Kim, and Y. Kim, "A study of measuring traffic congestion for urban network using average link travel time based on DTG big data," *The Journal of The Korea Institute of Intelligent Transport Systems*, vol. 16, no. 5, pp. 72–84, 2017.

[24] L. Song, K. Zhang, T. Liang, X. Han, and Y. Zhang, "Intelligent state of health estimation for lithium-ion battery pack based on big data analysis," *Journal of Energy Storage*, vol. 32, Article ID 101836, 2020.

[25] H.-Q. Zhang and Z. Li, "Intelligent travelling visitor estimation model with big data mining," *Enterprise Information Systems*, vol. 15, no. 1, pp. 1–14, 2021.

[26] S. Nesmachnow, S. Baña, and R. Massobrio, "A distributed platform for big data analysis in smart cities: Combining intelligent transportation systems and socioeconomic data for Montevideo, Uruguay[J]," *EAI Endorsed Transactions on Smart Cities*, vol. 2, no. 5, Article ID 153478, 2017.

[27] Z. Wang, X. Li, X. Zhu, J. Li, F. Wang, and F. Wang, "Big data-driven public transportation network: A simulation approach [J]," *Complex & Intelligent Systems*, vol. 6, pp. 1–13, 2021.

[28] G. Gang Xiong, F. Fenghua Zhu, X. Xisong Dong et al., "A kind of novel ITS based on space-air-ground big-data," *IEEE intelligent transportation systems magazine*, vol. 8, no. 1, pp. 10–22, 2016.

[29] T. Zheng, G. Chen, X. Wang, C. Chen, X. Wang, and S. Luo, "Real-time intelligent big data processing: Technology, platform, and applications[J]," *Science China Information Sciences*, vol. 62, no. 8, pp. l–12, 2019.

[30] Z. Hou, Y. Zhou, and R. Du, "Special issue on intelligent transportation systems, big data and intelligent technology," *Transportation Planning and Technology*, vol. 39, no. 8, pp. 747–750, 2016.

[31] M. G. Kibria, K. Nguyen, G. P. Villardi, O. Zhao, K. Ishizu, and F. Kojima, "Big data analytics, machine learning, and artificial intelligence in next-generation wireless networks," *IEEE access*, vol. 6, pp. 32328–32338, 2018.

[32] X. Zhang and Z. Yuan, "The GPS trajectory data research based on the intelligent traffic big data analysis platform," *Journal of Computational Methods in Sciences and Engineering*, vol. 17, no. 3, pp. 423–430, 2017.

[33] J. Chang, S. Nimer Kadry, and S. Krishnamoorthy, "Review and synthesis of Big Data analytics and computing for smart sustainable cities," *IET Intelligent Transport Systems*, vol. 14, no. 11, pp. 1363–1370, 2020.

[34] M. V. Moreno, F. Terroso-Sáenz, A. González-Vidal et al., "Applicability of big data techniques to smart cities deployments[J]," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 2, pp. 800–809, 2016.

[35] A. Fonzone, J.-D. Schmöcker, and F. Viti, "New services, new travelers, old models? Directions to pioneer public transport models in the era of big data," *Journal of Intelligent Transportation Systems*, vol. 20, no. 4, pp. 311–315, 2016.

[36] M. Mohammadi and A. Al-Fuqaha, "Enabling cognitive smart cities using big data and machine learning: pproaches and challenges," *IEEE Communications Magazine*, vol. 56, no. 2, pp. 94–101, 2018.

Hindawi
Scientific Programming
Volume 2022, Article ID 3220986, 7 pages
https://doi.org/10.1155/2022/3220986

Research Article

Forecast of Foreclosure Property Market Trends during the Epidemic Based on GA-BP Neural Network

YaoChing Fang

Institute of China and Asia-Pacific Studies, National Sun Yat-sen University, Kaohsiung 80424, Taiwan

Correspondence should be addressed to YaoChing Fang; d106070005@nsysu.edu.tw

Received 27 December 2021; Accepted 12 February 2022; Published 31 March 2022

Academic Editor: Baiyuan Ding

Due to the epidemic, foreclosure, as a special transaction commodity, has also been affected to a certain extent. This paper uses the sample data of foreclosure auctions combined with the genetic algorithm multi-layer feedforward neural network to apply in the forecast of the foreclosure market trend. First of all, to select the indicators that affect the price fluctuation of foreclosed houses, use factor analysis technology to reduce the dimensionality of the original data to obtain two common factors; then combine the BP neural network and the genetic algorithm to propose a GA-BP neural network The prediction model of the foreclosure housing market takes the influencing factors as the input of the model and the housing price as the output. Finally, the prediction results are compared with the multiple linear regression and the extreme learning machine, which shows that the GA-BP neural network model is in It is a good choice for research, and it provides an effective method for predicting housing prices.

1. Introduction

As a circulating commodity, foreclosed houses are also in line with the laws of market operation. The supply of foreclosed houses and market demand determine their market position [1–5]. The two are related, but on another level, foreclosed houses are born with distinctive features that are different from other houses. The characteristic, that is, the national judicial coercive force and enforcement power granted by itself, to a certain extent, occupy a great advantage, that is, the convenience and uniqueness of the sales and transfer links [6, 7]. Therefore, the epidemic cannot have a subversive impact on the foreclosure. More viewpoints point out that in recent years, the domestic market has been sluggish and lacklustre, many companies and individuals are facing bankruptcy, and banks and financial institutions have increasingly defaulted. Affected by the epidemic at the beginning of the year, the economic situation forecast makes people unable to be optimistic [8–11]. It is believed that foreclosure properties will soon usher in a peak supply period. The supply increases and the market purchasing power has not followed up. Then the advantages of foreclosure properties cannot be reflected. With the decrease in investment or rigid demand, prices will inevitably fall, and the market price of foreclosure housing will undoubtedly hit the bottom [1, 12–15].

How long will it take to transform from ordinary real estate to forensic auction, from litigation to the final execution of the auction? Two years as fast as possible, and three to five years as slow as possible [16, 17]. The duration of this cycle determines the stability of the supply of foreclosures. From this, it is not difficult to know that from the analysis of the supply source of foreclosure houses, the probability that the rate of foreclosure auctions will surge or decrease in a short period of time is extremely low. As we all know, since the foreclosure auction room has been converted from offline to online, the China Auction Platform, Taobao, JD.com and other online platforms have joined hands with the people's courts, and professional auction institutions have settled in to assist, and the national court system's online auction work has become more standardized and functioning well., Overall, the volume of foreclosure auctions has steadily increased year by year [5, 18–21].

With the various conveniences of online foreclosure transactions being widely accepted by home buyers, foreclosures are undoubtedly a beacon in the long night for the dream of buying a house blocked by the purchase restriction policy, bringing it to many buyers and investors Hope. For middle-class and above families, whether it is for the next generation to deploy in big cities in advance to help them settle down, or to make family assets safer and more appreciation and preservation, buying a house should be a relatively safe and feasible way. The foreclosure room just meets the needs of these families. The shining sign of "Unlimited Purchase" has attracted the attention of the whole society, and all foreclosures have a formal trading platform, and the foreclosures are endorsed by the people's court, the operation is standardized, and the bidding is transparent, which can interpret the auction to the greatest extent." The connotation of fairness, impartiality and openness; more importantly, professional auxiliary agencies provide a variety of thoughtful and thoughtful services to dispel doubts and worries. For those who really understand the foreclosure, it is undoubtedly a "leakage," although "Leak-picking" is a speculative behavior, but it can also be understood as a market behavior, and a large number of rigidly needed groups have always existed. The huge market demand and gap determine the rationality of the market. Affected by the epidemic, which affects all levels of society, foreclosure as a special transaction commodity should be listed among them. As for the extent and depth of its impact, there is still no conclusion [12]. Up to now, great progress has been made in housing price forecasting research methods. Hu Feng analyzed the historical data from 2006 to 2015 by establishing a regression model, and obtained the main factors affecting housing prices. Zhang Shuangni established a housing price prediction model based on stepwise regression. Liu Cong established a regression analysis model based on principal components when analyzing the influencing factors of housing prices, and established a time series model when predicting housing prices.

In order to further improve the accuracy of the foreclosure house price prediction model, this paper attempts to use the foreclosure information data combined with the genetic algorithm multi-layer feedforward neural network (genetic algorithm-back propagation, GA-BP) to be applied in the foreclosure market trend forecast. First, select the indicators that affect the fluctuation of the house price for foreclosure, and use factor analysis technology to reduce the dimension of the original data to obtain two common factors; then, use the GA-BP neural network to combine the BP neural network and the genetic algorithm., put forward a prediction model for the foreclosure housing market based on GA-BP neural network, which takes the influencing factors as the input of the model and the housing price as the output; finally, the prediction results are compared with multiple linear regression and extreme learning machine, which proves that the method is suitable for foreclosure auctions. Advantages in housing price forecasts.

2. Determination of the Influencing Factors of Foreclosure House Price Based on Factor Analysis

2.1. Factor Analysis Method Selection Index. The factor analysis method is a multivariate statistical method that extracts common factors from variable groups, thereby realizing dimension reduction of multiple variables. Factor analysis was first used in the field of psychology to study the fundamental factors affecting human intelligence, and now it has been widely used in various disciplines. This paper uses the number of auction houses and transactions in the court system of Yulin City (including districts and counties), counts the number of auction houses and transactions in the court system from 2018 to 2021 as a sample, and selects indicators that affect the fluctuation of foreclosure house prices. Using SPSS19 for factor analysis, the indicators are shown in Table 1.

(1) Applicability test. Enter the relevant variables of the foreclosure room into SPSS19, and the KMO value is 0.765, which is greater than 0.5, so there is a correlation between the variables. In the Bartlett sphere test, the P value is 0, which is less than 0.05, and it is considered that there is no significant difference between the variables. Therefore, common factors can be extracted, which is suitable for factor analysis.

(2) Factor rotation and interpretation. Two common factors are calculated: the first factor has a higher load on x1, x2, x3, x4, and x5, and it is named the foreclosure market impact factor; the second factor has on x6, x7, and x8 The higher the load, it is named the real estate market impact factor. According to the characteristic value greater than 1 and the cumulative variance contribution rate reached 92.689%, the original 8 variables were finally reduced in dimensionality.

(3) Calculate factor function. Get the factor function according to the score coefficient matrix, Among them, z_i ($i = 1, 2, 8$) is the standardized variable of x_i.: $F_1 = 0.237z_1 + 0.003z_2 + 0.221z_3 + 0.176z_4 + 0.192z_5 + 0.271z_6 + 0.099z_7 + 0.166z_8$, $F_2 = -0.081z_1 + 0.266z_2 - 0.045z_3 + 0.036z_4 + 0.01z5 - 0.177z6 - 0.403z7 - 0.459z8$.

2.2. Data Preprocessing. The data used in the research comes from the foreclosure housing price data set of a provincial court system. Then the following data processing methods were adopted according to the data of the data set. Data normalization is an important technique. Generally speaking, before using the neural network model, it must be used to preprocess the data. It can standardize the range of each feature in the input data set. If the model is not preprocessed before training, the model may not be able to converge to the optimal solution. In the multi-index evaluation method, due to the different nature of each evaluation index, it usually has different dimensions and orders of magnitude. If the initial data is used directly, the data of different dimensions will have different degrees of influence on the network, which

Table 1: Impact indicators.

Serial number	Index
x_1	Participate in registration
x_2	Buyer market
x_3	Total price
x_4	Appraisal price
x_5	House area
x_6	GDP
x_7	Seller's taxes
x_8	Present value of house

will eventually make it difficult to learn the model. . Therefore, it is necessary to perform a series of standardized processing on the original data, and then input it into the neural network model, which can reduce the difficulty of subsequent model training and improve the performance of the network model. Data standardization is to scale the data to a smaller specific interval. It is usually used in the comparison or evaluation of the indicators to remove the unit limit of the original data, eliminate the magnitude of different data, and is suitable for comprehensive comparison and evaluation of indicators of different dimensions.

2.2.1. Duplicate value processing. When obtaining data, there are often some repeated data, and the repeated data will affect the statistical results and mislead the decision-makers. Use the duplicated function that has been written in python, which will compare all the columns. If the value of each column in two rows is the same, it will be marked as a duplicate value.

2.2.2. Outlier handling. The Z-SCORE standardization method is to standardize the data according to the mean and standard deviation of the original data. All data can be converted to an area close to zero. This method is suitable for situations where the maximum range and minimum range of attribute data are uncertain or exist. In the case of outliers, the processed data will satisfy a normal distribution with a mean of 0 and a standard deviation of 1.

$$Z = \frac{(X - \mu)}{\sigma}. \tag{1}$$

Among them: μ is the overall average, X-μ is the deviation from the mean, and σ is the standard deviation. The absolute value of z represents the distance between the original score within the standard deviation and the overall mean. Determine whether the Z-score score is greater than 2.2 (here 2.2 represents an empirical value), if it is, it is True, otherwise it is False. Directly delete the data row whose judgment result is true.

2.3. Model Evaluation Method. It is very important to evaluate the performance of a predictive model. It is generally measured from the accuracy of the model and training time, and an appropriate cost function is selected to measure the results of the model to evaluate the pros and cons of the model. If the loss is large, it means that the predicted value deviates far from the actual value, and the predicted result is not ideal; the smaller the loss, the better the model effect; if the predicted value is equal to the true value, it means that there is no loss. There are different methods for selecting appropriate and effective error analysis for different models. The 4 evaluation methods of MSE, MAE, EV, and R^2 are used to evaluate the training effect of the model.

2.3.1. Mean square error (MSE)

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(f(x_i) - y_i\right)^2. \tag{2}$$

The range is [0, +∞). When the predicted value is exactly the same as the true value, it is called a complete model; otherwise, the greater the error, the greater the value.

2.3.2. Mean absolute error (MAE)

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|f(x_i) - y_i\right|. \tag{3}$$

The range is [0, +∞). When the predicted value is exactly the same as the true value, it is called a complete model; otherwise, the greater the error, the greater the value.

2.3.3. Root mean square error (RMSE)

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}\left(f(x_i) - y_i\right)^2}{n}}. \tag{4}$$

The formula, also known as the standard error, represents the expected value of the square of the error and can reflect the degree of dispersion of a data set. n is the number of model measurements.

2.3.4. Mean absolute percentage error (MAPE)

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{f(x_i) - y_i}{f(x_i)}\right| \times 100\%. \tag{5}$$

In the range [0, +∞), a map of 0% indicates a perfect model, and a MAPE greater than 100% indicates an inferior model. From the formula point of view, MAPE and MAE have similarities, the difference is mainly in the position of the denominator. When the true value has data equal to 0, there is a problem of dividing the denominator by 0, and this formula is not available.

2.3.5. Absolute coefficient (R^2)

$$R^2 = 1 - \frac{\sum \text{Squared\_Residuals}/N}{\text{Variance}_{Y_{\text{true}}}}. \tag{6}$$

The closer to 1, the better the effect. The meaning of R2 is from the perspective of least squares (that is, the second-order variance), indicating how much of the variance of the actual y value is explained by the predicted y value.

2.3.6. Difference Interpretation Score (EV)

$$EV = 1 - \frac{\text{Variance}\left(Y_{\text{true}} - Y_{\text{pred}}\right)}{\text{Variance}_{Y_{\text{true}}}}. \quad (7)$$

When the mean of the residual is 0, it is the same as R2.

3. 2GA-BP Neural Network Algorithm

Genetic algorithm is a kind of global search evolution algorithm that refers to biological evolution process. Analogous to the biological evolution process, genetic algorithms simulate biological evolution through selection, inferiority, and mutation to produce next generation solutions. By calculating and ranking the fitness of individuals, individuals with low fitness are eliminated, and the proportion of individuals with high fitness is increased. After several iterative calculations, the individual solution with the best fitness function value is found. The genetic algorithm's heuristic optimal solution search method combined with the BP neural network can effectively solve the problem that the BP neural network is easy to fall into the local optimal solution, and improve the accuracy of housing price prediction. The GA-BP neural network process is shown in Figure 1.

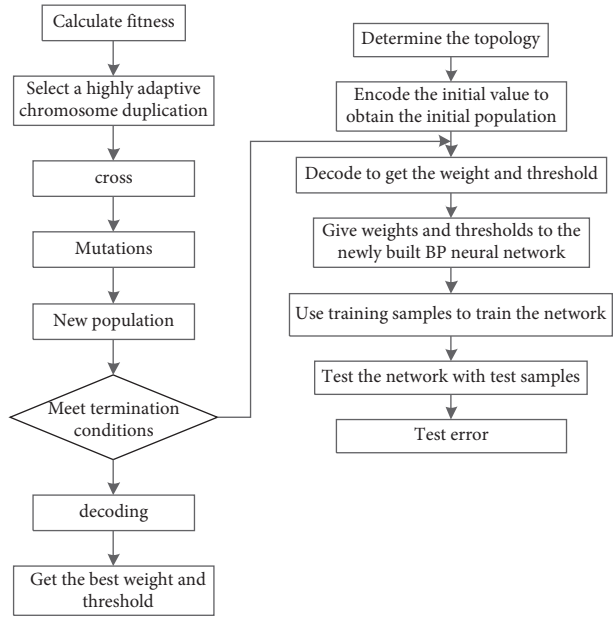


Figure 1: GA-BP network flow chart.

3.1. Genetic algorithm initialization. The genetic algorithm needs to decode the solution data of the solution space into genotype string structure data before using it. The commonly used encoding method is the binary encoding method, and different string structure data constitute different points. The N string structure data generated by encoding represent different individuals. Through the representation of different individuals, the total group of births is finally formed, also known as the initial population.

3.2. Fitness assessment. The fitness indicates the individual's ability to adapt, and also indicates the individual's strengths and weaknesses. The formula for calculating the individual's adaptive ability is

$$F = \text{mse}(Y - O) = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - o_i\right)^2,$$
$$F = \text{mse}(X - O) = \frac{1}{m}\sum_{i=1}^{m}\left(x_i - o_i\right)^2. \quad (8)$$

In the formula, m is the total number of samples; x_i is the output result of the genetic algorithm; O_i is the actual output result of the genetic algorithm, and mse represents the mean square error function.

3.3. Genetic manipulation. The principle of survival of the fittest is used in the group selection of genetic algorithm to select good individuals for the next generation of genetic reproduction. Individuals with strong adaptability can pass on excellent genes to their offspring through inheritance. In this paper, a proportional selection strategy is adopted, and the population data amount is set as M, then the fitness formula of individual i is

$$P_i = \frac{k/F_i}{\sum_{i=1}^{N} k/F_i}. \quad (9)$$

In the formula, k is the genetic coefficient.

In the genetic algorithm, the next generation of individuals is reproduced through crossover operations, so as to obtain new individuals with new characteristics. The crossover operation is used to exchange genetic information, and the mutant individual randomly changes the value of a string in the string structure data with a certain probability, which has the characteristic of small value.

3.4. Combination of genetic algorithm and BP neural network. The optimal value search using genetic algorithm can reduce the calculation amount of BP neural network and optimize the calculation process of BP neural network.

BP neural network draws on the model structure of neural network, and builds a highly complex learning system through the interconnection of a large number of neurons. It has strong self-organization, self-adaptation and self-learning capabilities, arbitrarily complex pattern classification capabilities and excellent multi-dimensional function mapping capabilities. BP network is composed of input layer, output layer and hidden layer. The hidden layer calculates data by setting one or more layers of neurons, and each layer of neurons can have several nodes. The BP neural network common language pattern classification problem such as double hidden has the advantage of fast classification speed, so the BP neural network structure with double hidden layers is used in this article. The neural network is adjusted through the weight matrix and error feedback between each layer to achieve the expected output result of the load. Compared with the traditional artificial neural

network, the double hidden layer BP neural network has improved in terms of parallel processing of massive data and accuracy. BP neural network is shown in Figure 2.

In this paper, the index factors that affect housing prices summarized above are used as input factors, and the predicted housing prices are used as output.

The input vector of BP neural network is $K = (k_1, k_2, k_3, \ldots k_n)^T$, The weight matrix of the input layer and the hidden layer is $S_{ni}(1<n<N,\ 1<i<I)$, and the weight matrix between the first hidden layer and the second hidden layer is $W_{ij}(1<i<I,\ 1<j<J)$, the output vector of the first hidden layer is $B = (b_1, b_2, b_3, \ldots, b_n)^T$, The threshold on node i is θi , The output vector of the second hidden layer is $C = (c_1, c_2, c_3, \ldots, c_n)^T$. The threshold at node j is θj , The weight matrix of the second hidden layer and the output layer is Wjn, The output vector of the output layer is $D = (d_1, d_2, d_3, \ldots, d_n)^T$, The threshold at node n is θ_n, f(x) is the activation function. Input the n-dimensional vector A, then the output of the first hidden layer node i is

$$b_i = f\left(\sum_{i=1}^{m} w_{mi} a_m - \theta_i\right). \tag{10}$$

The output of the second hidden layer node j is

$$c_j = f\left(\sum_{j=1}^{I} w_{ij} b_i - \theta_j\right). \tag{11}$$

The output result of the nth node in the output layer is

$$d_n = f\left(\sum_{n=1}^{J} w_{jm} c_j - \theta_n\right). \tag{12}$$

In order to improve the convergence speed of the BP neural network, the input data is normalized to reduce the range of changes and increase the flexibility of interval selection. The formula is as follows:

$$t_i' = \frac{t_i - t_{\min}}{t_{\max} - t_{\min}}, \tag{13}$$

where: t_j is the input value of the neural network, t_i is the normalized value of the interval [0, 1], and $t_{\min}$ is the minimum and $t_{\max}$ maximum value of the sample, respectively.

The activation function is as follows:

$$f(x) = \frac{1}{1 + e^{-x}}. \tag{14}$$

In view of the above analysis, by selecting the factors that affect the foreclosure property, the GA-BP neural network-based foreclosure property market prediction model is established, which can effectively predict the price of the property.

4. Experiment Analysis

In this paper, Matlab platform is used to design and implement a foreclosure market prediction model based on GA-BP neural network. According to the results of factor analysis, the obtained F1 and F2 are used as the model input, and the housing price as the output, to establish a housing price prediction model. The experimental data comes from the number of auctions and transactions of the court system in a certain province from 2018 to 2021 as sample data. This paper divides the data into training set and test set, the numbers are 580 and 480 respectively. The GA algorithm parameters are set as follows: the number of iterations is 200, the selection probability is 0.8, the crossover probability is 0.6, and the mutation probability is 0.08.

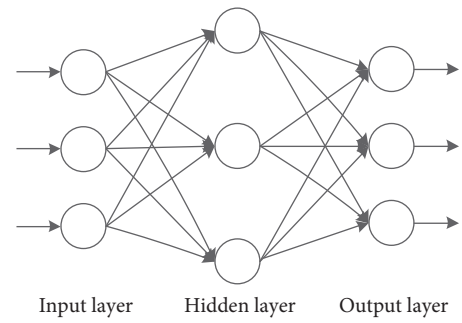


Figure 2: BP network topology.

When the model first iterated, the fitness value of the population individual was far from the optimal fitness value, and the individual fitness value increased significantly. At the later stage of the model iteration, due to the continuous convergence of the model, the population individual fitness value became more and more Close to optimal fitness. After completing the above work, the global optimal initial weight threshold searched by the genetic algorithm can be obtained, and the initial weight threshold can be brought into the network to train the BP network. Set the training parameters of the BP network: the number of neurons in the input layer is 6, the number of neurons in the hidden layer is 9, the number of neurons in the output layer is 1, and the unipolar sigmoid function is selected as the transfer Function, select the gradient descent BP training algorithm function training as the training function, the maximum number of training times is 200, the learning rate is 0.1, and the minimum mean square error of learning is 0.003. According to the selected 12 input vectors, the output data is normalized. The summary steps for establishing a housing price prediction model are as follows:

(1) Enter the original index data in SPSS19

(2) Use SPSS19 to do factor analysis to achieve dimensionality reduction of indicators, and to obtain new variable values

(3) Use the data from 2018 to 2020 as the training sample, and the data from 2017 to 2018 as the test sample

(4) Establish a GA-BP neural network prediction model based on training samples

(5) Based on test samples, verify the accuracy of the model

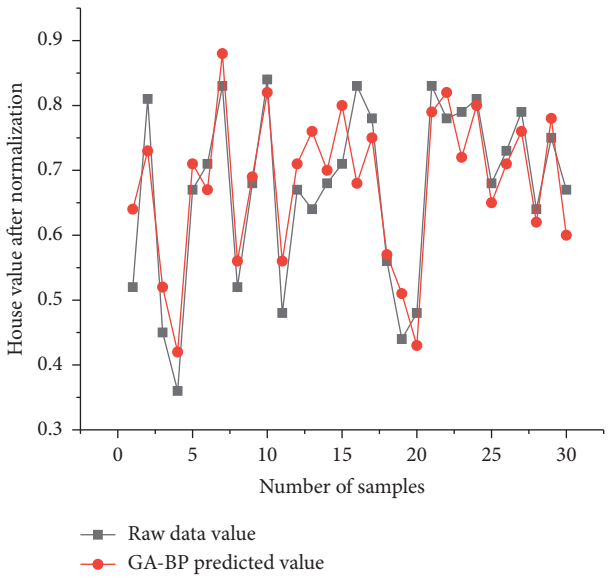


Figure 3: Forecast results.

Table 2: Model evaluation error analysis table.

Algorithm	MSE	MAE	EV	R^2
Multiple linear regression	0.0067	0.065	0.71	0.71
Extreme learning machine	0.0065	0.069	0.74	0.74
BP neural network	0.0051	0.073	0.68	0.67
GA-BP	0.0015	0.031	0.94	0.94

According to the constructed model, Matlab is used to establish a BP neural network model with two hidden layers, input the normalized data and iterate until the output results meet the prediction accuracy required by the experiment. The prediction results of training and testing samples obtained through GA-BP neural network algorithm analysis are shown in Figure 3, and samples are selected for display.

Figure 3 is a fitting diagram of the model and the training data. The dotted line is the original training set, and the solid line is the prediction result of the GA-BP model trained with the training set on each data in the training set. It can be seen that the prediction results can be better The trend of housing price changes is tracked, the prediction error is small, and the fitting effect is good.

It can be seen from Figure 3 that the GA-BP algorithm's prediction has a certain effect, but it is impossible to draw accurate conclusions only by observing the diagram. Therefore, a model evaluation algorithm needs to be used to evaluate the model. 4 evaluation methods such as MSE, MAE, EV, and R^2 are used to evaluate the training effect of the model. The error analysis details of the model evaluation are shown in Table 2.

It can be seen from Table 2 that the better model is the GA-BP model. When the model is used to fit the training data, the other three models are slightly less effective. When the GA-BP algorithm trains the model, it learns errors many times, so the prediction effect on the training data is better.

5. Conclusion

This article first selects the indicators that affect the price fluctuations of foreclosed houses, uses factor analysis technology to reduce the dimensionality of the original data, and obtains two common factors; then, combines the foreclosed house sample data with the method of GA-BP model to determine the The court system's foreclosure housing prices are predicted, and the predicted value is relatively close to the actual value. Moreover, by comparing the prediction results with multiple linear regression and extreme learning machine, it proves the superiority of this method in the accuracy of foreclosed house price prediction, which shows that the GA-BP neural network model is effective in the research of foreseeable house price prediction. A good choice, provides an effective method for the forecast of housing prices. [22].

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] P. Yue, R. Bai, and Y. Wang, "Review on second hand house price forecast," *Journal of Trend in Scientific Research and Development*, vol. 5, no. 1, 2020.

[2] K. Zhang and X. Zhu, "Analysis of Z city house price forecast based on application of multiple linear regression," in *Proceedings of the First International Symposium on Economics,Management,and Sustainable Development(EMSD 2019)*, pp. 74–77, Clausius Scientific Press, Hangzhou, China, 2019.

[3] Y. Liu and K. Li, "Research on house price forecast based on grey system GM(1, 1)," in *Proceedings of 2019 5th International Conference on Finance,investment,and Law(ICFIL 2019)*, pp. 208–214, Francis Academic Press, Hangzhou, China, 2019.

[4] J. Lai, L. Zhang, C. F. Duffield, and L. Aye, "Engineering reliability analysis in risk management framework: development and application in infrastructure project," *IAENG International Journal of Applied Mathematics*, vol. 43, no. 4, 2013.

[5] C. Pierdzioch, J. Christoph Rülke, and S. Georg, "House price forecasts, forecaster herding, and the recent crisis," *Italian Journal of Food Science*, vol. 1, no. 1, 2012.

[6] Y. Zhu, "Research on house price forecast based on multiple linear regression," *Basic and Clinical Pharmacology and Toxicology*, p. 128, 2021.

[7] J. Hu, Y. Wang, S. Wang, and A. Li, "Real estate warning based on BP neural network," in *Proceedings of the 2020 International Conference on Financial Economics and Investment Management*, Hangzhou, China, 2020.

[8] H. Oliver and R. Horst, "Euro area house prices and unconventional monetary policy surprises," *Economics Letters*, vol. 205, 2021.

[9] G. McGrevy Ryan and K. Sorensen, "A spatial model averaging approach to measuring house prices," *Journal of Spatial Econometrics*, vol. 2, no. 1, 2021.

[10] S. Fu, "Prediction and analysis of wuhan commercial housing prices based on grey prediction," *Modern Judaism*, vol. 10, no. 06, 2021.
[11] C. Zhou, "House price prediction using polynomial regression with Particle Swarm Optimization," *Journal of Physics: Conference Series*, vol. 1802, no. 3, 2021.
[12] Y. Sun, "Investigation on house price prediction with various gradient descent methods," *Journal of Physics: Conference Series*, vol. 1827, no. 1, 2021.
[13] A. R. Demong Nur, J. Lu, and K. Hussain Farookh, "An adaptive personalized property investment risk analysis method based on data-driven approach," *International Journal of Information Technology and Decision Making*, vol. 20, no. 02, 2021.
[14] N. H. Zulkifley, S. Abdul Rahman, N. H. Ubaidullah, and I. Ibrahim, "House price prediction using a machine learning model: a survey of literature," *International Journal of Modern Education and Computer Science*, vol. 12, no. 6, 2020.
[15] H. Zhang, K. Wang, M. Li, X. He, and R. Zhang, "House price prediction with an improved stack approach," *Journal of Physics: Conference Series*, vol. 1693, no. 1, 2020.
[16] Z. Wu, "Prediction of California house price based on multiple linear regression," *Academic Journal of Engineering and Technology Science*, vol. 3, no. 7.0, 2020.
[17] *UK House Price Index to return*, M2 Presswire, England, 2020.
[18] B. Nazemi and M. Rafiean, "Forecasting house prices in Iran using GMDH," *International Journal of Housing Markets and Analysis*, 2020.
[19] Y. Niu, "Linear regression model of house price in boston," *Science Discovery*, vol. 8, no. 3, 2020.
[20] R. Bala, K. Surya, T. Chandravas, and J Manikandan, "A machine learning based advanced house price prediction using logistic regression," *International Journal of Computer Application*, vol. 176, no. 28, 2020.
[21] M. Thamarai and S. P. Malarvizhi, "House price prediction modeling using machine learning[J]," *International Journal of Information Engineering and Electronic Business*, vol. 12, no. 2, 2020.
[22] A. N. Alfiyatin, R. E. Febrita, H. Taufiq, and W. F. Mahmudy, "Modeling house price prediction using regression analysis and particle swarm optimization case study: Malang, East Java, Indonesia," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 8, 2017.

Hindawi
Scientific Programming
Volume 2022, Article ID 8280036, 7 pages
https://doi.org/10.1155/2022/8280036

Research Article

Prediction of News Popularity Based on Deep Neural Network

Yan Cai[1] **and Zhiqiang Zheng**[2]

[1]*Shanghai University of Sport School of Media and Arts, Shanghai 200438, China*
[2]*Shanghai Dongfang Network Video Co., Ltd, Shanghai 200438, China*

Correspondence should be addressed to Yan Cai; caiyan@sus.edu.cn

Received 26 January 2022; Revised 5 March 2022; Accepted 7 March 2022; Published 31 March 2022

Academic Editor: Baiyuan Ding

Due to the timeliness and short life cycle of news, the postrelease prediction is limited, and the prerelease prediction also faces huge challenges due to the diversity and difficulty of defining influencing factors. This paper uses the foreclosure. This paper proposes a news popularity prediction method based on GRU deep neural network. Firstly, a web crawler was designed to obtain news data of different types and structures from 10 information security portal websites in China. After data preprocessing, the Word2Vec method was used to extract features, extract key news sentences and construct a subset of content features. Establish a GRU neural network regression prediction model to predict hot news on the Internet. The experimental results show that, compared with the traditional processing method, the model can process the multi-source rough data set in this paper and greatly reduce the prediction error. At the same time, because the threshold recurrent unit structure used in this paper is simpler than the long-short-term memory network structure, it can shorten the prediction time and improve the computing performance.

1. Introduction

The Internet is subtly changing the way of life of human beings, and has transformed from the original entertainment to an indispensable part of people's daily work and life [1]. News popularity prediction is an estimate of the number of page views or shares obtained after news is released. By predicting the number of news page views or shares, it can help content producers to better evaluate the quality of news in advance, and then help to rank news and recommended [2–4].

Popularity is one of the important characteristics of online news, it represents the spread of news and the possible social impact of news. By studying the relationship between online news and its popularity, it is possible to understand the reasons why online users forward and comment on news, and to dig out the key factors that determine news popularity, which is helpful for the reasonable release of online information and the grasp of public opinion trends [5, 6]. News with higher popularity tends to gain more attention from the public and become hot news, which has a greater impact on society and people's lives. News with less popularity tends to lack attention and have less impact. Major news websites release all kinds of news all the time. Due to the limited time and energy of the public, they cannot read every news. Only a very small number of news will attract widespread public attention and become hot news. Therefore, it is timely. It is necessary to predict which news will become hot [7, 8]. In addition, online hot news prediction has important application value; firstly, it can enable the government to grasp the trend of public opinion in a timely manner, which is convenient for the government to manage public opinion and grasp and handle sudden public events; secondly, it can help news websites manage the release locations of different news, Put hot news in the area that users pay more attention to, thereby increasing the influence of news websites; at the same time, it promotes the public to pay attention to the current hot news in a timely manner, and triggers thinking about daily life from the news, thereby improving the quality of life. For example, when hot news related to telecommunication fraud occupies the homepage of major fine-textured websites, it can increase people's attention to and beware of telecommunication fraud, and help people learn the relevant knowledge of

preventing telecommunication fraud [9–12]. Network news has become the main source of network waves and public opinion, and it is of great theoretical and applied value to accurately predict hot news and attract public attention and discussion [13–15].

With the development of Web 2.0 and various self-media and online social networks at home and abroad, it has greatly changed the way users generate and consume content. Whether it is for content consumers, or for companies, content providers and self-media content producers, online content is a valuable asset and a major attraction and competitiveness on the Internet. At the same time, user-generated content has exploded due to the ease and lower cost of content creation. For example, users around the world send more than 300,000 tweets on Twitter, share more than 680,000 pieces of content on Facebook, and upload 100 hours of video every minute [16]. In this case, determining the quality of the content will become very important. Online users overwhelmed with information can reduce clutter and focus on the information most relevant to them, helping content consumers focus on the most valuable resources in the online world. For content distributors, they can rely on the popularity prediction after content distribution to actively allocate resources according to the needs of future users. The news popularity prediction studied in this paper is to predict the number of pageviews or retweets that may be obtained after the news is released. It can help journalists to better evaluate the quality of news and rank news, so as to conduct news delivery more reasonably [17–20]. Online news popularity prediction is an extremely challenging task. Due to the richness and difficulty of mining text content, it is important for researchers to measure various factors that affect popularity (such as the quality of content or relevance to users) and how to choose an appropriate predictive model that can adapt to different datasets and It is difficult to fit the popularity value after news release more accurately [21–26].

Aiming at the problems in this field, this paper proposes a news popularity prediction model based on deep neural network, taking financial hot and nonhot news crawled from the Sohu News website as the research object, extracting features from the content of online news, and establishing a GRU neural network. The news popularity prediction regression model combined with the network and the fully connected layer, thereby shortening the prediction time and improving the computing performance.

2. Word Vector Representation of Text

2.1. Word Vector Representation. A word vector refers to a low-dimensional real number vector (usually 50-dimensional or 100-dimensional) to represent words. For example, the word "Xi'an" can be represented as a vector of the form [0.692, −0.877, −0.157, 0.119, −0.532, . . .], the text "I go to school in Xi'an" is represented by word segmentation as "I,", "zai," "Xi'an," and "going to school," where each word can be trained as a $1 * n$-dimensional tensor, and the subtext can be Represented as a $4 * n$-dimensional tensor.

Traditional text representations are based on bag-of-words models. The shortest answer is the one-hot model, which arranges all the words in the corpus in a column. For a certain document, the position of the words it contains in the corpus is k, then the k-th position is set to 1, for this document The words that are not included are set to 0; in order to show the importance of different words, the TF-IDF weight is used instead of 0 or 1, which is the traditional text space vector representation. The space vector model of text classification often results in the problem of too large text vector dimension and sparse space vector. The use of principal component analysis and singular value decomposition can achieve the purpose of reducing the dimensionality. However, in the process of dimensionality reduction, important information is often lost, such as Word order information. In addition, the spatial vector representation in this paper considers that any two words are completely independent, but adjacent words in text data are often associated. Using word vectors to represent text can not only solve the problem of high dimension and sparse vectors, thereby reducing the training difficulty of the model, but also word vectors can better reflect the semantic characteristics of words, reduce the distance between similar words or synonyms, and at the same time It can better represent the text content considering the order of words in the text.

2.2. Word2vec Theory. The vectorized representation of words includes One-Hot representation and word vector representation. The One-Hot representation represents each word as a long vector containing only 0s and 1s. The dimension of the vector is the size of the dictionary, and the position of 1 corresponds to the position of the word in the dictionary. There are some problems with this approach:

(1) Since the dimension of the vector is determined by the size of the dictionary. Then when the dictionary is large, it will cause the dimension of the vector to be too high and cause dimension disaster.

(2) When the vectorized representation of each word is performed, most of the vector values are 0, which will cause the word vector to be a high-dimensional sparse vector.

(3) In addition, the One-Hot method does not consider the semantic and contextual relationship between words.

To solve the above problems, Google in 2013 proposed a word vector representation that uses contextual relationships to map words to low-dimensional, dense vectors—Word2vec. Word2vec is a deep learning based tool released by Google in 2013. The tool mainly adopts two model architectures—CBOW model (Continuous Bag-of-Words) and Skip-Gram (Continuous Skip-Gram) model to learn the vector representation of words. The CBOW model predicts the word output at the current position based on the words in the window around the current word; Skip-Gram, on the contrary, predicts the word output around it based on the words at the current position. Converting the text dataset

into the input form accepted by the two models of Word2vec, CBOW and Skip-Gram, can get its corresponding word vector as the output. The basis of the vectorization of the two models of Word2vec is to construct a word list based on the training text data set, and learn the vector representation of the words in the word list. Word vectors can perform mathematical operations and are widely used in natural language processing and other related fields, such as calculating the similarity and semantic correlation between two words. The features learned through the generated word vectors can be used in fields such as text classification, named entity extraction, clustering, and sentiment analysis.

$$P\left(w_t | w_{t-k}, w_{t-(k-1)}, \ldots, w_{t-1}, w_{t+1}, w_{t+2}, \ldots, w_{t+k}\right). \quad (1)$$

The training goal of the CBOW model is to find the word that maximizes the probability shown in formula (1), where k is the window size, and the current word is predicted according to the context word.

2.3. Word Vector Training. Word2Vec is Google's open source natural language processing training tool. It represents all words in the corpus with vectors, so that the relationship between words can be quantitatively measured by distance measurement. Word2Vec uses the Distributed representation to replace the One-hot model. Through training, each word is mapped to a specified lower-dimensional vector, and then the relationship between words can be studied by ordinary statistical methods. The words in the document The use of vector representation can fully consider word semantics and word order. Word2Vec is often used for some problems in the field of natural language processing, such as machine translation, relation mining, part-of-speech tagging, etc. It is computationally efficient and can train a large number of corpora. Word2Vec generally uses a neural network without hidden layers for training, and the definitions of input and output include Skip-Gram and CBOW training models. The structure of the CBOW model is shown in Figure 1.

The input information of the CBOW model is the word vector corresponding to the context-related words of the target word in the vocabulary, and the output is the word vector of the target word. For example, the words "related," "network," "hot spots," "news," "of," "research," "has," "important," and "meaning," the target word is "news," and the word vector of the word is output; when the window size is 5, the context value is 2, then the word vectors of the first two words "network," "hot spot" and the last two words "de" and "research" adjacent to the target word are the input of the model; the training goal is to make The softmax probability of the target word in the training sample is the largest. In the CBOW neural network model corresponding to the above example, there are 4 neurons in the input layer, and the number of neurons in the output layer is equal to the number of words in the vocabulary. The parameters of the model can be obtained by the BP algorithm, and the word vectors corresponding to all the words in the vocabulary are obtained at the same time. In the figure, $w(t)$ represents the

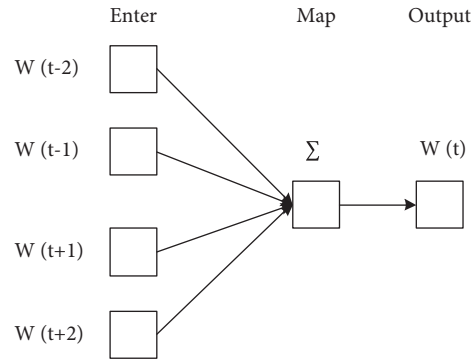


Figure 1: CBOW model structure.

position t of the target word in the sentence. In the window (the window size is 5), the upper and lower words except the target word together constitute the context.

3. GRU Neural Network

The LSTM network structure can solve the long-distance dependence problem of the traditional RNN network, but the internal structure of the LSTM is very complex, the training time is long, and there are many parameters. Therefore, various variants have appeared on the basis of LSTM. For example, in order to improve the training speed, the SRU (Simple Recurrent Unit, SRU) proposed by Lei et al. A widely popular variant of LSTM is that in 2014, CHOK et al. proposed a simplified structure of LSTM—GRU (Gated Recurrent Unit, GRU).

GRU and LSTM also use the gating mechanism, but compared with LSTM, GRU only retains two gating units reset gate (reset) and update gate (update). At the same time, the cell state and output gate are abandoned, and the cell state is combined with the hidden state. Compared with the LSTM structure, the update gate in the GRU plays the role of the input gate and the forgetting gate, controlling how much of the memory of the previous moment is retained at the current moment, that is, z_t in Figure 2 where $\widehat{h}_t$ is the memory state at the current moment, and h_t is the hidden state at the current moment.

Similar to the LSTM structure, the external input x_t at the current moment and the hidden state h_t at the previous moment jointly generate the input of the reset gate and the update gate. The update gate plays the role of "connecting the past and linking the future," it controls the discarding of the state at the previous moment and the information retained at the current moment; the reset gate controls the proportion of the information at the previous moment being forgotten. The cooperation of the reset gate and the update gate enables the learning of long-term memory. Since the GRU structure was proposed, it has been favored by researchers at home and abroad, and has been widely used in different fields. Various variants have also emerged based on their basic structure. The GRU structure has an advantage over the LSTM structure in terms of computational performance due to its simpler structure and fewer parameters, on the premise of ensuring the same effect as the LSTM.

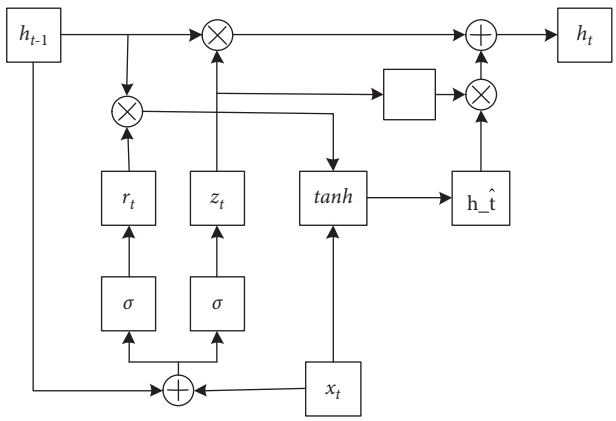


Figure 2: GRU network structure diagram.

The GRU structure not only overcomes the gradient dispersion problem existing in the traditional RNN structure and can learn long-distance dependencies, but also has a simpler structure and fewer parameters than LSTM. These advantages of GRU make it one of the widely used variants of the traditional RNN structure.

$$\widehat{h}_t = \tanh\left(W^x x_t + U\left(r \otimes h_{t-1}\right)\right). \tag{2}$$

4. News Popularity Prediction Model based on GRU Neural Network

The main process of the framework to achieve news popularity prediction is as follows:

(1) Data preprocessing: there is a lot of "dirty data" in the news data set obtained by designing web crawlers. Data preprocessing needs to be carried out first, which mainly includes deduplication of data and judgment of empty data. In the preprocessing of data It also includes word segmentation and stop word removal operations for news headlines to prepare for text feature extraction.

(2) Feature extraction process of surface information: after the data preprocessing in step (1), the news is decomposed into title, category, author, release time and text, and the custom features are derived from the title, category, author and release time. Sentiment polarity analysis and named entity extraction are performed on news headlines as "text features;" news headline lengths, publishing time conversions, author scores and category scores are calculated to construct a set of "metadata features."

(3) Key sentence extraction and content features: before using the key sentence extraction algorithm to extract key sentences from the news text, it is necessary to perform preprocessing of sentence segmentation, word segmentation and stop word removal. The preprocessed news uses the key sentence extraction algorithm to complete the extraction of key sentences. The named entity feature extraction of the key sentences of the news mainly includes the number, place name and organization name to construct a subset of content features.

(4) Feature set construction: after steps (2) and (3), the surface feature extraction and key sentence feature extraction process of news are completed, and text features, metadata features and content features are obtained respectively. The three feature subsets are connected and fused into a feature set, that is, the final feature set participating in model training.

(5) Model training: in this paper, the regression model is trained with the simplified structure GRU of LSTM, and the fully connected layer is added after the GRU network layer to output the prediction result. In order to prevent overfitting, a dropout layer is added to control the random deactivation of neural units and the L2 regular term during model training, and the ReLU function is used as the activation function. Take MSE as loss.

(6) Result output: based on this framework, the multi-feature extraction and fusion of news are realized, and the GRU structure is used to connect the fully connected layer to train a regression prediction model to predict news popularity.

4.1. Data Set. This paper describes the popularity prediction of news as a regression prediction problem, which predicts the number of pageviews received in the future after the news is published. Design a web crawler based on the crawler framework Scrapy, and use xpath to match by viewing the source code of the web page to obtain data and save it as a csv file. This article obtains some news from 2014 to 2018 through the crawler, and uses the number of news views as the prediction target of this article. In order to achieve faster convergence during training, logarithmically transform the news pageviews, and the transformed pageviews obey a Gaussian distribution.

4.2. Algorithm Evaluation Index. The research on the news popularity prediction framework in this paper belongs to regression prediction, so the MAE (Mean Absolute Error, MAE), MSE (Mean Squared Error, MSE), R^2 (R-squared, R^2) and RMSE (Root Mean Squared Error, RMSE) as metrics to measure the performance of the framework.

$$\begin{aligned}
\mathrm{MSE} &= \frac{1}{n}\sum_{i=1}^{n}\left(f\left(x_i\right) - y_i\right)^2,\\
\mathrm{MAE} &= \frac{1}{n}\sum_{i=1}^{n}\left|f\left(x_i\right) - y_i\right|,\\
\mathrm{RMSE} &= \sqrt{\frac{\sum_{i=1}^{n}\left(f\left(x_i\right) - y_i\right)^2}{n}},\\
R^2 &= 1 - \frac{\left(\sum \mathrm{Squared_Residuals}/N\right)}{\mathrm{Variance}_{Y_{\mathrm{true}}}}.
\end{aligned} \tag{3}$$

Among them, MAE, MSE, and RMSE indicators reflect the prediction error of the framework and are the main evaluation indicators; R^2 is the discriminant coefficient, which is used to indicate that the interpretability of the framework is a secondary indicator. Lower values of MAE, MSE and RMSE indicate smaller prediction errors of the framework.

5. Analysis of Results

The data set used in this article is the news data set of the security news portal website obtained through web crawlers. The experimental environment is the Python3 integrated environment Anaconda 3. The tools used mainly include the Python scientific computing package Numpy, Pandas, the word segmentation tool stuttering word segmentation and visualization Toolkit Matplotlib, Seaborn.

The construction of the model is based on Keras with Theano as the backend, performing 5 repeated experiments and taking the average result of the 5 experiments as the final result. At the same time, in order to more accurately evaluate the prediction effect of the framework during training, the data set is divided into training set, test set and validation set according to the ratio of 7 : 2 : 1. The experimental results shown in this article are the test set results. When building the GRU network, the Keras sequential model is used. In order to prevent overfitting, a Dropout layer is added, and the random deactivation ratio of neurons in the hidden layer is set to 0.5. At the same time, the L2 regularization is added to the GRU layer as a penalty term for model complexity; the ReLU function is used to calculate the model complexity. As an activation function, because the ReLU function does not involve a large number of exponential operations compared to the sigmoid or tanh function, the model converges faster. At the same time, ReLU can randomly inactivate some neurons, that is, "dead neurons," which can alleviate overfitting and There will be no gradient dispersion problems. The loss function used in this paper is MSE, stochastic gradient descent (SGD) is used as the optimization method, and the epoch is determined to be 6 by using grid search. Figure 3 shows the training process of the network. In the early stage of network training, with the increase of epoch, the loss function of the training set and the validation set decreases rapidly, indicating that the network learning speed is faster at this time; when the epoch continues to increase, the loss function decreases slowly, the network learning speed slows down; until epoch = 6, the loss function of the training set and the validation set basically does not change, and the network almost stops learning.

The results obtained by repeating the experiment three times and taking the average of the results are shown in Table 1.

The GRU and multi-feature fusion method proposed in this paper outperforms LSTM plus custom feature combination, ordinary linear regression-based, and random forest-based prediction methods in regression prediction, three indicators MSE, RMSE and MAE. The deep learning-based security news popularity prediction framework proposed in this paper shows low prediction error on rough and irregular

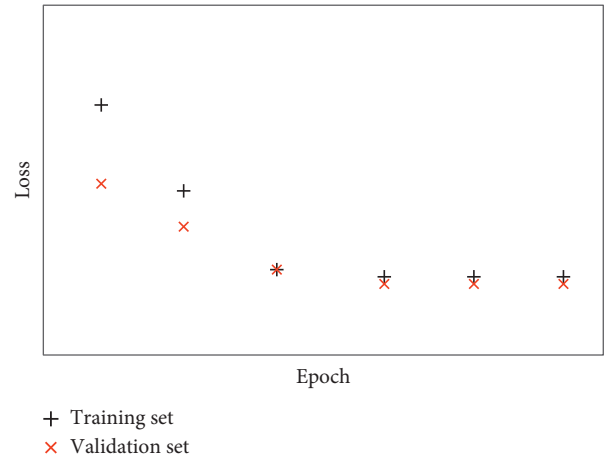


Figure 3: Model training loss comparison chart.

Table 1: Comparison of experimental results.

Model	MSE	MAE	RMSE	R^2
GRU	0.0152	0.0865	0.1098	0.6001
LSTM	0.2114	0.3352	0.4576	0.6298
Linear regression	0.2976	0.3621	0.5423	0.2651
Random forest	0.3521	0.4108	0.5901	0.1365

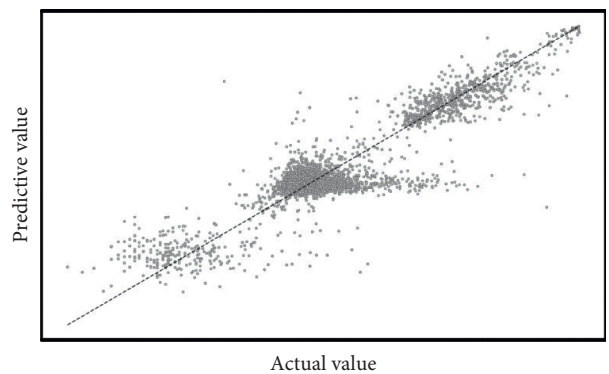


Figure 4: Model fitting effect diagram.

Table 2: Comparison of computing performance.

Index	GRU	LSTM
Time	10689	12351

datasets, while the prediction method of LSTM plus custom features does not achieve good results when dealing with rough multi-source datasets effect. At the same time, the discriminant coefficient R^2 of the prediction method proposed in this paper is 0.60349, which is slightly lower than the prediction method of LSTM plus custom features. R^2 is a secondary indicator, indicating that the model has better data fitting ability. Figure 4 is the fitting result of the predicted value and the actual value.

Compared with LSTM, GRU requires less training time due to its simpler structure and fewer parameters. Table 2 shows the comparison of computing performance using GRU structure and LSTM structure respectively. When

faced with larger-scale data sets, the advantage of GRU structure in computing performance will be more obvious.

6. Conclusion

News popularity prediction before news release can predict the number of clicks, comments or reposts of news in the future. Based on the prediction of popularity, news quality evaluation, news ranking, news recommendation and news retrieval can be carried out. At the same time, news popularity prediction plays an important role in alleviating the information explosion and information overload caused by the rapid development of the Internet and social media. Prerelease forecasting also faces huge challenges due to the diversity and difficulty of defining influencing factors. Compared with the existing work, this paper proposes a news popularity prediction framework that can process multi-source rough data sets and greatly reduce the prediction error by extracting multi-feature fusion and combining with neural network structure threshold recurrent unit training regression prediction model.

This paper proposes a news popularity prediction method based on GRU deep neural network. Firstly, a web crawler was designed to obtain news data of different types and structures from 10 information security portal websites in China. After data preprocessing, the Word2Vec method was used to extract features, extract key news sentences and construct a subset of content features. Establish a GRU neural network regression prediction model to predict hot news on the Internet. The experimental results show that, compared with the traditional processing method, the model can process the multi-source rough data set in this paper and greatly reduce the prediction error. At the same time, because the threshold recurrent unit structure used in this paper is simpler than the long-short-term memory network structure, it can shorten the prediction time and improve the computing performance.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] Y. Yang, Y. Liu, X. Lu, J. Xu, and F. Wang, "A named entity topic model for news popularity prediction," *Knowledge-Based Systems*, vol. 208, 2020.

[2] U. Douglas Zanatta, B. J. Luiz, M. C. Bonato, and S. Eusebio, "Exploring the effects of different clustering methods on a news recommender system," *Expert Systems with Applications*, p. 183, 2021.

[3] B. Ş. KemalAkyol, "Modeling and predicting of news popularity in social media sources," *Computers, Materials & Continua*, vol. 61, no. 1, 2019.

[4] A. S. Wicaksono and A. A Supianto, "Hyper parameter optimization using genetic algorithm on machine learning methods for online news popularity prediction," *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 12, 2018.

[5] T. T. Esther, J. A. Kumar, P. Narayanasamy, and J. Vidya, "Attention-based C-BiLSTM for fake news detection," *Applied Soft Computing Journal*, vol. 110, 2021.

[6] N. Moniz, L. Torgo, M. Eirinaki, and P. Branco, "A framework for recommendation of highly popular news lacking social feedback," *New Generation Computing*, vol. 35, no. 4, 2017.

[7] D. Chen, F. Qin, K. Cai, and Y. Shen, "Detecting and classifying typhoon information from Chinese news based on a neural network model," *Sustainability*, vol. 13, no. 13, 2021.

[8] P. Saigal and V. Khanna, "Multi-category news classification using Support Vector Machine based classifiers," *SN Applied Sciences*, vol. 2, no. 2, 2020.

[9] S. A. Kanekar, A. Sharma, G. S. Patkar, K. Amey, and S. Tilve, "Building semantically annotated corpus for text classification of indian defence news articles," *International Journal of Information Technology*, 2021.

[10] M. A. Mouriño-García, R. Pérez-Rodríguez, L. Anido-Rifón, and M. Vilares-Ferro, "Wikipedia-based hybrid document representation for textual news classification," in *Proceedings of the 2016 3rd International Conference on Soft Computing & Machine Intelligence (ISCMI)*, Dubai, UAE, November 2018.

[11] A. Abid, A. Ansar, K. Adel, F. Muhammad Shoaib, R. Iqbal, and F. Uzma, "An architectural framework for information integration using machine learning approaches for smart city security profiling," *International Journal of Distributed Sensor Networks*, vol. 16, no. 10, 2020.

[12] Y. Zhang and K. Lin, "Predicting and evaluating the online news popularity based on random forest," *Journal of Physics: Conference Series*, vol. 1994, no. 1, 2021.

[13] M. Naseri and H. Zamani, "Analyzing and Predicting News Popularity in an Instant Messaging Service," in *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, New York, NY, USA, July 2019.

[14] H. Young Baek, D.-K. Kim, and J. W. Kim, "Management earnings forecasts and adverse selection cost: good vs. bad news forecast," *International Journal of Accounting and Information Management*, vol. 16, no. 1, 2008.

[15] P. Kittisak and V. Peerapon, "Stock trend prediction using deep learning approach on technical indicator and industrial specific information," *Information*, vol. 12, no. 6, 2021.

[16] J. L. O. Hui, G. K. Hoon, and W. M. N. W. Zainon, "Effects of word class and text position in sentiment-based news classification," *Procedia Computer Science*, vol. 124, 2017.

[17] L. N. H. Nam and H. B. Quoc, "The hybrid filter feature selection methods for improving high-dimensional text categorization," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 25, no. 2, 2017.

[18] A. K. Tiwari, *Deep Learning and its Applications*, Nova Science Publishers, Inc., Suite N Hauppauge, NY, USA, 2021.

[19] S. Desai Prathamesh, "News sentiment informed time-series analyzing AI (SITALA) to curb the spread of COVID-19 in houston," *Expert Systems with Applications*, vol. 15, p. 180, 2021.

[20] F. Majeed, A. Muhammad Waqas, A. Hassan Muhammad, S. Ali, and M. I. Lali, "Social media news classification in healthcare communication," *Journal of Medical Imaging and Health Informatics*, vol. 9, no. 6, 2019.

[21] Z. Zhang, G. Zhang, and C. Zhao, "PeikeWang. Online Chinese news classification system based on Python reptile and convolutional neural network," *International Journal of*

Intelligent Information and Management Science, vol. 7, no. 6, 2018.

[22] D. Mehta, D. Aniket, P. Arunabha, and M Anand Kumar, "A transformer-based architecture for fake news classification," *Social Network Analysis and Mining*, vol. 11, no. 1, 2021.

[23] D. Dangol, R. Dahi Shrestha, and A. Timalsina, "Automated news classification using N-gram model and key features of Nepali language," *Scitech Nepal*, vol. 13, no. 1, 2018.

[24] K. S. Kyaw and S. Limsiroratana, "Optimization of text feature selection process based on advanced searching for news classification," *International Journal of Swarm Intelligence Research*, vol. 11, no. 4, 2020.

[25] Y. Zhang, Y. Dang, H. Chen, M. Thurmond, and C. Larson, "Automatic online news monitoring and classification for syndromic surveillance," *Decision Support Systems*, vol. 47, no. 4, 2009.

[26] Z. Wang and B. Song, "Research on Hot News Classification Algorithm Based on Deep learning," in *Proceedings of the 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, pp. 2376–2380, Chengdu, China, March 2019.

Hindawi
Scientific Programming
Volume 2022, Article ID 1126020, 10 pages
https://doi.org/10.1155/2022/1126020

Research Article

Design of Automatic Motion Capture Algorithm for Yao's Long Drum Dance Based on Multieye Machine Vision

Yan Meng

Wuzhou College Teacher Education College Dance Teaching and Research Section, Wuzhou, Guangxi 543002, China

Correspondence should be addressed to Yan Meng; my@gxuwz.edu.cn

Received 9 December 2021; Revised 30 December 2021; Accepted 15 January 2022; Published 27 March 2022

Academic Editor: Baiyuan Ding

Aiming at the problems of low accuracy and long time consumption in the existing dance motion automatic capture methods, this study designs an automatic capture algorithm of Yao's long drum dance based on multieye machine vision. The images of Yao's long drum dance are collected by the multieye machine vision measurement system, and the image signals under different lighting conditions are processed; on this basis, by determining the dance action image threshold and analyzing the change law of action image signal, the dance action image extraction of Yao's long drum is completed. By converting the dance action images from different angles, the Euler angle of the images from different angles is determined and fused. On this basis, the observation items of the observation part of the Yao long drum dance action image are set, and the a priori conditions and a posteriori probability of the Yao long drum dance action image preprocessing are determined, so as to obtain the best result of the dance action image, and dance action image preprocessing is completed. According to the preprocessed image, the capture algorithm of area range is designed, the remaining areas to be captured are determined by means of classification, and the automatic capture algorithm design of Yao's long drum dance is completed. The experimental results show that the motion image captured by the automatic motion capture algorithm of Yao's long drum dance based on multieye machine vision has high accuracy, and the capture time is long, which is feasible.

1. Introduction

Yao's long drum dance is an important cultural inheritance in China. Its dance movements reflect the integrity and historical flavor of the Chinese nation. Therefore, it is very important to inherit and innovate this dance [1]. This dance movement has its special national flavor. Its movement is quite different from the general dance movement, and its movement details and difficulty are relatively large [2]. In the training of this dance, the effective identification and capture of its movement will contribute to the rapid improvement in its dance training [3]. Therefore, relevant researchers have done a lot of research on the automatic capture of dance movements and achieved some results.

Reference [4] proposed a human motion image recognition method under high-intensity motion. This method is applied to the action capture of Yao's long drum dance. In the research of this method, the threshold of the action image is calculated by double convolution, and the action characteristics are determined through the calculation of the value. On this basis, the key targets of the action are determined combined with the Gaussian distribution, the background of the image is processed, and the recognition model is constructed by the Bayesian classification algorithm to complete the recognition of the action. This method can track the action well, but the analysis accuracy of the changes in the process of action transformation is poor, which needs to be further improved. Reference [5] proposed a research on aerobic decomposition action image recognition based on feature extraction. In this method, the image is divided into multiple segments through the pyramid, and the noisy image is effectively preprocessed, and the action is analyzed at different heights. Different optical flows in the image are solved by Laplace, and the similarity points in the image are effectively fused. Finally, the action is recognized according to the set threshold. In the research of this method, the

threshold calculation is more accurate, but the operation process is more complex and has some limitations, which need further improvement. In reference [6], a motion recognition method based on global spatiotemporal feature convolution neural network is proposed. The multiframe fusion method is used to improve the accuracy and deeply learn the moving image in the global time stream. Finally, the two streams are combined to identify human motion. However, this method is only suitable for small sample data, resulting in insufficient practical application performance of this method.

To solve some shortcomings of the existing methods, this study designs an automatic capture algorithm of Yao's long drum dance based on multieye machine vision to improve the automatic capture of the dance project. Machine vision refers to the use of machines instead of manual visual inspection to make measurement and judgment. Multicamera machine vision system refers to collecting images through multiple machine vision products and then transmitting the images to the processing unit. Through digital processing, it can judge the size, shape, and color according to the pixel distribution, brightness, color, and other information. Then, the on-site equipment action is controlled according to the discrimination result. With the development of computer technology and fieldbus technology, machine vision technology is becoming more and more mature. It has become an indispensable product in modern processing and manufacturing industry.

The images of Yao's long drum dance are collected by the multieye machine vision measurement system, and the image signals under different lighting conditions are processed; on this basis, by determining the dance action image threshold and analyzing the change law of action image signal, the dance action image extraction of Yao's long drum is completed. By converting the dance action images from different angles, the Euler angle of the images from different angles is determined and fused. On this basis, the observation items of the observation part of the Yao long drum dance action image are set, and the a priori conditions and a posteriori probability of the Yao long drum dance action image preprocessing are determined, to obtain the best result of the dance action image, and dance action image preprocessing is completed. According to the preprocessed image, the capture algorithm of area range is designed, and the remaining areas to be captured are determined by means of classification, and the automatic capture algorithm design of Yao's long drum dance is completed. The main technical route of this study is as follows:

(1) The images of Yao's long drum dance are collected by the multieye machine vision measurement system, and the image signals under different lighting conditions are processed;

(2) On this basis, by determining the dance action image threshold and analyzing the change law of action image signal, the dance action image extraction of Yao's long drum is completed. By converting the dance action images from different angles, the Euler angle of the images from different angles is determined and fused. On this basis, the observation items of the observation part of the Yao long drum dance action image are set, and the a priori conditions and a posteriori probability of the Yao long drum dance action image preprocessing are determined, to obtain the best result of the dance action image, and dance action image preprocessing is completed.

(3) According to the preprocessed image, the capture algorithm of area range is designed, and the remaining areas to be captured are determined by means of classification, and the automatic capture algorithm design of Yao's long drum dance is completed.

(4) Experimental analysis is given.

(5) Conclusion is presented finally.

2. Yao's Long Drum Dance Action Image Extraction and Preprocessing

2.1. Motion Image Extraction of Yao's Long Drum Dance Based on Multieye Machine Vision. To improve the effectiveness of the Yao long drum dance action automatic capture algorithm, it is necessary to extract the Yao long drum dance action image. To extract the Yao long drum dance action in more detail, this study first extracts the Yao long drum dance action with the help of multieye machine vision. The pipeline measurement method based on multivision uses the camera to capture the two-dimensional gray images of Yao's long drum dance at different angles [7], obtains the pipeline centerline through feature extraction, feature fitting, and other technologies, and reconstructs the three-dimensional space points of the centerline using the stereovision reconstruction method, to realize the reconstruction of Yao's long drum dance. When the measurement system is calibrated well, the processing result of Yao's long drum dance image is an important factor affecting Yao's long drum dance, and obtaining high-quality Yao's long drum dance image is a powerful guarantee for the successful processing of pipeline image [8].

This study analyzes the lighting mode adopted by the multivision system for the extraction of Yao's long drum dance. On this basis, Yao's long drum dance image is preprocessed to eliminate the influence of optical fiber on Yao's long drum dance, to improve the quality of Yao's long drum dance image to be processed [9].

In machine vision measurement system, the light source is an important part. The light source projects the light onto the surface of the measured object in an appropriate way and highlights the contrast of the features to be measured and is captured by the camera, to realize the extraction of the features to be measured. The selection of light source lighting mode has a great impact on the measurement system. A good lighting mode can improve the resolution of the whole measurement system and highlight the characteristics of the object to be measured, to reduce the pressure of subsequent image processing. If the lighting mode is improperly selected, it will have a certain adverse impact on the

subsequent processing of the measurement system. Therefore, it is necessary to select different lighting methods for different measurement objects [10].

In the Yao long drum dance action image extraction, the multivision light source and the camera are located on the same side of the dancer to be tested. The camera distinguishes the features to be tested and irrelevant features by receiving the reflected light of the Yao long drum dance action [11], as shown in Figure 1.

In the action image of Yao's long drum dance collected by the multieye machine vision system, the detailed information of the action is obtained through the reflection of the light source [12]. The action image obtained by the multicamera machine vision centralized camera is carried out through the Lambert body reflection model. The light intensity in the action image can be expressed as follows:

$$G(I) = a_I \times S, \tag{1}$$

where a_I represents the action attribute vector of point I and S represents the parameter value of the multiocular machine visual illumination attribute .

When the gray level of the collected Yao long drum dance action image changes to a certain extent, the gray level of the Yao long drum dance action image in the Lambert body reflection model [13] is expressed as follows:

$$E(I) = \mathrm{MAX}(a_I \times S, O). \tag{2}$$

Among them, O represents the grayscale value of the Yao long drum dance action images.

According to the determined gray level of Yao's long drum dance action image, the signal in Yao's long drum dance action image fluctuates greatly from a natural action of Yao's long drum dance action to the contraction state [14]. In this study, the signal in the stable Yao's long drum dance action image is used as Yao's long drum dance action image data collected in this study [15].

In the signal in the Yao long drum dance action image, the maximum value of the image signal is set as $\mathrm{MAX}_t\{x_i\}$ and the average value is x, and the key threshold [3] of the acquired image can be obtained as follows:

$$H = \begin{cases} \dfrac{3}{v}\displaystyle\sum_{v=1}^{L}\left|x_i\right| \mathrm{MAX}_t\{x_i\} > \dfrac{3}{v} \\ \\ \mathrm{MAX}_t\{x_i\}/3, \text{other.} \end{cases} \tag{3}$$

Among them, v represents the discrete Yao long drum dance action image signal, t represents the length of the Yao long drum dance action image signal, and $\{x_i\}$ represents the range of values.

The Yao long drum dance action image signal is obtained. In the process from the signal in the intermediate state to stability, the real Yao long drum dance action image signal value [16] is as follows:

$$D(a,b) = u\frac{1}{k}\sum_{i=1}^{t}\left[\frac{1}{n}\sum H\right]^2. \tag{4}$$

Among them, $D(a,b)$ represents the signal of the dance action at the actual moment, u represents the number of channels sampled, and k represents the wavelength of the signal.

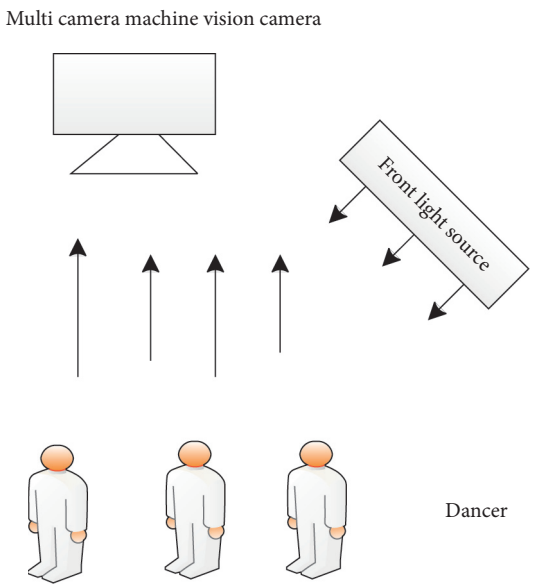


Figure 1: Schematic diagram of front light source of multicamera machine vision and camera on the same side of the dancer to be tested.

The initial to final changes in Yao's long drum dance action image signal are shown in Figure 2.

In the Yao long drum dance movement image extraction, the Yao long drum dance movement image is collected by the multieye machine vision measurement system, and the image signals under different lighting conditions are processed; on this basis, by determining the dance action image threshold and analyzing the change law of action image signal, the dance action image extraction of Yao's long drum is completed.

2.2. Image Preprocessing of Yao's Long Drum Dance. On the basis of the above collected Yao long drum dance action image extraction, due to the change and variability of action angle in the Yao long drum dance action image, there is a certain noise in the Yao long drum dance action image. Therefore, this study carries out multiangle preprocessing in the image preprocessing of Yao's long drum dance [17]. For any reference system of Yao's long drum dance action image, its orientation is determined by the rotation order of three Euler angles relative to the reference system. First, the coordinate system of the Yao long drum dance action image is rotated around the z-axis by 20, then rotated around the y-axis by 30, and finally rotated around the x-axis by 40°, and finally, the state is obtained after the coordinate system is combined with the original reference system, as shown in Figure 3.

In the image preprocessing of Yao's long drum dance, the whole process is regarded as the rotation of a bone joint in the movement process; that is, the rotation process in Yao's long drum dance is decomposed into three angles.

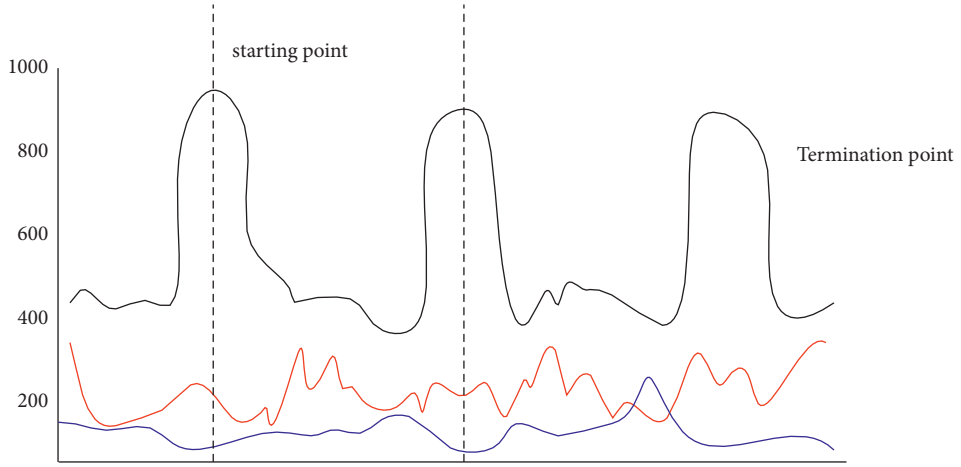


Figure 2: Change characteristics of Yao's long drum dance action image signal.

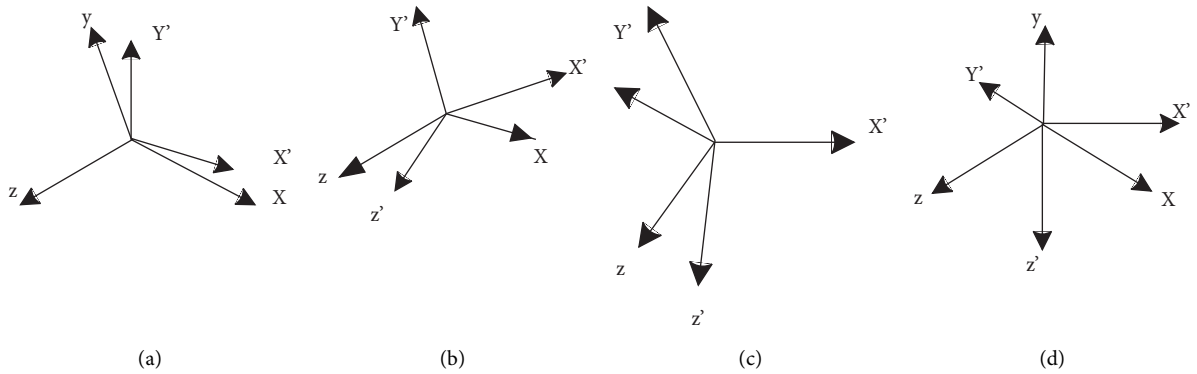


Figure 3: Rotation diagram of Yao's long drum dance action image in different states.

In different coordinate systems of Yao's long drum dance action images, the Euler angle is the key to measure its action amplitude. This angle represents the rotation degree of three coordinate axes. In this dance action, one rotation is transformed into the rotation angle of different coordinates [18].

To represent the angle of the Yao long drum dance movement image, the angle $q(\theta)$ is set after the rotation around the z-axis, $y(\alpha)$ represents the angle after the Y movement, and $x(\beta)$ represents the rotation after the horizontal axis. At this time, the Euler angle of the Yao long drum dance action images is expressed as follows:

$$q(\theta) = \begin{bmatrix} \cos\theta \ \sin\ \theta, 0 \\ -\sin\ \theta \ \cos\ \theta, 0 \\ 0, 0, 1 \end{bmatrix},$$

$$y(\alpha) = \begin{bmatrix} 10, 0 \\ 0, co\alpha \ \sin\ \alpha, 0 \\ 0, -\sin\ \alpha, \ \cos\ \alpha \end{bmatrix}, \quad (5)$$

$$x(\beta) = \begin{bmatrix} \cos\ \beta \ \sin\ \beta, 0 \\ -\sin\ \beta \ \cos\ \beta, 0 \\ 0, 0, 1 \end{bmatrix}.$$

After determining the above different angles, basically the angle after rotating the Yao long drum dance action image is rotated. At this time, the obtained angle is as follows:

$$P = \begin{bmatrix} \cos\ \theta \ \sin\ \theta, 0 \\ -\sin\ \theta \ \cos\ \theta, 0 \\ 0, 0, 1 \end{bmatrix} \begin{bmatrix} 10, 0 \\ 0, co\alpha \ \sin\ \alpha, 0 \\ 0, -\sin\ \alpha, \ \cos\ \alpha \end{bmatrix} \begin{bmatrix} \cos\ \beta \ \sin\ \beta, 0 \\ -\sin\ \beta \ \cos\ \beta, 0 \\ 0, 0, 1 \end{bmatrix}. \quad (6)$$

According to the motion data of the Yao long drum dance action image, it can be described by a rotation matrix, which effectively avoids the different final orientations of the joints due to different rotation sequences. Although the Euler angle can accurately describe the movement image of Yao's long drum dance, a large number of matrix operations will be involved in the process of the Euler angle processing, which will undoubtedly consume system space and time. At the same time, objects rotating around three angles in three-dimensional space will lose the autonomy of any axis due to rotation [19]. Therefore, to reduce the complexity of image preprocessing of Yao's long drum dance, further preprocessing is required.

According to the change in the rotation angle of the Yao long drum dance action image, the three-dimensional action

of the rotated action image is set in a known space, and a synchronous observation image of the Yao long drum dance action image is obtained. At this time, the observation item [6] of the observation part of the corresponding Yao long drum dance action image is expressed as follows:

$$\phi(A,B) = \aleph \prod_{i=1}^{n} \phi\left(F_k, v_k\right). \tag{7}$$

Among them, $\phi(A,B)$ represents the observation site of the action image of Yao's long drum dance, (F_k, v_k) represents the action decomposition value, and $\aleph$ represents the key noise value [7].

Setting table: the independence of Yao's long drum dance action images exists:

$$\varepsilon(A,B) = \tau \prod_{K=1}^{M} \varepsilon(A,B)\sigma. \tag{8}$$

Among them, τ represents the observation input value of Yao's long drum dance action under different visual angles, and σ represents the projection value of Yao's long drum dance action node position under machine vision.

Considering the a priori constraint of Yao's long drum dance action image [20], this is also the key to image preprocessing. Therefore, in the Yao long drum dance action image, the action image reflects the length of the action in the image through a graph structure model. At this time, the set constraint conditions [21] are as follows:

$$\rho = \gamma \prod_{(V_J, V_I)}^{M} \varepsilon(A,B)\mu. \tag{9}$$

Among them, γ is the relationship function representing the Yao long drum dance action image, and μ represents the length of the image action.

According to the determined a priori conditions for the image preprocessing of Yao's long drum dance action, the posture of dance action in the image will be analyzed by a posteriori probability from different perspectives [22], and the following results are obtained:

$$H(g) = \prod_{V_J, V_I}^{n} \rho \prod_{q} p\left(\vartheta | c_i\right), \tag{10}$$

Among these, $H(g)$ represents the postural posterior probability of the dance action, and c_i represents the conversion rate of the action at different angles.

Finally, the best result obtained from the preprocessed dance action image is as follows:

$$S^* = \text{Targmax} p \prod_{i=1}^{n} \prod \prod B\left(\vartheta | c_i\right). \tag{11}$$

Among them, T is the lowest equivalence problem representing the dance action image, and $B(\vartheta|c_i)$ represents the preprocessed dance action image.

In the dance action image preprocessing, it is preferred to convert the dance action images at different angles, determine the Euler angle of the images at different angles, and fuse them. On this basis, the observation items of the observation position of the Yao long drum dance action image are set, and the a priori condition and a posteriori probability of the Yao long drum dance action image preprocessing are determined, the best result obtained from the dance action image [23]isacquired, and the dance action image preprocessing is completed.

2.3. Design of Automatic Capture Algorithm for Yao's Long Drum Dance. Based on the above preprocessed Yao long drum dance action images, an automatic capture algorithm is designed to realize the automatic capture of Yao's long drum dance action and provide help to improve the accuracy of Yao's long drum dance action [24].

The area of Yao's long drum dance action image is large, and there is a certain gap. In the design of automatic capture algorithm in this study, firstly, the local action is effectively captured, and on this basis, the global over diffusion research is carried out. In this study, the local region capture method of window is studied. The basic process of local area capture of Yao long drum dance action image is as follows:

Step 1. Determining the cost function setting of the local area of the Yao long drum dance action image and determining its change law.

Step 2. Matching the aggregate value of the cost in the window according to the determined cost function [25].

Step 3. Finding the optimal value of the local area of the Yao long drum dance action image [26].

Step 4. Performing correction calculation in the local area of the dance action image of the Yao long drum.

It is determined that the image on one side of the regional range of the Yao long drum dance action image is the control object, and the image on the other side is regarded as the key value in the region. The cost function in the regional range of the Yao long drum dance action image obtained after the search is as follows:

$$\text{COST}_i = \left[R_i\left(d_i, d_j\right)\right] + \left[R_j\left(d_i, d_j\right)\right]^2. \tag{12}$$

Among them, R_i represents the pixels to be matched within the regional range of the Yao long drum dance action images, and R_i represents the parallax values.

W_i is set to change the central point within the area range, the image range window is built, and the cost is effectively unified, and we get the following:

$$W_i = \frac{1}{k \times j} \left\{ \sum_{j=1}^{n} \left[R_i\left(d_i, d_j\right)\right] + \left[R_j\left(d_i, d_j\right)\right]^2 \right\}. \tag{13}$$

According to the above determined image area range, the value of the point of the unified parallax value is determined, that is, the capture of the action within the area range of the Yao long drum dance action image, and the following is obtained:

$$d_i = \operatorname{argmin}\left\{y\left[R_j\left(d_i, d_j\right)\right]^2, \left(d_i, d_j\right), m\right\}. \quad (14)$$

Among them, y represents the capture results of the movements within the regional range of the Yao long drum dance action images, and m indicates the range.

After the movement within the region of the above Yao long drum dance movement image is automatically captured, the overall Yao long drum dance movement needs to be automatically captured [22, 27–29]. The remaining areas that cannot be captured directly and automatically are marked, the areas where different actions are located are determined, and then the overall action capture research is realized. The schematic diagram of remaining range division is shown in Figure 4.

Firstly, the remaining areas of the automatic capture of Yao's long drum dance action are labeled, and the area where the action is located is determined. The following results are obtained:

$$U_i = \left\{x | \exists_{ij} > 0, i = 1, 2, 3.... j \neq i\right\}. \quad (15)$$

Among them, $\exists_{ij}$ represents different labels.

Assuming that the Yao long drum dance action points belong E_i, it is classified and [19, 30, 31] is divided into the area with the largest number by voting:

$$E_i(x) = i = 1 \sum_{i=1}^{n} sig\left(E_i(x)\right). \quad (16)$$

On this basis, the track points of Yao's long drum dance in the remaining area are captured.

The running track of its action to L is set, and the track point is as follows: $o_1, o_2, ..., o_L$, and the capture motion feature point in the remaining area is as follows:

$$S(D) = \sum_{i=1}^{1} \sqrt[2]{s_{i+1}^2 - s^2}. \quad (17)$$

The captured motion feature points are in the remaining regions of the generated table, and s represents the midline pixels.

The overall flow of the design of the Yao long drum dance motion automatic capture algorithm is shown in Figure 5.

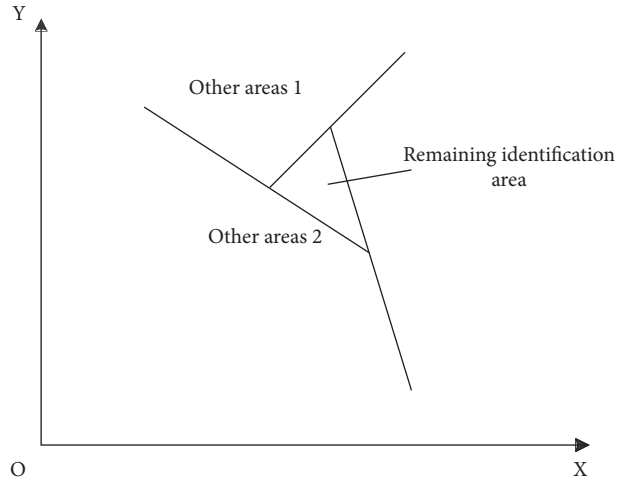


Figure 4: Schematic diagram of remaining range division.

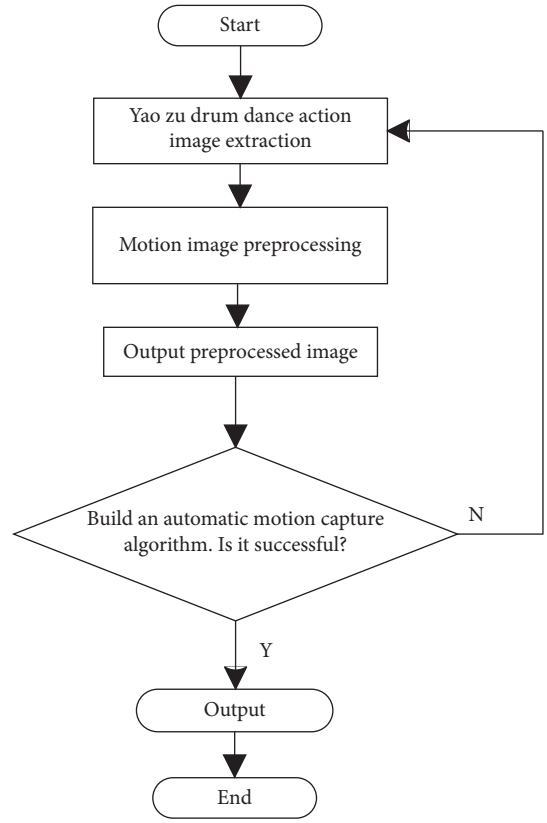


Figure 5: Overall flow chart of Yao's long drum dance motion automatic capture algorithm design.

3. Experimental Verification

To verify the practical application effect of the proposed automatic capture method of Yao's long drum dance based on multieye machine vision, a comparative verification experiment is carried out.

3.1. Experimental Preparation. To verify the effectiveness of the capture method in this study, the MySQL database is taken as the research object in the experiment, and the Yao long drum dance action images are selected as the research object. A total of 100 sample images are selected to train the image training set, and the most typical images are selected for research and analysis. The size of the image is 256 * 256, and the pixels are large. The sample image is shown in Figure 6.

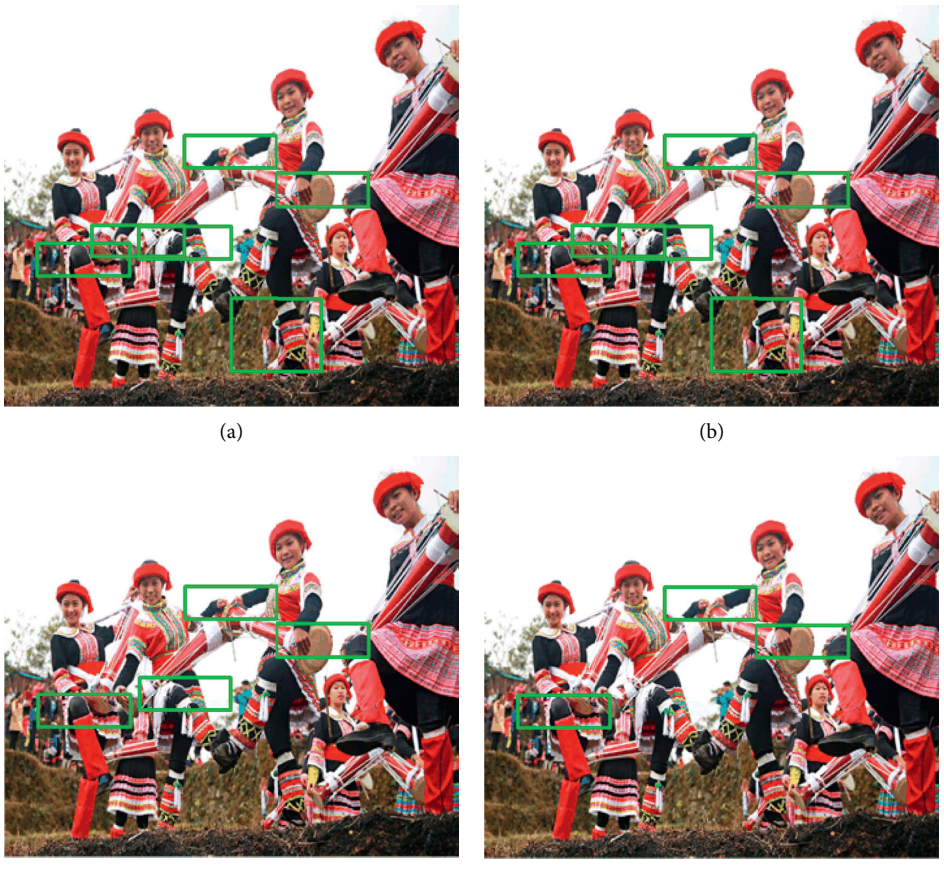
The experimental environment parameters are shown in Table 1.

3.2. Experimental Index. In the experiment, the accuracy and time consumption of sample image action capture are taken as the experimental indicators. The higher the

Figure 6: Experimental sample image.

Table 1: Experimental environment parameters.

Name	Parameter
Computer system	Windows 10
Development tools	C++
Processor	Intel (R)Core(TM)i7-4710HQ CPU @ 2.50 GHz
Memory size	8 GB
Simulation software	MATLAB 7.2



(a) (b)

(c) (d)

Figure 7: Analysis of sample image action recognition effect of different methods.

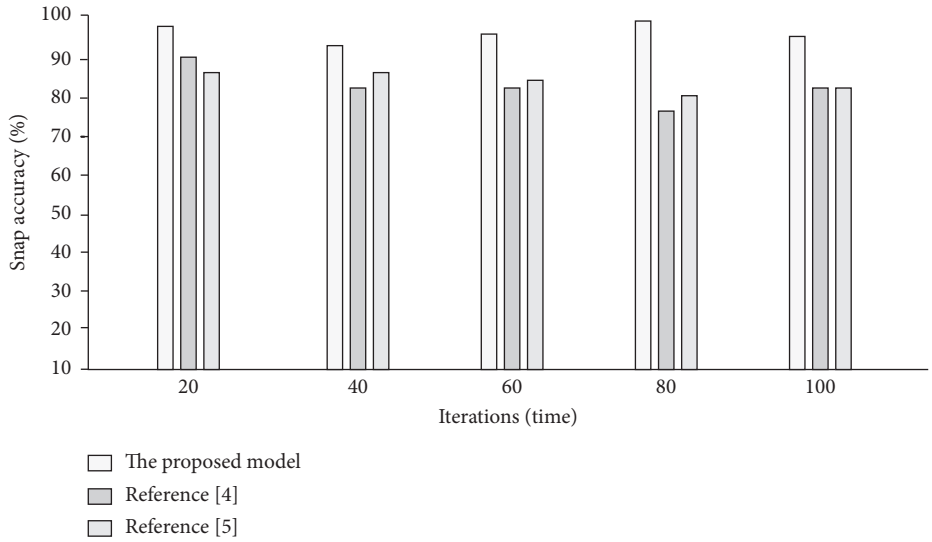


Figure 8: Analysis of accuracy results of sample image action recognition by different methods.

Table 2: Time consumption analysis of Yao's long drum dance movement image capture by different methods.

Capture times	The method of this study	Reference [4] method	Reference [5] method
20	1.2	2.3	2.4
40	1.3	2.5	2.6
60	1.2	2.8	2.8
80	1.3	3.0	3.1
100	1.2	3.2	3.3

accuracy of sample image action capture in the experiment, the better the effectiveness of the representative method.

3.3. Analysis of Experimental Results. In the experiment, the accuracy of sample image action recognition by this method, reference [4] method, and reference [5] method is compared. The recognition effect is shown in Figure 7.

By analyzing the experimental results in Figure 7, it can be seen that there are some differences in the effect of sample image action recognition using the methods in this study, literature [4], and literature [5]. Among them, the action identified by this method is consistent with the ideal value. There are some deficiencies in the identification of reference [4] method and reference [5], which is better than that of this method. The reason for the above experimental results is that this method determines the Euler angle of the images from different angles by converting the dance action images from different angles and fuses them. On this basis, it sets the observation items of the observation part of the Yao long drum dance action image and determines the a priori condition and a posteriori probability of the Yao long drum dance action image preprocessing, the best result of dance action image is obtained, and the action recognition effect of your method in this study is better.

The accuracy results of sample image motion recognition by this method, reference [4] method, and reference [5] method are shown in Figure 8.

By analyzing the experimental results in Figure 8, it can be seen that there are some differences in the motion recognition accuracy of sample images using the methods in this study, reference [4], and reference [5]. Among them, the recognition accuracy of this method is higher and always higher than 90%, while the capture accuracy of the other two methods is lower. This is because this method collects the Yao long drum dance action image through the multieye machine vision measurement system and processes the image signal under different illumination conditions; on this basis, by determining the dance action image threshold and analyzing the change law of action image signal, the Yao long drum dance action image extraction is completed, and then, the research accuracy is improved.

While ensuring the accuracy of sample image motion capture, the time consumption of sample image motion recognition is experimentally analyzed. The results are shown in Table 2.

By analyzing the experimental results, it can be seen that the three methods have different results in recognition time. With the continuous change in iteration times, the maximum time consumption of this method is about 1.3s, while the time consumption of the other two methods is always higher than that of this method. This is because this study designs the area range capture algorithm according to the preprocessed image, determines the remaining areas to be captured with the help of classification, completes the design of the automatic capture algorithm of Yao's long drum dance, and increases capture speed.

4. Conclusion

To improve the accuracy of dance motion capture, an automatic capture algorithm of Yao's long drum dance motion based on multieye machine vision is designed in this study. The performance of the algorithm is verified from both theoretical and experimental aspects. This method has high motion recognition accuracy and short recognition time when automatically capturing the movements of Yao's long drum dance. In particular, compared with the method using double convolution to capture, the motion recognition accuracy of this method is significantly improved, which is always higher than 90%; compared with the capture method based on feature extraction, the capture time of this method is significantly reduced, and the longest time is 1.3s. Therefore, it fully shows that the proposed capture algorithm based on multieye machine vision can better meet the requirements of automatic dance motion capture.

Data Availability

The raw data supporting the conclusions of this article will be made available by the author, without undue reservation.

Conflicts of Interest

The author declares that there are no conflicts of interest regarding this work.

References

[1] S. Chen, W. Wei, B. He, S. Chen, and J. Liu, "Action recognition based on improved deep convolutional neural network," *Application Research of Computers*, vol. 36, no. 3, pp. 945–949+953, 2019.

[2] E.-q Chen, M.-y Zhen, and J.-k Duan, "Human action recognition based on ResNeXt," *Journal of Graphics*, vol. 41, no. 2, pp. 277–282, 2020.

[3] L. Jiang, L. Cong, and L. Che, "Human motion recognition using ultra-wide band radar based on two-dimensional wavelet packet decomposition," *Journal of Electronic Measurement and Instrument*, vol. 32, no. 8, pp. 69–75, 2018.

[4] H. Zhang, "Research on human motion image recognition method under high intensity motion," *Computer Simulation*, vol. 36, no. 9, pp. 469–472, 2019.

[5] Fu-xiang, "Adaptive recognition method of Aerobics decomposition action image based on feature extraction," *Science Technology and Engineering*, vol. 19, no. 7, pp. 148–153, 2019.

[6] K. Wang, J. Wu, Z. Tianxiang, and R. Li, "An action recognition method based on global spatial-temporal feature convolutional neural networks," *Journal of Huazhong University of Science and Technology (Nature Science Edition)*, vol. 46, no. 12, pp. 36–41, 2018.

[7] R. Ding and X. Li-min, "Human action recognition using multi-frequency analysis of critical point trajectories," *Computer Engineering and Design*, vol. 38, no. 9, pp. 2546–2550, 2017.

[8] Z. Chen and W. Yue, "Spatio-temporal two-stream human action recognition model based on video deep learning," *Journal of Computer Applications*, vol. 38, no. 3, pp. 895–899+915, 2018.

[9] Z. Song, Y. Zhou, J. Jia, S. Xin, and Y. Liu, "Local feature fusion temporal convolutional network for human action recognition," *Journal of Computer-Aided Design & Computer Graphics*, vol. 32, no. 3, pp. 418–424, 2020.

[10] F. Cai and Z. Li, "Human action recognition based on coupled multi-Hidden Markov model and depth image data," *Journal of Computer Applications*, vol. 38, no. 2, pp. 454–457, 2018.

[11] Z. Li and L. Huang, "Hand-motion recognition based on improved BP neural network," *CAAI Transactions on Intelligent Systems*, vol. 13, no. 5, pp. 848–854, 2018.

[12] Li Qing-hui, Li Ai-hua, Zheng Yong, and Fang Hao, "Action recognition using geometric features and recurrent temporal attention network," *Optics and Precision Engineering*, vol. 26, no. 10, pp. 2584–2591, 2018.

[13] M. Zhi, "Human action recognition based on convolutional neural networks," *Computer Engineering and Design*, vol. 40, no. 4, pp. 1161–1166, 2019.

[14] F. Yun-lu and R. Miao, "Motion recognition algorithm based on time warping coupled linear discriminate analysis," *Computer Engineering and Design*, vol. 39, no. 11, pp. 3468–3474+3499, 2018.

[15] X. Yun-jiang and H. Xiong-bo, "Recognition method of human activity using double-layer classification model," *Computer Engineering and Design*, vol. 39, no. 12, pp. 3860–3866, 2018.

[16] J. Yin, J. Wei, L. Wang, and Y. Wu, "Human action recognition based on large margin nearest neighbor," *Robot*, vol. 40, no. 2, pp. 178–187, 2018.

[17] S. Huang, X. Fan, L. Sun, Y. Shen, and X. Suo, "Research on classification method of maize seed defect based on machine vision," *Journal of Sensors*, vol. 2019, no. 1, 9 pages, Article ID 2716975, 2019.

[18] Y. Qin, Q. Na, F. Liu, H. Wu, and K. Sun, "Strain gauges position based on machine vision positioning," *Integrated Ferroelectrics*, vol. 200, no. 1, pp. 191–198, 2019.

[19] Q. Wang, M. Cheng, A. Noureldin, and Z. Guo, "Research on the improved method for dual foot-mounted Inertial/Magnetometer pedestrian positioning based on adaptive inequality constraints Kalman Filter algorithm," *Measurement*, vol. 135, no. 12, pp. 189–198, 2019.

[20] R. Cai and P. Zhu, "Face tracking with multi-feature based on markov random field," *Laser & Optoelectronics Progress*, vol. 54, no. 2, pp. 21–25, 2017.

[21] L. Bai, X. Yang, and H. Gao, "Corner point-based coarse-fine method for surface-mount component positioning," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 3, pp. 877–886, 2018.

[22] P. Lin, L. Xiaoli, and D. Li, "Rapidly and exactly determining postharvest dry soybean seed quality based on machine vision technology," *Scientific Reports*, vol. 9, no. 1, pp. 15–20, 2019.

[23] X. Huang, F. Wang, J. Zhang, Z. Hu, and J. Jin, "A posture recognition method based on indoor positioning technology," *Sensors*, vol. 19, no. 6, pp. 259–264, 2019.

[24] K. Hao, "Multimedia English teaching analysis based on deep learning speech enhancement algorithm and robust expression positioning," *Journal of Intelligent and Fuzzy Systems*, vol. 39, no. 3, pp. 1–13, 2020.

[25] X. Wu, Q. Cao, and Y. Li, "A research on wireless sensor networks' node positioning mechanism based on Narrowband Internet of Things data linking," *International Journal of Distributed Sensor Networks*, vol. 14, no. 12, pp. 36–42, 2018.

[26] Ó. de Francisco Ortiz, M. E. Amestoy, H. T Sánchez Reinoso, and J. Carrero-Blanco Martínez-Hombre, "Enhanced positioning algorithm using a single image in an LCD camera

system by mesh elements' recalculation and angle error orientation," *Updates in Production & Manufacturing*, vol. 19, no. 5, pp. 1–16, 2020.

[27] C. Shi, J. Qian, S. Han, B. Fan, X. Yang, and X. Wu, "Developing a machine vision system for simultaneous prediction of freshness indicators based on tilapia (*Oreochromis niloticus*) pupil and gill color during storage at 4 °C," *Food Chemistry*, vol. 243, no. 15, pp. 134–140, 2018.

[28] L. Fernández-Robles, G. Azzopardi, and E. Alegre, "Identification of milling inserts in situ based on a versatile machine vision system," *Journal of Manufacturing Systems*, vol. 45, no. 1, pp. 48–57, 2017.

[29] Q. Y. A, J. F. A, and J. T. B, "Development of an automatic monitoring system for rice light-trap pests based on machine vision," *Journal of Integrative Agriculture*, vol. 19, no. 10, pp. 2500–2513, 2020.

[30] Z. Zhang and H. Min, "Analysis on the construction of personalized physical education teaching system based on a cloud computing platform," *Wireless Communications and Mobile Computing*, vol. 2020, no. 3, 8 pages, Article ID 8854811, 2020.

[31] C. Jh, Z. Jj, and G. Rj, "Research on modified algorithms of cylindrical external thread profile based on machine vision," *Measurement Science Review*, vol. 19, no. 15, pp. 11–21, 2020.

Hindawi
Scientific Programming
Volume 2022, Article ID 4664388, 10 pages
https://doi.org/10.1155/2022/4664388

Research Article

Research on Badminton Teaching Technology Based on Human Pose Estimation Algorithm

Zhang Xipeng,[1] **Zhao Peng**,[1] **and Cao Yecheng**[2]

[1]*Qingdao Huanghai University, Qingdao 266000, Shandong, China*
[2]*Shijiazhuang University, Shijiazhuang 050035, Hebei, China*

Correspondence should be addressed to Zhang Xipeng; zhangxp@qdhhc.edu.cn

Received 28 January 2022; Revised 15 February 2022; Accepted 28 February 2022; Published 24 March 2022

Academic Editor: Baiyuan Ding

Human pose estimation is an important task in physical education, which can provide a valuable reference for teachers and students. We propose a human pose estimation method based on part affinity field. Firstly, the correlation of position information and orientation information between limb regions is maintained by part affinity field. Then the key points of limb pose are localized by part confidence map, and finally, the part affinity field is integrated to correlate all the acquired feature key points to obtain the human pose estimation. With the aid of computer vision technology, the students' training movements can be compared with the standard movements. It enables the students to feel the standard movements and badminton hitting points more intuitively. In the experiment, we set up a comparison experiment to compare the teaching mode of the method in this paper with the traditional teaching mode. The experimental results prove that through the teaching mode of our method, students have more standard strokes, more smooth skill switching between badminton serves and strokes, and higher badminton stroke scores. At the same time, such a teaching system adds a lot of fun to the course and makes the students' participation higher.

1. Introduction

Badminton is a popular sport around the world and is within our reach in our lives. As a sport that requires quick reaction and moving to catch the ball [1]. It requires a synergistic stroke variation between both badminton participants. Therefore, badminton is also a sport that tests the agility of the human body, due to its fast hitting and strategic tracking and prediction of badminton landing points, which turns badminton into a high-spectator sports competition. On the other hand, from a professional point of view, badminton is an extremely complex sport [2] with high demands on the physical and mental strength of the players. According to statistical reports from badminton professional bodies, it is estimated that there are about 150 million badminton enthusiasts and about 10 million professional badminton players worldwide [3, 4].

The key performance indicators of badminton are mainly reflected in the use of court areas, stroke distribution, technical movements, and stroke effectiveness. Like tennis [5–7], squash [8–10], and table tennis [11–13], the key performance indicators are mainly reflected in tactics. A successful badminton matchup will hold the opponent in spatial pressure to hit the badminton to the opponent's most disadvantageous area. The athlete widens the gap with the opponent when hitting the ball, inducing the opponent to create a large distance between the shots, making it impossible to return the ball in the next round, and at the same time consuming the opponent's physical strength [14, 15]. A professional badminton player should learn to fully control the integration of space, time, and tactics. Coaches, on the other hand, can obtain training details in the data of each game, which can be of great reference in future training [16]. For badminton teaching, teaching professional badminton skills movements in a decomposed manner can achieve a good teaching effect. as shown in Figure 1. Then human posture estimation can better correct the wrong badminton strokes.

With the popularity of badminton, badminton has become a physical education course in colleges and universities. At present, most of the badminton teaching adopts the traditional physical education class mode. The physical

Figure 1: Badminton stroke breakdown.

education teacher will first explain the rules system, action points, tactical skills, and scoring techniques of badminton through theory. Then take a practical face-to-face teaching, mainly to explain some action specifications, bucketing techniques, and sports injury prevention. This traditional didactic course does not allow students to truly feel the details of the action, and students' blind imitation is likely to lead to problems of sports strains and muscle injuries. This leads to ineffective physical education [17, 18]. With this teacher-led learning style, students can only learn badminton by imitating to feel the badminton action and feel the badminton skills through repeated practice later. This makes the quality of badminton teaching half the effort. In the long run, students will be dependent on the physical education teacher's demonstration guidance, in their practice, often will not get the point. Through this traditional way of teaching badminton, students often do not feel the main points of badminton personally, which reduces their motivation and participation, and they do not learn the main points of badminton to think after class [19]. Relevant studies have proven that there is a direct link between students' motivation and their sense of learning experience. Using smart strategy instruction in physical education can better increase students' engagement and interest [20].

A researcher has conducted solid three-dimensional modeling of player positions and stroke trajectories in badminton and analyzed the relationship between stroke trajectories and scoring points. This is a meaningful and highly expressive study. The presentation model of badminton is transformed from a pure spectator perspective to a data perspective. Data visualization can better assist badminton teaching and training [21]. Some researchers also started with videos of past badminton events and dissected the details of badminton skills from a video analysis perspective. Through the action decomposition algorithm, each frame of action is perfectly presented, and the students' attention is transferred from the audience's point of view to the skill learning link, so that students have an intuitive understanding of serving and hitting skills [22]. Chu et al. documented the link between space and stroke types in badminton through video analysis and proposed a classification detection model to analyze stroke technique characteristics. Although this analysis method of badminton players' performance patterns is able to present scoring details in complex matches, spatial and opponent-related movement details are not analyzed in an integrated manner [23, 24]. Badminton, as a high-intensity sport, has high demands on the quick reaction ability of athletes. Therefore, the technical details of opponents in space should be combined with before and after frames for detailed reference in multivariate performance analysis, and it is more informative to integrate the analysis of the matchups between athletes and opponents [25].

With the rapid development of technology, the sports industry urgently needs the intervention of artificial intelligence technology. The traditional badminton teaching mode is difficult to mobilize students' enthusiasm and participation. It makes it impossible for students to quickly acquire badminton action essentials and the quality of badminton teaching is low. With the upgrade of computer vision technology, stance estimation technology has the ability to be applied in badminton teaching. The development of these technologies has directly promoted badminton teaching and improved students' interpretation and interaction with badminton movements. The introduction of new technical systems for badminton instruction and practical practice is directly related to the quality of badminton instruction. Understanding the technical movements of badminton is quite important, the number of movements that occur in badminton matches is very high, and it is difficult for badminton teaching to explain the movements in a decomposed manner, making students understand the technical movements and thus comprehend the key points.

In this paper, we analyzed the current situation of badminton and found that the quality of badminton teaching is not optimistic. In order to further improve the quality of badminton teaching and students' participation. We consider the introduction of artificial intelligence technology into physical education. With the assistance of AI technology, the quality of physical education will be greatly improved. So we analyzed the research results related to human posture estimation and considered integrating it into physical education. Finally, we constructed a human-computer interactive badminton teaching system integrating computer vision technology and neural network algorithm. We propose a human pose estimation algorithm to correlate all the key points of acquired human features based on part affinity fields. Thus, human pose estimation and prediction are obtained. Finally, we demonstrate through comparison experiments with the traditional teaching model that the teaching model through our method results in more standardized strokes, higher badminton stroke scores, and more smooth skill switching between badminton serves and strokes. In addition, through our method, badminton teaching is more intuitive and interesting, which can greatly motivate students to learn.

2. Related Work

Ong et al. take inspiration from the flower pollination algorithm (FPA) and use FPA in real-time tracking of video actions of athletes' events. Real-time tracking of the athlete's center-of-mass pixel coordinates is achieved, and the search window is set following the adaptive law. The detection accuracy of the method is verified by experimental tests to meet the accuracy requirements [26]. To address the problem that the predicted range of the athlete's pose estimation does not enable energy-efficient point-to-point motion, Wang developed a linear time-invariant system in 2012. The system is able to control the optimal time to complete the point-to-point motion control and reduce the cost due to equipment capacity loss by energy optimization strategy, achieving an accuracy of 85% in the final system attitude estimation experiment [27]. Rutten et al. embarked on the study of badminton interactive robots and proposed an evolutionary operational derivation approach. The optimal pattern of energy is ensured from the time level to ensure the optimal badminton stroke trajectory. And by simulating the badminton player's stroke, the opponent's stroke trajectory is predicted as a way to achieve a comeback [28].

The biggest challenge facing human posture is the variation of human appearance, which is challenging in predicting the coordinates of key points in human space [29, 30]. Human pose estimation is a necessary joint technology for many industries, and the subject has been invested in research and development in various industries in recent years. Human pose estimation has different cut-offs in terms of the number of people, and single-person pose estimation has better accuracy and stability than multi-person pose estimation. In single-person pose estimation, only one person in the image is specified and localized, and then its key points are obtained. In multi-person pose estimation, it is necessary to first specify the number of people to be detected, then extract the individual with the highest human recognition weight from the image, and then localize its key points to capture the spatial coordinates. Obviously, multi-person pose estimation is more challenging, especially involving the effects of unstructured environments, such as crowds, occlusions, and multi-person interactions. The main structure of the human pose estimation algorithm is a convolutional neural network. In this paper, to estimate the pose of badminton action using this method, we choose multi-person pose estimation. In multi-person pose estimation, there are two main approaches: top-down and bottom-up.

The top-down approach will first detect the human body through the target recognition CNN network and then label it in the form of a rectangular box. The rectangular border is then used as a boundary to locate the human center of mass points in the rectangular box. The same operation is then used to locate other people appearing in the image in the same number as the specified number of people. Thus, the computational cost of multi-person pose estimation is closely related to the number of people specified. This type of algorithm uses mainly human detector techniques for the annotation of rectangular boxes and then applies a pose estimator to each detected person. So the detection results of multi-person pose estimation algorithms depend heavily on the accuracy of human detectors [31]. The human pose estimation based on convolutional neural networks can all be presented in the form of a heat map, where a deeper presentation represents a richer nodal capture. A convolutional pose machine is added to the output stage of each network, which has arrived at a supervised training role for the network, thus solving the gradient disappearance problem [32]. Besides, the stacked hourglass network can effectively expand the perceptual field and is a necessary continuous structure for pooling and upsampling in human pose estimation networks [33]. Of course, some researchers have also adopted the structure of feature pyramids to deal with feature maps of different resolutions [34]. However, in this paper, a cascaded pyramid network [35] is adopted to pool all the feature maps together to achieve multi-scale detection and obtain multi-scale human features. The literature [36] proposed a combined algorithm for the human pose keypoint prediction method, which can accurately predict the offset of the pose heat map and feature points and then perform offset correction to obtain accurate human spatial keypoint coordinates. Other researchers have tried end-to-end algorithms for bounding box regression and keypoint estimation for the human pose, such as Mask RCNN [37] and Faster R-CNN [38].

The bottom-up approach is significantly different, as it first detects all keypoint coordinates of multiple people in the full map and then performs keypoint combination on an individual basis to obtain pose estimation results. The literature [39] takes inspiration from the integer linear programming (ILP) algorithm and divides the detection population into clusters. The clusters are then matched with the labeled individuals as a way to obtain the final pose

estimation coordinate information. The literature [40] incorporates ResNet [41] based on the former to perform pairwise matching operations on adaptive images through conditional constraints. Both of these methods are representatives of ILP methods, but the computational cost of this type of method is too high for many researchers to adopt. Later, Newell et al. proposed fractional map and pixel-by-pixel embedding methods to successfully match pose key points to the corresponding people in different clusters, thus obtaining pose estimation information [42]. The literature [31] profoundly investigated a balanced approach between performance and speed, which was effective in saving computational costs in later studies. Just because of its high stability and adaptability and low computational cost, this algorithm is chosen as the basis of our human pose estimation method in this paper.

3. Method

3.1. Part Affinity Fields. In the input of the pose estimation algorithm, the input image is assumed to be of size $w \times h$. The input is fed into a convolutional neural network for the prediction of body joints to obtain a confidence map H and a part affinity field (PAF) L, L for each limb. The confidence map $H = (H_1, \ldots, H_J)$ indicates the existence of J confidence maps in the whole pose estimation, where $H_j \in \mathbb{R}^{w \times h \times 2}$, $j \in \{1, \ldots, J\}$. H_j^{GT} denotes the average position of the key points of each human body in the image, in terms of the whole without distinguishing individuals, and its label information is generated by a Gaussian distribution. PAFs $L = (L_1, \ldots, L_c)$, C denotes the number of vector fields, and each limb is assigned a vector field, where $L_c \in \mathbb{R}^{w \times h \times 2}$, $c \in \{1, \ldots, C\}$. L_c^{GT} denotes the unit vector real information generated independently for each limb, and each unit vector corresponds to a key point pair j_1 and j_2, and all vector directions within the rectangular box boundaries are j_1 pointing to j_2. Given the true label $y_{GT} = (H^{GT}, L^{GT})$ and the model prediction $P = (H, L)$, the model is trained using the mean square error $E_{L2}(P, y_{GT})$, defined as follows:

$$E_{L2}(P, y_{GT}) = \sum_{j=1}^{J} \sum_{p} W(p) \left\| H_j(p) - H_j^{GT}(p) \right\|_2^2 + \sum_{c=1}^{C} \sum_{p} W(p) \left\| L_c(p) - L_c^{GT}(p) \right\|_2^2, \tag{1}$$

where P denotes the two-dimensional coordinates of the key points, W denotes the binary mask, mainly for regions with nonstructural factors (e.g., overlapping occlusion cases like crowds), and $W(p) = 0$. The purpose of the mask is mainly to prevent the true predictions from being deleted by mistake during the training process. In addition to this, supervised units are added in the middle of each stage to periodically compensate for the gradient and prevent the occurrence of gradient disappearance [43]. The overall objective is f.

$$f = \sum_{t=1}^{T_p} f_L^t + \sum_{t=T_p+1}^{T_p+T_c} f_S^t. \tag{2}$$

In the parsing process, we use a set of two-assignment methods in order to better match the body part candidate matches. Firstly, we set the detection candidate region of the key points to the maximum position of the confidence map. Then, we match the linear integrals between the key points and then obtain the confidence scores of the body joints by calculation. Finally, all confidence scores are combined, and the final result of pose estimation is obtained by greedy correlation law.

3.2. Part Confidence Map. In order to evaluate the f in equation (2), we chose to start in two dimensions and generate confidence maps S^* from the annotated key points. Each confidence map corresponds to a corresponding limb site, and each limb site has a corresponding pixel coordinate representation in 2D coordinates. Ideally, if one person is detected during the detection process, a peak appears in the confidence map for each of its limb parts. If more than one person is detected, a peak appears similarly for each visible part j of each person k.

Suppose that each person k generates a corresponding individual confidence map $S_{j,k}^*$, where $x_{j,k} \in \mathbb{R}^2$ denotes the true position of body part j of each person k in the image. Then the value of position $p \in \mathbb{R}^2$ in $S_{j,k}^*$ is defined as follows:

$$S_{j,k}^*(p) = \exp\left(- \frac{\left\| p - x_{j,k} \right\|_2^2}{\sigma^2} \right), \tag{3}$$

where σ is the controller of the peak size; the convolutional neural network first predicts to obtain a single confidence map and then aggregates through multiple maximal operators to obtain the final ground truth confidence map as a single confidence map.

$$S_j^*(p) = \max_k S_{j,k}^*(p). \tag{4}$$

Throughout the network prediction process, we took the average value of the confidence map instead of the maximum value in order to maintain the same accuracy of the nearby peaks. As shown in Figure 2, in our experimental tests, when we predict confidence maps, we tend to choose the no-maximal suppression method to obtain independent body part candidates. In the comparative experiment, we selected Gaussian curves with different P values, namely Gaussian 1 and Gaussian 2. From the experimental results, it can be seen that the intersection between Gaussian 1 and Gaussian 2 is the optimal range of nonmaximum

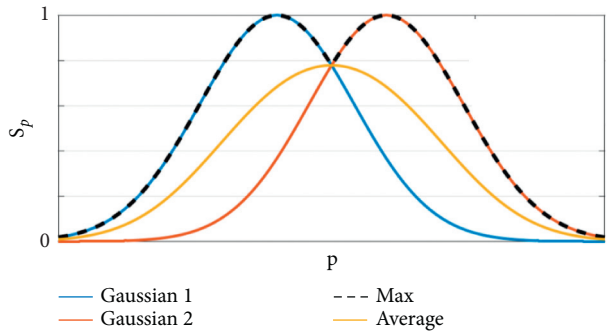


Figure 2: Confidence values for different peaks.

suppression because we take Gaussian 1 and Gaussian 2 to define the upper bound for nonmaxima suppression.

3.3. Detection and Association. The images are first fed into the convolutional neural network to generate the feature maps F required in the first stage, and then the VGG-19 initial layers are adjusted and initialized. In the first stage, $L^1 = \phi^1(F)$, represents each set of partial affinity fields (PAFs), where ϕ^1 represents the first inference result after network initialization. In the subsequent stages, each round of original image and feature F predictions will be based on the results of the previous stage, and then the PAF logic will be computed to obtain the accurate predictions.

$$L^t = \phi^t\left(F, L^{t-1}\right), \forall 2 \le t \le T_p, \tag{5}$$

where ϕ^t denotes the t stage of inference in the convolutional neural network and T_p denotes the total number of PAF stages. After each round of T_p iterations, the latest PAF predictions are passed through the confidence graph detection module.

$$\begin{aligned} S^{T_p} &= \rho^t\left(F, L^{T_p}\right), \forall t = T_p, \\ S^t &= \rho^t\left(F, L^{T_p}, S^{t-1}\right), \forall T_p \le t \le T_p + T_c, \end{aligned} \tag{6}$$

where ρ^t is the CNN used for inference at stage t and T_c is the number of total confidence map stages.

In contrast to the approach mentioned in the literature [31], the PAF and confidence maps are refined in each convolutional neural network inference stage. Thanks to the refinement process, the number of parameters in each inference stage is halved. Our preliminary experimental validation shows that the affinity field prediction is proportional to the confidence result. More generally, the output fusion results of each channel of PAF can be used to infer the corresponding body part. However, if we are only given fragmentary information about the body parts, we instead obtain PAF channel information and the associated confidence map results.

The effect of the refinement of the affinity field at different stages is shown in Figure 3. All the confidence map results unfold the predictions within certain PAF bounds, and this also creates a problem that there is no significant variation in variability between all the stage confidence maps. In the iterative process of the neural network, in order to ensure the correct matching of body joint point features. We add the loss function L2 at the end of each stage, which are applied to the PAFs of the body parts in the first branch and the confidence maps in the second branch. Finally, the real graphs and fields are matched. In addition to this, to address the drawback of the dataset, we used a spatial loss function weighting.

3.4. Human Pose Estimation Network. The structure of our proposed human pose estimation network is shown in Figure 4. The coded partial prediction is performed by an iterative approach, which is mainly divided into two parts: part affinity field ϕ^t and detection confidence map ρ^t. Based on the network structure mentioned in the literature [32], we optimize the stage prediction by designing its confidence map as a continuous prediction based on this, where $t \in \{1, \ldots, T\}$; each stage contains supervised units.

Compared with the network structure in the literature [31], the human pose estimation network in this paper has more layers. The initial conv 7×7 is replaced by three consecutive conv 3×3, such an improvement is inspired by the inception structure proposed by Google, and after such a structural optimization, the number of parameters of the network can be greatly reduced, thus reducing the computational cost. In addition, we also refer to the method of DenseNet [44], which takes three convolutional kernels as groups and concatenates the output of each group. Such an operation can increase the number of nonlinear layers and also extract higher- and lower-level features from the maximum perceptual field.

4. Experiments

4.1. Data Set. Badminton is a sport, and there is no dedicated badminton stance data set in the world. In order to verify the performance of our method, we requested data from badminton matches from relevant authorities. Then we perform manual classification based on badminton teaching points, then use video editing software for badminton action segmentation, and then perform manual annotation for each classified action. Finally, all the collected data were separated into a training set and test set, and a total of 30,000 movements were produced for the training set and 5,000 movements for the test set. All the following experimental results analysis is based on this data set.

4.2. Training. All experiments in this article were performed on Ubuntu 16.04 with python 3.7 configured as the programming language environment version. The experimental hardware environment uses an RTX 3080 Ti GPU, an Intel i7-7700 CPU, and 50 GB of RAM. The normalization criterion is used during training. It can optimize the model and improve the overall robustness and stability of the model, and a method based on stochastic policy estimation is proposed as an alternative. To avoid the difficulty of overfitting during the training of the deep network, the early stopping method is used in the training. The relevant training parameters are shown in Table 1.

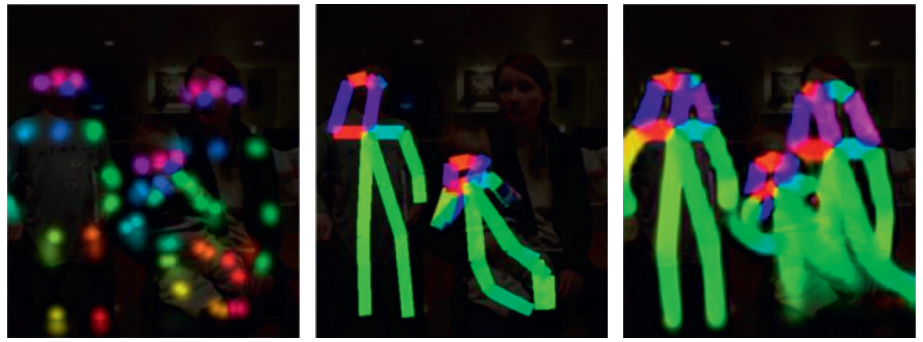


Figure 3: The effect of part affinity fields at different stages.

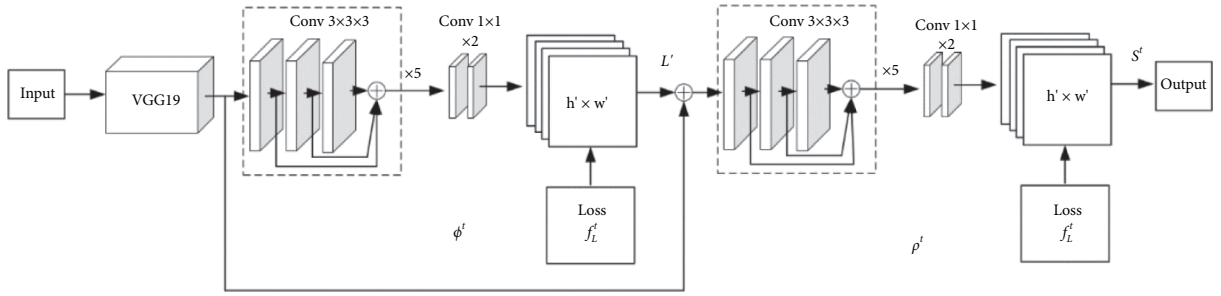


Figure 4: Human pose estimation network.

Table 1: Training parameter settings.

Parameter	Value
Epoch	30
Regularization	0.001
Initial learning rate	0.02
Hidden unit number	100
Weight attenuation coefficient	0.001
Momentum	0.9

4.3. Experimental Procedure. According to the physical education task set by the school, badminton class is held once a week for 90 minutes. For this reason, we set up a control group in the experiment: group A for the teaching method applying the human pose estimation method of this paper and group B for the traditional teaching method. For this purpose, we made detailed planning of the badminton teaching schedule as shown in Table 2.

The students in group A follow the human-machine interaction model. The class environment is in the machine vision arrangement of the field. We used the depth camera Kinect DK as a computer vision sensor to capture the badminton teaching situation through the depth camera and then feed it to the algorithm processing unit as shown in Figure 5.

Firstly, a human pose estimation algorithm is used to perform action decomposition of badminton strokes and then according to the main points explained. Then let the students in the machine vision field and capture the students comprehend the main points of badminton action; students can also visualize their own action details on the big screen and the standard action for comparison; students can correct their own action according to the standard action, so as to achieve a teaching purpose. At the end of the class, the teacher will comment on individual student movements with large differences and make comments. The video of the lesson with the human posture estimates is then given back to the students in the form of a video. Students can comprehend and think about badminton movements again after the class. The operation flow of the human-computer interaction system in group A teaching mode is shown in Figure 6.

The students in group B who follow the traditional teaching model learn badminton completely. The teacher will first verbally introduce the main points related to badminton, then demonstrate the main points of each badminton action through practice, and then let the students then practice in a cut and feel each badminton action. The teacher will give individual instruction to students with substandard posture. After the lesson, students are required to reflect on the instructional video and then write a summary.

4.4. Badminton Serve Evaluation. Regarding the assessment of badminton serve quality, this assessment guideline was followed for both groups A and B. Due to a large number of badminton skill movements, the assessment guidelines for each movement had subtle differences. In order to reflect the comparative performance of this study, we chose two skill movements, backhand and forehand, for the comparative experimental analysis. We assessed the pre- and post-test scores of the serve in four main areas: contact point, movement fluency, stroke trajectory, and stroke score. We

Table 2: Comparing the badminton teaching schedule of the experimental group.

Schedule	A	B
Week 1–4	Course introduction, human-computer interactive badminton action explanation, and learning	Course introduction and basic badminton skills learning
Week 5	Automatic pre-testing of badminton serve quality and serve correctness with feedback under machine vision	The pre-test of badminton serve accuracy and badminton serve quality
Week 6–9	Decompose difficult badminton skill movements through the human pose estimation algorithm and learn its key points	Learned difficult badminton skills, including ball control, footwork, swing, and stroke
Week 10	Automatic post-test of badminton serve quality and correctness based on machine vision and give feedback report	The post-test of badminton serve accuracy and badminton serve quality, a self-reflection report
Week 11	Badminton teaching effect test	Badminton teaching effect test

Figure 5: Computer vision environment construction.

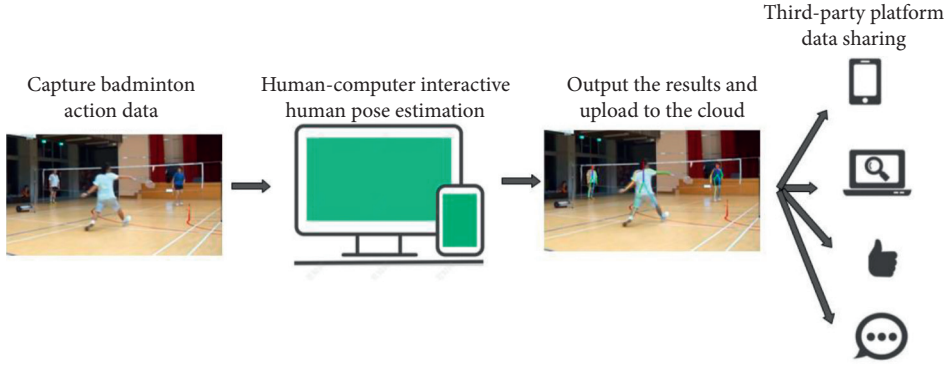


Figure 6: Human-computer interaction system operation flow for human posture estimation.

have summarized the scoring points of badminton strokes, and the specific scoring points are shown in Figure 7. The scoring rules of badminton are only divided according to the badminton court, and the extra points on the difficulty of the batter's actions are counted separately. We mainly study the scoring situation of badminton on the court.

We also assessed the variance of each weighted score, and the test of equality of variances implied satisfaction with the movement skills of the two groups assessed. To prevent differences in scores due to differences in student fitness, we also used ANCOVA analysis, which was primarily used to compare one variable in two or more aggregates, while considering other variables. The specific assessment results are shown in Table 3.

From the experimental results, it can be seen that group A was better than group B in serve quality assessment overall, the mean value of contact points in backhand action was higher in group A than group B by 0.35, the fluency of action in group A was more than group B by 0.31, the stroke trajectory in group A was better than group B by 0.25, and the average total score in group A was 1.09 ahead of group B. The experiment proves that in badminton teaching, the human-computer interaction teaching mode using the human posture assessment method, students' serving quality and scoring are excellent. Students' mastery of badminton movements was more precise. Students were more fluent in overall movement coherence for badminton sparring.

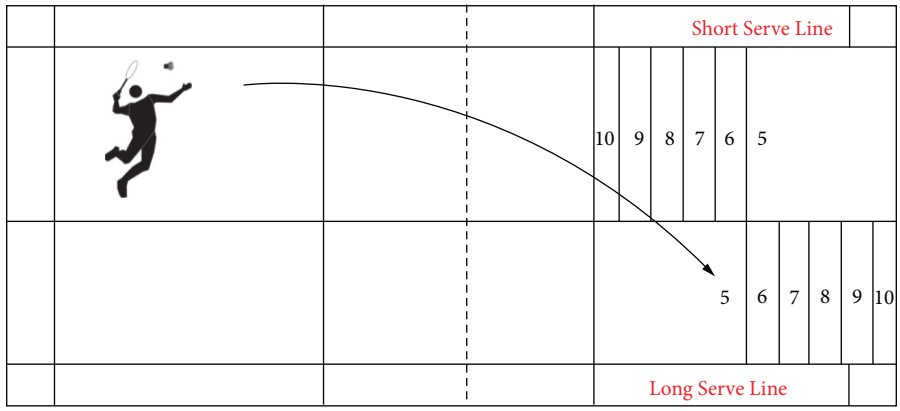


Figure 7: Badminton stroke scoring points.

Table 3: A and B group experimental group badminton stroke results.

Badminton skill	Evaluation options	Mean		SD		η^2	
		A	B	A	B	A	B
Backhand	Contact point	2.13	1.78	0.33	0.42	0.11	
	Movement fluency	2.16	1.85	0.41	0.36	0.12	
	Stroke trajectory	2.23	1.98	0.41	0.38	0.13	
	Stroke score	6.63	5.54	0.78	0.66	0.18	
Forehand	Contact point	2.23	1.66	0.43	0.54	0.24	
	Movement fluency	2.41	1.84	0.33	0.43	0.23	
	Stroke trajectory	1.34	1.85	0.61	0.61	0.15	
	Stroke score	6.97	5.36	0.98	0.99	0.31	

5. Conclusion

In this paper, we analyzed the current situation of badminton and found that the quality of badminton teaching is not optimistic. In order to further improve the quality of badminton teaching and students' participation. We integrate the human pose estimation algorithm into badminton teaching and construct a human-computer interactive badminton teaching system integrating computer vision technology and neural network algorithm. In this paper, we propose a human pose estimation algorithm, which firstly maintains the correlation of position information and orientation information between limb regions through part affinity fields. Then the center-of-mass key points are localized by part confidence map for the limb pose, and finally, all the feature key points obtained are correlated according to the part affinity field. Then the key points in each frame are integrated to obtain the features in the temporal dimension, and the pose estimation and prediction can be obtained. With our approach, real-time stance estimation can be achieved for badminton teaching, and students can observe the gap between their movements and the standard ones in the human-computer interface, thus feeling more deeply about the essentials of badminton strokes. In addition, such a teaching system makes badminton lessons more vivid and interesting and fully mobilizes the curiosity and motivation of students. Finally, we set up a comparison experiment to compare the teaching mode of this paper's method with the traditional teaching mode. The experimental results proved that through the teaching mode of our method, students' strokes were more standardized, the skill switching between badminton serves and strokes was more fluent, and badminton strokes scored higher.

Due to the wide range of badminton skill movements, the currently constructed dataset contains only two skill movements. For the neural network algorithm, the number of data sets determines the performance of the model. Therefore, in future research, we will further expand more professional badminton movements and improve the generalization ability of the model.

Data Availability

The data set can be accessed upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

[1] W. Chen, T. Liao, Z. Li et al., "Using FTOC to track shuttlecock for the badminton robot," *Neurocomputing*, vol. 334, pp. 182–196, 2019.

[2] A. Fabisch, C. Petzoldt, M. Otto, and F. Kirchner, "A survey of behavior learning applications in robotics-state of the art and perspectives," 2019, https://arxiv.org/abs/1906.01868.

[3] S. Lupashin, M. Hehn, M. W. Mueller, A. P. Schoellig, M. Sherback, and R. D'Andrea, "A platform for aerial robotics research and demonstration: the flying machine arena," *Mechatronics*, vol. 24, no. 1, pp. 41–54, 2014.

[4] C. Z. Shan, *Sensor-Based Assessment Using Machine Learning for Predictive Model of Badminton skills*, Universiti Teknologi Malaysia, Johor Bahru, Malaysia, 2018.

[5] R. Martínez-Gallego, J. F. Guzmán, N. James, J. Pers, J. Ramón-Llin, and G. Vuckovic, "Movement characteristics of elite tennis players on hard courts with respect to the direction of ground strokes," *Journal of Sports Science & Medicine*, vol. 12, no. 2, p. 275, 2013.

[6] N. S. Kolman, T. Kramer, M. T. Elferink-Gemser, B. C. H. Huijgen, and C. Visscher, "Technical and tactical skills related to performance levels in tennis: A systematic review," *Journal of Sports Sciences*, vol. 37, no. 1, pp. 108–121, 2019.

[7] C. Martin, B. Bideau, P. Touzard, and R. Kulpa, "Identification of serve pacing strategies during five-set tennis matches," *International Journal of Sports Science & Coaching*, vol. 14, no. 1, pp. 32–42, 2019.

[8] S. Murray, N. James, J. Perš, R. Mandeljc, and G. Vučković, "Using a situation Awareness Approach to Identify differences in the performance Profiles of the World's top two squash players and their opponents," *Frontiers in Psychology*, vol. 10, p. 1036, 2019.

[9] G. Vučković, N. James, M. Hughes et al., "A new method for assessing squash tactics using 15 court areas for ball locations," *Human Movement Science*, vol. 34, pp. 81–90, 2014.

[10] G. Vučković, N. James, M. Hughes, S. Murray, Z. Milanović, and J. Perš, "The effect of court location and available time on the tactical shot selection of elite squash players," *Journal of Sports Science and Medicine*, vol. 12, no. 1, p. 66, 2013.

[11] M. Fuchs, R. Liu, I. Malagoli Lanzoni et al., "Table tennis match analysis: a review," *Journal of Sports Sciences*, vol. 36, no. 23, pp. 2653–2662, 2018.

[12] J. Wang, "Comparison of table tennis serve and return characteristics in the London and the Rio Olympics," *International Journal of Performance Analysis in Sport*, vol. 19, no. 5, pp. 683–697, 2019.

[13] G. Munivrana, G. Furjan-Mandić, and M. Kondrič, "Determining the structure and evaluating the role of technical-tactical Elements in basic table tennis playing systems," *International Journal of Sports Science & Coaching*, vol. 10, no. 1, pp. 111–132, 2015.

[14] P. Abián, A. Castanedo, X. Q. Feng, J. Sampedro, and J. Abian-Vicen, "Notational comparison of men's singles badminton matches between Olympic Games in Beijing and London," *International Journal of Performance Analysis in Sport*, vol. 14, no. 1, pp. 42–53, 2014.

[15] C. L. Ming, C. C. Keong, and A. K. Ghosh, "Time motion and notational analysis of 21 point and 15 point badminton match play," *International Journal of Sports Science and Engineering*, vol. 2, no. 4, pp. 216–222, 2008.

[16] M. Hughes, M. T. Hughes, and H. Behan, "The evolution of computerised notational analysis through the example of racket sports," *International Journal of Sports Science and Engineering*, vol. 1, no. 1, pp. 3–28, 2007.

[17] J. Zeller, "Reflective practice in the ballet class: bringing progressive pedagogy to the classical tradition," *Journal of Dance Education*, vol. 17, no. 3, pp. 99–105, 2017.

[18] M. Xie, "Design of a physical education training system based on an intelligent vision," *Computer Applications in Engineering Education*, vol. 29, no. 3, pp. 590–602, 2021.

[19] P. Petsilas, J. Leigh, N. Brown, and C. Blackburn, "Creative and embodied methods to teach reflections and support students' learning," *Research in Dance Education*, vol. 20, no. 1, pp. 19–35, 2019.

[20] T. H.-C. Chiang, S. J. H. Yang, and C. Yin, "Effect of gender differences on 3-on-3 basketball games taught in a mobile flipped classroom," *Interactive Learning Environments*, vol. 27, no. 8, pp. 1093–1105, 2019.

[21] W. T. Chu and S. Situmeang, "Badminton video analysis based on spatiotemporal and stroke features," in *Proceedings of the 2017 ACM on International Conference on Multimedia Retrieval*, pp. 448–451, New York, NY, USA, 2017.

[22] S. Vial, J. Cochrane, A. J. Blazevich, and J. L. Croft, "Using the trajectory of the shuttlecock as a measure of performance accuracy in the badminton short serve," *International Journal of Sports Science & Coaching*, vol. 14, no. 1, pp. 91–96, 2019.

[23] G. Torres-Luque, Á. I. Fernández-García, J. C. Blanca-Torres, M. Kondric, and D. Cabello-Manrique, "Statistical differences in set Analysis in badminton at the RIO 2016 Olympic Games," *Frontiers in Psychology*, vol. 10, p. 731, 2019.

[24] M. Phomsoupha and G. Laffaye, "The science of badminton: game characteristics, anthropometry, physiology, visual fitness and biomechanics," *Sports Medicine*, vol. 45, no. 4, pp. 473–495, 2015.

[25] M. D. Hughes and R. M. Bartlett, "The use of performance indicators in performance analysis," *Journal of Sports Sciences*, vol. 20, no. 10, pp. 739–754, 2002.

[26] P. Ong, T. K. Chong, K. M. Ong, and E. S. Low, "Tracking of moving athlete from video sequences using flower pollination algorithm," *The Visual Computer*, vol. 38, pp. 939–962, 2022.

[27] X. Wang, J. Swevers, J. Stoev, and G. Pinte, "Energy optimal point-to-point motion using model predictive control," *American Society of Mechanical Engineers*, vol. 45301, pp. 267–273, 2012.

[28] K. Rutten, J. De Baerdemaeker, J. Stoev, M. Witters, and B. De Ketelaere, "Constrained online optimization using evolutionary operation: a case study about energy-optimal robot control," *Quality and Reliability Engineering International*, vol. 31, no. 6, pp. 1079–1088, 2015.

[29] S. Johnson and M. Everingham, "Clustered pose and nonlinear appearance models for human pose estimation," in *Proceedings of the British Machine Vision Conference*, Aberystwyth, UK, September 2010.

[30] B. Sapp and B. Taskar, "Modec: Multimodal decomposable models for human pose estimation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3674–3681, Portland, OR, USA, June 2013.

[31] Z. Cao, T. Simon, S. E. Wei, and Y. Sheikh, "Realtime multi-person 2D pose estimation using part affinity fields," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7291–7299, Honolulu, HI, USA, 2017.

[32] S. E. Wei, V. Ramakrishna, T. Kanade, and Y. Sheikh, "Convolutional pose machines," in *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pp. 4724–4732, Las Vegas, NV, USA, June 2016.

[33] A. Newell, K. Yang, and J. Deng, "Stacked hourglass networks for human pose estimation," in *Proceedings of the European Conference on Computer Vision*, pp. 483–499, Springer, Amsterdam, Netherlands, 2016.

[34] T. Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2117–2125, Honolulu, HI, USA, 2017.

[35] Y. Chen, Z. Wang, Y. Peng, Z. Zhang, G. Yu, and J. Sun, "Cascaded pyramid network for multi-person pose estimation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7103–7112, Salt Lake City, UT, USA, June 2018.

[36] G. Papandreou, T. Zhu, N. Kanazawa et al., "Towards accurate multi-person pose estimation in the wild," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4903–4911, Honolulu, HI, USA, 2017.

[37] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2961–2969, Venice, Italy, October 2017.

[38] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," *Advances in Neural Information Processing Systems*, vol. 28, pp. 91–99, 2015.

[39] L. Pishchulin, E. Insafutdinov, S. Tang, B. Andres, M. Andriluka, and B. Schiele, "DeepCut: joint subset partition and labeling for multi person pose estimation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4929–4937, Las Vegas, NV, USA, 2016.

[40] E. Insafutdinov, L. Pishchulin, B. Andres, M. Andriluka, and B. Schiele, "DeeperCut: a deeper, stronger, and faster multi-person pose estimation model," in *Proceedings of the European Conference on Computer Vision*, pp. 34–50, Amsterdam, Netherlands, 2016.

[41] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, Las Vegas, NV, USA, June 2016.

[42] A. Newell, Z. Huang, and J. Deng, "Associative embedding: end-to-end learning for joint detection and grouping," 2016, https://arxiv.org/abs/1611.05424.

[43] I. Radosavovic, P. Dollár, R. Girshick, G. Gkioxari, and K. He, "Data distillation: towards omni-supervised learning," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4119–4128, Salt Lake City, UT, USA, 2018.

[44] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4700–4708, Honolulu, HI, USA, July 2017.

Hindawi
Scientific Programming
Volume 2022, Article ID 9769044, 12 pages
https://doi.org/10.1155/2022/9769044

Research Article

Research on Travel Reimbursement Behavior Management Based on Deep Learning in Financial Sharing Mode

Qian Zhang[1] **and Bowei Feng**[2]

[1]*Jilin International Studies University, Jilin Changchun 130117, China*
[2]*State University of Information Technology, Mechanics and Optics, St. Petersburg 197376, Russia*

Correspondence should be addressed to Qian Zhang; zhangqian@jisu.edu.cn

Received 9 January 2022; Revised 24 January 2022; Accepted 16 February 2022; Published 22 March 2022

Academic Editor: Baiyuan Ding

The standardization, transformation and upgrading of financial management plays an important supporting role in promoting the standardized management and healthy operation of corporate expense reimbursement behaviors. This paper starts with the behavioral portrait of enterprise personnel travel expense reimbursement. Based on the problems of lengthy process and complex financial accounting in most reimbursement behaviors at this stage, an efficient and efficient expense reimbursement processing method is proposed, that is, reimbursement through collection Information images, using the convolutional neural network algorithm under deep learning, to detect target images and identify valid information, and rely on the financial shared service center's screening function of expense reimbursement behavior portraits under big data, the integrity of expense reimbursement data. Compliance is judged, thus forming a complete set of standardized and procedural personal self-service reimbursement model, which optimizes the management of enterprise expense reimbursement behavior, avoids subsequent losses and resource waste caused by human risks, and improves enterprise operation efficiency and expenses. Control has certain value and reference significance.

1. Introduction

In order to adapt to the changing requirements of the development strategy and management system of the enterprise, further promote the healthy and rapid development and transformation and upgrading of the enterprise, promote the standardization of financial management, and a virtuous circle of standardized management of expense reimbursement behavior. The first task is to diagnose the problem of employee reimbursement behavior, that is, reimbursement content and expense data, to ensure the compliance and effectiveness of reimbursement behavior. The data terminal based on the enterprise financial sharing center has the information identification function of the user's consumption operation behavior, which can realize the transformation of enterprise expense reimbursement to intelligent management and control and value creation; the construction of risk prediction system to highly controllable, full-process supervision and early warning will become a Focus on the focus, so how to manage the expense reimbursement behavior from user reimbursement to reimbursement is the focus of this paper. With the rapid development of computer vision technology and deep learning technology, more and more work is done by humans instead of by computers. Based on the current situation of travel reimbursement behavior, this paper will use the convolutional neural network image processing method under deep learning to carry out business travel through the functions of whole-process accounting, centralized standardized supervision and real-time data support provided by the Financial Shared Service Center. Research on the detection and identification of reimbursement picture information, and research on the compliance and validity of the reimbursement content list and expense data, trying to solve the problems of manual submission of expense reimbursement vouchers and the tedious verification of ledger

information by financial personnel, and formed a complete set of standardized research. And programmed personal self-service reimbursement mode, in order to achieve standardized and efficient reimbursement behavior management of user expense reimbursement behavior.

2. Related Work

With the advancement of science and technology and the development trend of business and financial integration, the concept of "financial sharing" was born. In order to achieve the goal of improving work efficiency, reducing operating costs, and improving service quality, the Financial Shared Service Center centralizes the processing of various scattered operating functions. Realize the centralized analysis of financial-related information of each business unit of the group enterprise, which greatly promotes the efficiency of corporate decision-making [1].

As one of the most common content of financial shared service application scenarios, travel reimbursement business has attracted the attention of academics and practical scholars. Cloud computing, OCR, electronic reimbursement information, etc. have brought development opportunities for expense reimbursement business, helping companies realize the goals of electronic bills, online approval, intelligent auditing, and automation of fund payment. As a key data source of accounting information, it is the beginning of corporate accounting. Link [2–5].

From the current situation and needs of travel reimbursement business, Guo Fang discussed the issue of expense reimbursement in the traditional branch management of group companies, focusing on the development of centralized reimbursement process organization, internal approval mechanism and information system construction [6–10]. Xing Ying analyzes the key issues that arise in the multiple business scenarios of financial reimbursement, and establishes a compliance control design through the discussion of the expense reimbursement system and reimbursement process [8]. With the gradual improvement of informatization construction, online reimbursement and reimbursement systems have been developed and applied successively to reshape the original reimbursement nodes and improve the efficiency of reimbursement and the level of financial management. However, in practice, online reimbursement still has a large amount of reimbursement documents, various types, low comprehensive level of reimbursement staff, and misconduct of reimbursement, which are the pain points under the current status of the reimbursement business market [11].

From the perspective of the reform and development trend of travel reimbursement, Qian Yijun and others believe that online reimbursement is a technical iteration of traditional reimbursement business from the perspective of mobile office needs. The three parties are directly connected to each other to realize the automatic processing of accounting and payment [12–15]. Luo Xiangming further mentioned that online expense reimbursement is operated through the network. Employees and leaders of an enterprise can interact with financial reimbursement in a modern office without being restricted by time, space and methods. After the original vouchers are verified, they can pass directly. Presigned banks conduct online payment settlement [16].

Expense reimbursement behavior management is the core content of enterprise reuse, and the formulation of individualized plans based on the classification of problem behaviors has high practical and practical significance. The optimization implementation involves intelligent budget planning and credit system construction, as well as the optimization concept of intelligent declaration path. The classification method of early warning mechanism based on behavioral characteristics and the level management based on scoring method are embedded, and financial intelligent information technology such as voice print is added to the declaration. This makes up for the lack of artificial intelligence technology research in the field of accounting. Focusing on business embedded technology, deepening joint multidisciplinary research will be an important direction in the future. Based on the perspective of deep learning, this paper will discuss the research on information identification and detection in travel reimbursement under the financial sharing model, and provide new ideas and directions for the in-depth development of standardized reimbursement behavior management and embedded technology.

3. Related Theories

3.1. Financial Sharing. Financial sharing is actually a process of reengineering and standardizing the financial business that is scattered in various projects, has high operational repeatability, and is easy to standardize. A financial management model that improves efficiency, reduces costs, allocates resources rationally, and increases the real-time and accuracy of information transmission. The financial shared service center is to comprehensively transform the traditional financial operation mode. The core principle is "professional stratification, business integration", relying on the shared service platform to create the integration of shared finance and business finance.

In terms of functional composition, the financial shared service center is supported by whole-process accounting, centralized standardized supervision and real-time data support. In addition, as a data support center, it can take advantage of its information advantages to provide enterprises with comprehensive, multi-dimensional and real-time accounting information services such as accounting and accounting reports. need. In the guarantee system supported by the financial shared service center, its core business platforms mainly include ERP system and SAP system, its auxiliary business platform includes the whole-process online reimbursement and payment system, and the information management system includes performance management system, expense management and control system, IT audit system, expense budget management system, etc. For the differences in business processing methods, the financial shared service center can provide a stable underlying interactive system and build a complete set of standardized systems, which is of great importance to realize the unified and centralized management of the financial

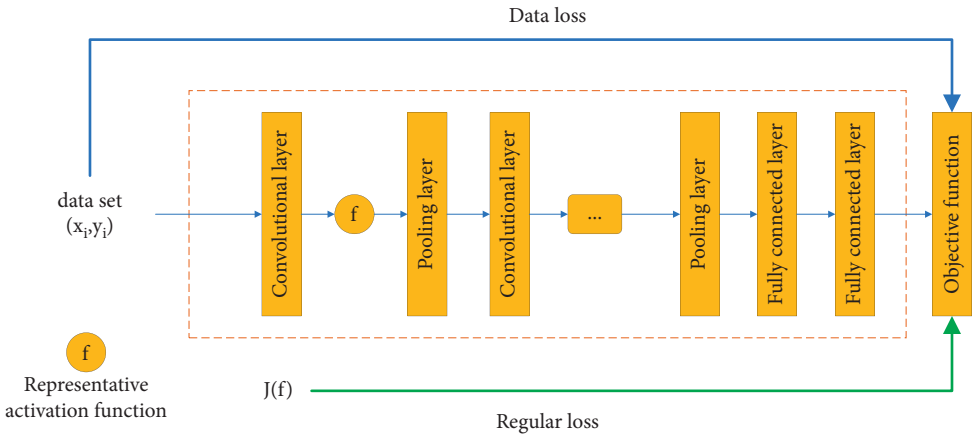


Figure 1: Convolutional neural network structure.

shared service business and ensure the compliance operation of various businesses. effect.

3.2. Convolutional Neural Network. Among the deep learning techniques that have emerged in recent years, the most widely used convolutional neural network is undoubtedly the convolutional neural network [17–21]. Convolutional neural network is a hierarchical model. It is mainly composed of an input layer, a convolutional layer, an activation function, a pooling layer, and a fully connected layer. Its input is similar to the original data of the image. Its feedforward operation is a process of acquiring and abstracting feature information layer by layer from the input layer through operations such as convolution, pooling, and nonlinear activation function mapping. Among them, different operations generally correspond to their related layers one-to-one. For example, the convolution operation corresponds to the convolution layer and so on [22–24]. Finally, its last layer converts target tasks such as classification and regression into objective functions, and uses back-propagation algorithm to feed back and update the difference between the predicted value and the true value from the last layer. Figure 1 shows the structure of the convolutional neural network.

3.2.1. Input Layer. The input layer is the beginning of the network model and is mainly used to input the data required for training. When the image is operated, it is generally stored and calculated in the form of a matrix for the color input image. Therefore, for a black and white input image, the input to the network is a two-dimensional matrix; like the number of channels of a color image, each channel corresponds to a two-dimensional matrix, so the input network is three two-dimensional matrices.

3.2.2. Design of DNQ Algorithm Structure Based on Pointer Network. In a convolutional neural network, each neuron in the hidden layer can be regarded as a convolution kernel, and each convolution kernel will perform a sliding convolution operation on the image. The convolution kernel is used to extract the features of the image, thanks to its sparse connection and weight sharing.

(1) Sparse Connection. For a general artificial neural network, neurons in adjacent layers are closely connected. That is to say, every neuron in the current layer maintains a connection with all neurons in the front and rear layers, and each connection means a weight that needs to be learned and saved. This will lead to when the network model has a deeper number of layers. The network is difficult to train and converge, and consumes a lot of memory, which makes it difficult to complete the task.

In a convolutional neural network, each neuron only needs to save the weight in its own structure. For example, for a convolution kernel with a size of 3×3, it only needs to store its own 9 weights. This means that the number of connections between each neuron and the previous layer is only related to its own size. Compared with the full connection, the convolution kernel has few connections, so it is called sparse connection.

(2) Weight Sharing. It can be seen from the sparse connection that each convolution kernel only stores a small amount of weights. For the same convolution kernel, in each round of training, the weights used in sliding convolution are the same. Only when one iteration is completed, will its internal weights be updated.

Therefore, for the same convolution kernel, in the same round of iteration, the weight of each convolution is constant, so it is called weight sharing.

The size of the image after the convolution operation is related to factors such as the size of the convolution kernel, the step size, and the pooling size. The specific formula is

$$O = \frac{I - F + 2P}{S} + 1. \tag{1}$$

Among them, I and O represent the size of the input and output respectively, F is the size of the convolution kernel, and P and S represent the size of the filled edge pixel and the convolution step size, respectively.

The size of the convolution kernel determines the size of the receptive field obtained after each convolution operation. The filled edge pixels can prevent the size of the image from becoming smaller and smaller due to the convolution operation and can avoid the loss of boundary information. Convolution step size The setting of directly causes the image size to change, and can also reduce the amount of calculation.

3.2.3. Pooling Layer. Usually several consecutive convolutional layers are used to extract more features, but this also means a large amount of calculation and parameters. Therefore, in order to reduce the amount of calculation and compress the image feature map, a pooling layer is generally added in the middle of the continuous convolutional layer. The operation of the pooling layer is very similar to that of the convolutional layer. Without filling, you only need to adjust the size of the convolution kernel to 2 in the convolution formula to realize that the size of the output image is half the size of the input image. According to different needs, there are two main operations of the pooling layer, namely maximum pooling and average pooling.

(1) Maximum pooling. Taking the case where the size of the pooling layer is 2×2 as an example, the maximum pooling is to take the largest pixel value in the 2×2 area as the pixel value of the area, and at the same time the size of the area changes from 2×2 to lxl.

(2) Average pooling. Similarly, taking the case where the size of the pooling layer is 2×2 as an example, the average pooling is to use the average value of all pixel values in a 2×2 area as the pixel value of the area, and the size of the area is changed from 2×2 to lxl.

It can be seen from the two pooling methods that the maximum pooling retains the maximum pixel value in the region, which is the texture feature in the image, while the average pooling retains the average value of the pixel values in the region, which is integral, Which is the overall characteristics of the image.

3.2.4. Active Layer. The essence of convolutional neural network training is to make the model have a good fit to the data, and at the same time have a good generalization ability. The convolution operation is essentially a linear operation. In order to make the model have better expressive ability, it is often necessary to add a certain nonlinearity, that is, add an activation layer after the convolution layer.

The activation layer structure is relatively simple, generally just an activation function, used to add nonlinearity to the output result of the convolutional layer. Commonly used activation functions include Sigmoid function, Tanh function and ReLU function.

(1) Sigmoid Function. The Sigmoid function is a typical S-shaped curve, and its function and derivative expressions are respectively

$$S(x) = \frac{1}{1 + e^{-x}},$$
$$S(x) = \frac{e^{-x}}{\left(1 + e^{-x}\right)^2} = S(x)(1 - S(x)). \tag{2}$$

From the Sigmoid function and its derivative, it can be found that the output result of the function is in. It is monotonic and has central symmetry, so it is very suitable as an activation function.

The characteristics of the sigmoid function are very suitable for use in binary classification problems, but because of the large amount of calculation of the function and the slow change of the derivative at both ends of the function, the gradient disappearance phenomenon is prone to occur during back propagation, which makes the deep model difficult train.

(2) Tanh Function. The Tanh function is the hyperbolic tangent function, and its function and derivative expressions are

$$T(x) = \frac{e^{x} - e^{-x}}{e^{x} - e^{-x}},$$
$$T(x) = 1 - T^{2}(x). \tag{3}$$

It can be found from the Tanh function and its derivative that it is very similar to the Sigmo id form, and the function image is very similar. It can be further found that the conversion relationship is

$$T(x) = 2.S(2x) - 1. \tag{4}$$

(3) ReLU Function. The ReLU function is the modified linear unit, which has the simplest function expression form, and its function and derivative expressions are

$$R(x) = \max(0, x),$$
$$R(x) = \begin{cases} 0\, x < 0 \\ 1\, x \geq 0 \end{cases} \tag{5}$$

From the ReLU function and its derivative, it can be found that the output of the ReLU function is 0 on the left side of the vertical axis, and the input itself is on the right side. It is only a function that takes the input and 0 to the maximum value, and the amount of calculation is very small. Observing the derivative of the ReLU function at the same time, it can be found that the derivative form is simpler, with a piecewise function with 0 as the boundary, and each segment is a constant, and the amount of calculation in the back propagation is also very small.

As shown in Figure 2, compared with the Sigmoid function and the Tanh function, it can be found that the ReLU function has a small amount of calculation, so that the model can converge faster; and the output on the left side of the vertical axis is 0, which means that it can play a certain role on the neuron. The most important thing is that the derivative of the ReLU function has nothing to do with the

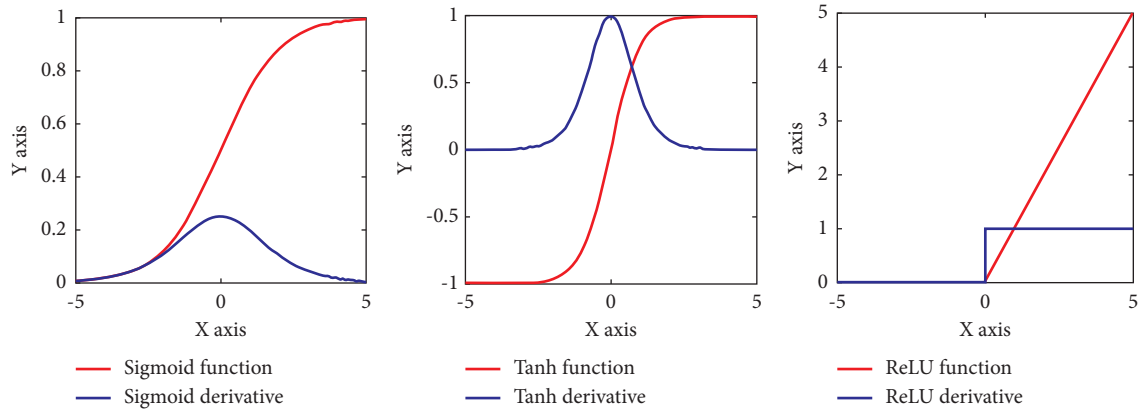


Figure 2: Comparison of three activation functions and their derivatives.

input and is a constant, thus eliminating the phenomenon of the disappearance of the gradient.

3.2.5. Fully Connected Layer. The function of the convolutional layer, the pooling layer and the activation layer can be seen as extracting features of the input image, that is, mapping the input original data into its feature space, and the function of the fully connected layer is to use the extracted features for classification tasks. That is, the feature is mapped to the label space. In other words, the function of the fully connected layer is to integrate the local features extracted from the previous layers to obtain overall feature information.

The operation performed by the fully connected layer is also very simple, that is, the basic matrix operation, and the result obtained is a one-dimensional vector. But because of its fully-connected nature, it often needs to store a lot of parameters and consumes a lot of memory.

In order to reduce the memory, considering the advantages of the convolutional layer, the convolutional layer can be used to replace the fully connected layer, thereby reducing the training difficulty and memory of the network model.

4. Construction of a Travel and Reimbursement Information Detection and Recognition Model Based on Convolutional Networks under Financial Sharing

4.1. The Overall Process of Travel Reimbursement under Financial Sharing. Electronic reimbursement information and paper reimbursement information. For electronic reimbursement information, due to its relatively regular form, uniform templates and printing standards, the information can be extracted and identified based on its location information using traditional methods. As for paper-based reimbursement information, with the massive use of mobile devices, more and more reimbursement information images are captured by mobile devices such as smart phones. Due to the different shooting habits and the shooting quality of mobile devices, the reimbursement information images taken by them are difficult to quickly identify with traditional methods. This section designs a method based on deep learning. In the reimbursement behavior, the reimbursement information images taken by smartphones, Perform target detection and information identification, synchronize the integrity of the identified information and data to the financial sharing service center for file integrity and data compliance determination, and through cost control, the final termination of the reimbursement behavior is completed. The overall process is shown in Figure 3:

The process framework is based on the identification of information for travel reimbursement, and then extends to the normative guidance of the financial reimbursement process, the three major management requirements of the discrimination of fraud and risk early warning management, and the support of deep learning algorithms under the big data platform. Provide intelligent technical solutions. Based on the difficulty in distinguishing the types of expense reimbursement behaviors faced by enterprises, and the difficulty in identifying and preventing fraudulent behaviors and operational risks, the reimbursement behavior management life cycle links are optimized and designed from the source.

4.2. Creation of Data Set of Travel and Reimbursement Information under Financial Sharing. The method based on deep learning is a kind of machine learning method. The essence is to let the model fit a certain data set and complete the corresponding task through its generalization ability. The convolutional neural network model used in this paper also needs to be trained to fit the travel and reimbursement data set. Therefore, the quality of the data set determines the upper limit of the task completion quality, and the training model is only constantly approaching this upper limit. In order to obtain enough travel reimbursement data image information and compliant travel reimbursement list information, use crawler technology, use keywords in Google pictures to obtain corresponding travel reimbursement information images and extract them according to the standardization of the Financial Shared Service Center The processed travel reimbursement list data information image. Because the amount of data is relatively small compared to the model, it causes overfitting problems. Therefore, it is

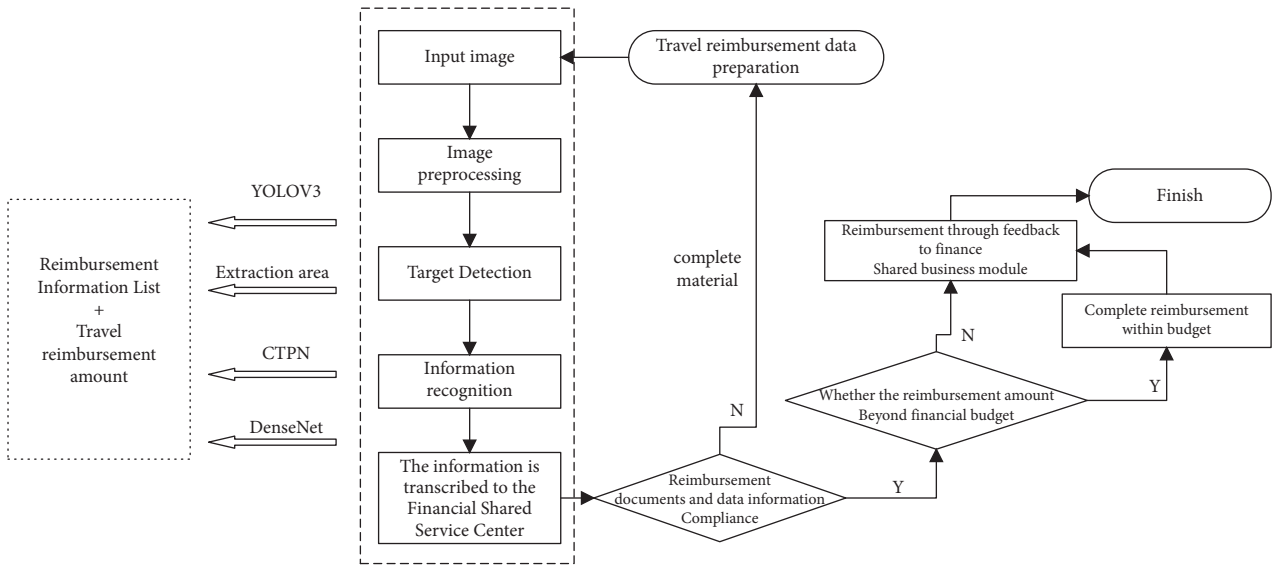


Figure 3: Overall flow chart of travel reimbursement under financial sharing.

necessary to perform data enhancement on the extracted image information, that is, use traditional image processing methods to expand the image. Since the features extracted by the convolutional neural network have nothing to do with the position, this section again uses Gaussian noise, horizontal and vertical flipping and rotation methods to expand the data set. After the data set is expanded, because it is a problem of target detection, different areas of the reimbursement information data set need to be labeled differently, that is, labeling. LabelImage software is used to label the reimbursement information images. After the data set is fully labeled, each piece Images are generated in the corresponding XML tag file, which contains the image size, the position information of the four rectangular boxes and their tag names.

There are a total of 1250 images in the data set. First, the data set is randomly divided according to a ratio of 9 : 1, of which 90% of the images, that is, 1125 images, are used as the training set. Then the remaining 125 images are randomly divided according to the ratio of 9 : 1, of which 90%, that is, 112 reimbursement information images, are used as the verification set, and the remaining 10%, that is, 13 images are used as the test. set. At the same time, the generated annotation data is divided into training set, verification set and test set accordingly.

4.3. Construction of Target Detection Model for Travel Expense Reimbursement Information under Financial Sharing. The POLO-V3 model is based on the Darknet-53 model and detects the outputs of the Darknet-53 model at three different depths respectively, so as to achieve multi-scale detection, and use the regression method to get the position of the detected target.

4.3.1. Darknet-53 Model. The probability that an image belongs to a certain category can be determined according to the value between 0 and 1. In the end, it is only necessary to find the category corresponding to the value with the highest probability, which is the classification result of the input image by the model. The basic model for extracting image features mainly uses Darknet-53, and its structure is shown in Figure 4.

	Type	Filters	Size	Output
	Convolutional	32	3 × 3	256 × 256
	Convolutional	64	3 × 3 / 2	128 × 128
1×	Convolutional	32	1 × 1	
	Convolutional	64	3 × 3	
	Residual			128 × 128
	Convolutional	128	3 × 3 / 2	64 × 64
2×	Convolutional	64	1 × 1	
	Convolutional	128	3 × 3	
	Residual			64 × 64
	Convolutional	256	3 × 3 / 2	32 × 32
8×	Convolutional	128	1 × 1	
	Convolutional	256	3 × 3	
	Residual			32 × 32
	Convolutional	512	3 × 3 / 2	16 × 16
8×	Convolutional	256	1 × 1	
	Convolutional	512	3 × 3	
	Residual			16 × 16
	Convolutional	1024	3 × 3 / 2	8 × 8
4×	Convolutional	512	1 × 1	
	Convolutional	1024	3 × 3	
	Residual			8 × 8
	Avgpool		Global	
	Connected		1000	
	Softmax			

Figure 4: Darknet-53 structure diagram.

It can be seen from the structure diagram that the Darknet-53 model contains a total of 53 convolutional layers, of which the first 52 convolutional layers are used to

extract features, and the last convolutional layer has a size of 1×1 and a step size of 1, which is used to replace the entire convolutional layer. The connection layer is classed to reduce the amount of calculation, so the model is named Darknet-53.

The core structure in the model is the residual unit. Each residual unit is mainly composed of two consecutive convolutional layers. At the same time, the input of the residual unit and the output of the convolutional layer are added, and the sum is used as the residual unit Output. The structure diagram of the residual unit is shown in Figure 5.

Among them, x is the input of the residual unit, ReLU is the activation function, Fix) is the output of the two convolutional layers, and $H(x)$ is the output of the residual unit, that is, the sum of $F(x)$ and x. It can be seen that the residual unit directly adds the input and the output as the final output to ensure that the worst case of its structure, that is, the performance of the previous layer can be maintained when the current convolutional layer has zero output. This type of gradient disappears and gradients explode. Therefore, the role in the deep convolutional neural network. At the same time, the deep mode is well resolved, and the residual unit plays an important role.

There are five residual modules in the model, and each module is a stack of several residual units to extract deeper features. The number of residual units in the five modules are 1, 2, 8, 8, and 4 respectively. After the residual module has extracted the features, it uses the global average pooling method to fuse the features to facilitate subsequent classification.

4.3.2. POLO-V3 Model. The model can correspond to different feature scales and can be used to detect objects of different sizes. The POLO-V3 model architecture is based on the Darknet-53 model. The first 52 convolutional layers of the Darknet-53 model are used to extract the features of the input image, and then the three feature maps of different scales in the extraction process are respectively performed Detection, using the idea of regression, so as to achieve the task of detecting and positioning the input image. The overall structure of the POLO-V3 model is shown in Figure 6.

It can be seen that the part marked by the red dashed box is the first 52 layers of the Darknet-53 model, that is, the feature extraction part. Using its feature extraction model, the detection is performed after the third, fourth and fifth residual modules of the feature extraction model. The size of the feature map is 52×52, 26×26 and 13×13, which correspond to different feature scales used to detect targets of different sizes.

4.4. Construction of Target Recognition Model of Travel Expense Reimbursement Information under Financial Sharing. Information recognition is divided into two parts: text detection and text recognition. The text detection part uses the CTPN model to initially detect the text area in the reimbursement information image, and then uses postprocessing methods to optimize the preliminary detection results; the text recognition part uses the DenseNet model Recognize the detected text area.

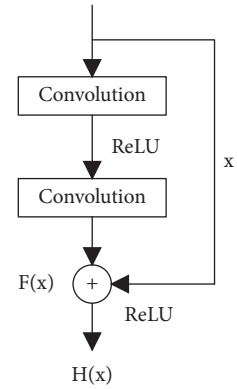


Figure 5: Residual unit structure diagram.

4.4.1. CTPN Model. The overall structure of the CTPN model includes the VGG-16 model, the BiLSTM model and the full convolutional layer. It combines the VGG-16 model and the BiLSTM model to extract both the spatial information of the image and the sequence features of the text, which is more suitable for detection For the text in the image, the fully convolutional layer is finally used to replace the fully connected layer, which simplifies the calculation and improves the efficiency. The overall result of the CTPN model is shown in Figure 7.

First, use the VGG-16 model to extract spatial features of the input image. Since only features need to be extracted and no classification is required, only the convolutional layer part is retained in the V GG-16 model, and its structure is shown in Figure 8.

The dimension of the feature map extracted by VGG-16 is NxCxHxW, where N is the number of output feature maps, C is the number of channels of the feature map, and H and W are the dimensions of the feature map, that is, height and width. Since the text information in the reimbursement information image is arranged horizontally from left to right, the 3×3 check feature map is used to slide row by row, and each pixel and the surrounding eight pixels are combined as a feature vector to obtain the dimension It is the feature map of Nx9CxHxW.

The size of the full convolution layer is set to 512, and the feature map dimension of the previous layer output is Nx256xHxW, then the output of the full convolution layer is the feature map dimension of Nx512xHxW. Three outputs can be obtained by classification and regression respectively, which are rectangular boxes. Position, rectangular box category score and rectangular box boundary calibration. Since the width of each rectangular box is fixed at 16 pixels by default, the position of the rectangular box is the ordinate and height of the center point of the detected rectangular box. The type of the rectangular box is whether the rectangular box contains text information. The boundary of the rectangular box is calibrated as a rectangle. The offset of the

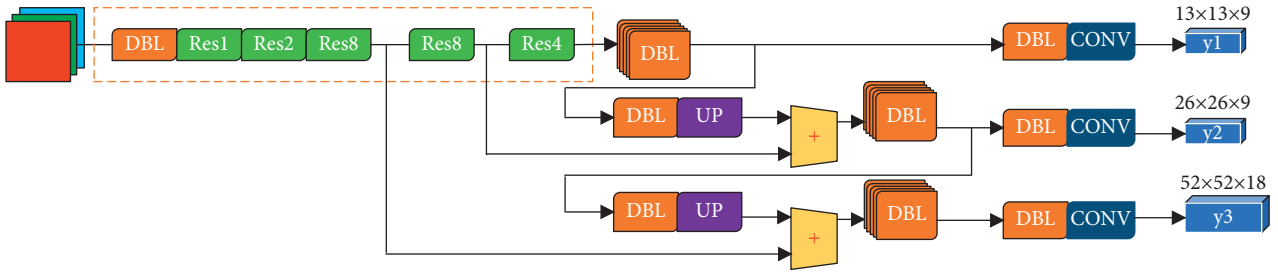


Figure 6: POLO-V3 model structure diagram.

frame in the horizontal direction. The loss function of the model is

$$L(S_i, V_j, O_k) = \frac{1}{N_x}\sum_i L_s^{cl}(S_i, S_i^*) + \frac{\lambda_1}{N_v}\sum_i L_v^{re}(V_j, V_j^*) + \frac{\lambda_2}{N_O}\sum_i L_o^{re}(O_k, O_k^*). \tag{6}$$

Among them, S_i, V_j, and O_k are the model category, the vertical coordinate and the horizontal offset of the L_s^{cl} is the cross-dimension loss function, which is used to calculate the category error, while L_v^{re} and L_o^{re} are both smooth L1 functions, which are used to calculate the vertical coordinate and horizontal offset errors of the rectangular box, respectively. N_s., N_v, and N_o respectively represent the category, the ordinate of the rectangular frame, and the total number of anchor frames used for the horizontal offset of the rectangular frame. λ_1 and λ_2 are the weights for balancing different tasks, which are set to 1 and 2, respectively.

As for the horizontally arranged text, 10 anchor boxes with the same width are used to locate the text more accurately. The width of each anchor frame is 16 pixels, and the heights are 283, 198, 139, 97, 68, 48, 33, 23, 16, and 11, respectively, decreasing by a ratio of 0.7. Correct the 10 anchor boxes detected in each group, leaving only one anchor box as the text detection box at the current position.

4.4.2. DenseNet Model. After obtaining the text detection frame, train the DenseNet model to recognize the text in the image in the detection frame, so as to obtain the text information in the image. The biggest feature of the DenseNet model structure is that the input of each layer comes from the output of all the previous layers. That is to say, the output of each layer takes into account the results of each layer before it, which greatly enhances the characteristics of the image. Delivery in the model. In order to facilitate training, the DenseNet model is simplified, and its structure is shown in Figure 9.

5. Simulation Experiment and Result Analysis

5.1. Simulation. After the data set is prepared and the model structure is determined, the model can be simulated and

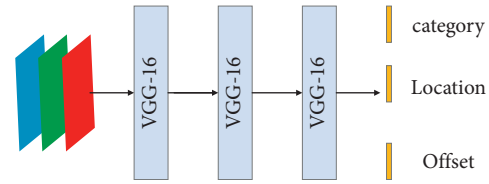


Figure 7: CTPN model structure diagram.

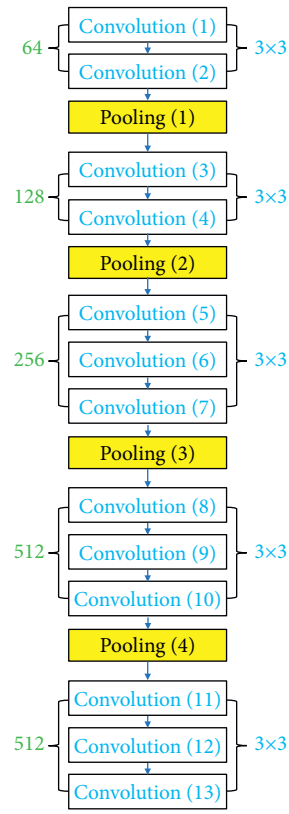


Figure 8: VGG-16 model structure diagram.

tested so that it can complete the task of automatically identifying the text detection frame of the reimbursement image. The experiment was run on a GeForce GTX 1080Ti GPU, and its video memory was about 11G.

Because text detection and text recognition use CTPN and DenseNet models respectively, the training process can be separated, and the test is performed sequentially. The specific parameters are set as follows.

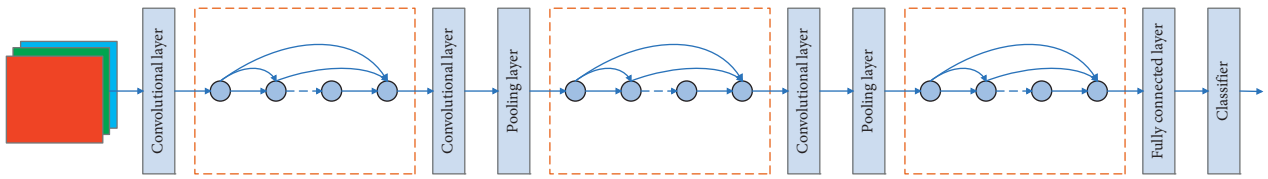


Figure 9: DenseNet model structure.

5.1.1. Input Image Size. When the CTPN model is trained, the image size is uniformly set to 900 × 900, that is, the width and height are both 600 pixels. When training the DenseNet model, the input image size is uniformly set to 280 × 32, in which the image width is 280 pixels, and the height is in pixels.

5.1.2. Training Batch. In this simulation experiment, a detection and recognition method of reimbursement information image information based on POLO-V3 model, CTPN model and DenseNet model is designed. In this section, the detection and recognition effect, accuracy and operating speed of the method are analyzed through experiments.

The CTPN model sets the maximum number of iterations to 50,000, and each batch trains 300 images. The DenseNet model sets the maximum number of training rounds to 10 rounds, and 128 images are trained in each batch.

5.1.3. Learning Rate. When the CTPN model is trained, the initial learning rate is 0.00001. When training 10,000 times, the learning rate is reduced to one-tenth of the current one. When the DenseNet model is trained, the initial learning rate is 0.0005, and every time it is trained, the learning rate is reduced to two-fifths of the current one.

5.1.4. Convolution Kernel. The specific structures of the CTPN model and the DenseNet model have been introduced in the previous section, and the specific convolutional layer parameter settings are detailed in the model structure introduction in the previous section.

5.1.5. Optimization Method. Both the CTPN model and the DenseNet model are trained using the Adam optimization method, that is, the adaptive moment estimation optimization algorithm.

5.1.6. Label Length. For the DenseNet model, since each image in the data set contains only 10 characters, the maximum length of its label is also set to 10.

5.2. Simulation Experiment Results and Analysis

5.2.1. Detection and Recognition Effect. This section will be based on the POLO-V3 model, C TPN model and DenseNet model of reimbursement image information detection and

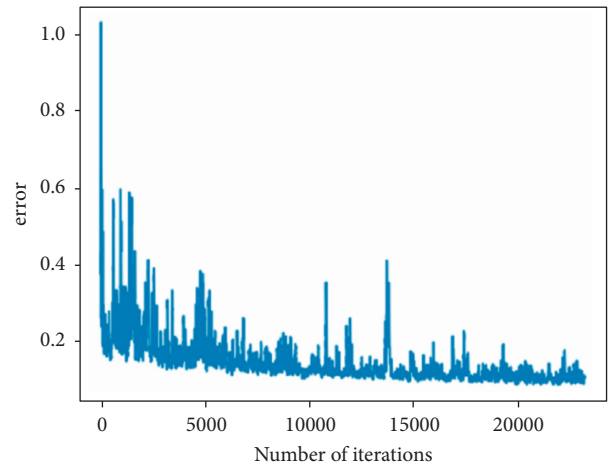


Figure 10: Training error graph.

recognition methods, through experiments to analyze the detection and recognition effect, accuracy and operating speed of the method. The input reimbursement information image is passed through the POLO-V3 model, the CTPN model and the DenseNet model in sequence, and the information area image, the text detection image and the text recognition result in the reimbursement information image are obtained respectively.

First, use the POLO-V3 model to perform target detection on the input reimbursement information image, detect the area where the relevant information is located in the reimbursement information image, and mark it with different colored rectangular boxes according to the category, and label the category label.

Then, based on the position information of the detected rectangular frame, the images of the relevant information area are respectively extracted.

Then, the CTPN model is used for text detection on the extracted information area images, and the positions of the text areas arranged in rows are obtained, and they are marked with a rectangular frame. It can be seen that the detection effect of the line text area is very good. For the "name" area, because the two characters are far apart, they are not all detected on the same line.

The DenseNet model is used for the recognition of the detected line text area image. The method designed in this chapter has a good effect on the detection and recognition of the reimbursement information image. It is very accurate except for the recognition of individual words.

5.2.2. Accuracy. In the process of detection and recognition of reimbursement information images, its accuracy is the

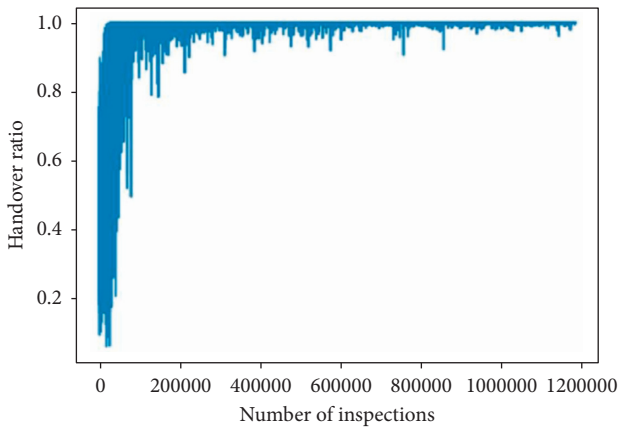


Figure 11: Training cross-combination diagram.

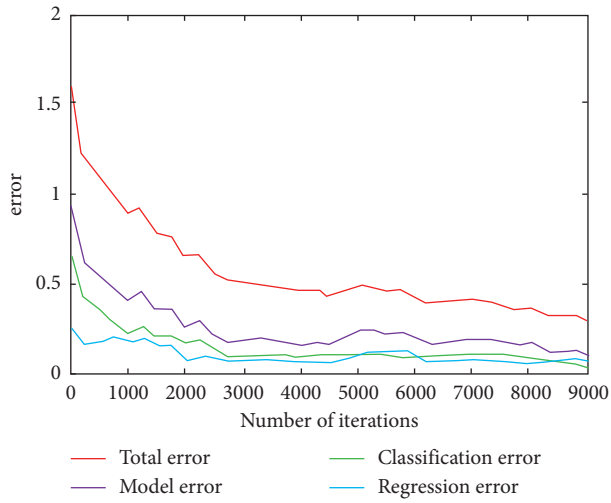


Figure 12: CTPN model training error diagram.

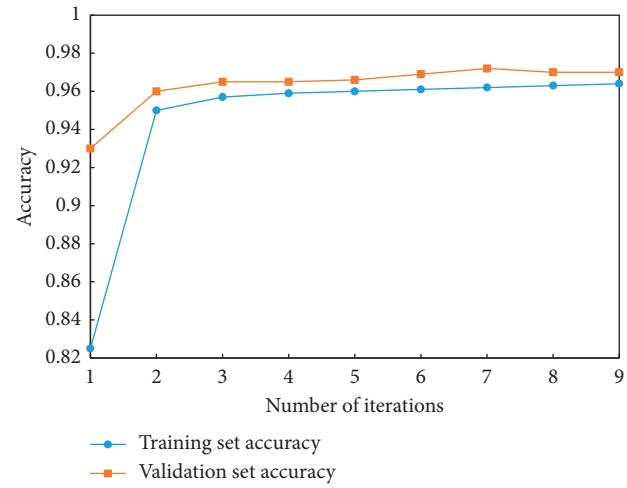


Figure 13: DenseNet model training accuracy curve.

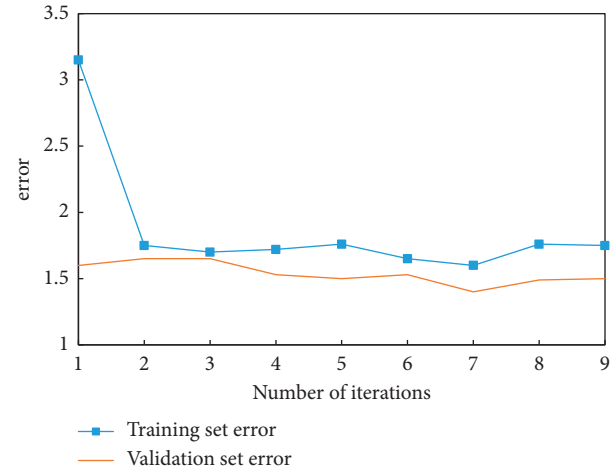


Figure 14: DenseNet model training error curve.

most important. This section conducts experiments and analyzes on the accuracy of detection and recognition of reimbursement information images.

For the problem of reimbursement information area detection, the training error loss and intersection ratio of the model are shown in Figures 10 and 11, respectively.

In the error loss diagram, the abscissa is the training batch, and the ordinate is the error loss. It can be seen that the final model converges and the error is small; The intersection ratio when the model converges is very high, close to 1, indicating that the accuracy of model detection is very high.

Since 16 images are trained in each batch, and the same image will be detected in three scales of large, medium and small, there is a multiple relationship between the number of detections and the training batch of 48, which can be obtained from the relationship between the abscissas of the two images Be confirmed.

For the recognition of the reimbursement information area, comparing the recognized characters with the original image of the reimbursement information, the accuracy rate is about 95.18%, and the recognition accuracy rate is better.

The error of CTPN model training is shown in Figure 12, where the overall error includes model error and horizontal offset error, and the model error includes classification error and regression error. It can be seen that after about 10,000 iterations of model training, the error curves all converge.

The accuracy and error of DenseNet model training are shown in Figures 13 and 14, respectively. It can be seen that the accuracy and error curves of the model converge after training, and the performance is very good.

5.2.3. Running Speed. In practical applications, high accuracy is often not enough, and a good user experience also has certain requirements for running speed. Since the detection and recognition methods of reimbursement information images are based on deep learning, the training of the model can be performed separately offline, so the training speed of the model is not considered, only the running speed of the tested model. The specific test results are shown in Table 1.

For the detection of the reimbursement information area, ten reimbursement information images were tested, and the average detection time for each reimbursement

Table 1: Running speed test.

	Detection time (s)	Recognition time per sheet (s)	Recognition time per line (s)
Image 1	0.68	1.346	0.056083
Image 2	0.63	1.312	0.50462
Image 3	0.66	1.264	0.046815
Image 4	0.63	1.107	0.04813
Image 5	0.71	1.233	0.047423
Image 6	0.59	1.273	0.05092
Image 7	0.67	1.126	0.048957
Image 8	0.62	1.09	0.047391
Image 9	0.68	4.192	0.182261
Image 10	0.7	1.578	0.068609
Average value	0.66	1.55	0.06

information was about 0.66 seconds. There are four information areas in each reimbursement information image that need to be detected, and the average detection of each area only takes about 0.17 seconds. Compared with the entire reimbursement information reimbursement process, the detection time is minimal, and it fully meets the requirements of real-time detection.

Therefore, the deep learning-based method designed in this section has a very fast running speed for the detection and recognition of reimbursement information images, and the detection and recognition of each area can meet the real-time requirements. In actual applications, the detection and recognition can be performed as needed. Modifying the identification area can greatly improve the efficiency of detection, while ensuring the compliance and effectiveness of reimbursement information, and enriching the data accumulation of the financial sharing service platform.

6. Conclusion

In order to transform and develop, enterprises need to use the financial sharing model to improve the implementation plan of internal control, so as to improve the implementation of informatization in financial management and deepen the goal of enterprise modernization and transformation. Therefore, this paper discusses the research on the validity detection and identification of travel reimbursement content and expense data in the enterprise financial management business. process design, the following conclusions can be drawn:

(1) According to the simulation of the training model, the CTPN model has a high success rate in detecting reimbursement vouchers. At the same time, the DenseNet model can effectively identify the detected content with high accuracy, which verifies the validity and feasibility of the model. sex.

(2) By testing the training speed of CTPN and DenseNnet, it is further verified that users will get a good experience in collecting fee vouchers independently.

(3) Based on the effective identification of expense reimbursement content and data, it provides guarantee and data support for the compliance determination of expense data information carried out by the Financial Shared Service Center, enabling the efficient operation of the service model of individual expense reimbursement.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] Y. Yang, Q. Liu, J. Song, and M. Zhou, "The influence mechanism of financial shared service mode on the competitive advantage of enterprises from the perspective of organizational complexity: a force field analysis," *International Journal of Accounting Information Systems*, vol. 42, Article ID 100525, 2021.

[2] Q. Yao, A. Au, T. Auang, and J. Hu, "TAe ResearcA of Implementing Enterprise Financial SAared Service Center Information System," in *Proceedings of the 2012 International Conference on Computing, Measurement, Control and Sensor Network*, pp. 388–391, Taiyuan, Shanxi, China, 2012.

[3] N. Dokoohaki and M. Matskin, "Reasoning about weighted semantic user profiles through collective Confidence analysis: a fuzzy evaluation," *Advances in Soft Computing*, vol. 67, pp. 71–81, 2010.

[4] B. Shi, X. Bai, and C. Yao, "An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 11, pp. 2298–2304, 2017.

[5] R Smith, "An overview of the Tesseract OCR engine," in *Proceedings of the International Conference on Document Analysis and Recognition(ICDAR)*, pp. 629–633, Curitiba, Brazil, September 2007.

[6] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[7] G. Hinton, L. Deng, D. Yu et al., "Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups," *IEEE Signal Processing Magazine*, vol. 29, no. 6, pp. 82–97, 2012.

[8] A Krizhevsky, I Sutskever, and G. E Hinton, "Imagenet classification with deep convolutional neural networks," in *Proceedings of the Advances in Neural Information Processing*

Systems, pp. 1097–1105, San Francisco, CA, United States, 2012.

[9] A Graves, A Mohamed, and G Hinton, "Speech recognition with deep recurrent neural networks," in *Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 6645–6649, Singapore, 2013.

[10] G. Hinton, "Deep belief networks," *Scholarpedia*, vol. 4, no. 5, p. 5947, 2009.

[11] C Szegedy, W Liu, Y Jia et al., "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1–9, Boston, MA, 2015.

[12] S. J. Kaplan, "Organization of behavior," *Yale Journal of Biology & Medicine*, vol. 23, no. 1, p. 79, 1950.

[13] F. Rosenblatt, "The perceptron: a probabilistic model for information storage and organization in the brain," *Psychological Review*, vol. 65, no. 6, pp. 386–408, 1988.

[14] E Karami, M Shehata, and A Smith, "Image Identification Using SIFT Algorithm: Performance Analysis against Different Image Deformations," 2017, http://arxiv.org/abs/1710.02728.

[15] H. M. Khan, "Automated breast cancer diagnosis using artificial neural network (ANN)," in *Proceedings of the Iranian Conference on Intelligent Systems and Signal Processing*, pp. 54–58, Shahrood, Iran, 2017.

[16] T. Li and H. Zhang, "Digital Image Enhancement System Based on MATLAB GUI," in *Proceedings of the International Conference on Software Engineering and Service Science (ICSESS)*, pp. 296–299, Beijing, China, 2017.

[17] A. Kumari, P. J. Thomas, and S. K. Sahoo, "Single image fog removal using gamma transformation and median filtering," in *Proceedings of the IEEE International Conference on INDICON*, pp. 1–5, Pune, India, 2014.

[18] G. A. Abdenova, "Scaling of Input-Output Data and Number of Conditionality of a Matrix of Grama," in *Proceedings of the Actual Problems of Electronics Instrument Engineering*, pp. 16–18, Novosibirsk, Russia, October 2012.

[19] S. H. Songe, M. Antonelli, T. W. K. Fung, B. D. Armstrong, A. Chong, A. Lo, and B. E. Shi, "Developing and assessing MATLAB exercises for active concept learning," *IEEE Transactions on Education*, vol. 99, pp. 1–9, 2018.

[20] K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask R-CNN," in *Proceedings of the IEEE International Conference on Computer Vision (ICC)*, pp. 2961–2969, Cambridge, MA, USA, 2017.

[21] E. Meng, S. Huang, Q. Huang, W. Fang, L. Wu, and L. Wang, "A robust method for non-stationary streamflow prediction based on improved EMD-SVM model," *Journal of Hydrology*, vol. 568, pp. 462–478, 2019.

[22] J Redmon, S Divvala, R Girshick, and A. Farhadi, "You only look once: unified, real-time object detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779–788, NV, USA, June 2016.

[23] J Redmon and A Farhadi, "YOL09000: better, faster, stronger," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7263–7271, Honolulu, Hawaii, 2017.

[24] J Redmon and A. Farhadi, "Yolov3: An Incremental Improvement," 2018, http://arxiv.org/abs/1804.02767.

Hindawi
Scientific Programming
Volume 2022, Article ID 6385972, 14 pages
https://doi.org/10.1155/2022/6385972

Research Article

Research on the Agricultural Machinery Path Tracking Method Based on Deep Reinforcement Learning

Hongchang Li,[1] Fang Gao,[1] and GuoCai Zuo [1,2]

[1]*Changzhou Vocational Institute of Mechatronic Technology, College of Vehicle Engineering, Changzhou, Jiangsu 213164, China*
[2]*Hunan Software Vocational and Technical University, Hunan Xiangtan 411100, China*

Correspondence should be addressed to GuoCai Zuo; zuoguocai@hnsoftedu.com

Received 19 December 2021; Revised 4 January 2022; Accepted 13 January 2022; Published 22 March 2022

Academic Editor: Baiyuan Ding

With the rapid development of information technology, industry and service industries have achieved rapid development in recent years. Then, looking at the development of agriculture, the popularity of informatization lags far behind industry and service industries, directly hindering the digital development of agriculture. Starting from the current agricultural machinery driving operation scene, this paper carried out a simplified research on the traditional agricultural machinery driving operation method through the agricultural machinery kinematics model, and based on the related theory of deep reinforcement learning to study the agricultural machinery path tracking in the agricultural operation scene, it carried out the controller design, built the agricultural machinery autonomous path tracking framework operating mechanism under deep reinforcement learning, and further researched through experimental design and found that the agricultural machinery autonomous path tracking control can achieve better automatic control after empirical learning. I-DQN algorithm enables agricultural robots to adapt to the environment faster when performing path tracking, which improves the performance of path tracking. It has important guiding significance for further promoting the automatic navigation and control of agricultural machinery to realize the efficient operation of agricultural mechanization.

1. Introduction

Automatic navigation control of agricultural machinery is a key technology to support precision agriculture. This technology can improve the working accuracy and efficiency of agricultural machinery, so that the driver can get rid of long-time tired and repetitive driving work and have enough time to monitor and operate agricultural machinery. Therefore, the automatic navigation control of agricultural machinery has broad development prospects.

The path tracking methods that are at the core of the automatic navigation control of agricultural machinery mainly include model-based control methods and model-independent control methods. In terms of model-based control methods, related scholars have separately studied the path following control methods based on the kinematics model and dynamics model of agricultural machinery [1–9]. However, among these methods, the method based on the kinematics model is mainly to approximate the model with a small angle linearization and design the controller under the assumption of constant speed. This introduces not only linearization error, but also the controller's performance when the speed changes. Robustness also deteriorates; while the control method based on the dynamic model can fully consider the dynamic characteristics of agricultural machinery, the dynamic model parameters are difficult to obtain online and in real time. In terms of model-independent control methods [10–15], the online adaptive determination of the forward-looking distance in the pure tracking method has not been well solved although the intelligent method has some human-like intelligence and incomparable traditional control methods. It has linear mapping ability, but its design requires certain experience knowledge and complex learning and training process. Aiming at the outstanding advantages of intelligent methods in agricultural machinery path control, this paper proposes

an agricultural machinery path tracking method based on deep reinforcement learning. The research of this method has certain practical significance for the development of efficient agricultural operation methods.

2. Related Theories

The research in this paper will use the deep combination of reinforcement learning and deep learning and make full use of the decision-making advantages of reinforcement learning and the perceptual advantages of deep learning [13, 16, 17] to carry out research. In deep reinforcement learning, reinforcement learning is used to define problems and optimization goals, deep learning is used to solve strategy functions or value functions, and backpropagation algorithms are used to optimize the objective function. To a certain extent, deep reinforcement learning has general intelligence to solve complex problems.

2.1. Deep Learning. Deep learning is derived from the idea of artificial neural network, which combines low-level features to form higher-level features and attribute categories. The most basic unit of artificial neural network is neuron, also known as perceptron. A deep neural network is called a multilayer perceptron. The difference from a single-layer perceptron is that it adds multiple hidden layers and can have multiple outputs. In the hidden layer, more complex feature information can be learned and multiple values can be output. It also enables the neural network model to solve more types of problems, such as classification, regression, dimensionality reduction, and clustering. At the same time, combining deep neural networks with different activation functions can further enhance the expressive ability of the model [13, 16, 17].

The deep neural network model is shown in Figure 1. The structure can be divided into input layer, hidden layer, and output layer. The input layer refers to information obtained through sensors or from the environment, such as radar data of agricultural intelligent harvesting vehicles. Each hidden layer is a feature level, in which each neuron represents a feature attribute. The output of the output layer is the required variables, such as the angular velocity and linear velocity of agricultural intelligent harvesting vehicles.

In DNN, each layer of neural network is fully connected; that is, the neurons in each i+1 layer are connected by the second layer of neurons. Assume that there are m neurons in the $l-1$ layer network, and W_{ij} represents the weight between the jth neuron in layer 1 and the kth neuron in $l-1$ layer, b_j^l is the bias of the kth neuron in the lth layer, and $\sigma(z)$ is the activation function. Then, for the output a_j^l of the jth neuron of the lth layer, there are

$$\begin{aligned} a_j &= \sigma\left(z_j^l\right) \\ &= \sigma\left(\sum_{k=1}^{m} w_{jk}^l a_k^{l-1} + b_j^l\right). \end{aligned} \tag{1}$$

The above process is the forward propagation of the neural network, but to optimize the parameters of the neural network, backpropagation is required. In order to calculate the error between the model output and the real training sample output, the neural network needs to first define the loss function for training, as defined in

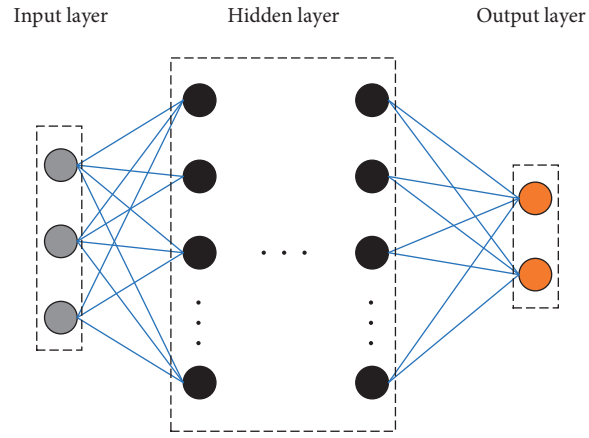


Figure 1: Deep neural network model.

$$J(W, b, x, y) = \frac{1}{2} a^L - y_2^2. \tag{2}$$

Finally, the error is used to update the weight of each neuron, and finally a better model is obtained. This is the process of backpropagation of the deep neural network.

2.2. Reinforcement Learning. Different from deep learning that focuses on perception and expression, reinforcement learning focuses on finding problem-solving strategies [18, 19]. Reinforcement learning is mainly composed of agent and environment. Since the interaction between the agent and the environment is similar to the interaction between the organism and the environment, it can be considered that reinforcement learning is a general learning framework, which represents the future development trend of general artificial intelligence algorithms [20, 21].

The basic framework of reinforcement learning is shown in Figure 2. Agents interact with the environment through states, actions, and rewards. Suppose that the state of the environment at time t in Figure 2 is denoted as s_t, and the agent performs a certain action a_t in the environment. At this time, the action a_t changes the original state of the environment and makes the agent reach a new state s_{t+1} at time t+1. In the new state, the environment generates a feedback reward r_t to the agent. The agent performs a new action a_{t+1} based on the new state s_{t+1} and the feedback reward r_{t+1} and iteratively interacts with the environment through feedback signals [22].

The ultimate goal of the above process is to maximize the cumulative reward for the agent. Equation (3) is the calculation process of the cumulative reward G.

$$G = r_1 + r_2 + \cdots + r_n. \tag{3}$$

In the above process, the rule of selecting actions according to the state s and the reward r is called the strategy

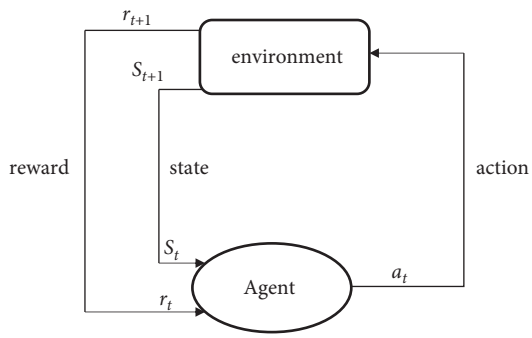


Figure 2: The basic framework of reinforcement learning.

π, where the value function v is the expectation of the cumulative reward.

Reinforcement learning is to continuously perform trial-and-error learning according to the feedback information of the environment and then adjust and optimize its own state information. The purpose is to find the optimal strategy or the maximum reward.

There are two types of environments in which an agent is located [23]: one is that the environment is known, which is called model-based; the other is that the environment is unknown, which is called model-free.

The relationship between model-based tasks and model-free tasks is shown in Figure 3. The line following agricultural robot shown in Figure 3(a) controls its walking by sensing the black course on the ground through sensors. Since the black route on the ground is planned in advance and the surrounding environment is also controllable and known, it can be regarded as a model-based task. Figure 3(b) is the autopilot system of a car. In the real traffic environment, many things cannot be estimated in advance, such as the behavior of passers-by, the trajectory of passing vehicles, and other emergencies, so it can be regarded as a model-free task.

2.3. Deep Q-Learning (DQN) Algorithm. The DQN algorithm is a famous work of the Google DeepMind team. They used reinforcement learning to propose a deep learning network model for solving control strategy problems, opening a new era of deep reinforcement learning [24–27].

The Q-learning algorithm stores Q values in the form of Q tables, as shown in Figure 4. This method of storing the Q value can handle maze problems when the state space and action space are very small, but when the problem has a large action or state space, the method of applying the Q table will cause a very large amount of data. The DQN algorithm combines Q-learning and deep learning algorithms, using deep convolutional nerves as shown in Figure 5.

The DQN algorithm has made the following improvements on the basis of the reinforcement learning algorithm:

(1) DQN uses a deep neural network to simulate the Q value function. The value function here corresponds to the weight θ of each layer in the convolutional neural network, that is, $Q(s,a;\theta) \approx Q^{\pi}(s,a)$. In this way, the update process of the Q value function is essentially an update of the weight θ of the neural network. When the parameter θ of the neural network is determined, the value function Q is also determined.

(2) Use experience playback technology to train neural networks. The deep neural network used by DQN is a supervised neural network model. The input data needs to be independent of each other and meet the same distribution. Since the data collected by the agent in the environment is continuous, there is a correlation between adjacent data. When the algorithm uses a set of continuous data for training, the direction of gradient descent will become the same. Calculating the gradient under the same training step size may cause the result to not converge. The experience playback mechanism puts the data collected by the agent into a memory bank, then uniformly randomly samples from the memory bank, and extracts the data from it for neural network training. By using experience replay, the behavior distribution can be averaged in its many previous states, thus smoothing the learning process and avoiding fluctuations or divergence of parameters. At the same time, assign priority to each conversion in the experience replay memory, which can greatly improve the learning efficiency compared with the uniform sampling from the experience replay memory.

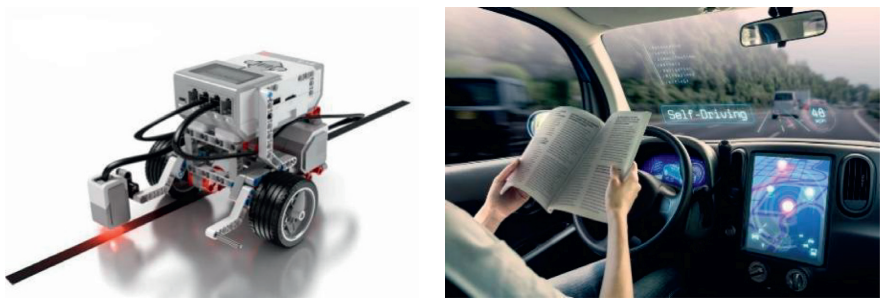
(3) The Q target network is set up to calculate the TD error. When using the convolutional neural network to approximate the Q value network, the parameter θ is processed by the gradient descent method, and the update process is

$$\theta_{i+1} = \theta_i + \alpha r + \gamma \max_{a'} Q(s', a'; \theta) - Q(s, a; \theta) \nabla Q(s, a; \theta). \tag{4}$$

In (4), $r + \gamma \max_{a'} Q(s', a'; \theta)$ is called the TD target, and the network used in calculating the TD target is called the target network. The neural network used to approximate the Q value function is called the estimation network. From the above formula, it can be seen that the parameters used by the target network are the same as the parameters of the estimated Q network, so that the results obtained by the calculation will have relevance. The training results of reinforcement learning are unstable. To solve this problem, the DQN algorithm expresses the parameters of the target network as θ^{-}. In the update of the neural network, the parameter θ of the estimated network is updated in real time, and the parameter θ^{-} of the target network is obtained by assigning the parameters of the estimated network to the target network after N rounds of iterations, so (4) changes to

$$\theta_{i+1} = \theta_i + \alpha r + \gamma \max_{a'} Q(s', a'; \theta^{-}) - Q(s, a; \theta) \nabla Q(s, a; \theta). \tag{5}$$

In the update of the neural network, the loss function is defined by the mean square error:



(a) (b)

Figure 3: Specific examples of model-based and model-free. (a) Line patrol agricultural machinery robot. (b) Autopilot system.

	a1	a2	a3	a4
s1	Q (1, 1)	Q (1, 2)	Q (1, 3)	Q (1, 4)
s2	Q (2, 1)	Q (2, 2)	Q (2, 3)	Q (2, 4)
s3	Q (3, 1)	Q (3, 2)	Q (3, 3)	Q (3, 4)
s4	Q (4, 1)	Q (4, 2)	Q (4, 3)	Q (4, 4)

Figure 4: Q table of Q-learning algorithm.

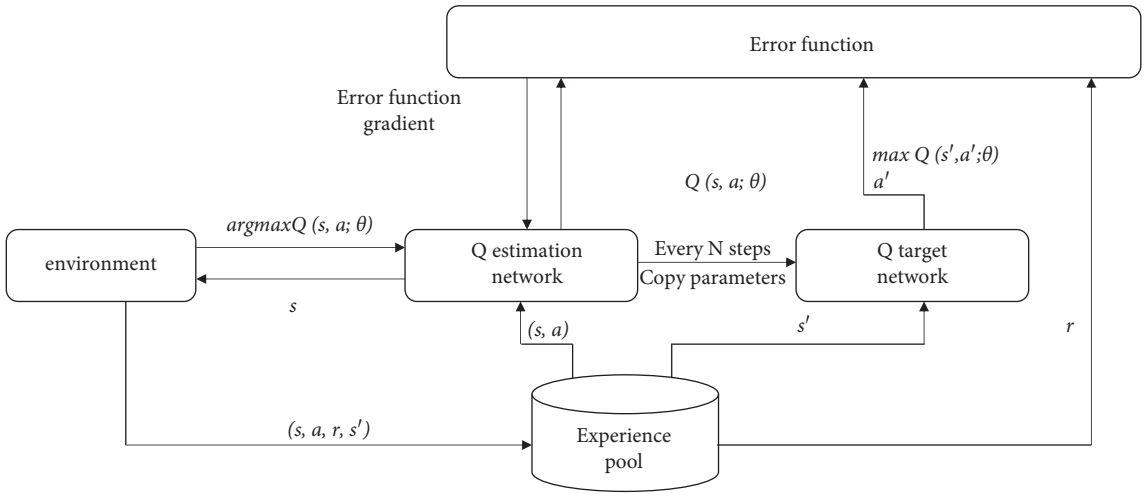


Figure 5: DQN algorithm principle diagram.

$$L(\theta) = E\left[\left(r + \max_{a'} Q(s', a'; \theta^{-}) - Q(s, a; \theta)\right)^2\right]. \quad (6)$$

Error function gradient:

$$\nabla L(\theta) = E\left[\left(r + \max_{a'} Q(s', a'; \theta^{-}) - Q(s, a; \theta)\right)\nabla Q(s, a; \theta)\right]. \quad (7)$$

After updating the network of (7) and obtaining the value of $Q(s, a; \theta)$, you can use $\nabla Q(s, a; \theta)$ to obtain the optimal Q value for the nerve of (5).

3. Agricultural Machinery Kinematics Model

3.1. Behavioral Learning Theory. Considering the application of agricultural machinery in actual agricultural land, agricultural machinery should have high flexibility and stability in complex environments. Therefore, this paper adopts a four-wheel agricultural machinery movement model, which provides power for the agricultural machinery movement through two rear wheels. The two front wheels adopt different steering angles to ensure the smooth steering of the mobile agricultural machinery. The movement model is shown in Figure 6.

When the agricultural machinery system is turning, its turning process can be simplified into a bicycle model as shown in Figure 7.

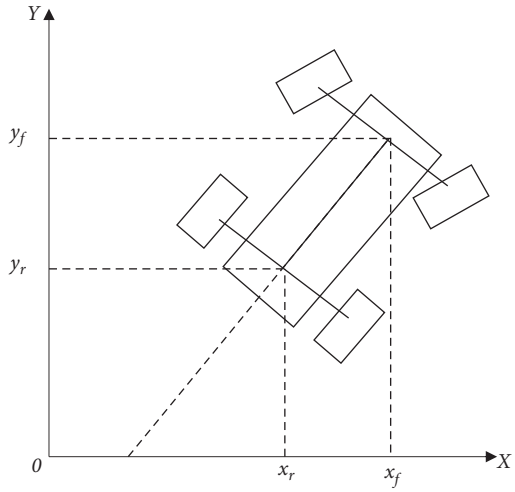


Figure 6: Schematic diagram of agricultural machinery steering.

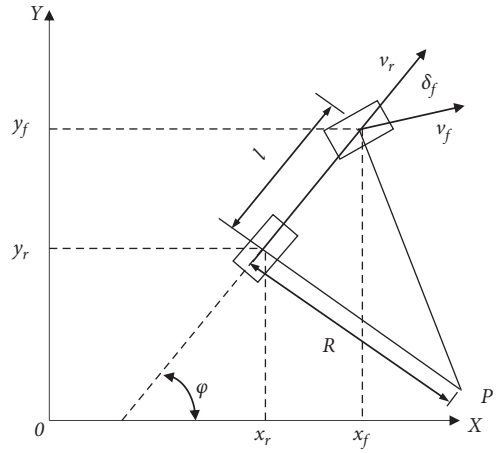


Figure 7: Model of the agricultural machinery steering bicycle.

In the map coordinate system, (x_r, y_r) and (x_f, y_f), respectively, represent the coordinates of the center position of the two rear wheels of the agricultural machine and the coordinates of the center position of the two front wheels, and v_r and v_f, respectively, represent the center position of the front wheel and the center of the rear wheel of the agricultural machine. The speed of the position, φ, is the heading angle of the agricultural machine in the map coordinate system, δ_f is the deflection angle of the front wheel of the agricultural machine, and l is the distance between the center position of the front wheel and the center position of the rear wheel. P is the instantaneous turning center of the rear wheel center position of the agricultural machinery during the turning process; R is the turning radius of the center point of the rear wheel of the agricultural machinery, assuming that the deflection angle of the center of mass of the moving agricultural machinery does not change during the turning process; that is, the instantaneous turning radius and the radius of curvature of the path are the same. Then the speed of the rear wheel center (x_r, y_r) of the agricultural machinery is v_r:

$$v_r = \dot{x}_r \cos\varphi + \dot{y}_r \sin\varphi. \tag{8}$$

I also know the kinematic constraints of the center of the front and rear wheels of agricultural machinery:

$$\begin{cases} \dot{x}_r \cos\varphi - \dot{y}_r \sin\varphi = 0, \\ \dot{x}_f \sin\left(\varphi + \delta_f\right) - \dot{y}_f \cos\left(\varphi + \delta_f\right) = 0. \end{cases} \tag{9}$$

Combining (8) and (9) can get

$$\begin{cases} \dot{x}_r = v_r \cos\varphi, \\ \dot{y}_r = v_r \sin\varphi. \end{cases} \tag{10}$$

According to the relationship between the center coordinates of the rear and front wheels (x_r, y_r) and (x_f, y_f):

$$\begin{cases} x_f = x_r + l\cos\varphi, \\ y_f = y_f + l\sin\varphi. \end{cases} \tag{11}$$

Incorporating (10) into (11) can reach the angular velocity ω when the agricultural machinery turns:

$$\omega = \frac{v_r}{l}\tan\delta_f. \tag{12}$$

ω is the angular velocity at which the agricultural machinery rotates around the instantaneous rotation center P. And the moving speed of the agricultural machinery v_r can get the turning radius R and the front wheel deflection angle δ_f:

$$\begin{cases} R = \dfrac{v_r}{\omega}, \\ \delta_f = \tan^{-1}\left(\dfrac{l}{R}\right). \end{cases} \tag{13}$$

Finally, the kinematics model of mobile agricultural machinery can be obtained as

$$\begin{bmatrix} \dot{x}_r \\ \dot{y}_r \\ \dot{\varphi} \end{bmatrix} = \begin{bmatrix} \cos\varphi \\ \sin\varphi \\ \tan\dfrac{\delta_f}{l} \end{bmatrix} v_r. \tag{14}$$

4. Design of the Control Strategy for the Agricultural Machinery Path following Deep Reinforcement Learning

4.1. Design of the Autonomous Path Tracking Framework for Agricultural Machinery. When agricultural machinery uses reinforcement learning to achieve autonomous path tracking and obstacle avoidance tasks in unknown environments, it must first meet the MDP model. When using MDP, it is necessary to define the state space, action space, and reward and punishment functions. When the agricultural machinery first interacts with the environment, it cannot distinguish between obstacles and targets. It can only

adjust its own strategy according to the reward and penalty values feedback from the environment in the process of exploring the environment and finally realize the task of path tracking. The framework of the path tracking algorithm is shown in Figure 8. The Autolabor four-wheeled vehicle is used to simulate the operation of agricultural machinery in the design of the path tracking framework.

In the above framework, the agricultural machinery obtains external information through the Lidar sensor and executes action a, tries different states S_t, and at the same time obtains the corresponding reward value r according to the set reward and punishment function. When exploring the environment, OU noise is added to increase the exploration degree of the action space, and the experience explored in the environment is stored in the form of tuples and placed in the experience playback pool. When training the network, the priority playback mechanism is used to sample and learn the important experience samples first, reducing the training time of the mobile agricultural machine, and finally the mobile agricultural machine learns to track autonomously in the environment. The following will design the state space, action space, and reward and punishment functions in the algorithm framework.

4.1.1. Agent State and Space Design. In order to simplify the path tracking model, it is assumed that the agricultural machine is moving at a fixed speed; that is, the agricultural machine has a fixed moving distance in each time step, so the steering angle φ of the machine is taken as the action space, and the dimension is 1.

In deep reinforcement learning training, the purpose of agricultural machinery is to move to the target path while avoiding obstacles. Therefore, the state space of agricultural machinery needs to include its own positional relationship with obstacles and target paths. This article defines the state space of agricultural machinery as follows:

$$S = \begin{cases} \dfrac{(x, y)}{k}, \\ \dfrac{\theta}{2\pi}, \\ \dfrac{d_{\text{obj}}}{k}, \\ \dfrac{\left(\left(x - x_{\text{obj}}\right), \left(y - y_{\text{obj}}\right)\right)}{k}, \\ \dfrac{d_{\text{aim}}}{k}, \\ \dfrac{\left(\left(x - x_{\text{aim}}\right), \left(y - y_{\text{aim}}\right)\right)}{k}. \end{cases} \tag{15}$$

Among them, (x, y) and θ represent the position and orientation of the agricultural machine in the current map, and k is the standardized coefficient; d_{obj} and d_{aim} represent the distance between the agricultural machine and the nearest obstacle and the target path; $(x - x_{\text{obj}})$, $(y - y_{\text{obj}})$ and $(x - x_{\text{aim}})$, $(y - y_{\text{aim}})$, respectively, represent the distance information of the agricultural machinery from the nearest obstacle and the target path.

In actual movement, the real-time pose of the agricultural machine in the environment can be obtained through SLAM technology, and the distance between the agricultural machine robot and the obstacle is obtained through sensors.

4.1.2. Reward Function Design. The reward and punishment value is the feedback signal given to the agent by the environment, which reflects the pros and cons of the actions performed by the agent during the task learning process. When the agricultural machinery obtains a higher reward value from the environment, it indicates that the current behavior of the agricultural machinery is more conducive to the path tracking task; on the contrary, if the mobile agricultural machinery receives a large penalty value in the environment, it means that the behavior performed by the mobile agricultural machinery is not good for the path tracking task and should be avoided as much as possible. Finally, the mobile farming opportunity adjusts its strategy according to the rewards and punishments in the environment. During the training of mobile agricultural machinery, when the agricultural robot reaches the target point or touches obstacles and walls, the agricultural robot is given a fixed reward. When the agricultural robot has not reached the target or touched an obstacle, the reward value contains two parts: one is the negative reward value of the distance information between the agricultural machine and the nearest obstacle; the second is the positive reward value of the distance information between the agricultural robot and the target path. The sum of the two parts of the reward value is used as the final reward value obtained by the agricultural robot after each action, set as follows:

$$\text{reward} = \text{reward}_{\text{att}} + \text{reward}_{\text{rep}} = \frac{1}{2^{d_{\text{aim}}/k}} - \frac{1}{2^{d_{\text{obj}}/k}}. \tag{16}$$

Therefore, the reward function of agricultural machinery action is

$$\begin{cases} 200 & d_{\text{aim}} < d_{\text{obj}}, \\ \dfrac{1}{2^{d_{aim}/k}} - \dfrac{1}{2^{d_{obj}/k}} & d_{\text{aim}} < d < d_{\text{obj}}, \\ -200 & d_{\text{aim}} > d_{\text{obj}}. \end{cases} \tag{17}$$

The rewards in the above reward and punishment function are divided into continuous rewards and instant rewards. Continuous rewards are rewards that are generated every time the agricultural robot takes an action; that is, rewards are rewards that are given immediately under certain circumstances.

4.1.3. Design of Autonomous Path Tracking Control for Agricultural Machinery. The path tracking process design of mobile agricultural machinery under the deep

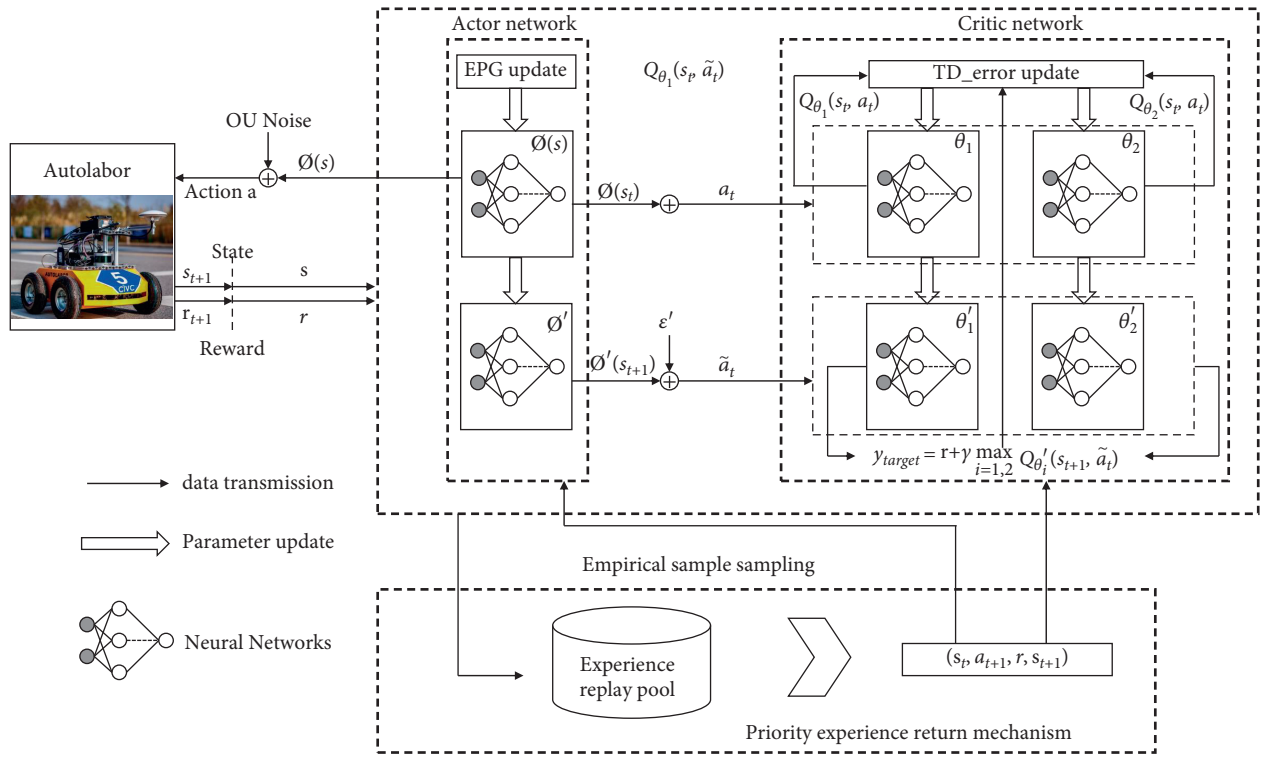


Figure 8: Agricultural machinery path tracking framework based on Autolabor.

reinforcement learning algorithm is shown in Figure 9. The agricultural machinery first obtains environmental information through sensors and calculates the orientation and distance of obstacles and targets and selects the corresponding action value according to the exploration noise and exploration attenuation rate. At the same time, it is judged whether it is the end state or the target state. If it is the end state, reset the environment and restart; otherwise continue to learn in the environment; if it is the target state, continue to judge whether the algorithm has converged; if it converged, the program ends; otherwise continue to generate target endpoints and interact with the environment until the end.

4.2. Deep Neural Network Structure Design. The deep neural network of agricultural machinery is based on the Actor-Critic framework. In the current state, the mobile agricultural machinery obtains and executes actions through the Actor network and interacts with the environment to reach the next state and obtain reward values. At this time, the Critic network takes the actions and state values output in the Actor network as input and outputs the evaluation of the current action value. This evaluation indicates the pros and cons of the action value of the mobile agricultural machine in the current state. The structure design of the network is shown in Figure 10.

In the Actor network, the input is the state S of the agricultural machinery robot. The number of neurons in the hidden layer is 400 and 300, the activation function is Relu, and the output layer is the linear velocity v and angular velocity v of the mobile agricultural machinery. Since the retreat of agricultural machinery is not considered, the linear velocity w has only positive values, and the angular velocity$w$is a vector, and the positive and negative values indicate the direction, so the Sigmoid and Tanh activation functions are used to output the action values in the continuous action space. In the Critic network, the hidden layer uses the same number of neurons and activation function as the Actor network. The Q value of the output layer does not require an activation function to perform a nonlinear transformation and directly performs a linear transformation. Finally, the smallest Q value is selected from two Critic networks of the same structure to avoid overestimation of the deviation. According to the set reward and punishment mechanism, network parameters will be continuously optimized, so that the Actor network can get a higher reward value after performing actions. In the Critic network, the value calculated in the Actor network is scored, and the score result is sent back to the Actor network. The Actor network will update according to the score result. The combination of the two networks can improve the efficiency of algorithm update.

5. Experimental Design and Results

5.1. Simulation Environment Settings. This chapter will adopt the mobile agricultural machinery model. Autolabor is a ROS-based mobile four-wheeled vehicle instead of agricultural machinery. It has programmable, SLAM mapping navigation, and motion control functions. At the same time, Autolabor software is also provided in open source form. In RVIZ, the models of agricultural machinery robots are

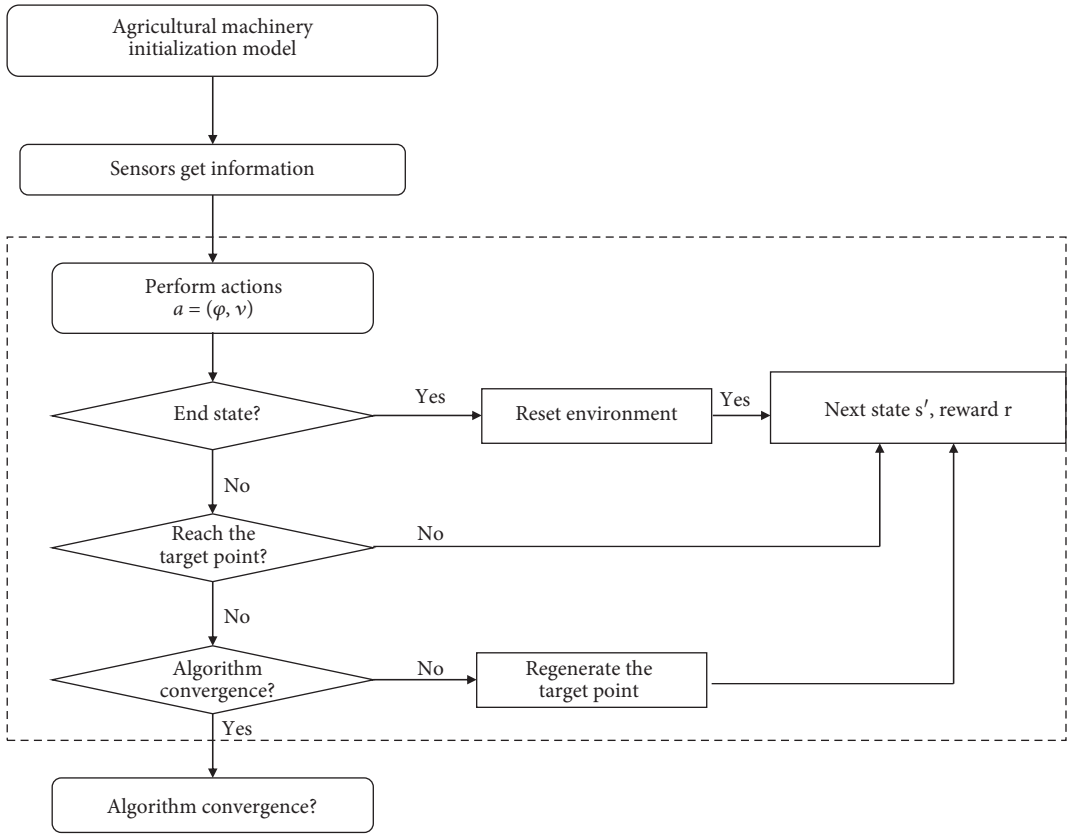


Figure 9: Design of autonomous path tracking control for agricultural machinery.

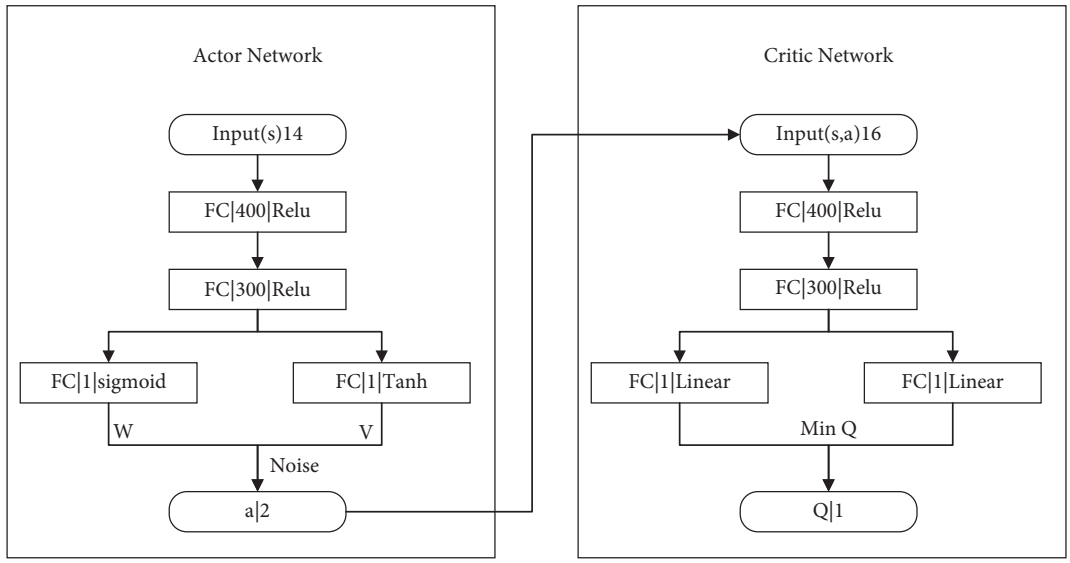


Figure 10: Actor-Critic network structure design.

commonly described in URDF and XARCO files, and their essence is in XML format. Autolabor's model files are shown in Figure 11.

After the model is built, start the model for testing. Create the file display.launch in the launch folder Figure 12. The first input parameter model is the path to the urdf file to be launched. The two input parameters gui specify whether to enable the joint rotation control panel window. Two parameters indicate describing the model description file to be started (urdf) and the joint to the control window (gui, corresponding to each joint), respectively. Three nodes are used to send joint information, robot control information, and rviz start.

Among them, Link and Joint can be compared to human skeletons and joints, which are the basis for describing the

```
<robot name="autolabor_description">
  <link name="base_link">
  <inertial>
    <origin
      xyz="0. 0. 0."
      rpy="0. 0. 0." />
    <mass
      value="0.251988675650349" />
    <inertia
      ixx="0.000595579869264794"
      ixy="5.99238175321912E-08"
      ixz="-1.98242615307314E-08"
      iyy="0.00102462329604677"
      iyz="-1.73115625503396E-05"
      izz="0.00060561972360446" />
  </inertial>
  <visual>
    <origin
      xyz="0. 0. 0.05"
      rpy="1.57 0. 1.57" />
    <geometry>
      <mesh
        filename="package://autolabor_description/meshes/base_link.stl" />
    </geometry>
    <material
      name="">
      <color
        rgba="0.792156862745098 0.819607843137255 0.933333333333333 1" />
    </material>
  </visual>
</link>
</robot>
```

Figure 11: Autolabor model file.

```
<launch>
  <arg name="model" />
  <arg name="gui" default="false" />

  <param name="robot_description" textfile="$(find autolabor_description)/urdf/autolabor_description.urdf" />
  <param name="use_gui" value="$(arg gui)" />

  <node name="joint_state_publisher" pkg="joint_state_publisher" type="joint_state_publisher" />
  <node name="robot_state_publisher" pkg="robot_state_publisher" type="state_publisher" />
  <node name="rviz" pkg="rviz" type="rviz" args="-d $(find autolabor_description)/urdf.rviz" />
</launch>
```

Figure 12: Model startup file.

structure of agricultural machinery and agricultural machinery robots and are constructed in a tree structure. The main body, wheels, and joints of the agricultural machinery and agricultural machinery robot are defined in the link, and some attributes are given: <visual> defines the appearance attributes of the link; <geometry> defines the shape of the structure; <inertial> and <collision> specify, respectively, inertial properties and collision properties. The final Autolabor model in RVIZ is shown in Figure 13.

Next, create a topographic map of agricultural land based on the topographic characteristics of agricultural land, as shown in Figure 14.

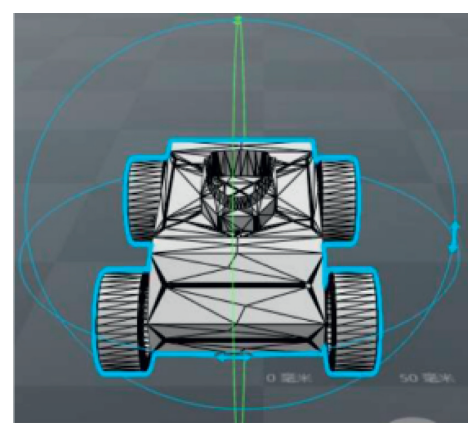


Figure 13: Autolabor model.

5.2. Training Process and Experimental Parameter Settings

5.2.1. Pretest Results and Analysis of Physical Fitness. When the mobile agricultural machinery is undergoing training experiments, it is essentially a process in which the agricultural robot explores the environment and adjusts its action strategy according to the feedback of the environment and finally realizes the path tracking and obstacle avoidance of the agricultural robot. During the training of agricultural robots, the starting point is the starting point, and the target end point is randomly generated in the set simulation environment. The same coordinate range as the obstacle collision area cannot be set as the target end point. When the agricultural machinery robot reaches the target, it means that it has successfully completed a path tracking task and uses this point as the starting point to continue to the next randomly generated target end position. When an agricultural robot fails to track the path, it is regarded as a terminal state. The terminal state includes that the

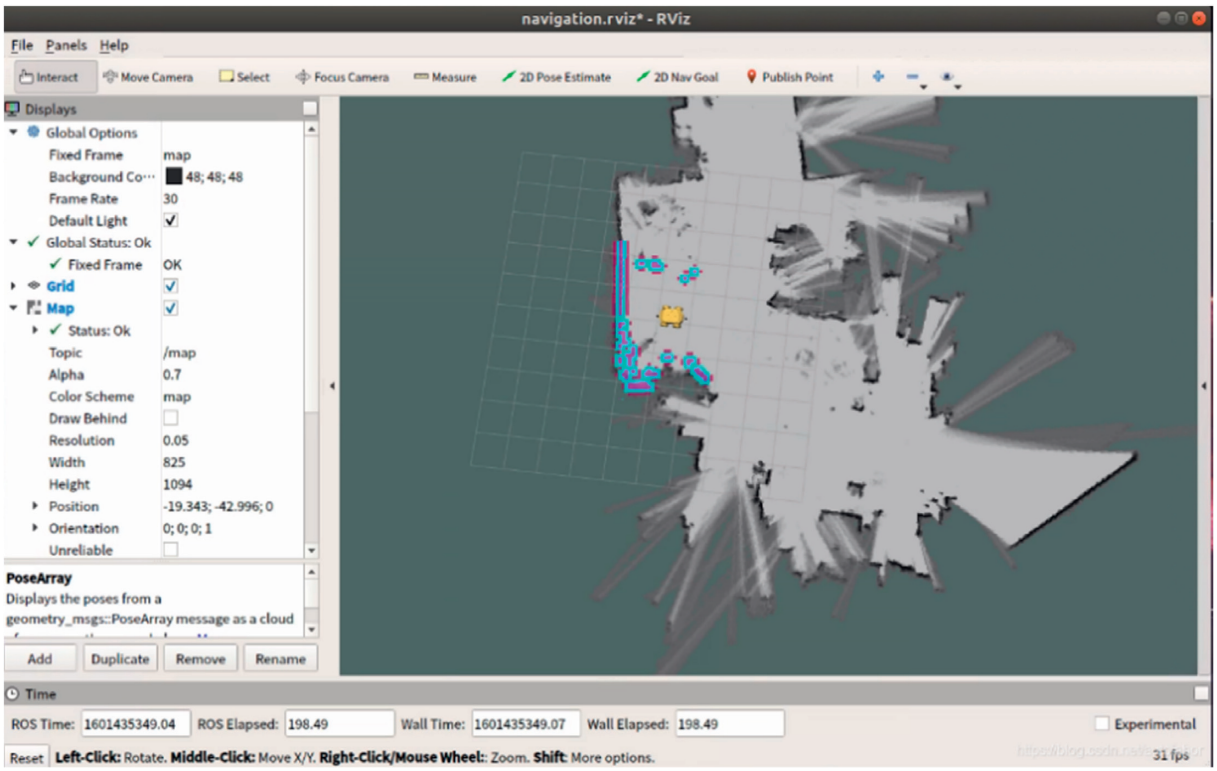


Figure 14: Topographic map of simulated agricultural land.

agricultural robot encounters an obstacle, a wall, or reaches the upper limit of the planned number of steps. At this time, the agricultural robot will start the next training from the planned starting point. Finally, the training is completed after reaching the set maximum number of training rounds. The training process is shown in Figure 15.

5.2.2. Experimental Parameter Settings. In order to improve the reliability of the experimental data, the experiments in this chapter are all completed under the environment ubuntu 6.04+cuda9.0+pytorch0.4.1, and the experimental hardware conditions are i7-8750H + GeForce GTX1060 + 16G. The specific settings of the experimental parameters are shown in Table 1.

5.3. Experiment and Result Analysis. In this section, experiments will be conducted on static obstacle scenes and dynamic obstacle scenes, respectively. In each scene, the path tracking results of the DQN algorithm and the agricultural robot proposed in this paper will be tested, and the results will be analyzed.

5.3.1. Static Obstacle Experiment. The reward value of the first 1000 training rounds is plotted as a reward curve, as shown in Figure 16. In the initial stage of training, since the agricultural robot has just started to interact with the environment, it will often drive away from the target and finally collide with obstacles or walls, so the penalty value is high, and the initial reward is basically around -500 to -400. In the 200 rounds before training, because the DQN algorithm cannot distinguish the importance of experience, it can only continue to explore and try to learn, and the curve fluctuates greatly. The agricultural machinery algorithm uses a priority playback mechanism, which will give priority to learning some important experiences. Compared with the DQN algorithm, it reduces the volatility of the curve, and the I-DQN algorithm starts to accelerate the convergence in about 100 rounds; however, the DQN algorithm does not start to increase the rewards until 250 rounds. As the training time increases, the I-DQN algorithm basically converges after 300 rounds, and the DQN algorithm gradually converges around 450 rounds. Therefore, in scenario 1 under the same training conditions, the I-DQN algorithm has better convergence and stability.

Figure 17 shows the path tracking success rate in the 1000 rounds before the training of the agricultural robot. The trend line of the success rate and the reward value curve are roughly the same. The I-DQN algorithm starts around 100 rounds, and the success rate is greatly improved, reaching 70% in 200 rounds. In about 300 rounds, the success rate of the I-DQN algorithm basically reached 90%; in contrast, the DQN algorithm had fewer successes in the early stage and lacked stability. In 200 rounds of training, there was only a 50% success rate until after 450 rounds. The success rate has gradually reached 90%. Therefore, the importance area of the experience samples in the experience pool can make the agricultural machinery robot better learn path tracking planning tasks and finally learn to use experience to avoid obstacles and reach the end.

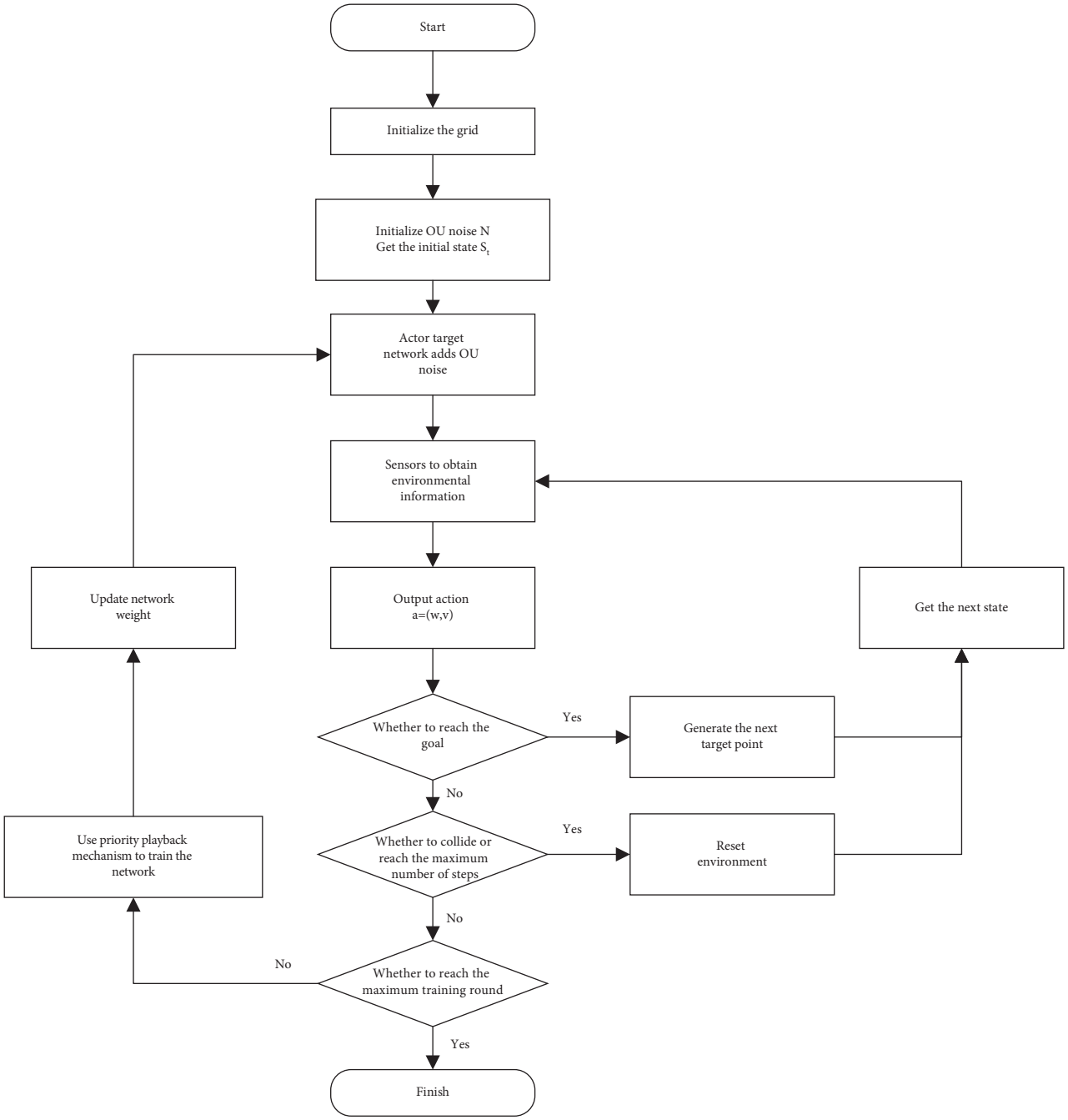


Figure 15: Experimental training process.

5.3.2. Dynamic Obstacle Experiment. In order to test the path tracking ability of the agricultural robot in the DQN and I-DQN algorithms under different types of obstacles, a dynamic obstacle path tracking test was performed in scenario 2. After the agricultural robot enters the termination state, the dynamic obstacle also returns to the original point. The agricultural machinery robot restarts path tracking. Analyze the reward value and success rate during training under the dynamic obstacle scene, and test the path length and planning time.

Figure 18 shows the reward value curve of the two-algorithm training under scenario 2. Similar to scenario 1, the agricultural robot is trying to learn how to avoid obstacles under the two algorithms in the early stage, because the dynamic obstacle avoidance process is more complicated, and the agricultural robot is more likely to collide with obstacles at first, and it takes longer to learn. After 150 rounds, the volatility of the I-DQN algorithm began to decrease, and the reward value increased rapidly in the subsequent 200 rounds, and finally the algorithm gradually converged around 400 rounds. The DQN algorithm fluctuated greatly in the first 200 rounds. The 250 rounds began to rise gradually and did not begin to converge until 550 rounds.

Table 1: Experimental parameter setting table of the mobile agricultural machinery robot.

Parameter name	Parameter assignment
Reward discount rate γ	0.9
Actor network learning rate	0.001
Critic network learning rate	0.001
Priority parameter α	0.6
Correction error parameter β	0.4
Target network delay update TAU	0.001
OU explores noise	$\sigma = 0.2, c = 0.15$
Experience playback pool capacity	200000
BATCH_SIZE	256
Optimizer	Adam
Maximum travel distance per round	3m
Total rounds	15000
Experience pool capacity	50000
Batch capacity	32

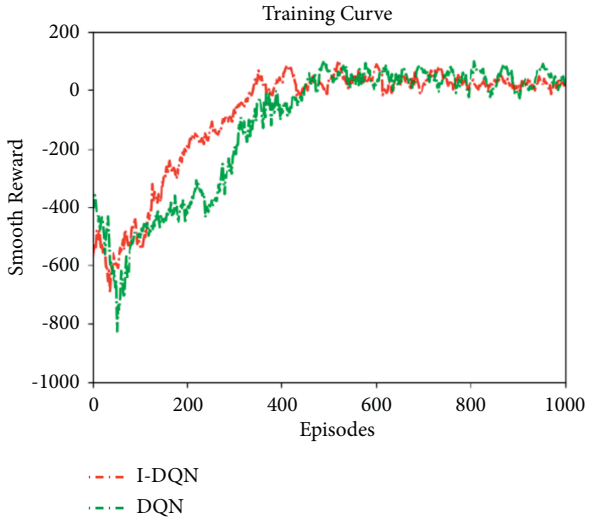


Figure 16: Reward curve of the static obstacle experiment.

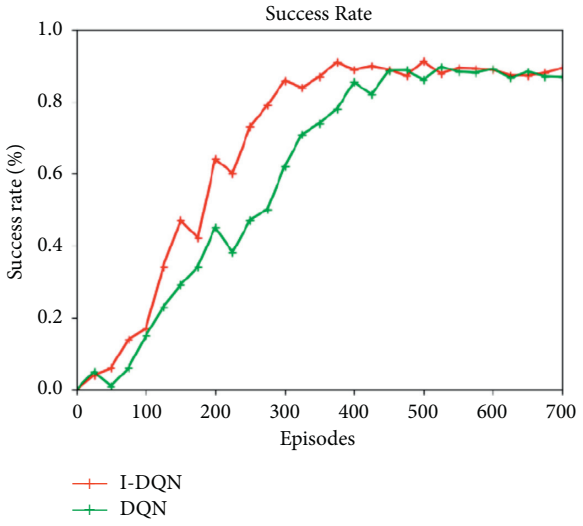


Figure 17: The success rate of static obstacle experiments.

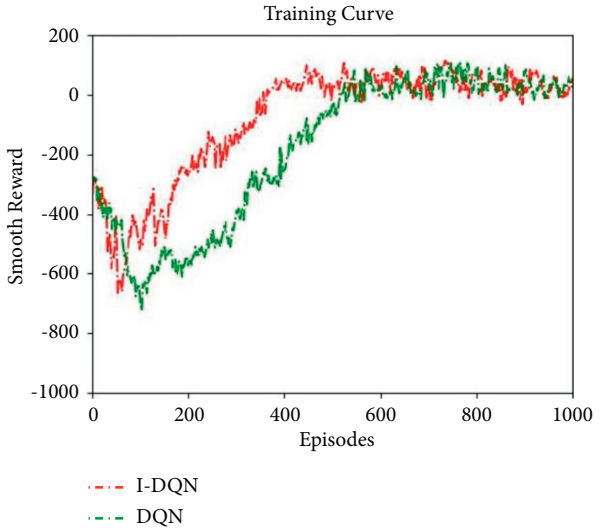


Figure 18: Dynamic obstacle experiment reward curve.

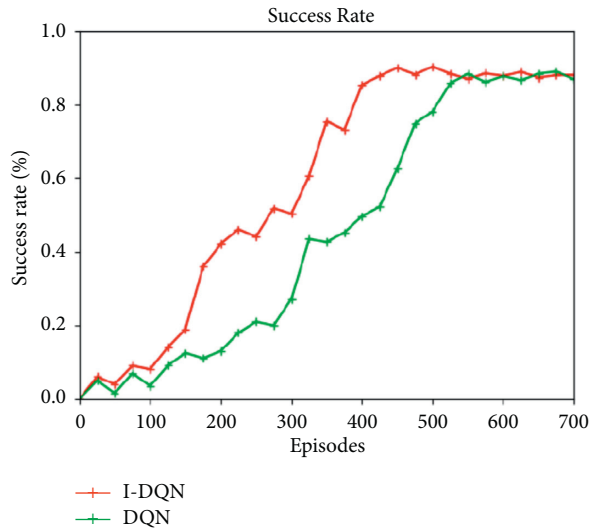


Figure 19: Success rate of the dynamic obstacle experiment.

The path tracking success rate results under scenario 2 are shown in Figure 19. It can be seen that the two algorithms have low success rates in the first 150 rounds, but the success rate of the I-DQN algorithm is greater than the DQN algorithm in the subsequent 200 rounds. The success rate of 350 rounds reaches 75%, which is about 30% higher than the success rate of the DQN algorithm. In 400 rounds, the success rate of I-DQN algorithm basically reached 90%, while DQN had the same success rate in 550 rounds, which proves that I-DQN is better than DQN in path tracking under dynamic obstacle scenarios.

After the training, the dynamic obstacle avoidance process of the mobile agricultural machine in the Gazebo environment is shown in Figure 20. The agricultural robot has been able to continuously reach different target paths while avoiding dynamic obstacles.

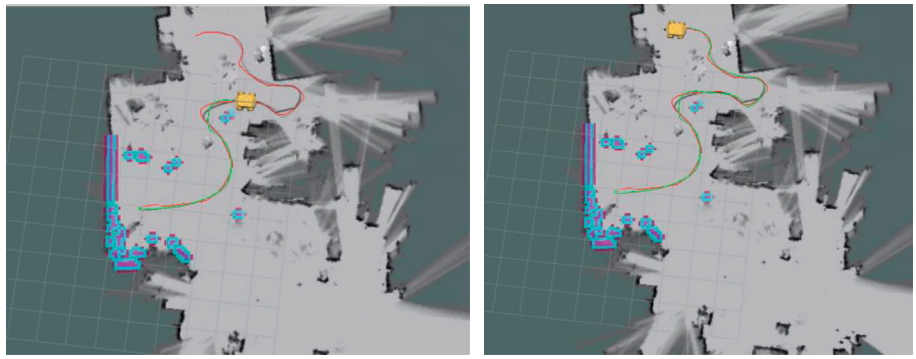


Figure 20: Continuous dynamic obstacle avoidance process of mobile agricultural machinery.

Table 2: I-DQN algorithm path tracking results.

	Target end	Path length (m)	Moving time (s)
1	(−1.13, 0.67)	1.41	7.73
2	(0.84, 1.04)	1.44	7.96
3	(−0.27, 0.34)	1.46	8.27
4	(−1.17, −0.96)	1.62	8.58
5	(0.86, 1.36)	1.69	9.10
6	(0.05, 1.66)	1.73	9.57
7	(−0.91, 1.41)	1.75	9.97
8	(−1.18, 1.33)	1.86	10.11
9	(−1.29, 1.50)	2.11	11.22
10	(1.44, 1.98)	—	—

Table 3: DQN algorithm path tracking results.

	Target end	Path length (m)	Moving time (s)
1	(−1.13, 0.67)	1.49	8.19
2	(0.84, 1.04)	1.49	8.21
3	(−0.27, 0.34)	1.61	8.78
4	(−1.17, −0.96)	1.73	9.16
5	(0.86, 1.36)	1.85	9.43
6	(0.05, 1.66)	1.89	9.90
7	(−0.91, 1.41)	1.95	10.60
8	(−1.18, 1.33)	1.86	10.83
9	(−1.29, 1.50)	2.17	11.56
10	(1.44, 1.98)	—	—

Because obstacle avoidance is more complicated in dynamic obstacle scenarios, the requirements for obstacle avoidance and path tracking of mobile agricultural machinery are higher. Therefore, in order to better test the algorithm, the path length and movement time of the path tracking in scene 2 are tested. Before the test starts, ten coordinate points are randomly generated as the target end point. In order to increase the reliability of the experiment, the coordinate range is set outside the obstacle bypass area, so that the agricultural robot must pass through the obstacle area and will not appear when tracking the path. The target path is very close to the agricultural robot. After the end point is set, perform ten experiments on the I-DQN and DQN algorithms in scenario 2, starting from the origin each time, and use ten randomly generated end points as the target path to perform path tracking, respectively. According to the target path, the straight-line distance length of the starting point is sorted, and the final moving results are shown in Table 2.

Comparing Tables 2 and 3, under the same target end point, the average path length and planning time of the I-DQN algorithm are shorter than those of the DQN algorithm, and the gap gradually increases as the target path distance increases, which proves that the agricultural machinery algorithm tends to make the agricultural machinery robot learn to take a shorter path in a dynamic obstacle scene, and the time is shorter, which improves the performance of path tracking. The average results of the ten times of I-DQN and DQN path tracking are shown in Table 4.

Table 4: Comparison table of path tracking movement results.

	Path length (m)	Moving time (s)	Success rate
I-DQN algorithm	1.67	9.17	90
DQN algorithm	1.76	9.63	90

Based on the analysis of the above experimental results, the DQN algorithm realizes the autonomous path tracking of mobile agricultural machinery in an unknown environment. At the same time, whether in static or dynamic obstacle scenarios, the I-DQN algorithm has a faster convergence speed, allowing agricultural robots to learn to avoid obstacles and reach the target destination faster, and the stability and path tracking performance are improved.

6. Conclusion

With the rapid development of information technology and the realization of smart agriculture, digital agriculture has become an inevitable trend in agricultural development now and in the future. Based on this background, this paper studies the automatic navigation control of agricultural machinery, adopts deep reinforcement learning theory, designs an autonomous path tracking control strategy for agricultural machinery, and conducts experimental simulations through two operating scenarios. The DQN and I-DQN algorithms are applied. In the path tracking task, a number of experiments were designed to verify and analyze

the results. The analysis of experimental results shows that the DQN algorithm realizes the autonomous path tracking of mobile agricultural machinery in unknown environments. At the same time, the I-DQN algorithm has a fast convergence speed. Whether in static or dynamic obstacle scenarios, it can make the agricultural machinery robot learn to avoid obstacles and reach the destination faster, so as to improve the stability and path tracking performance. This research simplifies the motion model and, to a certain extent, does not achieve the true restoration of the actual scene. It has certain limitations for practical applications, but the ideas provided have laid a theoretical foundation for subsequent practical application research.

Data Availability

The dataset can be accessed from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

Natural Science Foundation of Hunan Province (No. 2020JJ7007).

References

[1] D. Chen, Z. Shi, P. Yuan et al., "Trajectory tracking control method and experiment of AGV," in *Proceedings of the AdvancedMotion Control (AMC), 2016 IEEE 14th International Workshop on*, pp. 24–29, IEEE, Auckland, April 2016.

[2] J. E. Normey-Rico, I. Alcalá, J. Gómez-Ortega, and E. F. Camacho, "Mobile robot path tracking using a robust PID controller," *Control Engineering Practice*, vol. 9, no. 11, pp. 1209–1214, 2001.

[3] X. Li, C. Luo, Y. Xu, and P. Li, "A Fuzzy PID controller applied in AGV control system," in *Proceedings of the Advanced Robotics andMechatronics (ICA RM), International Conference on*, pp. 555–560, IEEE, Macau, China, August 2016.

[4] C.-F. Juang and Y.-C. Chang, "Evolutionary-group-based particle-swarm-optimized fuzzy controller with application to mobile-robot navigation in unknown environments," *IEEE Transactions on Fuzzy Systems*, vol. 19, no. 2, pp. 379–392, 2011.

[5] Y.-H. Lee, G.-G. Jin, and M.-O. So, "Level control of single water tank systems using Fuzzy-PID technique," *Journal of the Korean Society of Marine Engineering*, vol. 38, no. 5, pp. 550–556, 2014.

[6] Y. Kanayama, Y. Kimura, F. Miyazaki, and T. Noguchi, "A stable tracking control method for an autonomous mobile robot," in *Proceedings of the Robotics and Automation IEEE International Conference on*, pp. 384–389, IEEE, Cincinnati, OH, USA, May 1990.

[7] N. Hung, J. S. Im, S. Jeong, H. Kim, and B. Sang, "Design of a sliding mode controller for an automatic guided vehicle and its implementation," *International Journal of Control, Automation and Systems*, vol. 8, no. 1, pp. 81–90, 2010.

[8] W. Wu, H. Chen, and Y. Wang, "Adaptive exponential stabilization of mobile robots with uncertainties," in *Proceedings of the Decision and Control 38th IEEE Conference on*, vol. 4, pp. 3484–3489, IEEE, 1999.

[9] T.-L. Bui, P.-T. Doan, D.-T. Van, H.-K. Kim, and S.-B. Kim, "Hybrid control of a tricycle wheeled AGV for path following using advanced fuzzy-PID," *Journal of the Korean Society of Marine Engineering*, vol. 38, no. 10, pp. 1287–1296, 2014.

[10] G. Campion, G. Bastin, and B. Dandrea-Novel, "Structural properties and classification of kinematic and dynamic models of wheeled mobile robots," *IEEE Transactions on Robotics and Automation*, vol. 12, no. 1, pp. 47–62, 1996.

[11] P. Wawzynski, "Control policy with autocorrelated noise in reinforcement learning for robotics," *InternationalJournal of Machine Learning and Computing*, vol. 5, no. 2, pp. 91–95, 2015.

[12] L. Samson, B. O. K. Intelligentie, and E. Gavves, *Deep Reinforcement Learning Applied to the Game Bubbleshooter*, University of Amsterdam, Amsterdam, Netherlands, 2016.

[13] D. Silver, G. Lever, and N. Heess, "Deterministic policy gradient algorithms," in *Proceedings of the International Conference on Machine Learning*, pp. 387–395, 2014.

[14] T. P. Lillicrap, J. J. Hunt, A. Pritzel et al., "Continuous control with deep reinforcement learning," in *Proceedings of the International Conference on Learning Representations*, pp. 1–14, 2016.

[15] A. Barreto, W. Dabney, R. Munos et al., "Successor features for transfer in reinforcement learning," in *Proceedings of the Advances in neural information processing systems*, pp. 4055–4065, 2017.

[16] F. Shoeleh and M. Asadpour, "Graph based skill acquisition and transfer Learning for continuous reinforcement learning domains," *Pattern Recognition Letters*, vol. 87, pp. 104–116, 2017.

[17] M. Fortunato, M. G. Azar, B. Piot et al., "Noisy networks for exploration," 2017, https://arxiv.org/abs/1706.10295.

[18] T. P. Lillicrap, J. J. Hunt, A. Pritzel et al., "Continuous control with deep reinforcement learning," *Computer Science*, vol. 8, no. 6, 2015.

[19] M. Volodymyr, P. B. Adrià, M. Mirza et al., "Asynchronous methods for deep reinforcement learning," in *Proceedings of the International Conference on Machine Learning*, 2016.

[20] W. Meng, Q. Zheng, Y. Shi, and G. Pan, "An off-policy trust region policy optimization method with monotonic improvement guarantee for deep reinforcement learning," *in IEEE Transactions on Neural Networks and Learning Systems*.

[21] A. Kendall, J. Hawke, D. Janz et al., "Learning to drive in a day," in *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 8248–8254, IEEE, Montreal, QC, Canada, May 2019.

[22] R. S. Sutton and A. G. Barto, *Introduction to Reinforcement Learning*, MIT press, Cambridge, England, 1998.

[23] Q. Cai and B. Zhang, "A reinforcement learning model and application research based on agent team," *Computer Research and Development*, vol. 37, no. 9, pp. 1087–1093, 2000.

[24] V. Mnih, K. Kavukcuoglu, D. Silver et al., "Playing atari with deep reinforcement learning," 2013, https://arxiv.org/abs/1312.5602.

[25] J. Li, Y. Chen, X. N. Zhao, and J. Huang, "An improved DQN path planning algorithm," *The Journal of Supercomputing*, vol. 78, pp. 1–24, 2021.

[26] Y. Wang, L. Chen, H. Zhou et al., "Flexible transmission network expansion planning based on DQN algorithm," *Energies*, pp. 488–491, 2021.

[27] Y. Liu and Y. Xu, "Free gait planning of hexapod robot based on improved DQN algorithm," in *Proceedings of the IEEE 2nd International Conference on Civil Aviation Safety and Information Technology (ICCASIT)*, October 2020.

Hindawi
Scientific Programming
Volume 2022, Article ID 8148972, 7 pages
https://doi.org/10.1155/2022/8148972

Research Article

Deep Neural Network-Based Sports Marketing Video Detection Research

Longcheng Xu,[1] Deokhwan Choi ,[2] and Zeyun Yang [2]

[1]*Physical Education Department of Jiangsu Vocational College of Finance and Economics, Huaian, Jiangsu 223000, China*
[2]*Physical Education Department of Woosuk University, Jeonju, Jeollabuk-do 55340, Republic of Korea*

Correspondence should be addressed to Deokhwan Choi; y553149602@163.com

Received 5 January 2022; Accepted 11 February 2022; Published 19 March 2022

Academic Editor: Baiyuan Ding

With the rapid development of short video, the mode of sports marketing has diversified, and the difficulty of accurately detecting marketing videos has increased. Identifying certain key images in the video is the focus of detection, and then, analysis can effectively detect sports marketing videos. The research of video key image detection based on deep neural network is proposed to solve the problem of unclear and unrecognizable boundaries of key images for multiscene recognition. First, the key image detection model of the feedback network is proposed, and ablation experiments are conducted on a simple test set of DAVSOD. The experimental results show that the proposed model achieves better performance in both quantitative evaluation and visual effects and can accurately capture the overall shape of significant objects. The hybrid loss function is also introduced to identify the boundaries of key images, and the experimental results show that the proposed model outperforms or is comparable to the current state-of-the-art video significant object detection models in terms of quantitative evaluation and visual effects.

1. Introduction

Vision is the main way humans receive information from the outside world, and according to research in the field of neuroscience, about 10^8 to 10^9 bytes of data enter the human eye every second [1]. This is because of the selective role of the visual attention mechanism, which allows the visual system to selectively ignore irrelevant information and pay attention to relevant information, just like separating the grains of wheat from the husk. In this Internet era where the amount of data is exploding, how to get the information of people's concern from the huge amount of information in a labor and material-saving way has gained a lot of attention. Therefore, introducing attention mechanisms into data processing tasks and prioritizing the allocation of data processing resources to more critical information can help improve the efficiency of processing information [2–6].

In 1998, Borji and Itti [7] proposed the first computational model of visual saliency based on Koch et al.'s theory and the classical feature integration theory of cognitive psychology [8] and the pointing search model [9], whose algorithmic process contains three main steps: extraction of three primary visual features: color, luminance, and orientation. Three types of key features are computed at multiple scales using central-peripheral contrast (key feature extraction); the feature maps are normalized and then synthesized (feature fusion), and the key targets in the images are labeled using the WTA mechanism. The algorithm has had a significant impact on subsequent research on computational models of visual criticality in the field of computer vision, especially since mainstream criticality detection algorithms used a similar framework before deep learning techniques were used on a large scale.

Early image salient object detection models [10] were mainly based on a bottom-up approach using different underlying visual features, such as color, edges. Since salient object detection is closely related to the human eye attention detection task and both model the human visual attention mechanism, the early salient object detection models also borrowed some basic theories of the human visual attention mechanism, including the classical contrast assumption, center-surround assumption. For example, both

assumptions were used by Liu et al. [11] and Achanta et al. [12], and a similar assumption was used by Cheng et al. [10], who considered color contrast information on both local and global scales, and the algorithm was concise and straightforward and received wide attention from the academic community. In addition, Yan et al. [13] proposed to complete the apparently consistent image representation at different scales by over-segmenting the image at different scales and to extract and fuse the salient features at different scales for optimization to obtain the final salient object detection results. Visual center bias is also a commonly used hypothesis based on human attentional mechanisms [13]. The hypothesis is based on the phenomenon that the human visual system has a tendency to assign higher attentional weights to the center of the scene when observing the scene. After that, the popular hypothesis is the background prior hypothesis, which was proposed by Wei et al. [14] in 2012. Unlike the center-periphery hypothesis and the visual center shift hypothesis, which attempt to define "what is more likely to be the salient region," this hypothesis attempts to define "what is more likely to be the background." This assumption is based on the observation that in most scenes, the parts around the edges of the image have a higher probability of belonging to the background. This assumption can be considered as a further development of the visual center bias assumption. Before the large-scale application of deep learning techniques, the background prior assumption was the most effective assumption in the field of saliency detection, and the majority of high-performing models [15–19] were based on this assumption. These works focus on how to further improve the accuracy of the background prior assumption and how to apply more advanced one-class classifiers. By the background prior assumption, which is equivalent to obtaining a class of (background) samples, the problem can be considered as a one-class classifier giving only one class of samples.

With the great success of deep learning techniques in image classification problems, the focus of research in the field of significant object detection has gradually shifted to deep learning-based models. Slightly earlier work used deep learning features as a more effective key representation and trained using fully convolutional neural networks. Lee et al. [20] used depth features as high-level information and Gabor-filtered response and color histogram as bottom-level features to fuse different levels of significant information for significant prediction. These models achieve better performance but have some drawbacks, such as the large number of parameters and loss of spatial information due to the use of fully connected layer-based classification networks and the high computational cost of these algorithms due to the need for significant/insignificant classification of each superpixel or target object alternative.

With the rise of fully convolutional neural networks, in recent years, significant object detection efforts based on deep learning have used or adapted full convolutional neural networks for pixel-level critical prediction. There is some work [21] inspired by the pixel-level semantic segmentation task, proposing the fusion of features from different neural network layers for critical object detection. Since the shallower layers of deep neural networks can retain more fine-grained underlying visual features, and the deeper layers can extract higher-level, semantic-level features, the fusion of features from different neural network layers can retain the original underlying spatial information and obtain higher-level semantic information. Currently, the main research focus of the work on significant object detection based on deep learning techniques is to explore more efficient network structures that can retain more spatial details. Wang et al. [22] proposed an ASNet model for detecting visually salient objects by means of visual attention prior. The model treats visual attention as a high-level understanding of the whole scene, which is learned through higher-level neural network layers, and the salient object detection task is considered as a more fine-grained, object-level saliency detection, with visual attention providing top-down guidance. The ASNet model is based on a stacked convolutional long and short-term memory neural network, which has a unique recurrent structure that can iteratively optimize saliency detection results. This work provides a deeper understanding of the visual attention mechanism and reveals the correlation between salient object detection and human eye attention detection. As a whole, the deep learning-based salient object detection model achieves much better performance than traditional models [23–26].

In response to the current research status, this paper investigates video salient object detection based on deep neural networks as follows, extracting richer spatial saliency information and better capturing the overall shape of salient objects. In this paper, an attention feedback network-based video salient object detection model is proposed. To further obtain clearer bounds, a new hybrid loss function is introduced in this paper based on the video salient object detection model and the attentional feedback network.

2. Deep Neural Networks

2.1. Convolutional Neural Network. When people read or watch a video, they perceive and understand the current content based on the text or images they have already observed before and do not completely forget what they have observed before, and their brain goes blank to understand the content that follows. Traditional neural networks cannot predict salient information in later frames based on the salient object regions in the previous video frames. The emergence of recurrent neural networks makes the network memorable, and its network structure is shown in Figure 1. Assuming that $\{X_t\}_{t=0}^{t}$ is a set of inputs with $(t+1)$ time steps and $\{H_t\}_{t=0}^{t}$ is the corresponding output of the network, network N receives at time step t not only X_t but also the value of the first $(t-1)$ value of the hidden state at a time step, that is, the network processes the current input with reference to the previous memory.

However, when the video sequence is long, the interval between the current video frame to be processed and the related video frame may be large, and at this time, the RNN may lose the memory of distant video frames due to problems such as gradient disappearance. To address the problem of long-term dependence, Hochreiter et al. [27]

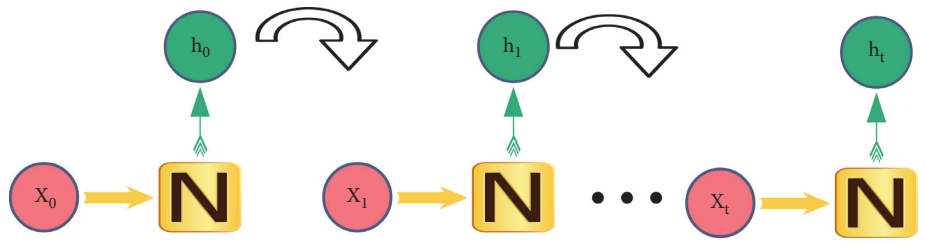


Figure 1: Network structure.

proposed a long-term and short-term memory network, as shown in Figure 2, where the contents of the three stages indicate the forgetting phase, updating state phase, and output phase, respectively.

All three stages contain a sigmoid layer that maps the input information to between [0, 1] and then selectively filters the useful information and forgets the useless information by a per-bit multiplication operation.

The forgetting stage is used to filter the useful information and forget the useless information. The current input is x_t, connecting x_t with the hidden state h_{t-1} of the previous moment, denoted as J_t, and ※ denotes the connection operation, as shown as follows:

$$h_t \text{※} x_t = J_t. \tag{1}$$

The sigmoid layer is then used to map J_t to between [0, 1] to obtain the output gate f_t, where W_f and b_f denote the weight and bias vector of the network layer, respectively, and σ denotes the sigmoid operation, as shown as follows:

$$f_t = \sigma\left(W_f \cdot J_t + b_f\right). \tag{2}$$

Then, the corresponding element multiplication operation $(\therefore)$ is performed with the cell state C_{t-1}, thus selectively filtering the useful information and forgetting the useless information, and the cell state at this point is noted as C_t^*.

$$C_t^* = f_t \therefore C_{t-1}. \tag{3}$$

The update cell state phase allows the control cell state to selectively absorb relevant information from J. J_t passes through the sigmoid layer and generates the input gate i_t.

$$i_t = \sigma\left(W_i \cdot J_t + b_i\right). \tag{4}$$

The information obtained by multiplying the feature obtained by J_t after the tanh layer with the corresponding element of it is the information added to the cell state, and the new cell state C_t^* is obtained by adding this information to the C_t obtained in the forgetting phase by bits.

$$C_t = C_t^* + i_t \therefore \tanh\left(W_c \cdot J_t + b_c\right). \tag{5}$$

The output phase controls what information is output at the current moment. J_t is inputted into the sigmoid layer to get an output gate O_t.

$$O_t = \sigma\left(W_o \cdot J_t + b_o\right). \tag{6}$$

Let O_t and the current cell state C_t be multiplied bitwise by the features obtained through the tanh layer to obtain the output at the current moment H_t.

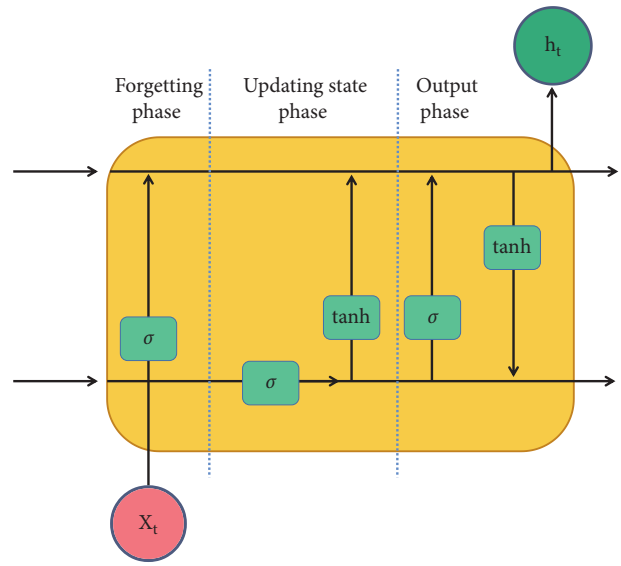


Figure 2: Memory network.

$$H_t = O_t \therefore \tanh\left(C_t\right). \tag{7}$$

2.2. Loss Function. When performing pixel-level salient object detection, it can be viewed as a binary partitioning problem, where pixels belonging to the salient object are labeled as 1 and pixels belonging to the background are labeled as 0. Assume that y_i denotes the label of sample x_i, the desired output, and $\overline{y}$ denotes the probability value of $y_i = 1$ for a given sample x_i.

$$\overline{y}_i = P\left(y_i = 1 | x_i\right). \tag{8}$$

$1 - \overline{y}$ denotes the probability value of $y_i = 0$ given sample x_i.

$$1 - \overline{y}_i = P\left(y_i = 0 | x_i\right). \tag{9}$$

When x_i occurs, the probability of y_i occurrence can be expressed by $P(y_i \mid x_i)$. From the perspective of maximum likelihood, $P(y_i|x_i)$ can be expressed in the following form.

$$P\left(y_i | x_i\right) = \overline{y}^{y_i} \cdot \left(1 - \overline{y}\right)^{1 - y_i}. \tag{10}$$

When the real mark $y_i = 0$, 1, and take the logarithm operation. Since the smaller the value of the loss function, the more favorable it is, and the log takes a negative value, and the loss function is calculated as follows:

$$L = \frac{1}{N}\sum_{i=1}^{N} -[y_i \log \overline{y} + (1 - y_i)\log(1 - \overline{y})]. \quad (11)$$

2.3. Feedback Network. In order to reduce the loss of necessary visual criticality information due to repetitive stride and pooling operations and to learn richer static criticality information, AFNet is used as the main skeleton of the static criticality module. Stimuli in Figure 3 show the input image frames, and the encoding and decoding networks consist of five convolutional blocks of VGG16 (denoted as E^i and D^i, respectively, $i \in \{1, 2, 3, 4, 5\}$), where the information transfer between the corresponding convolutional blocks is controlled by the attention feedback module.

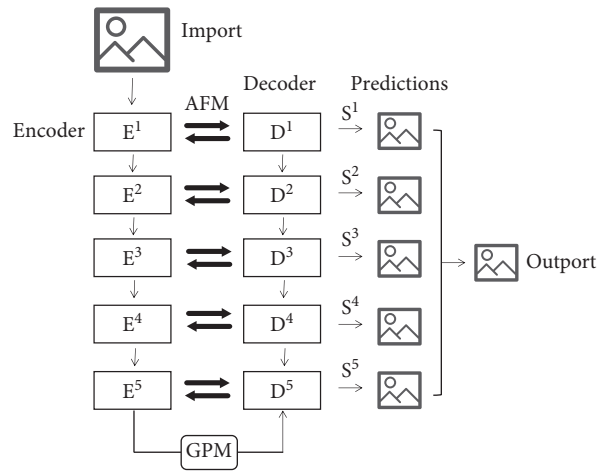


Figure 3: Feedback network model.

3. Design of Deep Neural Network

3.1. Feedback Network Detection Model. The NHM model is proposed to capture richer spatial criticality information and thus better capture the overall shape of key images. The NHM model uses the attentional feedback network as the backbone of the static criticality module to reduce the loss of visually critical information caused by scale-space issues and to guide the correct fusion of multiscale features from coarse to fine scales. The multiscale feature maps extracted from the five decoding blocks of the attentional feedback network are then fused and fed to the pyramidal expansion convolution module to retain more spatial visual critical information. After that, the time-critical information is captured using a key object transfer-aware convolutional long short-term memory network in consideration of attention-aware transfer, and finally, the parameters of the model are optimized by gradually reducing the value of the loss function through continuous iterations. The algorithm is divided into three parts: extraction of multiscale spatial features, integration of spatio-temporal critical information, and loss minimization.

To mitigate the negative effects such as the loss of visual information generated by the scale-space problem, the backbone of the static criticality detection module consists of AFNet and PDC modules connected together. AFNet as a novel codec forms the design of a fully convolutional network, its encoding and decoding network consists of five convolutional blocks, and E^i and D^i denote the encoder and decoder blocks, respectively, where $i \in \{1, 2, 3, 4, 5\}$, indicating that E^i and D^i each contain five convolutional blocks, where each layer of the encoder block transmits its critical information through the feedback module in AFNet to the corresponding decoder block. The feedback module uses a two-step iterative learning approach, where the time steps are denoted by $i \in \{1, 2\}$, which helps to correct inaccurate predictions generated in the previous network by simulating a feedback mechanism that multiplies the ternary map pixel by pixel with the obtained feature map, thus helping to capture the overall shape of the key object. Facing the global spatial criticality detection problem, AFNet uses the global perception module to overcome the problem that the fully connected operation ignores local information and generates redundant data. A multiscale segmentation strategy is used to divide the feature map into 4, 16, and 36 parts, which are then stacked and reorganized for global convolution operation to make full use of the global and intraregional saliency information.

The key image in the dynamic scene is detected directly by the image key object detection model. The key object detection can only detect the spatial differences of color contrast, direction contrast, brightness contrast, and so on. However, in dynamic scenes, the temporal factor is usually used as an important clue for the criticality detection. Second, detection only on each individual frame without reference to the criticality information contained in previous frames may be highly incoherent, because the target and background may differ significantly in appearance in different frames, which will lead to incoherent detection results between frames. Finally, video content often contains significant redundancy, as consecutive video frames require enough similar content to provide a smooth viewing experience. Simply ignoring content redundancy can lead to higher computational costs. Therefore, VSOD needs to consider both temporal and spatial saliency information, so a dynamic saliency detection module is used to integrate temporal and spatial saliency information. In order to better simulate the perceptual function of the human visual system, temporal saliency information is learned, and the process of attentional perceptual transfer is captured, and this paper uses SSLSTM as a dynamic saliency detection module, which combines the powerful spatio-temporal feature extraction capability of ConvLSTM with the attentional transfer mechanism.

Deep neural networks gradually optimize the network by iteratively minimizing the loss function. The loss function measures the difference between the value predicted by the model and the true value, and the weights of the network are updated by gradient descent.

$$L = \sum_{t=1}^{T} \left(l(I_t) \cdot l^{Att}(A_t, F_t) + l^{\mathrm{VSDD}}(S_t, M_t) \right). \quad (12)$$

The meaning of each symbol is shown in Table 1, because the video significant object detection dataset contains relatively

Table 1: Compare other model.

Metric		LNSM	MBNM	PDBM	SSAV
ViSal	S	0.993	0.883	0.884	0.933
	MAE	0.025	0.01	0.026	0.022
	Max f	0.993	0.873	0.866	0.929
DAVSOD	S	0.779	0.627	0.67	0.714
	MAE	0.14	0.149	0.097	0.082
	Max f	0.68	0.51	0.575	0.593
UVSD	S	0.825	0.688	0.891	0.85
	MAE	0.053	0.069	0.008	0.015
	Max f	0.701	0.54	0.853	0.791
VOS	S	0.86	0.732	0.794	0.808
	MAE	0.122	0.089	0.061	0.063
	Max f	0.779	0.66	0.728	0.732

few human eye focus annotations, so l_t to indicate whether the dataset contains human eye focus annotations, when the dataset does not contain human eye focus annotations, the loss function at this time does not contain the $l_t A_t$ term, the error will not be back-propagated. The meaning of each symbol is shown in Table 1. Since the video important target detection data set contains relatively few eye focus annotations, it is used to indicate whether the data set contains eye focus annotations. When the data set does not contain eye focus annotations, the loss function at this time does not contain ItAt term, and the error will not be propagated back.

3.2. Loss Function Design. A novel hybrid loss function is proposed based on the boundary enhancement loss, and the function consists of the loss L^a of the predicted attention-perception feature map, the loss L^v of the final key object prediction result, and the loss L_b^v of the final predicted target boundary.

$$L = \omega_1 \cdot \left(L^a + L^v\right) + \omega_2 \cdot L_b^v, \tag{13}$$

where ω_1, ω_2 are used as the learning rate parameters for object-level loss and object-boundary loss of the control target, respectively, and let $\omega_1 : \omega_2 = 1 : 10$ to emphasize the learning of the target boundary.

The dataset used for part of the training does not contain human eye focus annotations, so the predicted loss L^a of the perceptual attention feature map can be divided into two parts: loss L_f^a calculated using human eye focus annotations and loss L_m^a calculated using salient object annotations.

$$L^a = \delta(1) \cdot L_f^a + (1 - \delta(1)) \cdot L_m^a, \tag{14}$$

$L^a = L_f^a$ when $\delta(1) = 0$, $L^a = L_m^a$ when $\delta(1) = 1$. The final key object prediction results are denoted by S_t. That is, the loss L^v can be calculated.

When $\delta(1) = 0$, $L^a = L_f^a$; When $\delta(1) = 1$, $L^a = L_m^a$. S_t is used to represent the prediction result of the final key object, and M_t represents the object level annotation of the key object. The loss L^v can be calculated as follows:

$$L^v = \sum_{t=1}^{T} \left(L^{VSO\ D}(S_t, M_t)\right). \tag{15}$$

The average pooling operation $\overline{P}$ can be used to extract smooth boundaries. Suppose it is necessary to extract the boundary $B(X)$ of the image X and take the absolute value after making a difference between X and $\overline{P}(X)$. The final predicted target boundary loss L_b^v is as follows:

$$L_b^v = \sum_{t=1}^{T} \left(L^{VSOD}(B(S_t), B(M_t))\right). \tag{16}$$

On the basis of NHM, a mixed loss function for capturing clear boundaries is added. The loss function is based on the boundary enhancement loss and is composed of the attention perception feature map predicted by the model, the prediction results of key images, and the prediction results of key image boundaries. The model is recorded as LNSM.

4. Experiments and Results

4.1. Experimental Design. The experiments were run on an Nvidia GTX1080TI GPU. The experiments in this paper were implemented using the Python language on Caffe's deep learning framework, and Matlab was used for quantitative evaluation of performance. The training set of DAVIS, DAVSOD, and FBMS and the validation set of DAVSOD were also used to train the proposed model, where the weights of the network model were initialised by the AFNet model, and video was processed per batch, and the number of time steps for the conLSTM network layer processing was set to 3. The training process was set up as follows: first, the static key model was pretrained with a base learning rate of 10^{-9}; then, the entire model was trained by setting the learning rate of the dynamic key module to 10^{-8} and the learning rate of the static key module to 10^{-10}; finally, the static key module weights were fixed, and the dynamic key module was fine-tuned with the learning rate set to 10^{-10}. The LNSM module was trained using 32 hours and 64 k iterations.

4.2. Compare Other Model. In this paper, the proposed LNSM is compared with four advanced video critical object detection models, MBNM, PDBM, and SSAV, on datasets created specifically for the VSOD task (the entire dataset for ViSal and UVSD, a test set for VOS, and a simple test set for DAVSOD), and the experimental results of the quantitative evaluation are shown in Table 1. It can be seen from Table 1 that the three indexes of the model proposed in this paper are better than other models on DAVSOD and ViSal datasets. Especially on the simple test set of DAVSOD, the *f*-value index and average absolute error based on pixel error and the structural index measuring the overall structural difference have improved the performance by 0.06, 0.03, and 0.064, respectively, compared with SSAV; advanced performance has also been achieved on other datasets. Moreover, ViSal is the first test benchmark especially designed for video key object detection; DAVSOD dataset takes into account the transfer of visual attention and its selectivity when labeling and can represent the real attention behavior of the human visual system in dynamic scene. These two

datasets are very representative. The experimental results show that the LNSM model has good performance for creating datasets especially for VSOD and DAVSOD datasets that mark key images according to human eye concerns.

5. Conclusion

This paper focuses on key image detection based on deep neural networks to complete the detection of sports marketing videos. For the detection of multiple scenes, a feedback network-based video off-image detection model and a hybrid loss function are proposed to solve the detection problem of key images. The LNSM model proposed in this paper is compared with the quantitative evaluation and visualisation results of the three state-of-the-art models on six representative datasets. The quantitative results demonstrate that LNSM outperforms other advanced models in all three evaluation metrics on the DAVSOD and ViSal datasets and achieves advanced performance comparable to other models on widely used datasets.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

[1] K. Koch, J. McLean, R. Segev et al., "How much the eye tells the brain," *Current Biology*, vol. 16, no. 14, pp. 1428–1434, 2006.

[2] V. R. S. Mani, A. Saravanaselvan, and N. Arumugam, "Performance comparison of CNN, QNN and BNN deep neural networks for real-time object detection using ZYNQ FPGA node," *Microelectronics Journal*, vol. 119, Article ID 105319, 2022.

[3] X. Liang, A. M. Javid, M. Skoglund, and S. Chatterjee, "Decentralized learning of randomization-based neural networks with centralized equivalence," *Applied Soft Computing*, vol. 115, Article ID 108030, 2022.

[4] M. Jian, J. Wang, H. Yu, and G.-G. Wang, "Integrating object proposal with attention networks for video saliency detection," *Information Sciences*, vol. 576, pp. 819–830, 2021.

[5] F. B. Chen, X. L. Wang, X. Li, Z. R. Shu, and K. Zhou, "Prediction of wind pressures on tall buildings using wavelet neural network," *Journal of Building Engineering*, vol. 46, Article ID 103674, 2022.

[6] J. Pilarz, I. Polishuk, and M. Chorążewski, "Prediction of sound velocity for selected ionic liquids using a multilayer feed-forward neural network," *Journal of Molecular Liquids*, vol. 347, Article ID 118376, 2022.

[7] A. Borji and L. Itti, "State-of-the-Art in Visual attention modeling," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 1, pp. 185–207, 2013.

[8] M. Carrasco, "Visual attention: the past 25 years," *Vision Research*, vol. 51, no. 13, pp. 1484–1525, 2011.

[9] C. E. Connor, H. E. Egeth, and S. Yantis, "Visual attention: bottom-up versus top-down," *Current biology : CB*, vol. 14, no. 19, pp. R850–R852, 2004.

[10] M.-M. Cheng, N. J. Mitra, X. Huang, P. H. S. Torr, and S.-M. Hu, "Global contrast based salient region detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 3, pp. 569–582, 2015.

[11] T. Liu, J. Sun, N. N. Zheng, X. O. Tang, and H. Y. Shum, "Learning to detect a salient objec," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8, Minneapolis, Minnesota, June 2007.

[12] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency-tuned salient region detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1597–1604, Miami, Florida, June 2009.

[13] Q. Yan, L. Xu, J. P. Shi, and J. Y. Jia, "Hierarchical saliency detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1155–1162, Long Beach, California, June 2013.

[14] Y. Wei, F. Wen, W. Zhu, and J. Sun, "Geodesic saliency using background priors," in *Proceedings of the 12th European conference on Computer Vision*, pp. 29–42, Florence, Italy, October 2012.

[15] C. Yang, L. H. Zhang, H. C. Lu, X. Ruan, and M. H. Yang, "Saliency detection via graph-based manifold ranking," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3166–3173, Portland, Oregon, USA, June 2013.

[16] B. W. Jiang, L. H. Zhang, H. C. Lu, C. Yang, and M. H. Yang, "Saliency detection via absorbing Markov chain," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1165–1672, Sydney, NSW, Australia, April 2013.

[17] Y. Qin, H. C. Lu, Y. Q. Xu, and H. Wang, "Saliency detection via cellular automata," in *Proceedings of the IEEE Conference on Computer Vision and Pattern recognition*, pp. 110–119, Boston, MA, USA, June 2015.

[18] J. M. Zhang, S. Sclaroff, Z. Lin, X. H. Shen, B. Price, and R. Mech, "Minimum barrier salient object detection at 80fps," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1404–1412, Santiago, Chile, December 2015.

[19] W. C. Tu, S. F. He, Q. X. Yang, and S. Y. Chien, "Real-time salient object detection with a minimum spanning tree," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2334–2342, Las Vegas, Nevada, USA, June 2016.

[20] Z. Chen, C. Wu, Z. Huang et al., "Dangerous driving behavior detection using video-extracted vehicle trajectory histograms," *Journal of Intelligent Transportation Systems*, vol. 21, no. 5, pp. 409–421, 2017.

[21] P. P. Zhang, D. Wang, H. C. Lu, H. Y. Wang, and X. Ruan, "Amulet: aggregating multi-level convolutional features for salient object detection," in *Proceedings of the IEEE Int'l Conf. on Computer Vision*, pp. 202–211, USA, 2017.

[22] W. G. Wang, J. B. Shen, X. P. Dong, and A. Borji, "Salient object driven by fixation prediction," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1711–1720, USA, June 2018.

[23] X. Wang, Y. Zhang, and C. Ning, "A novel visual saliency detection method for infrared video sequences," *Infrared Physics & Technology*, vol. 87, pp. 91–103, 2017.

[24] M. Li, E. Zhang, L. Wu, and J. Duan, "A multistage and multiresolution deep convolutional neural network for inverse halftoning," *Expert Systems with Applications*, vol. 191, Article ID 116358, 2022.

[25] Xu Zhang, Z Yuanyuan, and L Shaoyuan, "Bayesian neural network with efficient priors for online quality prediction," *Digital Chemical Engineering*, vol. 2, Article ID 100008, 2021.

[26] B. S. Kronheim, M. P. Kuchera, and H. B. Prosper, "TensorBNN: Bayesian inference for neural networks using TensorFlow," *Computer Physics Communications*, vol. 270, Article ID 108168, 2022.

[27] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

Hindawi
Scientific Programming
Volume 2022, Article ID 4576457, 14 pages
https://doi.org/10.1155/2022/4576457

Research Article

The Application of Adaptive Analytic Hierarchy Process Driven by Multisource Big Data in the Training of School-Enterprise Joint Engineering Ability

Liqing Zhang

School of Mechanical and Vehicle Engineering, Changchun University, Changchun, Jilin 130022, China

Correspondence should be addressed to Liqing Zhang; zhanglq80@ccu.edu.cn

Received 27 December 2021; Revised 18 January 2022; Accepted 3 February 2022; Published 18 March 2022

Academic Editor: Baiyuan Ding

The Chinese government has pointed out in relevant literature that by 2020, China is in short supply of high-level innovative technology talents, and China's engineering education will face unprecedented challenges. Based on the needs of China's future economic development, the Ministry of Education urgently launched the "Excellent Engineer Education and Training Program" in 2010, which aims to cultivate a large number of solid theoretical foundations, strong practical skills, and high innovation capabilities for the industry, the world, and the future. An outstanding engineer is the one who can quickly adapt to the needs of economic and social development. The program is divided into three levels of undergraduate, master, and doctoral training. The undergraduate level focuses on training in a series of high-quality applications capable of not only being competent in the management and marketing of the production site but also in the design, development, and operation of engineering projects. Talent: to achieve this training goal, the in-depth cooperation between schools and enterprises has become the key. In order to achieve the undergraduate-level training goals of the "Excellence Program," various local colleges and universities have adopted the method of school-enterprise cooperation, combined with their own school positioning, trying to explore a training path suitable for their own development. This research takes China's local application-oriented universities as the research object and uses the literature method to sort out and summarize the status quo of the school-enterprise cooperation in training outstanding engineers in China's local application-oriented universities. Select two different application-oriented universities located in areas with large economic development gaps, analyze their common problems, and try to put forward countermeasures and suggestions.

1. Introduction

Since China's reform and opening up, after more than 20 years of unremitting development, the economy has been rapidly improved, and the pace of industrialization is accelerating. By the end of the 20th century and the beginning of the 21st century, China's basic economic conditions have truly changed from an agricultural country to a big industrial country [1–3]. At the same time, China's engineering education has bathed in spring breeze, has grown rapidly, and has made great progress. However, after long-term development, the shortcomings of China's engineering education have gradually emerged, and the international community and enterprises have rarely recognized the engineers trained by the school. China has many workers, but not many qualified engineers. Why does this happen [4–6]? In 2009, the report of Chinese Academy of Engineering pointed out that there are many problems in the hierarchy and adaptability of China's engineering education. (2) The problems of lack of engineering and weak practical links in engineering education have not been solved for a long time. (3) The evaluation system is oriented to emphasize thesis, neglect design, and lack practice. (4) Industrial, educational, and political cooperation is not in place. (5) Enterprises do not pay attention to the process of participating in talent training [3, 7–10].

The education sector has separated from the industry and the business sector for a long time, and the lack of

enthusiasm of enterprises to participate in the cultivation of talents in universities has caused the contradiction between supply and demand in the human resources market to become increasingly prominent. Enterprises are worried about not being able to recruit qualified employees, university graduates are becoming more and more anxious about not finding suitable jobs, social employment competition is becoming fiercer, and students' employment pressure is increasing. International competition, in the final analysis, is the competition of talents. Charles West once said that the country with the best engineering talent occupies the core position of economic competition and industrial advantage. It means that in the future international competition, whoever can train the best engineering talents will have a certain advantage in international competition. According to a McKinsey research report [11–13], by 2020, global high-tech companies will be facing a shortage of about 40 million technical talents, and China will face a shortage of 22 million engineering and technical talents. The "Outline of the National Medium and Long-term Talent Development Plan (2010–2020)" mentioned that by 2020, in China, the quantity demand of high-level innovative technology talents needed in key economic areas will reach more than 5 million.Intermediate and senior technical talents account for about 5% of the employees, while technical talents account for only 20% of the total human resources. This means that due to the needs of future economic development, China urgently needs a group of people who can master high-tech technologies [14–16].

In order to conscientiously implement the "National Medium and Long-term Education Reform and Development Plan Outline (2010-2020)" and the "National Medium and Long-term Talent Development Plan Outline (2010-2020)" and other major reform projects, to achieve the goals of engineering education reform. In June 2010, the Ministry of Education urgently launched the "Excellent Engineer Education and Training Program" (hereinafter referred to as the "Excellence Program"). Future-oriented high-quality compound talents for outstanding engineers will continue to improve China's innovation capabilities, enhance international competitiveness, and strive to develop China into a country with talents and skills [10, 17–19]. The "Excellence Program" divided into three levels of undergraduate, master, and doctoral training. The three levels of training highlight that school-enterprise cooperation is a key factor in the success of the program. In order to achieve the undergraduate level training standards, various local application-oriented universities actively carry out extensive and then they create technological innovation alliances, form a community of interests, and build a platform for industry-university cooperation. A school-enterprise cooperation road is suitable for long-term development [8, 20, 21].

At present, there are few educational research results on the use of school-enterprise joint training of engineers at the undergraduate level of engineering education in China, and most of them are concentrated on the cooperative education of industry-university-research in the relevant high and secondary vocational colleges. Therefore, a systematic review of the current situation of China's local application-oriented universities using school-enterprise joint training of outstanding engineers, analysis of its existing problems, and suggestions are helpful to improve the undergraduate level education of local application-oriented universities and school-enterprise joint training of outstanding engineers [22–25]. The research logical structure of this paper is shown in Figure 1.

2. Existing Research Results and Literature Review

2.1. Joint Cultivation of Postgraduates in a Collaborative Innovation Environment. The term "environment" is rich in meaning, including not only material or natural factors represented by the atmosphere, soil, etc. but also nonmaterial or social factors such as culture, concepts, and systems. Different disciplines also have different definitions of "environment." In recent years, some scholars have conducted preliminary research on the "innovation environment" or "collaborative innovation environment" with synergy characteristics and believe that it will have a profound impact on the scientific and technological system and promote the elements of scientific and technological innovation from "isolation, decentralization, and closure" to convergence and integration. We think that an innovation environment with synergistic characteristics formed under the long-term effects of various factors and a stable network system dedicated to improving innovation capabilities. Its subelements include infrastructure level, financial environment, and entrepreneurial level, etc. [26–28]. In addition to having the basic characteristics or basic attributes of "collaborative innovation environment" a knowledge and information exchange space must also be constructed by multiple subjects in interactive cooperation. The narrow sense of collaborative innovation environment refers to the collaborative innovation center and universities in our country and is composed of various participants [29–31].

The important feature of the multisubject cooperation space created by undertaking specific national major scientific and technological research and development projects or cultural inheritance innovation projects is that it is substantive and presented in a dot-like manner according to the distribution location of collaborative innovation. A broad collaborative innovation environment: it refers to the knowledge sharing. The cooperation philosophy and collaboration culture of complementary resources, organizational collaboration, and multiple win-win results not only include the physical interaction space, based on the collaborative innovation center, but also include those inspired by the concept of collaborative innovation, widespread and recognized by all parties to the cooperation values and codes of conduct. Since collaborative innovation is becoming a collective behavior for universities and governments at all levels to promote the reform of the scientific and technological system and accelerate the cultivation of innovative talents, the collaborative innovation concept has become a ubiquitous shared value beyond the "collaborative innovation center" itself. Therefore, the "collaborative innovation" in this research "Environment" is a broad sense of

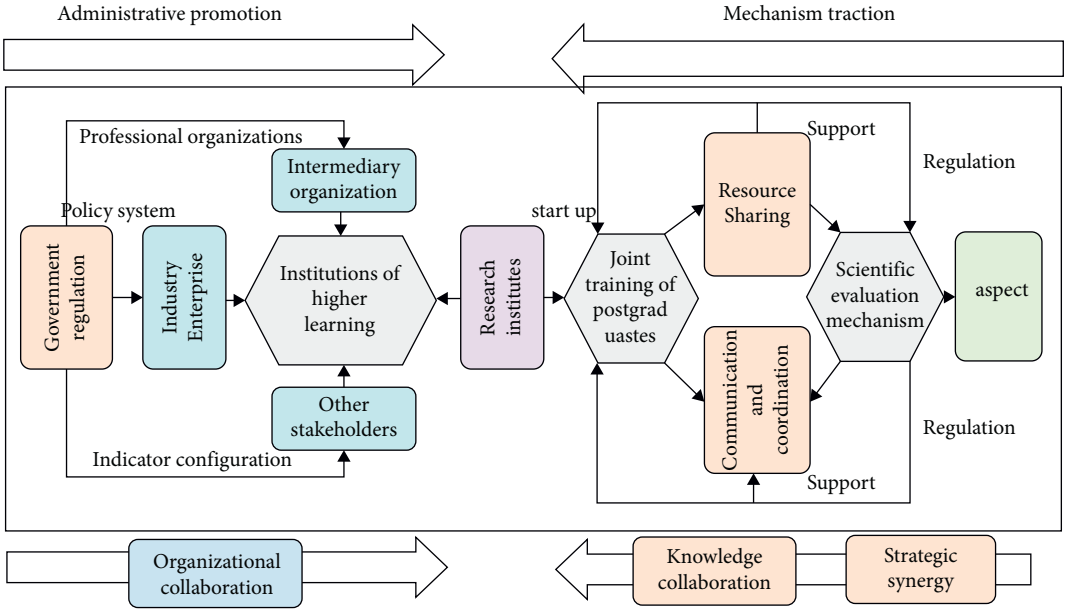


Figure 1: The research logical structure of this paper.

collaborative innovation environment. It believed that the joint training of research and students under the collaborative innovation environment is a high-end talent cultivation paradigm formed under the guidance of the collaborative innovation concept and guided by a sound and scientific mechanism [32]. Model analysis results are shown in Figure 2.

2.2. Reform of Joint Training Mechanism for Graduate Students under the Environment of Collaborative Innovation. The word "Mechanism" originated from Greece, and its original meaning is the relationship between the principle of mechanical braking and the interaction between mechanical internal components and the realization of certain functions through mechanical operation. Generally speaking, the "mechanism" has the following definitions. The structure and working principle, such as the mechanism of the computer; the structure, function and relationship of the organism. The mechanism of arteriosclerosis shows the physical and chemical laws of certain natural phenomena. Such as the mechanism of optimizing objects in the optimization method, also called mechanism. The process of the interaction shows the organization work system, such as market mechanism and competition mechanism. The "mechanism" in the ordinary sense is mostly the fourth interpretation of it in the "Modern Chinese Dictionary." In social science disciplines such as psychology, sociology, political science, and management, "mechanism" generally refers to the internal structure and mode of action that cause and restrict the movement, transformation, and development of things, including the coherent relationship of internal factors of things, and the mutual interaction of various factors such as the form of action, the procedure of function action, and the opportunity for change. The "mechanism" can also be understood from the perspective of the system, which means that the working system is formed by the internal laws of things and their connections with external things. In this system, the study believes that the joint training mechanism under the collaborative innovation environment is defined as the realization of the self-interest of the training subject, and Shanghai is cooperating to cultivate high-end talents with a broad interdisciplinary knowledge background and capable of solving major scientific and strategic issues in national development.

Mechanism reform is the core and key to ensuring the sustainable development of joint training. The reform of the postgraduate joint training mechanism guided by the concept of collaborative innovation pays more attention to the diversity of high-end talent training subjects, the nonlinearity of the training process, and the integration of training resources; more attention paid to opening up the system and mechanism of multiple subjects to solve the impact. This paper analyzes the key issues of student educational development. This paper studies in-depth institutional reform and scientific mechanism design. At present, science and technology, economy, and cultural undertakings have been developed in a highly coordinated manner. The development of science and technology has promoted the long-term linkage between various disciplines and improved the comprehensive quality and innovation ability of postgraduates. The reform of the mechanism is the key to the success or failure of the joint training of industry, university, research, and other subjects. These contents are closely related to the improvement of talent quality and innovation ability in various disciplines. The connotation of the reform of the joint training mechanism under the collaborative innovation environment mainly reflected in four basic aspects: in the development path, from the past administrative-driven development; the type of training transformed into a new type of cross-integration training; in the cooperation paradigm, from the past homogeneous and

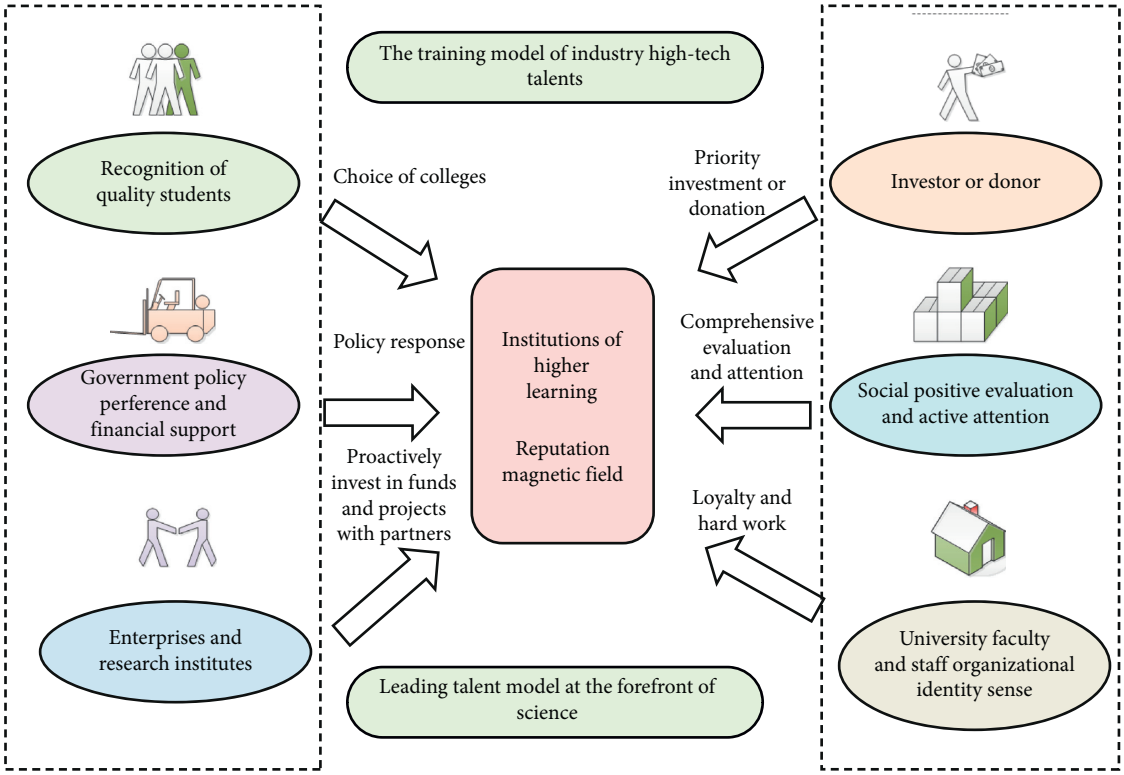


Figure 2: The magnetic field effect on the reputation of colleges and universities.

decentralized cooperation to a new type of large-scale linkage cooperation; in the distribution of benefits, from the past competition for resource acquisition to a new type of create a community of interests. Taking the concept of collaborative innovation as the starting point, it is undoubtedly of important strategic significance to carry out adequate system reform and perfect mechanism design in the multisubject joint training research and student career. The interrelationships between model elements are shown in Figure 3.

2.3. The Reform of Joint Training Mechanism Is the Only way to Realize the Strategic Task of Collaborative Innovation. The strategic task of the "plan" is to "accelerate the reform of colleges and universities mechanism. Transform the innovation methods of colleges and universities, gather and cultivate a group of top innovative talents, produce a group of major landmark achievements, and give full play to higher education as the first productive force of science and technology and the number one talent. The unique role of the important integration point of resources to make greater contributions to the country's innovation and development." As an important part of the reform of the system and mechanism of universities, the reform of the joint training mechanism can fully release the comprehensive disciplines and strong academic atmosphere in universities. We will build the platform for joint training and capability innovation of graduate students with universities as the center and bring together innovative forces such as enterprises Based on equality, integrity, and mutual benefit, resource sharing mechanism, evaluation mechanism, and other systemic institutional innovations. The core of the collaborative innovation strategy is high-end talents, and the core of joint training is high-end talents. Colleges and universities can only seize the important combination of high-end talents, vigorously promote the reform of the postgraduate joint training mechanism, and give play to the synergy of scientific research, industry and education. Collaboration and industry-university collaboration jointly promote the innovation of postgraduate joint training concepts, training models, and training processes, and cultivate high-quality research students suitable for scientific research, industry, and national and regional development needs. Transformation of actual productivity, cultivate top-notch innovative talents, promote the development of interdisciplinary and emerging disciplines, enhance the comprehensive innovation capabilities of universities in knowledge creation, knowledge dissemination and knowledge transfer, and promote the transformation of university development methods.

2.4. The Reform of the Joint Training Mechanism Is to Promote the Needs of Education and Scientific Research. Institutions of higher learning have obvious advantages in subject clusters, academic environment, mentor teams, international exchanges, and knowledge inheritance. Scientific and technological leaders of scientific research institutes, scientific research teams, R&D and manufacturing equipment and innovation platforms, major national topics and industry-leading projects, etc. One aspect is its outstanding

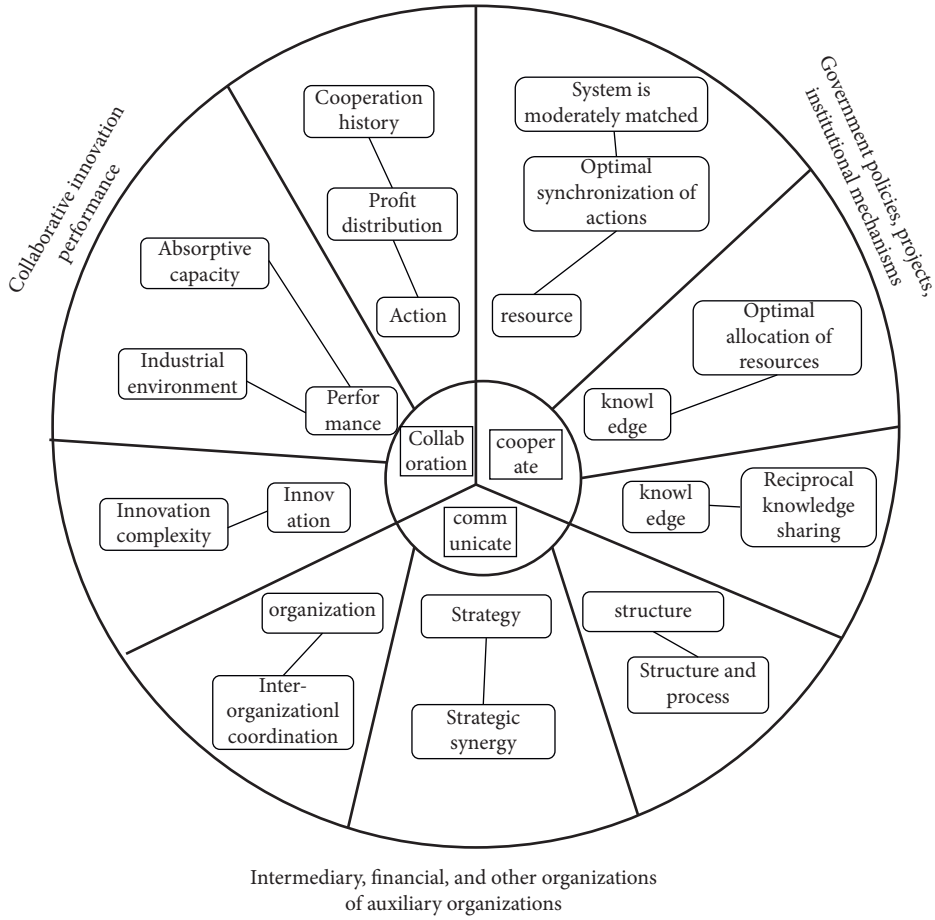


Figure 3: Schematic diagram of the ecological allocation of talent resources for postgraduate joint training.

strengths. The large- and medium-sized national backbone enterprises have abundant resources and strong demands in applied talent teams, R&D funds, demand for high-end talents, technological transformation, and application demands. The basic starting point of the joint training of research and students is to integrate the advantageous resources of the main body of industry-university-research training and jointly cultivate high-end talents that meet the needs of the country, region, and industry. However, in the practice of joint training, "the level of industry-university-research cooperation is not high, the depth of industry-university research cooperation is insufficient, the funds for industry-university-research cooperation are insufficient, the motivation for industry-university-research cooperation is insufficient, and the phenomenon of industry-university-research disconnection still exists" has always existed. The fundamental reason is that the existing joint training mechanism fails to give full play to the educational advantages of colleges and universities, the R&D advantages of scientific research institutes, and the market advantages of industry enterprises.

Therefore, to develop the joint training of research and students under the collaborative innovation environment, we must start from the basic design of the collaborative innovation strategy, take the reform of the joint training mechanism as the starting point, and focus on the major national and regional needs and the core tasks of high-end talent training. The reform of the joint training mechanism is an inherent requirement to improve the quality of high-end talent training. High-end talents have a strong sense of social responsibility, noble personality, and innovative spirit, able to stand on the commanding heights and forefront of international scientific and technological development, gather cross-cultural and interdisciplinary resources, and solve national development. Leading talents and innovation leaders on major scientific and strategic issues "all countries in the world" take research and education as a strategic choice for cultivating high-end talents, achieving national development goals, maintaining international competitive advantages, and a powerful way to seize the commanding heights of science and technology, education, and talent. Now when we actively promote interdisciplinary, cross-organization, cross-field, and cross-border joint training of research students, the quality of joint training models and mechanisms directly determines the quality of national high-end talent training. However, the "quality of postgraduate training in my country, including joint training, is not optimistic, which is mainly reflected in the quality of conditions to be improved, process quality to be improved, and structural quality to be improved." At present, high-end talents lack performance in innovative thinking and innovative results. It mostly regulated by the society, which

closely related to the failure of the existing joint training mechanism for the main body to effectively support high-end talent training.

3. The Applicable Analysis and Model Construction of Joint Training Mechanism Reform

Collaborative innovation strategy is a major project to improve the quality of my country's higher education, especially the quality of high-end talent training. It takes the frontiers of science, cultural inheritance, industry, and regional development as the main body of collaboration and the direction of the convergence of innovative resources and uses the collaborative innovation platform to integrate government, industry, university, research, and application. The advantages of collaborative education of other subjects are in the direction.

3.1. Synergetic Innovation Theory Overview and Connotation Analysis. The conditions and institutions ensure the smooth progress of the joint training of graduate students and provide strong conditions and main directions for the reform of the joint training mechanism. Scientific mechanism design and effective mechanism operation are one of the fundamental guarantees for the quality of high-end talent training. For joint training subjects, only through systematic mechanism incentives and regulations can they stimulate their endogenous motivation to train first-class talents, for graduate students. The sound external and internal mechanism is the guarantee of the quality of learning and scientific research. It imperceptibly draws research students to devote themselves to academic or R&D careers and guides and urges graduate students to become high-end talents. Model analysis results are shown in Figure 4.

The core idea of collaborative innovation theory is "integration" and "interaction." Since the concept of "collaborative innovation" was put forward in 2000, etc. (he has conducted in-depth discussions on the theory of collaborative innovation from the perspectives of "integration" and "interaction," among which collaborative innovation from the perspective of "integration" mainly refers to knowledge, resources, and actions. These involve the integration between different performance factors. The concept of collaborative innovation mainly refers to the sharing of knowledge and the optimal allocation of resources among innovation subjects. These ideas continue to be implemented. Existing studies have continuously refined the innovation process according to the unreasonable positioning of the main elements of innovation in the collaborative process. In this theoretical framework, "communication and coordination" is the basis of "collaborative innovation theory," and its corresponding elements are knowledge and resources. Its meaning means that the development of collaborative innovation activities must first be constructed and perfected.

Shape the general functional relationship between the output y of the injury model and the input $x_1, x_2, \ldots, x_n$. The Kolmogorov–Gabor polynomial is as follows:

$$y = f(x_1, x_2) = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_1^2 + a_4 x_2^2 + a_5 x_1 x_2. \tag{1}$$

And treat each of the monomials as m input models in the original structure of the modeling network:

$$\begin{aligned} v_1 &= a_0, \\ v_2 &= a_1 x_1, \\ v_3 &= a_2 x_2, \ldots, \\ v_6 &= a_5 x_1 x_2. \end{aligned} \tag{2}$$

The final information $i_t \times C_t'$ is expressed as the value that can be obtained C_t from the output information of the joint forgetting gate:

$$C_t = f_t * C_{t-1} + i_t * C_t'. \tag{3}$$

The calculation method is

$$\begin{aligned} O_t &= \sigma\left(W_o \cdot [h_{t-1}, x_t] + b_o\right), \\ h_t &= o_t * \tan h\left(C_C\right). \end{aligned} \tag{4}$$

As a generalization of ordinary linear model, GLM introduces connection function in the model in order to fit some nonlinear relationships. The model can be expressed as

$$g(\xi) = g(\sigma) + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n, \tag{5}$$

where $g(\sigma)$ is the connection function, $\sigma = E(Y)$.

$$g(\xi) = a + f_1\left(X_1\right) + f_2\left(X_2\right) + \cdots + f_n\left(X_n\right). \tag{6}$$

The function of forgetting gate is to determine the part discarded from the input information h_{t-1} and x_t and output a value between 0 and 1. The larger the value, the more information is retained. The output of forgetting gate is calculated as follows:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right). \tag{7}$$

The "communication and coordination" mechanism of knowledge and resources is followed by collaborative innovation. Participating organizations' cooperation at the action level will eventually achieve synergy at the performance level. Through the aforementioned research, we know that the results of "collaborative innovation" include science, technology, culture, etc. The output of tangible innovation results also includes the collaborative training and joint forging of innovative talents. Therefore, from this perspective, "joint training" can regarded as a sub-part of collaborative innovation activities in a broad sense. Joint development carried out in a collaborative innovation environment. Cultivation activities must follow the basic framework of the collaborative innovation theory, and the construction of the "communication and coordination" mechanism is the primary task of the joint training of graduate students. "Communication and coordination" is the joint training activities of graduate students in a collaborative innovation environment to achieve resource

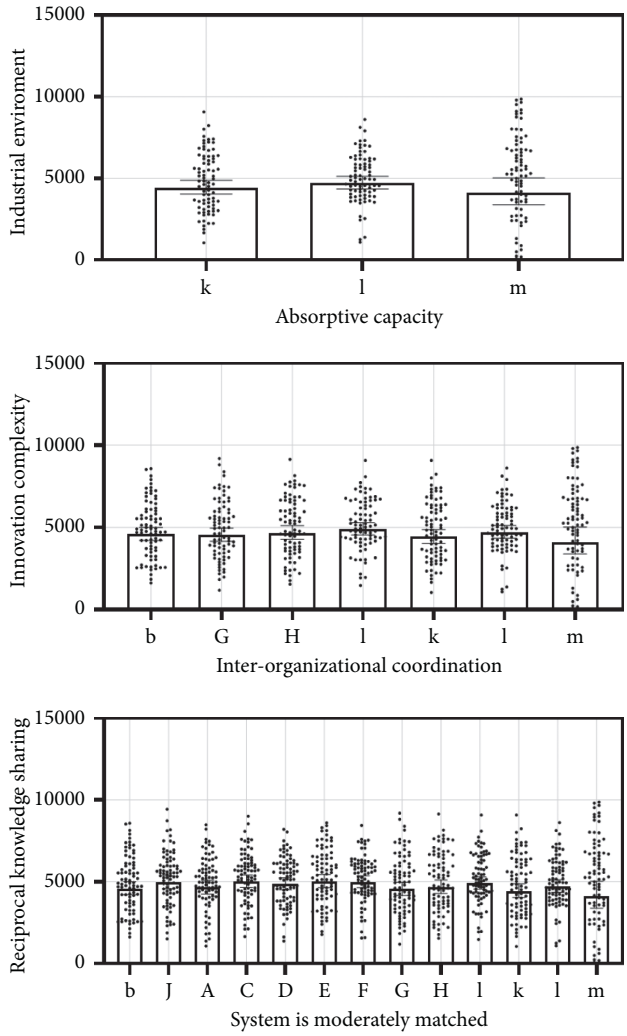


Figure 4: The applicability of collaborative innovation theory in joint training.

sharing and knowledge transfer. Model analysis results are shown in Figure 5.

The "collaborative innovation" formed by enterprises, universities, and scientific research institutions in the promotion of the flow, sharing and integration of knowledge and resources between organizations must carried out on the three levels of strategic coordination, knowledge coordination, and organizational coordination. The core layer: it is the synergy of the elements of strategy, knowledge, and organization. Government policy guidance, project promotion, and institutional incentives are the support layer. Intermediaries, financial institutions, and other organizations can regarded as the support layer. Due to the different cooperation motivations, fundamental tasks, and resource capabilities between enterprises and universities. The purpose of its cooperation with enterprises is to obtain corporate funding to produce academic results that have a boosting effect on human society. From this point of view, the application culture of the enterprise and the research culture of universities and research institutes. It is difficult to integrate its cooperation, and it is difficult to carry out automatically. Model analysis results are shown in Figure 6.

3.2. Synergetic Innovation Theory Overview and Connotation Analysis. However, this kind of "difficult to integrate" or "mutual exclusion" is inevitable. The real obstacle to the cooperation of all parties is the lack of appropriate collaborative performance evaluation tools or evaluation systems. In the process of "knowledge synergy," collaborative innovation activities generally have the following characteristics: systematic. In this ecosystem, the integration or blending process of various elements is not a simple superposition of functions, but an effective reconciliation of the functions of each part. The form of sex and integration expresses the ways, functions, and goals of each part after cooperation. The level of "collaborative innovation" manifested in two aspects: one is the level and step of knowledge flow. Generally, the higher level knowledge layer flows to the lower level knowledge layer, and knowledge layers of the same level and different properties have mutual influence. The second is the level and style of organizational characteristics. Organizations participating in collaborative innovation activities must have a relative advantage in a certain field, and the organizational style is open. So as to ensure the

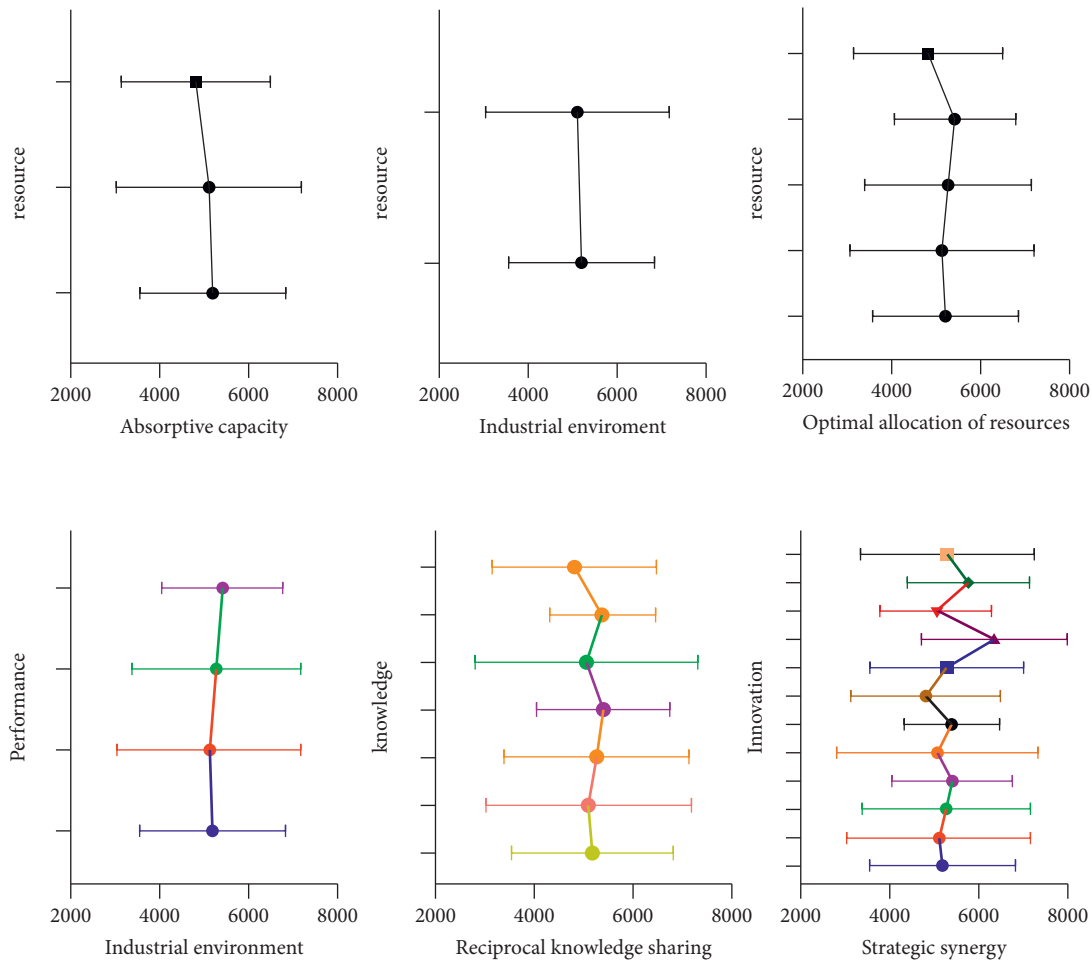


Figure 5: Model of joint training mechanism for graduate students in an innovative environment.

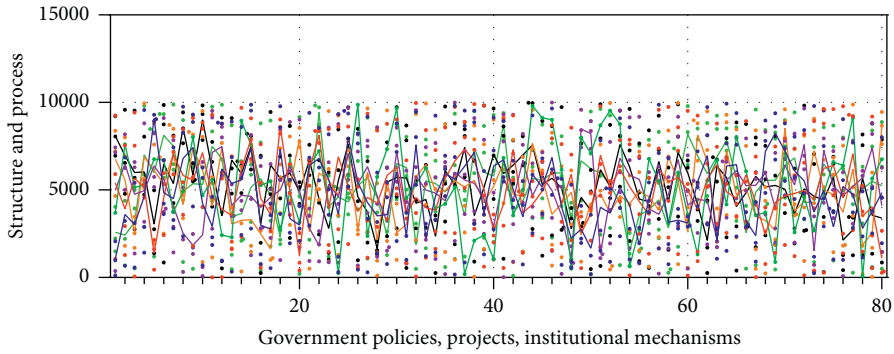


Figure 6: New public management theory in jointly cultivating government regulation mechanism.

interaction and complementarity of "collaborative innovation." Model analysis results are shown in Figure 7.

Dissipative: the innovation ecosystem will exchange information, energy, and material with the outside. "Collaborative innovation" is a new stable and orderly structure formed under certain conditions by an open system that is far from equilibrium and exchanging material and energy with the outside world. There are nonlinear interactions within the system. Dynamic: in collaborative innovation, there are a variety of interweaving and compatible behaviors of new ideas, new concepts, continuous updates, rapid response, flexible adaptation, and creative innovation.

The government, industry, university, and their innovative elements are diverse, and there are complex interactions and mutual influences among them. Various emergencies may also occur. It is unpredictable by the participating organizations. In addition, the results of innovation include both tangible and intangible results. The

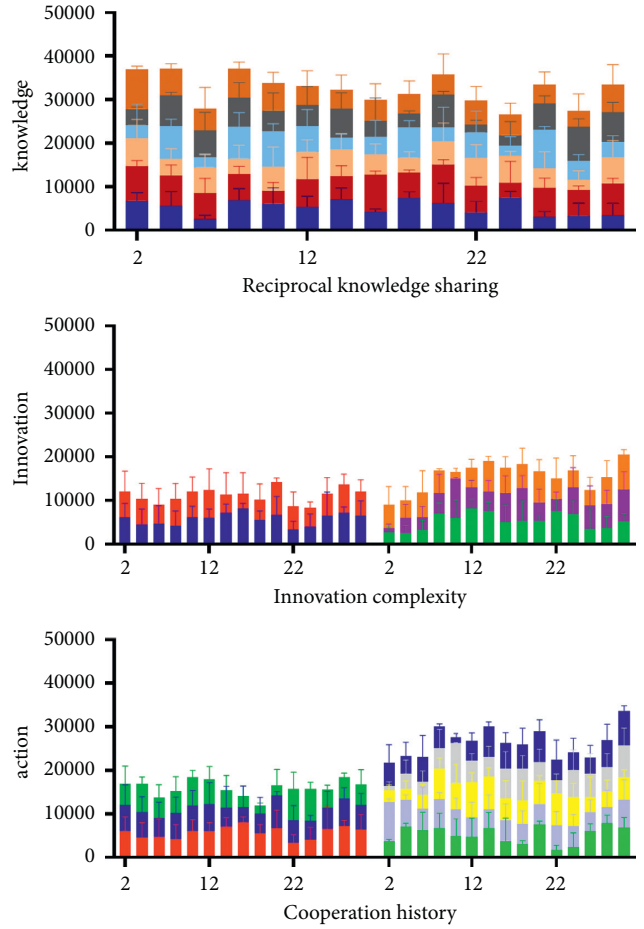


Figure 7: Systematic integration theory in joint cultivation, communication, and coordination mechanism.

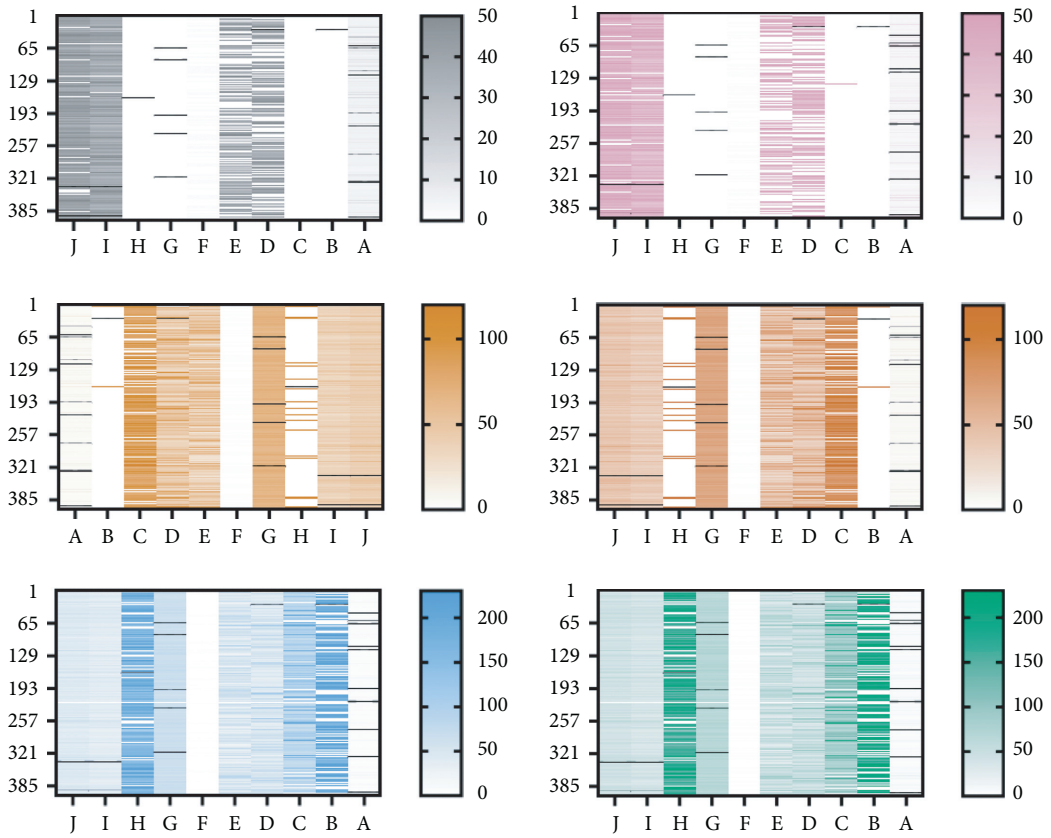


Figure 8: Resource dependence theory and its applicability in joint training resource sharing mechanism.

operation of collaborative innovation requires a complete mechanism system. Judging from the existing research results, the collaborative innovation mechanism system includes an incentive mechanism centered on the fair distribution of benefits, a sharing mechanism based on knowledge exchange, resource complementation, and subject trust. The platform based on the collaborative innovation center as the main body. The performance evaluation mechanism based on the reasonable setting of evaluation indicators and the government policy system aiming at adapting innovation policies to the environment. The setting of the mechanism system also profoundly affects the mechanism reform path of joint training activities under the environment it forms. Model analysis results are shown in Figure 8.

Organizational collaboration refers to collaboration and integration between the main bodies of industry, academia, and research (outside the organization) between science and technology R&D. The talent training business units (inside the organization), with clear technological innovation, R&D, and application, The talent training strategy and scientific governance process ensure the internal and external coordination of each subject and the process or state of continuous cyclic execution.

The calculation of the statistics of the hypothesis of the single-body sample is as follows:

$$t = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{\left((n_1 - 1)S_1^2 + (n_2 - 1)S_2^2/n_1 + n_2 - 2\right)\left((1/n_1) - (1/n_2)\right)}}. \tag{8}$$

Set two random sequences X and Y, Pearson correlation coefficient between the two sequences is r, then

$$r = \frac{\operatorname{cov}(X, Y)}{\sqrt{\sigma_x^2}\sqrt{\sigma_y^2}} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 (y_i - \overline{y})^2}}. \tag{9}$$

In order to ensure the accuracy of the results, this paper uses two evaluation indexes, mean absolute error and root mean square error, to evaluate the optimization effect of the model. The specific calculation formulas are as follows:

$$\begin{aligned} \mathrm{MAE} &= \frac{1}{s}\sum_{i=1}^{s}\left|\hat{y}_i - y_i\right|, \\ \mathrm{RMSE} &= \sqrt{\frac{1}{s}\sum_{i=1}^{s}\left[\hat{y}_i - y_i\right]^2}. \end{aligned} \tag{10}$$

The calculation formula of single-sample statistics is as follows:

$$t = \frac{\overline{X} - \mu}{\left(\delta_x/\sqrt{n-1}\right)}. \tag{11}$$

Over the years, the differences in the value positioning, strategic goals, cooperation motivation of the joint training subjects, and the contradictions of their internal systems in the organizational structure, power structure, and functional structure have caused the sustainable development of cooperative education to face a series of difficulties. The fundamental reason is that the "modular" organizational operation and cooperative ecology cannot adapt to the collaborative needs of joint graduate training under the new situation. Model analysis results are shown in Figure 9.

3.3. The Organizational Collaborative Ecosystem Layer in the Collaborative Innovation Theory. Based on the above-mentioned current situation, the collaborative innovation strategy is to pay more attention to the main position and linkage effect of the partners under the government's macro-control and focus on breaking through the institutional mechanism of joint training of various main organizations. Implementation of high-level performance in assessment, international cooperation, and innovative culture construction. The internal reform of the school, through the establishment of a coordination mechanism, a resource sharing mechanism and a benefit distribution mechanism in which participants have common goals, internal motivation, and direct communication to achieve external collaboration in the organization. The "collaborative innovation center" used as a common value platform for cooperation in running schools. Education maintenance, industry maintenance, service maintenance, and public maintenance (government maintenance) form a network system for technological innovation, R&D reference and talent training, and mutual knowledge sharing, resource optimal allocation, optimal synchronization of actions, optimal matching of organizations, and open information sharing Win the joint cultivation of "ecosystem." Model analysis results are shown in Figure 10.

Strategic synergy refers to the integration of talent, capital, information, technology, market, and other innovative elements of joint training subjects. The method improves the efficiency and overall value of talent training, in the establishment of organizational structure, talent quality assurance, teacher resource allocation, and innovation platform application. The unified deployment and overall planning of technological innovation, R&D, and application. It is easy for the subjects of joint training to reach a consensus on the "organizational synergy" at the meson level and the "knowledge synergy," while it is difficult to form a unified opinion on the macro level "strategic synergy." Universities, research institutes, industries, or companies in finance. The six dimensions of technology, strategy, education, politics, and theory have different motivations for cooperation in forty-six indicators. More importantly, the lack of a scientific evaluation mechanism has led to joint training of multiple organizations at the level of talent training quality and work performance. There are often differences at the level. Model analysis results are shown in Figure 11.

Collaborative innovation strategy provides a suitable institutional environment for the joint cultivation of strategic

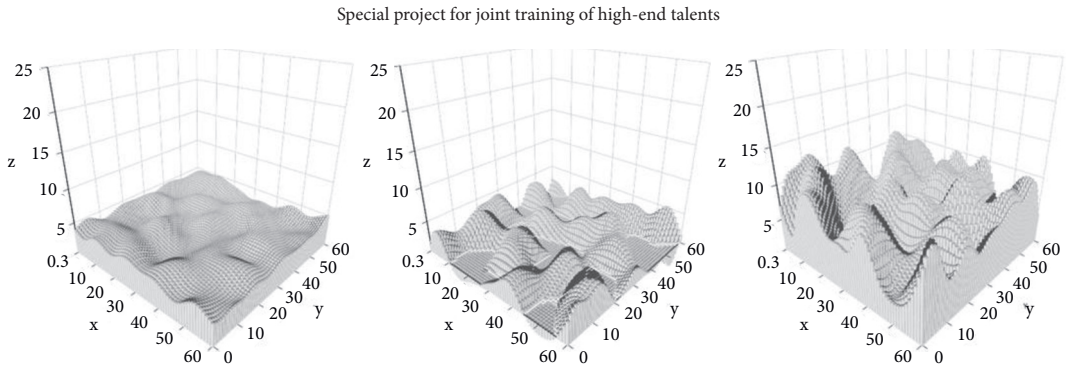


Figure 9: Scientifically formulate joint training enrollment plan and distribution plan.

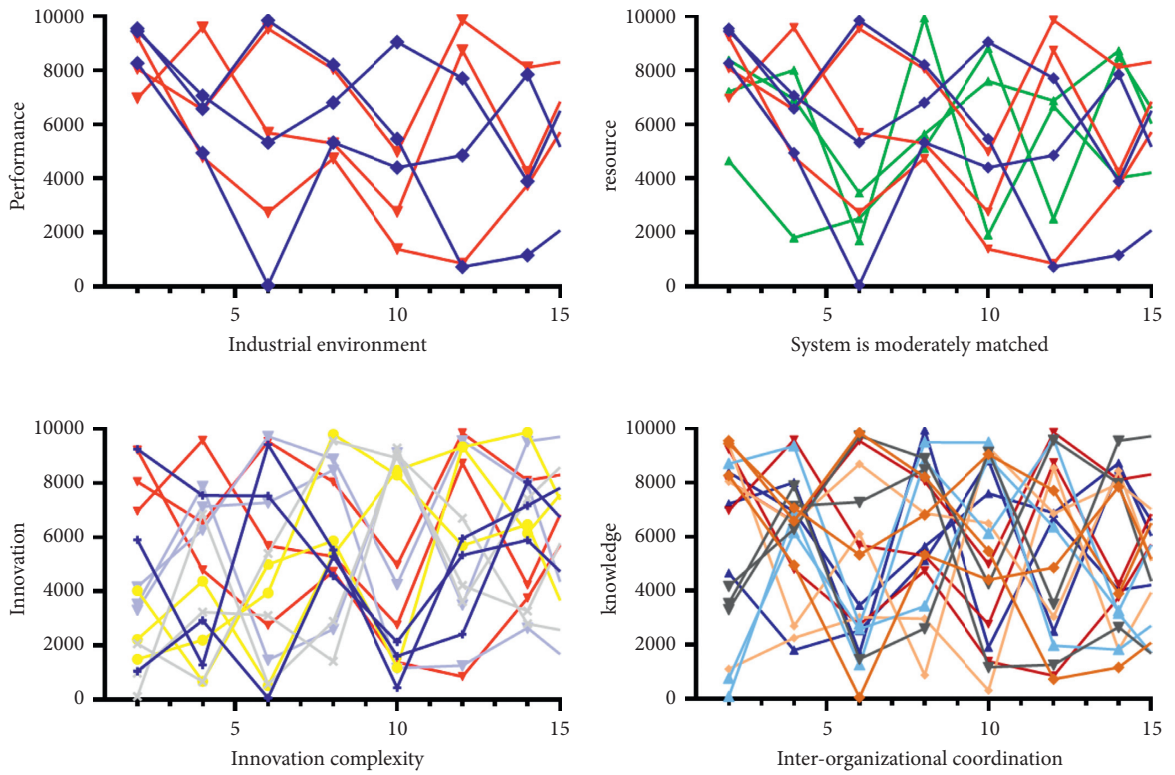


Figure 10: Scientifically plan the subprojects and types of joint training.

synergy among multiple organizations. First, it uses the collaborative innovation center as the platform and the cultivation of top-notch innovative talents as the carrier. It is committed to promoting the integration of the research culture of universities and enterprises, which create a value identity combination for the joint training of graduate students. This lays the conceptual foundation for the implementation of the scientific evaluation mechanism in multiple organizations. Second, the goal of the collaborative innovation strategy to bring together innovative resources is to serve the major needs of the country's science and technology. Support and build a long-term mechanism for sharing results to achieve a win-win cooperation among all subjects in terms of scientific and technological innovation, talent training, and economic development. This provides a driving force for the benefit of collaborative education work, which provides a scientific evaluation mechanism within a diverse organization. The implementation laid the foundation for power.

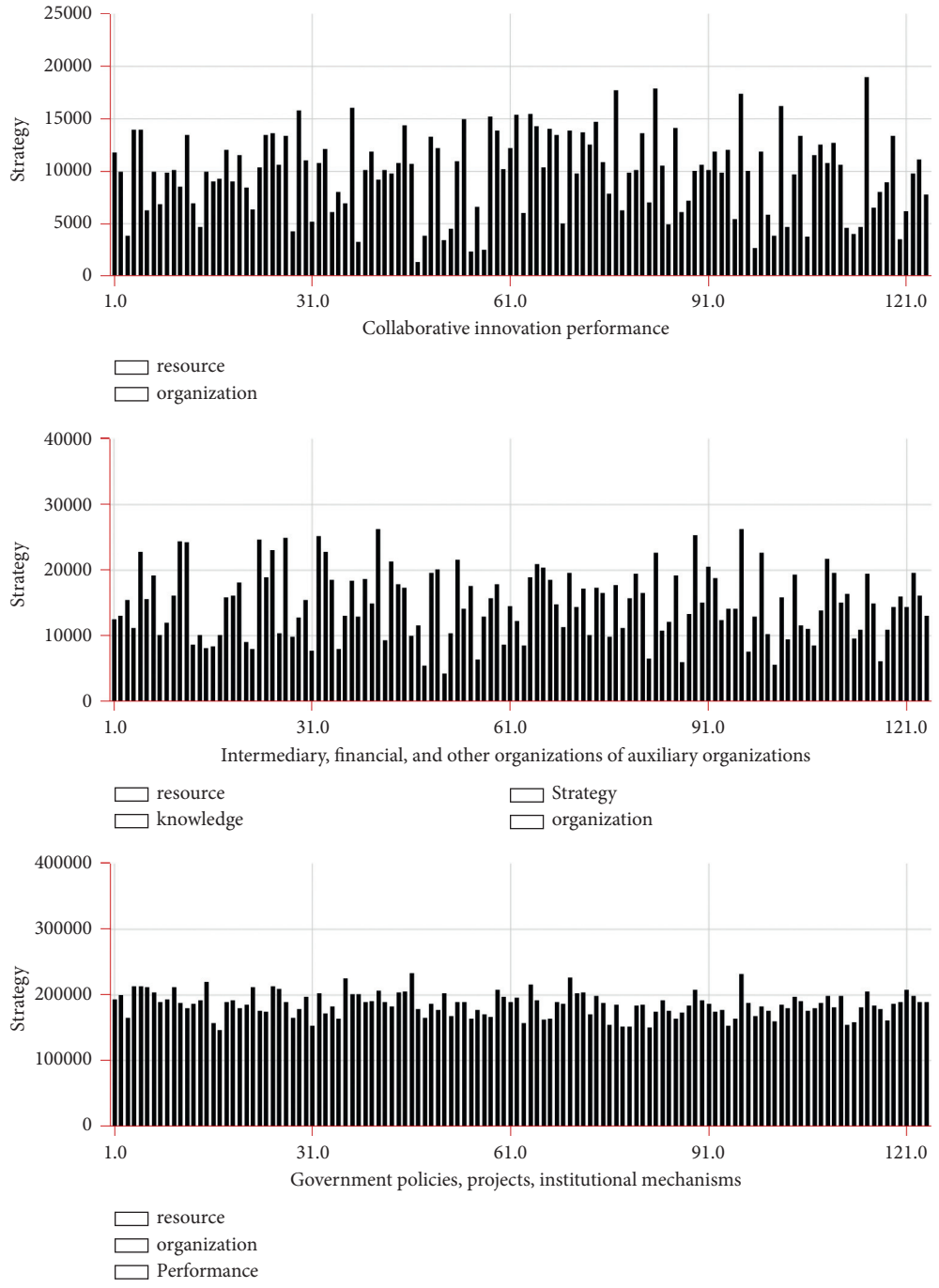


Figure 11: Implement dynamic management of joint training special projects.

4. Conclusion

After combing the development history, evolution trend and the basic requirements of the joint training of Chinese graduate students under the environment of collaborative innovation in the whole process, this study found that there are still some problems in the process of mechanism construction that are not suitable for research under the new situation. The endogenous development of the joint training cause urgently needs to be guided by a systematic theory to construct a complete reform plan for government regulation, communication and coordination, resource sharing, and scientific evaluation.

(1) Conducive to the reform and development of local applied universities. Local application-oriented colleges and universities developed to a certain stage of economic and social development in order to meet

the needs of China's economic modernization and the needs of higher education popularization. Its appearance is a microcosm of the transition from elite education to popular education. Different from research universities, applied universities insist on implementing applied undergraduate education, always based on market needs, rooting in industries, serving localities, and making contributions to the regional economy. However, after long-term development, the college itself has exposed many problems, such as unclear understanding of the school's philosophy, fuzzy positioning, inaccurate positioning, convergence of development models, backward teaching methods, improper teaching methods, and teacher team lack of engineering practice experience and single evaluation system. The talents cultivated by the school consistently show insufficient innovation ability, weak hands-on ability, insufficient learning ability, and fail to keep up with the new needs of the development of the enterprise (industry). To solve the above problems, local application-oriented colleges and universities need to develop in-depth cooperation and exchanges with enterprises, understand the trends of enterprises in the first place, and be familiar with what kind of talents are needed for the development of enterprises, and what types are needed for the development of society. Then, relying on these needs to formulate the school's talent training standards. It will help the school cultivated people to be used by the society, enterprise engineering, and technical experts are invited as consultants to participate in the school's teaching plans and curriculum settings.

(2) Conducive to improving the quality of training of outstanding Chinese engineers. The training of outstanding engineers is concentrated on the three levels of undergraduate, master, and doctoral degrees. The purpose of implementing the "Excellence Plan" at the undergraduate level is to pass practical training and theoretical teaching so that students can adapt to the development requirements of the society in the shortest time, with strong learning ability and practice. High-level engineering and technical talents with strong hands-on ability serve the country to take a new road to industrialization, implement the strategy of strengthening the country with talents, and build an innovative country. School-enterprise cooperation is the key to the success of the "Excellence Project." The joint training of talents by schools and enterprises is also a new demand for social development. Accurately grasp the social demand for engineering talents, from the basic qualities of engineers, engineering awareness, learning ability, and innovation Ability, practical ability and management ability, etc. start to train outstanding engineers to improve the overall quality of outstanding engineers in China.

(3) Conducive to the continuous development of students' careers. Since China's accession to the WTO, China's economic development has shown a prosperous scene. With the development of China's economy, applied-skilled talents who can only master technology and serve as the front line of production can no longer meet the needs of society. What society needs more are innovative engineering science and technology talents who can use the principles and methods of technological cross-convergence, the use of technological integration and innovation technology to carry out product creative design and development, and at the same time take into account engineering management and consulting. These talents not only have a solid theoretical foundation knowledge and basic scientific and humanistic quality but also can combine the actual operation of the enterprise to propose and solve the actual problems of the enterprise and carry out technological innovation. The school-enterprise cooperation to build a training base is precisely to allow students to have more opportunities to have access to the company's advanced machinery and equipment, understand and master the company's most advanced production technology, and continue to accumulate practical experience in the process of internship and training. Continuously improve their comprehensive practical ability and employment competitiveness.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

The authors thank Social Science Project "THE 13TH FIVE-YEAR PLAN" of Jilin Education Department Construction of practical teaching system of industrial engineering major in provincial universities under resource constraints (No. JJKH20191224SK).

References

[1] J. E. Varajao, "A new process for success management-bringing order to a typically ad-hoc area," *Mod. Proj. Manag*, vol. 5, no. 3, pp. 94–99, 2018.

[2] C. M. Kang, S.-H. Lee, and C. C. Chung, "Multirate lane-keeping system with kinematic vehicle model," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 10, pp. 9211–9222, 2018.

[3] K. Alexiou and J. Wiggins, "Measuring individual legitimacy perceptions: scale development and validation," *Strategic Organization*, vol. 17, no. 4, pp. 470–496, 2019.

[4] J. Barrena-Martinez, M. López-Fernández, and P. M. Romero-Fernández, "The link between socially responsible human resource management and intellectual capital," *Corporate Social Responsibility and Environmental Management*, vol. 26, no. 1, pp. 71–81, 2019.

[5] H. Aguinis, Y. H. Ji, and H. Joo, "Gender productivity gap among star performers in STEM and other scientific fields," *Journal of Applied Psychology*, vol. 103, no. 12, pp. 1283–1306, 2018.

[6] S. Banerjee and S. Venaik, "The effect of corporate political activity on MNC subsidiary legitimacy: an institutional perspective," *Management International Review*, vol. 58, no. 5, pp. 813–844, 2018.

[7] A. Edrees, H. Abdelhamed, S. W. Nho et al., "Construction and evaluation of type III secretion system mutants of the catfish pathogenEdwardsiella piscicida," *Journal of Fish Diseases*, vol. 41, no. 5, pp. 805–816, 2018.

[8] M. G. Mayhew, J. Gardner, and N. M. Ashkanasy, "Measuring individuals' need for identification: scale development and validation," *Personality & Individual Differences*, vol. 49, no. 5, pp. 356–361, 2010.

[9] T. Fischer and C. Krauss, "Deep learning with long short-term memory networks for financial market predictions," *European Journal of Operational Research*, vol. 270, no. 2, pp. 654–669, 2018.

[10] S. K. Dwivedi, R. Amin, and S. Vollala, "Blockchain-based secured event-information sharing protocol in Internet of vehicles for smart cities," *Computers & Electrical Engineering*, vol. 86, 2020.

[11] S.-M. Hosseininasab, S.-N. Shetab-Boushehri, S. R. Hejazi, and H. Karimi, "A multi-objective integrated model for selecting, scheduling, and budgeting road construction projects," *European Journal of Operational Research*, vol. 271, no. 1, pp. 262–277, 2018.

[12] Z. Yang and L. S. C. Pun-Cheng, "Vehicle detection in intelligent transportation systems and its applications under varying environments: a review," *Image and Vision Computing*, vol. 69, pp. 143–154, 2018.

[13] M. Guo and N. Arunkumar, "Construction of employee training program evaluation system of three exponential forecast based on sliding window," *Cluster Computing*, vol. 22, no. 3, pp. 6865–6870, 2019.

[14] F. Sadile, A. Bernasconi, F. Carbone, F. Lintz, and G. Mansueto, "Histological fibrosis may predict the failure of core decompression in the treatment of osteonecrosis of the femoral head," *International Journal of Surgery*, vol. 44, no. Spec, pp. 303–308, 2017.

[15] S. Schnelle, J. Wang, R. Jagacinski, and H.-j. Su, "A feedforward and feedback integrated lateral and longitudinal driver model for personalized advanced driver assistance systems," *Mechatronics*, vol. 50, pp. 177–188, 2018.

[16] E. M. A. Ahmed, "A hydrologic-economic-agronomic model with regard to salinity for an over-exploited coastal aquifer," *Journal of Geosciences*, vol. 12, no. 12, pp. 1–12, 2019.

[17] L. Ye and T. Yamamoto, "Modeling connected and autonomous vehicles in heterogeneous traffic flow," *Physica A: Statistical Mechanics and its Applications*, vol. 490, no. 40, pp. 78–81, 2018.

[18] P. Alessio and C. Peter, "Hallmark robert.prolonging the lifetime of old steel and steel-concrete bridges: assessment procedures and retrofitting interventions," *Structural Engineering International*, vol. 29, no. 4, pp. 507–518, 2019.

[19] Z. Khan and S. Amin, "Bottleneck model with heterogeneous information," *Transportation Research Part B: Methodological*, vol. 112, no. 1, pp. 157–190, 2018.

[20] M. D. Moreno, "Translation quality gained through the implementation of the iso en 17100:2015 and the usage of the blockchain," *Babel*, vol. 1, no. 2, pp. 1–9, 2020.

[21] T. Van Asch, W. Dewulf, F. K. Ivan, C. Eddy, and V. d. Voorde, "Cross-border e-commerce logistics-Strategic success factors for airports . Research in Transportation Economics," pp. 167–192, 2017.

[22] Á. Valarezo, T. Pérez-Amaral, T. Garín-Muñoz, I. Herguera García, and R. López, "Drivers and barriers to cross-border e-commerce: e," *Telecommunications Policy*, vol. 42, no. 6, pp. 464–473, 2018.

[23] A. Jazairy, J. Lenhardt, and R. von Haartman, "Improving logistics performance in cross-border 3PL relationships," *International Journal of Logistics Research and Applications*, vol. 20, no. 5, pp. 491–513, 2017.

[24] G. Alexandridis, G. Siolas, and A. Stafylopatis, "Enhancing social collaborative filtering through the application of non-negative matrix factorization and exponential random graph models," *Data Mining and Knowledge Discovery*, vol. 6, pp. 1–29, 2017.

[25] Y. Qingwen, "The construction mechanism and algorithm of cross border E-commerce export logistics mode from the perspective of value chain," *Journal of Intelligent and Fuzzy Systems*, vol. 37, no. 3, pp. 3393–3400, 2019.

[26] H. R. Boveiri, R. Khayami, M. Elhoseny, and M. Gunasekaran, "An efficient Swarm-Intelligence approach for task scheduling in cloud-based internet of things applications," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 9, pp. 3469–3479, 2019.

[27] M. P. André Marchand, "Automated product recommendations with preference-based explanations," *Journal of Retailing*, vol. 7, no. 1, pp. 48–52, 2020.

[28] S. Kant and T. Mahara, "Merging user and item based collaborative filtering to alleviate data sparsity," *International Journal of System Assurance Engineering & Management*, vol. 9, no. 1, pp. 1–7, 2018.

[29] W. Alnumay, U. Ghosh, and C. Pushpita, "A trust-based predictive model for mobile ad hoc network in internet of things," *Sensors*, vol. 7, no. 1, pp. 142–146, 2019.

[30] A. Marchand and P. Marx, "Automated product recommendations with preference-based explanations," *Journal of Retailing*, 2020.

[31] X. Wang and P. Lei, "Does strict environmental regulation lead to incentive contradiction? - e," *Journal of Environmental Management*, vol. 269, Article ID 110632, 2020.

[32] S. Yue, R. Lu, H. Chen, and J. Yuan, "Does financial development promote the win-win balance between environmental protection and economic growth?" *Environmental Science and Pollution Research*, vol. 25, no. 36, pp. 36438–36448, 2018.

Hindawi
Scientific Programming
Volume 2023, Article ID 9791531, 1 page
https://doi.org/10.1155/2023/9791531

Retraction

Retracted: Design of Packaging Design Evaluation Architecture Based on Deep Learning

Scientific Programming

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] L. Shi, "Design of Packaging Design Evaluation Architecture Based on Deep Learning," *Scientific Programming*, vol. 2022, Article ID 4469495, 8 pages, 2022.

Hindawi
Scientific Programming
Volume 2022, Article ID 4469495, 8 pages
https://doi.org/10.1155/2022/4469495

Research Article

Design of Packaging Design Evaluation Architecture Based on Deep Learning

Lei Shi

Dalian Polytechnic University, Dalian 161034, Liaoning, China

Correspondence should be addressed to Lei Shi; shilei@dlpu.edu.cn

Received 9 November 2021; Revised 20 January 2022; Accepted 21 February 2022; Published 16 March 2022

Academic Editor: Baiyuan Ding

Most researchers use visual communication symbols to achieve the purpose of information dissemination, which is also a very important marketing tool for the current era of packaging design. And the use of visual communication technology to make better product packaging design has become one of most the important means for major enterprises to sell their products and construct a good brand image. In this paper, we use a deep CNN-based aesthetic classification method for splash screens and a deep learning-based NIMA neural network to predict the aesthetic evaluation distribution of splash screen images, respectively. The connotation of visual communication and packaging design and the impact of the role of visual communication technology on packaging design are analyzed.

1. Introduction

In essence, packaging design is a kind of visual symbol transmission. It not only gives products a better aesthetic effect but is also an important means of product promotion, with both instrumental and rational characteristics. In this day and age, visual communication techniques are highly valued and are gradually showing diversified development. Packaging design is also an important means of product marketing [1]. As people's pursuit of material wealth continues, consumers' view of consumption has become more and more open, and how major enterprises can attract consumers' attention and capture their hearts, packaging design has become an important marketing tool for products [2]. Therefore, the application of visual communication technology in product packaging design is particularly important, and visual communication can be directly and effectively shown to have the effect of visual information transmission, which also determines the first impression of consumers of the product [3].

Visual communication is the direct purpose of visual communication design, is through the logo, typography, painting, graphic design, illustration, color, electronic devices, and other two-dimensional image performance to the public to convey a variety of visual information. Visual communication is more inclined to interactive design, focusing on interactive experiences and interactive feelings. Its focus is on functionality, but it also has a graphic design, color matching, and so visual communication also covers the visual aesthetics of the content [4, 5].

Packaging is a complete reflection of the brand concept, product characteristics, consumer psychology, and meeting the consumer's desire to buy [6]. Therefore, packaging design is a combination of art and natural science, applied to the protection and beautification of product packaging. Packaging design is not a broad sense of "art", nor is it just decoration, but contains a multifunctional embodiment of science, art, materials, economics, psychology, and market and other comprehensive elements. Packaging design includes the following three aspects: packaging design, packaging structure design, and packaging decoration design.

Excellent packaging design is the organic unification of the above three. Only the organic unification of the three can give full play to the role of packaging design, and packaging design not only involves two fields of technology and art, but it also involves other related disciplines in their respective fields. Therefore, to design good packaging [7]. We should

apply visual communication to package design and grasp consumer psychology for design [8].

Based on the existence of such a consensus tendency, an emerging field of computer vision, computable image aesthetics, has emerged, whose research aims to enable computers to simulate human vision and aesthetic thinking, thereby making aesthetic decisions about images and building a bridge between computers and visual artworks [9]. Through the calculation and evaluation of image aesthetics, it can predict the aesthetic feelings of users when using visual interactive systems and then help designers to judge and obtain aesthetic expressions that match users' psychological feelings, which are important for achieving positive human-computer interaction. In this paper, we take splash screen images as the research object, use the user's subjective aesthetic rating of splash screen images as the basis and use a deep learning method to simulate the user's aesthetic perception of images and verify the feasibility of evaluating the aesthetics of works through computer image aesthetics evaluation to assist designers.

1.1. The Importance of Visual Communication Technology in Packaging Design. Nowadays, in order to highlight the freshness and personalized features of the products, most of them will choose some colors related to the product development trend as the main color and show the characteristics of the products through the main colors related to the products, while adding some other colors as auxiliary colors so as to set off the freshness of the products. Monotonous and uniform color schemes make it difficult for consumers to be impressed by the product, and it is easy for people to ignore the cultural elements conveyed by the product and produce visual fatigue. Thus packaging color will have a greater impact on the development of the product [10]. Therefore, the color design of the packaging is a prerequisite for consumers to see the superiority of the product. The current product packaging design needs to be more bold and innovative in color matching, and designers should continue to inspire themselves and broaden their creative thinking to capture consumers' emotional tendencies through packaging color and improve the rendering power of the product [11].

The graphic design of the package is the most prominent design in the whole package design. Now many products in the market will be carefully considered before the product packaging pattern design. At the same time, this aspect of the excellent designer is also relatively scarce. Not only in the packaging graphics and patterns, but also in the overall LOGO of the product, is a major focus. The visual communication design will use LOGO and product patterns and the combination of the entire cover with more personalized and visually appealing graphics to highlight the theme of the product. The use of LOGO and packaging patterns will show some specific things in order to attract consumers' attention, so that people associate and deepen the impression of the product [12, 13].

2. Research Status

Among mainstream methods for image aesthetic quality assessment, they can be divided into traditional aesthetic assessment methods based on artificial design features and the currently popular aesthetic assessment methods based on deep learning.

In the method of evaluating aesthetics based on artificial design features, image aesthetics are mainly evaluated by expert manual design of low-level visual features, high-level aesthetic features, and composition aesthetic features (see Figure 1(a)). As a pioneer first proposed the relationship between computer vision features and image aesthetics, based on the basic aesthetic principles such as color matching and contrast of images, images were classified into two categories of high and low aesthetics by methods such as support vector machine and regression. Wu et al. [14] used low-level features to learn classification models to distinguish photographic images of professional photographers from those of ordinary users. Han et al. [15] developed a method to assess the aesthetic quality of images based on color coordination. Kumar [16] selected high-quality images based on image layout, scene, and natural lighting conditions. Domestic scholars have also made many contributions to image aesthetics assessment. Liu et al. [17] extracted low-level visual features, high-level aesthetic features, and visual area features from the overall area and visually critical areas of an image and established an image aesthetic classifier and an aesthetic score assessment model.

In such evaluation methods, which usually involve training and test sets consisting of high-quality and low-quality images, regression analysis of the extracted features against a human aesthetic quality score is required to distinguish high-quality images [18]. However, this requires the researcher to have expertise in photographic aesthetics such as composition and color.

In recent years, with the rise of deep learning techniques, researchers have introduced convolutional neural networks (CNN) to solve the related problems in image aesthetic evaluation tasks. Due to its powerful automatic learning capability, it can automatically extract high-level abstract features from a large amount of image data without requiring researchers to have specialized aesthetic knowledge (see Figure 1(b)), and has become a mainstream approach to solve image aesthetic evaluation problems [19]. They adapted convolutional neural networks to make them applicable to solving different image aesthetic evaluation problems [20] and proposed a deep convolutional neural network with RS-CJS. Fudan University proposed an aesthetic image reviewer model, NAIR, based on CNN and recurrent neural network (RNN), which not only predicts aesthetic ratings but also generates semantic evaluations. These studies have shown good performance in image aesthetic evaluation.

Previous research has mostly focused on photographic images as the main object of aesthetic evaluation, and researchers have developed various algorithms and programs to improve the accuracy of evaluation and help users filter and optimize photographic images. However, for designers,

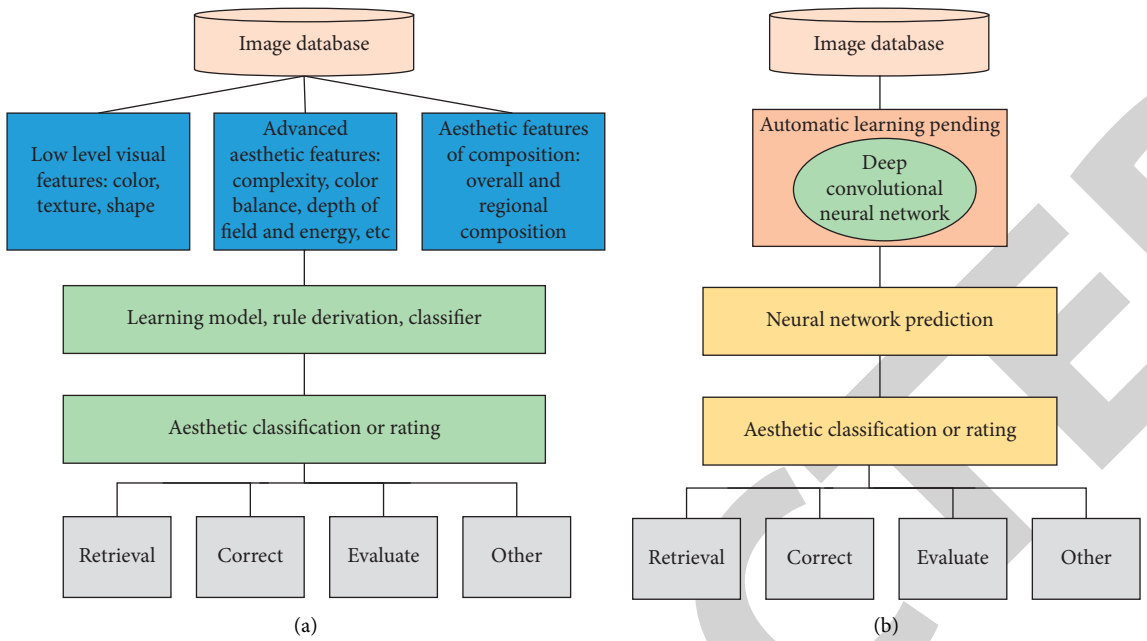


Figure 1: Two aesthetic quality assessment methods: (a) aesthetic evaluation method based on artificial design feature and (b) aesthetic evaluation method based on deep learning.

it is more meaningful to understand the precise aesthetic tendencies of user groups than to get an image aesthetic classification or rating. The method used in this paper differs from other methods in that instead of simply determining the image aesthetics as high or low, the statistical distribution of human ratings is used as the prediction result, so that the prediction result has a higher correlation with human ratings. In addition, other studies mainly use AVA as the mainstream dataset, and the evaluation results mainly represent the aesthetics of Westerners [21].

3. Aesthetic Evaluation Experiment of Splash Screen Design Based on Deep Learning

3.1. Splash Screen Image Data Acquisition. A total of 1002 samples of APP splash screen image data were collected through various methods, including screenshots and Internet downloads. The participants were recruited through a WeChat group, taking into account their age, gender, educational background, and APP usage experience. A total of nine participants were recruited, including five females and four males, aged 17 to 37 years old. These participants used more than ten different APPs on a daily basis and had knowledge of aesthetics such as color and composition [22]. No compensation or fees were provided to the participants for this study. Each participant was scored independently on a 5-point Likert scale (5 for very good looking, 4 for good looking, 3 for average, 2 for bad looking, and 1 for very bad looking) [23].

The data distribution of the aesthetic evaluation of splash screen images is shown in Figure 2. From Figure 2, it can be seen that the data samples for each rank are uneven, and most of the labeled data falls between rank 2 and rank 3, with less data in the high and low ranks and an overall Gaussian distribution. Therefore, 90% of the sample data in each rank is

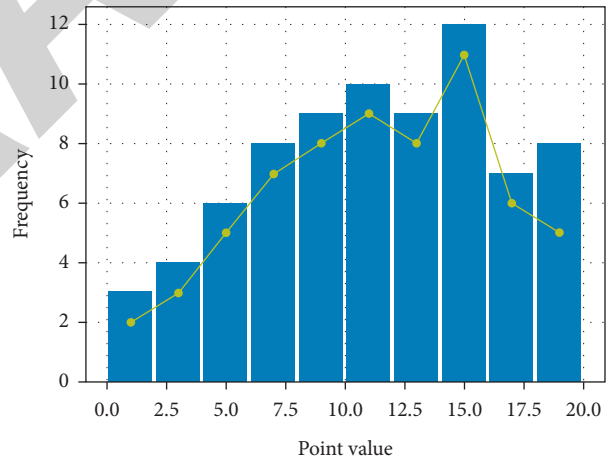


Figure 2: Data distribution of aesthetic evaluation.

randomly sampled as the training data set, and the remaining 10% are used as the test data set. Finally, 903 training data sets and 99 test data sets were obtained. The data distribution of training and test sets is shown in Figure 3.

The method is to classify the aesthetic quality of the splash screen image into good and poor grades. Those with an average score greater than or equal to 2.6 are judged to be of "good" aesthetic quality, and those with less than 2.6 are judged to be of "poor" aesthetic quality. The rules for classifying the aesthetic quality of splash screens are as follows:

$$\begin{cases} 1.0 \leq \text{socre} < 2.6, \ \text{difference}\,(482), \\ 2.6 \leq \text{socre} \leq 5.0, \ \text{good}\,(421). \end{cases} \quad (1)$$

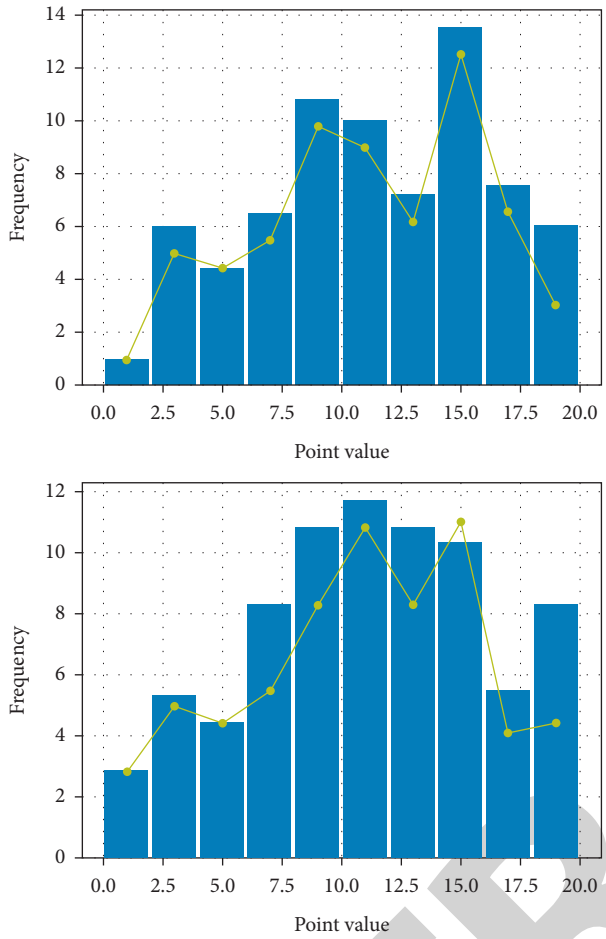


Figure 3: Data distribution of the training and test sets. Aesthetic classification of splash screens based on deep convolutional neural networks.

That is, the average score of a splash screen image rating is counted as

$$\text{score} = \frac{1}{n}\sum_{i=1}^{n} \text{person}_i. \tag{2}$$

Then, a deep neural network is designed to perform binary classification learning on the splash screen image data, and the neural network structure of the model is shown in Figure 4. A splash screen image is input, and the model obtained by pretraining the inception-ResNet-v2 CNN on the ImageNet dataset [24] is used to perform migration learning on the splash screen image to extract high-level abstract aesthetic features, and then the classifier network is passed through the final fully connected layer of the classifier network outputs two-dimensional probabilities (good/bad probabilities) for aesthetic binary classification.

The inception-ResNet-v2 CNN utilizes residual connections and convolutional operations with a large number of small kernels to make the network deeper and smaller, achieving the best current performance in the ILSVRC image classification benchmark test [25]. Therefore, the powerful feature extraction ability of the inception-ResNet-v2 convolutional neural network is utilized to achieve the aesthetic quality classification of splash screen images. The overall recognition rate of the experimental results reaches 64.7% (see Figure 5), and the overall recognition rate of the splash screen aesthetic classification is shown in Table 1.

4. Aesthetic Distribution Prediction Method for Splash Screens Based on Deep Learning NIMA

The difference between the aesthetic evaluation method NIMA proposed by Google and the above aesthetic classification method is that the above aesthetic classification method is to classify aesthetics into good and poor, and the predicted grade is to represent the average level of this image, which is displayed as the result of the predicted classification category. While the NIMA method [26] is to predict the probability distribution of a human's aesthetic evaluation of an image by CNN, the obtained probability distribution map can more accurately understand the concentration trend of a user's evaluation of an image and can more accurately guide how many people in the population find an image good looking to what degree. The distribution of splash screen aesthetic evaluation is shown in Figure 6.

From the distribution, it can be seen that 44% of people think this splash screen poster has an aesthetic rating of 3, and 11% think this splash screen poster has a poor aesthetic design and it has an aesthetic rating of 1. NIMA is designed to generate a histogram of the probability distribution of rating, i.e., the probability value of each rating, for any one image by predicting this probability distribution of human assessment of image aesthetics, which is similar to the human aesthetic rating system generates a histogram of aesthetic probability distributions that is formally compatible with the histogram of aesthetic probability distributions generated by the human aesthetic rating system. Therefore, the prediction results of NIMA [27] are closer to those of human evaluation and more representative of the aesthetics of the public.

The true distribution of human ratings for an image can be expressed as an empirical probability mass distribution function:

$$p = \left[p_{s_1}, p_{s_2}, \ldots, p_{s_i}, \ldots, p_{s_N}\right]. \tag{3}$$

where p_{s_i} represents the probability of level S_i. The goal of the NIMA method is to predict the probability distribution of the aesthetic rating of a given image.

The structure of the deep learning NIMA-based method for predicting the aesthetic distribution of splash screens is shown in Figure 7 [28]. The probability distribution of the aesthetic evaluation of the splash screen with five levels is obtained.

If we obtain the probability distribution of the aesthetic rating of the splash screen p [29], then the mean value of the aesthetic quality rating of the splash screen can be defined as

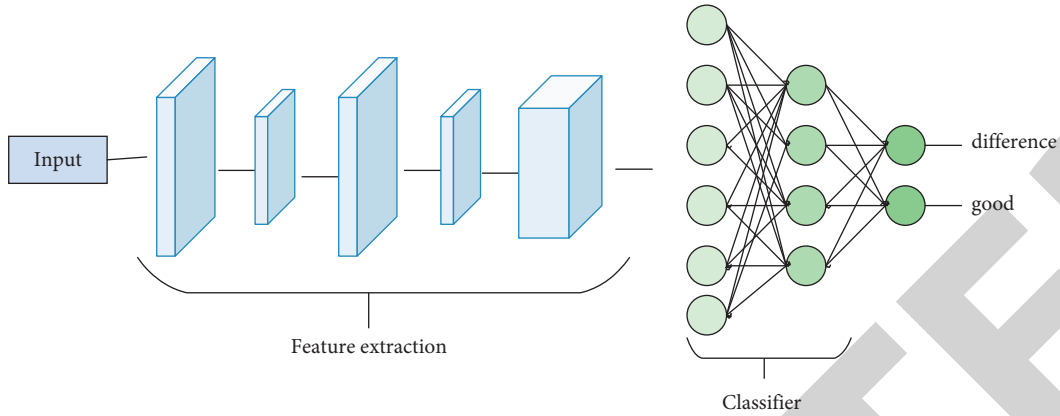


Figure 4: Aesthetic classification structure of splash screen based on CNN.

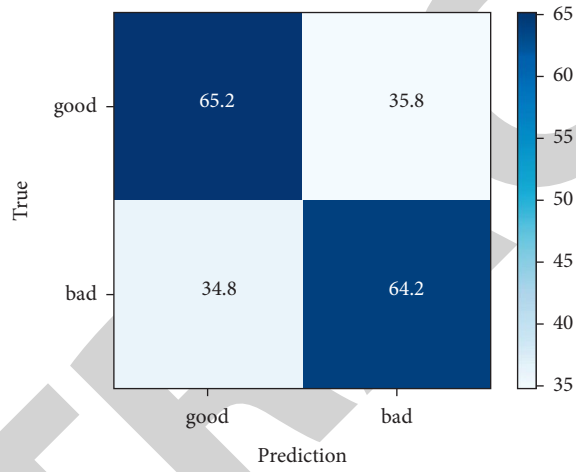


Figure 5: Confusion matrix of experimental results.

Table 1: Overall recognition rate of splash screen aesthetic classification.

Grade	Predicted as "poor" level	Predicted as "good" level
"Poor" level	30	16
"Good" level	19	34
Average overall recognition rate	64.7	

$$\mu_{\text{score}} = \sum_{i=1}^{N} s_i \times p_{s_i}. \tag{4}$$

The standard deviation of the aesthetic quality rating of the splash screen is

$$\sigma_{\text{score}} = \sqrt{\sum_{i=1}^{N} \left(s_i - \mu_{\text{score}}\right)^2 \times p_{s_i}}. \tag{5}$$

The aesthetic quality of the splash screens could then be compared qualitatively by the mean and standard variance of the splash screen aesthetic quality ratings. To compare the correlation between the predicted quality rating distribution p and the participants' labeled quality rating distribution q, the Pearson correlation coefficient was used to measure the correlation between two variables X and Y with values between −1 and 1, calculated as follows:

$$r = \frac{\sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n} (X_i - \overline{X})^2} \sqrt{\sum_{i=1}^{n} (Y_i - \overline{Y})^2}}. \tag{6}$$

The final results showed that the Pearson correlation coefficient value on the 99 test sets, trained by learning from the training data, was 0.516, with a moderate correlation agreement between its predicted and participant-labeled values. The predicted aesthetic data are shown in Figure 8. The mean value of participants' ratings for Figure 8(a) was 2.888 and the machine predicted 2.257, with a difference of 1.053; the mean value of participants' ratings for Figure 8(b) was 2.667 and the machine predicted 2.495, with a difference of 0.919. The two values differed but showed a moderate

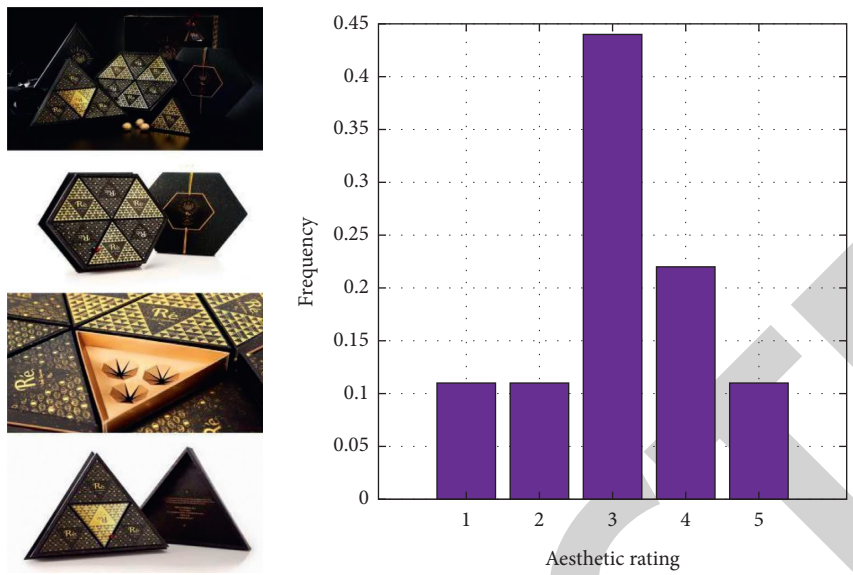


Figure 6: Distribution of splash screen aesthetics evaluation.

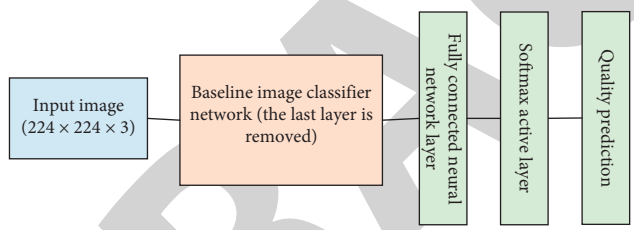


Figure 7: Structure of the flash screen aesthetic distribution prediction method is based on deep learning NIMA.

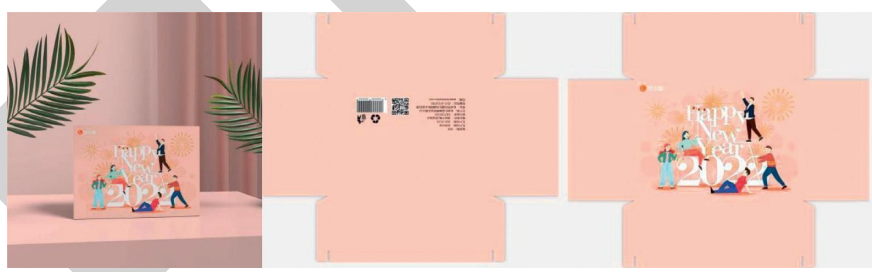


Figure 8: Predicted aesthetic data: (a) participant: 2.888 machine: 2.257 (± 1.053) and (b) participant: 2.667 machine: 2495 (± 0.919).

correlation. Due to the limited amount of data and labeled data in the "domestic mobile splash screen image aesthetics dataset," if there are enough training data, the Pearson correlation of the deep learning NIMA-based splash screen aesthetics distribution prediction method on the test set will reach a strong correlation, and the predicted results will be more representative of human aesthetic standards.

5. Experiment

Nowadays, when consumers buy products, they also pay attention to the external image of the products, which means that consumers will get different emotional experiences from them. In this way, the external image of the product is enhanced, and even the external image of the product itself becomes a consumer product. The external image of a product must be expressed through a series of graphic elements, which gives rise to certain functional and artistic visual communication techniques in the consumer field [30].

The evaluation method and results were provided to three visual designers with more than ten years of work experience, who felt that presenting the distributed evaluation results would provide a clearer view of the public's concentration on the image ratings and could be used as important supporting evidence for the evaluation of visual works within the team.

It is found that deep learning-based image aesthetic evaluation distribution can help designers and companies in

Figure 9: Two images with an aesthetic score of 3.

two dimensions [31]. First, the splash screen aesthetic distribution prediction method can help designers predict the user aesthetic evaluation distribution of their design work and establish an objective aesthetic evaluation. Based on this, future design teams can develop aesthetic parameter evaluation standards as a reference for visual evaluation and reduce the subjectivity of evaluation. Secondly, through the constructed "splash screen image aesthetics dataset," designers can more accurately understand the aesthetic characteristics perceived by users, obtain the aesthetic tendency of target users, and make forward-looking visual designs to provide users with a pleasant experience and realize precise marketing for enterprises, as shown in Figure 9.

The application of visual communication technology in packaging design is extremely important. The importance of color and pattern to package design was analyzed above, and the application strategy of these elements is now analyzed. The designer has to present the cultural connotation in the product packaging design through some artistic visual symbols or words to achieve a better information transfer effect. In general, most of the visual symbols people choose are design graphics or symbols in two-dimensional space, and these visual symbols are different from the traditional visual symbols. Based on this premise, designers have to realize that whatever visual design and whatever colors are used, they have to make it easy to understand for ordinary consumers. Color, graphics, and text are the basic elements of packaging design. Therefore, the aesthetic quality is evaluated as shown in Figure 10.

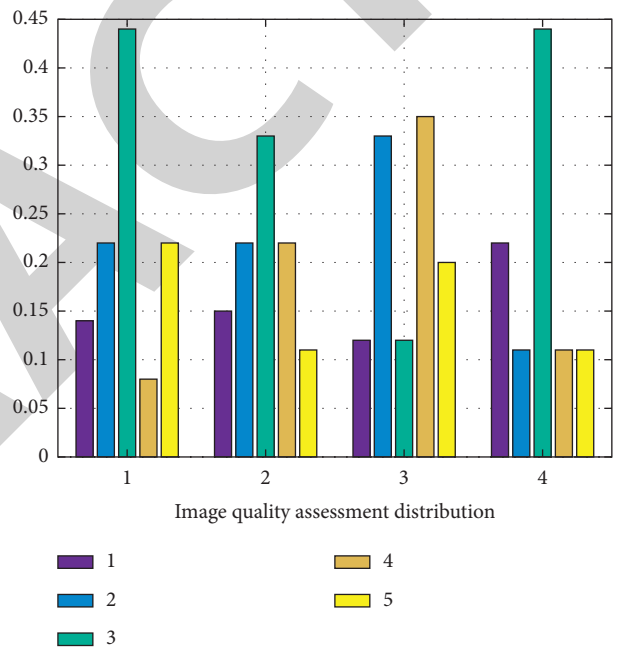


Figure 10: Distribution of aesthetic quality evaluation.

6. Conclusions

This study investigates the creative and emotional splash screen images designed by designers and uses NIMA as the main evaluation method to effectively predict the aesthetic evaluation distribution of splash screen images. The feasibility and effectiveness of applying deep convolutional neural networks to the aesthetic evaluation of interface design are verified. We could choose the visual symbols that can attract consumers' attention and enhance their desire to purchase, and then carry out a series of packaging designs so as to create a win-win situation in which quality design and healthy consumption promote each other.

Data Availability

The data used in this paper are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest regarding this work.

References

[1] T. Li, "Packaging design of Geriatric products based on human - machine evaluation system[J]," *Paper Asia*, vol. 34, no. 4, pp. 39–43, 2018.

Hindawi
Scientific Programming
Volume 2023, Article ID 9893506, 1 page
https://doi.org/10.1155/2023/9893506

Retraction

Retracted: LSTM-Based Attentional Embedding for English Machine Translation

Scientific Programming

Received 29 August 2023; Accepted 29 August 2023; Published 30 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] L. Jian, H. Xiang, and G. Le, "LSTM-Based Attentional Embedding for English Machine Translation," *Scientific Programming*, vol. 2022, Article ID 3909726, 8 pages, 2022.

Hindawi
Scientific Programming
Volume 2022, Article ID 3909726, 8 pages
https://doi.org/10.1155/2022/3909726

Research Article

LSTM-Based Attentional Embedding for English Machine Translation

Lihua Jian,[1] **Huiqun Xiang**,[2,3] **and Guobin Le**[2,3]

[1]*School of International Education, Hunan University of Medicine, Huaihua 418000, Hunan, China*
[2]*Changsha Vocational and Technical College, Changsha 410200, Hunan, China*
[3]*School of Foreign Languages, Huaihua University, Huaihua 418000, China*

Correspondence should be addressed to Lihua Jian; 202110080102@hunnu.edu.cn

Received 23 December 2021; Revised 9 January 2022; Accepted 17 January 2022; Published 16 March 2022

Academic Editor: Baiyuan Ding

In order to reduce the workload of manual grading and improve the efficiency of grading, a computerized intelligent grading system for English translation based on natural language processing is designed. An attention-embedded LSTM English machine translation model is proposed. Firstly, according to the characteristics of the standard LSTM network model that uses fixed dimensional vectors to represent words in the encoding stage, an English machine translation model based on LSTM attention embedding is established; the structure level of the English translation scoring system is constructed. A linguistic model of the English translation scoring system is established, and the probability distribution of a particular sentence sequence or word sequence of the translated text is statistically calculated using the model. The results show that the English machine translation model based on LSTM attention embedding proposed in this study can enhance the representation of the source language contextual information and improve the performance of the English machine translation model and the quality of the translation compared with the English machine translation models constructed by existing neural network structures, such as standard LSTM models, RNN models, and GRU-Attention translation models.

1. Introduction

With the development of computer technology and the maturity of artificial intelligence technology, machine translation is gradually replacing human translation and occupying a larger proportion in the translation field. At present, there are four main types of machine translation [1–3]. Among them, neural network-based machine translation models can alleviate the problem of feature design of high-dimensional data and improve the expressiveness of the model by building neural network classifiers when dealing with high-dimensional complex data, which has become the most popular and effective language translation model nowadays [4, 5].

In the literature [6], an English translation scoring system based on hidden Markov model is used, combining Markov model and Viterbi comparison system to input similar words between the translation and the reference translation, match the similar words to calculate the proximity between them, and then compare the similarity between the translated utterances, and according to the comparison results, achieve the translation scoring [7]. The accuracy of the scoring results of this system is high, but the computation is large and time-consuming. Corpus-based English translation scoring system designed in [8] obtains word alignment ratios by analyzing word collocations in the structure of corpus materials, compares the word collocations and structure of the input translations, and scores the translations. The scoring results of this system have large errors and the process of word collocation analysis is complicated. Translation models such as those based on LSTM, RNN, and GRU-Attention neural networks have been widely used in the field of English machine translation [8–11] using neural networks with different structures to study the translation effect of English machine translation in the field of component products and other areas and to

achieve intelligent English machine translation. The results of English machine translation in areas such as component products were studied using different structures of neural networks, and intelligent English machine translation was achieved. However, the abovementioned English machine translation models based on neural network structures all suffer from the problem of unsatisfactory translation results due to the loss of long-distance information in the process of transmission due to long-distance dependence and therefore need to be improved [12].

To address the problems in existing marking systems, an intelligent computerized marking system for English translation based on natural language processing is designed. Through simulation experiments, the system is compared with the current scoring system and manual scoring method, and it is verified that the designed scoring system has high operational stability and accuracy, and the overall performance is better than the current scoring system.

2. Design of a Computerized Intelligent Scoring System for English Translation

2.1. Hierarchical Construction of English Translation Scoring System. The hierarchical relationship of each module is shown in Figure 1.

At the initial stage of the system, students' English translations are entered through the translation data collection module and processed by the collection module to produce a standardised format of the database file [13].

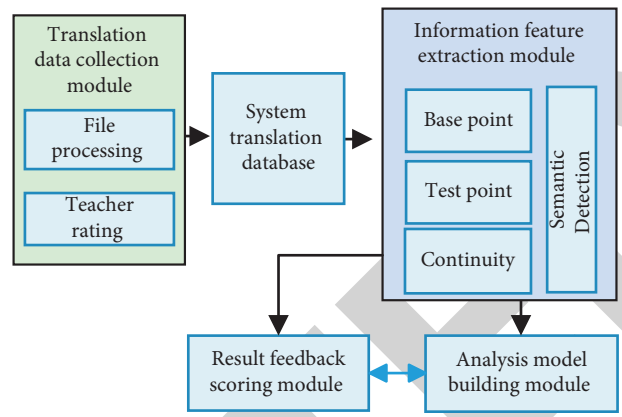


Figure 1: Rating system hierarchy.

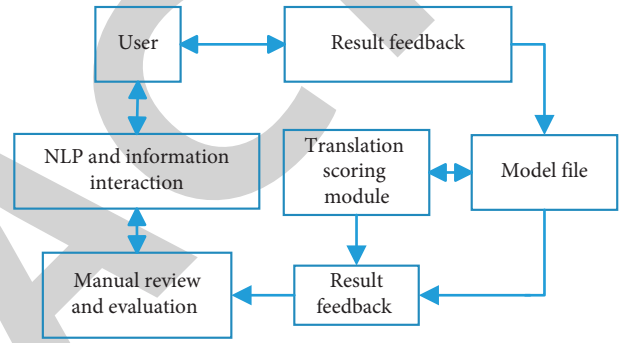


Figure 2: Overall architecture of the translation scoring system.

2.2. English Translations Scoring System. The overall framework of the natural language processing-based English translation system is shown in Figure 2. The user uploads a translation through the user side and, after the computer's natural language intelligence processing and information interaction, inputs it into the system's English translation scoring model.

2.3. Models in This Paper. LSTM is a special recurrent neural network model that solves the long sequence dependence problem in recurrent neural networks by adding memory units, input gates, output gates, and forgetting gates and improves the ability of recurrent neural networks to process long sequence data [14].

The transformer model also consists of an encoder group and a decoder group. An encoder or decoder group consists of multiple encoder modules or decoder modules stacked on top of each other. Each module consists of a multi-head attention and a fully connected feed-forward layer. Since the RNN is abandoned, another method is needed to remember the location information of the input sequence. A positional embedding is used in the transformer model to add a relative position to each element of the input sequence, and this position information is then used as a representation of each word [15, 16].

According to the above analysis of the LSTM network model, the output vector in the coding stage of the LSTM network model has a fixed dimension, so it uses the same dimensional vector for any length of the source language sequence to encode. In actual English machine translation, the input English sequences are of variable length, which makes it easy to use the standard LSTM model for English machine translation, and the model does not fit the English input sequences perfectly, thus making the translation effect unsatisfactory. Moreover, due to the different focus of translation, the use of a fixed dimensional representation of the input model sequence, i.e., the same level of attention to the sequence, is obviously not conducive to improving the quality of the translation. Therefore, in order to solve the above problems, an attention mechanism is embedded in the LSTM network [17], and an English machine translation model based on LSTM attention embedding is proposed.

First, a set of multiple vectors is used instead of a fixed dimension for representing the source language sequence. Then, by dynamically selecting the background vectors during the target sequence generation process, the translation model is improved to pay more attention to the parts with high relevance to the source language during the translation process, which in turn improves the translation performance of the model [18]. The LSTM English machine translation model embedded with attention mechanism consists of three parts: encoder, decoder, and attention mechanism, as shown in Figure 3.

The next hidden state at the target side of the model is calculated in the same way as the LSTM decoder part, as in the following equation:

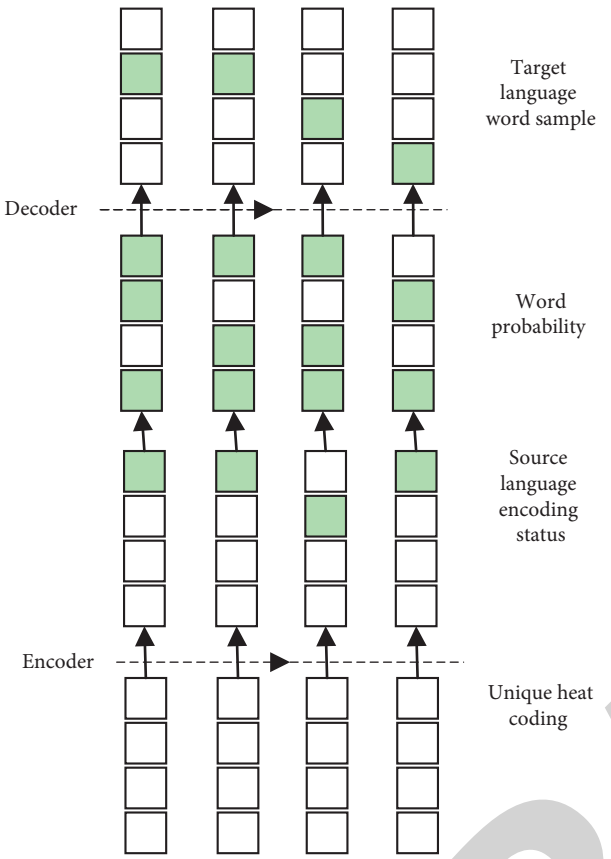


Figure 3: LSTM translation model with embedded attention mechanism.

$$Z_i + 1 = \sigma(c_i, u_i, z_i), \tag{1}$$

where u_i denotes the i-th word in the target language sequence and c_i denotes the background vector of word i. Since the background vectors of the LSTM model with the embedded attention mechanism are a set of multiple vectors, rather than being uniformly fixed [19], each word in the target language sequence can find a unique background vector corresponding to it.

Let the state of the implicit layer at encoder j be h_j, then its corresponding background vector can be calculated by

$$C_i = \sum_{j=1}^{r} a_{ij} h_j, \tag{2}$$

where a_{ij} represents the weight, i.e., the attention value of the i-th word in the target language sequence to the j-th word in the source language sequence, which can be calculated by the following equations:

$$a_{ij} = \frac{\exp(e_{ij})}{\sum_{k=1}^{T} \exp(e_{ik})}, \tag{3}$$

$$e_{ij} = a(z_i, h_j), \tag{4}$$

where a is a function that measures the match between the current hidden state z_i of the target language sequence and the hidden state h_j of the source language sequence and can be calculated by

$$e_{ij} = v^T \tanh(W_{zzi} + W_h h_j), \tag{5}$$

where v, W_z, and W_h denote the model parameters to be learned.

By embedding an attention mechanism in the LSTM network, the model can be weighted with different weights on the source language side, which solves the long-range dependency problem of standard LSTM models and thus improves the model performance.

3. Implementation of an English Translation Scoring System

3.1. Language Models for English Translation Scoring Systems. Statistical language models can give the probability distribution of a particular sentence sequence or word sequence in a translation [20–22]. To simplify the computation and reduce the complexity, a ternary model is introduced. Let the preferred set embedded in the ternary language model be V and the ternary combination be (u, v, w), corresponding to a parameter $q(w \mid u, v)$ with full $w \in v \cup \{\text{STOP}\}$ and u, $v \in v \cup \{*\}$. $q(w \mid u, v)$ represents the probability that a single word w follows a word u and v when the binary combination is known. The probability distribution of the ternary language model for a given translated sentence $x_1 x_2 \ldots x_n$ is given by

$$p(x_1 x_2 \ldots x_n) = \prod_{i=1} q(x_i \mid x_{i-2} x_{i-1}). \tag{6}$$

The restrictions that need to be met are

$$q(w \mid u, v) \geq 0 \quad \sum_{w \in v \cup |\text{STOP}|} q(w \mid u, v) = 1. \tag{7}$$

The maximum likelihood estimation algorithm is used to solve for $q(w \mid u, v)$, which corresponds to the following equation:

$$q(w \mid u, v) = \frac{c(u, v, w)}{c(u, v)}, \tag{8}$$

where $c(u, v, w)$ represents the frequency of occurrence of (u, v, w) in the translation training set and $c(u, v)$ is the frequency of occurrence of (u, v) in the translation training set [23].

To address the problem that not all ternary combinations that do not appear in the translation training set have a probability of zero, a smoothing algorithm is introduced to obtain the descriptive formula for the language model as

$$q(w \mid u, v) = \lambda_1 * q(w \mid u, v) + \lambda_2 * q(w \mid v) + \lambda_3 * q(w), \tag{9}$$

where λ_1, λ_2, and λ_3 represent the smoothing factors and satisfy $\lambda_1, \lambda_2, \lambda_3 \geq 0$, $\lambda_1 + \lambda_2 + \lambda_3 = 1$; $q(w \mid v)$ represents the probability of word w occurring after the word v when word v is known; and $q(w)$ represents the total probability of word w occurring.

3.2. Similarity Calculation and Scoring of English Translations. In order to calculate the similarity between the user's translation result and the standard answer, the similarity of keywords is introduced and the word similarity is calculated by the following formula [24]:

$$\text{sim Word}(A, B) = \frac{\text{Same}(A, B)}{\text{Num}(A) + \text{Num}(B)}, \tag{10}$$

where sim Word (A, B) is the word similarity between sentences A and B, Same (A, B) represents the number of identical words in sentences A and B, and Num (A) and Num (B) represent the number of words in sentences A and B, respectively.

The characteristic keyword similarity is calculated, the particle swarm optimized BP network is used to fit the calculation, and the calculation result is compared with the set scoring standard.

4. Experimental Results and Analysis

4.1. Experimental Environment and Parameter Settings. In order to verify the effectiveness of the proposed LSTM attentional embedding-based English machine translation model, the study built an LSTM English machine translation system on the TensorFlow framework [25]. The parameters of the LSTM neural network are set as follows: the vocabulary size is 30 000, the word vector dimension and the number of nodes in the implicit layer are 512, the number of LSTM network layers is 2, the column search width is 3, the learning rate is 0.1, the dropout is 0.5, and the batch size is 128. The decoding stage is based on the column search algorithm.

4.2. Dataset Sources and Preprocessing. The study chose the International Spoken Language and its Translation Review Contest (IWSLT) 2019 data, which has a small data size, as the dataset for this experiment, including 220,000 Chinese-English parallel utterance pairs, pairs of test set data, and pairs of development data [26]. Since the LSTM attentional embedding-based English machine translation model could not be trained and learned directly on the IWSLT 2019 dataset, word vector transformation of the dataset was also required [27]. The study performed a word separation process on the data and then used CBOW to factorize the separated data.

4.2.1. Split Word Processing. As (IWSLT) 2019 dataset contains Chinese and English parallel utterance pairs, the Chinese and English word separation methods are different; therefore, the study carried out word separation for the Chinese and English of the experimental dataset separately [28]. For Chinese word separation, a statistical-based word separation method was used. Firstly, a word is regarded as a combination of several fixed words according to the composition form of Chinese words; then, the probability of word generation is judged according to the frequency of co-occurrence between words in the context of an utterance, i.e., the credibility of the word; finally, a threshold is set according to the credibility of the word to form the word composition condition and determine the word separation. In the case of English, since the basic unit of English is the word, it is only necessary to split the word directly according to the space. However, since English sentences contain stop words, they also need to be deactivated during the word separation process [29]. The English deactivation process consists of three main steps: firstly, capitalisation of the English language, then space splitting of the words and symbols at the end of the sentence, and finally, generalisation of the word sentence using the special noun special bond method.

4.2.2. Word Vector Representation. Victorian representation of words means digitising linguistic symbols so that language numbers can be fed into a model for training and learning [30]. The study uses CBOW for the factorized representation of words. Suppose the size of the dictionary is v, and an index set 111 of one-to-one correspondence between the word and the integers in the dictionary is established. If there exists a test sequence with length T, time window size m, and word J at time t, the probability of CBOW maximizing the background work to generate a central word is given by

$$J = \prod_{t=1}^{t} P\left(w^{t-m}, \ldots, w^{t-1}, w^{t+1}, \ldots, w^{t+m}\right). \tag{11}$$

Taking the negative logarithm of the above equation gives the loss function, i.e., the maximum likelihood estimate of equation (11) can be calculated by minimizing the following equation:

$$J = \sum_{t=1}^{r} \log P\left(w^{t} \mid w^{t-m}, \ldots, w^{t+1}, \ldots, w^{t+m}\right). \tag{12}$$

Assuming that the background word vector is denoted as v and the central word vector is denoted as u, then by CBOW training, for each word indexed as i in the lexicon, the vector of that word as a background word is obtained (v_i) and the vector as a central word can be denoted u_i.

4.3. Evaluation Indicators. BLEU value is selected as the index to evaluate the translation quality of the translation model. The larger its value is, the higher the translation quality is. The calculation method of the BLEU value is shown as follows:

$$\text{BLEU} = BP \times \exp\left(\sum_{n=1}^{N} w_n \log_{10} p_n\right), \tag{13}$$

where BP is the penalty factor; N is the longest tuple length, usually 4; n is the number of tuples; w_n is the tuple n weight; and p_n is the tuple n ratio [12].

4.4. Model Validation. In order to verify the performance of the proposed LSTM attention-embedded translation model, the experimental dataset was trained with the standard LSTM model and the attention-embedded LSTM model,

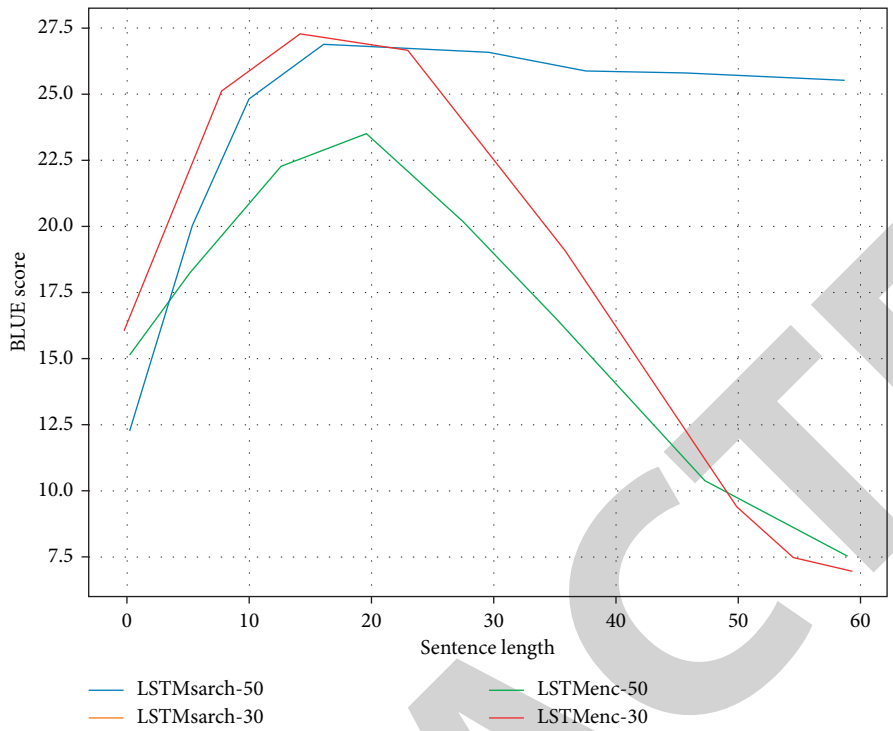


Figure 4: Performance of the attention mechanism (RNN to LSTM).

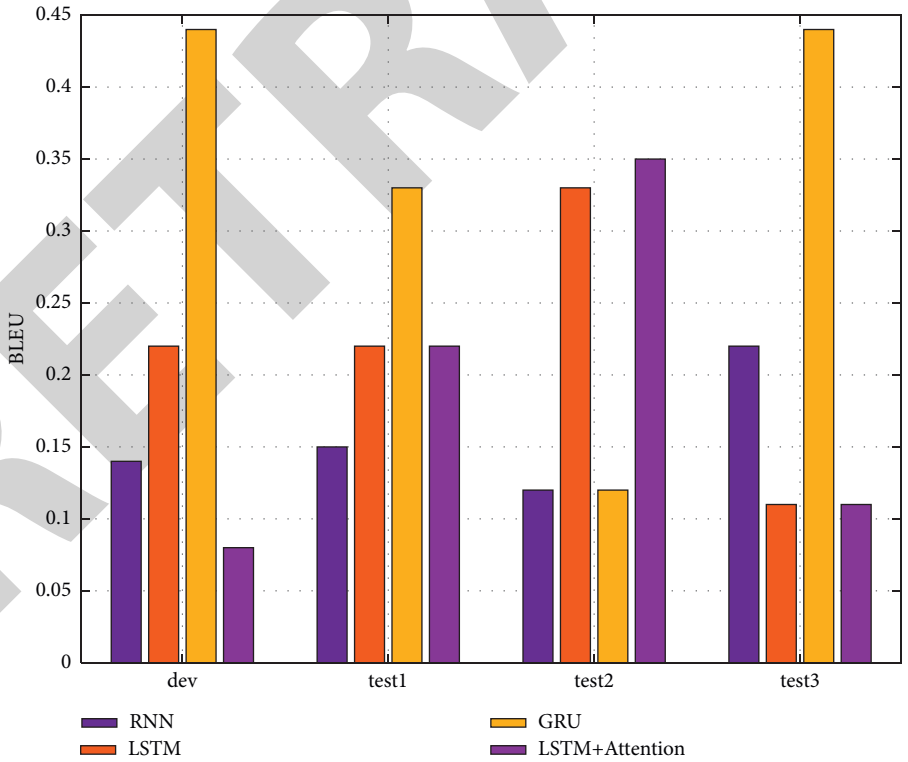


Figure 5: BLEU values of different translation models on the experimental dataset.

and the results are shown in Figure 4. The numbers in Figure 4 represent the different network layers in the network model we designed. Compared with the standard LSTM model, the attention-embedded LSTM model has a higher BLEU value, indicating that the attention-embedded

Table 1: Comparison of scoring values by scoring method.

DE	RM	SC/C
1	RA	87.4
	RB	84.7
	RC	87.6
2	RA	74.6
	RB	76.9
	RC	74.7
3	RA	72.6
	RB	69.8
	RC	72.9

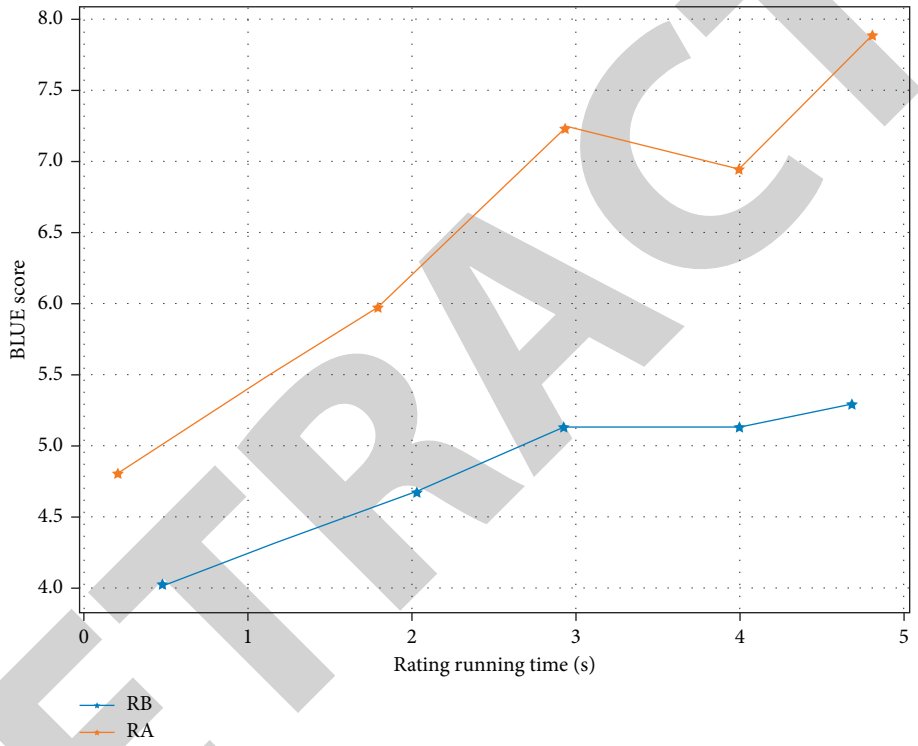


Figure 6: Comparison of scoring system run times.

LSTM model is more effective in translating long sentences and the model performance is effectively improved.

4.5. Model Comparison. In order to verify the translation effectiveness of the proposed English machine translation model of LSTM attention, the experimental analysis results are shown in Figure 5. As can be seen from the figure, the BLEU values of the proposed English machine translation model based on LSTM attention embedding on both the development machine and the test set are higher than those of the traditional LSTM, RNN, and GRU-Attention English machine translation models, indicating that, compared to the comparison translation models, the proposed translation model improves translation by our network. The proposed translation model improves the performance and translation quality by our network.

4.6. Scoring Effect. In Table 1, DE denotes the English translation document to be scored; RM denotes the scoring method; RA, RB, and RC denote the designed system, the existing scoring system, and the manual scoring method, respectively; SC denotes the score in points and is denoted by the letter C.

According to the data in Table 1, the scoring results of the designed system are closer to the manual scoring results, with a minimum difference of 0.1 C and a maximum difference of 0.3 C. This indicates that the scoring error of the designed English translation scoring system is smaller and the scoring performance is better. Experiments were conducted using the designed system and the existing scoring system to compare the running time of the scoring process, and the experimental results are shown in Figure 6. In Figure 6, RA and RB denote the runtime of the designed system and the existing scoring system, respectively.

According to Figure 6, the range of fluctuation of the scoring runtime curve of the designed system is smaller than that of the runtime curve of the existing scoring system, which indicates that the designed system is more stable in operation. For the translation sample, the scoring time of the designed system was 4.7 s, while that of the existing scoring system was 6.1 s. For the translation sample, the scoring times of the designed system and the existing scoring system were 4.9 s and 5.9 s, respectively, which indicates that the scoring time of the designed system was significantly lower than that of the existing scoring system for the same translation sample, indicating that the scoring efficiency of the designed system was higher.

5. Conclusions

The proposed English machine translation model based on LSTM attention embedding is innovative in that it enhances the representation of source language contextual information by introducing an attention mechanism into the standard LSTM English translation model, thereby improving the performance of the English machine translation model and the quality of the translated text. The result is better than the standard LSTM model and the traditional RNN and GRU-Attention English machine translation models and can be used in real English machine translation. The experimental results show that the overall performance of the designed system is better than that of the traditional system, indicating its strong practicality.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] C. Su, H. Huang, S. Shi, P. Jian, and X. Shi, "Neural machine translation with Gumbel tree-LSTM based encoder," *Journal of Visual Communication and Image Representation*, vol. 71, Article ID 102811, 2020.

[2] Y. Ma, J. Yu, B. Ji, J. Chen, S. Zhao, and J. Chen, "Three-way decisions based RNN models for sentiment classification," in *Proceedings of the International Joint Conference on Rough Sets*, pp. 247–258, Springer, Bratislava, Slovak Republic, September 2021.

[3] N. A. Mohamed, M. A. Zulkifley, A. A. Ibrahim, and M. Aouache, "Optimal training configurations of a CNN-LSTM-based tracker for a fall frame detection system," *Sensors*, vol. 21, no. 19, p. 6485, 2021.

[4] J. Wang, J. Zhu, and Z. Yu, "Person re-identification based on attention clustering and long short-term memory network," *Journal of Electronic Imaging*, vol. 30, no. 3, Article ID 033014, 2021.

[5] C. H. Cao, Y. N. Tang, D. Y. Huang, G. WeiMin, and Z. Chunjiong, "IIBE: an improved identity-based encryption algorithm for wsn security," *Security and Communication Networks*, vol. 2021, Article ID 8527068, 8 pages, 2021.

[6] A. K. Mandal, R. Sen, S. Goswami, and B. Chakraborty, "Comparative study of univariate and multivariate long short-term memory for very short-term forecasting of global horizontal irradiance," *Symmetry*, vol. 13, no. 8, p. 1544, 2021.

[7] C. Shi, S. Liu, S. Ren et al., "Knowledge-based semantic embedding for machine translation," in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, vol. 1, pp. 2245–2254, Berlin, Germany, August 2016.

[8] M. A. Hasan, F. Alam, S. A. Chowdhury, and N. Khan, "Neural machine translation for the Bangla-English language pair," in *Proceedings of the 2019 22nd International Conference on Computer and Information Technology (ICCIT)*, pp. 1–6, IEEE, Dhaka, Bangladesh, December 2019.

[9] V. Goyal and D. M. Sharma, "LTRC-MT simple & effective Hindi-English neural machine translation systems at WAT 2019," in *Proceedings of the 6th Workshop on Asian Translation*, pp. 137–140, Hong Kong, China, November 2019.

[10] Y. Xia, T. He, X. Tan, F. Tian, D. He, and T. Qin, "Tied transformers: neural machine translation with shared encoder and decoder," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 1, pp. 5466–5473, 2019.

[11] L. Wang, C. Zhang, Q. Chen et al., "A communication strategy of proactive nodes based on loop theorem in wireless sensor networks," in *Proceedings of the 2018 Ninth International Conference on Intelligent Control and Information Processing (ICICIP)*, pp. 160–167, IEEE, Wanzhou, China, November 2018.

[12] M. Sato, J. Suzuki, and S. Kiyono, "Effective adversarial regularization for neural machine translation," in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pp. 204–210, Florence, Italy, July 2019.

[13] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, "LSTM: a search space odyssey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 10, pp. 2222–2232, 2016.

[14] P. An, Z. Wang, and C. Zhang, "Ensemble unsupervised autoencoders and Gaussian mixture model for cyberattack detection," *Information Processing & Management*, vol. 59, no. 2, Article ID 102844, 2022.

[15] K. Chen, R. Wang, M. Utiyama, E. Sumita, and T. Zhao, "Syntax-directed attention for neural machine translation," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 32, no. 1, New Orleans, LA, USA, April 2018.

[16] C. Ma, A. Tamura, M. Utiyama, T. Zhao, and E. Sumita, "Forest-based neural machine translation," in *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, vol. 1, pp. 1253–1263, Melbourne, Australia, July 2018.

[17] Q. Xia, Z. Li, M. Zhang et al., "Syntax-Aware neural semantic role l," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 1, pp. 7305–7313, 2019.

[18] J. Yang, R. Yang, H. Lu, C. Wang, and J. Xie, "Multi-entity aspect-based sentiment analysis with context, entity, aspect memory and dependency information," *ACM Transactions on Asian and Low-Resource Language Information Processing*, vol. 18, no. 4, pp. 1–22, 2019.

[19] M. Nguyen, G. H. Ngo, and N. F. Chen, "Hierarchical character embeddings: learning phonological and semantic representations in languages of logographic origin using recursive neural networks," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 461–473, 2019.

[20] Y. Yin, J. Su, H. Wen, J. Zeng, Y. Liu, and Y. Chen, "POS tag-enhanced coarse-to-fine attention for Neural Machine

Hindawi
Scientific Programming
Volume 2022, Article ID 4862806, 10 pages
https://doi.org/10.1155/2022/4862806

Research Article

Methods on Visual Positioning Based on Basketball Shooting Direction Standardisation

Yiqun Huang and XiaoYun Liu

QuanZhou Medica College, QuanZhou, Fujian 362000, China

Correspondence should be addressed to XiaoYun Liu; 1988005@qzmc.edu.cn

Received 16 January 2022; Accepted 14 February 2022; Published 15 March 2022

Academic Editor: Baiyuan Ding

The existing basketball training shooting direction correction method has the problems of low correction accuracy and poor self-adaptability, and proposes a basketball training shooting direction correction method based on visual perception. A visual localisation algorithm for tracking feature points of object targets is used as the basis for the process of visual robot localisation and its effects, from camera calibration, template matching, background modelling and foreground target separation to feature point extraction, motion estimation, and Kalman filtering. An in-depth understanding and analysis of the traditional corner point detection algorithm is presented, on the basis of which improvements are proposed. An accurate tracking method based on improved Harris corner point extraction is introduced, which builds on the traditional Harris feature point detection by using the changing relationship between the gradient of the grey value of the pixels near the corner point, using simple operations and analysis to exclude some pseudocorner points and noncorner points, and further processing the retained points to derive the correct feature points. The code of this algorithm is written to finally achieve its detection effect, and compared with the traditional algorithm, it is concluded that this algorithm can then extract more accurate corner points in a shorter time, which lays the foundation for the next step of accurate basketball tracking, reflecting the practicality of this algorithm.

1. Introduction

In the context of increasingly sophisticated computer vision and image processing technologies, machine vision is used to recognise images, analyse the video parameters of captured sports images, and feed back into human-computer interaction systems and expert systems to achieve guidance in sports training. According to this idea, the basketball training shooting angle correction method is studied, the shooting angle information feature quantity of basketball training is extracted, the basketball training shooting angle parameters and training movement characteristics are analysed, and the correction improvement of basketball training movements is guided. For the basketball training shooting angle visual features modelling difficulties and other problems, combined with machine vision analysis, to achieve basketball training shooting angle correction, related basketball training shooting angle correction method research in the field of sports and computer vision parameter analysis have been better applied [1–4].

Visual positioning technology is divided into monocular and multiocular in terms of the form of image acquisition. Single vision refers to the use of a single camera for the acquisition of environmental information, while multivision refers to the use of multiple cameras, and what we currently call multivision generally stands for dual vision [5–8]. Monocular vision positioning uses a single camera to collect information about the target and calculate its motion parameters, offering the advantages of simple computing, low cost, and ease of installation. However, the limitations of the monocular acquisition window result in its inability to effectively accomplish localisation in complex environments, and monocular systems do not have access to height information about the target environment, so they cannot perform the task of reconstructing three-dimensional space [9, 10]. Binocular stereo vision can simultaneously obtain

two images about the target and compute and process them, and through triangulation can obtain three-dimensional information about the surrounding environment to reconstruct three-dimensional space [11–14]. The Zhurong exploration rover is China's Mars Exploration Robot for 2020, which uses vision processing and positioning technology to move around the surface of Mars and collect information about its surroundings. Visual localisation algorithms have been proposed since the 1980s. This method uses the feature points of the target being located to achieve object tracking, so the problem of extracting target feature points has become the key to visual localisation technology. Vision robots have been increasingly used in aerospace, industrial and agricultural production, military, and other fields, mainly due to the rapid development of image processing technology in recent years as well as the decreasing hardware cost of computers and the rapid increase in their computing speed. Vision robotics incorporates core elements from many fields and is the perfect combination of today's scientific and technological developments and advanced industrial operations [10, 11, 15, 16]. The ability to autonomously recognise its surroundings is now a mainstream direction in robotics research, as it has greater plasticity than traditional robots and can successfully perform the tasks we require in a variety of environments and situations. How to identify the target to be operated in the vision window is a prerequisite for a vision robot to work properly, so vision-based localisation and tracking technology is the core and key to robot systems. The study of its algorithms has therefore also become a problem for many researchers [17]. There are two general positioning methods: one is Global Positioning System (GPS) and the other is Voyage Position Projection. The first method is limited by the fact that positioning is not possible at all when GPS is not available, while the second method can cause the robot to be inaccurately positioned after a long period of operation due to environmental influences. Vision-based robot positioning systems are well placed to avoid the positioning problems caused by both of these methods and have therefore become a common international research priority. The advantages of low energy consumption, a wide field of application, and the small size of the equipment are also reasons why visual positioning systems have replaced the usual methods [18–20].

Therefore, this article will focus on the standardisation of basketball shooting direction based on visual localisation, analyse the traditional implementation equipment and algorithms, and design a visual localisation system with more real-time effect and accurate effect.

2. Image Processing for Visual Positioning

2.1. Overall System Structure. The system includes a camera, decoder, field-programmable gate array (FPGA), synchronous dynamic random memory (SDRAM), digital signal processor (DSP), and display, as shown in the schematic diagram in Figure 1. In order to reduce the workload of the FPGA, the system is designed to use a division of labour between the FPGA and the DSP, and the use of multiple core chips in conjunction with each other makes the system more scalable and more powerful. This system not only meets the requirements of real-time image processing but also allocates the various functional modules of the FPGA effectively. A single core chip allows the FPGA to complete the function of carrying on and processing image data in the whole system [21–23].

2.2. Image Acquisition. This system focuses on the automatic positioning of vision robots in a complex and changing application environment, which requires improved accuracy and stability of data acquisition. Charge-coupled device (CCD) cameras are superior to complementary metal-oxide semiconductor (CMOS) cameras in terms of sensitivity, resolution, and noise immunity, so CCD cameras are used. The visual positioning of the basketball does not require high colour resolution of the image, and in terms of sensitivity to infrared light, black and white cameras are superior to colour cameras and have a certain night vision effect, so based on the above advantages, choose black and white CCD cameras.

The design of the system focuses on the automatic positioning of basketballs, which involves many external environments and is difficult to achieve good results without further processing of the general image acquisition. The image information captured is optimised by selecting the most suitable processing algorithm according to the interference of the external environment. In order to adapt the system to different working environments, two distinct image processing algorithms, smoothing and sharpening, have been chosen for this design. The image sharpening algorithm takes into account the fact that the positioning system does not need to focus too much on the details of the acquisition screen, but more on obtaining the contour information of the scene in order to extract the correct corner points; thus, the Sobel sharpening operator is chosen as the edge detection algorithm. The image smoothing process is based on the general median filtering algorithm, which is improved according to the FPGA processing characteristics, in order to better adapt to the FPGA design requirements and reflect the real-time.

The focus of image processing is to achieve a better representation of the real scene and to enhance the efficiency of image transmission and storage. Image processing consists of image enhancement and restoration, image transformation, image coding and compression, and image segmentation. Image enhancement and restoration is used to restore or enhance image information to the maximum. By enhancing the high-frequency and low-frequency parts of the image to improve the clarity of the contours of the objects in the image and to remove noise, the image is able to reflect its true structure more completely. Image transforms are a fundamental tool in the study of complex algorithms and have become an extremely important part of the research process. The use of mathematical mapping methods such as the Fourier transform, wavelet transform, and discrete cosine transform to solve the problem of large amounts of information in image processing facilitates the extraction and understanding of further image information.

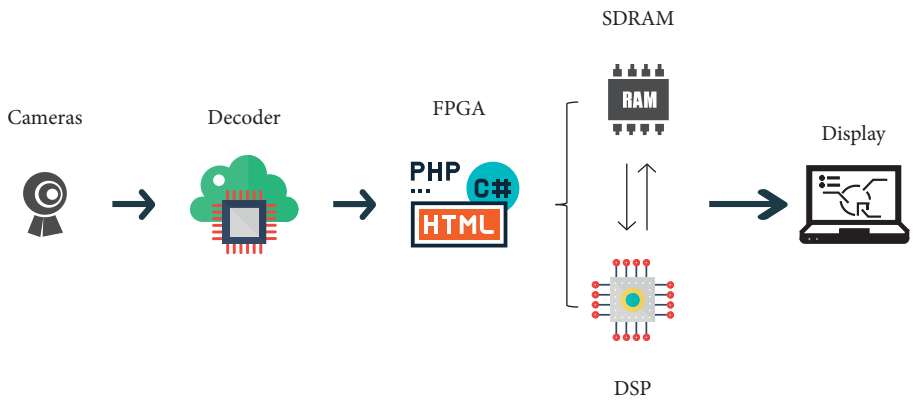


Figure 1: Overall system structure.

Image encoding and compression is the conversion of an image into a certain format and compression to increase the transmission speed and reduce the storage capacity of the image, while ensuring the quality of the image. Image segmentation extracts important parts of the image information such as edges, regions, and other characteristic parts.

2.3. Improved Median Filtering. The linear filtering method is low-pass, which leads to the removal of noise but also to the loss of image information. Median filtering, on the other hand, is a nonlinear filtering algorithm, so this algorithm can remove noise while preserving the original effect of the image. The algorithm for median filtering is to select a scan window X with a noneven number of pixel points, read the individual pixels of the image on the subwindow, arrange the scanned pixel points in this window by the size of the grey value, and finally replace the grey value in the centre of the window with the middle-most grey value in the arrangement. Any odd set of data is arranged in the descending order.

$$x_{i1} \le x_{i2} \le x_{i3} \le \cdots \le x_{\text{in}},$$

$$y = \text{Med}\left(x_1, x_2, x_3, \cdots, x_n\right) = \begin{cases} x_{i((n+1)/2)}, & n \text{ is odd}, \\ \frac{1}{2}\left[x_{i(n/2)} + x_{i((n+1)/2)}\right], & n \text{ is even}, \end{cases} \quad (1)$$

Generally, windows acquire image pixel values in a left-to-right, top-to-bottom sequence. Processing a 256×256 image with a $3 * 3$ window using traditional median filtering would result in 65536 pixel values being calculated, and each pixel value would need to be compared 36 times. Such a large amount of computation is obviously unsatisfiable in some areas with high real-time requirements [24, 25]. The use of FPGAs for image processing makes use of their parallel data processing characteristics to better perform median filtering. Instead of finding the median by ranking the magnitude of a pixel and its surrounding grey values in order of magnitude, the improved median filtering algorithm takes full advantage of the essence of the sorting algorithm to find the median and exclude other nonmedian values as quickly as possible in order to achieve greater efficiency.

The median filtering algorithm requires only 7 calls to the 3-value comparator to perform 21 comparisons, which significantly reduces the amount of computation, and the parallel processing of this algorithm using an FPGA will result in a significant increase in processing speed. The specific process of the algorithm is shown in Figure 2.

The improved median filter hardware circuit (Figure 3) is divided into four main parts: the $3 * 3$ filter window generation module, the row counter module, the median filter module, and the implementation module. The execution module is divided into a reset signal (rst) and a clock signal (clk), which together control the input of the image information. din (7 : 0) is the 8-bit image data input to the scan window, DOUT (7 : 0) is the image data after processing by median filtering, and DV is the output signal valid flag. (Figure 4)

Taking the $3 * 3$ filter window as an example, in order to ensure that the 9 captured pixels are output at the same time, 2 FIFO memories are used, with each FIFO memory storing the pixel data of the first captured row, waiting for the last row of data to be captured, and then outputting it at the same time as the data of the first two rows to form the $3 * 3$ template.

Once the filter hardware design is complete, the effect of the improved median filtering is verified using FPGA-related software. The improved median filter code is written in verilog and simulated jointly on MATLAB and ModelSim to compare the results and draw conclusions. MATLAB first reads the image data, processes it appropriately, and saves it as a data file, and then, ModelSim reads the data file, performs median filtering, and writes the results to a new data file.

2.4. Edge Detection Algorithms. Image processing is generally carried out through image scan windows, so in the edge extraction algorithm in this section, detection extraction can be achieved using the $3 * 3$ scan window in the median filtering described above. Through the analysis of the algorithm of edge detection, it can be concluded that the process is divided into three parts, namely, the calculation of

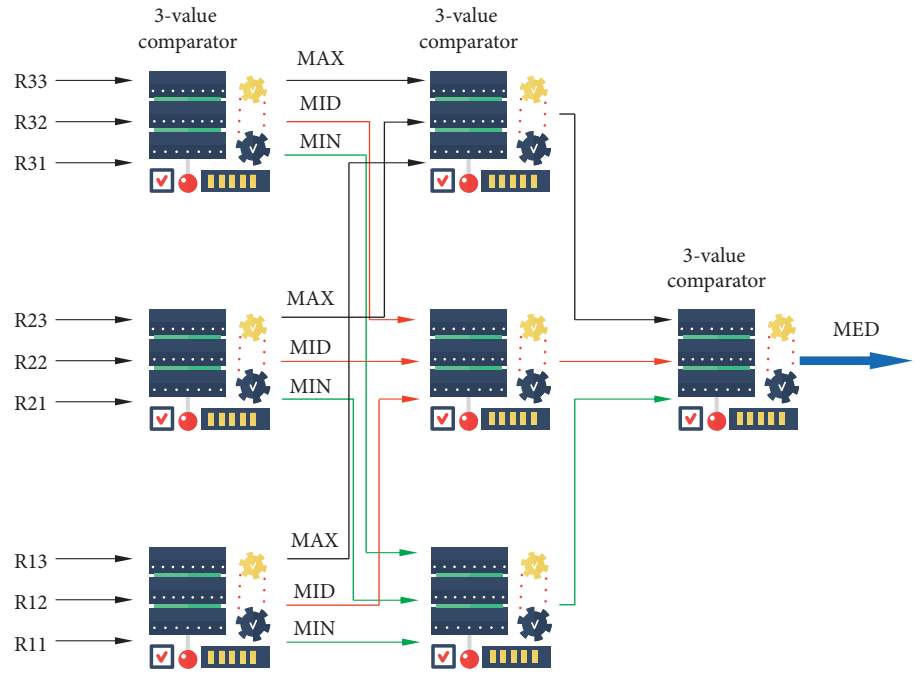


Figure 2: Median filtering algorithm.

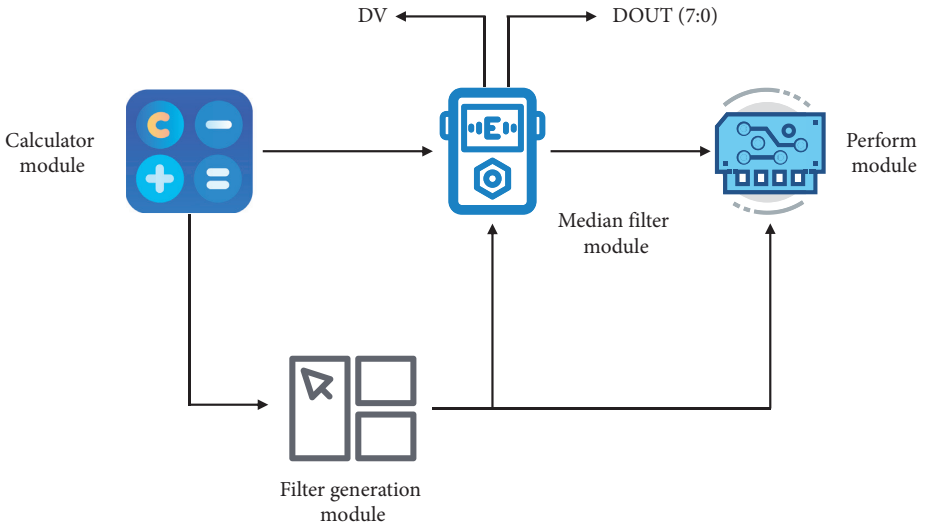


Figure 3: Improved median filter hardware circuit.

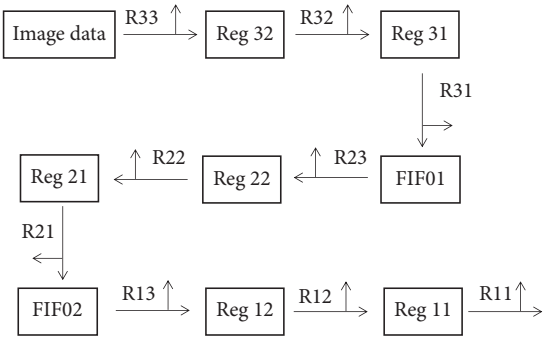


Figure 4: Algorithmic logic circuit diagram.

gradient amplitude, gradient intercomparison, and threshold comparison; the process of the edge detection algorithm is shown in Figure 5.

As can be seen from Figure 5, the gradient values in each direction are calculated separately and comparing the maximum value with the threshold value is the focus of edge extraction, so the computational analysis of the gradient is the main part of the algorithm implementation, and the gradient comparison structure can be divided into two parts, that is, comparing each other to arrive at the larger value and then comparing it to arrive at the maximum value.

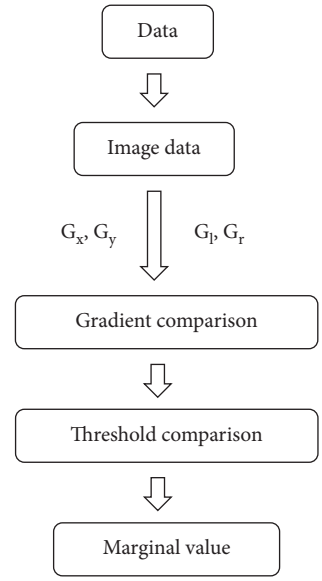


Figure 5: Edge detection algorithm.

3. Visual Positioning Algorithms

3.1. Binocular Stereo Vision Cameras. Digital imaging technology transforms the surrounding scene into an image and accurately reproduces the real 3D scene by means of specific algorithms, the interrelationships of which are shown in the following equation:

$$Z_C \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha_u & \gamma & u_o & 0 \\ 0 & \alpha_v & v_o & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0_3^T & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}, \tag{2}$$

where a point is defined as $(x_w, y_w, z_w, 1)$ in the world coordinate system, $(x_c, y_c, z_c, 1)$ in the camera coordinate system, and $(u, v, 1)^T$ in the digital image coordinate system. The image centre coordinates are u_o and v_o, and the information about the camera itself represented by the matrix $A = \begin{bmatrix} \alpha_u & \gamma & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}$ is called the camera internal reference matrix, where α_u and α_v represent the scale factor of the u and v axes of the image plane, respectively, and γ is the tilt of the image coordinate axis factor. $R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$, $T = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix}$ represents the rotation and translation between the world coordinate system and the camera coordinate system during the entire imaging process. Define the matrix $[R|T]$ as the external parameters of the camera and the matrix $A[R|T]$ as the camera perspective projection matrix P. The calibration of the camera is the process of calculating the internal and external parameters by obtaining the surrounding information in a certain way.

Generally, in stereo vision systems, we need to calibrate two cameras, unlike monoculars where not only the internal and external parameters of each camera are required, but also the relative positions of the two cameras. By using the above method, the internal and external parameters of the individual camera are prepared for the next step of the calculation. The results of the external parameters are expressed in the matrices $[R_l|T_l]$ and $[R_r|T_r]$. Let x_w, x_{cl}, and x_{cr} be the world coordinates and the left and right camera coordinates of a point, respectively:

$$\begin{aligned} x_{cl} &= R_l x_w + T_l, \\ x_{cr} &= R_r x_w + T_r. \end{aligned} \tag{3}$$

The relative positions of the binocular cameras are represented by R_0 and $T_{0,}$ and the coordinates are collated as follows:

$$\begin{aligned} R_o &= R_l R_r^{-1}, \\ R_o &= T_l - R_r^{-1} T_r. \end{aligned} \tag{4}$$

It can be concluded from the above equation that the external reference of the binocular camera can be calculated from the template of the same orientation, and the coordinates can be obtained by substituting the result into (4), where it should be noted that the position of both cameras cannot change when obtaining the template information.

3.2. Separation of Targets. The similarity between the contours of two objects can be quickly obtained by calculating the contour moments, which are generally a feature of the object that provides an approximate representation of the target's contour:

$$mp, q = \sum_{t=1}^{n} I(x, y) x^p y^q, \tag{5}$$

where p is the moment in the x-direction and q is the moment in the y-direction. The algorithm is an integration operation for each point on the target contour. If p and q are both equal to 0, then m_{00} is the sum of the number of pixel points on the target contour.

To solve the problem of degradation of matching accuracy due to size and rotation, a new method of moments, H-moment, is proposed. It is a normalised central moment with size and rotation invariance, ensuring matching accuracy.

Centre moment:

$$mp,q = \sum_{t=1}^{n} I\left(x - x_{\text{avg}}\right)\left(y - y_{\text{avg}}\right)x^p y^q. \tag{6}$$

Normalised matrix:

$$\eta_{p,q} = \frac{\mu_{p,q}}{m_{00}^{(p+q)/2+1}}. \tag{7}$$

The H-moment is a linear combination of the normalised centre distances.

In order to be able to avoid the inaccurate identification of targets due to the actual influence factor, a model is built for each point or group of points in the background with respect to time. It is effective in resolving time-dependent changes, but only at the cost of a large memory footprint. An analysis of the existing hardware equipment shows that the processing of the captured video information in the way described above is very computationally intensive and does not achieve the required processing speed, making it impractical to model the background in this way.

The analysis and study of video compression techniques has led to a background modelling algorithm that can perform similarly to the above method, that is, constructing a codebook (codebook) to represent the state in the background. The simplest way to do this is to compare the current value of a pixel with its past value; if the two values do not differ significantly, then it is defined as interference under the corresponding pixel. If the pixel difference is large, then it can generate a range of colours corresponding to it. To effectively address the problem of changing backgrounds, we have introduced codebook background modelling. It consists of a number of boxes containing pixels that are constant over time. The method models each pixel or group of pixels through changes in time, observing the value of this pixel at the corresponding position at different points in time and deriving a curve for each pixel with respect to time, which is then encoded and used to construct a background model.

The encoded pixel is defined as symbol code_elements, and then, all symbols are gathered together to form a codebook with the same scale as the background image. The extraction of information about changes in the background is achieved by using the function update_codebook() for all pixels. This process can be updated continuously, while the clear_stale_entries() function is used to train backgrounds that may be foreground targets. This is to remove those background variations caused by real foreground targets. When extracting foreground targets using the codebook method, it is first necessary to add an adjustment value maxMod and minMod to the two borders of each codebookbox. If the pixels in the box are made to add maxMod to the high part of each channel or subtract minMod from the low part, the value of matchChannel is added by one. When the matchChannel and the number of channels are the same, each dimension is searched and a match is known to have been made. If a pixel point is in the box for training at this point, then 255 is returned (deciding that this point is the foreground target), or 0 if it is not in the box (deciding that this point is the background).

The implementation of codebook foreground target separation using OpenCV is usually divided into the following steps:

(1) Using the function update_codebook() to construct an original background model within a certain time frame

(2) Call the clear_stale_entries() function to clear the stale index

(3) Setting the appropriate minMod and maxMod to achieve accurate separation of the identified foreground targets

(4) Maintaining a higher level scene model

(5) Using the function background_diff() to separate foreground targets

(6) Periodically update the learned background pixels

3.3. Target Tracking and Identification. Recognition refers to the extraction of the desired target from the observed environment. The moment mentioned in template matching can help identify the target object to be concerned. The most commonly used method of tracking unknown objects is to extract the visual feature points of the target, track these features, and then track the whole target. OpenCV contains two methods to track key points: Lucas Kanade and horn Schunk methods. These two methods represent the commonly mentioned sparse and dense optical flow.

Dense optical flow uses the relationship between each point in the captured image information and velocity, or the relative movement of the same point before and after the motion of the target object to estimate the trajectory of the object. This method of motion estimation is all achieved by the relationship between the point and the velocity of the motion. It is not feasible to use dense optical flow for motion estimation of the target, and it is quite computationally intensive, so we found an alternative method, sparse optical flow. Estimation using sparse optical flow is predicated on first providing a series of target-specific points. OpenCV can help us to find the most suitable tracking angle point. From the above presentation, it can be analysed that the sparse optical flow method is much superior to the dense optical flow in terms of computational speed and complexity.

The use of corner point detection can quickly obtain the features of the acquired image and is widely used in target tracking, motion estimation, template matching, and other fields because of its fast and stable characteristics, also known as feature point detection. The pixels around a corner point should exist on at least two different boundaries, and the corner point can be said to be the intersection of two boundaries. However, in practice, the corner points extracted by the detection methods used are generally

feature points that represent the target features and are not always just "corner points." The Harris feature point detection algorithm was obtained by improving the Moravec algorithm. It incorporates a Gaussian filter function in order to be able to avoid the effect of noise on the image during detection:

$$w(x, y) = \frac{1}{2\pi\sigma^2} e^{-\left(x^2+y^2\right)/2\sigma^2}. \tag{8}$$

Moravec corner detection only calculates feature points 45 degrees apart, whereas the Harris algorithm uses the Taylor formula to extract feature points in each direction:

$$\begin{aligned} E(u, v) &= \sum_{x,y} w(x, y)[I(x+u, y+v) - I(x, y)]^2 \\ &= \sum_{x,y} w(x, y)\left[I_x u + I_y v + O\left(u^2, v^2\right)\right]^2. \end{aligned} \tag{9}$$

Matrix form:

$$\begin{aligned} M &= \sum_{x,y} w(x, y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \\ &= w(x, y) * \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}, \\ E(u, v) &= [u, v] M \begin{bmatrix} u \\ v \end{bmatrix}. \end{aligned} \tag{10}$$

where I_x is the difference in the x-direction, I_y is the difference in the y-direction, and $w(x, y)$ is a Gaussian function.

The Harris feature point determination method is not available in traditional algorithms. Since the eigenvectors x_1 and x_2 of the autocorrelation matrix M are proportional to their main tendency to change, Harris uses x_1 and x_2 to indicate the orientation of the pixel values in terms of how quickly they change; that is, if x_1 and x_2 are close to each other and both are large, the pixel is a corner point; if x_1 and x_2 are very different, the point is an edge; and if x_1 and x_2 are both small, the point must not be a corner point.

Obtaining the eigenvectors involves a great deal of computation, and it is known from linear algebra that the trace of a matrix is the sum of the eigenvalues of the matrix and their product is equal to the determinant of the matrix. Therefore, thefollowing equation is used to select the most probable eigenpoints:

$$R = \Delta M - k(\text{trace}M)^2. \tag{11}$$

3.4. Kalman Filters. The Kalman filtering method was introduced in the 1960s and has since become an indispensable method in signal processing research. What the Kalman filter was originally intended to achieve was that if a set of convincing assumptions existed, when measurements were obtained for the entire target tracking process, then a model could be constructed to verify the probability of the current calculated value being correct. The previous measurements in this model do not need to be stored for a long time; that is, the content is always updated. The requirements for hardware equipment are therefore reduced, thus increasing the breadth of utilisation of this method.

The Kalman model is a linear function F of the target state. The model will be related to the combination of the first- and second-order derivatives of the previous motion step. Controlling the processing of the control input u_k in the model will result in a more realistic observed model Z in which only a few model state variables need to be measured and there is no direct link between the measured values and the state variables. If the current estimate has a large jump, then the predicted value from the previous movement will be used instead of the current measurement. Conversely, if the previous prediction is not accurate, then a more accurate measurement needs to be obtained and the result considered accurate. If both the current measurement and the previous prediction are stable, the expectation of the current position must exist somewhere in between them. The above discussion is consistent with our expectations. Figure 6 represents how uncertainty changes over time in response to new observations.

Since updates are sensitive to uncertainty, some further notation needs to be introduced in order to solve this problem. The state at time k is introduced as a function of the state at time $k-1$:

$$x_k = Fx_{k-1} + Bu_k + w_k, \tag{12}$$

where x_k denotes an n-dimensional vector of state elements and the transfer matrix F is an $n * n$ matrix multiplied by x_k. Vector u_k is a new addition, which serves to allow external control to be applied to the system and consists of a c-dimensional vector representing the input control. B is an $n * c$ matrix linking input control and state change. The variable w_k is a random event or external force that directly affects the state of the system. The elements of x_k are assumed to have a Gaussian distribution $N(0, Q_k)$ and $n * n$ covariance moments Q_k. It is often not entirely possible to determine whether z_k is a direct measure of the state variable x_k:

$$z_k = H_k x_k + v_k, \tag{13}$$

where H_k is the matrix of $m * n$ and v_k is the measurement error, also assumed to have a Gaussian distribution $N(0, Q_k)$ and $m * m$ covariance matrix R_k.

3.5. Model Testing. Here is an example of measuring the motion of a basketball. The motion of the basketball is represented by two directions x and y and two velocities v_k and v_y. From this information, the basketball's motion state vector x_k is formed:

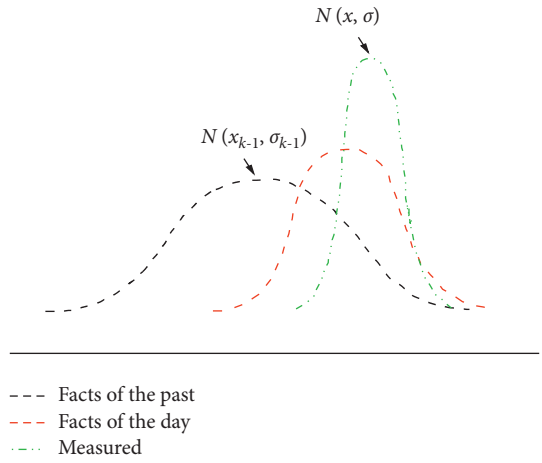


Figure 6: Curve of uncertain change.

$$x_k = \begin{bmatrix} x \\ y \\ v_x \\ v_y \end{bmatrix}, \quad F = \begin{bmatrix} 1 & 0 & dt & 0 \\ 0 & 1 & 0 & dt \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{14}$$

The change of relative position of the basketball movement can only be represented in the camera and only the position variable can be obtained:

$$z_k = \begin{bmatrix} z_x \\ z_y \end{bmatrix}_k. \tag{15}$$

Therefore, H can be expressed as the following structure:

$$H = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}. \tag{16}$$

The basketball is not moving at a uniform speed, so a value Q_k is needed to reflect this. The current position of the basketball is estimated by the method mentioned earlier, and then, the choice of R_k is made based on its accuracy.

Embedding the above expression into the broader update equation requires only the computation of an a priori estimate x_k^- for the next state:

$$x_k^- = Fx_{k-1} + Bu_{k-1} + w_k. \tag{17}$$

The above formula gives an idea of the value to be expected in the next step from the results already obtained. This leads to the Kalman update rate or mixing ratio, which provides the most useful predictive information for the current measurement:

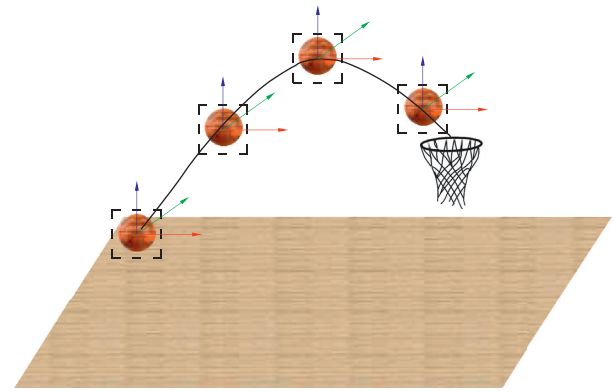


Figure 7: Basketball movement and actual movement predicted by the system.

$$K_k = P_k^- H_k^T \left(H_k P_k^- H_k^T + R_k\right)^{-1}. \tag{18}$$

If the one-dimensional case of a motion variable is obtained directly, H_k is just an identity matrix with one dimension. So, if the measurement error is σ_{2k+1}, then R_k can only be an identity matrix with one dimension. And similarly, P_k is exactly this covariance, σ_{2k}:

$$K = \frac{\sigma_k^2}{\sigma_k^2 + \sigma_{k+1}^2}. \tag{19}$$

When the current measurement value is obtained through the update rate, the optimal current value of x_k and P_k can be obtained:

$$\begin{aligned} x_k &= x_k^- + K_k\left(z_k^- - H_k x_k^-\right), \\ P_k &= \left(I - K_k H_k x_k^-\right), \end{aligned} \tag{20}$$

where I is the identity matrix; when the system enters the $k+1$ state, P_k is P_k^- in (18), so the algorithm can be repeated.

Taking the basketball moving in the shooting process as an example to observe the effect of Kalman filtering, Figure 7 is the graphical representation of basketball movement and actual movement predicted by the system. The box in the figure indicates the position predicted by the system.

It can be seen from the above experimental results that the Kalman filtering method can achieve better and more accurate tracking and positioning of moving objects.

4. Conclusion

Aiming at the shortcomings of shooting angle correction methods, this article proposes a shooting direction correction method based on visual perception in basketball training. From visual robot localisation, template matching, background modelling and foreground object separation to target point extraction, motion estimation, and Kalman filtering, the visual target localisation and its effect process are based. The traditional corner detection algorithm is deeply understood and analysed, and some improved methods are proposed. The experimental results show that this method has a good effect on shooting tracking in basketball training and can basically accurately reflect the position of basketball movement, which is helpful to basketball players' shooting training.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] J. Xiong, Z. He, R. Lin et al., "Visual positioning technology of picking robots for dynamic litchi clusters with disturbance," *Computers and Electronics in Agriculture*, vol. 151, pp. 226–237, 2018.

[2] Y. Chen, S. Wang, J. Bi, and H. Ai, "Study on visual positioning and evaluation of automatic evisceration system of chicken," *Food and Bioproducts Processing*, vol. 124, pp. 222–232, 2020.

[3] V. Margarita, "Sotnikova, control system design for visual positioning of a ship based on NMPC and multi-objective structure," *IFAC-PapersOnLine*, vol. 51, no. 32, pp. 445–450, 2018.

[4] Y. Bian, X. Fang, M. Yang, and J. Yang, "Underwater visual position measurement system for high-accuracy beam installation," *Optik*, vol. 127, no. 8, pp. 964–3968, 2016.

[5] J. Fan, S. Bi, R. Xu, L. Wang, and L. Zhang, "Hybrid lightweight Deep-learning model for Sensor-fusion basketball Shooting-posture recognition," *Measurement*, vol. 189, Article ID 110595, 2022.

[6] Y. Wang, M. Sun, and L. Liu, "Basketball shooting angle calculation and analysis by deeply-learned vision model," *Future Generation Computer Systems*, vol. 125, pp. 949–953, 2021.

[7] R. R. Oudejans, R. W. van de Langenberg, and R. I. Hutter, "Aiming at a far target under different viewing conditions: visual control in basketball jump shooting," *Human Movement Science*, vol. 21, no. 4, pp. 457–480, 2002.

[8] J. Gun, "Basketball action recognition based on FPGA and particle image," *Microprocessors and Microsystems*, vol. 80, Article ID 103334, 2021.

[9] W. Ding, J. Gu, S. Tang, Z. Shang, E. A. Duodu, and C. Zheng, "Development of a calibrating algorithm for Delta Robot's visual positioning based on artificial neural network," *Optik*, vol. 127, no. 20, pp. 9095–9104, 2016.

[10] Z. Wang, H. Li, and X. Yang, "Vision-based robotic system for on-site construction and demolition waste sorting and recycling," *Journal of Building Engineering*, vol. 32, Article ID 101769, 2020.

[11] L. Bergamini, M. Sposato, M. Pellicciari, M. Peruzzini, S. Calderara, and J. Schmidt, "Deep learning-based method for vision-guided robotic grasping of unknown objects," *Advanced Engineering Informatics*, vol. 44, Article ID 101052, 2020.

[12] T. P. Nguyen and J. Yoon, "A novel vision-based method for 3D profile extraction of wire harness in robotized assembly process," *Journal of Manufacturing Systems*, vol. 61, pp. 365–374, 2021.

[13] Z. Lu, M. Zhao, J. Luo, G. Wang, and D. Wang, "Design of a winter-jujube grading robot based on machine vision," *Computers and Electronics in Agriculture*, vol. 186, Article ID 106170, 2021.

[14] S. Mishra and S. Jabin, "Recent trends in pedestrian detection for robotic vision using deep learning techniques," in *Artificial Intelligence for Future Generation Robotics*, N. S. Rabindra, G. Ankush, E. B. Valentina, and B. Monica, Eds., Elsevier, Amsterdam, Netherlands, pp. 137–157, 2021.

[15] H. Yang, L. Chen, Z. Ma et al., "Computer vision-based high-quality tea automatic plucking robot using Delta parallel manipulator," *Computers and Electronics in Agriculture*, vol. 181, Article ID 105946, 2021.

[16] W.-S. Kim, D.-H. Lee, Y.-J. Kim, T. Kim, W.-S. Lee, and C.-H. Choi, "Stereo-vision-based crop height estimation for agricultural robots," *Computers and Electronics in Agriculture*, vol. 181, Article ID 105937, 2021.

[17] R. Wang, A. Wu, X. Chen, and J. Wang, "A point and distance constraint based 6R robot calibration method through machine vision," *Robotics and Computer-Integrated Manufacturing*, vol. 65, Article ID 101959, 2020.

[18] A. Anna, "The use of vision systems in the autonomous control of mobile robots equipped with a manipulator," *Transportation Research Procedia*, vol. 40, pp. 132–135, 2019.

[19] W. Fang, F. Chao, L. Yang et al., "A recurrent emotional CMAC neural network controller for vision-based mobile robots," *Neurocomputing*, vol. 334, pp. 227–238, 2019.

[20] M. D. Hazrat Ali, K. Aizat, K. Yerkhan, T. Zhandos, and O. Anuar, "Vision-based robot manipulator for industrial applications," *Procedia Computer Science*, vol. 133, pp. 205–212, 2018.

[21] M. Badawy, H. Khalifa, and H. Arafat, "New approach to enhancing the performance of cloud-based vision system of mobile robots," *Computers & Electrical Engineering*, vol. 74, pp. 1–21, 2019.

[22] L. Xiao, Y. Zhao, L. Gong, C. Liu, and T. Wang, "Dual-arm cooperation and implementing for robotic harvesting tomato using binocular vision," *Robotics and Autonomous Systems*, vol. 114, pp. 134–143, 2019.

[23] S. V. Shavetov, I. I. Merkulova, A. A. Ekimenko, O. I. Borisov, and V. S. Gromov, "Computer vision in control and robotics for educational purposes," *Computer Vision in Control and Robotics for Educational Purposes∗∗This work was supported by the Ministry of Education and Science of Russian Federation*, vol. 52, no. 9, pp. 127–132, 2019.

[24] K. Asadi, H. Ramshankar, H. Pullagurla et al., "Vision-based integrated mobile robotic system for real-time applications in

construction," *Automation in Construction*, vol. 96, pp. 470–482, 2018.

[25] R.-J. Halme, M. Lanz, J. Kämäräinen, R. Pieters, J. Latokartano, and A. Hietanen, "Review of vision-based safety systems for human-robot collaboration," *Procedia CIRP*, vol. 72, pp. 111–116, 2018.

Hindawi
Scientific Programming
Volume 2022, Article ID 9022563, 11 pages
https://doi.org/10.1155/2022/9022563

Research Article

CAD-Aided 3D Reconstruction of Intelligent Manufacturing Image Based on Time Series

Liming Zhang, Lei Wang, Xu du, and Fanbo Meng

College of Information Science and Electronic Technology, Jiamusi University, Jiamusi 154007, China

Correspondence should be addressed to Lei Wang; wangl@jmsu.edu.cn

Received 1 December 2021; Revised 25 December 2021; Accepted 7 January 2022; Published 11 March 2022

Academic Editor: Baiyuan Ding

To improve the three-dimensional (3D) reconstruction effect of intelligent manufacturing image and reduce the reconstruction time, a new CAD-aided 3D reconstruction of intelligent manufacturing image based on time series was proposed. Kinect sensor is used to collect depth image data and convert it into 3D point cloud coordinates. The collected point cloud data are divided into regions, and different point cloud denoising algorithms are used to filter and denoise the divided regions. With the help of CAD, FLANN matching algorithm is used to extract feature points of time-series images and complete image matching. Three-dimensional reconstruction of sparse point cloud and dense point cloud is carried out to complete 3D reconstruction of intelligent manufacturing images. The experimental results show that the image PSNR of this method is always above 52 dB, and the maximum reconstruction time is 4.9 s. The 3D reconstruction effect of intelligent manufacturing image is better, and it has higher practical application value.

1. Introduction

In recent years, in the field of information technology and industry, significant changes have taken place, such as large data, cloud computing, mobile Internet, 3D printing, and industrial robots, and these changes have brought a new round of revolution in global manufacturing, including intelligent manufacturing as a product of informatization and industrialization depth fusion, but also got great progress. It has been widely concerned and valued by the governments of various countries [1], such as the "Advanced Manufacturing National Strategy" of the United States, the "New Industry France Plan" of France, the "New Industry Creation Strategy" of Japan, and the "Industry 4.0" of Germany [2]. An important reason why developed countries attach importance to intelligent manufacturing is that the financial crisis in 2008 fully exposed the fragility of the virtual economy, so many countries have re-examined the important role of manufacturing in social development, taking intelligent manufacturing as a breakthrough to improve the country's comprehensive competitiveness [3]. Intelligent manufacturing helps to improve the level of China's traditional manufacturing industry, achieve high-end technological innovation, relieve energy pressure, promote the transformation of new production modes, and accelerate the transformation from a manufacturing country to a manufacturing power. However, due to the continuous development of manufacturing technology, information technology, and network technology, the concept and connotation of intelligent manufacturing are also in constant change, enrichment, and improvement [4]. At this stage, the field has generally recognized the definition of intelligent manufacturing and believes that intelligent manufacturing integrates a variety of modern technologies, which can realize the monitoring and decision-making in the process of intelligent production and improve the intelligent degree and production efficiency of product production [5]. Therefore, in order to cope with the competitive market environment, ensure the clarity of intelligent manufacturing images, and enhance the competitiveness of enterprises.

Image 3D reconstruction has become the focus of current business and academic circles. So it is very important to study the 3D reconstruction method of intelligent manufacturing image.

There is not much existing image design 3D reconstruction methods in intelligent manufacturing, so we just need to migrate to other areas of image 3D reconstruction method, for example by Zhuang and Wan [6], who proposed a framework based on P2M single-image reconstruction of 3D model of the improved method, mainly because the traditional 3D reconstruction method cannot rebuild object's invisible part, it takes a long time to reconstruct, and it is also difficult to reconstruct objects without texture. Three-dimensional reconstruction is carried out around a single image. The main work is to replace the backbone of feature establishment, improve the feature extraction network by VGG-16 network, and propose an improved DenseNet network to obtain better two-dimensional feature points. To improve the reconstruction effects, Liu et al. [7] proposed a single-image 3D reconstruction method based on vanishing point optimization. First, the image is processed, the long lines in the image are extracted, the characteristics of the lines are analyzed, and the lines in different directions are grouped. Then, according to the linear distribution relationship of the lines in each direction, the linear model of the line parameters is obtained by the improved regression algorithm, the error lines are eliminated, and the extinction point is solved by the least square method. After the accurate extinction point is obtained, the internal and external parameters of the camera are obtained according to the properties of the extinction point. The minimum two-dimensional points of the image are obtained through interaction. The 3D reconstruction of the object is realized by calculating the 3D coordinates of the camera parameters and the geometric features of the object. Zhao et al. [8] proposed a 3D reconstruction method based on sea ice scene image classification. Aiming at the problem of large amount of computation and easy accumulation of errors in 3D reconstruction process, this method uses a classification network to screen suitable image sequence through training and uses 8-neighborhood filling algorithm to segment the reconstructed region of the obtained specific category sea ice scene image and finally carries out the 3D reconstruction of the segmented image. Yu et al. [9] take flange parts as an example to propose a reconstruction method of flange parts based on working drawing images. This method takes the scanned images of engineering drawings as the research object; carries out noise reduction, segmentation, and refinement processing; and then uses Harris corner recognition algorithm and Hough circle detection algorithm to extract and make statistics on their upper features. Back-propagation (BP) neural network was used to identify and extract the necessary parameters of parts reconstruction, and finally realized the three-dimensional reconstruction of flange model. Liu et al. [10] proposed a 3D reconstruction method of fuzzy image edge contour based on machine learning. Through the detection of image edge contour, the method of stereo matching is used to find the corresponding points in the detected image edge, so as to obtain 3D coordinates. According to the obtained 3D coordinates, the 3D point cloud of the object surface is obtained, and the 3D reconstruction of fuzzy image edge contour is realized based on machine learning.

But the aforementioned methods' image denoising processing effect is poor, the reconstruction effect is poorer, and the reconstruction takes longer, so this article proposes to solve the aforementioned problems existing in the method as the research target, to access the sensor as the tool of depth image data acquisition, combined with ordinary filter algorithm and bilateral filtering algorithm as the theoretical basis, to deal with the noise in point cloud data. Based on FLANN matching algorithm, feature points of time-series images were extracted, and 3D reconstruction of sparse point cloud and dense point cloud was carried out by CMVS algorithm and PMVS algorithm, respectively, so as to complete 3D reconstruction of intelligent manufacturing images.

2. CAD-Aided Intelligent Manufacturing Image 3D Reconstruction Method Design

2.1. Point Cloud Data Collection. As a 3D depth sensor, Kinect has the functions of real-time dynamic capture, voice recognition, microphone input, and intelligent interaction in addition to obtaining color images and depth images applied in this paper. The appearance is shown in the figure below. The external structure is simple and light, and the internal device is complex and powerful [11]. Kinect sensors emit infrared light using color and depth cameras to collect raw data, which are combined to generate 3D data [12, 13]. The imaging principle of Kinect camera is similar to that of pinhole imaging. In this paper, the pinhole model is used to introduce the camera imaging principle. Assume that P represents any object in space, the pinhole plane is the vertical plane where the camera view is located [14], and the imaging plane after P passes through the optical center is called the image plane. The horizontal ray passing through the optical center and perpendicular to the two planes is called the optical axis, and the intersection between the optical axis and the pinhole plane is called the main point. The distance between the image plane and the pinhole plane f is the focal length of the camera, the horizontal distance between P and the pinhole plane is Z, the real height of the object is X, the imaging height is x, and the imaging point is p [15]. According to the similar triangle principle:

$$\frac{x}{f} = -\frac{X}{Z}. \tag{1}$$

Considering the image coordinate system, axis u represents the rows of the image and axis v represents the columns of the image [16]. Then, the coordinate conversion relation of $p(u, v)$ and $P(x, y)$ is as follows:

$$\begin{cases} u = \dfrac{x}{k} + u_0, \\ \\ v = \dfrac{y}{l} + v_0. \end{cases} \tag{2}$$

The relationship between camera coordinate system and image coordinate system can be expressed by the following formula:

$$\begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = \frac{1}{Z_c} \begin{bmatrix} f & -f\cot\theta & 0 & 0 \\ 0 & \dfrac{f}{\sin}\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix}. \tag{3}$$

In the aforementioned formula, θ represents the included angle between two coordinate systems, and f represents the focal length of the camera. Combining the aforementioned two equations, the relationship between camera coordinate system and image pixel coordinate system can be expressed by the following formula:

$$\begin{aligned} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} &= \frac{1}{Z_c} \begin{bmatrix} \dfrac{f}{k} & -\dfrac{f}{k}\cot\theta & u_0 & 0 \\ 0 & \dfrac{f}{l \cdot \sin\theta} & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \\ &= \frac{1}{Z_c} A P_c. \end{aligned} \tag{4}$$

In the aforementioned formula, (k, l, u, v, f, θ) represents camera parameters and A represents parameter matrix [17]. For the same pixel, the relationship between the world coordinate system and the camera coordinate system can be expressed by the following formula:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R_{3\times3} & t_{3\times1} \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}. \tag{5}$$

In the aforementioned formula, $R_{3\times3} = (r_x r_y r_z)$ is the rotation matrix, $t_{3\times1} = (t_x t_y t_z)^T$ is the translation vector, and (R, t) is the external parameter matrix of the camera.

Combining the aforementioned two equations, the relationship between the image pixel coordinate system and the world coordinate system can be obtained, which can be described by the following formula:

$$s\tilde{m} = A(R, t)\tilde{M}. \tag{6}$$

In the aforementioned formula, $\tilde{m} = (u, v, 1)^T$ represents the coordinates of pixel points, and its corresponding world coordinate system is $\tilde{M} = (X_w, X_w, X_w, 1)^T$, and s represents the scale coefficient.

The camera coordinate system consists of optical center o, coordinate axes X_c, Y_c, and Z_c, and the optical axis is represented by axis Z_c. The u and v planes of the image coordinate system constitute the image plane and are perpendicular to axis Z_c. The origin is the intersection of the optical axis and the image plane. The projection of point $P(x, y, z)$ in space onto the image plane is $p(u, v)$. Firstly, the points in the world coordinate system are transformed into the camera coordinate system, and then the camera coordinates are transformed into the image physical coordinate system according to the similar triangle principle, and finally the coordinates $p(u,v)$ in the image pixel coordinate system are obtained.

The original data collected by the Kinect sensor cannot be used directly but need to be converted into data that can be understood and directly applied. The raw output data generated by the sensor device are usually mapped to the grayscale space in the form of a depth image. The image takes the origin of depth image, and the unit is pixel. Each pixel represents the depth distance from Kinect. The function R Depth Gen.Convert Projective To Real World in Open Source library Open NI is used for coordinate conversion. The following relationship exists between the world coordinate system and the depth image coordinate system:

$$\begin{aligned} X &= (u - 300) \cdot Z \cdot \frac{1}{f_{xz}}, \\ Y &= (v - 300) \cdot Z \cdot \frac{1}{f_{yz}}. \end{aligned} \tag{7}$$

In the aforementioned formula, X, Y, and Z, respectively, represent the X, Y, and Z axes in the world coordinate system with Kinect device as the origin, and u and v, respectively, represent the horizontal and vertical coordinate axes of the depth image coordinate system. $f_{xz}f_{yz}$ is the focal length of the corresponding depth camera. The X and Y coordinates of the camera center are 320 and 240, respectively (image resolution 640×480). The coordinates after transformation are the real-world coordinates. Coordinate values of X, Y, and Z coordinate form 3D point cloud, and the depth image data are converted into 3D point cloud coordinates for subsequent processing.

The specific conversion process is as follows:

Since the origin of the world coordinates and the origin of the camera coincide, that is, there is no rotation and translation, the 3×3 rotation matrix R and 3×1 translation matrix t are calculated as follows:

$$\begin{aligned} R &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} t \\ &= \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \end{aligned} \tag{8}$$

At this point, the coordinate origin of the camera coordinate system and the world coordinate system coincide, so the same object under the camera coordinate system and the world coordinate system has the same depth, namely $Z_c = Z_w$. Then there is

$$Z_c \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{bmatrix} \frac{f}{\mathrm{dx}} & 0 & u_0 \\ 0 & \frac{f}{\mathrm{dy}} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}. \tag{9}$$

From the aforementioned transformation matrix formula, the coordinate transformation formula of depth image point conversion into 3D point cloud can be calculated:

$$\begin{cases} x_w = z_c \cdot (u - u_0) \cdot \dfrac{\mathrm{dx}}{f}, \\ y_w = z_c \cdot (v - v_0) \cdot \dfrac{\mathrm{dy}}{f}, \\ z_w = z_c. \end{cases} \tag{10}$$

2.2. Point Cloud Data Filtering and Denoising. To obtain better denoising results of point cloud model, the features of the model can be preserved while removing noise. This paper first divides the scanned point cloud data into regions, and then adopts targeted point cloud denoising algorithms for the divided regions [18]. The specific region division method is as follows: the local surface fitting method is used to calculate the differential geometric information of the point cloud, and the local feature weights are set according to the mean curvature of the model point cloud obtained by calculation. Then the local feature weight threshold is given, and then the neighborhood feature weight of the sampling point is judged with the set threshold. Thus, the region of the point cloud data of the model is divided, and the flat region with less feature information and the region with more feature information are obtained [19].

Regional division is carried out by calculating the local feature weights of points. Given that the average curvature at sampling point p_i is H_i, the local feature weights of sampling points in k-nearest neighbor neighborhood are defined as follows:

$$s(p_i, k) = \sqrt{\frac{1}{k} \sum_{i}^{k} \left(\left| H_{p_i} \right| - \overline{H} \right)} + \sqrt{(H_i - \overline{H})}. \tag{11}$$

where $\overline{H}$ is the average curvature in the neighborhood of the sampling point [20]:

$$\overline{H} = \frac{1}{n} \sum_{i=1}^{n} H_i. \tag{12}$$

If the local feature weight $s(p_i, k)$ at sampling point p_i is less than the threshold value $s_{\max}$ of the set feature weight, then sampling point p_i is judged to be a point in the flat region with less feature information [21]. If the local feature weight value $s(p_i, k)$ at sampling point p_i is greater than the threshold value $s_{\max}$ of the set feature weight value, then sampling point p_i is judged to be the point in the area with more feature information.

For point cloud data with well-divided feature types, targeted filtering denoising algorithm is adopted [22–24]. The neighborhood distance average filtering algorithm and the adaptive bilateral filtering algorithm are used to denoise the two different feature areas.

2.2.1. Point Cloud Denoising in Flat Area. Owing to the characteristic information of the few curvature changes more gently on the surface of the flat area of the model, the low complexity of denoising methods, namely using sampling point k-nearest neighborhood interior point to the average distance between neighboring points of statistical filtering algorithm for discrete points and noise points in the area to remove noise effectively after receiving new model of point cloud data [25], as follows:

$$P' = \left\{ p_i \in P \middle| (\mu_n - \alpha\sigma_n) \le \overline{d}_i \le (\mu_n + \alpha\sigma_n) \right\}, \tag{13}$$

where $\overline{d}_i$ is the average distance between the inner point and the nearest neighbor point in the neighborhood μ_n of the sampling point K, μ_n is the average distance of the whole flat region, and σ_n is the standard deviation. The calculation formula for each parameter is as follows:

$$\overline{d}_i = \frac{1}{k} \sum_{j=1}^{k} \left\| p_{ij} - p_j \right\|. \tag{14}$$

$$\begin{aligned} \mu_n &= \frac{1}{n} \sum_{i=1}^{N} \overline{d}_i, \\ \sigma_n &= \sqrt{\frac{1}{N} \sum_{i=1}^{n} \left(\overline{d}_i - \mu_n \right)^2}. \end{aligned} \tag{15}$$

The algorithm steps are as follows:

(1) For each sampled data point p_i in the flat region with small curvature changes, all points in its k-nearest neighbor neighborhood are searched;

(2) Formulas (10) to (12) were used to calculate the average distance from k-nearest neighbor to nearest neighbor $\overline{d}_i$ of each sampling point p_i, and the average distance and mean square error of all sampling points in point cloud of flat region $\overline{d}_i$.

(3) The mean-variance theory is used to determine the noise points and discrete points to be removed [26], that is, formula (9) is used to judge. If the average distance $\overline{d}_i$ of sampling point p_i exceeds the given threshold, the point is removed.

2.2.2. Point Cloud Smoothness in Feature-Rich Regions. Compared with other denoising algorithms, bilateral filtering algorithm has more advantages in the noise removal process of 3D scattered point cloud. It was initially used in image processing to preserve the image contour. Owing to its ability to maintain high-frequency features, this paper chooses the point cloud bilateral filtering method to denoising the feature-rich region of point cloud containing noise [27]. The specific method is shown in formula (16):

$$p' = p + \beta \cdot \overrightarrow{n}, \tag{16}$$

where p' is the filtered point data, p is the initial point cloud data, $\overrightarrow{n}$ is the normal vector, and β is the bilateral filtering factor. Its calculation formula is shown as follows:

$$\beta = \frac{\sum_{i=1}^{k} z_1\left(\left\|p - p_i\right\|\right) z_2\left(\langle p - p_i, n\rangle\right)\langle p - p_i, n\rangle}{\sum_{i=1}^{k} z_1\left(\left\|p - p_i\right\|\right) z_2\left(\langle p - p_i, n\rangle\right)}, \tag{17}$$

where k represents the sampling points in the neighborhood nearest to the sampling point [28], z_1 and z_2 are Gaussian filtering functions in the spatial domain and frequency domain of bilateral filtering functions, respectively, and their specific forms are shown as follows:

$$\begin{aligned} z_1(x) &= e^{y^2/2\sigma_1^2}, \\ z_2(y) &= e^{-y^2/2\sigma_2^2}. \end{aligned} \tag{18}$$

In the aforementioned formula, σ_1 represents the effect factor of the distance between sampling point p_i and points in its neighborhood on point p_i. The larger the value of σ_1 is, the more neighborhood points will be selected. In this case, the denoising result will be better [29, 30], but the effect of point cloud feature preservation will be reduced. σ_2 is the action factor of logarithmic point p_i on the projection of the distance vector from sampling point p_i to the point in its neighborhood on the normal direction $\overrightarrow{n}$ of the point, which mainly plays a role in regulating the degree of feature retention of point cloud data. The larger the value of p_i is, the better the feature retention effect of the point cloud model will be. In general, σ_1 is the neighborhood radius of the point, and σ_2 is the standard deviation of the neighborhood point.

The specific algorithm steps are as follows:

(1) Search k-nearest neighbor points of each data point p_i in the region with drastic curvature changes in the point cloud model, that is, in the feature-rich region;

(2) For all its nearest neighbor points, the value of $x = \|p - p_i\|$ in parameter $z_1(x)$ and $y = \langle p - p_i, n\rangle$ in parameter $z_2(y)$ are calculated;

(3) Calculate $z_1(x)$ and $z_2(y)$ according to the calculation formula of 6 and 7 and the selected value of 8;

(4) The values of $z_1(x)$ and $z_2(y)$ calculated in Step (3) are substituted into the calculation formula of bilateral filter factor β to calculate bilateral filter factor β;

(5) Then, formula (14) is used to move the points in the feature-rich region normally to obtain the filtered new point cloud data;

(6) After all points are calculated in turn and new point cloud data are obtained through transformation, end.

2.2.3. D Image Reconstruction Based on Time Series. In the real world, the collection of a large number of data is related to time, and the data are time related. These kinds of data are called time series. Strictly speaking, time series refers to a group of numerical sequences formed by sequential observation values of the same phenomenon at different times. It is a limited sequence of real values recorded in the direction of the time axis, reflecting the characteristics of attribute values in the time order. If the data series is continuous, it is called continuous time series [31]. If the data sequence is discrete, it is called a discrete time series. In daily life, a large amount of time-series data will be generated in different fields, which can be referred to as "time-series data." By collecting, recording, and organizing these time-series data and using advanced data mining tools, we can find a lot of valuable knowledge from time series, which can be used to guide our work and life. Therefore, 3D reconstruction of intelligent manufacturing image is carried out in this paper with the help of CAD.

By analyzing the structure of CAD-shaped DXF file, the corresponding object information can be extracted. In intelligent manufacturing image object data, the starting point coordinates of a line are the corresponding following values after block codes 10, 20, and 30, and the end coordinates are the corresponding following values after block codes 11, 21, and 31. In the entity extraction process, the first judgment is whether the group code is 10. If so, it indicates that the subsequent group value is the entity coordinate. The extracted object is projected (converted) onto the image according to the transformation parameters, and the projected value is taken as the initial value of the least square template matching. As the objects in some images may occlude each other, it is necessary to eliminate the lines that occlude each other. On this basis, FLANN matching algorithm is used to extract the feature points of intelligent manufacturing time-series image and complete the image matching. AKAZE algorithm introduces FED, a fast display diffusion mathematical framework, to solve partial differential equations quickly, which improves the solving speed and quality.

The nonlinear diffusion filter describes the evolution of image brightness and controls the diffusion process using the scaling parameter as the divergence factor of thermal diffusion function. Usually, partial differential equations are used to solve the problem [32]. Owing to the nonlinear nature of differential equations, the scale space is constructed through the diffusion of image brightness. The classical nonlinear diffusion equation is as follows:

$$\frac{\partial L}{\partial t} = \operatorname{div}\left(c(x, y, t) * \nabla L\right). \tag{19}$$

In the aforementioned formula, L is the image brightness matrix, and div and ∇ represent the solution operation of divergence and gradient, respectively. Owing to the introduction of conduction function c into the diffusion equation, the diffusion can be adaptive to the local structure characteristics of the image. The conduction function depends on the local difference structure of the image and can be scalar or constant. The time parameter t corresponds to the scale factor and is controlled by image gradient in the diffusion process [33].

Assuming that the number of layers and towers are represented by O and S, respectively, the scale space is constructed by performing time t_i of the diffusion function:

$$\sigma_i(o, s) = 2^{(o+s/S)}. \tag{20}$$

AKAZE feature detection is also achieved by looking for Hessian local maximum points normalized at different scales [34, 35]. In the same scale, response values of eight field points and 18 upper-- and lower-level points are identified as extreme points, namely feature points. Hessian matrix formula is as follows:

$$L_H^i = \left(\frac{\sigma_i}{2^{oi}}\right)^2 * \left(L_{xx}^i L_{yy}^i - \left(L_{xy}^i\right)^2\right). \tag{21}$$

where σ_i represents the scale parameter, o is the corresponding group, L_{xx} and L_{yy} are the transverse and longitudinal differential, and L_{xy} represents the second-order cross-differential.

Based on the successful extraction of image feature points, FLANN matching algorithm is used to complete the rough matching process [36]. The specific process is as follows:

(1) Find the feature point P_1 in the image R_1 and the feature point P_1 in the image R_2, and get the feature set (P_1, P_2), traverse the images of all feature points R_1, be made up of all minimum distance matching points for collection, calculate the minimum distance from the collection of min d, $T = \mu \times \min d$, setting the threshold value if $d < T$, p_2 as p_1 candidates will match point.

(2) Perform step (1) for all feature points in image R_2 to obtain FLANN matching point pair (p_2, p_3) from figure R_2 to figure R_1. If $p_3 = p_1$, the matching is successful; otherwise, the point is removed, and the cycle continues until all image matching is completed.

According to the geometric constraints between images, the point cloud model is reconstructed by triangulation. On the basis of the camera motion, the 3D spatial position of feature points can be estimated. However, the depth information of pixels cannot be obtained only from a single image, so the method of triangulation (triangulation) is adopted to estimate. The principle of trigonometry is to determine the distance of a point by observing the angle of the same point in two places.

Consider that there are two images I_1 and I_2, and their camera light centers are O_1 and O_2. Taking the image on the left as a reference, the transformation matrix of the image on the right can be determined as T. Assuming that feature point P_1 in figure I_1, and I_2 of the feature points P_1 form corresponding, regardless of other factors, under ideal conditions, O_1P_1 and O_2P_2 will intersect at one point in the space, which is two feature points of the map point position in the 3D scene but cannot be totally eliminated due to the effects of noise, the two straight lines generally cannot smooth rendezvous in the actual situation, so the least square method is used to find the point closest to the intersection of the two lines. According to the definition in polar geometry theory, if x_1 and x_2 are set as the normalized coordinates of the two feature points, the relationship between them is as follows:

$$s_1 x_1 = s_2 R x_2 + t. \tag{22}$$

In the aforementioned formula, s_1 and s_2 are the depths of the two feature points. Multiply the left and right sides of the aforementioned formula by $\widehat{x}_1$ to obtain the following formula:

$$\begin{aligned} s_1 \widehat{x}_1 x_1 &= 0 \\ &= s_2 \widehat{x}_2 r x_2 + \widehat{x}_1 t. \end{aligned} \tag{23}$$

According to the formula, we can solve for s_1 and s_2 in turn. In practice, the influence of noise cannot be completely eliminated, so solving R and t may not make the result of the aforementioned equation exactly 0. Therefore, in practice, it is more common to replace the zero solution with the least square solution.

The point cloud results obtained by SFM algorithm can be optimized by means of clustering multiview stereo (CMVS) algorithm for image clustering classification. CMVS can reduce the time of dense matching, and on this basis, dense point clouds can be obtained through patch-based multiview STEREO (PMVS) processing. There are three necessary conditions for clustering: the clustering should be as small as possible under the condition that every cluster can be reconstructed; maintain density while removing mismatches and redundancies; and the result of reconstruction should be relatively complete. The overall process of the algorithm is shown in Figure 1.

The purpose of SFM screening is to screen out some points and take out the average value of each field location of feature points. Select images with high resolution and remove images that do not meet the requirements. The standard image segmentation method was adopted to segment the screened images, and the SFM feature points that were not included were constructed to construct image clusters. The efficiency basis was set, and the maximum efficiency was regarded as the behavior of the added clusters. The process of clustering classification and increasing images is repeated to obtain clusters that meet the requirements.

PMVS is a dense matching based on facet model, which can reconstruct the cluster of CMVS independently. Patches are generated based on sequential images and feature points,

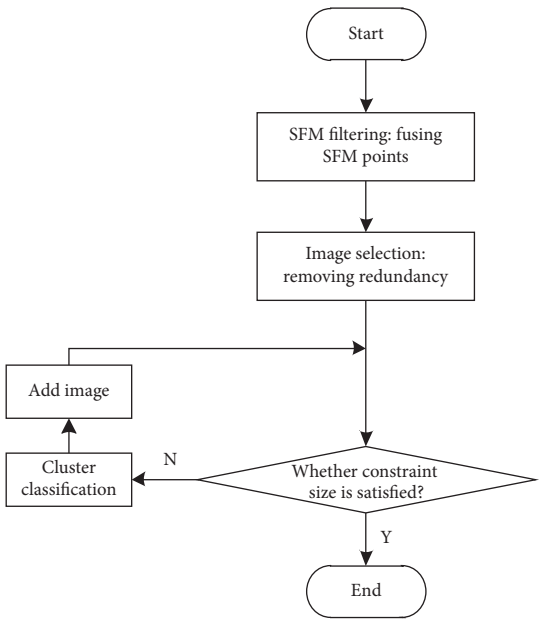


Figure 1: CMVS flow chart.

which are called seed points. The algorithm flow is shown in Figure 2.

It can be found that the algorithm consists of three parts: the first part is the generation of seed surface, the second part is the diffusion of surface, and the third part is the filtering of surface. The last two parts are iterated three times according to experience. After three iterations of surface diffusion and surface filtering, the number of outpoints in the diffusion reconstruction can be reduced, and the result is better after the third iteration. The normal section of a part of the surface of the target object represents a plane, and the seed plane refers to the sparse reconstruction of the spatial point cloud plane. First, corner detection is carried out on the sequence, and the obtained feature corner points are matched and 3D information is recovered according to triangulation measurement, and finally the face is generated. This process is called sparse reconstruction of 3D point cloud. For a given facet, it is impossible to ensure that all image blocks are included in the acquisition of facet, so the results of sparse reconstruction are extended to ensure that each image block has at least one facet. In this process, there will be wrong surfaces, which need to be filtered. The filtering is based on gray consistency and geometric consistency.

To sum up, this paper mainly uses Kinect sensor to collect depth image data and convert it into 3D point cloud coordinates. The collected point cloud data are divided into regions and the 3D point cloud data are filtered and denoised. With the help of CAD, FLANN matching algorithm was used to extract the feature points of time-series images, and then the images were matched. Combined with the image matching results, 3D reconstruction of sparse point cloud and dense point cloud was carried out to complete 3D reconstruction of intelligent manufacturing images.

3. Experimental Design and Result Analysis

3.1. Experimental Design. To verify the 3D reconstruction effect of CAD-aided intelligent manufacturing image based on time series designed in this paper, an experimental design is carried out. To ensure the reliability of the experimental results, this experiment needs to be carried out in the same environment, and the specific experimental environment parameters are shown in Table 1.

Multiple intelligent manufacturing images were collected, and the collected images were integrated to construct the experimental sample set. The experimental sample set was input into the computer, and the data in the sample set were processed using the reconstruction method based on P2M framework improvement, the reconstruction method based on vanishing point optimization and the reconstruction method based on time series. In this way, 3D reconstruction results of intelligent manufacturing images with different methods are obtained.

3.2. Analysis of Experimental Results. The PSNR of intelligent manufacturing image is one of the important indicators to verify the 3D reconstruction effect of intelligent manufacturing image. The higher the PSNR is, the better the 3D reconstruction effect of the image is. The specific comparison results are shown in Figure 3.

The analysis of the data in Figure 3 shows that with the increase of the amount of experimental data, the PSNR of images reconstructed by different methods changes in a fluctuating manner. Among them, the PSNR of images reconstructed by the improved P2M framework changes between 37.5 dB and 44.1 dB. Image reconstruction method based on vanishing point optimization PSNR between 33.9 dB and 40 dB changes, and reconstruction method based on improved P2M framework, based on the reconstruction of the vanishing point optimization method, the reconstruction method based on time-series images of the PSNR is always above 52 dB, reconstruction after using this method of intelligent manufacturing high image quality, intelligent manufacturing image 3D reconstruction effect is better.

The 3D reconstruction time of intelligent manufacturing image is one of the important indicators to verify the 3D reconstruction efficiency of intelligent manufacturing image. The lower the 3D reconstruction time of intelligent manufacturing image, the higher the 3D reconstruction efficiency of intelligent manufacturing image. The specific comparison results are shown in Figure 4.

By analyzing the data in Figure 4, it can be seen that with the increase of the amount of experimental data, the time of 3D reconstruction of intelligent manufacturing images by different methods presents an increasing trend. When the amount of experimental data is 800 MB, the reconstruction methods based on P2M frame improvement, vanishing point optimization, and

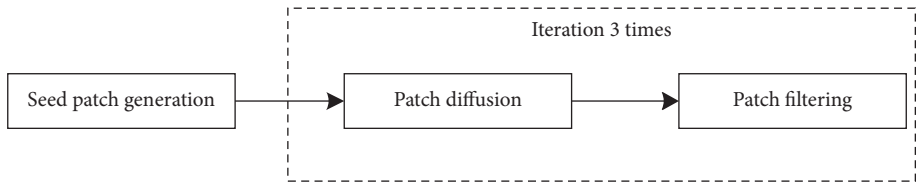


Figure 2: Flow chart of PMVS algorithm.

Table 1: Experimental environmental parameters.

Environmental parameters	Configuration	Parameter
Hardware environment	CPU	Intel(R)Core(TM)i5-9400
	Frequency	2.90 GHz
	RAM	16.0 GB
Software environment	Operating system	Windows 10
	Version	18,362.1082 pro
	Digits	64 bit
	Analog software language	APDL
	Simulation software	MATLAB 7.2

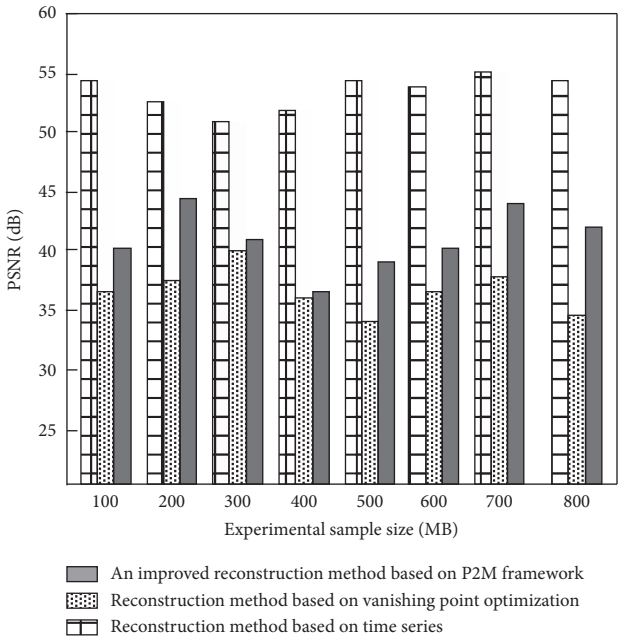


Figure 3: Comparison of PSNR of different methods.

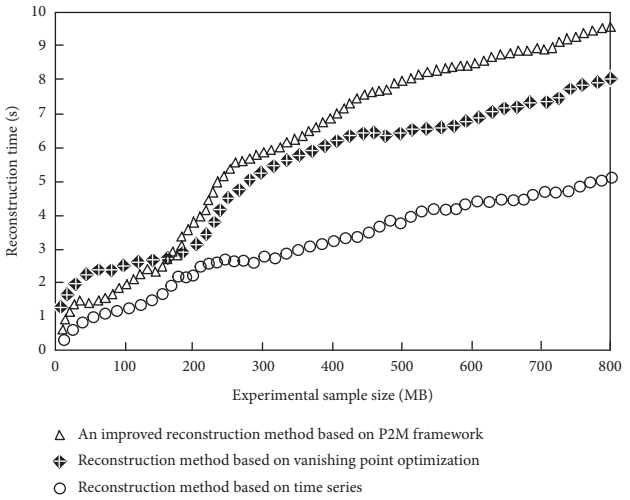
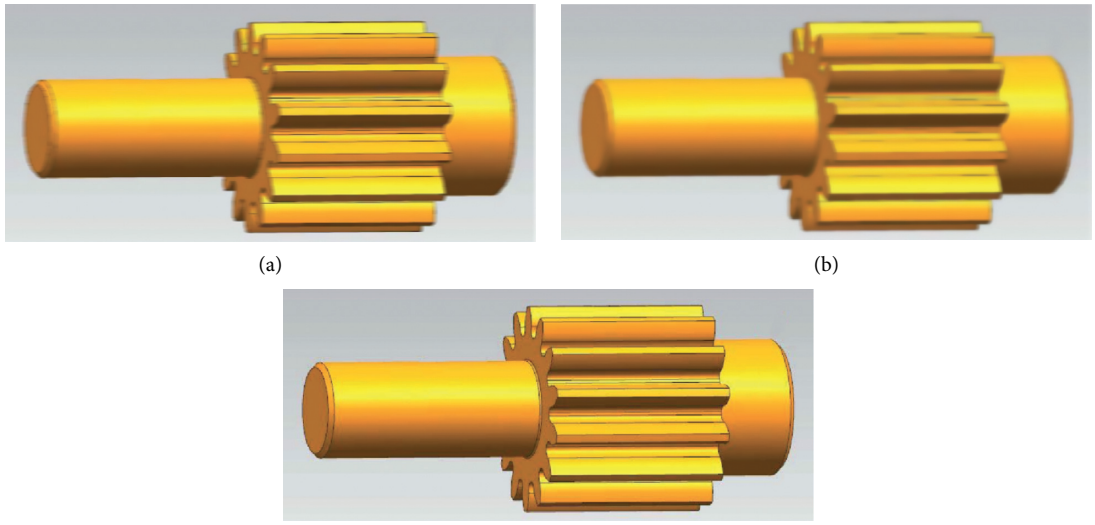


Figure 4: Comparison of reconstruction time between different methods.

time series have reached the maximum time of 3D reconstruction of intelligent manufacturing image. Among them, the maximum time of 3D reconstruction of intelligent manufacturing image based on the improved reconstruction method of P2M frame is 9.3 s, vanishing point optimization is 7.7 s, and time series is 4.9 s. Compared with the reconstruction methods based on P2M frame improvement and vanishing point optimization, the reconstruction time of image 3D reconstruction based on time series always remains at a low level, indicating that this method has a very high efficiency of 3D reconstruction of intelligent manufacturing images.

To further verify the 3D reconstruction effects of intelligent manufacturing images by different methods, two images are taken as examples to verify the 3D reconstruction of intelligent manufacturing images by different methods, and the results are shown in Figures 5 and 6.

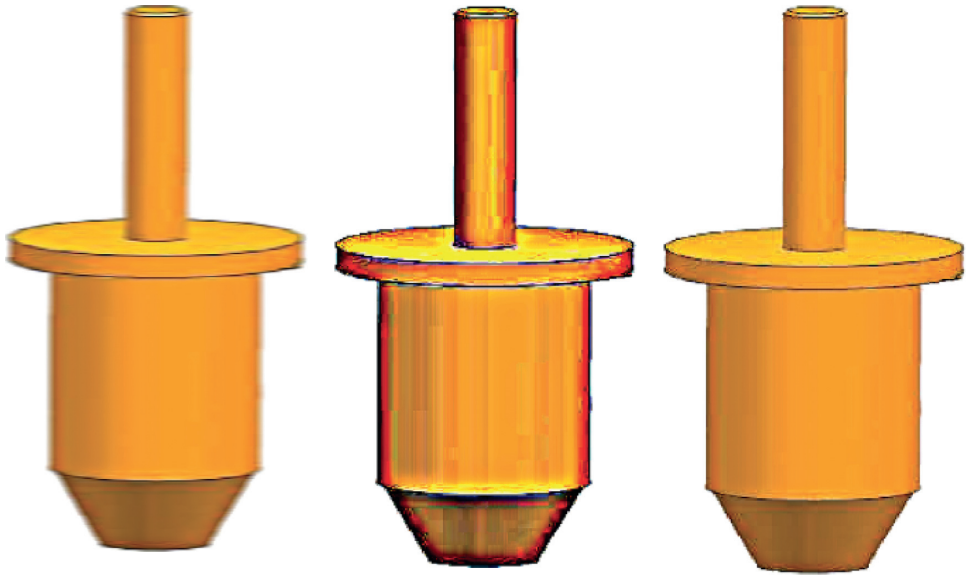
By analyzing the 3D reconstruction results of CAD assisted intelligent manufacturing images in Figures 5 and 6, the reason is that this method uses different point cloud denoising algorithms to filter and denoise the divided area. With the help of CAD, FLANN matching algorithm is used to extract feature points of time-series images and complete image matching. Three-dimensional reconstruction of sparse point cloud and dense point cloud is respectively carried out to complete 3D reconstruction of intelligent manufacturing images.



(a) (b)

(c)

Figure 5: 3D reconstruction effect of image A: (a) improved reconstruction method based on P2M framework, (b) reconstruction method based on vanishing point optimization, and (c) reconstruction methods based on time series.



(a) (b) (c)

Figure 6: 3D reconstruction effect of image B: (a) improved reconstruction method based on P2M framework, (b) reconstruction method based on vanishing point optimization, and (c) reconstruction methods based on time series.

4. Conclusion

As the economic development situation enters the new normal, the manufacturing industry faces increasing external constraints such as resources and environment, and the manufacturing cost of enterprises keeps increasing. At this time, the extensive development path that mainly relies on input of factor resources is difficult to continue. Turn the current situation, structure, methods, promoting development, and it is urgent to form a new economic growth, shape the new international competition advantage focus on manufacturing, and in the process of manufacturing production, image clarity is to ensure that all production can proceed smoothly, the basis of which can accurately control the production progress and the problems existing in the production process, so to improving the quality of industrial production. Research of intelligent image 3D reconstruction method is of great significance, so in this paper, based on the time series of CAD auxiliary intelligent manufacturing image 3D reconstruction method, the experimental results show that the method of image peak signal-to-noise ratio is always above 52 dB, intelligent manufacture image 3 d reconstruction takes a maximum of 4.9 s, and intelligent manufacture image 3D reconstruction effect is better.

Data Availability

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding this work.

Acknowledgments

This work was supported by Bureau Level Heilongjiang Provincial Higher Education Institution Basic Scientific Research Business Expenses Projects) "*Dimension reduction of high dimensional data based on low rank matrix decomposition preserving structure*"(2021-KYYWF-0579).

References

[1] A. K. Bashir, S. Mumtaz, V. G. Menon, and K. F. Tsang, "Guest editorial: cognitive analytics of social media for industrial manufacturing," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 4, pp. 2899–2901, 2021.

[2] D. Zhu, Z. Xu, X. Xu, Q. Zhao, L. Qi, and G. Srivastava, "Cognitive analytics of social media services for edge resource pre-allocation in industrial manufacturing," *IEEE Transactions on Computational Social Systems*, vol. 8, no. 2, pp. 500–511, 2021.

[3] P. A. Bucur, P. Armbrust, and P. Hungerländer, "On the propagation of quality requirements for mechanical assemblies in industrial manufacturing," *Expert Systems with Applications*, vol. 174, no. 15, pp. 114608–114619, 2021.

[4] F. Ji, "Research on the construction of an intelligent manufacturing maturity model based on the three basic paradigms of "digitalization networking and intelligence"," *Modern Management*, vol. 11, no. 7, pp. 761–773, 2021.

[5] D.-D. Li, W.-M. Zhang, Y.-S. Li, F. Xue, and J. Fleischer, "Chatter identification of thin-walled parts for intelligent manufacturing based on multi-signal processing," *Advances in Manufacturing*, vol. 9, no. 1, pp. 22–33, 2021.

[6] Y. F. Zhuang and W. G. Wan, "3D reconstruction of single image based on p2m framework," *Electronic Measurement Technology*, vol. 43, no. 9, pp. 61–64, 2020.

[7] Z. Liu, J. Y. Wang, and L. J. Chen, "3D reconstruction of single image based on vanishing point optimization," *Journal of Zhejiang University of technology*, vol. 47, no. 2, pp. 180–185, 2019.

[8] C. H. Zhao, R. C. Zhao, and S. Feng, "A 3D reconstruction method based on sea ice scene image classification," *Journal of Engineering of Heilongjiang University*, vol. 11, no. 2, pp. 64–73, 2020.

[9] H. Yu, J. M. Yang, and X. G. Wang, "Research on flange 3D reconstruction method based on engineering drawing image," *Mechanical design and manufacturing*, no. 11, pp. 221–223, 2020.

[10] X. Liu, M. Y. Zeng, Y. C. Duan, and G. Fan, "Research on 3D reconstruction method of fuzzy image edge contour based on machine learning," *Electronic testing*, vol. 17, no. 1, pp. 39–41, 2019.

[11] E. Hannink, T. Shannon, H. Dawes, and K. Barker, "The validity of the kinect sensor for the measurement of sagittal spine curvature against the gold standard lateral spinal radiograph," *Physiotherapy*, vol. 107, no. 5, pp. e28–e29, 2020.

[12] Y. M. Lee, S. Lee, K. E. Uhm, G Kurillo, J. J Han, and J Lee, "Upper limb three-dimensional reachable workspace analysis using the kinect sensor in hemiplegic stroke patients: a cross-sectional observational study," *American Journal of Physical Medicine & Rehabilitation*, vol. 99, no. 5, pp. 397–403, 2020.

[13] Y. Tian, G. Wang, L. Li, T. Jin, F. Xi, and G. Yuan, "A universal self-adaption workspace mapping method for human-robot interaction using kinect sensor data," *IEEE Sensors Journal*, vol. 20, no. 99, pp. 1–10, 2020.

[14] C. L. Ramirez, C. Escobedo, J. Edwin, D. Miranda, D. Menotti, and G. Chávez, "A multimodal LIBRAS-UFOP Brazilian sign language dataset of minimal pairs using a microsoft Kinect sensor," *Expert Systems with Applications*, vol. 167, no. 1, pp. 114179–114183, 2020.

[15] T. Guzsvinecz, V. Szucs, and C. Sik-Lanyi, "Suitability of the kinect sensor and leap motion controller—a literature review," *Sensors*, vol. 19, no. 5, pp. 1–25, 2019.

[16] Y. Ma, D. Liu, and L. Cai, "Deep learning-based upper limb functional assessment using a single kinect v2 sensor," *Sensors*, vol. 20, no. 7, pp. 1903–1921, 2020.

[17] T. L. Banh and V. B. Bui, "First experiences with microsoft kinect V2 for 3D modelling of mechanical parts," *Applied Mechanics and Materials*, vol. 889, no. 1, pp. 329–336, 2019.

[18] J. B. Qu, Y. Wang, and Q. Zhao, "Application of DBSCAN clustering and improved bilateral filtering algorithm in point cloud denoising," *Surveying and Mapping Bulletin*, vol. 12, no. 11, pp. 89–92, 2019.

[19] J. Wang, L. I. Huajian, and H. Wang, "Research on tunnel point cloud denoising based on centerline fitting," *Journal of Shenyang Jianzhu University(Natural Science)*, vol. 35, no. 2, pp. 676–683, 2019.

[20] P. H. Casajus, T. Ritschel, and T. Ropinski, "Total denoising: unsupervised learning of 3D point cloud cleaning," *IEEE*, vol. 8, no. 1, pp. 1–11, 2019.

[21] Y. Zhou, R. Chen, Y. Zhao, X. Ai, and G. Zhou, "Point cloud denoising using non-local collaborative projections," *Pattern Recognition*, vol. 120, no. 81, pp. 108128–108139, 2021.

[22] M. A. Irfan and E. Magli, "Exploiting color for graph-based 3D point cloud denoising," *Journal of Visual Communication and Image Representation*, vol. 75, no. 1, pp. 103027–103039, 2021.

[23] W. Hu, X. Gao, G. Cheung, and Z. Guo, "Feature graph learning for 3D point cloud denoising," *IEEE Transactions on Signal Processing*, vol. 68, no. 99, pp. 1–10, 2020.

[24] Y. Guo, S. Guo, K. Guo, and H. l. Zhou, "Seismic data denoising under the morphological component analysis framework by dictionary learning," *International Journal of Earth Sciences*, vol. 110, no. 5, pp. 963–978, 2021.

[25] L. Zhang, L. Han, A. Chang, and J. Fang, "Seismic data denoising using double sparsity dictionary and alternating direction method of multipliers," *Journal of Seismic Exploration*, vol. 29, no. 1, pp. 49–71, 2020.

[26] X. Wang and J. Ma, "Adaptive dictionary learning for blind seismic data denoising," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 7, pp. 1273–1277, 2020.

[27] L. Li, M. Cai, X. Guan, and D. Chu, "Piecewise adaptive-norm trend filtering method for ICESat/GLAS waveform data denoising," *IEEE Access*, vol. 8, no. 1, pp. 168965–168979, 2020.

[28] M. Gu, R. Xie, and L. Xiao, "A novel method for NMR data denoising based on discrete cosine transform and variable

length windows," *Journal of Petroleum Science and Engineering*, vol. 25, no. 1, pp. 1–12, 2021.

[29] Z. Peng, S. Peng, L. Fu et al., "A novel deep learning ensemble model with data denoising for short-term wind speed forecasting," *Energy Conversion and Management*, vol. 207, no. 6, pp. 112524–112535, 2020.

[30] J. W. Zheng and Z. Deli, "Seismic data denoising based on data-driven tight frame dictionary learning method," *Global Geology*, vol. 23, no. 04, pp. 45–50, 2020.

[31] M. Cao, L. Zheng, and X. Liu, "Single view 3D reconstruction based on improved RGB-D image," *IEEE Sensors Journal*, vol. 20, no. 20, pp. 12049–12056, 2020.

[32] S. Croix, P. Deckers, and S. M. Sindbæk, "Recasting a Viking warrior woman from Ribe: 3D digital image reconstruction compared," *Journal of Archaeological Science: Report*, vol. 32, no. 2, pp. 102455–102469, 2020.

[33] C. Liu and G. Liu, "Characterization of pore structure parameters of foam concrete by 3D reconstruction and image analysis," *Construction and Building Materials*, vol. 267, no. 1, pp. 120958–120972, 2020.

[34] M. Shahbazi, P. Ménard, G. Sohn, and J. Théau, "Unmanned aerial image dataset: ready for 3D reconstruction," *Data in Brief*, vol. 25, no. 1, pp. 103962–103975, 2019.

[35] H. X. Wu and Z. Zhang, "3D image reconstruction method based on absolute conic image," *Computer simulation*, vol. 38, no. 8, pp. 203–206, 2021.

[36] S. T. Wan, B. L. Zhang, and T. Yin, "Design of 3D reconstruction software for X-ray nondestructive testing image," *Chinese Journal of Construction Machinery*, vol. 18, no. 5, pp. 425–429, 2020.

Hindawi
Scientific Programming
Volume 2022, Article ID 7340552, 8 pages
https://doi.org/10.1155/2022/7340552

Research Article

Design of Air Passenger Travel Choice Intention Prediction System Based on Deep Learning

Wei Wei and Wang Cheng

Shanghai University of Engineering Science, College of Air Transport, Shanghai 201620, China

Correspondence should be addressed to Wei Wei; 08140002@sues.edu.cn

Received 13 January 2022; Revised 27 January 2022; Accepted 9 February 2022; Published 11 March 2022

Academic Editor: Baiyuan Ding

Under the Beijing-Tianjin regional comprehensive transportation system, the flow of air passengers between multiple airports in the region is more frequent. The fundamental reason for the flow of air passengers is that there are differences in the level of service quality provided by airports and airlines in the region. Passengers' choice intention is the consumption and purchase decision of passengers on aviation services. By constructing a Logit model, this paper analyzes the degree of influence on the travel choice intention of air passengers in the Beijing-Tianjin region from five aspects: individual passenger demographic characteristics, travel purpose, ground transportation characteristics, airport operation capacity, and airport soft power. Passengers can effectively predict the choice of air travel mode in the Beijing-Tianjin region. The results show that Beijing Capital Airport is favored by business travelers; Beijing Daxing International Airport is favored by travelers because of its fast security check-through speed; for Tianjin Binhai International Airport, the convenience of getting in and out of the airport by car and the speed of airport security check-through are two significant factors. Indicators do not affect the selection of airports; reasonable follow-up arrangements when airport flights are delayed are the only significant but negatively correlated factor; the research design results provide new ideas for the analysis of passenger travel mode selection behavior in multiple airport areas, enriching the data-driven research on transportation choices.

1. Related Introduction

Since the birth of the basic theory of artificial intelligence, artificial intelligence technology and applications have developed rapidly, and machine learning is an important form of artificial intelligence. Machine learning has experienced two waves from shallow machine learning to deep learning. With the continuous development of machine theory and applications, various shallow machine learning models have been proposed one after another. Typical machine learning models include support vector machines invented by Cortes and others. At present, the main methods for airport passenger throughput include two aspects, linear prediction and nonlinear prediction. Linear-based methods include time series models, grey models. Although these methods have achieved good prediction results, they cannot reflect nonlinear trends, and the prediction accuracy needs to be improved; nonlinear-based methods include BP neural network, recurrent neural network, long short-term memory network, support vector regression, and other models; in addition, the combined model also achieved good prediction results. This type of model can fit the nonlinear relationship between input and output and has strong fault tolerance. It is also a commonly used research model for airport passenger throughput forecasting. In a word, machine learning has become one of the key research areas of artificial intelligence at present, and it is applied in many fields, including speech processing, computer vision, and natural language processing. Machine learning has also achieved good results in regression prediction. In recent years, domestic and foreign scholars have achieved certain results in the research on the travel choice intention of air passengers. Different from previous studies, this paper not only focuses on the travel choice intention of air passengers but also can more

intuitively reflect the travel choice intention of air passengers after calculation based on the consumption and purchase decision-making of air services by passengers. This paper belongs to a research proposition that considers comprehensive influencing factors. It does not just focus on the travel choice intention of air passengers but also studies the choice intention of passengers through the consumption and purchasing power of air services from the side. The model selected is mainly based on the Logit model. Finally, the feasibility and effectiveness of the model are verified [1–9].

2. Related Theoretical Methods

A deep neural network model framework is built, as shown in Figure 1. First, the transmission parameters of the data are explained, including time step, which indicates how many historical input values are used to predict future values; learning rate, which indicates the learning step size in gradient descent; input dimension, output dimension; batch size; batch number; hidden the number of neurons in the layer, the number of neurons in the GRU layer, and the number of GRU layers. The data are divided into batches before entering the neural network, and each batch is divided into time steps. The data first pass through the FIR filter layer, then enter the input layer, then pass through the hidden layer, then pass through several GRU layers for time series prediction, finally pass through the hidden layer and the output layer to output data, and train the parameters of each part of the network by comparing with the label data and BPTT propagation. During the training process, if the setting of the learning rate is too large, the loss value will oscillate around the local optimal solution. If the setting of the learning rate is too small, the convergence will be slow in the gradient descent process [10–14].

3. Model and Variable Selection

3.1. Model Selection. For the airport selection research of passenger travel, the deep neural network model and the Logit model are the most commonly used methods. The Logit model is the earliest discrete choice model, and it has gradually formed a complete discrete choice model system, such as the probit model, the NL model, and the mixed Logit model.

In this paper, the standard polynomial Logit model is used to study the probability of airport selection for Beijing-Tianjin air passengers. The theoretical basis of the standard polynomial Logit model is the stochastic utility theory and the utility maximization hypothesis. Therefore, the effect function 1.1 is introduced first [15–19].

$$T_{ki} = x_k \beta_i + \varepsilon_{ki}. \tag{1}$$

Among them, T_{ki} is the utility of individual k choosing i scheme (such as choosing Beijing Daxing International Airport), x_k is the set of influencing factors, such as gender, and age, β_i is the generation of influencing factors in choosing the i-th scheme. The estimated parameter vector, ε_{ki}, is the error term.

When $T_{ki} > T_{kj}$ ($i \neq j, i, j \in J$, J represents the set of airports that can be selected), the individual will choose the scheme i. Therefore, the probability formula for individual k to choose plan i is 1.2.

$$P_k(i) = P\left(T_{ki} > T_{kj}\right), \quad (i \neq j). \tag{2}$$

Substituting formulas (1) into (2), and assuming that ε_{ki} obeys the generalized extreme value distribution, formula (3) of the standard polynomial Logit model can be derived, that is, the probability that individual k chooses the i airport. The standard polynomial Logit model needs to assume that one item is selected as the benchmark group (this paper takes Beijing Capital International Airport as the benchmark group). The standard polynomial Logit model is not only fast and stable but also has the assumption that each airport in Beijing and Tianjin is independent and irrelevant, which is conducive to the analysis of each indicator [20].

$$P_{ki} = \frac{e^{\beta x_{ki}}}{\sum_{j \in J}^{J} e^{\beta x_{kj}}}. \tag{3}$$

In expression (3), P_{ki} represents the probability of choosing an airport i for air passenger k; β_x represents the deterministic utility generated when choosing an airport for air passenger travel and using the influencing factors that air passengers experience when choosing an airport. The linear combination of $i, j \in J$, J represents the set of airports that can be selected.

In this paper, Logit regression is performed on each influencing factor through Stata software to explore the significant degree of influence of each factor on the choice of each airport. Then, the multinomial Logit model is solved to obtain the selection probability of each individual for each airport, and then, the selected probability of each airport is summed and averaged to finally obtain the passenger occurrence probability of each airport.

3.2. Data Acquisition and Analysis. This paper extracts the demographic characteristics, travel purpose, ground transportation characteristics, airport operation capacity, airport soft power, and other factors that may affect the passenger's choice behavior at the individual level of passengers and establishes a quantitative analysis model to describe the passenger's choice behavior, as shown in Table 1. By exploring the effect on choice behavior, the selected elements and dimensions contain the relevant intentions at the individual level of air tourist travel.

3.2.1. Demographic Characteristics. Based on network data and past research, it is found that women's travel frequency is generally higher than that of men for shopping, leisure, entertainment, and other reasons, and the travel frequency of young people is generally higher than that of middle-aged and elderly people. It can be seen that demographic characteristics have a certain degree of influence on passengers' travel choices. However, due to the protection of the

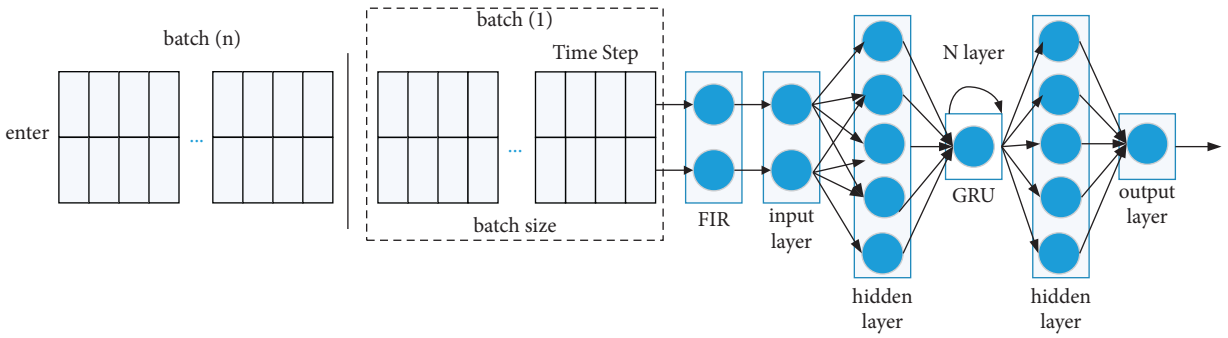


Figure 1: Deep neural network model framework.

Table 1: Variable names, categories, attributes, and definitions.

Variable	Variable category	Variable properties	Definition
Gender	Demographic characteristics	Numeric variables	Dependent variable takes a value of 1 if the gender is male; otherwise, it is 0
Age		Categorical variables	18–29 years old, take the value 1; 30–39 years old, take the value 2; 40–49 years old, take the value 3; 50–59 years old, take the value 4; Over 60 years old, take the value 5
Whether or not to travel for leisure tourism		Numeric variables	For the dependent variable, if the passenger travels for leisure tourism, the value is 1; otherwise, 0
Whether the purpose of travel is to visit relatives and friends	Purpose of travel	Numeric variables	For the passenger's travel purpose of visiting relatives and friends, the value is 1; otherwise, 0
Whether you are traveling for business purposes		Numeric variables	For the passenger, if the passenger travels for business purposes, the value is 1; otherwise, 0
Access to the airport is quick and easy to get in and out of the airport	Ground transportation features	Categorical variables	Great influence, take the value 1; larger influence, take the value 2; general influence, take the value 3; slightly affect, take the value 4; no effect, take the value 5
The airport parking facilities are well built		Categorical variables	Great influence, take the value 1; larger influence, take the value 2; general influence, take the value 3; slightly affect, take the value 4; no effect, take the value 5
Check-in at the airport is easy and fast	Airport operational capacity	Categorical variables	Great influence, take the value 1; larger influence, take the value 2; general influence, take the value 3; slightly affect, take the value 4; no effect, take the value 5
The airport has a high punctuality rate of flights		Categorical variables	Great influence, take the value 1; larger influence, take the value 2; general influence, take the value 3; slightly affect, take the value 4; no effect, take the value 5
Reasonable follow-up arrangements in the event of flight delays at the airport		Categorical variables	Great influence, take the value 1; larger influence, take the value 2; general influence, take the value 3; slightly affect, take the value 4; no effect, take the value 5
Airport security passes quickly		Categorical variables	Great influence, take the value 1; larger influence, take the value 2; general influence, take the value 3; slightly affect, take the value 4; no effect, take the value 5
The airport terminal environment is clean	Airport soft power	Categorical variables	Great influence, take the value 1; larger influence, take the value 2; general influence, take the value 3; slightly affect, take the value 4; no effect, take the value 5
There are many airport-related supporting entertainment facilities		Categorical variables	Great influence, take the value 1; larger influence, take the value 2; general influence, take the value 3; slightly affect, take the value 4; no effect, take the value 5
The airport flight information is prominently located and the information is correct		Categorical variables	Great influence, take the value 1; larger influence, take the value 2; general influence, take the value 3; slightly affect, take the value 4; no effect, take the value 5

individual privacy of respondents, complete demographic characteristics cannot be obtained. Therefore, extracting characteristics such as gender and age helps to explore the demographic characteristics in the Beijing-Tianjin area.

3.2.2. Purpose of Travel. Combined with the economic and social construction conditions in the Beijing-Tianjin region, the purpose of travel also has an impact on travelers' travel choices. Therefore, the vast majority of tourists choose

Beijing's two airports for business and leisure purposes. Therefore, in this paper, the purpose of travel is divided into three categories: leisure travel, visiting relatives and friends, and business office, to help explore the impact of travel purpose in the Beijing-Tianjin region on the travel choice of air passengers [21].

3.2.3. Characteristics of Ground Traffic. With the development of social construction planning, the availability of ground transportation tends to be average, and airport options are basically available. Therefore, this paper focuses on the impact of ground transportation convenience on travelers' travel choices. The convenience of entering and leaving the airport by car and the construction of airport parking lot facilities are used as two indicators to reflect the convenience of ground transportation.

3.2.4. Airport Operation Capability. The impact indicators of airport operation capacity on passengers' choice of an airport are mainly reflected in the convenience of airport procedures, the on-time rate of airport flights, the rationality of airport follow-up arrangements when flights are delayed, and the rate of airport security checks. Excellent airport operation ability can strongly attract passengers to choose. Therefore, this paper incorporates various indicators of airport operation ability into the exploration factors to explore the impact of airport operation ability on air passengers' travel choices in the Beijing-Tianjin region.

3.2.5. Airport Soft Power. Under the condition that the hard demands are basically satisfied, the soft power of the airport is the characteristic competitiveness of the airport, which can improve the passenger stickiness, enhance the travel experience of passengers, and increase the loyalty and frequency of choice. This research uses the airport environment, the number of airport supporting entertainment facilities, and the accuracy and convenience of airport flight information transmission as indicators of airport soft power to explore the impact of airport soft power on air passengers' travel choices in the Beijing-Tianjin region [22, 23].

4. Empirical Analysis

Based on the Beijing-Tianjin regional air passenger travel choice data obtained from the questionnaire, the relevant data sources are mainly obtained through random interviews in the form of Richter scales, and Stata16.0 was used to carry out descriptive statistics, linear relationship analysis, and multiple Logit model solutions for variables. In order to make the running results clearer and more beautiful, this paper encodes the variables in the form of pinyin: gender (xb), age (nl), whether the travel purpose is for leisure travel (xxly), whether the travel purpose is to visit relatives and friends (tqfy), whether the purpose of travel is for business office (swbg), the airport is convenient and fast to travel by car (ccbj), the airport parking lot facilities are well constructed (tcch), the airport check-in procedures are convenient and fast (sxbj), and the airport flight has a high punctuality rate (zdl), reasonable follow-up arrangements when airport flights are delayed (aphl), fast airport security check (ajk), clean and tidy airport terminal environment (hj), more airport-related entertainment facilities (ylss), airport flight information is eye-catching. The information is correct (xxzq). The selection of relevant variables is mainly based on the most common variables of air travel.

4.1. Running Results. In order to better observe the variables, this study carried out descriptive statistics and linear relationship analysis on the variables. The descriptive statistics of the variables are shown in Table 2, where the first column obs represents the observed quantity, the second column mean represents the mean, the third column Std.Dev. represents the standard deviation, the fourth column Min represents the minimum value, and the fifth column Max represents the maximum value.

In order to reduce the deviation of the influence of the respective variables on the dependent variable, this paper also explores whether there is a linear relationship between the respective variables and uses Stata16.0 to test the linear relationship between the variables, as shown in Table 3. It can be found that the VIF values of the respective variables are all less than 3, indicating that there is no significant linear relationship between the respective variables, which proves that the subsequent Logit model regression solution value deviation is small, and the results are efficient.

In this paper, the multinomial Logit model is solved, and the choice of airport for air passenger travel is taken as the dependent variable (select Beijing Capital International Airport, value 1; Beijing Daxing International Airport, value 2; Tianjin Binhai International Airport, value 3), and various indicators of factors such as demographic characteristics, travel purpose, ground traffic characteristics, airport operating capacity, and airport soft power are used as independent variables, and Stata software is used to solve the model to obtain the degree of influence of each factor on each choice, as shown in Tables 4-6.

4.2. Analysis of Results. It can be seen from the results in Tables 4-6. The first column in Table 4 to Table 6 is each factor variable; the second column is the coefficient corresponding to each variable; the positive or negative of the coefficient is logically related to the correlation of each factor, and a positive number indicates a positive correlation; otherwise, it is A negative correlation; the importance of each factor can be obtained by the absolute value of the coefficient; the third column is the standard error of each factor; the degree of dispersion of the sample can be judged by its size; the fourth column is the corresponding value of each variable Z values; the fifth column is the standard error for each factor. This column is the P value corresponding to each variable. Usually, these two values are used to judge the significance of each variable. Under the 90% confidence level, if the absolute value of z is greater than 1.645 and the P-value is less than 0.10, then the variable is selected by passengers.

Table 2: Descriptive statistics for variables.

Variable	Observational measurement	Average value	Standard error	Minimum	Maximum
Gender (xb).	855	0.3883041	0.4876496	0	1
Age (nl).	855	1.762573	0.913599	1	5
Whether to travel for leisure tourism (xxly).	855	0.6584795	0.4744972	0	1
Whether to visit relatives and friends for the purpose of travel (tqfy).	855	0.0467836	0.2112987	0	1
Whether it is for business office purposes (swbg).	855	0.1918129	0.3939571	0	1
Airport selection	855	1.906433	1.496489	1	9
Easy and quick access to the airport (ccbj).	855	2.466667	1.328308	1	5
The airport parking facilities are well built (tcch).	855	3.274854	1.265841	1	5
Check-in at the airport is quick and easy (sxbj).	855	2.378947	1.352083	1	5
Airport flights have a high punctuality rate (zdl).	855	2.100585	1.328551	1	5
Reasonable follow-up arrangements in the event of airport flight delays (aphl).	855	2.269006	1.342918	1	5
Airport security passes quickly (ajk).	855	2.582456	1.2976	1	5
The airport terminal is clean and tidy (hj).	855	2.823392	1.253721	1	5
There are many airport-related supporting facilities (ylss).	855	3.633918	1.212936	1	5
The airport flight information is prominently located and the information is correct (xxzq).	855	2.511111	1.299272	1	5

Table 3: Variable linear relationship test.

Variable	Bright	1/VIF
Whether the purpose of travel is to visit relatives and friends (tqfy).	2.80	0.357124
Whether it is for business office purposes (swbg).	2.67	0.374038
Airport selection	2.48	0.403751
Easy and quick access to the airport (ccbj).	2.44	0.410381
The airport parking facilities are well built (tcch).	2.33	0.429666
Check-in at the airport is quick and easy (sxbj).	2.26	0.442728
Airport flights have a high punctuality rate (zdl).	2.16	0.463237
Reasonable follow-up arrangements in the event of airport flight delays (aphl).	2.15	0.464106
Airport security passes quickly (ajk).	2.00	0.500020
The airport terminal is clean and tidy (hj).	1.62	0.617394
There are many airport-related supporting facilities (ylss).	1.41	0.711014
The airport flight information is prominently located and the information is correct (xxzq).	1.28	0.779130
Whether the purpose of travel is to visit relatives and friends (tqfy).	1.10	0.911822
Whether it is for business office purposes (swbg).	1.09	0.916891
Mean VIF	1.98	

Table 4: Logit model regression results for selecting Beijing Capital International Airport.

Choice 1 Variable	Quasi $R2 = 0.0266$			
	Regression coefficients	Standard error	With	p
Gender (xb).	−.2165614	0.1562371	−1.39	0.166
Age (nl).	0.0859551	0.0847011	1.01	0.310
Whether to travel for leisure tourism (xxly).	0.0969474	0.2420128	0.40	0.689
Whether the purpose of travel is to visit relatives and friends (tqfy).	0.1350224	0.3990109	0.34	0.735
Whether it is for business office purposes (swbg).	0.8002119	0.2929658	2.73	0.006
Easy and quick access to the airport (ccbj).	0.1194681	0.0838746	1.42	0.154
The airport parking facilities are well built (tcch).	−0.0071703	0.0733454	−0.10	0.922
Check-in at the airport is quick and easy (sxbj).	0.0790751	0.0912878	0.87	0.386
Airport flights have a high punctuality rate (zdl).	−0.0826886	0.0863616	−0.96	0.338
Reasonable follow-up arrangements in the event of airport flight delays (aphl).	0.1619173	0.0805046	2.01	0.044
Airport security passes quickly (ajk).	−0.2033971	0.0831281	−2.45	0.014
The airport terminal is clean and tidy (hj).	−0.0440311	0.082426	−0.53	0.593
There are many airport-related supporting facilities (ylss).	0.0193179	0.0682984	0.28	0.777
The airport flight information is prominently located and the information is correct (xxzq).	−0.0848699	0.0855178	−0.99	0.321

Table 5: Logit model regression results for selecting Beijing Daxing International Airport.

Choice 2	Quasi $R2 = 0.00775$			
	Regression coefficients	Standard error	With	p
Gender (xb).	0.0422496	0.3583026	0.12	0.906
Age (nl).	−0.6052624	0.2657307	−2.28	0.023
Whether to travel for leisure tourism (xxly).	−0.1346333	0.495872	−0.27	0.786
Whether the purpose of travel is to visit relatives and friends (tqfy).	0.0549319	0.7601636	0.07	0.942
Whether it is for business office purposes (swbg).	0.0565374	0.586965	0.10	0.923
Easy and quick access to the airport (ccbj).	−0.1697648	0.1847487	−0.92	0.358
The airport parking facilities are well built (tcch).	0.097763	0.1638026	0.60	0.551
Check-in at the airport is quick and easy (sxbj).	−0.0899377	0.2017813	−0.45	0.656
Airport flights have a high punctuality rate (zdl).	0.1975662	0.1815875	1.09	0.277
Reasonable follow-up arrangements in the event of airport flight delays (aphl).	−0.0473564	0.1714052	−0.28	0.782
Airport security passes quickly (ajk).	0.5481134	0.184276	2.97	0.003
The airport terminal is clean and tidy (hj).	−0.2903814	0.1798466	−1.61	0.106
There are many airport-related supporting facilities (ylss).	0.1305016	0.1604658	0.81	0.416
The airport flight information is prominently located and the information is correct (xxzq).	−20.772141	0.1917859	−1.00	0.320

Table 6: Logit model regression results for selecting Tianjin Binhai International Airport.

Choice 3	Quasi $R2 = 0.0255$			
	Regression coefficients	Standard error	With	p
Gender (xb).	0.0496101	0.1994063	0.25	0.804
Age (nl).	−0.0250308	−0.1079822	−0.23	0.817
Whether to travel for leisure tourism (xxly).	0.1335351	0.3157477	0.42	0.672
Whether the purpose of travel is to visit relatives and friends (tqfy).	0.5174501	0.4818168	1.07	0.283
Whether it is for business office purposes(swbg).	−0.4837559	0.3899762	−1.24	0.215
Easy and quick access to the airport (ccbj).	−0.2100871	0.107067	−1.96	0.050
The airport parking facilities are well built (tcch).	0.0822015	0.0915746	0.90	0.369
Check-in at the airport is quick and easy (sxbj).	0.020125	0.1152669	0.17	0.861
Airport flights have a high punctuality rate (zdl).	0.0813616	0.1112021	0.73	0.464
Reasonable follow-up arrangements in the event of airport flight delays (aphl).	−0.1190637	0.1030323	−1.16	0.248
Airport security passes quickly (ajk).	−0.2009067	0.1059133	−1.90	0.058
The airport terminal is clean and tidy (hj).	0.0614114	0.104984	0.58	0.559
There are many airport-related supporting facilities (ylss).	−0.0119767	0.0845989	−0.14	0.887
The airport flight information is prominently located and the information is correct (xxzq).	0.1267582	0.1076092	1.18	0.239

4.2.1. Beijing Capital International Airport. As shown in Table 4, the Z-values of the three indicators are more than 1.645, and the P-values are all less than 0.10, indicating that these three indicators are important for the influence of Beijing-Tianjin Airlines passengers, and choosing Beijing Capital International Airport is significant. From the perspective of travel purpose, it can be seen that most of the passengers who choose Beijing Capital International Airport are for business and office travel purposes. From the perspective of airport operation capability, Beijing-Tianjin Airlines passengers pay great attention to the rationality of the follow-up arrangements for flight delays at Beijing Capital International Airport and the speed of security inspection. These two items will directly affect whether Beijing-Tianjin Airlines passengers choose Beijing Capital International Airport.

4.2.2. Beijing Daxing International Airport. As shown in Table 5, the absolute values of the Z-values of the two indicators of age and the speed of passing through the airport security check are both greater than 1.645, and the P-values are both less than 0.10, indicating that these two indicators have a high level of significance. However, the corresponding coefficient of age is −0.605, indicating that age is negatively correlated with the choice of Beijing Daxing International Airport by air passengers, which reflects that age has no profound influence on the choice of Beijing Daxing International Airport. In terms of airport operation capability, only the speed of passing the airport security check has a significant impact, which reflects the degree to which the speed of passing the airport security check has an impact on the travel choices of Beijing-Tianjin air passengers.

4.2.3. Tianjin Binhai International Airport. According to Table 6, the absolute value of the Z value of the two indicators of the convenience and speed of entering and leaving the airport and the speed of airport security inspection is greater than 1.645, and the P-value is less than 0.10, indicating that the two indicators have a high level of significance. The coefficient corresponding to the convenience of entering and leaving the airport by car is −0.210, and the corresponding coefficient of the airport security check passing speed is −0.200, both of which are negative numbers, indicating that these two indicators have a negative correlation with the selection of Tianjin Binhai International Airport. Tianjin Airlines passengers have little influence on the choice of Tianjin Binhai International Airport.

5. Conclusion

According to the analysis of the results in chapter 4.2, it can be seen that the most important factor when choosing an airport is the airport's operational capability, especially the speed of passing through the airport security check. Efficiency needs. Therefore, the following conclusions can be drawn: (1) The airport is appropriately expanded, and the security check channel is increased. (2) Artificial intelligence to speed up the identification of dangerous goods is introduced, and the quality and efficiency of security inspections are improved. (3) Green passages are built to allow passengers without security inspection requirements to pass quickly. In this way, the travel time cost of Beijing-Tianjin Airlines passengers can be reduced, and the travel efficiency of passengers can be improved. According to the results of the model solution, it can also be concluded that the current regional airports with less passenger distribution have less influencing factors than the current Beijing-Tianjin hub airports, which indicates that if the regional airports want to attract more. There are many Beijing-Tianjin Airlines passengers while enhancing the airport's operational strength, and they need to focus on the corresponding soft power and combine regional characteristics and passenger characteristics to implement special services to increase the travel experience of passengers, so as to improve the Beijing-Tianjin Airlines passengers' interest in the airport. As for Beijing Capital International Airport, which is currently saturated with passengers, the first thing to consider is the passenger diversion strategy, and it will be completed jointly with Beijing Daxing International Airport and other airports in the two places. This requires the government to strengthen the air space between Beijing and Tianjin. The railway combined transport mode makes the transportation between the three places more accessible and convenient, reduces the gap between regions, and eliminates the inertial thinking of Beijing-Tianjin air passengers on previous travel choices, so as to make the allocation of air passengers between Beijing and Tianjin more convenient and reasonably average.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] R. Buehler, "Determinants of transport mode choice: a comparison of Germany and the USA," *Journal of Transport Geography*, vol. 19, no. 4, pp. 644–657, 2011.

[2] E. Cascetta, A. Papola, F. Pagliara, and V. Marzano, "Analysis of mobility impacts of the high speed Rome-Naples rail link using withinday dynamic mode service choice models," *Journal of Transport Geography*, vol. 19, no. 4, pp. 635–643, 2011.

[3] D. Albalate, G. Bel, and X. Fageda, "Competition and cooperation between high-speed rail and air transportation services in Europe," *Journal of Transport Geography*, vol. 42, pp. 166–174, 2015.

[4] R. R. Clewlow, J. M. Sussman, and H. Balakrishnan, "The impact of high-speed rail and low-cost carriers on European air passenger traffic," *Transport Policy*, vol. 33, no. 2, pp. 136–143, 2014.

[5] B. Baker, O. Gupta, N. Naik, and R. Raskar, "Designing neural network architectures using reinforcement learning," 2016, https://arxiv.org/abs/1611.02167.

[6] Z. Xiao, X. Li, L. Wang, and Q. Yang, "Using convolution control block for Chinese sentiment analysis," *Journal of Parallel and Distributed Computing*, vol. 116, pp. 18–26, 2017.

[7] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proceedings of the Advances in Neural Information Processing Systems*, pp. 1097–1105, New York, NY, USA, December 2012.

[8] P. Sermanet, K. Kavukcuoglu, and S. Chintala, "Pedestrian detection with unsupervised multi-stage feature learning," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3626–3633, New York, NY, USA, December 2013.

[9] D. Ciresan, U. Meier, J. Masci, and J. Schmidhuber, "Multi-column deep neural network for traffic sign classification," *Neural Networks*, vol. 32, pp. 333–338, 2012.

[10] V. Mnih, A. P. Badia, M. Mirza, A. Graves, and T. Harley, "Asynchronous methods for deep reinforcement learning," in *Proceedings of the International Conference on Machine Learning*, pp. 1928–1937, New York, NY, USA, June 2016.

[11] A. E. Sallab, M. Abdou, E. Perot, and S. Yogamani, "Deep reinforcement learning framework for autonomous driving," *Electronic Imaging*, vol. 2017, no. 19, pp. 70–76, 2017.

[12] Y. Duan, X. Chen, R. Houthooft, J. Schulman, and P. Abbeel, "Benchmarking deep reinforcement learning for continuous control," in *Proceedings of the International Conference on Machine Learning*, pp. 1329–1338, New York, NY, USA, June 2016.

[13] A. E. Sallab, M. Abdou, E. Perot, and S. Yogamani, "End-to-end deep reinforcement learning for lane keeping assist," 2016, https://arxiv.org/abs/1612.04340.

[14] V. Mnih, K. Kavukcuoglu, D. Silver et al., "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[15] M. Hausknecht and P. Stone, "Deep recurrent q-learning for partially observable mdps," *AAAI Fall Symposium Series*, vol. 90, pp. 18–26, 2015.

[16] H. Van, A. Guez, and D. Silver, "Deep reinforcement learning with double qlearning," in *Proceedings of the 13th AAAI*

conference on artificial intelligence, Phoenix, AZ, USA, February 2016.

[17] Z. Wang, T. Schaul, and M. Hessel, "Dueling network architectures for deep reinforcement learning," in *Proceedings of the 33rd International Conference on Machine Learning*, New York, NY, USA, June 2016.

[18] T. Schaul, J. Quan, I. Antonoglou, and D. Silver, "Prioritized experience replay," *Computer Science*, vol. 86, pp. 23–25, 2015.

[19] T. Hester, M. Vecerik, O. Pietquin et al., "Learning from demonstrations for real world reinforcement learning," 2017, https://arxiv.org/abs/1704.03732.

[20] D. Horgan, J. Quan, D. Budden et al., "Distributed prioritized experience replay," 2018, https://arxiv.org/abs/1803.00933.

[21] G. Farquhar, T. Rocktäschel, M. Igl, and S. Whiteson, "Treeqn and atreec: differentiable tree-structured models for deep reinforcement learning," 2017, https://arxiv.org/abs/1710.11417.

[22] X. Zhu and D. Dunson, "Stochastic lipschitz Q-learning," 2019, https://arxiv.org/abs/1904.10653.

[23] M. G. Bellemare, W. Dabney, and R. Munos, "A distributional perspective on reinforcement learning," in *Proceedings of the 34th International Conference on Machine Learning*, pp. 449–458, New York, NY, USA, June 2017.

Hindawi
Scientific Programming
Volume 2022, Article ID 5663824, 10 pages
https://doi.org/10.1155/2022/5663824

Research Article

Performance Optimization of CNC Machine Tool System Based on Sensor Data

Yanfang Wu, Ning Yue, and Kangkang Qian

Zibo Technical College, Zibo, Shandong 255000, China

Correspondence should be addressed to Yanfang Wu; huailaochengnxx@163.com

Received 29 December 2021; Revised 22 January 2022; Accepted 6 February 2022; Published 10 March 2022

Academic Editor: Baiyuan Ding

CNC machine tools have been popularized in the development of the manufacturing industry because of their high precision, high speed, high efficiency, and safe and reliable processing. The numerical control system and the measurement system of the machine tool are the key components of the modern numerical control machine tool, especially the measurement system of the machine tool, which is a prerequisite to ensure the high precision of the machine tool. The sensor is an important part of the measurement system. Sensors are devices that can sense the measurement objects within the measurement range and convert them into output signals and information according to certain rules. They are mainly used to detect the operating objects, system status, and operating environment status of the machine tool system and effectively control the normal operation of the machine tool system. Accurate and reliable information is provided. This study takes the CNC machine tool control system as the research object and studies the optimization measures of the sensor data to its performance. First, a model of the CNC machine tool control system was built based on MATLAB/Simulink, and then, based on this model, an open-loop (without sensor) and closed-loop (with speed and position sensors) control systems were built and were finally verified by simulation. The research results show that when there are sensors, the CNC machine tool control system has better stability and robustness.

1. Introduction

The sensor is a measuring device that converts the measurement to a certain physical quantity with a certain degree of accuracy and is convenient for application. The input quantity of a commonly used sensor is a certain measured quantity, which may be a physical quantity and can be a physical quantity of gas, light, or electricity, mainly electrophysical quantity. There are many types of sensors [1–7].

The requirements of CNC machine tools for sensors include (1) high anti-interference and strong reliability; (2) meeting the requirements of high precision and high speed; (3) convenient use and maintenance, suitable for the operating environment of the machine tool; and (4) low price and cost low. In a broad sense, the servo system refers to an automatic control system whose output can follow the change in the input with certain accuracy. Its control is also a motion control with high precision and high speed as the goal. It mainly includes a spindle servo drive system and a feed servo system. [8, 9]. Figure 1 shows the CNC machine tool system.

The machine tool industry is the industry that uses the most servo products, so the development of the machine tool industry is closely related to the development of servo technology. The servo system used in CNC machine tools should meet the basic requirements of good stability, high positioning accuracy, fast response without overshoot, large torque, and wide speed range. In order to meet these requirements, the detection devices related to the servo system must also have high accuracy and precision. The development of the servo system is closely connected with the speed regulation theory of AC motors and the development of power electronic technology, micromotor technology, microcomputer technology, sensor technology, etc. and has experienced the development of open loop, hydraulic, electrical, DC motor servo, etc. Later, the current electrical servo system driven by AC motors has become the mainstream of the development of modern high-performance

Figure 1: CNC machine tool system.

servo systems, and the system is mainly composed of semiclosed loop, fully closed loop, and other servo control forms [10–15].

The AC servo system used in CNC machine tools generally includes servo motors, drives (power amplifiers), motion controllers, and some feedback detection devices. Broadly speaking, they should also include transmission devices and work tables. The driver mainly completes the control of the motor speed and torque (current), which is the core of the servo system; the motion controller completes the more complex position motion control. Servo motors used in AC servo systems include synchronous motors and asynchronous motors. Permanent magnet synchronous motors have the characteristics of small size, high efficiency, excellent low-speed performance, wide speed range, high efficiency, and easy realization of field weakening control, etc., which can meet the requirements of high-performance control such as CNC machine tools, and with the performance of magnetic materials constantly increasing, declining prices, and easy control, etc., permanent magnet synchronous motors are more and more widely used [16–23].

According to the application and the different control performance requirements, the structure of the CNC servo system includes an open-loop position servo system, a semiclosed-loop position servo system, and a fully closed-loop position servo system.

(1) The position servo system for an open loop is a kind of servo system used in early CNC machine tools without position feedback. Due to the low positioning accuracy and large transmission error, it is generally suitable for simple CNC systems that do not require high control accuracy and low movement speed.

(2) The semiclosed-loop position servo system has the functions of position detection and feedback, but this position detection uses the displacement value integrated by the measured speed of the pulse encoder coaxially connected to the servo motor as the position feedback signal. In the case of bars and other transmission links, the displacement value measured in this way may have a certain deviation from the final table position, so it can only be called a semiclosed loop. With the improvement of manufacturing level, technology, and processing accuracy, the quality of the functional transmission components that constitute the servo system is getting better and better, which greatly improves the error of the transmission part, and the system can also use backlash compensation and pitch error compensation. The servo system of this structure also occupies a large proportion of the market share in the current mid-range CNC machine tools [24–28].

(3) The fully closed-loop position servo system is directly installed on the worktable of the machine tool using high-precision measuring devices such as grating rulers. The precise displacement of the actual movement of the machine tool worktable can be obtained, so the precise position information of the worktable movement can be returned. In the above, this kind of control is the most ideal, with the smallest error. However, because the mechanical transmission chain of the machine tool itself is included in the position of the closed loop, the gap of the transmission part of the system, the nonlinearity of the friction characteristic, and the rigidity of the transmission will cause the system to be unstable, causing the system to resonate and crawl at low speed, which also gives control to the system. The tuning of system parameters brings great difficulties. The basic servo system is shown in Figure 1.

Most of the servo system manufacturers at present provide supporting servo motors and servo drives for different users to choose and use. The supporting drive equipment makes the drive parameter setting more convenient and reliable. Generally speaking, most of the system parameters can be set using the initial parameters provided by the drive manufacturer, and some parameters that have a greater impact on the control performance of the machine tool and are related to the load, such as position loop gain value, speed gain value, compound position feedforward coefficients, and other parameters that affect the accuracy of the machine tool need to be manually adjusted and set by professional engineers with experience [29]. However, the control system of the CNC machine tools is still a problem and is to be covered in detail.

Reasonable setting of the servo system parameters is an important measure to improve the machining performance and work stability and reliability of the machine tool. For the same type of machine tools, the gap in processing performance between China and foreign countries is a very important aspect of the fact that the domestic optimization of

servo parameters is not enough or the parameters are too conservative in the setting. Of course, stability and reliability are the prerequisites, but on the basis of improving machine tool manufacturing performance, we should try to set more reasonable parameters. Due to the wide variety of CNC machine tools produced by the machine tool factory, and the manufacturing process and load conditions of the same model of machine tools may be different, this is a serious problem for the staff of servo optimization. In order to be helpful to the debugging personnel of the machine tool, this study will study the debugging process of some commonly used servo parameters and put forward some ideas and measures to improve the optimization performance of the servo. Specifically for the FANUC CNC servo system used on a horizontal machining center, with the help of its servo adjustment software Servo Guide and the machine tool circular error measurement suite Ballbar (ballbar) produced by Renishaw, some errors of the machine tool are measured and analyzed. The performance changes such as the accuracy of the machine tool after the adjustment of the parameters that have a greater impact are expected to provide some ideas for the parameter setting of the CNC machine tool servo system and the improvement of the machine tool performance.

The feed servo system is the intermediate connection part of the CNC numerical control device and the mechanical transmission part. Its performance directly affects the tracking and positioning accuracy of the servo axis movement, the surface quality of the processing, the productivity, and the working reliability, and other technical indicators, so the servo system is improved. Reliability and motion control performance are of great significance to CNC machine tools.

2. Mathematical Model of CNC Machine Tool Control System

The performance of the permanent magnet synchronous motor servo system is closely related to the structure and performance of the motor itself and the selected control strategy called direct torque control. As the motion execution unit of the servo system, the AC permanent magnet synchronous servo motor itself has many advantages and can well meet the performance requirements of the servo system, so it has become a motor widely used in the servo system of the high-precision transmission field. Starting from the basic structure of the permanent magnet synchronous motor, this study establishes the dynamic mathematical model of the motor in each analysis coordinate system, which lays the foundation for the subsequent analysis of different motor control strategies and methods.

What needs to be emphasized here is that before the mathematical modeling of the AC motor, the equations satisfied by the three-phase composite flux linkage of the AC motor must be first solved. The total flux linkage of the turns of each phase winding of the AC motor is mainly the flux linkage generated by the armature magnetic field and the flux linkage generated by the permanent magnet magnetic field. Therefore, we have obtained the flux linkage equation of the AC motor winding, as shown in formula (11):

$$\begin{cases} \psi_{ma} = \psi_{m0} + \psi_m \cos(\theta_e) \\ \psi_{mb} = \psi_{m0} + \psi_m \cos(\theta_e - 120^\circ), \\ \psi_{mc} = \psi_{m0} + \psi_m \cos(\theta_e + 120^\circ) \end{cases} \tag{1}$$

where ψm0 is the DC component of the permanent magnet flux linkage, ψm is the fundamental wave amplitude of the permanent magnet flux linkage, and θe is the position angle of the AC motor mover.

In the actual modeling process, some specific mathematical transformations, such as Parker transformation, are usually used to transform the coordinate system of each physical quantity of the AC motor:

$$P = \frac{2}{3}\begin{bmatrix} \cos\theta_e & \cos(\theta_e - 2\pi/3) & \cos(\theta_e + 2\pi/3) \\ -\sin\theta_e & -\sin(\theta_e - 2\pi/3) & -\sin(\theta_e + 2\pi/3) \\ 1/2 & 1/2 & 1/2 \end{bmatrix}. \tag{2}$$

Therefore, after transforming the formula (1), the following form can be obtained as follows:

$$\begin{bmatrix} \psi_{m\,d} \\ \psi_{mq} \\ \psi_{m0} \end{bmatrix} = p\begin{bmatrix} \psi_{ma} \\ \psi_{mb} \\ \psi_{mc} \end{bmatrix}. \tag{3}$$

In the formula, ψmd is expressed as the d-axis permanent magnet flux linkage, ψmq is expressed as the q-axis permanent magnet flux linkage, and ψm0 is expressed as the 0-axis permanent magnet flux linkage.

Substituting formulas (1) and (2) into formula (3), the following relationship can be obtained as follows:

$$\begin{cases} \psi_{m\,d} = \psi_m, \\ \psi_{mq} = 0, \\ \psi_{m0} = \psi_0. \end{cases} \tag{4}$$

In addition, each element in the inductance matrix of the AC motor and the specific expression of the permanent magnet flux linkage of the AC motor are obtained, and finally, the flux linkage equation of the AC motor can be obtained. The next two steps start with the elements of the inductance matrix. Generally speaking, the specific expressions of the self-inductance and mutual inductance of an AC motor are as follows:

$$\begin{cases} L_{aa} = L_0 + L_m \cos(2P_r\theta_r) \\ L_{bb} = L_0 + L_m \cos(2P_r\theta + 120^\circ), \\ L_{cc} = L_0 + L_m \cos(2P_r\theta - 120^\circ) \end{cases} \tag{5}$$

$$\begin{cases} M_{ab} = M_0 - M_m \cos(2P_r\theta_r - 120^\circ) \\ M_{bc} = M_0 - M_m \cos(2P_r\theta) \\ M_{ca} = M_0 - M_m \cos(2P_r\theta + 120^\circ). \end{cases} \tag{6}$$

What needs to be explained here is that the torque discussed here is derived from the electromagnetic torque equation of the AC motor under the condition of constant amplitude. At the same time, the frequency and phase angle of the current of the AC motor also remain constant, so the average electromagnetic torque in the steady-state expression is as follows:

$$T_{pm} = \frac{P_{em}}{\omega_r} = T_{pm} + T_r. \tag{7}$$

In the formula, Pem is expressed as the power of the AC motor, and Tpm is expressed as the permanent magnet torque generated by the permanent magnet flux linkage of the AC motor and the armature current, which satisfies the following:

$$T_{pm} = \frac{e_{ma}i_a + e_{mb}i_b + e_{mc}i_c}{\omega_r}. \tag{8}$$

If the above formula is subjected to a certain mathematical transformation, then we can get the following:

$$T_{pm} = \frac{3}{2}\frac{E_m I_m \cos\beta}{\omega_r} = \frac{3}{2}P_r\psi_m I_m \cos\beta. \tag{9}$$

Figure 2 is the control vector principle diagram of the CNC machine tool control system when direct torque control is used. As can be seen, both the static and dynamic coordinate systems exist when modeling the control system since both the static and dynamic parts contain. The coordinate system is the AC motor stator two-phase static coordinate system. When it needs to be explained here, the selected axial and stator A1-phase windings here, the axis directions of the D and q coordinate systems are the same; the rotating coordinate system of the D and q coordinate systems is considered to be fixed on the rotor side of the AC motor, the axial direction of the stator is the positive direction of the D axis, and the angle between the D axis and A1 winding is θe, us and is are each the stator voltage and current vector of the AC motor, and ψs and ψm are the stator armature flux and stator permanent flux. If the stator resistance is ignored, δ is the torque angle of the motor. In direct torque control, the stator flux is integrated with the stator voltage. The torque is estimated based on the inner product of the estimated stator flux vector and the measured current vector. The magnetic flux and torque will be compared with the reference value. If the error between the magnetic flux or torque and the reference value exceeds the allowable value, the power crystal in the inverter will switch so that the error of the magnetic flux or torque can be reduced as soon as possible.

It can be seen from the figure that the torque equation of AC motor can be expressed as follows:

$$T_{em} = \frac{3P_r|\psi_s|}{4L_dL_q}\left[2\psi_m L_q \sin\delta - |\psi_s|(L_q - L_d)\sin 2\delta\right]. \tag{10}$$

Since the Ld and Lq of the AC motor are approximately equal, the formula (10) can be written as follows:

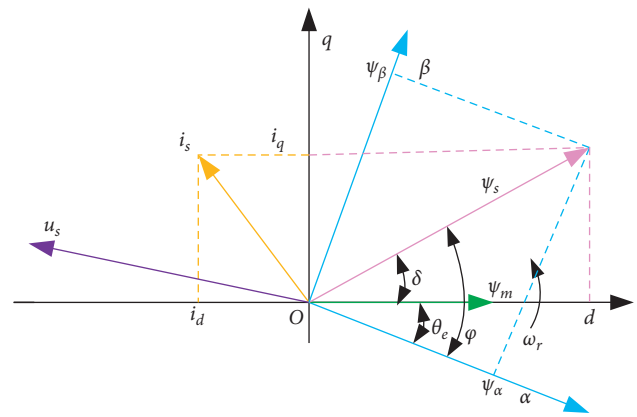


Figure 2: Vector diagram of direct torque control of AC motor.

$$T_{em} = \frac{3}{4L_s}p_r|\psi_s|\psi_m \sin\delta. \tag{11}$$

Among them, δ is expressed as the angle formed between the stator armature flux linkage of the AC motor and the corresponding permanent magnet flux linkage. However, because the ψm obtained by the permanent magnets of the AC motor stator is a constant amount, we can adjust the stator armature flux, the value, and the included angle δ between the stator armature flux linkage of the AC motor and the stator permanent magnet flux linkage, so as to achieve the purpose of controlling the electromagnetic torque Tem of the AC motor system. Therefore, direct torque control can also be regarded as a hysteresis or relay control.

For AC motors, the corresponding stator flux linkage vector can be expressed by formula (11):

$$\psi_s = \int (u_s - R_s i_s)dt. \tag{12}$$

Since in the control circuit, when the control switch is turned on or off, each voltage vector is a constant quantity, so the formula (12) can also be expressed as follows:

$$\psi_s = u_s t - R_s \int i_s dt + \psi_{s|t=0}. \tag{13}$$

Among them, ψ_s is expressed as the initial armature flux linkage vector of the AC motor stator. Since the AC motor uses a permanent magnet motor, there is also a flux linkage when the motor is not energized. That is to say, what needs to be noted here is that these two quantities have not only magnitude but also direction, Therefore, it is necessary to know the initial position of the AC motor rotor during actual operation. The stability of the AC control system control will depend to a certain extent on the judgment of the rotor position.

If the stator resistance is neglected, then according to formula (12), it can be known that the stator flux linkage of an AC motor can be directly expressed as integrating the voltage space vector:

$$\psi_s = \int u_s dt. \tag{14}$$

Therefore, we can achieve the control of the flux linkage by appropriately selecting the space voltage vector so that its trajectory is close to a perfect circle.

In an AC control system, when the angle between the applied voltage vector and the current flux linkage vector of the AC motor is <90°, the corresponding flux linkage amplitude will accordingly increase. When the angle between the applied voltage vector and the current flux linkage vector of the AC motor is >90°, the corresponding flux linkage amplitude will correspondingly decrease due to the effect of the vector.

Therefore, in order to facilitate our selection of voltage vector, we can divide the vector plane into six regions as shown in Figure 3. For example, when the stator flux linkage runs in the I area and rotates counterclockwise, we can choose the voltage space V6 to increase the amplitude of the flux linkage or select the voltage space V5 to reduce the amplitude of the flux linkage. In the same way, if the stator flux linkage of the AC motor runs clockwise, we can choose V3 to increase the flux linkage amplitude or select V2 to decrease the flux linkage amplitude. Therefore, we can choose an appropriate space vector to control the amplitude of the AC motor's stator flux linkage so that its amplitude is basically constant.

Direct torque control (DTC) is a way for the inverter to control the torque of a three-phase motor. The method is to calculate the estimated value of the motor magnetic flux and torque according to the measured motor voltage and current, and after controlling the torque, the speed of the motor can also be controlled.

3. Optimization of CNC Machine Tool System Based on Speed Sensor and Position Sensor

In order to ensure that the machine tool reaches the best state when it leaves the factory to obtain the best steady-state performance and dynamic performance, after the servo drive completes the basic configuration of the drive and the motor, further control parameters need to be adjusted for the system with load to make the electrical parameters that are matched with the mechanical structure to improve the accuracy of the system speed and position control, which is called the optimization of the servo system.

The optimization and adjustment of controller parameters should not only follow the general principles of controller parameter settings but also combine some experience in actual debugging. Different types of machine tools (or even the same type of machine tools) often have different settings according to the load of the servo system. The value of this is indeed a complicated matter for the engineers who optimize and debug the machine tool. This study specifically combines the fully closed-loop permanent magnet synchronous motor feed position servo system used in a horizontal machining center to study the servo optimization method. The machining center uses FANUC permanent magnet synchronous servo motor and matching servo drives. Combining some basic theories of servo system design, a large number of experiments are carried out on the optimization of servo system parameters, and certain optimization conclusions and methods are analyzed and obtained in order to help the machine tool factory's servo optimization and debugging work.

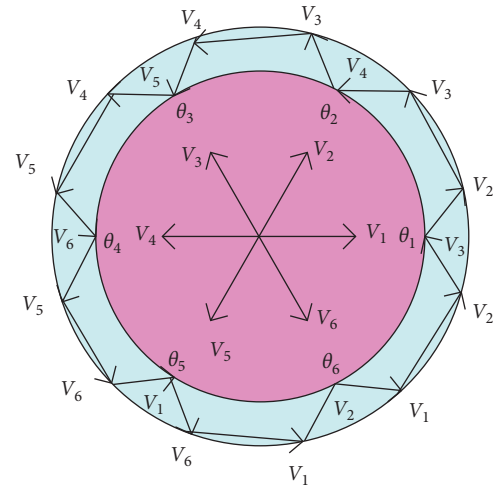


Figure 3: Plane control diagram of voltage vector.

The closed-loop servo system usually includes three feedback loops: position loop (outer loop), speed loop (middle loop), and current loop (inner loop). The general structure diagram is shown in Figure 4. As can be seen, using the closed loops can help to improve the precision, which will be further discussed in the following. For a position servo system controlled by three loops, the response speed of the inner loop current loop should be higher than the response speed of the outer loop. In order to avoid the vibration and poor response of the servo system, the general servo optimization process is to first optimize the current loop, then optimize the speed loop, and finally adjust the position loop parameters.

3.1. Current Loop. The main function of the current loop is to play a timely antiwinding effect against the fluctuation of the grid voltage, and second, it can limit the armature current when the motor is locked and play a role in automatically and quickly protecting the motor. The AC servo system achieves good control of motor torque and speed through high-performance control of the current link. The control parameters of the current link are related to the used inverter power electronic devices, current detection devices, motor stator inductance, and other parameters. The parameters of the current controller are determined according to the physical parameters of the motor and the load. The parameters of the motor play an important role in the parameter setting of the current loop. In the current general-purpose servo products, the servo driver can identify the serial number of the matching motor and automatically set the current controller parameters to ensure that the current loop has a good response. For users, there is generally no need to manually adjust current loop parameters. The torque comparison with/without the current loop is shown in Figure 5. Besides, the torque map is compared in Figure 6.

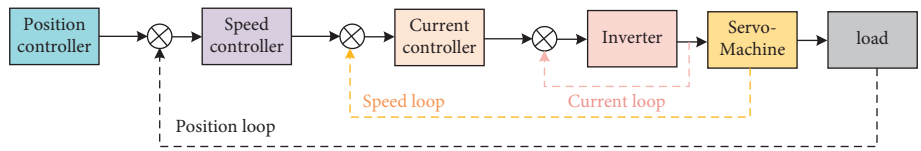


Figure 4: The general structure diagram.

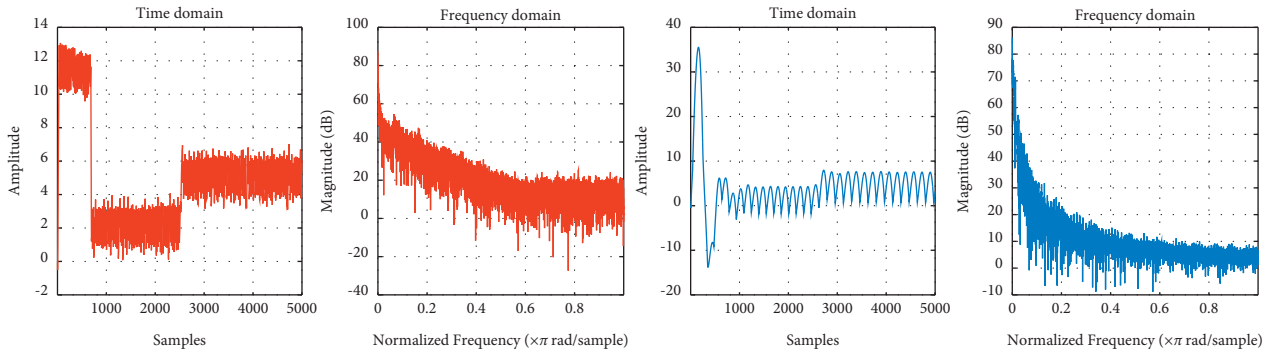


Figure 5: Torque comparison with/without current loop.

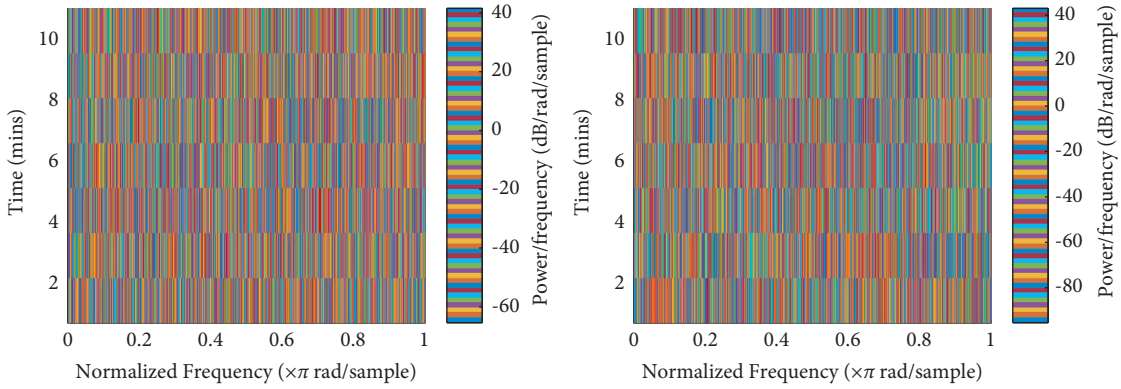


Figure 6: Torque map comparison with/without current loop.

3.2. Speed Loop. The speed loop is an important part of the machine tool servo system to realize dynamic tracking. The basis of the servo system's fast positioning and accurate tracking is that its speed loop has good dynamic response capability, wide speed range, and excellent antidisturbance performance. In engineering practice, the speed loop controller is usually set as a PI regulator.

The problems of unstable operation, slow response, vibration, and howling of the servo motor in the machine tool are closely related to the performance of the speed loop servo. The basic principle of speed controller parameter tuning is to maintain the stability of the system. On this basis, the proportional and integral gains of the speed loop are adjusted to adjust the dynamic performance of the system to the best. Generally, high-performance servo drives have some automatic optimization functions, but these parameters fully consider the stability of the system, and the settings are relatively conservative. Through the measurement and analysis of the frequency response of the speed loop, the parameters can be adjusted to make the bandwidth of the speed loop large enough to improve the dynamic response speed of the speed loop. Speed and torque comparisons with/without speed loop are shown in Figures 7 and 8.

3.3. Position Loop. The function of the position loop is to use the difference between the set target position and the actual position to generate the motor degree command through the position regulator. The regulator of the position loop is usually designed as a pure proportional regulator, which can easily obtain stable, non-overshoot position control and good positioning accuracy, but it is inevitable that there will be a steady-state position following error, which affects the accuracy of the machine tool processing.

The analysis results of the position following error of CNC machine tool linear cutting, taper cutting, arc cutting, etc. show that the tracking error of the position servo system is closely related to the position loop gain of the system and presents a certain inverse proportional relationship, that is, the position loop. The higher the gain, the smaller the tracking error of the system. However, the increase in the

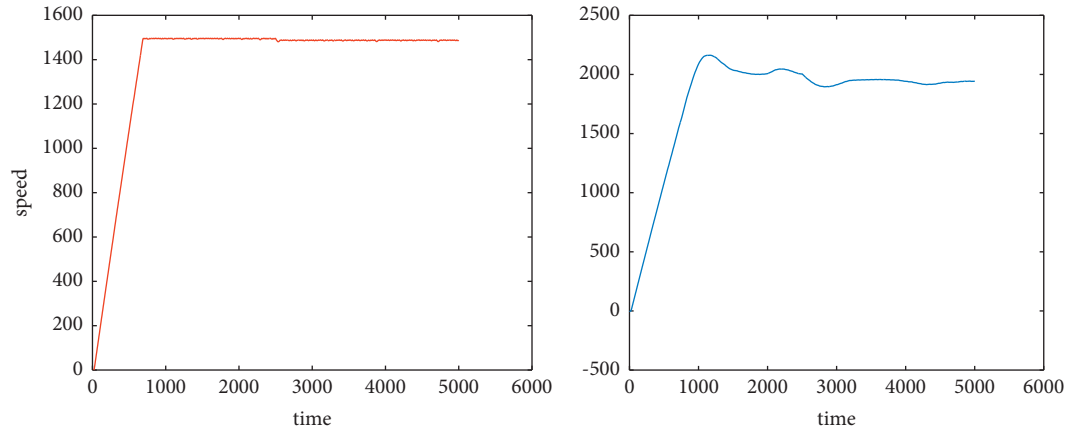


Figure 7: Speed comparison with/without speed loop.

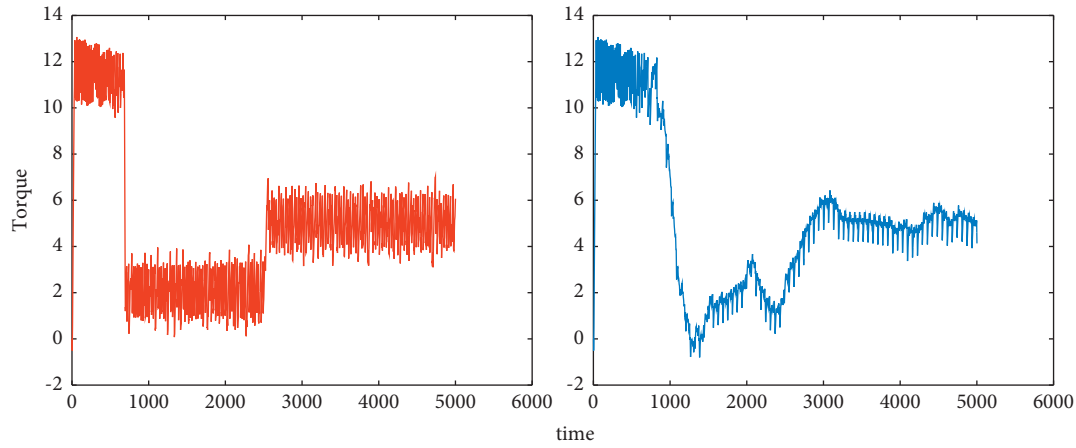


Figure 8: Torque comparison with/without speed loop.

position loop gain is limited by the cut off frequency of the speed inner loop. It cannot be indefinitely increased; otherwise, it will affect the stability of the system. For this reason, the method of compound feedforward control is often used to reduce the position following error.

The servo optimization of the position loop is mainly to adjust the position loop control gain and the position loop compound control feedforward gain coefficient to reduce the following error of the position loop. Since the three closed loops are employed in this method, less error may be obtained when compared with the conventional methods.

Combined with the FANUC drive control system adopted by the machine tool, this debugging intends to adjust the following main servo parameters, and the adjustment of auxiliary parameters should be set according to the needs, without too much research. First, with the help of the FANUC system servo debugging guide tool—Servo Guide software, the servo axis parameters are adjusted to optimize the single-axis servo performance. Using Servo Guide software to collect test signals can easily make judgments on the servo performance after servo parameter adjustment and guide the optimization of servo parameters to determine the best state of servo adjustment. Subsequently, the QC10 ballbar is used to simulate the arc, analyze the dual-axis linkage operation performance after adjusting the parameters, and use the error analysis function provided by the software to further guide the optimization of the servo performance parameters of each servo axis. The machine tool is made to meet the needs of actual processing.

On the premise that the mechanical system does not vibrate, the larger the parameter setting, the better the command tracking of the speed loop, and the stronger the rigidity of the servo system. FANUC system No. 2021 parameter speed gain, also known as load inertia ratio, is the most critical parameter of the speed loop. Its setting is related to the actual load inertia ratio of the axis servo system, but the two are relatively independent and different concepts. For convenience, the 2021 parameter is called as the load inertia ratio parameter.

There is a conversion relationship between the actual saved value of the No. 2021 load inertia ratio parameter and the speed gain display value. The higher the load inertia ratio parameter setting, the faster the response speed of the system, so this parameter should be set as high as possible. But the setting value is not as high as possible. If the speed gain is set too high, vibration and whistling will occur when the machine axis moves. The predicted values are shown in Figure 9.

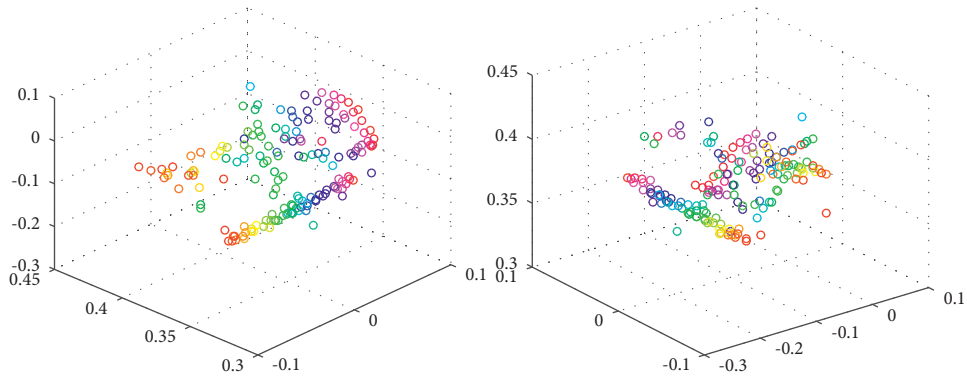


Figure 9: Predicted values.

(1) Measurement of the frequency response of speed loop interference. First, the position gain is set to a lower value (such as 1000), and then, the speed gain is gradually increased without generating abnormal response and vibration. At each speed gain value, the servo debugging software Servo Guide is used to input the frequency-varying sinusoidal disturbance to the torque command, the sine wave disturbance input is measured as 1 and the torque command output 1 is output by the speed controller, and then, the speed loop interference frequency can be obtained as a response. The best speed gain is selected with the help of the frequency response graph obtained by the measurement and whether there is abnormal vibration and howling during the measurement process.

(2) Software automatic gain adjustment. The Servo Guide software adjustment navigator provides the function of automatically optimizing the speed gain. The speed closed-loop frequency response graph (called the speed feed frequency response graph) is generated through the feed motion. The measurement principle is to make the machine tool servo axis feed at a low speed. The frequency response of the transfer function between the input speed command (input 2) and the motor feedback speed (output 2) is measured, as shown in Figures 5 and 6. But after experimentation, it is found that when the initial setting value is low (such as 100), the software will recommend a higher speed gain value, which may cause the machine tool to significantly oscillate. When setting a slightly higher gain value (such as 150), the value recommended by the software is reasonable, but the recommended value is also conservative. It can be adjusted upward on the basis of this recommended value. After selecting the appropriate gain, you can adjust the filter and reduce the gain value at the resonance frequency through the notch filter.

The adjustment of the position loop gain is mainly based on two principles: the principle of stability and the principle of small tracking error. Generally, the position loop controller in the industrial application field is a constant gain, that is, the proportional gain, which has the characteristics of simple control and fast response speed. When the position loop gain is small, the servo system is easy to stabilize, but the tracking error of the servo system is large, which will form a large contour error when processing the workpiece; the position loop gain value is high, the position tracking error is small, but the input feed speed suddenly changes. When its output drastically changes, the mechanical load has to withstand greater impact. Therefore, the setting of the specific position gain should comprehensively consider the two factors of stability and position following error.

In order to maintain a smooth transition at the start, stagnation, trajectory turning, and speed changes in the machine tool, the machine tool must be accelerated and decelerated in accordance with a given smooth law. Commonly used acceleration and deceleration control laws are linear and exponential. There are two control modes of acceleration and deceleration: acceleration and deceleration before interpolation and acceleration and deceleration after interpolation. The acceleration and deceleration before interpolation are to control the speed along the trajectory direction, which will not cause trajectory error but require complicated path calculation along the arc direction, and the acceleration and deceleration method after interpolation is based on the coordinate direction of each axis to the end point. The difference is controlled by changing the system loop gain, which may cause trajectory errors due to the inconsistent servo characteristics of each coordinate axis of the machine tool.

The interpolation time constant of the servo axis cutting feed is used to control the acceleration of the interpolation axis during cutting motion. The smaller the time constant after interpolation is set, the shorter the acceleration and deceleration time when the motor is running, and the greater the acceleration, which may cause acceleration overshoot and excessive axis movement. Therefore, it is not advisable to set the time constant too small. The reasonable setting of this parameter value can effectively reduce the time to accelerate to a given cutting speed without mechanical impact and reduce the contour error in the acceleration and deceleration phases. This method does not require complex coordinate transformation but directly calculates the magnitude of the flux linkage and torque on the motor stator coordinates and realizes PWM pulse width modulation and high dynamic

performance of the system through the direct tracking of flux linkage and torque.

4. Conclusions

CNC machine tools have been popularized in the development of the manufacturing industry because of their high precision, high speed, high efficiency, and safe and reliable processing. The numerical control system and the measurement system of the machine tool are the key components of the modern numerical control machine tool, especially the measurement system of the machine tool, which is a prerequisite to ensure the high precision of the machine tool. The sensor is an important part of the measurement system. Sensors are devices that can sense the measurement objects within the measurement range and convert them into output signals and information according to certain rules. They are mainly used to detect the operating objects, system status, and operating environment status of the machine tool system and effectively control the normal operation of the machine tool system. Accurate and reliable information is provided. This study takes the CNC machine tool control system as the research object and studies the optimization measures of the sensor data to its performance. First, a model of the CNC machine tool control system was built based on MATLAB/Simulink, and then, based on this model, an open-loop (without sensor) and closed-loop (with speed and position sensors) control systems were built and finally verified by simulation. The research results show that when there are sensors, the CNC machine tool control system has better stability and robustness.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

[1] Y. Yu, C. Yang, Q. Deng, T. Nyima, S. Liang, and C. Zhou, "Memristive network-based genetic algorithm and its application to image edge detection," *Journal of Systems Engineering and Electronics*, vol. 32, no. 5, pp. 1–9, 2021.

[2] Y. Ishida and S. Hashimoto, "Asymmetric characterization of diversity in symmetric Stable marriage problems: An example of Agent evacuation," *Procedia Computer Science*, vol. 60, no. 1, pp. 1472–1481, 2015.

[3] P. Zoha and R. Kaushik, "Image edge detection based on swarm intelligence using memristive networks," *IEEE trans. On CAD of Integrated Circuits and Systems*, vol. 37, no. 9, pp. 1774–1787, 2018.

[4] J. Pais, "Random matching in the college admissions problem," *Economic Theory*, vol. 35, no. 1, pp. 99–116, 2018.

[5] J. J. Jung and G. S. Jo, "Brokerage between buyer and seller agents using constraint satisfaction problem models," *Decision Support Systems*, vol. 28, no. 4, pp. 291–384, 2020.

[6] Y. Liu and K. W. Li, "A two-sided matching decision method for supply and demand of technological knowledge," *Journal of Knowledge Management*, vol. 21, no. 3, 2017.

[7] J. Byun and S. Jang, "Effective destination advertising: matching effect between advertising language and destination type," *Tourism Management*, vol. 50, no. 10, pp. 31–40, 2015.

[8] A. N. Nagamani, S. N. Anuktha, N. Nanditha, and V. K. Agrawal, "A genetic Algorithm-based Heuristic method for test Set generation in Reversible circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 37, no. 2, pp. 324–336, 2018.

[9] C. Koch and S. P. Penczynski, "The winner's curse: c," *Journal of Economic Theory*, vol. 174, pp. 57–102, 2018.

[10] C. K. Karl, "Investigating the winner's curse based on decision making in an Auction environment," *Simulation & Gaming*, vol. 47, no. 3, pp. 324–345, 2016.

[11] D. Ettinger and F. Michelucci, "Creating a winner's curse via jump bids," *Review of Economic Design*, vol. 20, no. 3, pp. 173–186, 2016.

[12] J. A. Brander and E. J. Egan, "The winner's curse in acquisitions of privately-held firms," *The Quarterly Review of Economics and Finance*, vol. 65, pp. 249–262, 2017.

[13] Z. A. Palmowski, "Note on var for the Winner's curse," *Economics/Ekonomia*, vol. 15, no. 3, pp. 124–134, 2017.

[14] B. R. Routledge and S. E. Zin, "Model uncertainty and liquidity," *Review of Economic Dynamics*, vol. 12, no. 4, pp. 543–566, 2009.

[15] D. Easley and M. O'Hara, "Ambiguity and nonparticipation: the role of regulation," *Review of Financial Studies*, vol. 22, no. 5, pp. 1817–1843, 2019.

[16] P. Klibanoff, M. Marinacci, and S. Mukerji, "A Smooth model of decision making under Ambiguity," *Econometrica*, vol. 73, no. 6, pp. 1849–1892, 2005.

[17] Y. Halevy, "Ellsberg revisited: An experimental study," *Econometrica*, vol. 75, no. 2, pp. 503–536, 2017.

[18] D. Ahn, S. Choi, D. Gale, and S. Kariv, "Estimating ambiguity aversion in a portfolio choice experiment," *Quantitative Economics*, vol. 5, no. 2, pp. 195–223, 2014.

[19] T. Hayashi and R. Wada, "Choice with imprecise information: an experimental approach," *Theory and Decision*, vol. 69, no. 3, pp. 355–373, 2010.

[20] K. Zima, E. Plebankiewicz, and D. Wieczorek, "A SWOT Analysis of the Use of BIM technology in the Polish construction industry," *Buildings*, vol. 10, no. 1, p. 16, 2020.

[21] S. Peng, L. Baobao, and S. Tao, "Injury status and strategies of female 7-a-side rugby players in Anhui Province," *Sports Boutique*, vol. 38, no. 03, pp. 72–74, 2019.

[22] P. Guild, M. R. Lininger, and M. Warren, "The Association between the Single Leg Hop test and lower-extremity injuries in female Athletes: A critically Appraised topic," *Journal of Sport Rehabilitation*, vol. 30, no. 2, pp. 1–7, 2020.

[23] U. G. Inyang, E. E. Akpan, and O. C. Akinyokun, "A Hybrid machine learning Approach for Flood Risk Assessment and classification," *International Journal of Computational Intelligence and Applications*, vol. 19, no. 2, p. 2050012, 2020.

[24] Q. Liu, S. Du, B. J. van Wyk, and Y. Sun, "Double-layer-clustering differential evolution multimodal optimization by speciation and self-adaptive strategies," *Information Sciences*, vol. 545, no. 1, pp. 465–486, 2021.

[25] H. R. Medeiros, F. D. B. de Oliveira, H. F. Bassani, and A. F. R. Araujo, "Dynamic topology and relevance learning SOM-based algorithm for image clustering tasks," *Computer Vision and Image Understanding*, vol. 179, pp. 19–30, 2019.

[26] Y. Deng, D. Huang, S. Du, G. Li, C. Zhao, and J. Lv, "A double-layer attention based adversarial network for partial transfer learning in machinery fault diagnosis," *Computers in Industry*, vol. 127, p. 103399, 2021.

[27] J. J. Chan, K. K. Chen, S. Sarker et al., "Epidemiology of Achilles tendon injuries in collegiate level athletes in the United States," *International Orthopaedics*, vol. 44, no. 3, pp. 585–594, 2020.

[28] W. Li, G. G. Wang, and A. H. Gandomi, "A survey of learning-based intelligent optimization algorithms," *Archives of Computational Methods in Engineering*, vol. 1, no. 2, pp. 1–19, 2021.

[29] G.-G. Wang, A. H. Gandomi, A. H. Alavi, and D. Gong, "A comprehensive review of krill herd algorithm: variants, hybrids and applications," *Artificial Intelligence Review*, vol. 51, no. 1, pp. 119–148, 2019.

Hindawi
Scientific Programming
Volume 2022, Article ID 3463413, 7 pages
https://doi.org/10.1155/2022/3463413

Research Article

Machine-Learning-Based Road Soft Soil Foundation Treatment and Settlement Prediction Method

Jing Zhai

School of Information Engineering, Yinchuan University of Science and Technology, Yinchuan, Ningxia 750021, China

Correspondence should be addressed to Jing Zhai; zhaijing@lidapoly.edu.cn

Received 13 January 2022; Revised 24 January 2022; Accepted 9 February 2022; Published 9 March 2022

Academic Editor: Baiyuan Ding

In order to effectively predict the settlement of soft soil foundation, improve the accuracy of road soft soil foundation settlement prediction, and improve the safety of the project, this paper proposes an optimized SVM-AR model and discusses the application scope of the SVM model and the time series AR model, respectively. The SVM-AR model is proposed by combining the respective advantages of the two types of models. Firstly, the prediction method of foundation settlement is analyzed and studied, and then the improved ABC algorithm is used to optimize the SVM model. Secondly, the optimized SVM model is combined with the AR model, the ABC-SVM model is used to predict the trend settlement, and the AR model is used to predict the random settlement and then combined to obtain the predicted settlement. The example verification shows that SVM-AR is more accurate than the SVM model prediction results and better reflects the settlement process of highway soft soil foundation.

1. Introduction

With the rapid development of my country's economy, a large number of rural population poured into cities, resulting in an increasing shortage of urban land. More and more civil and industrial high-rise buildings have attracted more and more attention, and the safety of high-rise buildings and some buildings with special requirements has become an urgent problem to be solved [1, 2]. My country has a vast territory and diverse geological and geomorphological conditions. In recent years, the construction of some airports, granaries, oil storage tanks, and large steel plants has shifted to soft soil areas with special geological environments. The most important thing to build structures in these soft soil areas is to solve the problems of foundation settlement and stability of structures. With the development and utilization of soft soil areas by government departments, the problem of foundation settlement of buildings in soft soil areas not only affects the project cost and project cycle, but also affects the quality of the entire project [3–6].

The most important mechanical properties of the soil after repeated loads are the strength of the soil and the deformation of the soil. The influence of the settlement of the foundation on the upper buildings can be accurately determined according to the deformation of the soil under the action of long-term cyclic loads. It is one of the main research topics in the field of civil engineering [7]. For the structures built on the soft soil foundation, the corresponding foundation treatment has been carried out before construction, such as overload preloading, cement mixing piles, compacted sandstone piles [8]. However, there are few studies on the settlement of the foundation under the action of repeated loads (the rise and fall of the oil level in the oil tank, the load change caused by the loading and unloading of grain in the granary) for the building above the foundation. Therefore, the prediction of soft soil foundation settlement is still very important in practical engineering practice [9–11].

Due to the differences in geological survey, simulation test, and manual calculation methods, the expected situation during design is often different from the actual foundation treatment and subgrade construction progress, and there is generally a large error in the actual subgrade settlement and its change process from the initial design stage [12–15]. Therefore, when necessary, it is necessary to accurately predict the later settlement and modify the original data through the field measured settlement data, that is, to carry

out dynamic design and construction control of the settlement of soft soil subgrade during actual construction. If the settlement amount during construction or post-construction settlement is too different from the design value, it will cause the bridge head to jump and the road surface to sink. Therefore, accurately predicting the settlement of highway soft soil foundation is an important geotechnical problem. With the continuous investment in highway construction and the increase in the calculation section of soft soil subgrade, a model is needed that can easily and accurately predict settlement. To improve prediction reliability and accuracy, various settlement calculation models have been proposed. At present, the subsidence prediction methods mainly include support vector machine, hyperbolic method, grey theory method, and neuron network [5, 16–19]. The above methods have improved the accuracy of soft foundation settlement prediction to a certain extent, but at the same time there are shortcomings. The foundation settlement not only has a certain regularity, but also has a relatively high performance due to the comprehensive effect of various influences on the subgrade settlement process. Because of strong randomness, the actual process of settlement often cannot be accurately reflected when a single prediction model is used.

Therefore, this paper proposes an optimized SVM-AR combination model, uses the improved ABC algorithm to optimize the SVM, and then uses the improved SVM model and AR model to predict the trend and random quantities of building settlement deformation to reflect the foundation settlement. The characteristics of regularity and randomness make the prediction results more accurate.

2. Analysis and Research on Prediction Method of Foundation Settlement

2.1. Deformation Mechanism of Foundation Settlement. The engineering community believes that the deformation mechanism of soft soil foundation settlement can be roughly divided into three stages: instantaneous settlement S_d, primary consolidation settlement S_c, and secondary consolidation settlement S_s. The general formula for foundation settlement calculation is

$$S(t) = S_d(t) + S_c(t) + S_s(t). \quad (1)$$

Instantaneous settlement is actually the settlement value of the foundation at the moment when the building is subjected to the load transferred from the outside and has no direct impact on the discharge of pore water in the foundation soil. The main factors affecting the instantaneous settlement are the loading rate and the loading method. Due to the different loading methods at different times, the effective stress in the soil is different, and the deformation modulus of the soil is also different; the main consolidation settlement is due to the pore water flowing from the soil. The flow out of the soil will cause the volume of the soil to gradually decrease, which will eventually lead to an increase in the settlement of the soft soil foundation. The stress increases; in other words, the settlement of the foundation increases [20]. The most important part of the three parts of foundation settlement is the main consolidation settlement; in general, people think that the subconsolidation settlement is the settlement amount that begins when the super void water pressure basically disappears. It is very small, and it takes a relatively long time. This is just what people think is that the foundation settlement is divided into three stages. In fact, the three stages of foundation settlement are basically generated at the same time from the beginning of loading. Because soils with different properties have different internal complex mechanical properties, the proportions of the three types of settlement in the total settlement are also different, and the dominant time in the entire settlement process is also different [21].

2.2. Foundation Settlement Prediction Method. The prediction of foundation settlement is very important, which not only has certain guiding significance for the construction in the early stage of the project but also can play a role in predicting the foundation settlement after the building is completed and put into use. The most important method of traditional foundation settlement prediction is the layered sum method to predict the foundation settlement [22]. In the calculation process of the layered sum method, the formula is relatively simple and the relevant parameters used are not too many, and the parameters are relatively easy to obtain. However, the precondition for the prediction of foundation settlement by this method is to only consider the influence of vertical load on foundation settlement and deformation. It is assumed that there is no influence of longitudinal load on foundation settlement and deformation. Both vertical load and longitudinal load have different degrees of influence on the settlement and deformation of the foundation. The assumed preconditions of this method are quite different from the actual influencing factors of foundation settlement, which will eventually cause a large error between the predicted results of foundation settlement and the actual settlement value. The settlement and deformation are roughly described, not as a reference for the actual construction and maintenance in the later stage. One of the traditional methods for predicting foundation settlement is the numerical analysis method [23]. Combine this model and some related parameters and then get the foundation settlement value we want. However, in the actual application process, due to the relatively many related parameters involved and the vagueness of the parameter values, there is no recognized normative value. Due to the influence of various aspects such as the environment, there is also a large error between the final data and the actual foundation settlement deformation value. Therefore, the numerical analysis method to calculate the foundation settlement value is also gradually eliminated [24].

In recent years, with the development of numerical analysis methods such as finite element methods in soil mechanics, the calculation parameters involved in their models are difficult to test, and many existing numerical calculation methods are due to researchers' understanding of soil constitutive models. The description is not clear

enough and the actual calculation conditions are not mature enough to make the obtained results unsatisfactory. Therefore, it is thought that the later settlement can be estimated based on the measured settlement data. The more common prediction methods are curve fitting method, grey system prediction method, Asaoka method, artificial neural network method, and some traditional settlement prediction methods.

(1) Curve fitting method (hyperbolic curve method, exponential curve method, and Poisson curve method) belongs to the prediction and analysis of the later settlement trend based on the existing experience combined with the settlement data observed in the previous period, among which there are the hyperbolic curve and logarithmic curve. In the prediction process, the model has some shortcomings and deficiencies, such as requiring the sample distribution to be studied to have specific properties, and there are many information parameters. It is also impossible to predict the foundation settlement in the later stage because the measured settlement time is relatively short.

(2) The gray system prediction method does not have very strict requirements on the measured settlement data, and the gray system prediction method itself is a dynamic prediction. In the actual prediction process, the model can be adjusted according to the newly added measured data. Adjustment: but the calculation program does not need to be changed, which overcomes the shortcomings of the traditional model such that the selected research data must be qualified typical data and have many information parameters. In order to use the original data for modeling analysis to predict the change of the later subsidence, we need the observation data recorded in the previous period to be exponential.

(3) The advantage of the Asaoka method (Shaogang method) is that it is possible to calculate the settlement value with a relatively low error by relying on a relatively small amount of settlement observation data, which can save a lot of time, and at the same time, it can also determine the time for the settlement to enter the subconsolidation stage. The biggest disadvantage of making judgments and calculations is that the estimated value of the final settlement depends too much on the division of the time interval. If the division of the time interval is not clear, it may cause the final prediction result to deviate greatly from the actual value.

(4) The artificial neural grid method can regard the traditional independent variables and dependent variables as input and output and use nonlinear mapping to replace the traditional complex functional relationship, which can deal with nonlinear problems well, and have good data simulation combined ability. The artificial neural network method has a very strong adaptability and can adjust the learning ability with the change of the external environment. On the other hand, the neural network has a relatively strong fault tolerance ability, the prediction and recognition speed is relatively fast, and there is no need to consider characteristic factors in the prediction process complex relationship with the target to be predicted. However, in the actual prediction process, the number of samples is determined based on experience because the neural network method does not have an effective criterion.

(5) The traditional method of predicting foundation settlement has more or less its own shortcomings and deficiencies: the preparatory work in the early stage is very heavy, not only need to deal with a large amount of information, but also a lot of repetitive work to make the work. The efficiency is relatively low, especially when encountering a lot of data and information; it becomes very difficult to process, and some data cannot be processed at all. In order to avoid these shortcomings, it is necessary to scientifically manage the settlement of soft soil foundations. A large amount of monitoring data information and reasonable selection and establishment of prediction models can improve work efficiency so as to better serve engineering construction.

3. Predictive Models and Theory

3.1. Support Vector Machines. Support vector machines (SVM) is a supervised learning model that analyzes data between classification and regression analysis. Its basic idea is to define a linear classifier with the largest margin in the function space. SVM classifiers also include kernel techniques that allow nonlinear classification. The learning strategy of the SVM classifier is the optimal classification hyperplane, in which this hyperplane must meet the classification requirements and maximize the blank space on both sides of the hyperplane while ensuring the classification accuracy. The main idea of SVM is as follows: given a set of data sets $T = \{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\}$, where $x_i \in R^n$, $y_i \in \{-1, +1\}, i = 1, 2, \ldots, N$, satisfying

$$y_i(w \cdot x_i + b) \geq 1, \tag{2}$$

make

$$\min_{w,b} \frac{w^2}{2}. \tag{3}$$

According to Lagrangian duality, the optimal solution can be obtained by solving the dual problem of the original problem, which is transformed into

$$\begin{cases} \max\limits_{\alpha} \left(\dfrac{1}{2} \sum\limits_{i=1}^{N} \sum\limits_{j=1}^{N} \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) + \sum\limits_{j=1}^{N} \alpha_i \right), \\ s.t. \quad \sum\limits_{j=1}^{N} \alpha_i y_i = 0, \\ \alpha_i \geq 0, \ i = 1, 2, \ldots, N. \end{cases} \tag{4}$$

After adding a negative sign to the target formula, the problem of solving the maximum value is converted into the problem of the minimum value. After conversion, it is

$$\begin{cases} \min\limits_{\alpha} \left(\dfrac{1}{2} \sum\limits_{i=1}^{N} \sum\limits_{j=1}^{N} \alpha_i \alpha_j y_i y_j \left(x_i \cdot x_j \right) - \sum\limits_{j=1}^{N} \alpha_i \right), \\ s.t. \quad \sum\limits_{j=1}^{N} \alpha_i y_i = 0, \\ \alpha_i \geq 0,\ i = 1, 2, \ldots, N. \end{cases} \tag{5}$$

After calculating the solution α, we further solve w and b according to α and obtain the maximum separation hyperplane and classification decision function.

3.2. Improved Artificial Bee Colony Algorithm. Aiming at the problem that the Artificial Bee Colony algorithm (ABC) algorithm has low search efficiency and is prone to generate local optimal solutions, this paper proposes an improved ABC algorithm based on two-dimensional unified population initialization and Euclidean distance update of food sources, thereby improving the search rate and efficiency of the ABC algorithm. Convergence: the specific steps are as follows.

Step 1: Initialize the population. According to the results of many experiments, the value of the kernel parameter γ is [0, 0.01], and the range of the penalty factor C is [1, 100]. Using the two-dimensional unified method, the values of γ and C were evenly divided into 25 squares, that is, 25 initial food sources, and each square represented the range of initial food sources. The schematic diagram of the initial food source range is shown in Figure 1. When the picker bee leaves the local optimal solution, it finds the square without optimal solution in all squares, randomly generates the optimal solution in the remaining squares, and uses the scout bee to search.

Step 2: Update the food source. First, optimize the penalty factor C and the kernel function parameter γ, and then express the Euclidean distance between the food source (C_1, γ_1) and the food source (C_2, γ_2), as shown in

$$d = \sqrt{(C_1 - C_2)^2 + (\gamma_1 - \gamma_2)^2}. \tag{6}$$

The traditional ABC algorithm generates new food sources as shown in

$$v_{ij} = x_{ij} + \varphi_{ij}\left(x_{ij} - x_{kj} \right). \tag{7}$$

The scout bee selects the food source as shown in

$$\mathrm{P}_i = \frac{fit_i}{\sum_{n=1}^{SN} fit_n}. \tag{8}$$

$$fit_i = \begin{cases} \dfrac{1}{1 + fit_i}, & fit_i \geq 0, \\ 1 + abs\left(fit_i \right), & fit_i < 0. \end{cases} \tag{9}$$

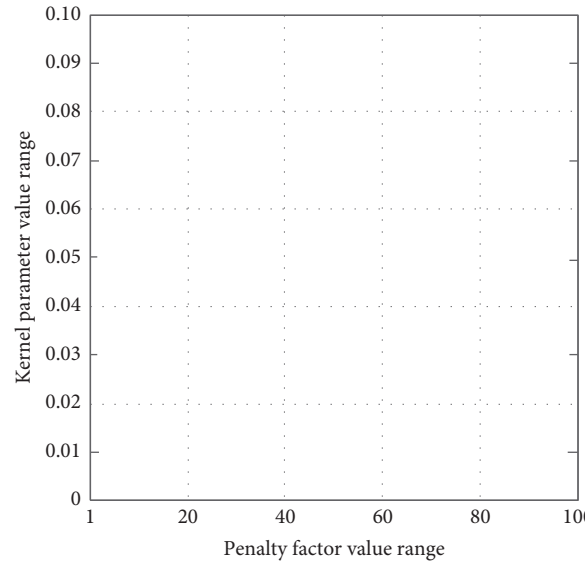


Figure 1: Schematic diagram of the range of initial food sources.

In the formula: $j \in [1,2,\ldots,D]$, $k \in [1,2,\ldots,SN]$, $\varphi\_ij \in [-1,1]$, these parameters are randomly selected; fit_i is the fitness value. If the value of φ_{ij} is small, it means that the search range of bees is small, and the algorithm does not converge or leads to early convergence; otherwise, it is easy to ignore the optimal solution, which affects the convergence on this basis, so it is improved.

Step 3: Define $\Delta_i = d_i / d_{\max}$ and take the value [0, 1], where d_i represents the distance between the current solution and the optimal solution. We substitute vertex (1, 0) and vertex (100, 0.1) into formula (5) to obtain $d_{\max}$. The updated solution of the food source is automatically adjusted by the value of Δ_{i}. If the value of Δ_{i} is small, it means that the range of finding solutions is also small; otherwise, the range is large. By using this method, the number of updates is effectively reduced. By substituting the Δ_{i} value, the updated food source is shown in

$$v_{ij} = x_{ij} + \Delta_{\mathrm{i}} \cdot \varphi_{ij}\left(x_{ij} - x_{kj} \right). \tag{10}$$

3.3. Improved ABC Algorithm to Optimize SVM. The value range of the kernel parameter γ is defined as [0, 0.01], and the penalty factor C is [1, 100]. Therefore, the improved algorithm flow is shown in Figure 2, and the steps are as follows:

Step 1: Initialize the parameters of the ABC algorithm. The parameters are as follows: food source, control parameter limit, the number of bee colonies, and the maximum number of updates.

Step 2: Initialize the parameters (C, γ) in the SVM model.

Step 3: Initialize the food sources in the search range of (C, γ) using a two-dimensional uniform pair, the obtained solution is used as the input of the SVM model, and the output is the algorithm fitness.

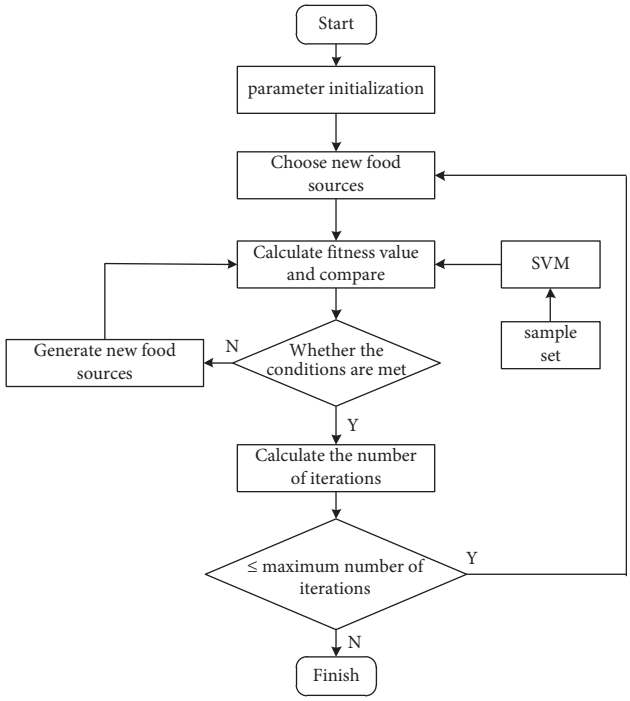


Figure 2: Optimization flow chart.

Step 4: According to formula (10), select the nearest food source to collect honey to obtain the fitness value and then compare with the previous value to retain the high fitness value.

Step 5: According to formula (8), the scout bee selects the food source. If the food source at this position has been selected by other bees, the scout bee continues to search for the food source, obtains the fitness value, and then compares it with the previous value.

3.4. Autoregressive Model. Autoregressive model (AR) models are also known as time series models. For a stable, normal, zero-mean time series $x(k)$, the autoregressive model is

$$x(k) = \sum_{i=1}^{n} a_i x(k-i) + v(k), \tag{11}$$

where $v(k)$ is a white noise sequence with zero mean.

The AR autoregressive model needs to determine the end n of the model and the parameter sequence $\{a_i\}$. The parameter series is generally estimated by the small square method. In formula (11), let $k = n+1, n+2,...,N$, and we can get formula

$$\begin{cases} x(n+1) = a_1 x(n) + a_2 x(n-1) + .. + v(n+1) \\ x(n+2) = a_1 x(n+1) + a_2 x(n) + .. + v(n+2) \\ \dots \\ x(N) = a_1 x(N-1) + a_2 x(N-1) + .. + v(N). \end{cases} \tag{12}$$

Then use the least squares method to get

$$A_N = \left(B_N^T B_N\right)^{-1} \cdot B_N^T \cdot C_N, \tag{13}$$

where

$$A_N = [a_1, a_2, \dots, a_n]^T,$$

$$B_N = \begin{bmatrix} x(n) & x(n-1) & \dots & x(1) \\ x(n+1) & x(n) & \dots & x(2) \\ \dots & \dots & \dots & \dots \\ x(N-1) & x(N-2) & \dots & x(N-n) \end{bmatrix}, \tag{14}$$

$$C_N = [x(n+1), x(n+2), \dots, x(N)]^T.$$

The parameter estimates are obtained to obtain the AR autoregressive model.

$$x(k) = \sum_{i=1}^{n} a_i x(k-i). \tag{15}$$

Use the above function to make predictions on the sequence. For the determination of the model end n, the minimum information criterion (AIC) is used, that is,

$$AIC(n) = p \ln \sigma^2 + 2n. \tag{16}$$

In the formula, p is the total number of sequence data; σ^2 is the residual equation of order. Let the minimum value of (16) corresponding to n be the optimal order.

3.5. ABC-SVM Combined with AR Model. From the perspective of subsidence analysis, the observed data can be divided into trend and random quantities, and the characteristics of a large number of surface subsidence monitoring data show that the subsidence is composed of two parts: the trend quantity and the random quantity of subsidence, that is,

$$s_i = a_i + b_i. \tag{17}$$

In the formula, s_i is the amount of settlement; a_i is the trend of settlement; b_i is the random amount of settlement. However, the core idea of the SVM model is to transform the input into a high-dimensional space through nonlinear transformation and obtain a unique optimal solution, so that the trend term can be accurately determined. The AR model is more suitable for analyzing stationary random quantities. Therefore, the combination of SVM and the respective characteristics of AR establish the SVM-AR model.

$$S_i = A_i + B_i. \tag{18}$$

In the formula, S_i is the sedimentation amount; A_i is the SVM predicted trend amount; B_i is the AR model predicted random amount.

The prediction steps of the SVM-AR model are as follows.

(1) Using n observations of sedimentation, an SVM model is established, the predicted trend value A is obtained, and the residual value x_i is calculated.

(2) Using the residual value x, find the minimum value of the minimum information criterion AIC, determine the optimal order n of the AR model, and use

Table 1: Measured value of typical subgrade section settlement.

Time (d)	Measured value (mm)	Time (d)	Measured value (mm)
215	74.9	221	134.2
232	86.3	253	145.2
271	96.2	294	148.3
302	104.8	314	154.7
321	108.1	336	159.9
359	118.6	375	167.8

Table 2: SVM model prediction results.

Time (d)	Measured value (mm)	SVM predictive value (mm)	Relative error (%)
253	145.2	141.8	2.31
294	148.3	145.2	1.67
314	154.7	150.7	2.08
336	159.9	156.1	2.31
375	167.8	163.9	2.27

the least squares method to obtain the model AR, thereby obtaining the settlement random item B_i.

(3) Using formula (18), the sedimentation amount is obtained.

4. Case Application Analysis

This paper is based on the Wujiang section of the Sujiahang Expressway. The length of the highway is 50KM. There are a lot of waters in this range, the altitude is low, and the terrain is flat. As much as 92% of the total length of the roadbed is soft soil foundation, the soft soil buried depth is generally more than 20 m, even up to 33 m, mostly silt or silt soil, with high compressibility and low strength. In this paper, a typical section is taken as an example of the average settlement of the top surface of K86 + 520. The observed data are shown in Table 1.

In this paper, the SVM model is used for conventional prediction first, and then the improved ABC algorithm is used to optimize the SVM model prediction for comparison. First, the first seven subsidence data were used as training samples to construct the model structure, and the last five subsidence data were used as test samples for the predicted value of the model. The SVM model was used to predict the data, and cross-validation was used to select the optimal parameters.

The prediction results of the SVM model are shown in Table 2.

$A_i = (80.15, 85.56, 92.14, 100.57, 104.78, 114.25, 123.41, 128.91)$ from the above SVM model. The corresponding residuals $x_i = (-4.34, 0.26, 1.57, 3.71, 2.08, 3.59, 3.98, 4.51)$. Using MATLAB programming to get the minimum AIC = 36.7, AR order $n = 2$, that is, the model AR (2), so as to get the prediction random item $B_i = (3.09, 3.41, 3.45, 3.45)$, the SVM-AR prediction result can be obtained; see Table 3.

In order to facilitate comprehensive analysis, the prediction results of the above example SVM and SVM-AR models are plotted, and the results are shown in Figure 3.

Table 3: SVM-AR model prediction results.

Time (d)	Measured value (mm)	SVM-AR predictive value (mm)	Relative error (%)
253	145.2	145.8	0.41
294	148.3	148.2	0.12
314	154.7	154.6	0.08
336	159.9	159.7	0.21
375	167.8	167.9	0.09

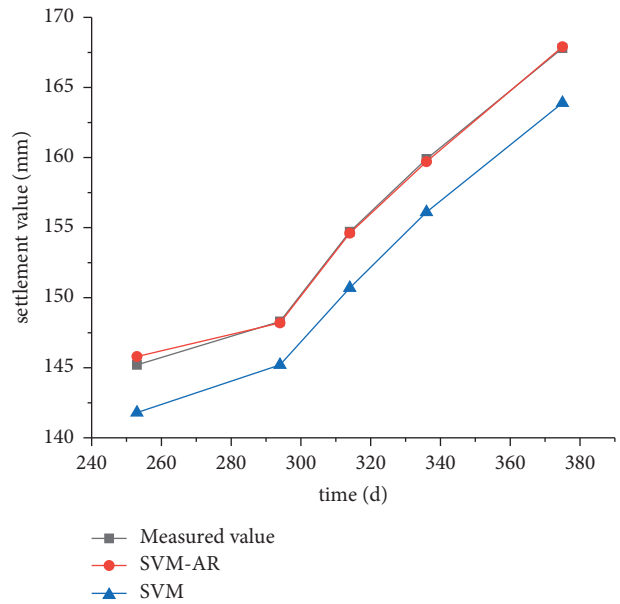


Figure 3: Comparison of predicted values between SVM-AR and SVM.

From the above calculation results, it can be seen that the SVM-AR model extracts the trend term of the subgrade settlement and conducts time series analysis on the residual, so the fitting results not only contain the trend of settlement development but also the randomness of the settlement process. It can be seen that SVM-AR (the model prediction accuracy) is better than the SVM model.

5. Conclusion

In this paper, the SVM-AR model is proposed by analyzing the characteristics of settlement data, which introduces a new viewpoint and method for predicting the settlement of highway soft soil foundation; the SVM and AR model based on ABC optimization are used to predict the foundation settlement, and the optimized SVM model extracts the trend item and the AR model extracts the random item; the foundation settlement can be accurately predicted with less parameters and a simple model form, and the prediction accuracy of the optimized SVM-AR model is significantly improved than that of the single model SVM model, so the optimized SVM-AR can better reflect the settlement process of highway soft soil foundation. The example verification shows that SVM-AR is more accurate than the SVM model prediction results and better reflects the settlement process of highway soft soil foundation. Soft soil foundations widely

exist in the engineering field. SVM-AR can accurately predict the settlement process of highway soft soil foundations. However, further research is needed on the accuracy of soft soil foundation settlement prediction in other projects.

Data Availability

The data set can be accessed upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

[1] J. R. Guo, X. C. Zheng, J. Q. Zhang, and Z. M. Zhao, "Calculation and analysis on the soft soil foundation settlement with elastic-visco-plasticity (EVP) model," *Advanced Materials Research*, vol. 1145, pp. 8–16, 2018.

[2] M. Zhu, S. Li, X. Wei, and P. Wang, "Prediction and stability assessment of soft foundation settlement of the fishbone-shaped dike near the estuary of the yangtze river using machine learning methods," *Sustainability*, vol. 13, no. 7, 2021.

[3] J. Luo, C. Wu, X. Liu, D. Mi, F. Zeng, and Y. Zeng, "Prediction of soft soil foundation settlement in Guangxi granite area based on fuzzy neural network model," *IOP Conference Series: Earth and Environmental Science*, vol. 108, no. 3, 2018.

[4] B. Yao, Li. P Zhao, and C. F Pan, "Research on influencing factors of composite foundation settlement with capped pipe piles based on centrifugal model test," *E3S Web of Conferences*, vol. 198, 2020.

[5] G. Zhang, X. Yue, J. Chen, J. Yang, and J. Wang, "Support vector machine for regression model of soft soil foundation settlement[J]," *International Journal of Applied Mathematics & Statistics*, vol. 52, no. 5, 2014.

[6] X. C. Zheng, G. U. O. Jing-rui, Z. M. Zhao, and . Q. Zhang, "Calculation and analysis on the soft soil foundation settlement with Elastic-Visco-Plasticity (EVP) model," in *Proceedings of the Abstracts of International Conference on Mechanics, Materials and Structural Engineering*, 2016.

[7] L. K. Ivana, P. Maja, and S. N. Vlasta, "Calibration OF numerical modeling and a new direct method for calculation OF shallow foundation settlements IN sand[j]," *E-Zbornik: Elektronički Zbornik Radova Građevinskog Fakulteta*, vol. 9, no. 17, 2019.

[8] H. Zhong and M. Yang, "Dynamic effect of foundation settlement on bridge-vehicle interaction," *Engineering Structures*, vol. 135, 2017.

[9] M. Nazarzadeh and S. Sarbishe-ee, "Probabilistic analysis of shallow foundation settlement considering soil parameters uncertainty effects[J]," *Open Journal of Geology*, vol. 7, no. 5, 2017.

[10] H. Siddesha and M. N. Hegde, "Structural damage detection in framed structures using under foundation settlement/rotation of bases[J]," *Structural Durability & Health Monitoring*, vol. 11, no. 1, 2017.

[11] A. M. Liu and A. J. ZhuGe, "Method of estimating the foundation settlement under different working conditions through settlement measurement[J]," *Applied Mechanics and Materials*, vol. 744, pp. 527–530, 2015.

[12] Y. Wang and G. Yang, "Prediction of composite foundation settlement based on multi-variable gray model[J]," *Applied Mechanics and Materials*, vol. 3307, pp. 580–583, 2014.

[13] Q. J. Wang, S. J. Zhang, C. Guo, and Z. R. Lu, "Centrifuge model study of high strength piles composite foundation settlement & instability under embankment in different kind soils by different pile spacing[J]," *Applied Mechanics and Materials*, vol. 3307, pp. 580–583, 2014.

[14] J. J. Guo, Z. Y. Yang, L. Cheng, X. Y. Li, and K. P. Chen, "Research on transitional pavement structures based on foundation settlement in goafs," *IOP Conference Series: Earth and Environmental Science*, vol. 304, no. 4, 2019.

[15] L. H. Zhang, H. B. Liu, and F. P. An, "Predicts based on multi modality support vector of the settlement of composite foundation," *Advanced Materials Research*, vol. 3149, pp. 919–921, 2014.

[16] S. Zhou and X. Yue, "Soft soil foundation settlement prediction and economic cost management analysis based on new algorithm," in *Proceedings of the International Conference on Education, Management, Commerce and Society*, 2015.

[17] J. Zhang and R. Du, "Study on calculation and analysis of foundation settlement in cooling tower," *IOP Conference Series: Earth and Environmental Science*, vol. 804, no. 2, 2021.

[18] X. M. Dong, "Research on prediction method of soft soil foundation settlement," *Applied Mechanics and Materials*, vol. 1448, pp. 97-98, 2011.

[19] M. Mariamme, S. Ahmad, A. A. Nadhir, and Y. Z. Mundher, "Shallow foundation settlement quantification: application of hybridized adaptive neuro-fuzzy inference system model," *Advances in Civil Engineering*, vol. 2020, 2020.

[20] K. Qian, M. Wang, A. B. Hu, and G. H. Liu, "Improvement of the formula of foundation settlement which considering three-dimensional transform[J]," *Applied Mechanics and Materials*, vol. 2974, pp. 501–504, 2014.

[21] P.-Y. Chen, H.-M. Yu, and G. Milovanović, "Foundation settlement prediction based on a novel NGM model," *Mathematical Problems in Engineering*, vol. 2014, no. 5, 2014.

[22] R. P. Li, "Nonlinear analysis of foundation settlement by modified secant modulus method," *Advanced Materials Research*, vol. 2091, pp. 594–597, 2012.

[23] G. Q. Bi, "Prediction of ground settlement based on immune genetic algorithm," *Applied Mechanics and Materials*, vol. 1682, no. 155-156, 2012.

[24] X. P. Liu, "Influences of horizontal strengthened storey on foundation settlement in high-rise frame tube structure," *Advanced Materials Research*, vol. 1615, pp. 446–449, 2012.

Hindawi
Scientific Programming
Volume 2022, Article ID 3652931, 10 pages
https://doi.org/10.1155/2022/3652931

Research Article

Research on Financial Risk Forecast Model of Listed Companies Based on Convolutional Neural Network

Weina Qin [1,2]

[1]*Guangxi Normal University for Nationalities, Chongzuo, Guangxi 532200, China*
[2]*Central Philippine University, Iloilo 5000, Philippines*

Correspondence should be addressed to Weina Qin; qinweina@gxnun.edu.cn

Received 31 December 2021; Revised 26 January 2022; Accepted 31 January 2022; Published 9 March 2022

Academic Editor: Baiyuan Ding

With the continuous improvement of China's market economy, many listed companies enjoy the unlimited development opportunities brought by the market economy environment but are also threatened by various potential risks. They may be labeled "ST" at any time due to financial risks. The label may even end up in danger of delisting. Most companies encountered serious financial crises or even bankruptcies in the later period because they did not pay enough attention to the financial problems that occurred in the early stage and did not take effective measures to deal with the crisis in a timely manner. This is extremely detrimental to the subsequent development of the company. Therefore, more and more attention has been paid to the research on the financial risk status of enterprises. Therefore, on the basis of analyzing the financial information of listed companies, this article extracts the characteristics of listed companies and images them and uses convolutional neural networks to construct a financial risk prediction model to improve the accuracy of risk prediction. Specifically, this article also compares and analyzes the financial risk prediction models of different types of listed companies, optimizes the index system, and uses the convolutional neural network method to construct a targeted financial risk prediction model with data characteristics. The actual operation data and actual risk data of the listed companies are verified, proving that it has strong adaptive ability to face different types of data, strong operability, and high prediction accuracy.

1. Introduction

With the continuous improvement of financial markets, the impact of financial conditions on the healthy development of enterprises has become particularly significant. Maintaining a good financial situation will help listed companies improve their corporate reputation and promote their faster development. As in the daily management of an enterprise, the emergence of financial risks will have certain signs in the early stage, which also provides the possibility for us to engage in research in this area. In an era where the capital market is so prosperous, more and more companies are gradually paying more attention to their own financial status, hoping to find more of their own shortcomings, so that the company can develop more healthily and also establish a better image in the market and bring more opportunities for development. Financial crisis can be directly reflected by the financial status and operating results of the enterprise. To a certain extent, this financial risk situation can be predicted by tracking and analyzing all aspects of the enterprise. Therefore, this predictable corporate financial risk situation provides the possibility for further analysis by scholars and related research institutions. The current global economic environment is unpredictable. In order to ensure sustainable and healthy development, more and more companies are focusing more on their own financial risk prediction and analysis. Effective financial risk prediction can not only sound the alarm for companies but also timely adjust the level of corporate financial risk and can help corporate managers to better manage and make rational decisions, avoid the emergence of financial crises, and make corporate development more long-term, stable, and healthy [1–9].

From the perspective of domestic and foreign research on financial risk prediction, foreign countries started earlier and the theory is relatively mature. The

main research directions and contents are as follows.: single variable judgment model, multiple linear judgment model, multiple logic probability judgment model, and fuzzy neural network judgment model. In the foreign dynamic and static financial early warning research, there are more static researches and less dynamic researches, while the domestic research still stays at the level of static research, and there is almost no dynamic research [10–15].

(1) The univariate judgment model. It uses a single variable and individual financial ratios to predict financial risks. The univariate model research originated in 1932. Professor Ftz Patrick conducted a research on 38 companies and found that the two indicators of shareholder equity ratio and debt-equity ratio have a strong ability to judge. Later, in 1966, Chicago Professor Beaver found the single financial ratio with the most differentiated ability and its critical value and proposed three indicators that are most effective in predicting the status of financial risks: debt protection, return on assets, and assets and liabilities. Rate.

(2) Multiple linear judgment model. The most successful models in this category are the *Z*-integral model and the Zeta model used in business. This model uses multiple variables and multiple financial indicators, adopts mathematical methods to construct multiple linear formulas, and predicts the financial risks of the enterprise through the discriminant values generated by weighted aggregation. The advantage of the *Z*-integral model and the *Z*eta model is that the prediction accuracy of the year before corporate bankruptcy is very high, up to 95%, and it is widely used. The main disadvantage is that the effect of horizontal comparison using this judgment is poor. The prediction accuracy rate of this model is high within two years before the bankruptcy of the enterprise, and the accuracy rate is poor if it exceeds two years [16–19]. The selected sample space and financial indicator variable requirements obey normal distribution.

(3) Multiple logic probability judgment model. This model uses the methods and principles of multivariate statistical mathematics to preset the judgment criteria, calculate the event probability based on the actual data of known factors, and then analyze the sample data to determine its classification. The main methods used in the study of financial risk early warning by using multiple logic probability judgment models are distance discrimination, Fisher discrimination, Bayes discrimination, Log it discrimination, Probit discrimination, and so on.

(4) Artificial neural network method. Beginning in the mid-1980s, with the maturity of artificial neural network technology and its successful application in various aspects, artificial neural network technology began to be used in financial risk forecasting research, and three-layer feedforward neural networks were generally used in the initial research. After that, some scholars used different neural network models to conduct financial risk forecasting research. These models are Multilayer Perceptron (MLP), Probabilistic Neural Network (PNN), Self-Organizing Mapping Neural Network (SOM), and so on.

(5) Research on dynamic financial early warning. There are four main categories: the inventory cash management model of Baumol and Tobin; the production cash management model of Friedman, Nadiri, and Coates; the wealth cash management model of Meltzer, Wallen, and Alessi; and Suvas's joint model (Corporate Model). The first three categories are all analyzed from the perspective of cash. The cash stock management model assumes that the cash holdings depend on the transaction volume; the product cash management model assumes that the cash holdings depend on the output of the product; the wealth cash management model uses wealth as the cash holding motivation; and the joint model dynamically describes the behavioral and financial characteristics of the enterprise by simulating the operation process of the enterprise. The purposes of these early warning models are similar. They all seek to balance the optimal cash holding, to minimize cash management costs and maximize the present value of future net cash flows, and to optimize the capital structure.

For the financial risk assessment problems of listed companies that belong to the same category, when analyzing the problems in the process of model construction, it is proposed that if the independent variable has predictive ability on the dependent variable, there must be a correlation between the independent variable and the dependent variable; the stronger the predictive ability, the stronger the correlation; the opposite is not necessarily the case. Therefore, the evaluation model constructed through simple analysis and processing of data has been unable to meet the needs of relevant entities for the increasing accuracy of customer credit evaluation. Therefore, new models or methods are needed to optimize the modeling process and achieve higher predictions [20–25].

The problem of financial risk prediction of listed companies is essentially a classification problem. It is the most suitable application scenario for neural networks to evaluate whether companies have risk conditions based on the monitoring index system. At present, there have been many documents that have studied the credit of listed companies based on neural networks. Risk prediction model. The deeper network layers of the neural network represent better performance, but as the number of layers deepens and the input nodes increase, the parameters of the ordinary fully connected deep network will increase sharply, slowing down the calculation speed and becoming prone to fitting problems. It leads to a convolutional neural network. The convolutional neural network uses the idea of local connection and weight sharing to greatly reduce the parameters that need to be trained in the network.

The concept of deep learning model races back to 2006; Geoffrey Hinton used neural network to complete the dimensionality reduction of data and published the results in "Science." Since then, the concept of deep learning has been continuously extended to other fields and has been successfully used. For example, three leading figures in the field of deep learning, Yann LeCun, Yoshua Bengio, and Geoffrey Hinton, published a review research article titled "Deep Learning" in Nature journal in 2015, in which they discussed deep learning. A detailed discussion was launched. In general, the main content of deep learning is that it mainly learns various feature expressions through a model composed of multiple cascaded network layers, and it also has the characteristics of multiple abstract levels. In addition, it also needs to use the back-propagation algorithm to guide the machine to self-learn by changing the internal variables and explore the deeper content contained in the data sample. In fact, this method of using back-propagation or hierarchical models to expand corresponding learning has been used in media such as images, videos, text, and audio. For now, the more successful training network types are Deep Belief Networks based on DBN algorithm, Generative Adversarial Networks based on model optimization training, Long Short-Term Memory to solve RNN and feedforward neural network (English name is convolutional neural networks), and so on. In addition, there are also many scholars who pay attention to the back-propagation algorithm of training the network. Therefore, people began to develop more efficient algorithms, including Adadelta, Adam, and RMSprop [26–30].

Convolutional neural network is the most common type in the field of ANN, because the neural network requires a lot of data in the initial stage for simulation training, and, for the computer itself, the hardware equipment requirements are high, so it is often difficult to obtain a network with relatively good performance through training. However, in recent years, with the continuous advancement of GPUs and corresponding labeled data, CNN has shown better and better results in dealing with image recognition or image classification problems. It is precisely because of this advantage of CNN that it is widely used in face recognition, object recognition, and other occasions. In recent years, the successful application of convolutional neural networks in image recognition has received widespread attention. Generally speaking, common image recognition methods can generally be divided into the three following types: decision theory recognition, syntactic pattern recognition, and fuzzy pattern recognition. Among them, a major feature of syntactic pattern recognition is the use of several structural features to form a single recognized object, which can accurately describe the characteristics of the image. Suppose that a picture is composed of lines, curves, polylines, and so forth. According to specific conventions, the knowledge of statistical decision-making in mathematical statistics is often combined to reconstruct the secondary space to achieve the purpose of image recognition. Commonly used methods include similar judgment method, similar analysis method, and function classification method.

Therefore, in order to effectively prevent the occurrence of corporate financial risks, this article takes listed companies as the research object, starting from multiple angles that affect the occurrence of corporate financial risks, constructs a comprehensive and effective financial risk prediction index system for listed companies, and uses some artificial intelligence related algorithms to construct an effective financial risk prediction model, which can effectively enhance the enterprise's risk management capabilities and improve the enterprise's risk prevention mechanism, and successfully apply it to the actual management of the enterprise to enhance enterprise risk management mechanism to promote the sustainable development of enterprises.

2. Convolutional Neural Networks

Convolutional neural network (CNN) is one of the most mature models for deep learning technology applications. On the one hand, because it inherits the advantages of deep learning to automatically extract features, in the experiment, the model automatically performs comprehensive processing operations on the original data for extraction. After effective feature information is trained and predicted, it is possible to extract and use data feature information with the maximum validity to a certain extent, thereby effectively reducing the intervention of human factors, achieving the unity of feature processing and model training, so it can solve traditional methods well. The "two-step" modeling process brings about the problem that the data dimension and model performance cannot be effectively balanced. On the other hand, the convolutional neural network model uses the theory of local receptive fields to perform convolution operations, which can reduce the number of training times by sharing weights, thereby greatly improving the efficiency of the model. It has been used in many related studies, such as that by Li Hui. In the convolutional neural network sentiment analysis method, the experiment achieved a high accuracy rate while maintaining good operating efficiency. However, the research on convolutional neural networks is mostly for nonnumerical data modeling and analysis, and there are relatively few studies on applying convolutional neural networks to numerical data. Hosaka tries to combine convolutional neural networks with financial early warning of listed companies. Financial ratio imaging is performed on the financial statement index data of listed companies, and then the convolutional neural network is used to build a model for the bankruptcy risk assessment research of listed companies, and methods such as Z-score, SVM, and MLP are compared and analyzed. The empirical results show that the new method has greatly improved the prediction accuracy rate compared with the traditional method. At the same time, the convolutional neural network is extended to the analysis of numerical data problems, which further expands the research methods and ideas of the same type of problems.

In summary, due to the complexity and diversity of actual customer data, customer credit risk is often not a single factor or a few single factors but the result of a combination of multiple factors. In the process of constructing traditional customer credit evaluation models, due to the intervention of human factors, it is not possible to

achieve a true "unified" modeling. This article draws on the method of image processing of numerical data by foreign scholars and the use of convolutional neural networks to establish risk prediction models.

As a special multilayer neural network, convolutional neural network uses back-propagation algorithm like other neural networks when training neural network. The difference lies in the network structure. The network connection of the convolutional neural network has the characteristics of local connection and parameter sharing. Local connection is relative to the full connection of ordinary neural networks, which means that a certain node of this layer is only connected to some nodes of the previous layer. Parameter sharing refers to the connection of multiple nodes in a layer sharing the same set of parameters. The core of a convolutional neural network is a multilayer network structure composed of an input layer, a convolutional layer, a pooling layer (also called a subsampling layer), and a fully connected layer. Among them, the convolutional layer and the pooling layer will generally take several, and through the alternate setting of these two structural layers in the network structure, the neural network's feature depth extraction and optimization of the input data are realized, and then it is linked to the fully connected layer and the final result is output.

Convolutional neural networks include the "input layer" of the original predictor variables, one or more "convolutional layers" that interactively or nonlinearly transform the predictors, and the "output layer" that aggregates the convolutional layers into the final result prediction, which is more complicated the structure of will also include a pooling layer to reduce parameter dimensions and a dropout composition to reduce neuronal activity. Similar to axons in a biological brain, the layers of the network represent groups of "neurons," and each layer is connected by "synapses" that transmit signals between neurons in different layers. In the convolutional layer, after the convolution operation is applied, the result of the convolution is passed to the next layer. In the convolutional layer, the number of parameters and the space size of the representation are reduced. In the final full connection, the data becomes a one-dimensional vector. In this way, as in the case of traditional classifiers, advanced decision-making can be performed. As a result, the previous layer of CNN actually performed implicit feature extraction. The general CNN structure and the corresponding working principle are shown in Figure 1.

2.1. Input Layer. The number of units in the input layer is equal to the size of the predictor. In the above figure, it is set to a sample size of 28×28. Generally speaking, each piece of index data of the stock market needs to be preprocessed before being sent to the model training. The main reason for preprocessing is to unify the unit. If the input data units are not the same, this will slow down the convergence speed of the neural network and reduce the convergence efficiency; at the same time, the input with a large data range accounts for too much weight in the training process, leading to the model ignoring the effect of other data; in addition, the limitation of the value range also requires preprocessing of the data before the training can continue.

2.2. Convolutional Layer. The convolutional layer incorporates more flexible predictor correlation items by adding convolution operations between the input and output. Each convolutional layer extracts information linearly from all input neurons. Then, each neuron applies a nonlinear activation function to activate the neuron. Before sending its output to the next layer, first it is restored to an aggregate signal. Using different convolution kernels (filters) to perform convolution operations can obtain a variety of feature information, so as to better measure the training target.

The features extracted through the convolution operation need to be converted into a two-dimensional or three-dimensional structure and input into the training model. Each feature represents a different information dimension of the sample data. Therefore, when the number of convolution kernels is increased, the convolutional neural network can enhance the structural performance and extract different information.

The basic two-dimensional convolution operation is shown in the equation. Here it is assumed that the subscript (i, j) of the output y of the convolution starts from (U, V).

$$y_{i,j} = \sum_{u=1}^{U} \sum_{v=1}^{V} w_{u,v} x_{i-u+1,j-v+1}, \tag{1}$$

where x is the sample matrix of a certain signal. The size of filter w is $U * V$; then the v output y is the convolution of the signal sequence x and filter w. Xiaotong's filter w can extract different information of signal samples.

A simple example of convolution operation is shown in Figure 2. The $5 * 5$ matrix on the left is convolved with the $3 * 3$ convolution kernel to obtain a $3 * 3$ matrix. First, suppose that the subscript of the output sample y starts from (3, 3); then,

$$\begin{aligned} y_{3,5} &= \sum_{u=1}^{3} \sum_{v=1}^{3} w_{u,v} x_{3-u+1,3-v+1} \\ &= -1. \end{aligned} \tag{2}$$

Take this as an example to get the output matrix on the right, which extracts the features of the sample matrix on the basis of maximizing edge information. Its general continuous form is

$$y(n) = \int_{-\infty}^{\infty} f(x) w(n-x) \mathrm{dx}. \tag{3}$$

The discrete form is

$$y(n) = \sum_{x=-\infty}^{\infty} f(x) w(n-x). \tag{4}$$

Function f is usually called the input function, while w is generally called the kernel function or weighting function, and output y is called the feature map. Each convolution operation extracts information linearly from the feature map

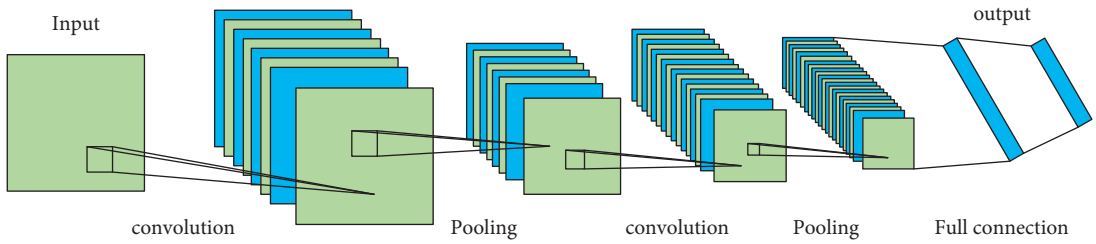


Figure 1: The general CNN structure.

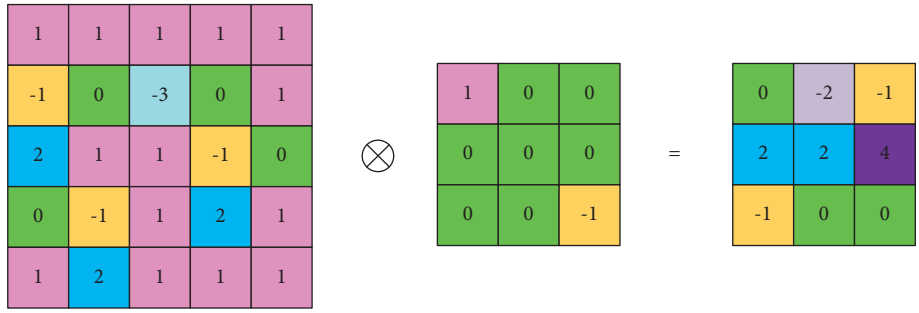


Figure 2: Convolution operation.

of the upper layer, performs a convolution operation on map x of the upper layer and filter w, and adds the bias term constant b to the convolution result to obtain the net input signal. Then the nonlinear activation function f is applied to activate the neuron signal and restore it to the aggregate signal. Finally, the results of each feature mapping are linearly summarized and input to the next layer. The calculation process is as follows:

$$\begin{aligned} z^l &= w^l \otimes x + b^l, \\ y^l &= f\left(z^l\right). \end{aligned} \tag{5}$$

2.3. Pooling Layer. The pooling layer uses pooling functions to measure the overall characteristics of data information, while ignoring unimportant subtle features. The pooling operation mainly reduces the feature dimension and enhances the network's robustness to image scaling and rotation. In the convolutional layer, the number of features is reduced, but the number of neurons is basically unchanged. Therefore, it is still necessary to perform a pooling operation at the pooling layer to reduce the feature dimension and avoid overfitting.

2.4. Activation Function. The function of the activation function is to activate the neurons to reduce the probability of overfitting by appropriately abandoning some of the meridians. After the net input z is activated by the nonlinear activation function f, the neuron activity value a can be obtained.

$$a = f\left(z^l\right) \tag{6}$$

In order to enhance the network's presentation ability and learning ability, while reducing overfitting, there are many commonly used activation functions, mainly Sigmoid-type functions and ReLU functions, as shown in Figure 3. Among them, the graph of Sigmoid-type function is S-type, and there are two commonly used forms, namely, Logistic function and Tanh function. The function definitions are

$$\begin{aligned} \sigma(x) &= \frac{1}{1+e^{-x}}, \\ \tanh(x) &= \frac{e^x - e^{-x}}{e^x + e^{-x}}. \end{aligned} \tag{7}$$

ReLU is actually a ramp function, defined as

$$\mathrm{ReLU}(x) = \begin{cases} x, & x \geq 0, \\ 0, & x < 0. \end{cases} \tag{8}$$

The last layer of the convolutional neural network is generally the Softmax layer, and the Softmax classifier is used to obtain the final classification results. The loss function of the logistic regression is in the following form:

$$J(\theta) = -\frac{1}{m}\left[\sum_{i=1}^{m}\sum_{j=0}^{1}\left\{y^{(i)} = j\right\}\log p\left(y^{(i)} = j|x^{(i)};\theta\right)\right], \tag{9}$$

where

$$p\left(y^{(i)} = j|x^{(i)};\theta\right) = \frac{e^{\theta_j^{T}x^{(i)}}}{\sum_{l=1}^{k} e^{\theta_j^{T}x^{(i)}}}. \tag{10}$$

3. Financial Risk Forecast Model of Listed Companies Based on Convolutional Neural Network

The enterprise financial risk prediction model based on convolutional neural network effectively realizes the

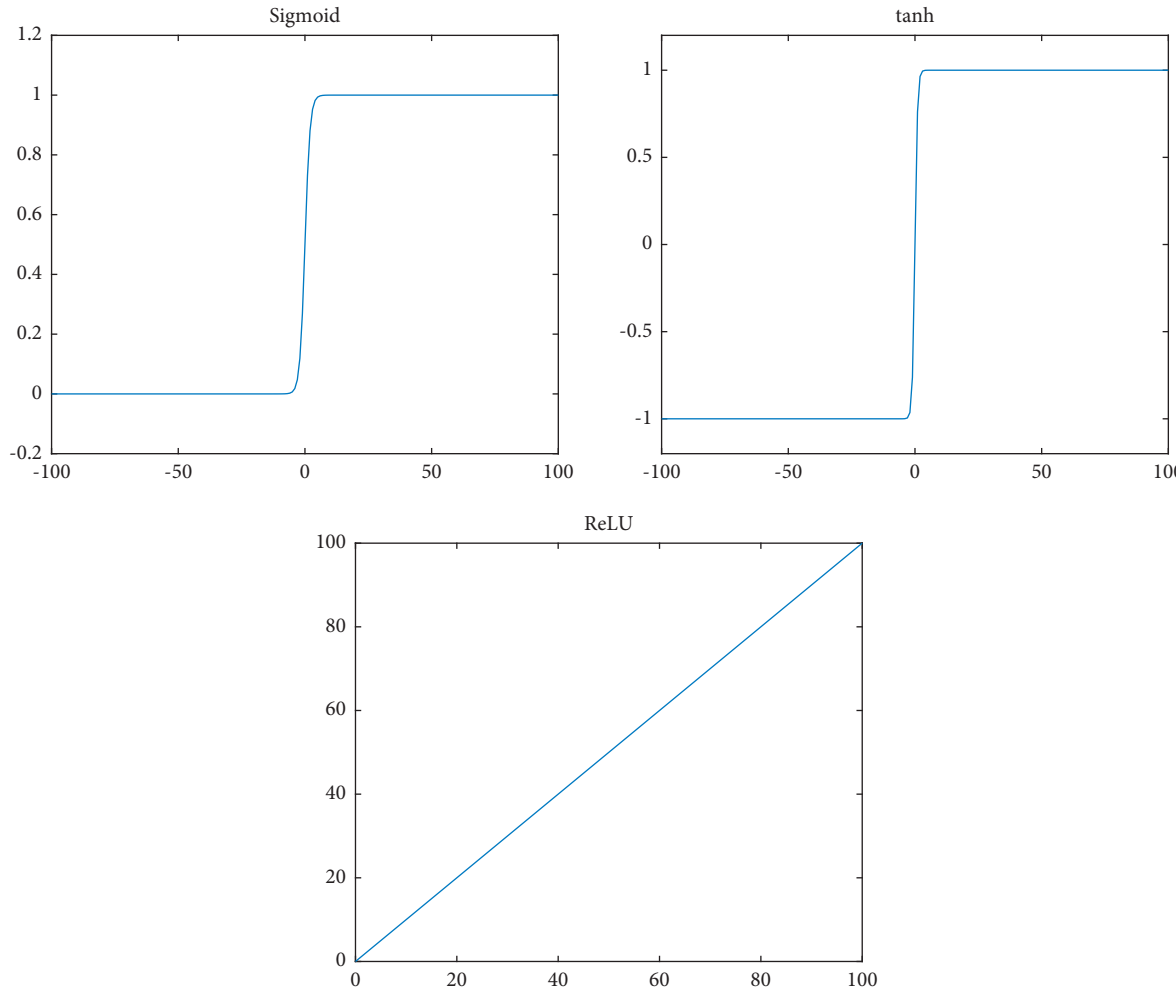


Figure 3: Function.

unification of index feature selection and model training, so the model input variables are not required when modeling, so this article only made simple data for the input data of the new model preprocessing such as classification, missing value interpolation, and cleaning. The evaluation model of the traditional method, due to the model itself and the complexity of the data source environment, cannot handle high-dimensional complex data, and further screening and processing of the data are required. Commonly used methods to eliminate redundancy and noise for index variable screening include linear models based on regularized loss functions, feature importance based on the output of machine learning models, and feature information degrees. According to the data situation, a representative method based on feature information degree is selected for feature screening. The convergence is shown in Figure 4; as can be seen, the third one is the best, since its convergence goes to zero. The financial risk prediction index system is the basis for companies to conduct financial risk assessment. This article adopts the expert interview method and the analysis method of the operational characteristics of listed companies, and the financial indicators constructed are as follows.

3.1. Solvency. Debt solvency refers to the ability of an enterprise to repay debts at maturity. Debt solvency is the basic prerequisite for ensuring the survival and sustainable development of an enterprise, as well as an important analysis indicator of enterprise credit. Debt solvency reflects the financial status and operating capacity of an enterprise. The stronger the solvency, the better the financial status and operating capacity of the enterprise. There are many financial indicators used to illustrate the solvency of a company. The indicators generally used to reflect the solvency are current ratio, express ratio, cash ratio, equity ratio, interest protection multiple, and net asset-liability ratio; this article intends to select the previously mentioned six indicators to measure debt solvency.

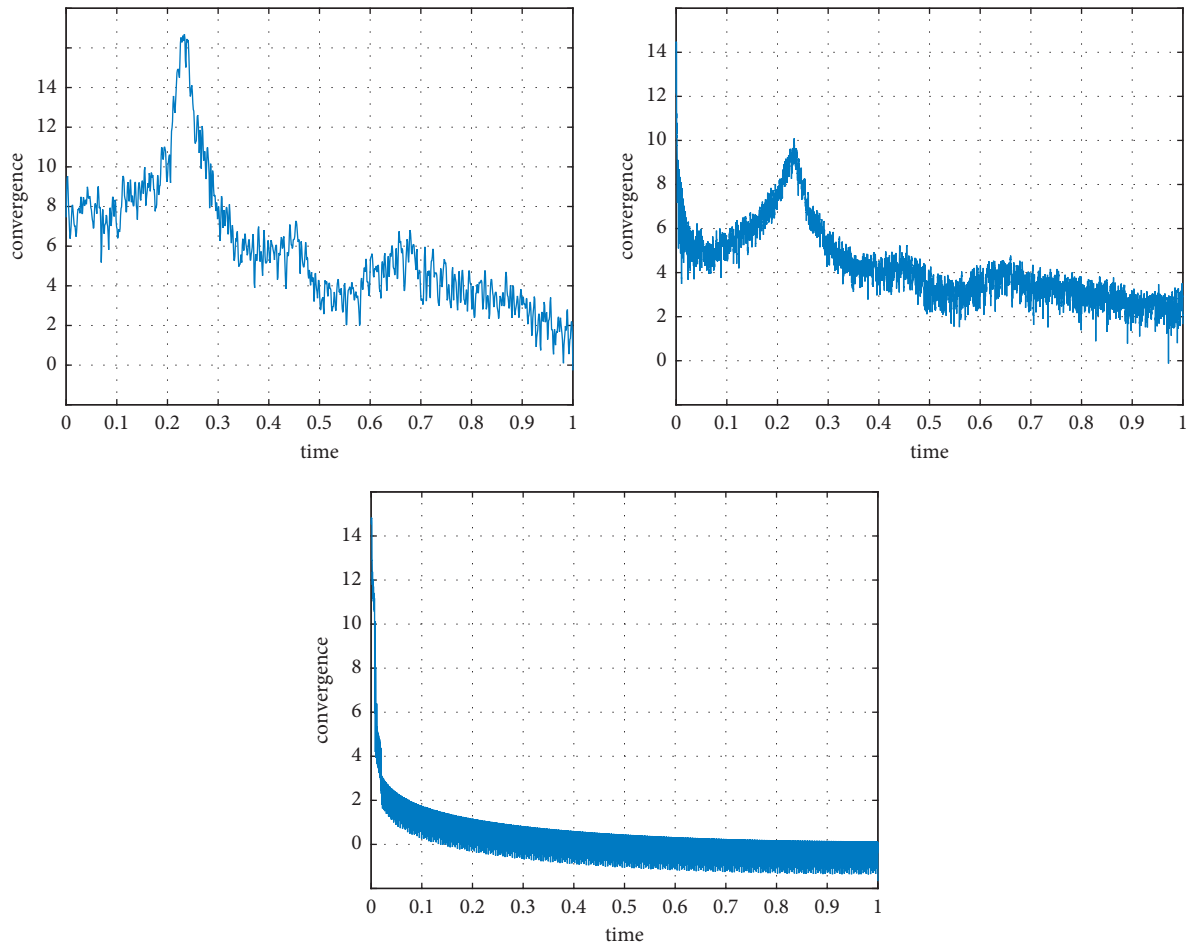


Figure 4: Convergence.

3.2. Profitability. Profitability refers to the capital appreciation ability of an enterprise to make profits. The stronger the profitability of an enterprise is, the higher the profit that the enterprise earns, and the enterprise can obtain stable survival and development. There is a strong positive correlation between profitability and solvency. The stronger the profitability, the higher the solvency. There are many indicators to illustrate profitability. Generally, the indicators used to illustrate profitability are return on net assets, return on total assets, net sales interest rate, cost and expense profit rate, total assets net interest rate, and operating net interest rate; this article intends to select the previously mentioned six indicators to measure profitability.

3.3. Operating Capability. Operational capability refers to the operational capability of an enterprise. Operating capability includes the ability to manage corporate funds. The strength of operating capability depends critically on the speed of capital circulation. The faster the capital circulation of listed companies is, the higher the efficiency of asset utilization is, the more profits the company can obtain in a certain period of time, and the stronger its operational capabilities are. The indicators that usually reflect operating capacity are inventory turnover rate, accounts receivable turnover rate, total asset turnover rate, current asset turnover rate, and accounts payable turnover rate; this article intends to select the previously mentioned five indicators to reflect operating capacity. The predicted data is shown in Figure 5.

3.4. Growth Ability. Growth ability is the ability reflected in the development process of listed companies. Compared with large companies, listed companies have smaller assets and lower risk resistance. Growth ability is the core indicator of listed companies' credit risk. This indicator is related to the future of the enterprise and can reflect the future development speed and future value of the enterprise. Therefore, an analysis of growth abilities should be added to the indicator system. The indicators that usually reflect the growth ability of a company include operating income growth rate, operating profit growth rate, net asset growth rate, total asset growth rate, and net profit growth rate; this article intends to select the previously mentioned five indicators to measure growth ability.

3.5. Ability to Obtain Cash. The ability to obtain cash mainly refers to the ability to obtain cash from operating activities in the current period. Having sufficient cash flow is the basis for

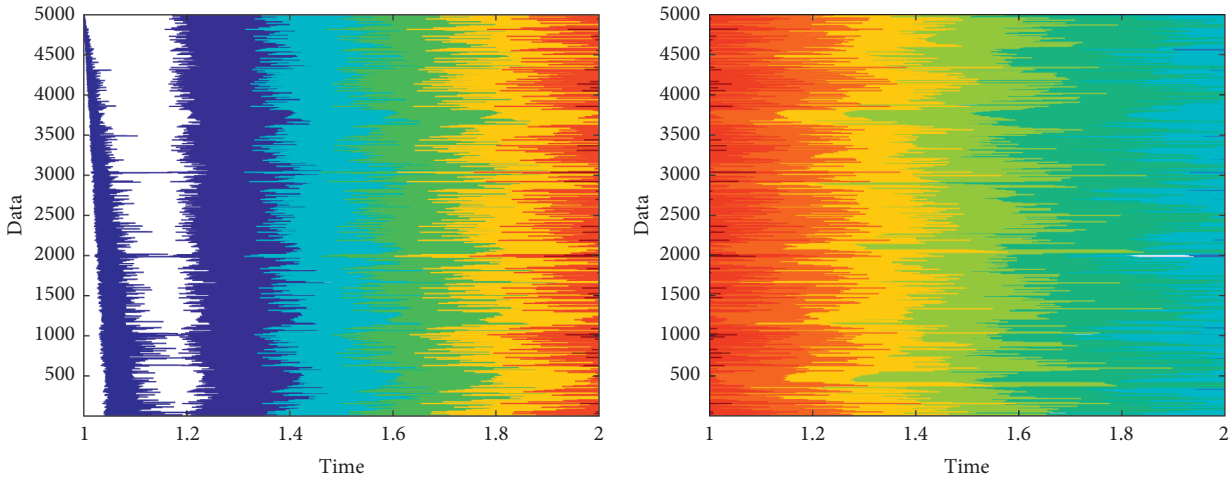


Figure 5: Predicted data.

listed companies to repay their debts, because cash flow is a direct source of debt repayment for companies. The ability to obtain cash is also a strong guarantee for the future development of listed companies and an important factor in the analysis of corporate credit. Therefore, the analysis of the ability to obtain cash should be included in the indicator system. The indicators that are usually used to reflect the ability of companies to obtain cash include the proportion of net cash flow generated from operating activities, the net cash content of operating income, and the net cash content of net profits; this article intends to use the previously mentioned three indicators to reflect the company's cash acquisition ability.

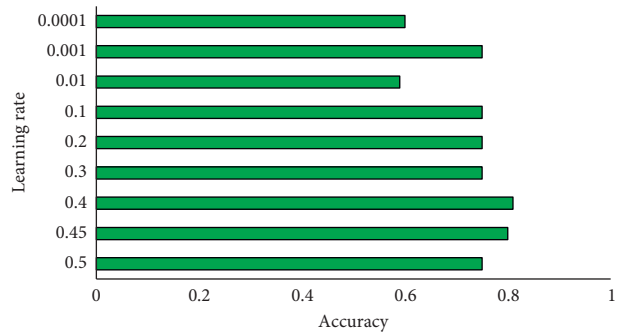


Figure 6: The results of the initial learning rate.

4. Empirical Analysis

This paper finally selected 75 non-ST companies in the Shanghai and Shenzhen listed companies as normal sample data in 2018 and 25 listed companies with the first ST from 2012 to 2018 as sample data with credit risk to construct the training data set. If there is a large difference between the sample proportions of normal and risky enterprises in the enterprise credit risk prediction and the proportions of the two types of enterprises in the actual population, the actual significance of the model will be greatly reduced, and the accuracy of the model's judgment may be overwhelmed. High estimate, so the sample selection in this article is reasonable.

The learning rate represents the speed at which information accumulates in the neural network over time. If the learning rate is set too low, the training will progress very slowly: because only a few adjustments are made to the weight of the network. However, if the learning rate is set too high, it may bring undesirable consequences in the loss function. In order to explore the impact of different learning rate settings on network performance, this paper conducted multiple experiments on the data set to set the initial learning rate, convolution kernel depth, cmV2 convolution kernel depth, and dropout ratio. The results of the initial learning rate are shown in Figure 6: the experimental results show that the selection of the initial learning rate has a direct impact on the experimental results. Too high or too low a learning rate reduces the accuracy of model training. According to the above experimental results, the learning rate is 0.4. At this time, the convolutional neural network model can quickly converge and reach a higher accuracy.

From the prediction results in Figures 7 and 8, we can see that the accuracy rates of the multivariate linear model, logistic regression model, BP neural network model, and convolutional neural network model for predicting whether there is a credit risk in the 68 test sample companies are 83.8%, 89.7%, 92.6%, and 97.1%, respectively; it can be seen that the prediction accuracy of the model based on convolutional neural network proposed in this paper is significantly higher than those of the other three models. At the same time, the accuracies of other methods are also higher, which shows that the credit risk indicator system we designed has certain reference significance for predicting the credit risk of listed companies.

The prediction accuracy of the convolutional neural network model is higher than those of the multivariate linear model, logistic regression model, and BP neural network model. Convolutional neural network is a nonlinear model, so, compared to the multivariate linear model, the linear method of logistic regression model can make a more appropriate description of the factors affecting corporate credit

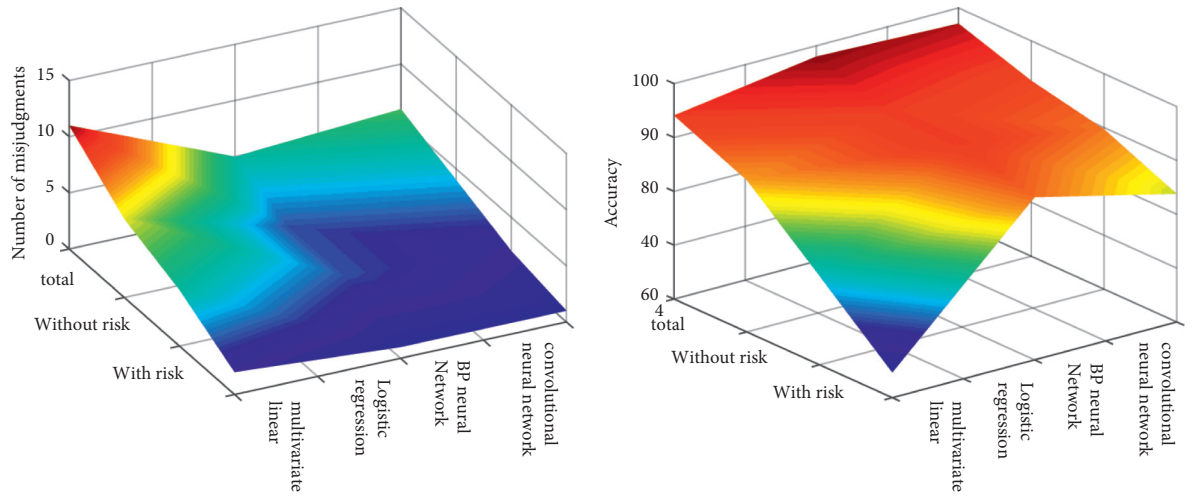


Figure 7: Comparison of the results of four risk prediction models.

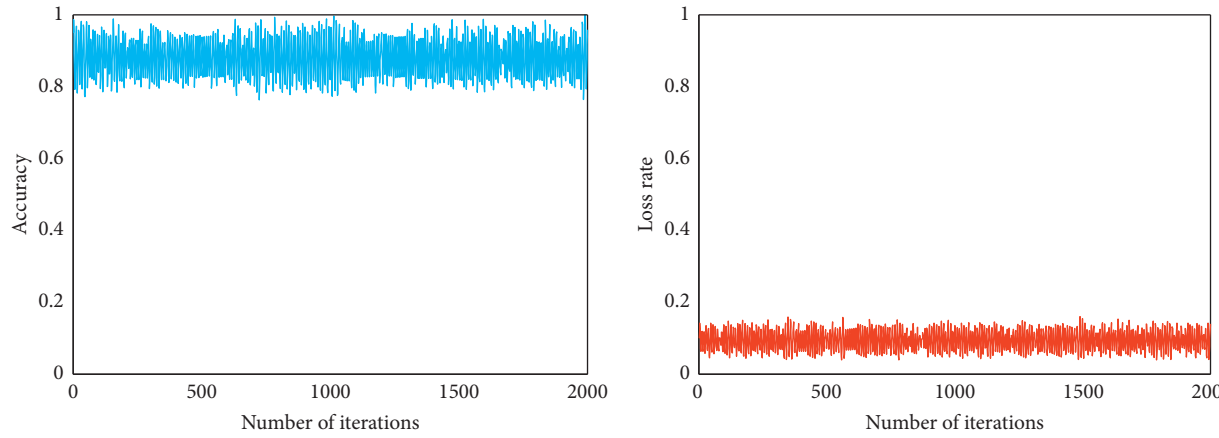


Figure 8: The prediction results.

risk. At the same time, the convolutional neural network itself has a strong self-learning ability. Compared with the BP neural network, the convolutional neural network is very good at dealing with the classification of matrix tensor data. The dynamic financial data constructed in this article is combined with nonfinancial data. The sample indicator system is taking advantage of this. The previous forecasting models all analyzed the impact of a single-year indicator of a company on whether a company's credit risk will occur in the future. The multiyear financial indicators combined with nonfinancial indicators are used to construct a corporate credit risk prediction model. Compared with the other three models, the model in this article also considers changes in corporate operating conditions. Therefore, compared with the other three statistical models, the convolutional neural network model can more accurately analyze and predict the credit risk of listed companies.

5. Conclusion

With the continuous improvement of China's market economy, many listed companies not only enjoy the unlimited development opportunities brought by the market economy environment but also are threatened by various potential risks. They may be labeled "ST" at any time due to financial risks. The label may even end up in danger of delisting. Most companies encountered serious financial crises or even bankruptcies in the later period because they did not pay enough attention to the financial problems that occurred in the early stage and did not take effective measures to deal with the crisis in a timely manner. This is extremely detrimental to the subsequent development of the company. Therefore, more and more attention has been paid to the research on the financial risk status of enterprises. Therefore, on the basis of analyzing the financial information of listed companies, this article extracts the characteristics of listed companies and images them and uses convolutional neural networks to construct a financial risk prediction model to improve the accuracy of risk prediction. Specifically, this article also compares and analyzes the financial risk prediction models of different types of listed companies, optimizes the index system, and uses the convolutional neural network method to construct a targeted financial risk prediction model with data characteristics. The actual

operation data and actual risk data of the listed companies are verified, proving that it has strong adaptive ability to face different types of data, strong operability, and high prediction accuracy. However, although the credit risk prediction model of listed companies based on convolutional neural network proposed in this paper has achieved certain success, there is still room for improvement in terms of training data set, index selection, algorithm optimization, and selection.

Data Availability

The data set can be accessed upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

[1] H. Hamada, S. Miki, and R. Nakatsu, "Automatic evaluation of English pronunciation based on speech recognition techniques," *IEICE - Transactions on Info and Systems*, vol. E76-D, no. 3, pp. 352–359, 1993.

[2] K. Truong, *Automatic Pronunciation Error Detection in Dutch as a Second Language: An Acoustic-Phonetic Approach*, Utrecht University, Utrecht, Netherlands, 2014.

[3] B. Dong, Q. Zhao, and Y. Yan, "Automatic scoring of flat tongue and raised tongue in computer-assisted Mandarin learning," in *Proceedings of the International Symposium on Chinese Spoken Language Processing (ISCSLP*, IEEE, Tianjin, China, October 2016.

[4] S. M Witt,S and J. Young, "Phone-level pronunciation scoring and assessment for interactive language learning," *Speech Communication*, vol. 30, no. 4, pp. 5–108, 2000.

[5] H. Chao, Z. Feng, F. K. Soong, M. Chu, and R. Wang, "Automatic mispronunciation detection for Mandarin," in *Proceedings of the IEEE International Conference on Acoustics*, March 2018.

[6] Y. B. Wang and L. S. Lee, "Improved approaches of modeling and detecting error patterns with empirical analysis for computer-aided pronunciation training," in *Proceedings of the 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2021.

[7] A. Neri, C. Cucchiarini, and H. Strik, "ASR-based corrective feedback on pronunciation: does it really work?" in *Proceedings of the International Conference on Interspeechicslp. DBLP*, Austin, Tx, USA, May 2016.

[8] Y. Ishida and S. Hashimoto, "Asymmetric characterization of diversity in symmetric stable marriage problems: an example of agent evacuation," *Procedia Computer Science*, vol. 60, no. 1, pp. 1472–1481, 2015.

[9] P. Zoha and R. Kaushik, "Image edge detection based on swarm intelligence using memristive networks," *IEEE Trans. on CAD of Integrated Circuits and Systems*, vol. 37, no. 9, pp. 1774–1787, 2018.

[10] W. Li, S. M. Siniscalchi, N. F. Chen, and C. H. Lee, "Improving non-native mispronunciation detection and enriching diagnostic feedback with DNN-based speech attribute modeling," in *Proceedings of the 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, March 2016.

[11] X. Qian, H. Meng, and F. Soong, *The Use of DBNHMMs for Mispronunciation Detection and Diagnosis in L2 English to Support Computer-Aided Pronunciation Training*, proc interspeech, 2021.

[12] K. Li, X. Qian, and H. Meng, "Mispronunciation detection and diagnosis in L2 English speech using multidistribution deep neural networks," *IEEE ACM Transactions on Audio, Speech, and Language Processing*, 2016.

[13] A. Lee, Y. Zhang, and J. Glass, "Mispronunciation detection via dynamic time warping on deep belief network-based posteriorgrams," in *Proceedings of the IEEE International Conference on Acoustics*, March 2020.

[14] Y. Hua, J. Zhao, and L. Jia, "Improve mispronunciation detection with Tandem feature," in *Proceedings of the International Symposium on Chinese Spoken Language Processing*, IEEE, Hong Kong, China, December 2020.

[15] J. Pais, "Random matching in the college admissions problem," *Economic Theory*, vol. 35, no. 1, pp. 99–116, 2018.

[16] J. J. Jung and G. S. Jo, "Brokerage between buyer and seller agents using constraint satisfaction problem models," *Decision Support Systems*, vol. 28, no. 4, pp. 291–384, 2020.

[17] Y. Liu and K. W. Li, "A two-sided matching decision method for supply and demand of technological knowledge," *Journal of Knowledge Management*, vol. 21, no. 3, 2017.

[18] J. Byun and S. Jang, "Effective destination advertising: matching effect between advertising language and destination type," *Tourism Management*, vol. 50, no. 10, pp. 31–40, 2015.

[19] A. N. Nagamani, S. N. Anuktha, N. Nanditha, and V. K. Agrawal, "A genetic algorithm-based heuristic method for test set generation in reversible circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 37, no. 2, pp. 324–336, 2018.

[20] C. Koch and S. P. Penczynski, "The winner's curse: conditional reasoning and belief formation," *Journal of Economic Theory*, vol. 174, pp. 57–102, 2018.

[21] C. K. Karl, "Investigating the winner's curse based on decision making in an auction environment," *Simulation & Gaming*, vol. 47, no. 3, pp. 324–345, 2016.

[22] D. Ettinger and F. Michelucci, "Creating a winner's curse via jump bids," *Review of Economic Design*, vol. 20, no. 3, pp. 173–186, 2016.

[23] J. A. Brander and E. J. Egan, "The winner's curse in acquisitions of privately-held firms," *The Quarterly Review of Economics and Finance*, vol. 65, pp. 249–262, 2017.

[24] Z. Palmowski, "A note on var for the winner's curse," *Economics/Ekonomia*, vol. 15, no. 3, pp. 124–134, 2017.

[25] B. R. Routledge and S. E. Zin, "Model uncertainty and liquidity," *Review of Economic Dynamics*, vol. 12, no. 4, pp. 543–566, 2009.

[26] D. Easley and M. O'Hara, "Ambiguity and nonparticipation," *The Role of Regulation*, vol. 22, no. 5, pp. 1817–1843, 2019.

[27] P. Klibano, M. Marinacci, and S. Mukerji, "A smooth model of decision making under ambiguity," *Econometrica*, vol. 73, no. 6, pp. 1849–1892, 2005.

[28] Y. Halevy, "Ellsberg revisited: an experimental study," *Econometrica*, vol. 75, no. 2, pp. 503–536, 2017.

[29] D. Ahn, S. Choi, D. Gale, and S. Kariv, "Estimating ambiguity aversion in a portfolio choice experiment," *Working paper*, vol. 5, no. 2, pp. 195–223, 2019.

[30] T. Hayashi and R. Wada, "Choice with imprecise information: an experimental approach," *Theory and Decision*, vol. 69, no. 3, pp. 355–373, 2010.

Hindawi
Scientific Programming
Volume 2022, Article ID 3639771, 10 pages
https://doi.org/10.1155/2022/3639771

Research Article

Design of "Immersive" Teaching Situation Simulation System for Law Course Based on Artificial Intelligence

Yanyan Tian

Shengda University of Economics, Business & Management, Zhengzhou 451191, Henan, China

Correspondence should be addressed to Yanyan Tian; 100720@shengda.edu.cn

Received 9 January 2022; Revised 24 January 2022; Accepted 9 February 2022; Published 7 March 2022

Academic Editor: Baiyuan Ding

The method of scene mapping provides a very effective technical support for classroom participants to roam freely in a virtual scene larger than the real space. When the virtual scene is larger than the area in the real space, the existing scene mapping method will cause distortion of the original virtual scene and affect the user's roaming experience and visual experience. In order to obtain the scene map with low isometric distortion when the virtual scene scale is much larger than the real space, this paper proposes the block map to divide and group the roads in the virtual scene. For each group of roadblocks, the mapping is obtained by minimizing the isometric distortion energy of the mapping on the premise that the adjacent road blocks are smooth and continuous at the junction. The segmented roadblocks are mapped to the interior of the real intersections, and then the mapping of the whole scene is optimized globally to further reduce the isometric distortion of the whole mapping. Based on the above theories, the design scheme of immersive virtual legal reality classroom is proposed, including the overall framework of immersive virtual legal reality classroom, interactive scheme design, interactive behavior design, and required technical scheme, etc. The application technology, educational application, and application experience are analyzed in combination with the constructed immersive virtual legal reality scene for corresponding innovation and improvement. By applying immersive virtual reality technology in the field of education, the results show that the simulation system has good practicability and stronger stability and can provide a better immersive experience for law class participants.

1. Introduction

The 21st century was a period of transition from a primary society ruled by law to a higher society ruled by law. Modern law classroom teaching not only served law students' learning today but also laid a foundation for the sustainable development of law students tomorrow [1]. Among them, the evaluation of law classroom teaching was the foundation and key to the study and development of law students [2]. Modern law education paid attention to the role and value of law students in the future society, and the "high quality" talent education with innovative spirit and innovative ability required that law students must be guaranteed the subject status and ensure that law students were the leading role in learning activities [3].

The standard of law course points out that law teachers were the guiders of law study and law undergraduate students were the masters of law study [4]. Law undergraduate students were the masters of law learning activities, with independent subject consciousness and clear goals, and autonomy. They can actively accept the influence of legal education through a series of autonomous learning activities and self-regulation and transform legal knowledge into their spiritual wealth through legal practice [5]. The classroom was the main channel of legal education, the main link in the process of forming the knowledge of law students, and the main position for law teachers to complete the teaching task of legal education. With the development of science and technology, the traditional classroom teaching model had been unable to meet the needs of students to acquire knowledge, which to some extent limits the students' pioneering thinking, making their thinking only follow the development of teachers. The students' learning model was roughly the same, which limits their self-concept initiative

[6]. Blended Learning combined the advantages of traditional learning with the advantages of E-learning [7]. However, most of the video resources of online learning were two-dimensional, and learners who were separated from the screen of the computer or mobile phone always study as a spectator, unable to fully integrate into the classroom behind the screen [8].

Education was changed by technology, and virtual reality was also applicable to education. As a technology made up for the defects, VR brought a new experience to education, cleverly avoiding the defects of two-dimensional video materials, changing the traditional two-dimensional into virtual three-dimensional, and giving students an immersive learning experience. Especially those who emphasized the spatial sense of education type, such as surgical medical station, sports competition tactical design, dance theater set coordination, a strong sense of interaction and impact of the law class; VR reproduces the teaching scene and accurately restores every detail to achieve an immersive experience. Nowadays, with the development of information technology, the problem of injecting technology into curriculum teaching and mobilizing students' subjective initiative could not be ignored. Motivation and behavior were not only the research emphasis of management but also was the key to the success or failure of education. In law classroom teaching, the student's learning motivation and learning investment was the key factor that influenced the learning result. Advanced information technology applications could promote the students' motivation; promoting students' learning investment behavior was the research focus on the reality of the problem. The full sensory and media richness of virtual reality technology broadens the dimension and deepens the level of technology-mediated perception. The influence of virtual reality teaching perception on teaching acceptance behavior, the effective mechanism, and the construction of the teaching model are the main contents of this study.

2. Related Work

The teaching context based on artificial intelligence, namely intelligent learning environment, had been expounded by domestic scholars in various fields from different perspectives. Zhong and Zhang [9] in 2007 defined an intelligent learning environment from the perspectives of constructivism learning theory, mixed learning theory, and modern teaching theory. They believed that intelligent learning environment was an intelligent, open, and integrated Virtual Reality (VR) learning space. Huang et al. [10] in 2012 put forward the system model and TRACE3 functional model of the intelligent learning environment and summarized the support of current technological development for the construction of intelligent learning environment from artificial intelligence, sensor technology, and communication technology. Nie et al. [11] 2013 constructed the ISMART model of the smart classroom from the perspective of system components, which consists of six systems including infrastructure, network awareness, visual management, augmented reality, real-time recording, and ubiquitous technology. Hwang [12] 2014, from the perspective of ubiquitous situational perception, believed that a smart learning environment was a technology-supported learning environment, which could be adjusted at an appropriate time and place according to learners' needs and provided appropriate support (such as guidance, feedback, and tools). Shi [13] 2017 explored the deep integration of information technology and smart classroom construction based on the web of things, cloud computing, VOIP, video surveillance, and other technologies and proposed the overall architecture diagram of a smart classroom system. Pan et al. [14] 2018 introduced speech processing technology into the construction of a smart classroom. Zero-intervention classroom data collection was achieved by mounting microphones, and acoustic wave recognition and speech recognition algorithms were designed to realize speaker identification and recording of the classroom discussion process, and the results were fed back to students and teachers in class in real-time. Zhang et al. [15] 2019 explored the mechanism of influencing factors of college students' learning engagement by using multiple regression methods from the perspectives of students' own factors, teachers and peers, and intelligent learning environment.

Foreign scholars had conducted in-depth research on intelligent learning environments. Kim et al. [16] in 2011 introduced the concept of Elastic Four Smarts (E4S) to build a smart learning cloud environment that could realize situational awareness, focused on predicting potential needs of learners and matched relevant learning content by collecting and analyzing behavioral data of learners, pushing it with the help of smart learning cloud environment, so as to provide smart learning services for learners. Singapore scholar Wu and Looi [17] 2012 emphasized the importance of learners' reflection in the intelligent learning environment, aimed to trigger students' double-cycle reflection on metacognitive strategies and beliefs. Specific prompts encourage students to reflect on a single cycle of related fields and task-specific skills. Malaysian scholar Cheng and Shan [18] 2013 believed that a smart learning environment should be learner-centered, with the ability to push learning content based on learners' knowledge base, learning style, and learning ability. It provided support for learners' lifelong learning and development. Information and communication technology was the foundation for building a smart learning environment. Thai scholar Temdee [19] pointed out in 2014 that the key to learning was to make learners in a learning environment with situational awareness function, and they would not even be aware of their own learning process. Based on a multiagent framework, a ubiquitous intelligent learning environment was developed which could adaptively push learning content in an appropriate way to promote learners' autonomous learning and collaborative learning. Spanish scholar Griol et al. [20] pointed out that artificial intelligence, immersive virtual environment, and natural language processing create more possibilities for the research and development of intelligent learning environments. They proposed a method to create an intelligent learning environment in Second Life, Open Simulator, and other games. Canadian scholar Kinshuk et al. [21] 2016 redesigned the basic structure and function of the current

learning environment to better combine emerging technologies with the required teaching transformation so as to completely transform the current learning environment into a smart learning environment. German scholars Le and Pinkwart [22] 2019 proposed an intelligent learning environment model based on constructivism learning theory. That wisdom was mainly reflected in the following: First, it can provide learners with real problems existing in reality; Second, it could provide information retrieval or collection tools to help learners in the problem-solving process; Third, problem generation tools could facilitate learners' reflection and thinking processes.

To sum up: in this paper, combined with virtual reality and immersive virtual reality of the core concepts, characteristics, and research status at home and abroad, put forward the VR and technology application in law class of immersive virtual reality simulation system, including the overall framework of immersive virtual reality classroom, interactive design, interaction design, and technical solutions required. The immersive virtual reality classroom was constructed with the real simulation theory of large scenes based on block mapping. The educational application and application experience were analyzed through experiments, and continuous innovation and improvement were made. The application of immersive virtual reality technology in the field of education and the analysis of the educational application value of immersive virtual reality classrooms were bold attempts of modern artificial intelligence technology in the direction of education, which was bound to promote the development of educational technology and provide the theoretical basis for the realization of the modernization of education.

3. Real Simulation Theory of Large Scene Based on Block Mapping

In order to roam the large scene in the small real space, it is necessary to control the user in the specific real space at all times and make the user's visual perception and motion perception maintain the same consistency with real life. In order to meet these requirements, researchers have proposed a variety of methods in recent years [50,128,153], such as redirecting walking technology, manipulating scene structure, and so on. The most widely used redirected walking technique involves forcing the user to interrupt his roaming and reset his direction of motion whenever he reaches the boundary of real space. Such operations greatly affect users' roaming experience. Recently, Sim [13] proposed a method based on scene mapping to realize real walking and roaming virtual scenes. The method calculates the smooth mapping from the virtual scene to the inner part of the real space so as to avoid the user touching the boundary of the real space in the process of walking and roaming and avoid the incoherent experience caused by forcibly interrupting the user's roaming process. However, this method is a global mapping method. When the scale gap between virtual scene and real space is large, the scene mapping generated by this method will cause great distortion of the scene, as shown in Figure 1. This will produce an unacceptable visual experience when the user is roaming and even lead to the failure of real walking and roaming. Computing scene maps with low distortion for large scenes remains challenging.

In order to map superlarge virtual scenes into small real Spaces with low isometric distortion, a new mapping method named block mapping is proposed in this paper. Different from the global mapping method proposed by Sim et al., this method is based on the divide-and-conquer strategy. It creates multiple fragmented local mappings for the virtual scene and maps each corresponding area to the interior of the real space. Firstly, the virtual scene is divided into small area blocks, and each area block is mapped to the interior of the real space by minimizing the mapping isometric distortion energy. In the process of optimizing local mapping, we need to keep smooth continuity between adjacent area blocks. After obtaining all local mappings, we can optimize the mapping of the whole virtual scene globally. In the process of global optimization, additional consideration should be given to the continuity of loop closure in the virtual scene and the obstacle avoidance requirement when the road in the whole scene is mapped to the interior of the real space. In addition, the block mapping algorithm can also map the specific areas of multiple virtual scenes to the same real location so that different virtual scenes can transfer and transform each other using transfer points. Compared with previous mapping methods, block mapping can generate high-quality scene mapping with low distortion for very large virtual scenes so as to provide a highly realistic immersive experience for virtual reality applications combined with real walking.

3.1. Block Mapping Algorithm. The algorithm input includes the plane plan of the virtual scene and plane plan of real space SV, and plane plan of real space S^R. $S^V > S^R$ and even $S^V \gg S^R$ in general. In practice, for a virtual scenario, only the area within which users can walk is considered, that is, the network within the S^V. For the sake of generality, the walkable area in the virtual space is taken as the actual input and is still denoted as S^V.

The problem of realizing real roaming using scene mapping can be abstracted as calculating a smooth mapping f: $S^V - S^R$ according to some conditions and constraints, which can map every point to point $(x, y) \in S^R$.

A global mapping from S^V to S^R may produce a large distortion. A divide-and-conquer strategy is proposed to calculate the mapping. S^V is first decomposed into K region blocks such that. Each region block P_k^V can be independently mapped to the interior of SR, resulting in a low distortion local map f_k: $P_k^V \longrightarrow S^R$. All local maps $\{f_k\}_{k=1}^K$ are assembled together smoothly to obtain the global map f: $S^V \longrightarrow S^R$.

There are two technical difficulties in stitching all local maps $\{f_k\}_{k=1}^K$ together to produce A smooth map f: $S^V \longrightarrow S^R$. First of all, it is very important to map large virtual scenes into small real Spaces and maintain low isometric distortion to provide a good experience for users to roam the virtual scenes, but this is a very difficult thing to do. In addition, the local mapping corresponding to two

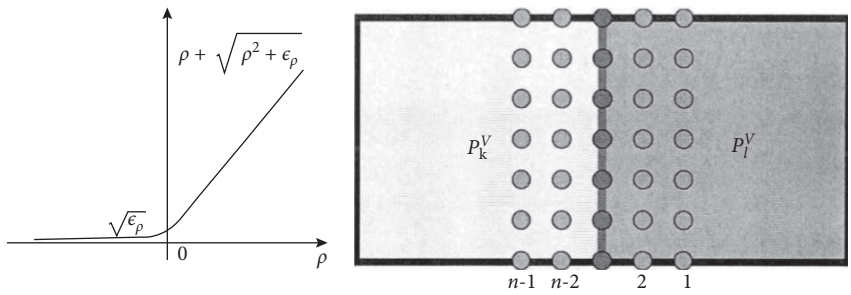


Figure 1: (a) Schematic diagram of the boundary cost function. (b) Smooth continuity diagram.

adjacent area blocks should be smooth enough at the junction; otherwise, users will experience unnatural jumps across area blocks in the process of roaming the virtual scene. Inspired by spline theory, local maps are expressed in the form of tensor product Bezier surface maps (assuming P_k^V is a rectangular region). This representation is very convenient to maintain the smoothness of the adjacency during the assembly of the local map. It is necessary to ensure that the control points of the regional block maintain corresponding linear constraints during the assembly. Other constraints on the map can be easily implemented in this representation.

3.1.1. Mapping Represents. Let's call the region block $P_k^V = [a_k, b_k] \times [c_k, d_k]$. And express the mapping as f_k: $P_k^V \longrightarrow S^R$ in the form of a tensor product Bezier surface mapping:

$$f_k(u, v) = \sum_{i=0}^{n} \sum_{j=0}^{m} c_{i,j}^k B_i^n(u^*) B_j^m(v^*), (u, v) \in P_k^V, \tag{1}$$

where $\{B_i^n(\cdot)\}_{i=0}^n$ and $\{B_j^m(\cdot)\}_{j=0}^m$ are Bernstein polynomial bases of order n and m, respectively. $u^* = u - a_k / b_k - a_k$, $v^* = u - c_k / d_k - c_k$ and $(n+1) \times (m+1)$ control points $\{c_{i,j}^k\}_{i=0,j=0}^{n,m}$ are the variables that define this mapping. To measure the quality of the mapping f_k. Uniformly sample N_k sample points $Z_k = \{z_{k,l}\}_{l=1}^{N_k}$ in P_k^V and calculate various cost functions at each sample point. After weighing the flexibility and size of the variables, $n = m = 7$ and $N_k = 3600$ are implemented.

3.1.2. Distortion Cost Function. Distortion cost using isometric energy measure $f_k(z_{k,l})$.

$$E_{K,l}^{\text{dis}} = \sum_{j=1}^{2} \left[\left(\sigma_{K,l}^j\right)^2 + \left(\sigma_{K,l}^j\right)^{-2} \right] = \left\| J_{K,l} \right\|_F^2 + \frac{\left\| J_{K,l} \right\|_F^2}{\left(\det J_{K,l}\right)^2}, \tag{2}$$

where $J_{K,l}$ represents the Jacobian matrix of f_k, $z_{K,l}$, $\sigma_{K,l}^1$, $J_{K,l}$ is the singular value of $J_{K,l}$, and $\| \cdot \|_F$ represents the Frobenius norm u. If $\det J_{K,l} > 0$, namely f_k local bijection, $E_{K,l}^{\text{dis}}$ reaches the minimum when all the singular values are 1. $E_{K,l}^{\text{dis}}$ approaches infinity as $\det J_{K,l}$ approaches 0 to prevent f_k from degenerating and flipping.

3.1.3. Boundary Cost Function. The directed distance that maps $f_k(P_k^V)$ completely into the j boundary of B, If $f_k(z_{k,l}) \in S^R, d_{k,l}^j > 0$, ε_d is a positive small quantity to compensate for the negative directed distance, when $d_{k,l}^j > 0$, the denominator of this energy becomes extremely small, making $E_{K,l}^{\text{dis}}$ extremely large.

3.1.4. Total Cost Function. The above two cost functions are combined to measure the quality of the local mapping f_k from P_k^V to S^R.

$$E\left(P_k^V, S^R\right) = \frac{1}{N_k} \sum_{l=1}^{Nk} \left(\omega_{\text{bnd}} E_{k,j}^{\text{dis}} + \omega_{\text{dis}} E_{k,j}^{\text{bnd}} \right), \tag{3}$$

where ω_{dis} and ω_{bnd} are the weights of the distortion cost function and the boundary cost function, respectively, and the default values are $\omega_{\text{dis}} = 5$ and $\omega_{\text{bnd}} = 1$.

3.1.5. Smooth Continuity Constraint. In order to ensure smooth continuity of the local mapping connection for two adjacent area blocks P_k^V and P_l^V with common edges, to ensure that the corresponding local mapping f_k and f_t have C^2 connectivity at the connection, the following constraints are applied to their control points:

$$\begin{aligned} c_{n,j}^k &= c_{0,j}^l \\ \frac{c_{n,j}^k - c_{n-1,j}^k}{b_k - a_k} &= \frac{c_{1,j}^l - c_{0,j}^l}{b_l - a_j} \\ \frac{c_{n,j}^k - 2c_{n-1,j}^k + c_{n-2,j}^k}{\left(b_k - a_k\right)^2} &= \frac{c_{2,j}^l - 2c_{1,j}^l + c_{0,j}^l}{\left(b_l - a_l\right)^2}. \end{aligned} \tag{4}$$

Among them $j = 0, 1, \ldots, m$. The above formula is linear for the control point $\{c_{i,j}^k\}_{i=0,j=0}^{n,m}$, so it can be expressed as follows:

$$AC = b, \tag{5}$$

where C is a vector composed of control points.

The local mapping f_k on the A region block P_k^V is obtained, and the local mapping force f_l on the adjacent region block P_l^V can be obtained while minimizing the constraint $AC = b$. For each region block, the degree of freedom of the

variable is not enough to satisfy the smooth continuity constraint and optimize the algorithm to obtain a map with low distortion energy. Therefore, we combine some adjacent area blocks together to form superpatches so that each superpatch can have enough degrees of freedom to obtain local mappings that meet the conditions through optimization. After the mapping calculation and assembly of all regional blocks are completed, we proceed to a global optimization to further reduce the mapping distortion:

$$\begin{cases} \min\limits_{\left\{c_{i,j}^{k}\right\}} \sum\limits_{k=1}^{K} E\left(P_k^V, S^R\right), \\ \text{s.t.} AC = b. \end{cases} \tag{6}$$

Some additional constraints need to be considered during global optimization. First, if the S^V contains a circular road, the constraint of smooth continuity between adjacent area blocks in the circular road needs to be added to the above equation. On the other hand, if there is an obstacle inside S^R, the mapping of S^V needs to avoid overlapping with the obstacle. The cost function in $E(P_k^V, S^R)$ the above formula is modified, and the internal obstacle punishment function is added. The penalty function added for each obstacle is expressed as a two-dimensional Gauss-based barrier function.

4. Construction Process of Immersion Law Classroom Situation

The construction of immersive virtual reality classroom mainly uses 3ds Max software for basic model modeling and material endowing, and Pano2VR was used to transform the virtual classroom into panoramic mode. The virtual reality classroom constructed this time did not realize real interaction. After all, a completely immersive virtual reality environment requires a mature technology research and development team. Before establishing a virtual reality classroom, it should know what objects there were in the classroom, what the size was, and what the specific values of the length, width, and height of the classroom were, which was the preliminary preparation for a basic scene modeling. Reference to the real classroom virtual classroom modeling, so that students in a real teaching situation could reduce the distraction of attention, more concentrate on learning.

4.1. Realization of Immersive Virtual Reality Classroom Modeling. Immersive virtual reality classroom is mainly composed of two parts. One is to create a virtual reality classroom environment using 3DS MAX modeling software and Unity 3D scene integration software. The second is the interaction with the virtual reality classroom environment. Teachers and students need to wear the helmet of virtual reality, auxiliary equipment necessary sensor immersed in the virtual classroom environment, the virtual reality system based on the teachers' and students' body displacement, gestures, eye location and tracking head, voice recognition, such as signal input is explained, and then to the virtual reality environment database updates, real-time adjustment of the virtual environment; Then the 3D scene information from the new viewpoint is immediately transmitted to the head-mounted display and other corresponding sensory experience.

Immersive virtual reality classrooms had tables, chairs, lecturers, blackboards, doors and windows, floors, and wall charts for advertising, just like traditional classrooms. The area of the classroom, the length, width, and height of the desks and chairs were modeled according to the standard size. Construct a classroom that could accommodate 30 people, 9 meters long, 6.6 meters wide, and 3 meters high. There were many ways to model, of course, the first choice was to use the standard basic body for modeling, cuboids, cylinders, such as walls, table legs, doors, window frames; Secondly, the angular cuboid and angular cylinder in the extended basic body were selected for modeling, such as desktop, chair, blackboard, wall chart frame, etc. Furthermore, the line, rectangle, circle, polygon in the spline line were selected, and the modeling was carried out by creating the lofting method in the composite object, such as the circular arc rod of the table and chair, curtain, etc. Finally, some more complex models, such as Windows and doors, could be completed by using two simple standard cuboids and Boolean operations. For the complex model that was difficult to construct, the object could be modeled by displaying the background image in the viewport background, tracing the two-dimensional basic shape of the object, and performing the "extrude" operation command in the editor or modifier.

The modeling of desks in the virtual classroom was completed in 3ds Max software. Before modeling the object, be sure to set the units. In the custom menu bar, set the system units to centimeters. Select the cuboid in the extended basic body of the command panel, set the parameters as length: 60 cm, width: 40 cm, height: 2 cm, and rounded corner: 50 cm. Drag the mouse in the top view to complete the modeling of the desk surface. Select the cuboid again and set the length as 38 cm; 36 cm wide, 17 cm high, 5 cm rounded, and converted to editable polygons. In the Edit modifier, select the polygon command, then select a side face to delete, add a "shell" to the object in the modifier list to increase the thickness, and the drawer modeling was complete. The circular pole on the side of the table is constructed by lofting. In the spline, select a line, draw a smooth curve in the left view as the lofting path, and adjust the radian and position of the curve; Then draw a cuboid with a length of 52 cm, a width of 2.5 cm, and a rounded angle of 1 cm as the lofting figure. Finally, click the curve to create a composite object, select Lofting, and select Get Graph in the creation method. The other legs of the desk are also made of cuboids in the standard basic body. Select all the parts of the desk, select groups from the menu bar, and then "group" operation, named Desk, and make them into a whole. In order to improve the computing efficiency of the computer, only the Max file with one desk was saved, and finally, when combined together, 30 desks were copied by the "array" method. The finished desk model is shown in Figure 2.

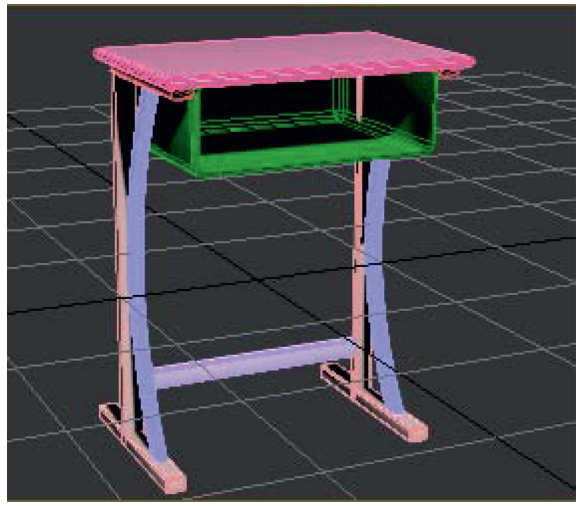


Figure 2: Classroom desk modeling diagram.

4.2. Material Endowing of Immersive Virtual Reality Classroom. After the basic modeling was completed, it also needed to assign materials to them, select the compact material mode, adjust the texture ball, bump, reflection, refraction, highlight, and other properties, and then assign the texture ball to the objects in the scene to complete the texture mapping operation. For some complex objects, you could choose to use UVW expansion or multidimensional subobjects to texture the object.

First, open the material editor and select the desired bitmap "desktop" in diffuse. JPG ", set the highlight level to 15 and glossiness to 20 in the reflection highlight. The material ball was then assigned to the tabletop in the scene to complete the material mapping of the tabletop. Select another material ball, set the highlight level to 35, glossiness to 10, and diffuse color to white, and assign it to the desk drawers, desk legs, etc. The final rendering of the desk material diagram is shown in Figure 3.

For classroom windows, doors, and other objects with glass, to give them material, but also on its transparency settings. Glass was also modeled using simple cuboids. In other words, glass was also a solid object, but it was transparent. Select a blank material ball in the Material editor, lower the opacity, and assign it to the window.

Figure 3: Classroom desk material rendering.

4.3. Lighting Setting of Immersive Virtual Reality Classroom. Once the material was finished, the next step was to lay out the lights for the scene. The scene had default lighting, but the default lighting could only be single rendered. So to make the virtual reality classroom look more realistic, "photometric" lighting was used to simulate sunlight in real life. Photometric lighting was relatively simple, and you could choose the appropriate bulb lighting template, exposure control parameters, lighting color, atmospheric effect, attenuation range, and other parameters that could be set to achieve the simulation of the real classroom light. Using photometric lighting for precise numerical Settings could make the scene more realistic. In the Create panel, select Photometry, click Target Light in the object type, and drag in the left view to finish creating the light. In the modify Settings bar, select Enable Shadow, set the creation parameters, select Exposure control in the "Render" menu, set the brightness, contrast, and other values of exposure control parameters in the dialog box that appears, add volume light to make the scene had the smoke effect of light. Create two more suitable free lights in the classroom and adjust the position of the lights in the view to complete the lighting of the scene. Free lighting was set up similarly. After modeling, material, and lighting operations, you could set up a camera in the scene, choose the appropriate position, set up a good perspective for the scene, easy to choose the perspective for observation. Figure 4 shows the virtual classroom rendered after lighting Settings.

4.4. Interactive Technology of Immersive Virtual Reality Classroom. The immersive interactive experience of virtual reality classrooms mainly comes from the realistic THREE-DIMENSIONAL display and accurate positioning and tracking so that teachers and students can have a feeling of being in the classroom. The key to interact with an immersive VR classroom was location tracking technology

Figure 4: Virtual classroom rendering after lighting setting.

such as head tracking, motion tracking, eye movement tracking, and recognition technology such as speech recognition and gesture recognition.

In order to simulate the natural visual transformation effect of human, the head-mounted display was usually equipped with a head tracking function. In terms of motion tracking, supporting controllers were required, such as Oculus Rift's Touch controller, SONY Playstation's PS Camera and Move controller, HTC Vive's laser sensor controller, etc. A large number of sensors were built into the controller. In order to achieve the movement effect of the scene. Supporting SDKS was also an essential resource for implementing interactive behavior. In immersive virtual reality classrooms, teachers and students control their vision through head movements and act as a mouse by capturing the fixation point or movement trajectory of the eyes to become spontaneous and more in line with the feeling of real life. FOVE was an eye-tracking device that used infrared tracking sensors built into a head-mounted display to track eye movements and enhance the sense of presence.

Leap Motion, Kinect, and other depth sensors were used to realize gesture recognition. These sensors could accurately detect the position and movement of the hand, obtain depth image information, and perform calculations to obtain data on the bending angle of the hand, finger, and so on. In terms of speech recognition, it was very difficult for a virtual reality environment to understand human language and interact with people in real time.

4.5. Interactive Realization of Immersive Virtual Reality Classroom. Due to the lack of necessary head-display equipment and other auxiliary equipment in the construction process of the virtual reality classroom, the real immersive virtual reality experience could not be realized. The virtual reality classroom could only generate web pages through Pano2VR software to complete the 360-degree panoramic roaming experience. In Pano2VR, hot area interaction was completed by adding hot spot interaction. Clicked the hot interactive button in the virtual classroom and entered the virtual classroom for law study by linking online resources. First, import a panorama image into Pano2VR, select add noninteractive hotspot and move the interactive hotspot button to the appropriate position. At this time, the hotspot button was red, indicating that no Settings had been carried out. Set the ID to Point01, type to URL, skin ID to the hotspot, title to Video, and you could add the corresponding main content description. The most important thing was to set the URL and destination address of the link target. Enter the target address of the online resource you wanted to link to here so that you can accurately open the externally linked resource when browsing the web page. The set hot button would turn blue. Students enter the virtual classroom from the master perspective of "first-person", immerse themselves in the classroom from the perspective of observation looked around the classroom freely, click the hot zone link, and wander in the virtual environment. Through the immersive experience of the virtual classroom, to understand the virtual classroom, to have a comparison of the real classroom felt the wonderful journey brought by the virtual world.

4.6. Panoramic Conversion of Immersive Virtual Reality Classroom. Finally, the whole virtual classroom scene was transformed into "panoramic mode." In order to render a 3D scene into a panorama, it needed to use the V-Ray rendering engine of 3ds Max plug-in. First, create a free camera in the virtual classroom scene and place it in the appropriate position of the scene. Select Render from the menu bar, click "Render Settings", change the output file size to 1000 width and 500 height in "Public", make sure the output image aspect ratio was 2 : 1, select the file save path and image format in the Render output. In "V-Ray", change the camera type to "spherical", check "Cover field", the Angle was 360°, click "Render" to output the panorama. Figure 5 shows the exported panorama. Import the rendered panoramic image into Pano2VR software and set the skin as simple. GGSK, select the output file format as HTML5, select the file output path, and click OK to generate dynamic panoramic image. You could view the panoramic VR effect of the virtual classroom through the web page version. In this way, the scene could be transformed into a VR panoramic effect, and you could use the mouse, keyboard, up, down, left, and right buttons or the hot button at the bottom of the web page to roam the virtual classroom in the "First Life perspective". You can also click the hot button to link to external resources for browsing. At this point, the entire construction process of an immersive virtual reality classroom was complete.

The constructed immersive virtual reality classroom exports a panorama through a V-Ray renderer. In Pano2VR, the panoramic pictures were converted to HTML5 format, and the web version was browsed. Students could choose their own observation angle and experience the strong visual impact brought by the virtual world through 360° rotation,

Figure 5: Panorama of the virtual classroom.

movement, zooming in, zooming out, and hot zone link, and felt themselves in it. The realistic, immersive virtual reality classroom allowed students to immerse themselves in it quickly. Classroom teaching in the state of immersion could improve students' interest in learning, promote students to explore and acquire knowledge actively, and help to maximize the learning effect.

4.7. Interview Data Analysis on the Application of Immersive Virtual Reality in Classroom Teaching. An immersive virtual reality classroom only provides a virtual space for students' learning or classroom teaching. A complete classroom needs to construct rich objects such as teaching AIDS, experimental equipment, and physical models in line with the curriculum content with the help of immersive virtual reality to fully present a complete classroom teaching. With the immersive, interactive, and visionary features of VIRTUAL reality, could the introduction of immersive virtual reality into the classroom really promote more effective teaching? Which courses are suitable for, and what difficulties will exist in the process of using them. Based on such doubts, through the form of interview, the author deeply discusses some attitudes, viewpoints, and views of the application of immersive virtual reality in classroom teaching and interviews five teachers, respectively. Before this, five teachers were briefly introduced to the knowledge related to immersive virtual reality, and five questions were designed, namely: Through what ways to learn about immersive virtual reality, what types of courses immersive virtual reality was suitable for, what advantages it had in teaching, what difficulties exist in the process of using it and what needs it had for future immersive virtual reality teaching, relevant data would be collected through their answers. VR in Table 1 refers to immersive virtual reality. Table 1 shows the collation of interview data results.

From the answers of the five teachers interviewed above, it could be seen that immersive virtual reality was widely used in education. All legal courses had their applications, providing an infinite virtual experience for students. Immersive virtual reality presents three-dimensional physical models. In order to verify experimental hypotheses, students could conduct simulation operations in the virtual reality space. Constantly explore and discover new knowledge and deepen the impression of knowledge so as to accelerate and consolidate the process of learning knowledge.

There were still many difficulties in the application of immersive virtual reality in teaching. Creating a scene was a big project that required a lot of effort and time. Complex model construction and interaction Settings require expertise. For an ordinary subject teacher, it was still very difficult to construct an appropriate virtual teaching scene by mastering multiple software. If the created virtual scene could not be widely used, it would also cause a huge waste of resources. Immersive virtual reality mainly provides users with a virtual space for scene experience and game entertainment, but it was still very rare to apply it to classroom teaching. If teachers only knew the basic functions and knowledge of immersive virtual reality and did not know how to integrate immersive virtual reality into the content of the course, the use value of immersive virtual reality would also be reduced. Schools, governments or other education departments had no relevant policy support and laws and regulations, making it difficult for immersive virtual reality to be introduced into the classroom.

According to the data of the questionnaire and interview, both students and teachers were optimistic about the prospect of an immersive virtual reality classroom. However, immersive virtual reality was still under development in terms of technology, hardware, and content. From the perspective of its functions, immersive virtual reality provides teachers and students with a realistic virtual reality classroom, as long as a set of virtual reality equipment can enter the exciting virtual reality world. Teachers and students through the helmet position tracking, intelligent speech recognition, gesture motion capture, close to the real natural interaction, and virtual objects interact in the real world, and at the same time in the virtual world, peers and teachers as "face to face" communication made teachers and students get the feel of real class participation that was exactly the same. At the same time, it also avoided

Table 1: Teacher interview form.

Teach	Q1 understanding VR approach	Q2 VR applied law courses	Q1VR advantages of application	Q4 difficulties in VR application in curriculum
1	News media, newspapers, and magazines	Intellectual property law	Provide a three-dimensional physical model	Resource development costs are high
2	Film and television works friends recommended	Administrative law	Save the cost of equipment purchase	The model is complex and difficult to construct
3	Venue experience, portal website, advertising	History of legal thought	Restore the thought scene, increase the sense of presence	The experience is not strong, and the hardware and software are not mature
4	Media campaigns, search engines	Jurisprudence of administrative procedure	In class, simulate cases and play roles to arouse students' interest	Lack of monitoring, unable to record the actual situation of students
5	TV news, website recommendations	Legal logic	Draw a 3D logic diagram to enhance the logic experience	Difficult to talk about realistic scenes without notes, which is not conducive to summary and induction

interference with external information, improved the attention of classroom teaching, and got unprecedented ultimate experience.

5. Conclusion

This paper proposes a divide-and-conquer partition mapping method, which can map a large scene into a small real space. Block mapping generates high-quality mapping, and AR technology is used to realize "immersive" teaching situation simulation of law course, and the following conclusions are drawn:

(1) A new representation is used for mapping, and a new formula is used for boundary constraints and distortion energy. The optimization problem is easy to solve and better optimization results are obtained.

(2) The divide-and-conquer strategy transforms global optimization into local optimization, which reduces the difficulty of optimization and is more suitable for the mapping of large scenes.

(3) Use 3ds Max software to build the virtual reality classroom of law and import the panorama rendering into Pano2VR to convert the whole scene into panorama mode. The key points of software applications are also emphasized.

(4) The application value of the constructed immersive virtual legal reality classroom in education is studied.

(5) Guide the constructed immersive virtual reality classroom panorama into HTML5 format, collect data, and analyze the questionnaire through teachers' experience and filling in the questionnaire. Understand the experience of immersive VR classrooms and the degree of acceptance of immersive VR classrooms. Deeply understand teachers' opinions on immersive virtual reality simulation and analyze the results.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The author declares that they have no conflicts of interest.

References

[1] N. Vaughan, V. N. Dubey, T. W. Wainwright, and R. G. Middleton, "A review of virtual reality based training simulators for orthopaedic surgery," *Medical Engineering & Physics*, vol. 38, no. 2, pp. 59–71, 2016.

[2] J. L. McGrath, J. M. Taekman, P. Dev et al., "Using virtual reality simulation environments to assess competence for emergency medicine learners," *Academic Emergency Medicine*, vol. 25, no. 2, pp. 186–195, 2018.

[3] L. L. Wang and H. N. Cheng, "Physical education image analysis based on virtual crowd simulation and FPGA," *Microprocessors and Microsystems*, vol. 79, Article ID 103319, 2020.

[4] S. L. Olivares Olivares, M. V. López Cabrera, and J. E. Valdez-García, "Aprendizaje basado en retos: una experiencia de innovación para enfrentar problemas de salud pública," *Educación Médica*, vol. 19, no. 3, pp. 230–237, 2018.

[5] G. S. Letterie, "How virtual reality may enhance training in obstetrics and gynecology," *Am Obstet Gynecol*, vol. 187, no. 3 suppl, pp. 37–40, 2002.

[6] L. Tang, C. Zhu, and H. Luo, "5G-oriented IoT coverage enhancement and physical education resource management," *Microprocessors and Microsystems*, vol. 80, Article ID 103346, 2020.

[7] T. Miki, T. Iwai, K. Kotani, J. Dang, H. Sawada, and M. Miyake, "Development of a virtual reality training system for endoscope-assisted submandibular gland removal," *Journal of Cranio-Maxillofacial Surgery*, vol. 44, no. 11, pp. 1800–1805, 2016.

[8] H. A. Davis, C. DiStefano, and P. A. Schutz, "Identifying patterns of appraising tests in first-year college students: implications for anxiety and emotion regulation during test taking," *Journal of Educational Psychology*, vol. 100, no. 4, pp. 942–960, 2008.

[9] G. X. Zhong and X. Z. Zhang, "Construction of a general intelligent learning environment model," *Computer Science*, vol. 1, pp. 170-171+197, 2007.

[10] R. H. Huang, J. F. Yang, and Y. B. Hu, "From digital learning environment to Intelligent learning environment-The change and trend of learning environment," *Open Education Research*, vol. 18, no. 1, pp. 75–84, 2012.

[11] F. H. Nie, X. L. Zhong, and S. Q. Song, "Smart classroom: conceptual features, system models and construction cases," *Modern Educational Technology*, vol. 23, no. 7, pp. 5–8, 2013.
[12] G.-J. Hwang, "Definition, framework and research issues of smart learning environments - a context-aware ubiquitous learning perspective," *Smart Learning Environments*, vol. 1, no. 1, p. 4, 2014.
[13] B. H. Shi, "Architecture design and implementation of Intelligent classroom in Internet + contemporary universities," *Journal of Central China Normal University (Natural Science edition)*, vol. 5, no. S1, pp. 91–95, 2017.
[14] L. X. Pan, W. B. Xu, and S. B. Li, "Construction of a research and discussion type wisdom classroom based on voiceprint recognition," *Experimental Technology and Management*, vol. 35, no. 07, pp. 245–250, 2018.
[15] Y. Zhang, Q. Hao, and B. L. Chen, "A study on college students' classroom learning engagement and its influencing factors under smart classroom environment-a case study of "Educational Technology Research Methods course"," *China Educational Technology*, vol. 1, pp. 106–115, 2019.
[16] S. Kim, S.-M. Song, and Y.-I. Yoon, "Smart learning services based on smart cloud computing," *Sensors*, vol. 11, no. 8, pp. 7835–7850, 2011.
[17] L. Wu and C. K. Looi, "Agent prompts: scaffolding for productive reflection in an intelligent learning environment," *Journal of Educational Technology & Society*, vol. 15, no. 1, pp. 339–353, 2012.
[18] M. Cheng and M. X. Shan, "A review of research on intelligent learning environment," *Modern Educational Technology*, vol. 23, pp. 25–28, 2013.
[19] P. Temdee, "Ubiquitous learning environment: smart learning platform with multi-agent architecture," *Wireless Personal Communications*, vol. 76, no. 3, pp. 627–641, 2014.
[20] D. Griol, J. M. Molina, and Z. Callejas, "An approach to develop intelligent learning environments by means of immersive virtual worlds," *Journal of Ambient Intelligence and Smart Environments*, vol. 6, no. 2, pp. 237–255, 2014.
[21] Kinshuk, N. S. Chen, N.-S. Chen, I.-L. Cheng, and S. W. Chew, "Evolution is not enough: revolutionizing current learning environments to smart learning environments," *International Journal of Artificial Intelligence in Education*, vol. 26, no. 2, pp. 561–581, 2016.
[22] N. T. Le and N. Pinkwart, "A Smart Problem Solving environment," 2019, https://cses.informatik.

Hindawi
Scientific Programming
Volume 2022, Article ID 5799354, 11 pages
https://doi.org/10.1155/2022/5799354

Research Article

Intelligent Processing and Classification of Multisource Health Big Data from the Perspective of Physical and Medical Integration

Haiou Tang [1,2]

[1]*Hunan City University, Yiyang, Hunan 413000, China*
[2]*Dept of Physical Education, Hoseo University, Asan, Chungcheongnam-do 336-795, Republic of Korea*

Correspondence should be addressed to Haiou Tang; tanghaiou@hncu.edu.cn

Received 17 December 2021; Revised 12 January 2022; Accepted 17 January 2022; Published 28 February 2022

Academic Editor: Baiyuan Ding

With the development of computer science and information technology, human society is gradually stepping into the Internet and big data. The medical and health industry can realize the integration and readjustment of existing resources, improve the operation efficiency of the industry, and tap the huge potential of the industry with the support of big data technology. However, the medical data in the new era has the characteristics of massive, high latitude, complex structure, and complex information, which is not conducive to the direct classification of health data. The preprocessing of health data can improve the quality of dataset, reduce the size of data, and improve the efficiency and accuracy of data classification. Based on this and according to the characteristics of health dataset and the existing pretreatment technology, this paper analyzes and improves the algorithm of abnormal data detection and data protocol in the process of reprocessing data cleaning. This paper analyzes and studies feature selection algorithms based on Bayesian inference algorithm and focuses on feature selection algorithms based on random forest. In order to solve the problem that the original algorithm ignored the relationship between the importance degrees of each feature in a single tree, a feature importance degree calculation method based on local importance degree was proposed. Through experimental analysis and comparison, the improved algorithm can select better feature subset and improve the performance of the classification model. Then, TAN classifier, BAN classifier, and MBN classifier were constructed based on preprocessed hypothyroidism data, and the performances of these four classifiers were compared through experiments. The final results show that BAN classifier has the best average classification effect.

1. Introduction

With the development of computer science and information technology, people have more and more opportunities and ways to get in touch with the Internet, and more and more network data are generated [1]. The massive increase and diversity of network data in the new era bring challenges to data analysis. In order to overcome the above difficulties, data mining technology emerges at the historic moment. Data mining is to discover the latent rules or knowledge which is not easy to be obtained directly through data observation from the massive data containing noise information redundancy or information loss data. Data mining has become one of the important directions of the development of contemporary computer science. The development of health and medical informatization is highly related to the development of computer technology [2, 3]. The development of computer technology and the popularization of computers in medical institutions bring a revolution to medical informatization [4–6].

However, medical big data and other types of big data have similar but different problems; that is, there is a large amount of missing data and there are a large number of repeated data outliers in the data, resulting in low data quality and seriously affecting the effect of data mining. In the medical field, there are various types of data, such as basic information, for example, medical treatment information, hospitalization information, physical examination information, and medical insurance information, and the data access mode is changeable [7]. For example, the common medical structure information input is uploaded to the cloud platform through intelligent testing

equipment and APP. Due to the different data structures of different information and different input methods and platform design, the structure of medical data is different, and many health data used in data mining have existed for many years; due to the input storage integration process error, which makes the situation of dataset more complex, the original data directly used in data mining will bring a large error, and it is difficult to meet the needs of products and applications [8–12].

Health big data copies the doctors on the infinite resources of high quality, make the limited distribution of medical resources in a more reasonable manner, and promote the grading clinical therapy to make it more reasonable. Through the analysis of large data on health at the same time, government agencies can realize rational pricing of drug products [13], discover the epidemic disease, and take relevant preventive measures [14, 15].

Therefore, semisupervised learning enhances the performance of learning assumptions by using labeled data and unlabeled data at the same time [16–19]. The initial assumptions are usually learned from labeled data and then updated and strengthened by unlabeled data information to complete the improvement of model performance. A semisupervised learning is actually a supervised learning and unsupervised learning in a compromised way; at the same time it combines the advantages of supervised learning and unsupervised learning, uses a lot of unlabeled data to help improve learning model in a small amount of labeled data generalization ability, and has become a current hot spot of machine learning field. More and more researches focus on semisupervised learning (SSL) [20–23].

2. Related Works

In the middle of the 20th century, after medical informatization started, machine learning technology gradually penetrated into the medical industry with the increasing popularity of Internet mobile devices and the increasing demand of social development and certain results were achieved in many aspects such as the development of auxiliary diagnostic drugs and health management. Foreign research on health big data started earlier, and the representative regions are the United States, Europe, and Japan. Their research on health big data mainly focuses on personal health, clinical decision support, medicine, disease prediction, public health, and other fields and has achieved a lot of results. The American Steward healthcare system is a community-based organization that provides basic care for community residents [24]. Every year, it treats more than 1 million patients in Massachusetts in the form of community hospital services. The Korea Biomedical Center plans to run the national DNA management system, which will combine patients' electronic health record data with system biology data, such as biological small molecules, genes, proteins, and other related data, to provide personalized diagnosis, treatment, and health management for patients, relying on the analysis and mining ability of medical and health big data. Google's Flu Trends APP, for example, helps people understand flu outbreaks in different parts of the world by checking health opinion keywords [25]. IBM developed the Healthcare Fraud Prevention and Abuse Management System (FAMS) to help health insurance payers, which can quickly identify healthcare fraud by mining health insurance payment history information. Artificial neural network can simulate the way of thinking of human brain. Considering that it can be as adaptive as the brain when dealing with nonlinear relations, it has very strong practical value [26]. BP neural network algorithm is applied to breast cancer data and improved it with particle swarm optimization algorithm. The results show that the BP neural network has better performance with fewer samples and more attributes. Support vector machine (SVM) maps the sample vector to the high-dimensional space according to the kernel function, and the mapped vector is relatively sparse in the high-dimensional space, which is conducive to finding the best separated hyperplane to complete the classification task [27]. Since it is very effective in the classification of small samples and nonlinear problems, it is also often applied in the medical and health field. In 2002, a variety of classification methods were applied to diagnose skin pigmentation diseases, and the results showed that SVM had the most reliable classification effect.

Although China's information-based medical treatment started late and there is a gap in the application scale of health big data classification technology compared with foreign countries, with the strong support for the development of health big data, the research on health big data classification technology has also received more and more attention [28]. The model was used for auxiliary diagnosis of breast cancer. The results of 5-fold cross verification showed that the detection reached 96.93 in 683 patients. The authors of [29] used SVM to obtain the highest classification ability and classification accuracy and could effectively conduct clinical differential diagnosis for sarcoidosis and *tuberculosis*. Li et al. [30] used artificial neural network (ANN) to perform auxiliary diagnosis of DMD in children with rare leg neuromuscular disease based on magnetic resonance images (MRI) of patients, alleviating the pain caused by traditional diagnosis and detection schemes. Shanghai built a municipal data center in 2018 to share medical data with all 500 public hospitals, with about 16 million pieces of data stored in the core database every day. Liu et al. [31] analyzed the characteristics and content of medical records text, proposed a preprocessing method aiming at these characteristics, and applied this method to coronary heart disease dataset, and the effect of data analysis was significantly improved. Due to the problems of abnormal data, redundant data, and missing data in the original physical examination dataset, it cannot be directly used for data analysis and information mining of diseases. In order to make better use of valuable information in physical examination data, different preprocessing methods are proposed for different purposes: to reduce the time and space complexity of preprocessing, datasets are compressed. Liu et al. [32] realized the consistency and continuity of physical examination data over the years through data transformation based on linear function [34, 33].

From the above analysis, we know that the above methods have studied the intelligent processing and classification of multisource health big data to some extent;

some problem still exists [35, 36]. For example, no scholar has applied the models to this field from the perspective of physical and medical integration till now, so the research here is still a blank, which has great theoretical research and practical application value for intelligent processing and classification of multisource health big data. In addition, almost all classification models have shallow structure framework.

The contributions of this paper are as follows: (a) It introduces the basic theory of Bayesian network, including probability theory, basic principle of Bayesian network, Bayesian network learning, and common Bayesian network classifier. (b) The improved Bayesian network structure learning algorithm in Chapter 3 was used to construct the data classification model of hypothyroidism, and the performances of different Bayesian network classifiers were compared.

This paper consists of five parts. The first and second parts give the research status and background. The third part gives the processing and classification of multisource health big data. The fourth part shows the experimental results and analysis. The experimental results of this paper are introduced and compared and analyzed with relevant comparison algorithms followed. Finally, the fifth part concludes the paper.

3. Processing and Classification of Multisource Health Big Data

3.1. Perspective of Physical and Medical Integration. In fact, before the two terms of sports appeared or became specific nouns, our ancestors have long given us precious historical and cultural heritage: the longevity of the Traditional Chinese guidance method of health preservation represented by the Five Birds Opera and Eight Duan Brocade is the result of historical inheritance. In the new era, people begin to constantly improve their health needs, and sports and medicine are different levels of solutions around health needs.

Health is a complex and multidimensional concept, covering the concept of human physiology, psychology, society, and many other fields. As different branches of physiology, sports technology and medical technology have the same root but different application directions. Simply speaking, medical science is to guarantee the safety of human life, just like food and clothing in life. Solve the problem of human health and sports is the goal of life to a higher level; for example, a well-off standard of living in the life health is also a relatively vague definition; it is difficult to accurately define, having different embodiment in different areas.

Many industries all around health in modern society in the development, such as the primary side of the agriculture and animal husbandry and fisheries, life cannot leave the food processing, such as industrial equipment and quality inspection again. Environmental protection, from this point of view, as mentioned above, sports play a more prominent role in health, while medicine is only to solve the negative impact of disease on health in life, sports method is also more intuitive, and medicine? Without the help of drugs, I believe that the ability of doctors to heal the wounded and save the dying will immediately decline.

3.2. Multisource Health Big Data System. Firstly, this paper designs a multisource health big data management system as shown in Figure 1.

(1) The system uses Bluetooth, network, WIFI, and other technologies for intelligent collection, covering a number of medical and health items such as blood analysis, biochemical analysis, urine analysis, and ECG monitoring.

(2) There is no liquid path or pipeline in the Chinese medicine testing equipment of the system, and it has wireless and wired network automatic data upload function.

(3) Based on B/S and C/S framework structure, the system built a data uploading platform to realize accurate and stable uploading of all medical and health data and a smart signing mobile APP was launched to make it meet the basic public health service requirements.

Based on the computer network communication technology and network technology, the B/S framework structure is used to realize the integration of the detected equipment detection data, and the collected data will generate dynamic health records and connect with other hospital systems to realize the computer monitoring and automatic management of medical examination and test process. The dynamic full-process closed-loop health management mode combining offline and online is adopted to collect medical data in real-time offline and analyze and manage the data in all directions online to achieve prediction, prevention, and personalized health maintenance.

3.3. Feature Extraction Strategy. With the development of the medical big data diagnosis and treatment technology, the realization of a more efficient medical image analysis can be complementary to help the doctor condition analysis, help doctors to determine treatment plan, and reduce the dependence on clinical experience in the diagnosis of misjudgment rate. Therefore, high-efficiency and highly accurate medical diagnosis model can provide quantitative and objective endoscopy diagnosis for doctors. It makes it easier for clinicians to notice suspicious pathological images, reduces the workload of eye screening, and helps doctors to make correct clinical medical decisions.

In order to improve the accuracy of medical diagnosis model and effectively extract medical image features, this chapter proposes an image feature extraction algorithm with rotation invariance, which is named TriZ. TriZ algorithm is improved from image feature extraction algorithm HOG and generated by the HOG algorithm with 378,434 features. In this section, it is proved experimentally that this algorithm has rotation invariance for three gastric diseases, namely, gastric polyp, gastritis, and gastric ulcer, and can achieve effective detection and classification of gastric diseases in the case of 10-fold cross validation. TriZ's classification accuracy reached 87.0% among the four classification problems of the three types of gastric diseases and the healthy control. The specific research process of

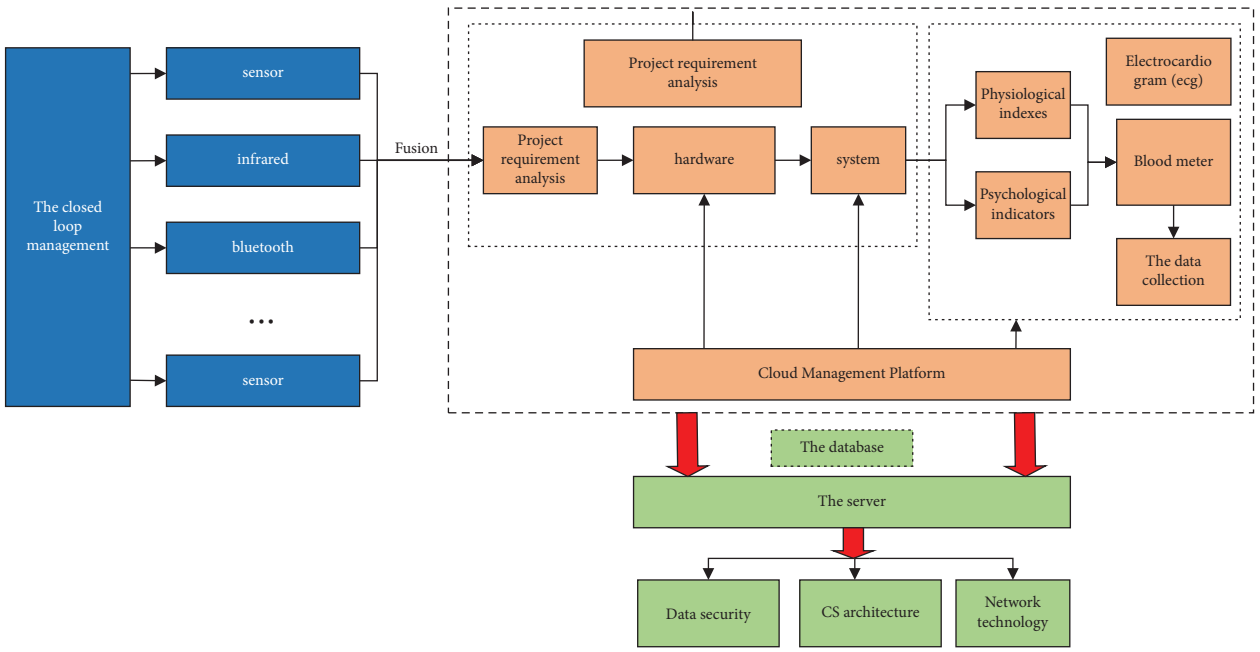


Figure 1: Multisource health big data system process.

extracting the diagnostic model of gastric diseases based on the medical image features of TriZ is shown in Figure 2.

3.4. Application of Naive Bayesian Network Classifier. The Primitive Bayesian network classifier adopts the Attribute Conditional independence assumption, which assumes that all nonclass attributes are independent of each other; that is, each attribute can independently affect the classification results, corresponding to the Bayesian network, each nonclass attribute node only has the category as its parent node, and then formula (1) is obtained.

$$P(F_1, F_2, \ldots, F_n, C) = P(C)\prod_{i=1}^{n} P(F_i|C). \tag{1}$$

In a similar way,

$$P(F_1, F_2, \ldots, F_n) = \prod_{i=1}^{n} P(F_i|\pi(F_i)) = \prod_{i=1}^{n} P(F_i), \tag{2}$$

where $F_1, F_2, \ldots, F_n$ denote the n variables and C is the center variable.

By integrating formulas (1) and (2), the posterior formula of calculation in naive Bayes classifier is

$$P(C|F_1, F_2, \ldots, F_n) = \frac{P(C)\prod_{i=1}^{n} P(F_i|C)}{\prod_{i=1}^{n} P(F_i)}. \tag{3}$$

In the above formula, $P(F_i)$ represents the attributes probability of F_i; $\prod_{i=1}^{n} P(F_i)$ is constant for every category. There are usually several attributes, so class attributes is the a posteriori probability of C and proportional to that of $P(C)\prod_{i=1}^{n} P(F_i|C)$; namely,

$$P(C|F_1, F_2, \ldots, F_n) \propto P(C)\prod_{i=1}^{n} P(F_i|C), \tag{4}$$

where $P(C)$ represents the prior probabilities of each category and $P(F_i|C)$ denotes the probability of occurrence of attribute F_i under the condition of known class C, which can be directly calculated by sample dataset.

Naive Bayesian network classifier is based on all the class attributes under the premise of mutual independence between complete classification tasks; although in real life it is often difficult to fully meet the conditions of datasets, there are still many researchers who use naive Bayesian network classifier as a kind of commonly used classification model. In this case, even if the dataset does not satisfy the conditional independence hypothesis, it still has good classification performance. Therefore, it is necessary to learn the tree structure between nonclass nodes, and the maximum weighted spanning tree is generally adopted. The weight between two nodes is expressed in conditional mutual information; its calculation formula is as follows:

$$I(X_i, X_j|C) = \sum_{x_i, x_j, c} P(x_i, x_j|c)\log\frac{P(x_i, x_j|c)}{P(x_i|c)P(x_j|c)}. \tag{5}$$

The main principle of ReliefF algorithm is as follows: firstly, a sample is randomly selected from the dataset as X, and then k samples closest to X are selected as H in the sample set of the same class as X according to Euclidean distance, and k samples closest to X are found in the sample set different from X, and then the above process is repeated M times according to formula (6) to update the weight of each feature and output the final weight of each feature:

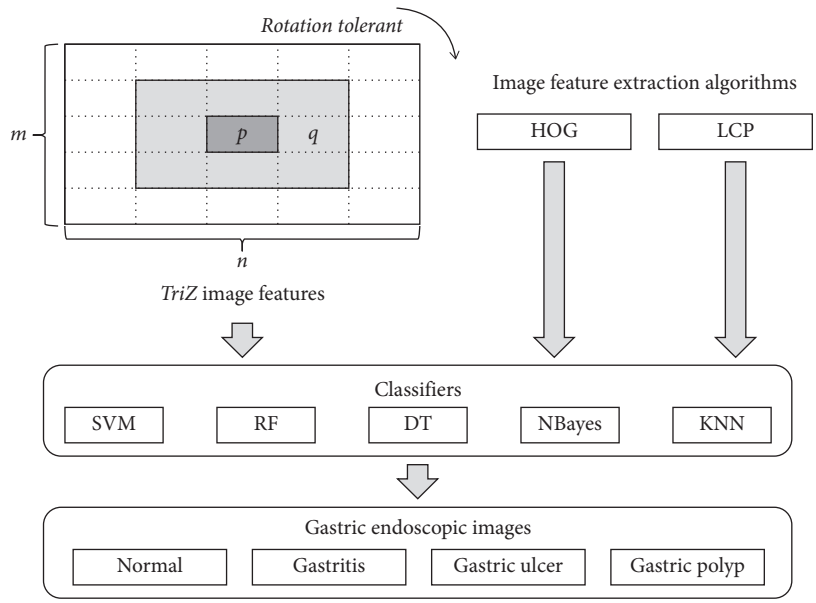


Figure 2: Feature extraction flow chart.

$$W(A) = \frac{\sum_{i=1}^{k} \operatorname{diff}(A, x, H_i)}{mk} + \frac{\sum_{C \in \operatorname{Class}(x)} \left[(P(C)/1 - P(\operatorname{Class}(x))) \sum_{i=1}^{k} \operatorname{diff}(A, x, M_i(C)) \right]}{mk}. \tag{6}$$

In the calculation of feature weight, ReliefF algorithm only considers the correlation between feature and class and ignores the possible redundancy between features. Therefore, this algorithm has certain limitations. In order to eliminate redundant attributes more effectively, this section introduces symmetric uncertainty in information theory based on ReliefF algorithm. Further eliminating redundant features, symmetrical uncertainty (SU) can measure the correlation between two variables. Suppose that the two variables are X and Y, and the formula for calculating the symmetrical uncertainty between two variables is as follows:

$$SU(X, Y) = \frac{2^* IG(X; Y)}{H(X) + H(Y)}, \tag{7}$$

where $H(X)$ and $H(Y)$, respectively, represent the information entropies of variables X and Y, and $H(X)$ is defined as follows:

$$H(X) = -\sum p(x) \log\, p(x), \tag{8}$$

where x represents different values of variable X and $IG(X; Y)$ represents information gain, also known as mutual information, which can be obtained by the following formula:

$$IG(X; Y) = H(X) - H(X|Y), \tag{9}$$

where $H(X|Y)$ denotes the conditional entropy of given variables X and Y, as defined below:

$$H(X|Y) = -\sum_{x \in X, y \in Y} p(x, y) \log\, p(x|y). \tag{10}$$

By synthesizing the above formulas, the symmetric uncertainty between variables X and Y can be obtained, and symmetry is full due to mutual information. It can be inferred that symmetric uncertainty is also symmetric, and, in order to make the magnitude of symmetric uncertainty comparable, it normalized the mutual information so that the symmetric uncertainty between features is between 0 and 1. When $SU(X, Y) = 0$, it means that variables X and Y are two independent variables; when $SU(X, Y) = 1$, it is indicated that variables X and Y are completely correlated.

$$IG(X; Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) = IG(Y; X). \tag{11}$$

Let X be the condition that satisfies Y_v; then variable Y_v satisfies the following equation:

$$p(Y_v|X, Y_u, u \neq v) = p(Y_v|X, Y_u, u:\ v), \tag{12}$$

where u and v represent two vertices contained in graph T, and then (X, Y) is a random field of strips.

4. Experimental Results and Analysis

4.1. Introduction to Experimental Environment and Dataset. The purpose of the experiment in this section is to test the effectiveness of the improved algorithm on liver disease detection data, which is mainly reflected from two aspects: the

efficiency of the algorithm execution and the detection accuracy of repeated data in the dataset. However, the actual data set accessed does not have a data mark about whether each data is a duplicate, so the performance of the improved algorithm cannot be tested with the original data set. Therefore, in order to measure the efficiency and scalability of the algorithm more comprehensively, the original numbers are standardized according to the centralized data in this paper, and 5000, 10000, and 20000 data points are distributed. The generation rules of repeated data detection for three datasets of different sizes are as follows: The data of each dimension of the original record is standardized. Each original data consists of 0–9 corresponding repeated pieces of data, and the number of repeated pieces of data follows Zipf distribution. Each repeated piece of data has 0–5 changes, and the similarity between the modified data and the original data is greater than or equal to the threshold value. The data in each dataset consists of two parts, 50% of which is the original data, and the other 50% is the repeated data modified according to the original data. All the models in this paper are coded by Python language, and all the experiments in this paper are carried out on a hardware device of NVIDIA 1080Ti GPU.

The standard to measure the performance of the repeated data detection algorithm is whether it is efficient and comprehensive to the repeated data detection in the dataset. According to the setting in this chapter, the detection of repeated data in the experiment in this chapter is essentially a binary algorithm, and the commonly used standards mainly include precision rate, recall rate, consumption time (Time), and AUC area.

4.2. Experimental Results Analysis. In order to verify the performance of classifier SVM, the best values of two parameters C and Gamma need to be filtered. The measured values of classification performance were calculated through 10-fold cross validation. Three heat maps were used to represent the data of four evaluation indicators: S_n, S_p, and Acc. The maximum, average, and minimum values of each measurement were represented in red, yellow, and blue, and the values of the color range were represented in gradient colors, as shown in Figure 3.

Parameter C is set to 20 with step size of 0.125 between 0.125 and 3.000, and parameter Gamma is set to {0.100, 0.178, 0.316, 0.562, 1.00, 1.334, 1.778} for grid search to find the best choice for these two parameters. The results show that when $C = 2.125$ and Gamma = 0.100, the SVM classifier is the best, and its classification accuracy is 97.2%. The algorithm integrates the morphological features of eyes and mouth in the face region and studies and discusses the fatigue detection problem from the aspects of feature number, classifier, and modeling parameters. The algorithm consists of three main steps. First, PCA algorithm is used to calculate the main components. Finally, the SVM model with RBF kernel is trained to classify the images. The experimental results show that the image recognition accuracy of this algorithm reaches 96.07%, and the operation time is only about 21 milliseconds, which can meet the requirements of real-time fatigue monitoring task with 30 frames per second.

In order to verify the stability of the proposed health big data classification method, we can choose the standard deep learning algorithm, whose data processing and parameter setting are roughly the same as the proposed algorithm. At the same time, all the above methods were cross-validated 10 times, and the average result of the test dataset is shown in Figure 4. It can be seen from Figure 4 that the robustness of the proposed method is best correlated with classification. It is worth noting that these experimental results were averaged over 20 times over 80000 datasets for more universality.

In order to verify the validity of the Bayesian network classifier based on the improved ReliefF algorithm, this classifier is compared with the BAN classifier based on ReliefF algorithm (ReliefF-BAN) and three other Bayesian network classifiers (NBC, TAN, and BAN).

Since ReliefF algorithm is needed to calculate the weights during the initial feature screening and the results of the initial feature screening with k larger than the initial value are selected, the difference of k will affect the feature subset finally obtained. If k is too large, for example, 28, or k is too small, it is likely to lead to the deletion of some features that are highly correlated with the class. In this section, $k = 27$, $k = 23$, and $k = 17$ are selected, respectively, as the preliminary screening results, and then further screening of feature subsets is completed according to different thresholds, and the performances of classifiers formed under different feature subsets are compared. The final results are shown in Figure 5.

In this section, Youden index is used to evaluate the Jorden index of each model under different proportions of labeled samples, as shown in Figure 6. After analyzing Figure 6, we can draw the following conclusion. When using the same base classifier to train the classification model, the Youden index of the optimized self-training model is higher than that of the standard self-training model and the supervised learning model. For example, taking naive Bayes as an example, the Youden index of the optimized self-training classification model is 58.90%, and the Youden index of the standard self-training classification model is 49.97%, while the Youden index is 48.84% when only naive Bayes algorithm is used for classification. This is because the self-training algorithm after optimization can learn more unlabeled sample information, and the information learned through repeated labeling strategy is more accurate. Therefore, the comprehensive performance of the optimization algorithm is better, which also proves the effectiveness of the algorithm.

Due to the introduction of mislabeled samples, the Youden index of the standard self-training classification model is not necessarily higher than that of the supervised learning classification model. For example, taking decision tree as an example, the Youden index of the standard self-training classification model is 53.99%. The Jorden index of decision tree classification model is 54.97%, and the comprehensive performance of decision tree classification model is better than that of standard self-training classification model. This illustrates the instability of the standard self-training algorithm.

In conclusion, we can know that the supervised classification algorithm performs well in the results that the training algorithm for standard test data classification does not, which may be due to the low classification performance of the base classifier when selecting unlabeled samples, which leads to the continuous accumulation of errors and weakens the

Sn		C																							
		0.125	0.250	0.375	0.500	0.625	0.750	0.875	1.000	1.125	1.250	1.375	1.500	1.625	1.750	1.875	2.000	2.125	2.250	2.375	2.500	2.625	2.750	2.875	3.000
Gamma	0.100	0.9946	0.9911	0.9902	0.9920	0.9911	0.9902	0.9911	0.9911	0.9911	0.9884	0.9884	0.9875	0.9875	0.9875	0.9875	0.9866	0.9875	0.9875	0.9866	0.9866	0.9866	0.9866	0.9857	0.9848
	0.178	0.9991	0.9929	0.9929	0.9911	0.9902	0.9911	0.9902	0.9911	0.9902	0.9902	0.9902	0.9875	0.9875	0.9848	0.9848	0.9839	0.9821	0.9804	0.9804	0.9795	0.9786	0.9786	0.9777	0.9768
	0.316	1.0000	0.9982	0.9973	0.9946	0.9938	0.9938	0.9920	0.9893	0.9857	0.9830	0.9830	0.9821	0.9821	0.9804	0.9795	0.9795	0.9804	0.9804	0.9795	0.9795	0.9795	0.9786	0.9786	0.9777
	0.562	1.0000	1.0000	0.9982	0.9982	0.9982	0.9964	0.9946	0.9946	0.9938	0.9929	0.9911	0.9911	0.9902	0.9902	0.9893	0.9893	0.9884	0.9884	0.9884	0.9884	0.9884	0.9884	0.9884	0.9884
	1.000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9991	0.9982	0.9973	0.9973	0.9973	0.9964	0.9964	0.9964	0.9964	0.9964	0.9964	0.9964	0.9964	0.9964	0.9964	0.9964	0.9964	0.9964	0.9964
	1.334	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9991	0.9982	0.9982	0.9982	0.9982	0.9982	0.9982	0.9982	0.9982	0.9982	0.9982	0.9982	0.9982	0.9982	0.9982	0.9982	0.9982
	1.778	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9991	0.9991	0.9991	0.9991	0.9991	0.9991	0.9991	0.9991	0.9991	0.9991	0.9991	0.9991	0.9991	0.9991	0.9991

Sp		C																							
		0.125	0.250	0.375	0.500	0.625	0.750	0.875	1.000	1.125	1.250	1.375	1.500	1.625	1.750	1.875	2.000	2.125	2.250	2.375	2.500	2.625	2.750	2.875	3.000
Gamma	0.100	0.632	0.766	0.818	0.852	0.864	0.877	0.9911	0.889	0.891	0.896	0.9884	0.895	0.895	0.896	0.898	0.900	0.907	0.907	0.904	0.905	0.904	0.904	0.904	0.904
	0.178	0.4196	0.6875	0.7750	0.8161	0.8446	0.8554	0.9902	0.8732	0.8750	0.8804	0.9902	0.8875	0.8893	0.8911	0.8911	0.8929	0.8929	0.8911	0.8911	0.8911	0.8911	0.8929	0.8946	0.8982
	0.316	0.1339	0.4518	0.6036	0.6875	0.7571	0.7804	0.9920	0.8179	0.8286	0.8393	0.9830	0.8500	0.8618	0.8500	0.8518	0.8518	0.8518	0.8536	0.8536	0.8536	0.8554	0.8536	0.8554	0.8554
	0.562	0.0000	0.1643	0.3250	0.4500	0.5393	0.5911	0.9946	0.6750	0.6964	0.7089	0.9911	0.7143	0.7143	0.7143	0.7161	0.7161	0.7179	0.7196	0.7196	0.7196	0.7196	0.7196	0.7196	0.7196
	1.000	0.0000	0.0000	0.0500	0.1357	0.2500	0.3196	0.9982	0.4179	0.4518	0.4768	0.9964	0.4821	0.4821	0.4821	0.4821	0.4821	0.4821	0.4821	0.4821	0.4821	0.4821	0.4821	0.4821	0.4821
	1.334	0.0000	0.0000	0.0036	0.0321	0.0946	0.1732	1.0000	0.2786	0.3268	0.3554	0.9982	0.3554	0.3554	0.3554	0.3554	0.3554	0.3554	0.3554	0.3554	0.3554	0.3554	0.3554	0.3554	0.3554
	1.778	0.000	0.000	0.000	0.009	0.039	0.063	1.0000	0.129	0.171	0.200	0.9991	0.204	0.204	0.204	0.204	0.204	0.204	0.204	0.204	0.204	0.204	0.204	0.204	0.204

Acc		C																							
		0.125	0.250	0.375	0.500	0.625	0.750	0.875	1.000	1.125	1.250	1.375	1.500	1.625	1.750	1.875	2.000	2.125	2.250	2.375	2.500	2.625	2.750	2.875	3.000
Gamma	0.100	0.874	0.916	0.933	0.945	0.949	0.952	0.955	0.957	0.958	0.958	0.958	0.957	0.957	0.957	0.958	0.958	0.961	0.961	0.959	0.960	0.959	0.959	0.958	0.958
	0.178	0.8060	0.8911	0.9202	0.9327	0.9417	0.9458	0.9470	0.9518	0.9518	0.9536	0.9548	0.9542	0.9548	0.9546	0.9546	0.9546	0.9524	0.9506	0.9506	0.9500	0.9494	0.9500	0.9500	0.9500
	0.316	0.7113	0.8161	0.8661	0.8923	0.9149	0.9226	0.9280	0.9321	0.9333	0.9351	0.9369	0.9381	0.9387	0.9369	0.9369	0.9369	0.9375	0.9381	0.9375	0.9375	0.9381	0.9369	0.9375	0.9369
	0.562	0.667	0.7214	0.7738	0.8155	0.8432	0.8613	0.8750	0.8881	0.8946	0.8982	0.8988	0.8988	0.8982	0.8982	0.8982	0.8982	0.8982	0.8988	0.8988	0.8988	0.8988	0.8988	0.8988	0.8988
	1.000	0.667	0.6667	0.6833	0.7119	0.7500	0.7726	0.7923	0.8042	0.8155	0.8238	0.8250	0.8250	0.8250	0.8250	0.8250	0.8250	0.8250	0.8250	0.8250	0.8250	0.8250	0.8250	0.8250	0.8250
	1.334	0.667	0.6667	0.6679	0.6774	0.6982	0.7244	0.7405	0.7589	0.7744	0.7839	0.7839	0.7839	0.7839	0.7839	0.7839	0.7839	0.7839	0.7839	0.7839	0.7839	0.7839	0.7839	0.7839	0.7839
	1.778	0.667	0.667	0.667	0.670	0.680	0.688	0.695	0.710	0.724	0.733	0.734	0.734	0.734	0.734	0.734	0.734	0.734	0.734	0.734	0.734	0.734	0.734	0.734	0.734

Figure 3: Optimization of the heat maps of parameters C and Gamma of SVM.

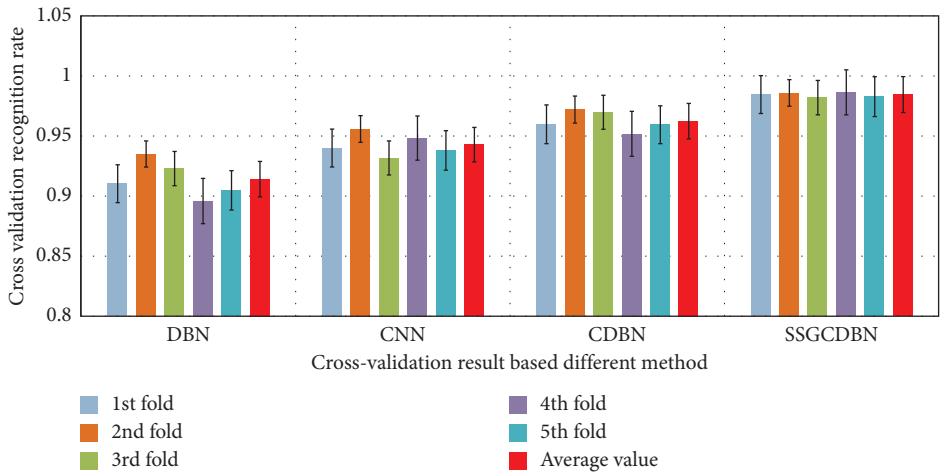


Figure 4: Cross validation based on different classification methods.

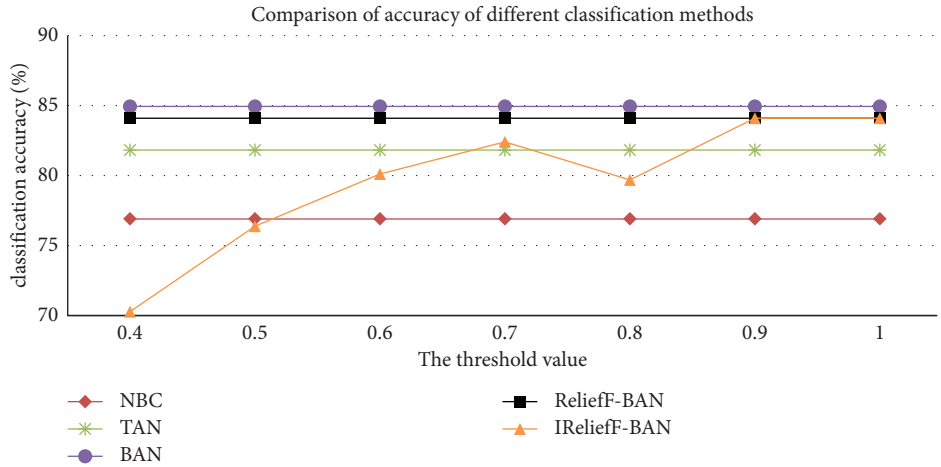


Figure 5: Comparison of the accuracies of different classifiers when $k = 27$.

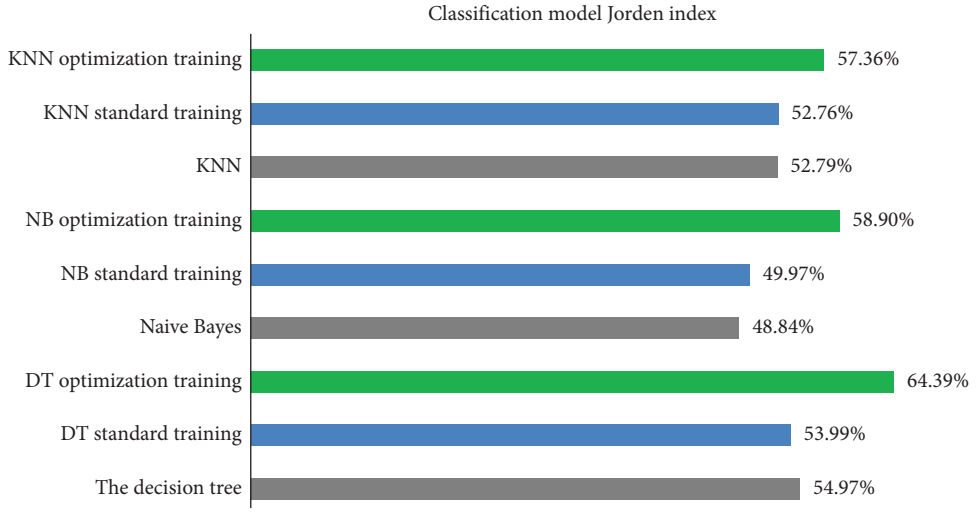


Figure 6: Comparison diagram of Jorden index for classification model.

R2	LR	SVC	KNN	GBDT	DT	RFC	GaussianNB	BernoulliNB	SGDC	PAC
MI	−0.7073	−0.0054	−0.6503	−0.1477	−0.8163	−0.3468	−0.8448	−0.8495	−1.5182	−1.0724
Ftest	−0.2804	0.0136	−0.4701	−0.0244	−0.7547	−0.2520	−0.7452	−0.7310	−1.2858	−0.8638
Ttest	−0.6456	−0.0291	−0.6219	−0.2994	−1.0392	−0.3563	−1.2052	−0.0670	−0.3279	−0.4322
PearsonR	−0.5413	−0.0054	−0.4701	−0.0718	−0.8638	−0.2899	−0.8068	−0.7736	−0.7310	−0.6930
MINE	−0.6883	0.0088	−0.4275	−0.1097	−0.9633	−0.3658	−0.7262	−0.6788	−1.1246	−1.1104
PSR	−0.5413	−0.0054	−0.4701	−0.0338	−0.8638	−0.2899	−0.8068	−0.7736	−0.7310	−0.6930
RankSum	−0.6409	−0.0386	−0.6456	−0.2757	−0.9586	−0.4512	−1.2005	−0.0623	−0.5602	−0.4831
LrRFE	0.2080	−0.0101	−0.5318	−0.2330	−1.0013	−0.4559	−1.1056	−1.2811	−0.6409	−0.3848
LassoRFE	−0.2852	−0.0101	−0.6314	−0.1382	−0.8400	−0.4227	−1.1198	−1.3190	−0.9112	−0.6029
SgdcRFE	−0.0907	−0.0101	−0.6503	−0.1809	−1.0155	−0.3468	−1.0108	−1.2194	−0.3658	−0.3136

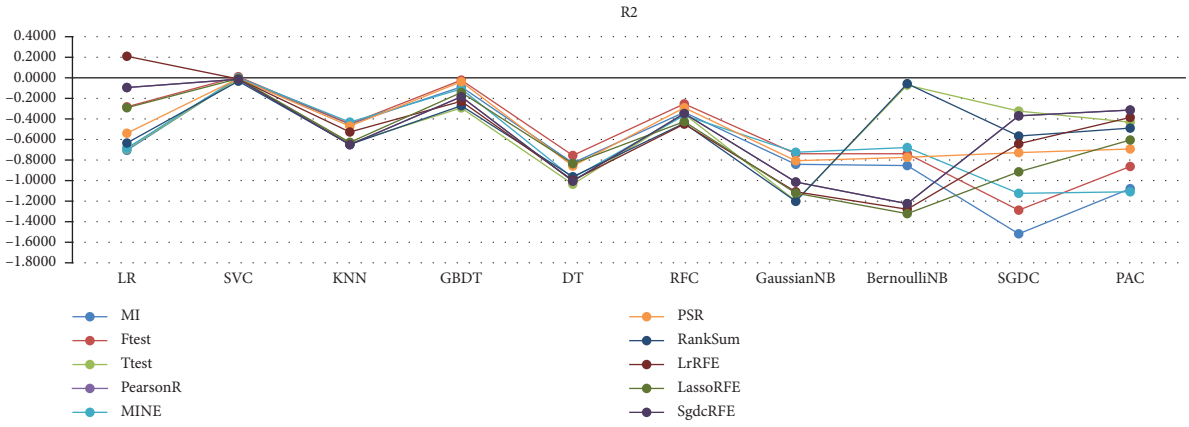


Figure 7: Comparison with the existing feature selection algorithms.

performance of the classifier. However, the classification performance of the self-training algorithm after optimization is generally superior to those of the supervised algorithm and the standard self-training algorithm, which proves the effectiveness of the optimization algorithm.

Figure 7 shows the comparison between the proposed regression biomarker detection algorithm and the 10 existing feature selection algorithms. The figure shows that 10 algorithms of R2 evaluation index calculate the classification accuracy of each feature subset under cross validation for 510 times and mark out the maximum accuracy. The horizontal axis lists the names of 10 classification algorithms.

As shown in Figure 8, the red regular triangle scatter points represent students with excellent physical fitness. At least two of the three physical test datasets of such students are excellent or good. The yellow regular triangle scatter points represent students with average physical fitness. There are few excellent blue inverted triangles in the data, indicating the students with poor physical fitness. The majority of these students are medium and unqualified in the three

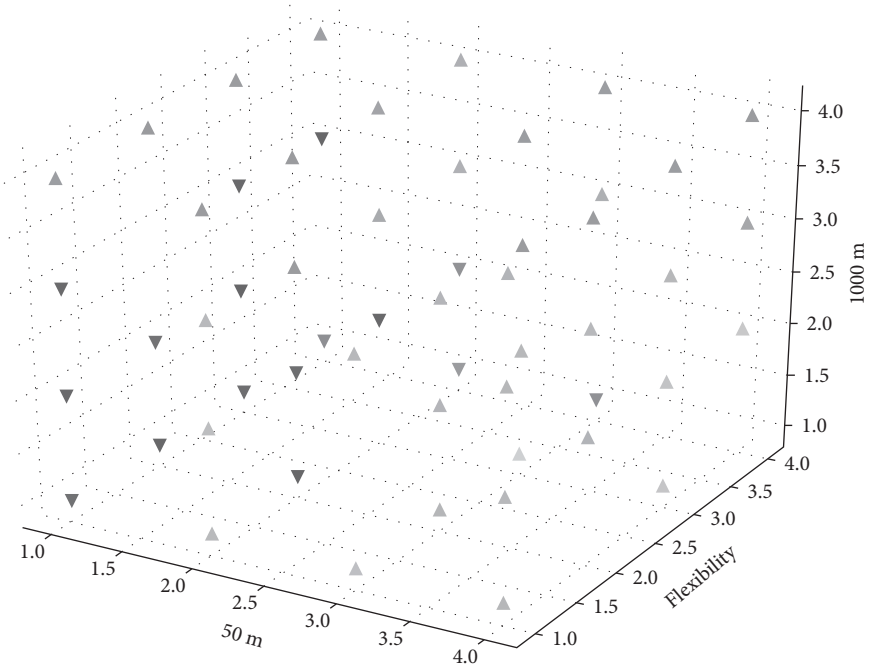


Figure 8: Health data clustering results.

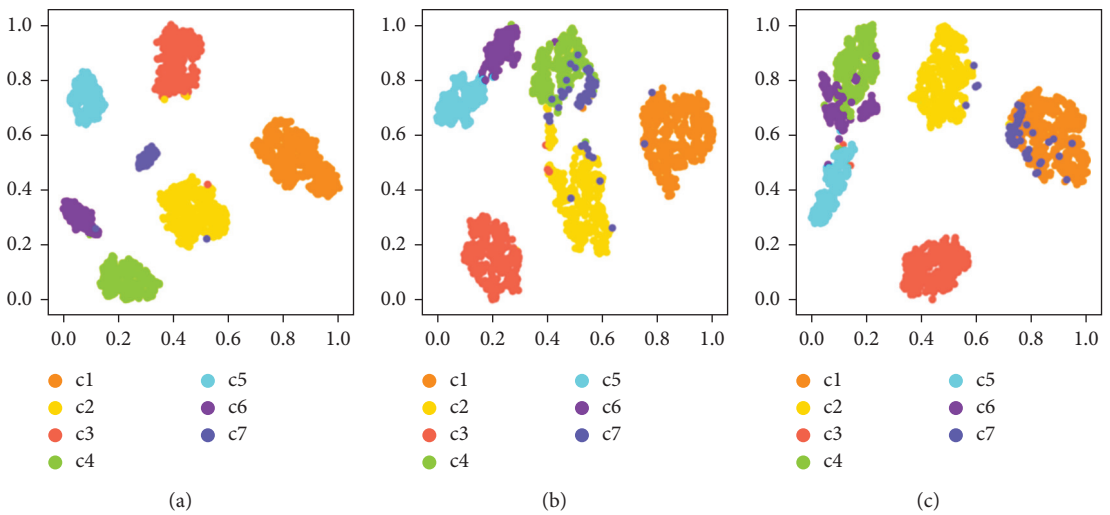


Figure 9: Projection results of the proposed method on different datasets.

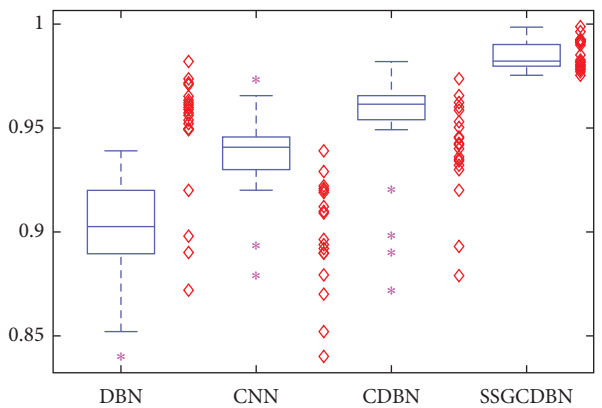


Figure 10: Boxplot of different health big data classification methods.

physical test datasets. Although one of them may be good or excellent, the overall physical fitness needs to be improved.

Figure 9 shows the projection results of the method in this paper on different data sets, including different categories represented by C. The three data sets of each category are classified into seven. It can be seen from the figure that the method proposed in this paper has good classification results on the three data sets, especially in data set 1, which shows that this method can deal with the big health problem of big data. The superiority of the proposed method is proved. In order to better prove the classification effect of the proposed method, 8000 datasets were processed and classified 20 times on average. Figure 8 shows the boxplot and scatter distribution of 20 mean diagnostic results of test samples under different models. As can be seen from Figure 10, the classification performance of the algorithm proposed in this paper is more stable and the classification accuracy is the highest.

5. Conclusion

In this paper, the relevant theories of Bayesian network are studied, and the classifier based on Bayesian network is applied to the data of hypothyroidism. Aiming at the key technologies needed in the application process, the improved ideas on methods are proposed and the specific contents are as follows.

The improved Bayesian network learning algorithm is applied to the classification of hypothyroidism. Firstly, the dataset of hypothyroidism is preprocessed to make it conform to the calculation requirements of the algorithm. Then, four Bayesian network classifiers are constructed for the preprocessed data, namely, naive Bayesian classifier (NBC), TAN classifier, BAN classifier, and Bayesian multi-network classifier. The network structures of different classifiers meet different degrees of dependence. Finally, BAN classifier was found to have the best effect on the classification of hypothyroidism data.

When diversity enters the transition stage, the combination operator method of fast convergence rate and competitive cross mutation can form good species quickly, but when diversity enters the mutation stage, it will be less. In order to avoid population convergence to local optimum, high-precision genetic combination operator and dynamic mutation rate method are used in this stage. Finally, experiments prove that the network structure of the improved algorithm is better.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

The author acknowledges General Project of Hunan Philosophy and Social Science Foundation in 2017, Research on the integrated health service model of human medicine for the elderly in poor areas under the background of targeted Poverty Alleviation (no. 17YBA063).

References

[1] J. Sun, H. Wang, Z. Song, J. Lu, P. Meng, and S. Qin, "Mapping essential urban land use categories in Nanjing by integrating multi-source big data," *Remote Sensing*, vol. 12, no. 15, Article ID 2386, 2020.

[2] I. D. Dinov, B. Heavner, M. Tang et al., "Predictive big data analytics: a study of Parkinson's disease using large, complex, heterogeneous, incongruent, multi-source and incomplete observations," *PLoS one*, vol. 11, no. 8, Article ID e0157077, 2016.

[3] F. Li, F. Li, S. Li, and Y. Long, "Deciphering the recreational use of urban parks: experiments using multi-source big data for all Chinese cities," *The Science of the Total Environment*, vol. 701, Article ID 134896, 2020.

[4] F. Lyu and L. Zhang, "Using multi-source big data to understand the factors affecting urban park use in Wuhan," *Urban Forestry and Urban Greening*, vol. 43, Article ID 126367, 2019.

[5] N. Niu, X. Liu, H. Jin et al., "Integrating multi-source big data to infer building functions," *International Journal of Geographical Information Science*, vol. 31, no. 9, pp. 1871–1890, 2017.

[6] J. Zhang, C. Li, Z. Sun, Z. Luo, C. Zhou, and S. Li, "Towards a unified multi-source-based optimization framework for multi-label learning," *Applied Soft Computing*, vol. 76, pp. 425–435, 2019.

[7] Y. Tu, B. Chen, W. Lang et al., "Uncovering the nature of urban land use composition using multi-source open big data with ensemble learning," *Remote Sensing*, vol. 13, no. 21, Article ID 4241, 2021.

[8] W. Hu, "On legal English translation from the perspective of legal linguistics," *Review of Educational Theory*, vol. 2, no. 3, pp. 6–10, 2019.

[9] X. Liu, N. Niu, X. Liu et al., "Characterizing mixed-use buildings based on multi-source big data," *International Journal of Geographical Information Science*, vol. 32, no. 4, pp. 738–756, 2018.

[10] P. Tschandl, C. Rosendahl, and H. Kittler, "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," *Scientific Data*, vol. 5, no. 1, pp. 1–9, 2018.

[11] X. Guo, H. Chen, and X. Yang, "An evaluation of street dynamic vitality and its influential factors based on multi-source big data," *ISPRS International Journal of Geo-Information*, vol. 10, no. 3, p. 143, 2021.

[12] D. Viorela-Valentina, "Translation practice–A means for enhancing student employability," *Dialogos*, vol. 22, no. 38, 215 pages, 2021.

[13] J. Prince, F. Andreotti, and M. De Vos, "Multi-source ensemble learning for the remote prediction of Parkinson's disease in the presence of source-wise missing data," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 5, pp. 1402–1411, 2018.

[14] A. R. Martinez, "Classification of covid-19 in ct scans using multi-source transfer learning," arXiv preprint http://arXiv.org/abs/2009.10474, 2020.

[15] Y. Zhang, Q. Li, W. Tu, K. Mai, Y. Yao, and Y. Chen, "Functional urban land use recognition integrating multi-source geospatial

data and cross-correlations," *Computers, Environment and Urban Systems*, vol. 78, Article ID 101374, 2019.
[16] P. Zhang, T. Li, G. Wang et al., "Multi-source information fusion based on rough set theory: a review," *Information Fusion*, vol. 68, pp. 85–117, 2021.
[17] E. M. Lalitha and L. Satish, "Wavelet analysis for classification of multi-source PD patterns," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 7, no. 1, pp. 40–47, 2000.
[18] T. Niu, Y. Chen, and Y. Yuan, "Measuring urban poverty using multi-source data and a random forest algorithm: a case study in Guangzhou," *Sustainable Cities and Society*, vol. 54, Article ID 102014, 2020.
[19] L. Zong, S. He, J. Lian et al., "Detailed mapping of urban land use based on multi-source data: a case study of lanzhou," *Remote Sensing*, vol. 12, no. 12, p. 1987, 2020.
[20] X. He, Y. Cao, and C. Zhou, "Evaluation of polycentric spatial structure in the urban agglomeration of the pearl river delta (PRD) based on multi-source big data fusion," *Remote Sensing*, vol. 13, no. 18, Article ID 3639, 2021.
[21] L. Yuan, Y. Wang, P. M. Thompson, V. A. Narayan, and J. Ye, "Multi-source feature learning for joint analysis of incomplete multiple heterogeneous neuroimaging data," *NeuroImage*, vol. 61, no. 3, pp. 622–632, 2012.
[22] H. Yang, M. Fu, L. Wang, and F. Tang, "Mixed land use evaluation and its impact on housing prices in beijing based on multi-source big data," *Land*, vol. 10, no. 10, Article ID 1103, 2021.
[23] H. Kuai, N. Zhong, J. Chen et al., "Multi-source brain computing with systematic fusion for smart health," *Information Fusion*, vol. 75, pp. 150–167, 2021.
[24] Y. Guo, C. Yin, M. Li, X. Ren, and P. Liu, "Mobile e-commerce recommendation system based on multi-source information fusion for sustainable e-business," *Sustainability*, vol. 10, no. 1, p. 147, 2018.
[25] X. Xu, H. Peng, M. Z. A. Bhuiyan et al., "Privacy-preserving federated depression detection from multi-source mobile health data," *IEEE Transactions on Industrial Informatics*, vol. 1, 2021.
[26] X. Zhou, Z. Guan, J. Xi, and G. Wei, "Public transportation operational health assessment based on multi-source data," *Applied Sciences*, vol. 11, no. 22, Article ID 10611, 2021.
[27] H. Wang, "Marine environment salinity measurement based on data classification system and features of business English translation," *Arabian Journal of Geosciences*, vol. 14, no. 15, pp. 1–14, 2021.
[28] S. Jiang, X. Qian, T. Mei, and Y. Fu, "Personalized travel sequence recommendation on multi-source big social media," *IEEE Transactions on Big Data*, vol. 2, no. 1, pp. 43–56, 2016.
[29] C. Qiu, M. Schmitt, L. Mou, P. Ghamisi, and X. Zhu, "Feature importance analysis for local climate zone classification using a residual convolutional neural network with multi-source datasets," *Remote Sensing*, vol. 10, no. 10, Article ID 1572, 2018.
[30] Y. Li, G. Wen, Y. Hu et al., "Multi-source Seq2seq guided by knowledge for Chinese healthcare consultation," *Journal of Biomedical Informatics*, vol. 117, Article ID 103727, 2021.
[31] S. Liu, L. Zhang, Y. Long, Y. Long, and M. Xu, "A new urban vitality analysis and evaluation framework based on human activity modeling using multi-source big data," *ISPRS International Journal of Geo-Information*, vol. 9, no. 11, p. 617, 2020.
[32] K. Liu, Y. Feng, and X. Xue, "Fault diagnosis of hydraulic retraction system based on multi-source signals feature fusion and health assessment for the actuator," *Journal of Intelligent and Fuzzy Systems*, vol. 34, no. 6, pp. 3635–3649, 2018.
[33] H. Chen, L. Huang, L. Yang, Y. Chen, and J. Huang, "Model-based method with nonlinear ultrasonic system identification for mechanical structural health assessment," *Transactions on Emerging Telecommunications Technologies*, vol. 31, no. 12, Article ID e3955, 2020.
[34] A. B. Stella, M. AjČeviĆ, G. Furlanis et al., "Smart technology for physical activity and health assessment during COVID-19 lockdown," *The Journal of Sports Medicine and Physical Fitness*, vol. 61, no. 3, pp. 452–460, 2021.
[35] S. Fatima, O. Schieir, M. F. Valois et al., "Health assessment questionnaire at one year predicts all-cause mortality in patients with early rheumatoid arthritis," *Arthritis & Rheumatology*, vol. 73, no. 2, pp. 197–202, 2021.
[36] W. Cheng, H. Xi, C. Sindikubwabo et al., "Ecosystem health assessment of desert nature reserve with entropy weight and fuzzy mathematics methods: a case study of Badain Jaran Desert," *Ecological Indicators*, vol. 119, Article ID 106843, 2020.

Hindawi
Scientific Programming
Volume 2022, Article ID 8264462, 9 pages
https://doi.org/10.1155/2022/8264462

Research Article

Research on Rapid Dynamic Rendering and Modeling Technology of Water Landscape Based on 3D Image Technology

Li Sun

School of Art, Anhui University of Finance & Economics, Bengbu, Anhui 233000, China

Correspondence should be addressed to Li Sun; 120081141@aufe.edu.cn

Received 17 December 2021; Revised 12 January 2022; Accepted 17 January 2022; Published 28 February 2022

Academic Editor: Baiyuan Ding

With the development of computer graphics technology, the application of 3D visualization simulation technology has become increasingly widespread. This paper develops a three-dimensional visualization simulation system for water bodies by studying the general model of regional landscape space and water quality purification. This paper mainly does research on rapid dynamic rendering and modeling technology of water landscape based on 3D image technology and uses the 3ds Max Material Editor to map the enhanced 3D landscape design to complete the design, and the designed landscape feature data will be saved in the system data statistics unit for later query. This article also expounds the entire process of alternative simulation construction and confirms the feasibility and superiority of alternative modeling methods through multiple case comparisons. Alternative modeling methods are used as the main way to construct waterfall scenes, that is, simple models are combined with texture mapping, particle effects, and other forms to replace complex models, while ensuring the simulation effect, and the complex models are presented in the most simplified way to make the virtual scene. It is well performed in terms of authenticity and real time. The research results show that the designed system design renderings are extremely clear and of good quality; they have a high score in the water landscape design. Moreover, the system application cost is low, energy consumption is low, and operation efficiency is high.

1. Introduction

Water body is a common natural scene in real life. Simulation research on water body has important theoretical significance and practical value in many fields such as landscape simulation, online games, film and television special effects, and water conservancy engineering. Due to the irregular shape and random variability of water bodies, water body simulation has become a research difficulty and hotspot in the simulation watershed. At the same time, it has also aroused the interest of countless researchers. The water body is simulated. However, due to the limitations of computer hardware equipment, there is still a contradiction between the authenticity and real-time nature of water simulation. Fish and bear's paw cannot be achieved at the same time. This requires a balance between authenticity and real time [1–5]. The 3D image technology is shown in Figure 1.

The application field of water body simulation requires simulation effects so that water body simulation has three characteristics: authenticity, real time, and interactivity. Authenticity refers not only to similarity in form, but also to having a sense of reality in terms of object features. At present, particle systems and fluid dynamics are mainly used to express the shape of water bodies, and rendering techniques such as environment mapping and ray tracing are used to realize the refraction of light by water bodies. In terms of real-time performance, technologies such as GPU hardware acceleration and LOD are used to increase the rendering speed and ensure the real-time performance of water body simulation. Interactivity refers to the dynamic effects of the interaction between the water body and other objects, such as waves caused by the ship when sailing and floods hitting buildings. Authenticity and real time are the goals that water simulation research and development are constantly pursuing, but these two characteristics have very high requirements on computer hardware. When pursuing high real-time effects, realism is often sacrificed. Water simulation modeling methods are controlled by the

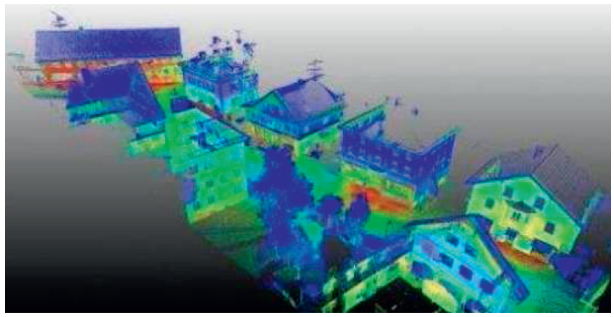


Figure 1: 3D image technology.

development of computer software and hardware. In the early days, texture mapping and parameter modeling were generally used to achieve simulation effects. The development of GPU technology enabled the use of particle systems, fluid dynamics, and other methods. The authenticity of water simulation effects, real time, and interactivity are also gradually increasing. At present, the common water simulation performance methods include texture mapping, surface parameters, particle systems, fluid dynamics, and wave functions. With the continuous improvement of water body simulation requirements, modeling methods tend to be integrated with multiple methods, and comprehensive treatment is adopted to complement them in order to achieve better water body simulation results [6–12].

Compared with other landscapes, water landscapes are more distinctive. They not only meet the objective requirements of human aesthetics, but also have aesthetics and spiritual contributions. Water landscape is the product of the combination of human aesthetic standards, culture, and water [13–18]. It is the evolution product of human civilization that perfectly integrates human mythology, literature, and other conceptual forms and tangible entities such as architecture and ornaments with the water landscape of a certain designated area. People use the principle of VISC conversion of computer graphics to realize the conversion of data to graphics, display the graphics on the screen, and finally meet their needs through interactive processing. From a long-term perspective, people will still meet their different needs from different data structures, algorithms, and mathematical theories, and they will also get a variety of three-dimensional visual models according to different computer software and hardware configurations. These calculated accurate models will be put into actual production. Regarding the actual requirements for the visualization of water bodies in the water system, the water quality model can be used to simulate the required water bodies [19–21].

Traditional methods usually use virtual design methods to improve the authenticity of water design works. However, the virtual design method has disadvantages such as large data collection error and picture distortion. With the rapid development of electronic technology, electronic drawing and computer visualization technology are gradually optimized. Since the 1960s, two-dimensional drawing has evolved into the current three-dimensional virtual simulation method [22–26]. The three-dimensional virtual simulation method can present the actual environment with high precision. It has important application value in the fields of medicine, urban planning, design, and manufacturing. Therefore, this article applies 3D image processing technology to landscape design, designing a landscape design system based on 3D image processing technology [27–33].

2. Water Landscape Design System Based on 3D Image Processing Technology

2.1. Hardware Design. Virtual reality technology is a computer-generated three-dimensional digital model of a scene model that exists or does not exist in reality and imports the virtual reality technology into the virtual reality software to generate a visual, auditory, and tactile model that can be viewed from all angles. The modeling technology of water landscape based on 3D image technology is shown in Figure 2, which is mainly used to store garden design examples, plant picture collections, plant graphic symbol libraries, and scenery picture collections, which are used to provide designers with landscape design materials. In the landscape planning of the model scene design unit, the scene design is mainly completed through the three-dimensional model. The structure of the three-dimensional model is shown in Figure 3. This module mainly contains the architectural landscape point elements, which are directly produced using ArcGIS 10.2. First, construct a three-dimensional geometric model, and introduce the basic building information and the two-dimensional CAD distribution map. Set real textures to the buildings to achieve a landscape design close to reality. The 3D landscape image can be preprocessed during design. Finally, the 3ds MAX material editor is used to the enhanced 3D landscape design drawing for texture mapping. The feature data such as landscape name, basic functions, land area, and elevation data are all optimized with ArcMap to construct a feature database for landscape design.

2.2. 3D Image Processing. The rendering of model colors in landscape 3D design must be implemented based on real landscape images. Decompose the covariance of L_w according to the relevant theory:

$$D_{L_w} = F\left[L_w^T\right] = \mathrm{FEF}^T, \tag{1}$$

where E represents the diagonal matrix of eigenvalues and F represents the orthogonal matrix. Then, the linear whitening matrix is

$$U = \frac{F^T}{\sqrt{E}}. \tag{2}$$

Through the above analysis, s covariance matrices can be obtained:

$$\begin{aligned} D_s &= F\left\{ss^T\right\} \\ &= UF\left\{L_w L_w^T\right\}U^T. \end{aligned} \tag{3}$$

In order to overcome the shortcoming, that is, the 3D image base is too difficult to roughly increase the computational difficulty, the 3D image features must be processed by dimensionality reduction. Distribute the eigenvalues in

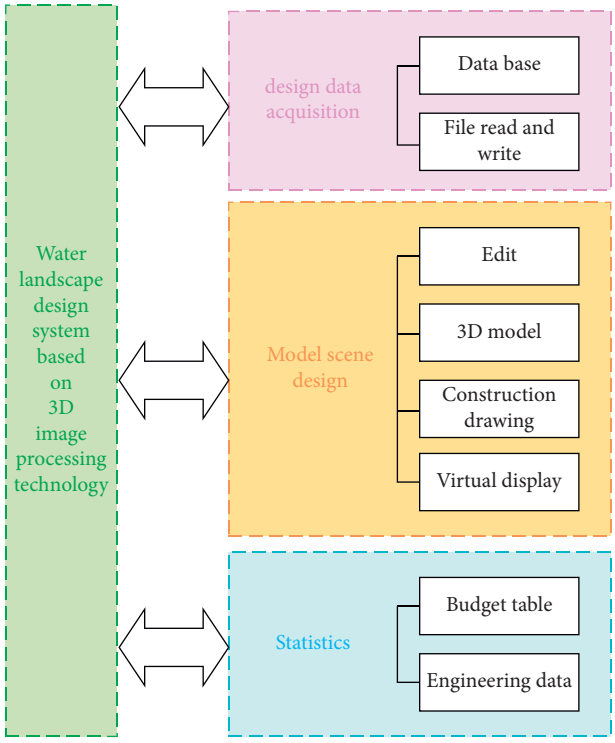


Figure 2: The landscape design system structure of 3D image processing technology.

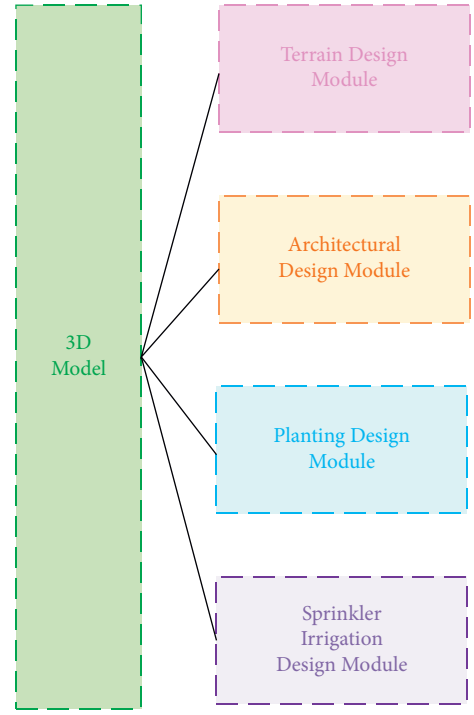


Figure 3: The structure of the three-dimensional model.

descending order, and select eigenvalues with large values to obtain the reduced-dimensional diagonal matrix; then,

$$U = \frac{1}{\sqrt{\overline{E}D^T}}, \tag{4}$$

where D describes the variance and U represents the 3D landscape image after dimensionality reduction. After the above operations, the observation vector of the 3D landscape image is

$$s(t) = \left[s_1(t), s_2(t), \ldots, s_a(t)\right]^T. \tag{5}$$

The above formula conforms to

$$s(t) = \widehat{U}L_w(t). \tag{6}$$

Then, the statistics of the independent components of the 3D landscape image are

$$z_j(t) = \left[\beta_{j1}, \beta_{j2} \ldots, \beta_{ja}\right] \begin{bmatrix} s_1(t) \\ s_2(t) \\ \cdots \\ s_a(t) \end{bmatrix}. \tag{7}$$

In the formula, βja describes the independent components of the 3D landscape image. In summary, selecting a reasonable image base for 3D landscape images can minimize the amount of calculation while not causing 3D image distortion. The auxiliary virtual environment gives people a sense of psychological reality and can directly act on all objects in this virtual environment. The human-computer interaction functions of virtual reality technology provide people with a platform for experience and simulation training anytime and anywhere. First of all, it plays a pivotal role in simulation training in the military and aerospace fields. The evaluated x and y are shown in Figure 4.

After the 3D landscape image is preprocessed, when enhancing the 3D landscape image, the 3D landscape image is first enhanced in the frequency domain, and the pixel grayscale of the image is converted into a new image grayscale based on a fixed function to obtain the image grayscale feature. Obtain the enhancement function of the gray contrast of the edge of the image, and realize the enhancement of the 3D landscape image features. Suppose $g''(a,b)$ is used to describe the preprocessed 3D landscape image, y_{jui}' describes the neighborhood of random pixels d_{sg}' in the 3D landscape image, v_{wer}' describes the number of values of each image feature, and c_{wer}' describes the total value of the number of image pixels; then, use the following equation to enhance the image in the frequency domain:

$$K_{\text{qwu}}' = \frac{c_{\text{wep}}' g''(a,b)}{y_{\text{jui}}' d_{sg}'} \left\{v_{\text{wer}}' \oplus h_{\text{gtu}}'\right\}. \tag{8}$$

In the formula, h_{gtu}' describes the gray value distribution of the 3D landscape image. Assuming that the number of occurrences of various gray values in a 3D landscape image is set to $\theta_{(a,b)}'$ which describes a random point in the image, $(a+n, b+m)$ describes the disturbance point of the image, and $(a+n, b+m)^{kl}$ describes the corresponding points of (a, b) and $(a+n, b+m)$; then, obtain the new image grayscale by the following formula:

$$c_{\text{poi}}{}' = \frac{(a+n,b+m)^{kl}}{(a,b)(a+n,b+m)} \oplus \frac{\theta'}{y_{\text{jui}}{}'}. \tag{9}$$

Assuming that the first-order differential function of the image is $\vartheta_{\text{poi}}{}'$, the inherent characteristic of the 3D landscape image is $u_{\text{wer}}{}'$, and the amplitude-frequency characteristic function is set to $e_{\text{sgh}}{}'$; then, the grayscale characteristics of the image can be obtained by the following formula:

$$r_{\text{yup}}{}' = \frac{e_{\text{sgh}}{}' \pm \vartheta_{\text{poi}}{}'}{u_{\text{wer}}{}'} \pm e_{\text{sgh}}{}'. \tag{10}$$

It is assumed that the variance of the original image block and each adjacent basic image block after merging is $\varphi_{\text{uip}}{}'$, and the ratio of the total image occupied by each image block is set to $\lambda_{\text{wepp}}{}'$. The enhancement function of the edge gray contrast of the 3D landscape image is

$$\varepsilon' = \frac{\lambda_{\text{wepp}}{}'\varphi_{\text{uip}}{}'}{g_{tu}{}'} g_{\text{rty}}{}'. \tag{11}$$

In the formula, $g_{tu}{}'$ is the mask operator and $g_{\text{rty}}{}'$ is the texture attribute in the low frequency range. The predicted value is shown in Figure 5.

Assume that the texture attribute weight space in the low frequency range is η and that the relevance of each pixel in the 3D landscape image signal when it is near in the spatial domain is set to $M_{\text{POL}}{}'$; then, the 3D landscape image feature enhancement method is

$$\varpi_{po}{}' = \frac{M_{\text{POL}}{}'\eta_{XZ}{}'}{Z_{\text{sdj}}{}' K_{\text{qwu}}{}'} \otimes_{\text{yup}}{}' \frac{\{\varpi_{po}{}''\} c_{\text{poi}}{}'}{r_{\text{yup}}{}'} \varepsilon'. \tag{12}$$

In the formula, $Z_{\text{sdj}}{}'$ describes the structural information of the 3D landscape image and $\varpi_{po}{}''$ describes the variance of the original image block and the basic image block of each neighbor. Based on the above processing, the 3D landscape image feature enhancement can be achieved, and the landscape design image clarity is improved.

3. 3D Visualization Simulation Technology

In recent years, with the continuous improvement of virtual reality technology and the rapid advancement of computer technology, people have become more and more aware of the importance of virtual reality technology, and it has been used in many industries, showing broad development prospects. Developed countries began to study the emerging research field of scientific visualization in the late 1980s. It can effectively obtain accurate data. The essence is to track the simulation calculation process through graphics and image processing, and the results are on the screen. The above shows that the interactive processing method closely integrates the graphics and image processing understanding technology and the human-computer interaction technology and finally produces complex multidimensional data image graphics. At present, the field of scientific computing visualization research mainly focuses on the research of computing environment, the research of virtual environment display equipment, and the research of scientific computing visualization technology. (1) As for research on the computing environment, in the high-end computing environment, massively parallel computers with distributed storage, symmetric multiprocessors with shared internal storage, and multiprocessors with distributed shared internal storage have come out one after another. In low-end aspects such as PCs, multiprocessor structures and parallel processing functions have also appeared to improve the quality and efficiency of scientific computing. (2) The research of virtual environment display devices and the development of wall-type large-screen displays are suitable for true three-dimensional display. The cave-like display device with immersive characteristics provides methods and means for the generation and performance of huge amounts of data. (3) There are currently two main types of scientific computing visualization technology that can be studied: one is the study of various application models, and the other is visualization technology. The research of various application models is based on simulation and design application models. It uses visual insight to interactively solve various problems and processes data from different disciplines such as medicine, geology, meteorology, physics, and biological sciences. Graphical information that can be intuitively accepted by human vision and the research of visualization technology, distributed, collaborative, and immersive technologies have become the current mainstream research directions. The prediction is compared in Figure 6.

Water simulation is an important research and development direction of virtual reality technology. In a virtual reality scene, in order to increase the richness and realism of the scene, it is often necessary to add a large number of natural landscapes. Focusing on the needs of the water purification system and in accordance with the requirements of project construction, the system can be roughly divided into three parts: a water purification module, a visual simulation module, and a data management platform. The water purification module includes four sub-function modules: water purification model parameter management, model establishment, model calculation, and result query. The visual simulation module includes six sub-function modules: scene management, modeling, channelization, topography, geophysicalization, and water quality. The data management platform includes six sub-function modules: database parameter management, module data management, channel data management, terrain data management, surface data management, and spatial data management.

1. The water purification module is one of the core modules implemented in this project. It is based on channel modeling, using relevant spatial information and model data provided by the data management platform to dynamically establish a water purification model, calculate and analyze the water quality based on system parameters setup, and then form a result query record to provide guiding information for decision making. 2. The visual simulation module is another core module of this project. It is based on the data management platform to extract channel data, terrain data, feature data, and related data information; use OpenGL technology to dynamically draw channels, terrain, features,

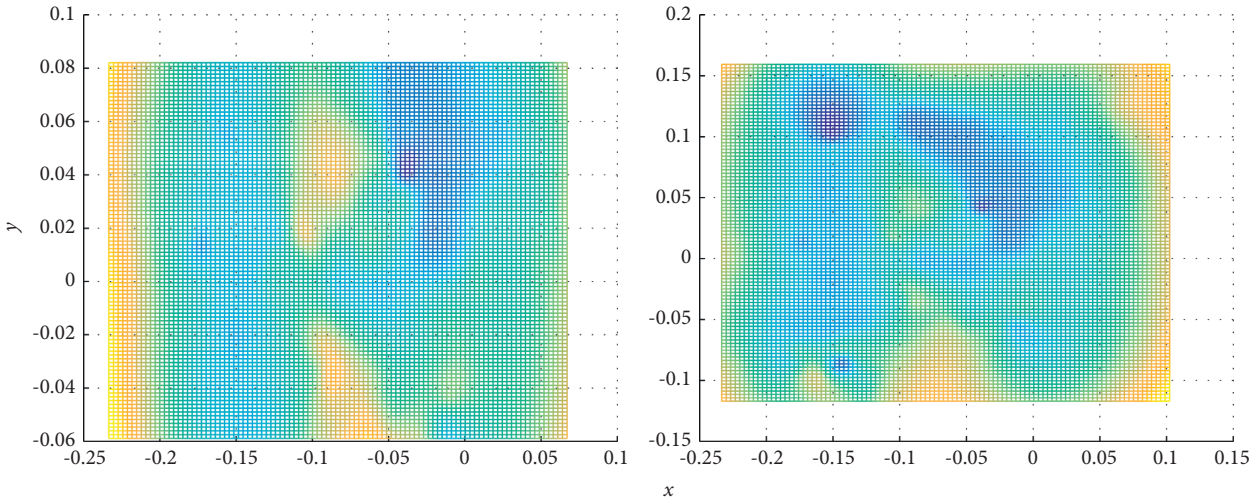


Figure 4: Evaluated x and y.

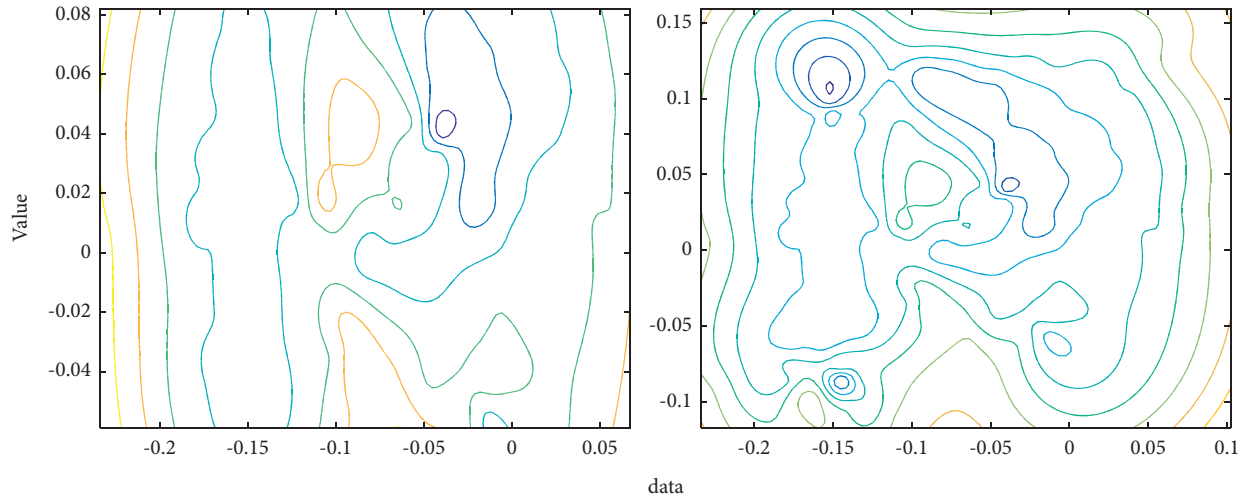


Figure 5: Predicted value.

etc. Objects represent the water channel and its spatial scene in a true three-dimensional manner; use human-computer interaction technology to dynamically set the scene expression mode to realize the scene's functions such as changes in lighting and conversion of observation angles. 3. The data management platform is the foundation of the other two parts, which centrally manages all kinds of data. This module is responsible for storing system initial data, model data, channel data, terrain data, various object data, etc.; it provides various data modification, addition, deletion, and query functions; at the same time, it provides data for the results of model calculations of the water purification system. The prediction is shown in Figure 7.

The system menu includes menu items such as "display initialization," "database initialization," "system initialization," "import model," "model save," and "exit." The "display initialization" menu item is mainly used to set the control parameters of the scene display, the "database initialization" menu item is used to complete the establishment of the database and the setting of the database parameters, and the "system initialization" menu item is mainly responsible for completing the reading of channel data and establishing the system water body. In water purification model, the "import model" menu item is used to complete the import of the established system water purification model into the system, the "model save" menu item is used to export the established model in the system to a file and save it, and the "exit" menu item is used to complete the system exit operation. The data management menu includes items such as "database setting," "river bank data management," "channel data management," "loading object management," "coordinate point management," "cross section management," and "terrain data management." The "database settings" menu item is mainly used to set the correspondence between the system and the tables in the database and the extraction of database parameters. The "river bank data management" menu item mainly maintains the spatial data information of the river bank; the "loading object management" menu item is mainly

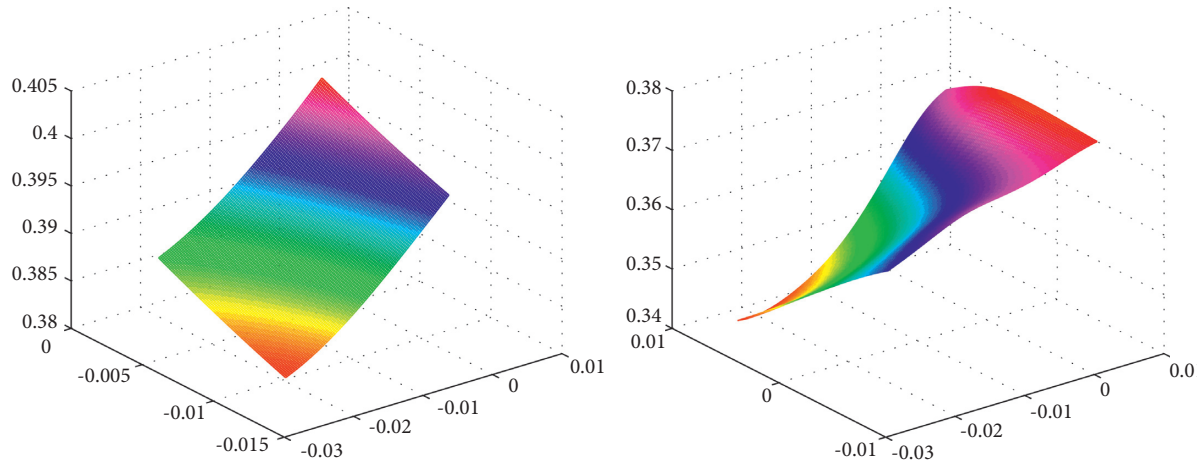


Figure 6: Prediction.

used to load 3ds Max objects into the system library; the "coordinate point management" menu item is used to complete the functions of adding, deleting, modifying, and querying all spatial coordinate points in the system; and the "cross section management" menu item is used to manage each item in the system. For the cross section data of the channel, the menu item "terrain data management" is responsible for the management of the terrain elevation data in this system. The Fourier analysis is compared in Figure 8.

The design of the terrain 3D visualization simulation subsystem is based on the construction of the terrain elevation model as the main axis; establishes the terrain category for management, drawing, and interpolation and other related spatial data; and designs and presents the terrain 3D visualization simulation system. 1. Realization of terrain elevation model. The terrain elevation model is based on discrete spatial elevation points, and the terrain model is created by constructing nonrepetitive triangles. The data in different position is shown in Figure 9.

The specific process is as follows. (1) Data preprocessing: The discrete spatial elevation point data, such as channel key point coordinates, ground feature key point coordinates, and actual measured elevation point coordinates, are sorted according to the spatial distribution. After abnormal points are eliminated, the maximum X, Y, Z values and the minimum X, Y, Z values of the display area are formed. Provide a reference for the display parameters for the simulation scene setting. (2) Spatial interpolation processing: The distance power inverse ratio method is used to interpolate the space where the known points are relatively sparse, so as to achieve a relatively uniform point density and lay the foundation for the later construction of the triangle model. (3) Construct a triangle, according to the original space point and the space coordinate value of the interpolation supplementary point, use the triangle to connect all space points, cover the entire terrain area, and form a terrain model. In the construction process, the system adopts the regular spatial interpolation method in the process of spatial interpolation and uses an interpolation radius of 5 m × 5 m to perform interpolation to obtain regular gridded spatial coordinate data. These gridded spatial data are stored in the system using a two-dimensional array for three-dimensional space rendering. 2. Rendering of terrain elevation model. In the process of terrain 3D visualization, the regular grid data established by the terrain elevation model established above is used to complete the 3D simulation rendering of terrain using triangle-based surface rendering technology. The process includes the following steps. (1) Establish the terrain class according to the preprocessing results and simulation needs, constructing the destructor and the initialization of the terrain, rendering, and some texture-related functions. (2) Obtain the model, associate the model with the terrain class, instantiate the terrain class to form a terrain object, and at the same time copy and store the two-dimensional array of the elevation model to obtain the elevation data of the terrain. (3) Draw the model. According to the grid vertex data constructed by the two-dimensional array, a grid is split into two triangles, and the triangles are drawn in a counterclockwise order. After drawing the triangle patch for the entire area, the preliminary three-dimensional visualization of the terrain model is basically completed. However, in the actual process, you need to further modify and add effects (such as color settings, texture settings, light settings, and material settings) to achieve a more realistic drawing effect.

The channel 3D visualization simulation subsystem includes several aspects such as cross-sectional data management, channel data management, purification system data management and calculation, and 3D channel modeling and drawing. 1. The organization of the channel model. The channel model is based on cross-sectional information, and the channel model is constructed through the combination of two adjacent cross-sections. The specific process is as follows. (1) Data preprocessing, extracting two adjacent section data, corresponding to vertices, and forming 8 ordered vertices. (2) Construct a channel model, compose a trapezoidal or rectangular quadrangular prism based on the coordinate information of the 8 vertices, corresponding to the channel data in the database, extract various parameter information of the channel, and establish a three-dimensional channel model. 2. Drawing the channel model. In the

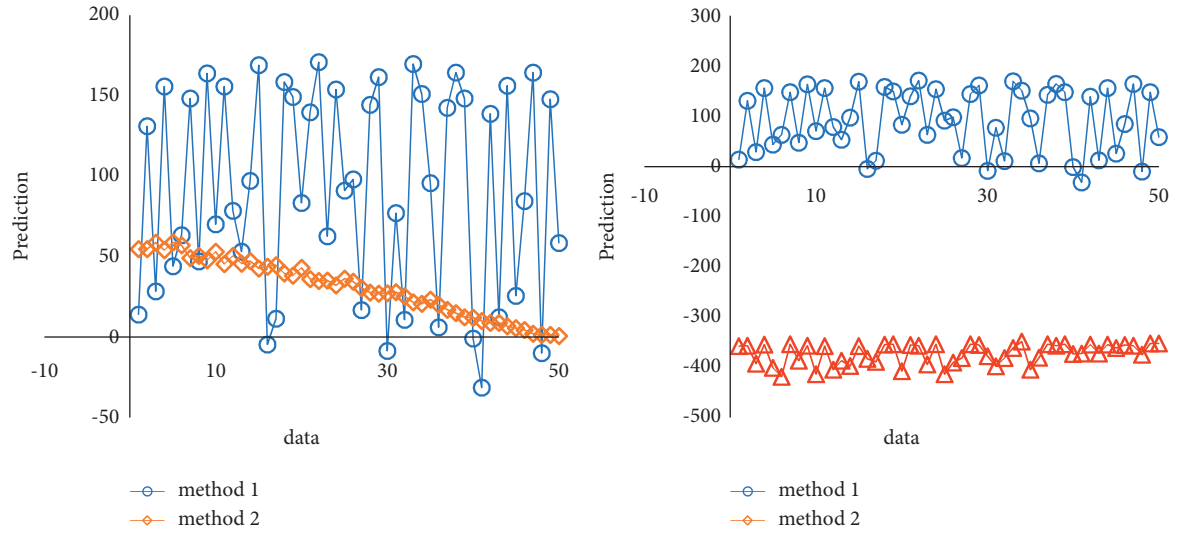


Figure 7: The prediction.

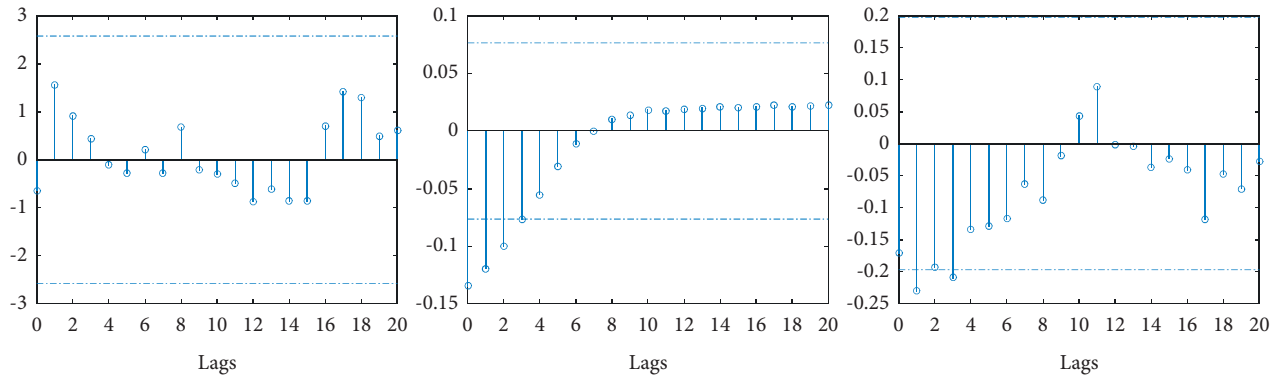


Figure 8: Fourier analysis.

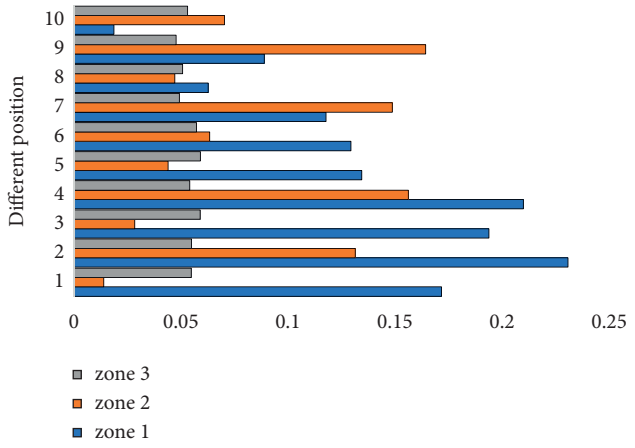


Figure 9: Data in different position.

process of channel three-dimensional visualization, the channel model established above is used to complete the three-dimensional drawing of each surface of the channel using quadrilateral-based surface rendering technology. The process includes the following steps. (1) Establish the channel class, and construct the channel's destructor, initialization, rendering, and some texture-related functions according to the preprocessing results and simulation needs. (2) Obtain the model, associate the model data with the channel class, instantiate the channel class to form the channel object, copy and store the channel's spatial data information and organizational structure information, and construct the left and right dams and a bottom model of the channel. (3) Draw the model. According to the coordinate information of each vertex of the channel model, draw the left and right dams and the bottom model in a quadrilateral manner, and give different texture feature maps to achieve a more realistic drawing effect, which is shown in Figure 10.

The first step is modeling of other objects in the simulation process of the 3D visualization scene for the water purification system. The drawn objects involve not only the main objects such as terrain and channels, but also objects such as ground objects, lawns, trees, buildings, dams, river banks, and other modified objects. The structure of these objects is relatively complex, and it is difficult to express with simple three-dimensional models. 3ds Max is an excellent 3D

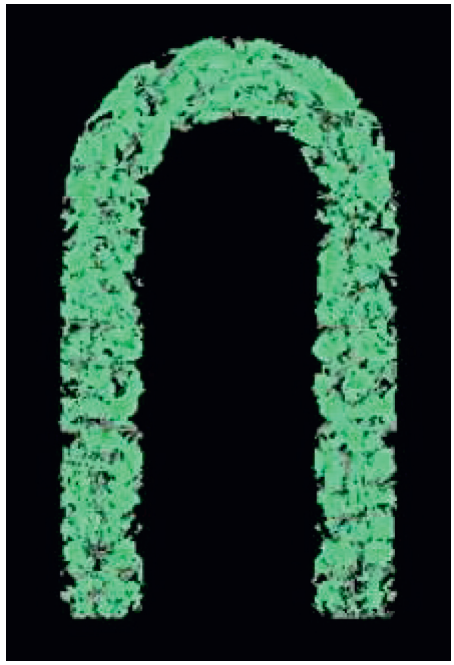


Figure 10: The effect.

modeling tool. It can easily build 3D models of various complex shapes. It is currently one of the most used static modeling tools. Therefore, 3ds Max modeling tools can be used to quickly create various complex objects and attach textures, materials, and other attributes to the objects to achieve a certain rendering effect. To read and draw objects, we can use the 3ds Max tool to build a variety of different three-dimensional models. The established model can be saved in the corresponding 3ds file. Therefore, as long as the system can read these modeling files, it can easily convert various objects and display them further. Here, the system uses the obj file storage format as a unified interface for object reading to read objects and their textures. In this type of file, the object is decomposed into several triangular patches for storage. Therefore, after installing its storage method and reading it into the system, the object can be remodeled and drawn in the form of triangular patches. The effect is shown in Figure 9. This example is to read a 3ds Max rosette architecture. After the system is introduced, you can complete the translation, zoom, and rotation of the object through the control interface to achieve better observation effects.

The scene includes three parts: channel, lake surface, and terrain. The specific steps for drawing are as follows. (1) Object loading: After modeling the above objects, load these objects into the system in turn. (2) Scene modeling: Establish a view class for integrated display of 3D object scenes; this class inherits from CsGL GL class. Implement Init(), Resize(), and other initialization functions in this class according to the scene initialization parameters, lighting model, and projection transformation. (3) Scene drawing: Implement GL's Display() drawing function in the view class. In this function, each model is drawn in the order of object loading, and the rotation angle of the object is set in the local coordinate system according to the relevant parameters of the object. While drawing the scene in 3D, the system adds simple human-computer interaction technology to provide technicians with a multiangle view of the 3D scene.

4. Conclusion

This paper designs a water landscape design system based on 3D image processing technology. Designers can obtain landscape design materials in the landscape design data acquisition unit and then, in the model scene design unit, realize the design based on the landscape design materials through the three-dimensional model.

(1) The design process can preprocess the 3D landscape image, remove noise information and redundant information, and use the 3D landscape image feature enhancement method to achieve 3D landscape image feature enhancement and improve the clarity of the landscape design image. The 3ds Max Material Editor is used to map the enhanced 3D landscape design to complete the design, and the designed landscape feature data will be saved in the system data statistics unit for later query.

(2) This paper also expounds the entire process of alternative simulation construction and confirms the feasibility and superiority of alternative modeling methods through multiple case comparisons. Alternative modeling methods are used as the main way to construct waterfall scenes; that is, simple models are combined with texture mapping, particle effects, and other forms to replace complex models, while ensuring the simulation effect, and the complex models are presented in the most simplified way to make the virtual scene. It is well performed in terms of authenticity and real time.

(3) The research results show that the designed system design renderings are extremely clear and of good quality; they have a high score in the water landscape design. In addition, the system application cost is low, energy consumption is low, and operation efficiency is high.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The author declares no conflicts of interest.

Acknowledgments

This work was supported by the Project of Anhui Province Federation of Social Sciences: "Study on the Integration Development of Urban Image of Shouxian Ancient City Environment and Calligraphy," no. 2021CX529.

References

[1] X. Zhang, J. Yan, S. Feng, Z. Lei, D. Yi, and S. Z. Li, "Water Filling: Unsupervised People Counting via Vertical Kinect Sensor," in *Proceedings of the IEEE Ninth International Conference on Advanced Video and Signal-Based Surveillance*, pp. 215–220, Beijing, China, October 2012.

[2] F. Heide, W. Heidrich, M. Hullin, and G. Wetzstein, "Doppler Time-Of-Flight imaging," *ACM Transactions on Graphics (ToG)*, vol. 34, no. 4, pp. 36–47, 2019.

[3] S. K. Dwivedi, R. Amin, V. Satyanarayana, and R. Chaudhry, "Blockchain-based secured event-information sharing protocol in internet of vehicles for smart cities," *Computers & Electrical Engineering*, vol. 86, no. 1, pp. 1–9, 2020.

[4] Z. Khan and S. Amin, "Bottleneck model with heterogeneous information," *Transportation Research Part B: Methodological*, vol. 112, no. 1, pp. 157–190, 2018.

[5] J. M. Cairney, K. Rajan, D. Haley et al., "Mining information from atom probe data," *Ultramicroscopy*, vol. 159, no. 1, pp. 324–337, 2020.

[6] J. Yu and P. Lu, "Learning traffic signal phase and timing information from low-sampling rate taxi GPS trajectories," *Knowledge-Based Systems*, vol. 110, no. 1, pp. 275–292, 2016.

[7] K. P. Wijayaratna, V. V Dixit, L. Denant-Boemont, and S. T. Waller, "An experimental study of the Online Information Paradox: does en-route information improve road network performance?" *PLoS ONE*, vol. 12, no. 9, pp. 184–191, 2017.

[8] Z. Wang, H. Ren, Q. Shen, W. Sui, and X. Zhang, "Seismic performance evaluation of a steel tubular bridge pier in a five-span continuous girder bridge system," *Structures*, vol. 31, no. 1, pp. 909–920, 2021.

[9] S. Nakayama and J. Takayama, "Traffic network equilibrium model for uncertain demands," in *Proceedings of the 82nd Transportation Research Board Annual Meeting*, Washington DC, USA, March 2021.

[10] H. Shao, W. H. K. Lam, and M. L. Tam, "A reliability-based stochastic traffic assignment model for network with multiple user classes under uncertainty in demand," *Networks and Spatial Economics*, vol. 6, no. 3, pp. 173–204, 2019.

[11] A. Chen, J. Kim, S. Lee, and Y Kim, "Stochastic multi-objective models for network design problem," *Expert Systems with Applications*, vol. 37, no. 2, pp. 1608–1619, 2020.

[12] H. Wang, W. H. K. Lam, X. Zhang, and H. Shao, "Sustainable transportation network design with stochastic demands and chance constraints," *International Journal of Sustainable Transportation*, vol. 9, no. 2, pp. 126–144, 2015.

[13] S.-M. Hosseininasab and S.-N. Shetab-Boushehri, "Integration of selecting and scheduling urban road construction projects as a time-dependent discrete network design problem," *European Journal of Operational Research*, vol. 246, no. 3, pp. 762–771, 2015.

[14] S.-M. Hosseininasab, S.-N. Shetab-Boushehri, S. R. Hejazi, and H. Karimi, "A multi-objective integrated model for selecting, scheduling, and budgeting road construction projects," *European Journal of Operational Research*, vol. 271, no. 1, pp. 262–277, 2018.

[15] M. Johnson, M. Schuster, Q. V. Le et al., "Google's multilingual neural machine translation system: enabling zero-shot translation," *Transactions of the Association for Computational Linguistics*, vol. 5, no. 1, pp. 339–351, 2017.

[16] M. D. Moreno, "Translation quality gained through the implementation of the iso en 17100:2015 and the usage of the blockchain," *Babel*, vol. 1, no. 2, pp. 1–9, 2020.

[17] X. Wang, X. Yu, L. Guo, F. Liu, and L. Xu, "Student performance prediction with short-term sequential campus behaviors," *Information*.vol. 11, no. 4, p. 101, 2020.

[18] Q. Guo, Z. Zhu, Q. Lu, D. Zhang, and W. Wu, "A dynamic emotional session generation model based on Seq2Seq and a dictionary-based attention mechanism," *Applied Sciences*, vol. 10, no. 6, pp. 1–10, 2020.

[19] H. Ren, Xi Mao, W. Ma, J. Wang, and L. Wang, "An English-Chinese machine translation and evaluation method for geographical names," *ISPRS International Journal of Geo-Information*, vol. 9, no. 3, pp. 193–201, 2020.

[20] J. Arús-Pous, T. Blaschke, S. Ulander, J.-L. Reymond, H. Chen, and O. Engkvist, "Exploring the GDB-13 chemical space using deep generative models," *Journal of Cheminformatics*, vol. 11, no. 1, pp. 20–29, 2019.

[21] T. Moon, T. In Ahn, and J. E. Son, "Long short-term memory for a model-free estimation of macronutrient ion concentrations of root-zone in closed-loop soilless cultures," *Plant Methods*, vol. 15, no. 1, pp. 1–12, 2019.

[22] N. Pourdamghani and K. Knight, "Neighbors helping the poor: improving low-resource machine translation using related languages," *Machine Translation*, vol. 33, no. 3, pp. 239–258, 2019.

[23] L. Bote-Curiel, S. Muñoz-Romero, A. Gerrero-Curieses, and J. L Rojo-Álvarez, "Deep learning and big data in healthcare: a double review for critical beginners," *Applied Sciences*, vol. 9, no. 11, pp. 1–11, 2019.

[24] J. Zhang and T. Matsumoto, "Corpus augmentation for neural machine translation with Chinese-Japanese parallel corpora," *Applied Sciences*, vol. 9, no. 10, pp. 1–12, 2019.

[25] Y. Chen, Y. Ma, X. Mao, and Q. Li, "Multi-task learning for abstractive and extractive summarization," *Data Science and Engineering*, vol. 4, no. 1, pp. 14–23, 2019.

[26] Pu. Zhou and Z. Jiang, "Self-organizing map neural network (SOM) downscaling method to simulate daily precipitation in the Yangtze and Huaihe River Basin," *Climatic and Environmental Research*, vol. 21, no. 5, pp. 512–524, 2016.

[27] X. Xiao, "Analysis on the employment psychological problems and adjustment of retired athletes in the process of career transformation," *Modern Vocational Education*, vol. 5, no. 12, pp. 216-217, 2018.

[28] S. Sahoo and M. K. Jha, "Pattern recognition in lithology classification: modeling using neural networks, self-organizing maps and genetic algorithms," *Hydrogeology Journal*, vol. 25, no. 2, pp. 311–330, 2016.

[29] Y. Zhou and B. Yang, "Sports video athlete detection using convolutional neural network," *Journal of Natural Science of Xiangtan University*, vol. 39, no. 1, pp. 95–98, 2017.

[30] J. Pang, "Research on the evaluation model of sports training adaptation based on self-organizing neural network," *Journal of Nanjing Institute of Physical Education*, vol. 16, no. 1, pp. 74–77, 2017.

[31] G. Querzola, C. Lovati, C. Mariani, and L. Pantoni, "A semi-quantitative sport-specific assessment of recurrent traumatic brain injury: the TraQ questionnaire and its application in American football," *Neurological Sciences*, vol. 40, no. 9, pp. 1909–1915, 2019.

[32] J. Wang, X. Luo, and H. Yan, "Correlation analysis between injuries and functional movement screening for athletes of the National Shooting Team," *Journal of Capital Institute of Physical Education*, vol. 5, no. 4, pp. 352–355, 2016.

[33] G. Ma, "Research on the design of juvenile football players' sports injury prediction model," *Automation Technology and Application*, vol. 277, no. 7, pp. 141–144, 2018.

Hindawi
Scientific Programming
Volume 2022, Article ID 8187680, 11 pages
https://doi.org/10.1155/2022/8187680

Research Article

Deep Neural Networks Algorithm for Vietnamese Word Segmentation

Kexiao Zheng[1] **and Wenkui Zheng**[2]

[1]*School of Language and Literature, Guilin University, Guilin 541006, China*
[2]*College of Computer and Information Engineering, Henan University, Kaifeng 475004, Henan, China*

Correspondence should be addressed to Wenkui Zheng; henuzwk@henu.edu.cn

Received 29 December 2021; Revised 26 January 2022; Accepted 31 January 2022; Published 28 February 2022

Academic Editor: Baiyuan Ding

Traditional Vietnamese word segmentation methods do not perform well in the face of Vietnamese ambiguity, in response to the enormous challenge posed by the scarcity of the Vietnamese corpus to language processing. We first investigated the most advanced deep neural network method. According to the ambiguity problem of Vietnamese word segmentation, we then proposed a Vietnamese word segmentation processing technology based on an improved long short-term memory neural network (LSTM), which is made up of an LSTM encoding and a CNN feature extraction portion. The previous important information is kept in the memory unit; the word segmentation processing task is refined into a classification problem and a sequence labeling problem, which can gain the useful features of the word segmentation character and word level automatically. The limitation of the local context window size is avoided, and the word segmentation processing task is refined into a classification problem and a sequence labeling problem. Finally, validated by a homemade Vietnamese news website crawler dataset, the experimental results show that, compared with the single LSTM, single CNN methods, and traditional methods, the performance improvement of our proposed method is more obvious. In the Vietnamese word separation task, the accuracy reaches 96.6%, the recall reaches 95.2%, and the F1 value reaches 96.3%, which is significantly better than the traditional methods CNN and LSTM.

1. Introduction

Under the current dual promotion of economic globalization and artificial intelligence. In the subject of machine translation, language processing has become an important technique, and language processing technology is based on language word segmentation. At present, in the field of linguistic information processing, there has been a lot of studies on word segmentation. The research findings are divided into three types: dictionary-based word segmentation methods, statistics-based word segmentation methods, and understanding-based word segmentation methods. The dictionary-based word segmentation method matches the character string to be studied with the entries of a machine dictionary that has been artificially created according to a strategy. If it is successfully matched with the string in the character dictionary, following that, word segmentation is carried out. The statistical method mainly performs statistical analysis on the words and phrases in the corpora and calculates the information about their mutual occurrence. The closeness of the characters' combination relationship is reflected in the mutual information. When the gap between the characters is greater than a specific threshold, this character combination can be deemed a word. Finally, by defining the mutual information of two characters, then calculating the probability of the two characters appearing next to each other, through algorithm design, the understanding-based word segmentation method allows the computer to imitate a human's understanding of the text, in order to reach the appearance of words being recognized. It is challenging to organize varied linguistic information into a form that can

be directly read by a machine due to the generality and complexity of language knowledge.

Currently, in the realm of linguistic analysis, using information retrieval techniques, various studies have been completed for common languages such as English and Chinese [1, 2], hand-crafted rules [3], or neural mechanisms [4–6]. However, there has been minimal research into Vietnamese word segmentation. Most research on Vietnamese is limited to detecting the meaning of word segmentation using traditional methods [7] or using connection matching to extract context [8]. As far as we know, in the study of Vietnamese word segmentation, research at home and abroad has just begun. Until now, there are no specific shared resources available for academic research. All language resources need to be built from scratch. As the basis of Vietnamese natural language processing, Vietnamese word segmentation requires the collection of Vietnamese corpus resources and processing them as required, which is a prerequisite for Vietnamese word segmentation. At present, the most widely used Vietnamese word segmentation tool is VnTokenizer [9], which was developed by the University of Hanoi in 2008 based on the maximum matching and N-Gram model [10]. Vietnamese word segmentation currently has two major problems: combination ambiguity and cross ambiguity. Although some researchers have used maximum entropy, SVM, and CRF [11] methods to segment Vietnamese words and have achieved certain success, they are all in the experimental stage. The accuracy is not stable enough. Based on previous work, combining Vietnamese word-formation features and language features, we propose a model based on long short-term memory neural network (LSTM), which is determined by input, output, and forgetting gates of how to use previous information to model and update the memory of previous information. Experiments show that the method presented in our research is capable of effectively resolving the ambiguity issue. This paper's contribution can be summed up as follows.

(i) Firstly, we introduced the relevant research work in the direction of language processing and then proposed the study of Vietnamese word segmentation in response to the scarcity of the Vietnamese corpus.

(ii) Although traditional methods are not effective in processing Vietnamese word segmentation ambiguity models, we have studied methods based on deep neural networks for Vietnamese word segmentation.

(iii) Although Vietnamese word segmentation is a special case, our model is an improvement of the LSTM method, which makes our method easy to generalize and apply to other sequence labeling tasks.

(iv) A Vietnamese word segmentation model based on an enhanced LSTM framework is presented, which is composed of an LSTM encoding part and a CNN feature extraction part.

(v) The findings of the experiments reveal that, when compared to the old method and the single LSTM and the single CNN method, our method shows a greater improvement in performance.

The remainder of this paper is arranged in the following manner. Section 2 discusses language processing-related work. Section 3 introduces the Vietnamese language features and ambiguity model and then describes in detail the relevant principles and implementation process of the improved LSTM Vietnamese word segmentation processing method. Section 4 reports the experimental dataset, experimental settings, evaluation indicators, and analysis of experimental results. Finally, Section 5 summarizes our research and reveals some further research works.

2. Related Work

To detect language segmentation, most studies use annotated corpora to train traditional classifiers by investigating different types of features [12, 13]. However, this technique necessitates a significant amount of feature engineering. Another choice is to learn discriminative features by using deep neural network algorithms. The specific process is shown in Figure 1. First, input the vector, then use the filter to extract the character features and automatically learn, and then output the prediction results through the pooling layer and the fully connected layer. For instance, Shi et al. [14] presented modeling multilevel nonlinear feature representations by stacking distinct CNN feature maps; Kato et al. [15] used recurrent autoencoders to spoken Japanese conversations based on smartphones systematic Japanese word segmentation classification of discourse; Goo et al. [16] proposed a slotted gate, which concentrates on understanding the link between word segmentation and slot attention vectors in order to achieve superior semantic framework results via overall optimization.

To infer the meaning of word segmentation through joint context, the majority of contemporary research makes use of IR technologies [17]. Ji et al. proposed an optimization method that extracts the most comparable word segmentation problem from a specified common context and responds by matching the most likely word segmentation as a morpheme. This technology is usually used to build open domain chat interface machine translation (for example, to provide small chat services) or to answer common questions in a given domain. In terms of e-commerce, to select the most appropriate response from the existing datasets, Cui et al. [4] use the word segmentation autonomous learning system to support small chats and comments on the research and development system. Yan et al. [6] proposed a generic approach to building a task-based dialogue system for online shopping. To obtain the PI requested by the customer, the system uses the DSSM model to match the question with the basic segmentation PI [18]. However, many researchers do not permit online booking by customers; in many practical uses, external data resources are challenging to manage. The ultimate goal of these researchers is to build a cross-language retrieval system, as shown in Figure 2.

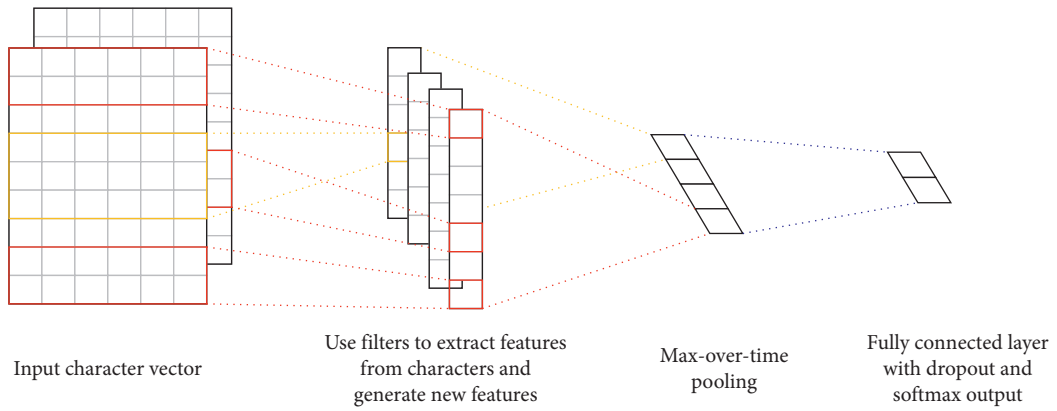


Figure 1: Use deep neural methods to learn character features.

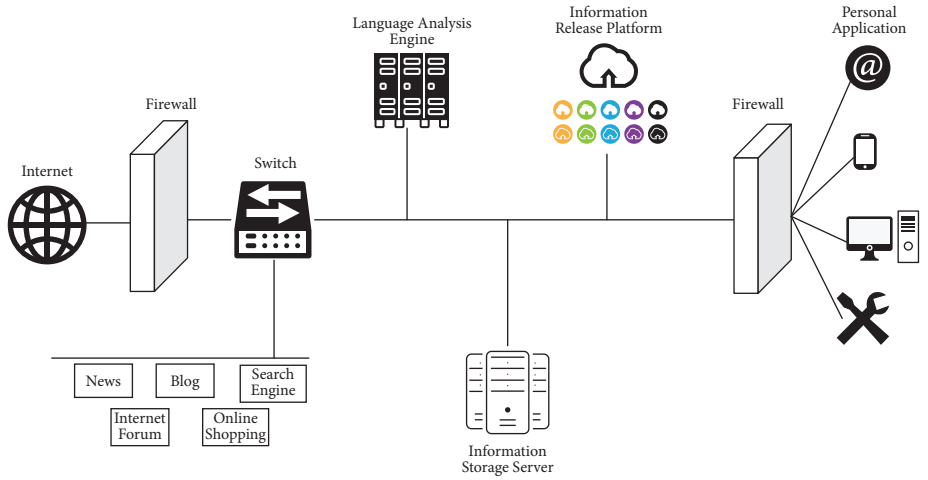


Figure 2: Cross-language retrieval system.

There has been very little research on Vietnamese word segmentation. The MaxEnt classifier is used by Ngo to allocate applications, and in order to detect sentence meanings, use the connected dictionary matching rule. [7]. This matching system is extremely expensive, necessitates domain expertise, and cannot handle Vietnamese ambiguity. In this work, due to simple problems being easier to analyze, they always focus on small problems to assist users in completing mobile phone interaction tasks. Tran et al. [8] investigated the recognition of named entities in spoken Vietnamese text. This research provides a simple machine learning model that incorporates a vast variety of handcrafted characteristics. However, designing these functions manually is extremely difficult; it also necessitates a high level of expertise in the realm of technology and the Vietnamese language.

There have been some research results in Vietnamese word segmentation methods [19–21]. However, there are still various challenges in understanding natural language, especially in the case of ambiguity. There are two main types of ambiguity in Vietnamese, which are combined ambiguity and cross ambiguity. Different participles will produce different ambiguities in Vietnamese expressions. In this paper, we concentrate on researching and comprehending the ambiguity in Vietnamese word segmentation to help language processing technology to better understand Vietnamese. As mentioned earlier, this language presents some challenges. To this end, we presented a solution using long and short-term memory networks. The solution uses input, output, and forget gates to determine how to use previous information to explicitly model and renew the memory of the previous feature. If a beneficial character from the input sequence is detected by the LSTM unit at a preliminary phase, it can easily change this feature to be carried long distances to capture potentially useful long-distance information. When new features are generated in the word segmentation traversal, they will be compared again, and then the corpus will be updated again to achieve a full understanding of the purpose of the Vietnamese sentence.

3. Method

3.1. Vietnamese Features and Ambiguity. The tones of Vietnamese are very similar to Chinese Pinyin. Each syllable is composed of initials, vowels, and tones, but the difference is that the initials of Vietnamese include the flat, sharp, profound, question, down, and accent. Like Chinese, although it lacks morphological changes, every syllable has

meaning. In addition, its composition is the Latin alphabet, phonetic characters, and punctuation marks. Morphemes, as the word-building units of Vietnamese, can be divided into five categories: monosyllable words, compound words, accented alliterative words, coupled words, and derivative works. The construction of Vietnamese phrases also plays a vital role in Vietnamese participles. Transforming the Vietnamese phrase structure tree into a dependency structure tree is the current standard processing approach. The labeling system of the Vietnamese dependency structure tree library is shown in Figure 3, which annotates the dependencies and types of dependencies between words in a sentence [22, 23].

The annotation of sentences in the Vietnamese phrase structure tree library is shown in Figure 4. It only identifies the phrase hierarchy and phrase type of each sentence and does not indicate the central subnode of each phrase. The most common way to determine the central subnode of a phrase is to use the central subnode filter table. The dependency tree structure is a supplement to the phrase structure tree. The advantage of this is that the scale of the target tree bank can be increased, and the ability of the dependency analyzer can be improved without changing the learning strategy of the syntactic analysis model. In other words, it is a way of learning syntactic knowledge using multiple treebanks. It has a good experimental effect in dealing with Vietnamese treebank conversion and Vietnamese dependency treebank expansion and solves the problem of Vietnamese dependency syntax analysis well.

The development of the central subnode filtering table is an important part of the whole work. Table 1 is a part of the central subnode filtering table, and each row contains three items (phrase type, search direction, and priority), where phrase type is the phrase symbol for nonterminal nodes; search direction is the direction to search for the central subnode within nonterminal nodes. When the value is L, the search begins on the left side of the phrase and works its way to the right, and when the value is R, the search starts from the right side of the phrase to the left; priority is to determine the priority order of each token subnode as the central node in the phrase. For example, based on an entry in the filter table <VP, L, VP; V; A; AP; N; NP; S;. ∗>, the central subnode of the VP phrase can be determined as follows: observe each VPs from left to right to find a subnode, and the subnode with the symbol V found first is the central subnode of VP; if no VP node is found, observe each subnode of VP from left to right again, and the subnode with the symbol V found first is the central subnode of VP; and so on, if there is no subnode in VP with the symbols VP, V, A, AP, N, NP, S,. ∗, then the leftmost subnode is the central subnode by default.

There are two main types of Vietnamese ambiguity: combined ambiguity and cross ambiguity [24–26]. In combined ambiguity, some words are combined into sentences and have different meanings from word morphemes, such as "Bàn là một dụng cụ để ăn (The table is a tool for eating)." The morpheme "Bàn" means "table," and "là" means "is," but the combination of these two morphemes "Bàn là" means "iron." This ambiguity is called combined ambiguity. Cross ambiguity means that both the current morpheme and its preceding and following morphemes can form words, for example: "Tốc độ truyền thông tin không tốt lắm (transmission speed is not very good)," where "truyền thông" means "media," and "thông tin" means "information." This kind of ambiguity is cross ambiguity. These two kinds of ambiguous information often occur, which brings huge challenges to Vietnamese language processing.

3.2. Vietnamese Character Feature Extraction. The selection of Vietnamese character features has a great impact on the result of word segmentation [27]. Combined with the morpheme characteristics of Vietnamese, we adopted a method based on a Markov random field method for character feature extraction. According to Tseng's research [10], in this paper, two basic features are selected, i.e., character N-gram feature and character repetition information feature, as shown in the following equation:

$$\begin{aligned} &\text{Character } N-\text{gram feature} \begin{cases} W_k\,(k=[-2,-1,0,1,2]) \\ W_kW_{k+1}\,(k=[-2,-1,0,1]) \end{cases}, \\ &\text{Character repetition information feature} \begin{cases} W_kW_{k+2}\,(k=[-1,0]) \\ W_kW_{k-2}\,(k=[0,1]) \end{cases}, \end{aligned} \tag{1}$$

where W represents the Vietnamese morpheme; W_0 represents the current morpheme, and k represents the position relative to the current morpheme. For example, in "Tôi thích nằm trên ghế và xem TV," if W_0 means the current Vietnamese morpheme "ghế," then W_{-1} means "trên"; W_{-2} means "nằm"; W_1 means "sofa." Repeat (W_0W_1) means that the current morpheme and the next morpheme are the same.

Aiming at the unregistered words that are prone to errors such as numbers, letters, and punctuation in Vietnamese, this paper defines Vietnamese morphemes into ten categories based on language characteristics: Sin, Pre, Suf, Pun, Dig, Let, Spe, Tim, Dat, and Oth [28]; related specific definitions are shown in Table 2.

3.3. Neural Model for Vietnamese Word Segmentation. Vietnamese word segmentation is generally considered to be based on character sequence labeling. We use the Begin,

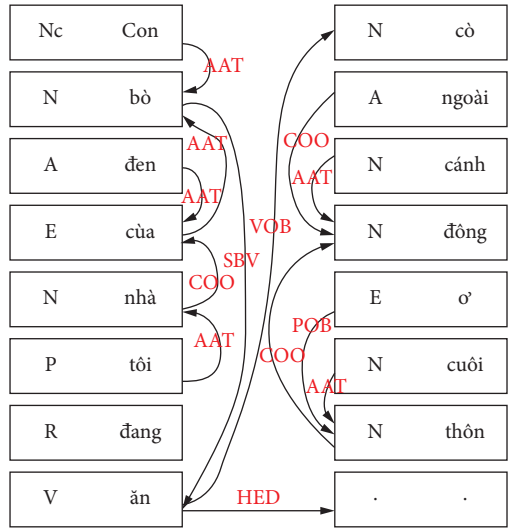


Figure 3: Example of Vietnamese dependency structure tree.

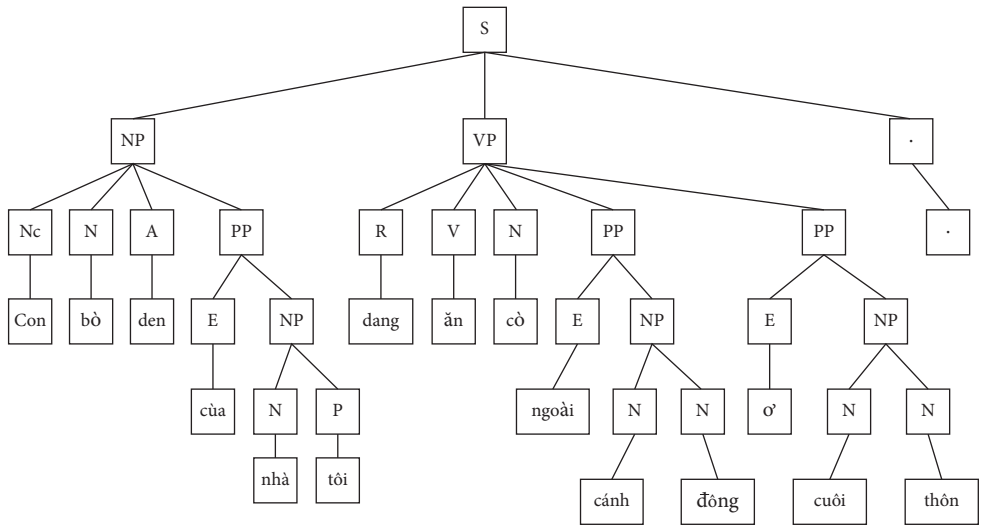


Figure 4: Example of Vietnamese phrase structure tree.

Table 1: Classification of Vietnamese characters.

Phrase type	Search direction	Priority
S	L	S; VP; AP; NP . *
SBAR	L	SBAR; S;VP; AP; NP;. *
SQ	L	SQ; VP; AP; NP;. *
NP	L	NP; Nc; Nu; Np; N;P. *
VP	L	VP; V;A; AP; N;NP; S . *
AP	L	AP; A;N; S . *
PP	L	PP; E;VP; SBAR; AP; QP . *
RP	R	RP; R;T; NP . *
XP	L	XP; X . *
MDP	L	MDP; T;I; A;P; R;X . *
UCP	L	. *
WHADV	L	R . *
WHVP	L	V . *
QP	L	QP; M . *

Table 2: Classification of Vietnamese characters.

Feature representation	Feature meaning	Example
Sin	Separate words	thi, ah
Pre	Start	c? m,xin
Suf	End	Sinh
Pun	Punctuation	,.!
Dig	Number	1,2,3
Let	Letters and combinations	A, a
Spe	Special identifier	@, %
Tim	Time	Minutes and seconds
Dat	Date	Date
Oth	Other	I II et al.

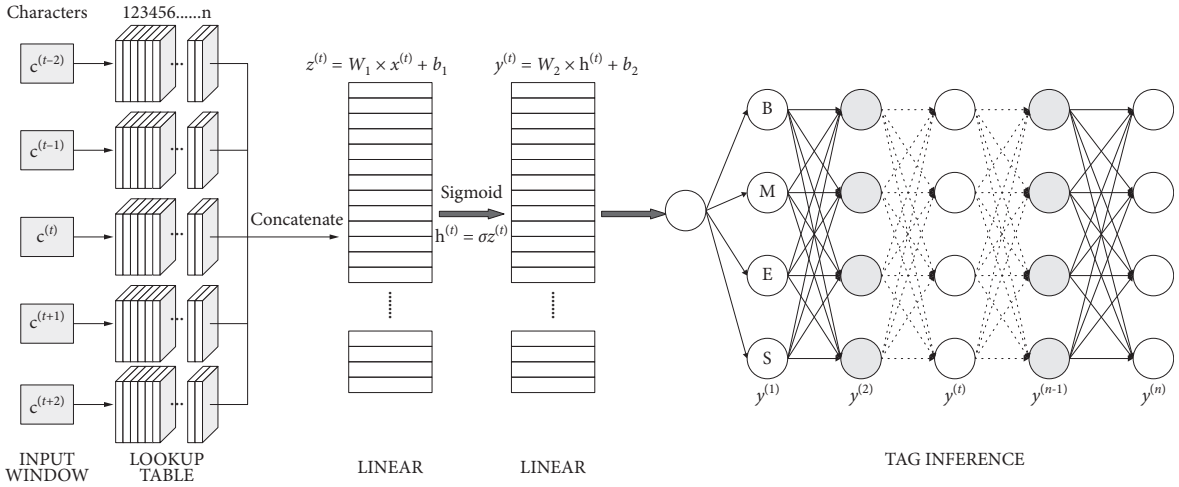


Figure 5: Vietnamese word segmentation neural model.

Middle, End, and Single of the multicharacter word segmentation to represent the Vietnamese characters and take the first letter to mark the sequence as {B, M, E, S}, and each Vietnamese character is marked as one of them, where the following neural model builds the foundation.

The marking method we usually use is the local window method. From the perspective of adjacent characters, it is assumed that the marked character has a close relationship with the weight of the neighbor. Given a sentence $c^{(1:n)}$ as input, a k-sized window will glide from character $c^{(1)}$ to $c^{(n)}$, where n represents the length of the sentence. k represents the window size, when k is 5; for each character $c^{(t)}\,(1 \le t \le n)$, the context characters $(c^{(t-2)}, c^{(t-1)}, c^{(t)}, c^{(t+1)}, c^{(t+2)})$ will be input into the lookup table layer for character matching. Characters beyond the boundary of the sentence are assigned to special symbols, specifically, the start and finish symbols. After that, the lookup table layer connects the character embedding to a single vector $x^{(t)} \in \mathbb{R}^{H_1}$, where $H_1 = k \times d$ represents the size of the first layer. Then $x^{(t)}$ is sent to the next layer, which transforms linearly. Finally, the sigmoid function $\sigma(x) = (1 + e^{-x})^{-1}$ is used as the activation function.

$$\begin{aligned} h^{(t)} &= \mathrm{g}\left(W_1 x^{(t)} + b_1\right), \\ y^{(t)} &= W_2 h^{(t)} + b_2, \end{aligned} \tag{2}$$

where $W_1 \in \mathbb{R}^{H_2 \times H_1}$, $b_1 \in \mathbb{R}^{H_2}$, $h^{(t)} \in \mathbb{R}^{H_2}$, $W_2 \in \mathbb{R}^{|\mathrm{T}| \times H_2}$, $b_2 \in \mathbb{R}^{|\mathrm{T}|}$, $y^{(t)} \in \mathbb{R}^{|\mathrm{T}|}$. H_2 represents the number of hidden units in the second layer, which exists as a hyperparameter. Given a set of tags T of size |T|, a linear transformation is carried out in a similar way, but it does not follow a nonlinear function. In Vietnamese word segmentation, the most commonly used tag set T is the character sequence set {B, M, E, S}.

In order to map the dependency between tags, we introduce the transition score A_{ij} proposed by Collobert et al. [29], which is used to measure the probability from the label $i \in \mathrm{T}$ to the label $j \in \mathrm{T}$. The transition score method can stably complete sequence tagging tasks such as Vietnamese word segmentation. The only disadvantage is that it can only use the context information of a limited-size window, and large-distance information may be ignored. The general architecture of the detailed Vietnamese word segmentation neural model is shown in Figure 5. It is mainly composed of three levels: the first layer is the character embedding layer, which inputs characters, then performs sequence tagging of the characters, and embeds morphemes; the second layer is a series of neural network layers, which performs feature learning and classification of the characters; and the last layer is the tag label inference layer, which performs label matching and inference on the predicted word segmentation characters. In particular, we adopt a bilinear structure, which examines the interaction of semantic information between different dimensions by computing the outer product of convolutional description vectors. Since different dimensions of the description vector correspond to different channels of convolutional features, and different channels extract different semantic features, the relationship between different semantic features of the input image can be captured simultaneously through bilinear operations.

3.4. Improved LSTM Model of Vietnamese Word Segmentation. The initial stage in processing symbolic input with neural networks is to represent it as distributed vectors; it is also known as embeddings [30, 31]. In the Vietnamese word segmentation task, we manually built a character dictionary. Extract a character dictionary from the training set, and map unknown characters to different morpheme characters. A real-valued vector is used to represent each character. A vector matrix is created by stacking the embedded characters. The lookup database retrieves the corresponding characters after embedding. The lookup surface layer is the projection layer, and each embedded context morpheme character is implemented according to its index through a lookup table operation.

The long short-term memory neural network (LSTM) was presented in 1997 by Hochreiter and Schmidhuber [32]. LSTM is a derivative of the recurrent neural network (RNN). Since 2010, it has been demonstrated that RNNs have been successfully applied to speech recognition [33], language

modeling [34], and text generation [35]. However, the disappearance of gradients and explosions makes RNN difficult to apply to long-term dynamics research. As an improved network of RNN, LSTM can handle this problem well. LSTM gives the network a great degree of freedom so that the network memory unit has an adaptive solution for learning and updating information. Therefore, LSTM neural network has an excellent performance in word segmentation tasks. The principle of the LSTM network is shown in Figure 6.

Assume that $X = (x_1, x_2, ..., x_n)$ represents an input sentence composed of word representations of n words. In every position t, the RNN produces a hidden layer h in the middle denoted as y_t, and the hidden state h_t uses a nonlinear activation function to update the previous hidden layers h_{t-1} and the input x_t, as shown below:

$$\begin{aligned} y_t &= \sigma\left(W_y h_t + b_y\right), \\ h_t &= f\left(h_{t-1}, x_t\right), \end{aligned} \tag{3}$$

where W_y and b_y are the parameter matrices and vectors learned during the training process, and σ represents the elementwise softmax function.

The LSTM unit includes an input gate i_t, a forget gate f_t, an output gate o_t, and a memory unit c_t to update the hidden state h_t, as shown below:

$$\begin{aligned} i_t &= \sigma\left(W_i x_t + V_i h_{t-1} + b_i\right), \\ f_t &= \sigma\left(W_f x_t + V_f h_{t-1} + b_f\right), \\ o_t &= \sigma\left(W_o x_t + V_o h_{t-1} + b_o\right), \\ c_t &= f_t \odot c_{t-1} + i_t \odot \tanh\left(W_c x_t + V_c h_{t-1} + b_c\right), \\ h_t &= o_t \odot \tanh\left(c_t\right), \end{aligned} \tag{4}$$

where $\odot$ is a kind of function which is similar to the multiplication operation, V represents a matrix related to weight, and b represents the learning vector. In order to improve the performance of the model, morpheme training was carried out on two LSTMs. The first one is a morpheme that begins on the left and works its way to the right; the next one is a reverse duplicate of a character. Before passing to the next layer, the outputs of the forward and reverse passes are combined in series. Finally, the activation function is used to obtain the prediction result.

In order to train the model, this paper takes advantage of pretrained word embeddings gathered through news website crawlers. In order to deal with the out-of-vocabulary (OOV) issues, we use a pretrained word embedding method and apply character-stage embedding from words. The character embedding is learned using the complete network after being randomly initialized. The overall architecture is shown in Figure 7. On the left side of the word segmentation, the forward LSTM estimates the description of the context, and the second LSTM calculates the morpheme character in the reverse direction and reads the sequence. All the representations are connected and linearly projected to the next layer, the size of which is the same as the amount of different word segmentation. Then we use the CRFs method to consider neighboring tags to generate the final context prediction for each word, in order to handle local information instead of the entire sentence in the long sentence domain, so as to enhance precision. Finally, in order to extract the local information between adjacent segments of the target word, we add an adaptive CNN layer at the end of the network.

4. Experiments

4.1. Datasets. Vietnamese is a scarce resource, and there is currently no publicly available dataset of Vietnamese subwords. For verifying the effectiveness of the method, we apply crawler technology to crawl Vietnamese news text data from the Vietnam Daily News website, use the word segmentation tool proposed by VnTokenizer to preprocess the crawled Vietnamese text data, and then obtain it through manual proofreading of 100,000 Vietnamese subword datasets. The process of making Vietnamese corpus dataset is shown in Figure 8. All the following experiments are based on this dataset for training. The test dataset contains 10,000 data and the training dataset contains 90,000 data. At the same time, based on constructing the word corpus, we further segment the vocabulary into syllables. All datasets are preprocessed by replacing Vietnamese corpus, continuous English characters, and numbers with unique signs. We use traditional scoring methods to calculate accuracy, recall, and F1 scores to evaluate the effect of Vietnamese word segmentation.

4.2. Training. All experiments in this paper are executed on the PyTorch framework and configure python 3.5 as the language environment version. The experimental hardware environment uses GTX 1080 Ti GPU, Intel i7-7700 CPU, 50 GB memory, and Pycharm Community 2020.3.2 as a development tool. Use the Max-Margin criterion during training. It concentrates on the model's decision boundary's robustness and proposes a probabilistic-based estimating approach as an alternative. So as to avoid the difficulty of overfitting during the training of the deep network, the dropout method presented by Srivastava et al. is also used in the training. The relevant training parameters are shown in Table 3.

4.3. Evaluation Metrics. In the experimental verification process, we choose precision, recall, and $F1$ score to evaluate the effectiveness of the method. Most classification and sequence labeling problems also adopt the above three evaluation factors. The equation is as follows:

$$\begin{aligned} F1 &= \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} \times \text{recall}}, \\ \text{precision} &= \frac{\text{TP}}{\text{TP} + \text{FP}}, \\ \text{recall} &= \frac{\text{TP}}{\text{TP} + \text{FN}}. \end{aligned} \tag{5}$$

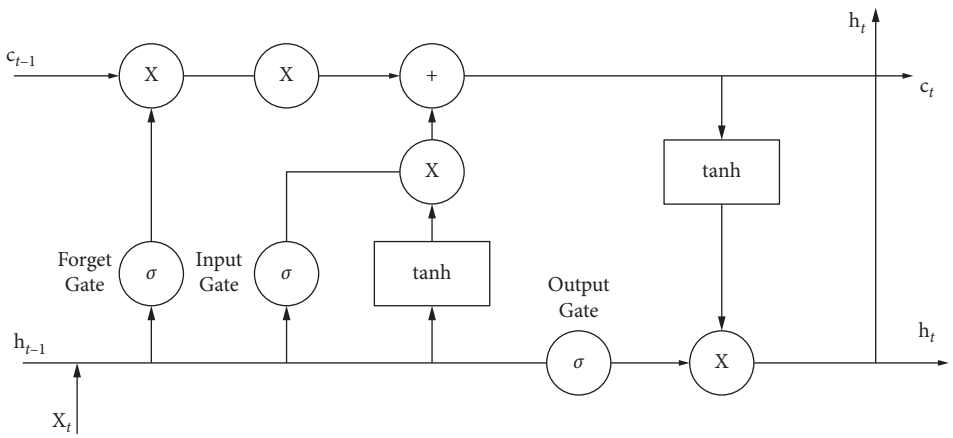


Figure 6: LSTM network function principle.

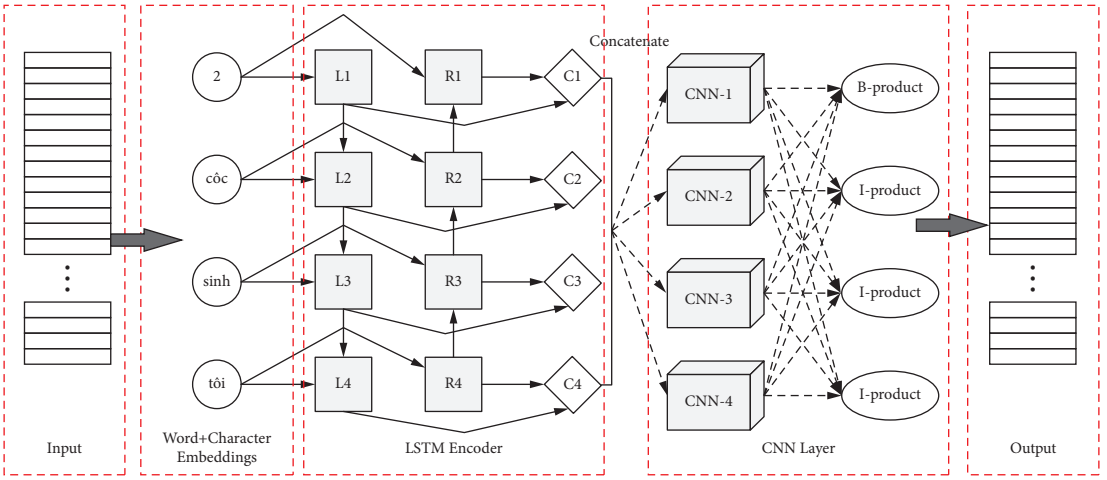


Figure 7: A framework of using improved LSTM for Vietnamese word segmentation.

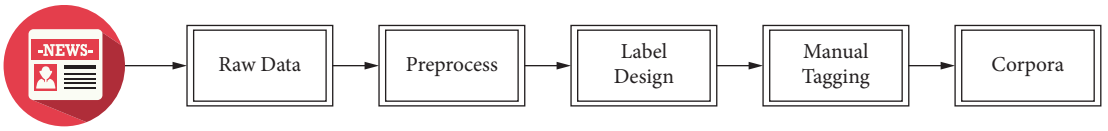


Figure 8: A framework of using improved LSTM for Vietnamese word segmentation.

Table 3: Training parameter settings.

Parameter	Value
Learning rate	0.15
Epoch	100
Context length	(−2,2)
Character embedding size	100
Regularization	0.0001
Margin loss discount	0.2
Dropout rate	0.5
Initial learning rate	0.2
Hidden unit number	200

In the above equation, TP (True Positive) is the number of properly recognized word segments. FP (False Positive) is the number of misrecognized word segments. FN (False Negative) indicates the number of unrecognized word segments.

4.4. Experimental Result. To verify how varying window widths affected the experimental results, we, respectively, compared the word segmentation effects under input sequence windows of different sizes. The comparison results are shown in Table 4.

Table 4: The influence of different window sizes on experimental results.

Window size	P	R	F
3	92.2	90.3	91.1
5	96.6	95.2	96.3
7	93.5	92.4	94.2
9	93.1	91.5	93.6

Table 5: The influence of different numbers of training corpora on experimental results.

Number of corpora	E1	E2
1	78.3	81.2
3	84.6	89.9
6	87.3	91.1
9	89.2	96.6

Table 6: Comparison of the performance of different methods on Vietnamese word segmentation.

	P	R	F
MaxEnt	84.3	82.5	82.9
biLSTM	92.6	90.1	91.5
CNN	91.3	88.2	90.7
Ours	96.6	95.2	96.3

Table 4 shows that the choice of window size has an impact on the result of word segmentation, and the experimental results are normally distributed. Among the common Vietnamese words, most Vietnamese words contain 2 to 5 syllables. When the window is 5, the experimental results are the best, which can not only avoid the current syllable's insufficient dependence on the upper and lower syllables but also avoid the excessively long window. The syllable features are redundant. Therefore, the syllable window in this experiment is set to 5.

For confirming the impact of varied training corpus scales on the experimental results and to test the influence of the Vietnamese syllable structure knowledge feature on the character embedding vector representation effect, we conducted a comparative experiment on different scales of the training corpus. The results are shown in Table 5. Among them, E1 represents the word segmentation accuracy of the method under different corpus scales, and E2 represents the word segmentation accuracy of the method under the same corpus scale.

Table 5 shows that when the corpus is small, more features can be provided during neural network training. When the corpus increased from 10,000 to 30,000, the accuracy of Vietnamese word segmentation was significantly improved. The main reason was that the problem of insufficient feature coverage in the case of the small-scale corpus was supplemented, which caused the deep network to be unable to learn more features. At the same time, the method in this paper is faster than the CNN network in terms of accuracy rate improvement and combines the representation method of the syllable structure knowledge vector. In the case of the same corpus, it can provide more special segmentation morphemes for neural network learning. When the number of corpora was increased from 60,000 to 90,000, the accuracy of Vietnamese word segmentation was increased by 5.5%. It can be seen that the integration of linguistic knowledge features in the Vietnamese word segmentation vector representation can effectively compensate for the impact of resource scarcity on the performance of Vietnamese word segmentation. Therefore, the number of training corpus in the final experimental verification of this paper is 90,000.

Previous work found that by pretraining character embedding on unlabeled data, the model's overall performance can be improved. In the pretraining character embedding work in the homemade Vietnamese corpus, we used the toolkit researched by Mikolov et al. [36]. The character search database is initialized with the acquired embedding character, which replaces the earlier random initialization process. In this section, we verify it on a self-made dataset and compare it with the most advanced methods MaxEnt, biLSTM, and CNN. The experimental results are shown in Table 6.

The experimental results show that, comparing with the CNN method, the accuracy of our method is boosted by 5.3%. Comparing with the biLSTM method, the accuracy of our method is boosted by 4%. Comparing with the MaxEnt method, the accuracy of our method is boosted by 12.3%. Regardless of whether it is from the recall rate or F1 score, the performance of the method proposed in this paper is even better, which proves the effectiveness of the method in this paper.

5. Conclusion

In this paper, we analyze the related work and research status of word segmentation in language processing and then lead to the study of Vietnamese word segmentation with a sparse corpus. The problem of combinational ambiguity and cross ambiguity in Vietnamese poses a great challenge to the task of Vietnamese word segmentation. In order to realize the Vietnamese word segmentation task, an improved LSTM neural network framework is proposed, and the Vietnamese word segmentation task is separated into a classification part and a sequence labeling part. We abandoned the traditional methods and chose the most advanced deep neural network to gain beneficial features at the character level. In order to test the method's performance, a verification experiment was conducted through a self-made Vietnamese word corpus. The experimental results show that, overall, the performance of the Vietnamese word separation system can be markedly increased by applying neural networks. In general, the accuracy of our method in the Vietnamese word segmentation task reached 96.6%, the recall rate reached 95.2%, and the F1 value reached 96.3%, which is significantly better than the traditional methods, CNN and LSTM methods.

In the future, we hope to use bidirectional recurrent neural networks to process sequences in two directions. Some adjustments are made to the method so that it can work well in other Southeast Asian language fields. Research on deeper semantic feature extraction may also be another

direction in the future. In addition, self-made datasets have imbalance problems; in our future research, more attention will be paid to the construction and balance of the dataset.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] Z. Ji, Z. Lu, and H. Li, "An information retrieval approach to short text conversation," 2014, https://arxiv.org/abs/1408.6988.

[2] Z. Yan, N. Duan, J. Bao et al., "Docchat: an information retrieval approach for chatbot engines using unstructured documents," in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, vol. 1, pp. 516–525, Berlin, Germany, January 2016.

[3] D. A. Ali and N. Habash, "Botta: an Arabic dialect chatbot," in *Proceedings of the COLING 2016, the 26th International Conference on Computational Linguistics: System Demonstrations*, pp. 208–212, Osaka, Japan, December 2016.

[4] L. Cui, S. Huang, F. Wei, C. Tan, C. Duan, and M. Zhou, "Superagent: a customer service chatbot for e-commerce websites," in *Proceedings of the ACL 2017, System Demonstrations*, pp. 97–102, Vancouver, Canada, 2017.

[5] C. Li, L. Li, and J. Qi, "A self-attentive model with gate mechanism for spoken language understanding," in *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pp. 3824–3833, USA, January 2018.

[6] Z. Yan, N. Duan, P. Chen, M. Zhou, J. Zhou, and Z. Li, "Building task-oriented dialogue systems for online shopping," in *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, San Francisco California, USA, February 2017.

[7] T.-L. Ngo, V.-H. Nguyen, T.-H.-Y. Vuong et al., "Identifying user intents in Vietnamese spoken language commands and its application in smart mobile voice interaction," in *Proceedings of the Asian Conference on Intelligent Information and Database Systems*, pp. 190–201, Springer, Da Nang, Vietnam, March 2016.

[8] P.-N. Tran, V.-D. Ta, Q.-T. Truong, Q.-V. Duong, T.-T. Nguyen, and X.-H. Phan, "Named entity recognition for Vietnamese spoken texts and its application in smart mobile voice interaction," in *Proceedings of the Asian Conference on Intelligent Information and Database Systems*, pp. 170–180, Springer, Da Nang, Vietnam, March 2016.

[9] N. T. M. Huyen, A. Roussanaly, and H. T. Vinh, "A hybrid approach to word segmentation of Vietnamese texts," in *Proceedings of the International conference on language and automata theory and applications*, pp. 240–249, Springer, Tarragona, Spain, March 2008.

[10] H. Tseng, P. C. Chang, G. Andrew, and D. Jurafsky, "A conditional random field word segmenter for sighan bakeoff 2005," in *Proceedings of the fourth SIGHAN workshop on Chinese language Processing*, Jeju Island, Korea, January 2005.

[11] N. X. Bach, N. D. Linh, and T. M. Phuong, "An empirical study on POS tagging for Vietnamese social media text," *Computer Speech & Language*, vol. 50, pp. 1–15, 2018.

[12] J. Hu, G. Wang, F. Lochovsky, J. T. Sun, and Z. Chen, "Understanding user's query intent with wikipedia," in *Proceedings of the 18th international conference on World wide web*, pp. 471–480, Madrid Spain, April 2009.

[13] M. Mendoza and J. Zamora, "Identifying the intent of a user query using support vector machines," in *Proceedings of the International symposium on string processing and information retrieval*, pp. 131–142, Springer, Saariselkä, Finland, Auguest 2009.

[14] Y. Shi, K. Yao, L. Tian, and D. Jiang, "Deep LSTM based feature mapping for query classification," in *Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: Human language technologies*, pp. 1501–1511, San Diego CA, USA, January 2016.

[15] T. Kato, A. Nagai, N. Noda, R. Sumitomo, J. Wu, and S. Yamamoto, "Utterance intent classification of a spoken dialogue system with efficiently untied recursive autoencoders," in *Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue*, pp. 60–64, Germany, Auguest 2017.

[16] C. W. Goo, G. Gao, Y. K. Hsu et al., "Slot-gated modeling for joint slot filling and intent prediction," in *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, vol. 2, pp. 753–757, Louisiana, NO, USA, January 2018.

[17] M. Qiu, F. L. Li, S. Wang et al., "Alime chat: a sequence to sequence and rerank based chatbot engine," in *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, vol. 2, pp. 498–503Short Papers), Vancouver, Canada, January 2017.

[18] P. S. Huang, X. He, J. Gao, L. Deng, A. Acero, and L. Heck, "Learning deep structured semantic models for web search using clickthrough data," in *Proceedings of the 22nd ACM international conference on Information & Knowledge Management*, pp. 2333–2338, San Francisco California USA, October 2013.

[19] N. X. Bach, K. Hiraishi, N. Le Minh, and A. Shimazu, "Dual Decomposition for Vietnamese part-of-speech tagging," *Procedia Computer Science*, vol. 22, pp. 123–131, 2013.

[20] D. Q. Nguyen, D. Q. Nguyen, S. B. Pham, P.-T. Nguyen, and M. Le Nguyen, "From treebank conversion to automatic dependency parsing for Vietnamese," in *Proceedings of the International Conference on Applications of Natural Language to Data Bases/Information Systems*, pp. 196–207, Springer, Montpellier, France, June 2014.

[21] Y. Li, J. Y. Guo, Z. T. Yu, C. Mao, and Y. Xian, "Constituent-to-Dependency conversion for Vietnamese[J]," *Journal of Frontiers of Computer Science and Technology*, vol. 11, no. 4, pp. 599–607, 2017.

[22] H. Nguyễn Đ, M. J. Alves, and H. C. Nguyễn, *Vietnamese[M]*, Routledge, England, UK, 2018.

[23] L. C. A. Thompson, *Vietnamese Grammar*, University of Hawaii Press, Honolulu, HI, USA, 1988.

[24] L. A. Michaelis, "Expectation contravention and use ambiguity: the Vietnamese connective cũng," *Journal of Pragmatics*, vol. 21, no. 1, pp. 1–36, 1994.

[25] H. V. Luong, "Plural markers and personal pronouns in Vietnamese person reference: an analysis of pragmatic ambiguity and native models," *Anthropological Linguistics*, pp. 49–70, 1987.

[26] C. Miller, "Structural ambiguity in the Vietnamese relative clause," *Mon-Khmer Studies*, vol. 5, pp. 233–267, 1976.

[27] B. Tesar, "Enforcing grammatical restrictiveness can help resolve structural ambiguity," vol. 21, pp. 443–456, in *Proceedings of the West Coast Conference on Formal Linguistics*, vol. 21, pp. 443–456, Cascadilla Press, USA, April 2002.

[28] M. S. Zhang, Z. L. Deng, W. X. Che, and T. Liu, "Combining statistical model and dictionary for domain adaption of Chinese word segmentation," *Journal of Chinese Information*, vol. 26, no. 2, pp. 8–13, 2012.

[29] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa, "Natural language processing (almost) from scratch," *Journal of Machine Learning Research*, vol. 12, no. ARTICLE, pp. 2493–2537, 2011.

[30] R. Collobert and J. Weston, "A unified architecture for natural language processing: deep neural networks with multitask learning," in *Proceedings of the 25th international conference on Machine learning*, pp. 160–167, Helsinki Finland, July 2008.

[31] Y. Bengio, R. Ducharme, P. Vincent, and C. Janvin, "A neural probabilistic language model," *Journal of Machine Learning Research*, vol. 3, no. Feb, pp. 1137–1155, 2003.

[32] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[33] O. Vinyals, S. V. Ravuri, and D Povey, ". Revisiting recurrent neural networks for robust ASR," in *Proceedings of the 2012 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pp. 4085–4088, IEEE, Kyoto, Japan, March 2012.

[34] T. Mikolov, M. Karafiát, L. Burget, J. Cernocky, and S. Khudapur, "Recurrent neural network based language model," in *Proceedings of the Interspeech Communication Association*, vol. 2, no. 3, pp. 1045–1048, chiba, Japan, September 2010.

[35] I. Sutskever, J. Martens, and G. E. Hinton, "Generating text with recurrent neural networks," in *Proceedings of the International conference on Machine Learnung ICML*, Washington, USA, July 2011.

[36] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," 2013, https://arxiv.org/abs/1301.3781.

Hindawi
Scientific Programming
Volume 2022, Article ID 5747647, 10 pages
https://doi.org/10.1155/2022/5747647

Research Article

Design of National Sports Action Feature Extraction System Based on Convolutional Neural Network

Yajun Pang [1,2]

[1]*College of Physical Education, Luoyang Institute of Science and Technology, Luoyang, Henan 471023, China*
[2]*Henan Province Engineering Research Center of Industrial Intelligent Vision, Luoyang, Henan 471023, China*

Correspondence should be addressed to Yajun Pang; yajun.pang@lit.edu.cn

Received 2 December 2021; Accepted 27 December 2021; Published 25 February 2022

Academic Editor: Baiyuan Ding

Human action recognition is one of the hotspots in computer vision research. Its purpose is to detect and recognize target actions from videos, so that computer systems can understand human actions, and thus it has great research significance. Based on the action features of famous sports, this paper proposes an action recognition scheme based on RGB-D video compression to establish action features and deep learning as a means of recognition. By establishing the connection between the bone data of the three-dimensional data type and the depth image data, the depth sequence is analyzed and expressed as a three-level structure diagram sequence, which is the overall figure sequence, partial figure sequence, and joint point figure sequence, and then passes through the two-way pool. The sorting algorithm extracts the action features in the three picture sequences and compresses and generates three types of structured images of the corresponding picture sequences, and these three types of structured images are used as the feature expression of the video. When constructing a three-level structure diagram sequence, the innovation of this paper is to splice the extracted key unit image blocks to obtain a three-level structured moving image based on the three-key unit splicing, so that the image is not only retained. In addition to time-space information, the structure information of the depth image is also strengthened, and the amount of calculation is reduced at the same time. Finally, the three types of structured images are input into the convolutional neural network, respectively, and the judgment and recognition results obtained are multiplicatively fused to obtain the final recognition rate of the action.

1. Introduction

In human-centered computer vision research (such as human detection, tracking, human posture estimation, and human motion recognition), human motion recognition is widely used; for example, video surveillance, human-machine interface, home assistance, human-machine interaction, and intelligent driving have become an important research direction in computer vision research [1–5].

According to the complexity and duration of the action, action recognition can be roughly divided into four types: gesture recognition, action recognition, interaction recognition, and group activity recognition. Specifically, gesture recognition is defined as expressing people's thoughts, opinions, and emotions through the basic movements or positions of hands, arms, human body, or head: Typically, "waves" and "nods." The posture duration is relatively short and the complexity is low. Action is defined as an activity completed by a single human body mobilizing multiple parts of the body; that is to say, an action is a combination of multiple postures, such as "walking" and "boxing." The interactive action is mainly completed by two subjects: people and things, and people and people [6–11]. This means that interaction expresses the interaction of people or characters, such as "hug" and "playing the guitar." Group activities are the most complicated type of action. It may combine three types of actions: posture, action, and interaction, involving two or more people or objects, such as "two teams play basketball" and " group meeting," shown in Figure 1.

The realization process of action recognition is generally divided into three stages: first, target detection, then feature extraction, and finally feature analysis and judgment and recognition. There are corresponding studies for each stage, and the goal is to achieve efficient action recognition. Moving target detection is not the focus of this

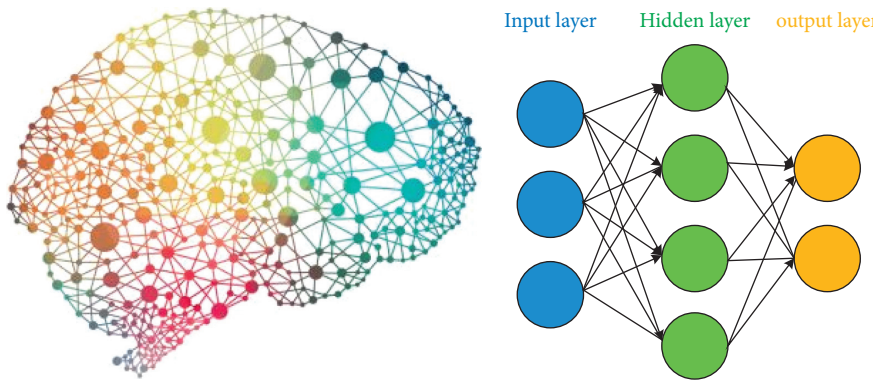


Figure 1: Neural networks.

article, so it will be not repeated here. In the feature extraction stage, the traditional method is to manually select features. However, in most cases, specific algorithm analysis is performed on the data characteristics. The generalization ability is poor, and there are often situations where the effect of different datasets is quite different. The complexity of the processing process is on different datasets. Moreover, with the advent of the era of big data, the development of datasets is gradually moving towards larger data volumes, more data categories, and wider data ranges, making it more difficult to analyze data and extract features. As it becomes more difficult to determine the features to be extracted, the step of feature extraction gradually abandons specific customization and tends to be general. People are more hopeful that the action features in the video can be extracted through nonmanual methods, which reduces the difficulty of qualitative extraction of features in the early stage and improves the efficiency of large-scale action data processing. With the deepening of research, researchers use models and algorithms to characterize features, such as image-based representation, model-based representation, time-space-based representation, and so on. In the last stage, when analyzing data features for classification, a classifier is needed for identification; that is, the extracted features are used as an abstract representation of the original input image and input into the classifier. By using the model training parameters in the classifier, a similar match with the input data is found. The main purpose of human action recognition is that, for the test data and the label data predicted by the classification model obtained through training, the higher the fit with the actual label data, the better. With the evolution of features from simple to complex, the classifier is also a process from simple to complex, from the linear binary classifier at the beginning to the logistic regression classifier, as well as a series of traditional excellent learning algorithms such as SVM and HMM. It is a typical classifier model. However, the common feature of these classifiers is that different types of features need to be matched with different classifiers, which greatly reduces the computational efficiency. Deep learning is a branch of traditional machine learning that can make up for the shortcomings of the above algorithms and is a popular algorithm among recent classification algorithms [12, 13].

Deep learning is an important branch of traditional machine learning. With the development of hardware equipment and big data, deep learning has gradually become one of the hot points in the research of computer vision research. Deep learning builds a hierarchical learning and training model and establishes a progressive learning mechanism between input and output data. With each step forward, the extracted feature dimension rises by one level, so that the final trained model can extract the original. The high-dimensional features of the image are conducive to the final recognition and judgment. Because deep learning has the characteristics of autonomous learning and does not require artificial design of related algorithms, it is a more efficient and generalized feature extraction method. So far, deep learning has good experimental results in the fields of target detection, target recognition, image processing, and so on. One of the more prominent advantages of deep learning is that its processing speed for big data is much faster than traditional hand-craft features. With the rapid development of the Internet, deep learning is more in line with the characteristics of the times and can play a role from massive amounts of data. Information makes the future development of deep learning very impressive [14–17].

In recent years, the development of the field of deep learning has provided new methods for the judgment and recognition of human actions in the later stages. Deep learning algorithms can extract higher-level action features and give better classification results. The system that applies deep learning for action recognition has achieved a high recognition rate. The good classification effect of deep learning has been recognized, and it has been practically applied in driverless cars, medical image recognition, and image retrieval. Convolutional neural networks are a common way to learn high-level features in deep learning. G.W. Taylor et al. proposed an action recognition framework based on convolutional neural networks, which directly process continuous image sequences and use convolutional neural networks to learn color texture maps. S. Ji et al. used a three-dimensional convolutional neural network to extract human action information that contains both time and space features. A. Karpathy et al. conducted a comparative study on a variety of convolutional neural networks in the time domain and provided an empirical evaluation of convolutional neural networks in high-

resolution video classification. At the same time, the article also pointed out that the use of convolutional neural networks to process each frame in the video sequence alone has almost the same result as processing a series of frames at the same time; that is, the convolutional neural network for comparative research in the article learns the characteristics of time and space [18–22]. The above is not well integrated. In order to better mine the spatial and temporal features in the action sequence, K. Simonyan et al. used convolutional neural networks to process spatial data streams and temporal data streams separately and finally used specific methods to perform learning results on different features [23–28]; see Figure 2.

In order to solve the above problems, this paper proposes a novel action recognition framework, which compresses the action sequence based on depth information, effectively compresses a video sequence into several pictures, and finally uses the convolutional neural network to improve the picture and learning and classification capabilities to complete the learning and classification of action sequences. This paper applies sorting pooling to the depth image sequence to obtain three levels of dynamic images, compress the depth video sequence, and then use three convolutional neural networks to judge and recognize the three types of images, respectively, and finally merge the results to obtain best effect. In the research, we introduced AlexNet, which has a good effect on human action recognition in deep learning and improved the accuracy of action recognition. Compared with other algorithms, the recognition effect and robustness of this method are far stronger than other methods, and the trained network can allow anyone to control the UAV by gestures according to regulations, which has good universality [29–33].

2. Convolutional Neural Network

One of the most common algorithms for image recognition is a model built based on a convolutional neural network (CNN), which can be regarded as a forward feedback neural network that includes convolution operations. After layers of convolution and pooling operations, the model can gradually extract the features of the target image, and the proposed features have translation invariance. The biggest feature of convolutional neural network is weight sharing and local perception, so when training, it needs to learn a huge number of feature parameters. What needs to be explained here is that each extraction of feature values requires multiple core convolution kernels. Different cores correspond to different features, as shown in Figure 3. At the same time, its corresponding two-dimensional spatial characteristics help it in computer vision tasks and demonstrate a good characterization ability on the problem. Therefore, the classification model based on CNN mainly includes the following modules.

2.1. Convolutional Layer. After inputting the current data features in the convolutional layer, by setting the appropriate number of convolutions, the size of the convolution kernel, and the convolution method, the feature expression of the input data is completed. After completing this operation, you can multiply the data and the convolution kernel to get each feature point on the corresponding feature map. Traverse all the data according to the convolution step length, and realize the convolution between the convolution kernel parameters and the input data in the receptive field. The process can be expressed as follows:

$$\begin{aligned}\mathbf{Z}^{l+1}(i,j) &= \left[\mathbf{Z}^{l} \otimes \mathbf{w}^{l+1}\right](i,j) + b \\ &= \sum_{k=1}^{K_l} \sum_{x=1}^{f} \sum_{y=1}^{f} \left[\mathbf{Z}_k^{l}(s_0 i + x, s_0 j + y)\mathbf{w}_k^{l+1}(x,y)\right] + b, \\ (i,j) \in \{0,1,\ldots,L_{l+1}\}, L_{l+1} &= \frac{(L_l + 2p - f)}{s_0} + 1,\end{aligned} \tag{1}$$

where b represents the corresponding deviation, $Z^{l+1}|_{,}Z^{l}$ are, respectively, used to represent the output and input values of the neural network. L_{l+1} represents the feature size of the output image, $Z(i,j)$ represents the feature value, and K represents the number of convolution kernels, that is, the number of channels; f, s_0, p represent the number of convolution operations involved in this operation, convolution kernel size.

When the convolution step size and the size of the convolution kernel are both 1, it is called unit convolution, and the convolutional layer composed of these unit convolutions is called the net in the net. In the case of unchanged input features, unit convolution can perform feature fusion on multiple channels, which helps to reduce the number of corresponding model parameters, thereby reducing a certain amount of calculation.

The convolutional layer has two characteristics of local perception and weight sharing at the same time. The neurons in the front and back layers are connected in pairs. For example, in a 1000*1000 image, the input layer has 106 nodes. Therefore, it needs to learn a huge number of feature parameters during the training process. The feature value is extracted when the convolution operation slides in the previous layer of the convolution kernel, and the weight of the core parameter needs to be kept unchanged during each operation, and it is only at the same time. Establish a connection with a part of the neurons in the previous layer. It can be seen that the corresponding local perception not only is more in line with human's cognitive characteristics of things from part to the whole, but also greatly reduces the number of characteristic parameters. The process of applying the learning information in a certain local area to other places is called weight sharing. In simple terms, the whole image is filtered on the image, which is also conducive to greatly reducing the parameters.

Each feature extraction needs to use multiple convolution kernels, the features obtained by different convolution kernel extraction are different, and each convolution kernel will extract a certain aspect of the feature. At the same time, its two-dimensional spatial characteristics help it have good characterization capabilities for computer vision tasks.

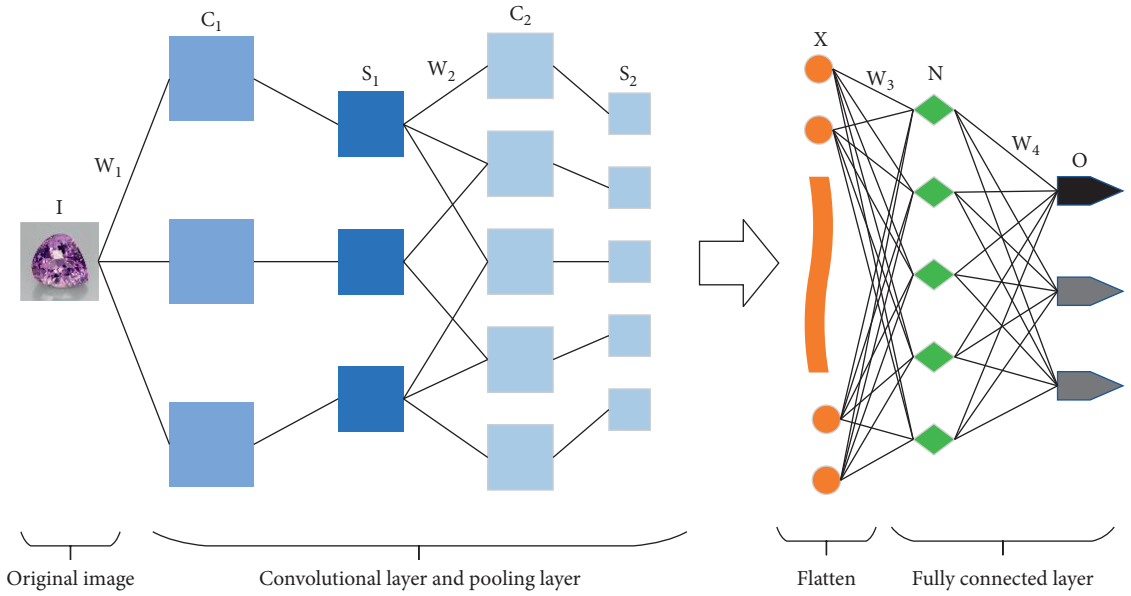


Figure 2: Convolutional neural network structure.

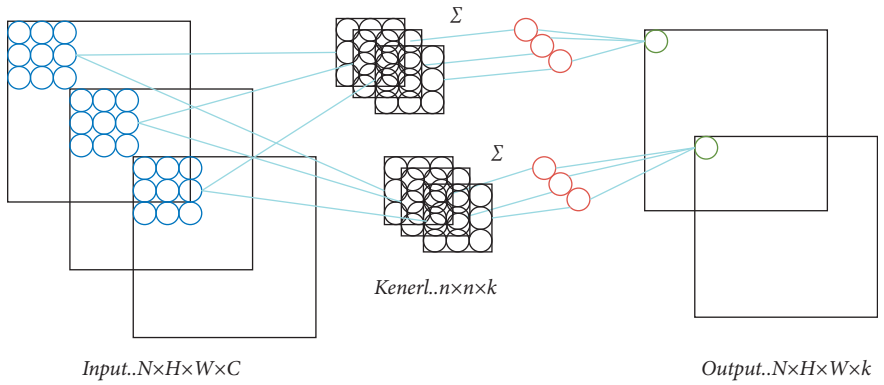


Figure 3: Process of convolutional neural network structure.

2.2. Pooling Layer. After the data input of the previous layer is completed, it is passed to the pooling layer. In this layer, the data filtering of the previous layer is mainly completed. The standard for filtering is the data extraction standard of the convolutional layer. This process can be understood, in order to imitate the human visual system to complete the dimensionality reduction of the data and finally obtain the image representation features of higher-level features. Pooling helps to reduce the redundancy of information, and at the same time it can improve the invariance of the model scale, thereby helping to avoid the adverse effects of the model due to overfitting. In general, the pooling layer can be further divided into mean pooling and maximum pooling. The advantage of the latter is that it can learn the edge and texture structure of the image. The advantage of the former is that it can effectively reduce the deviation of the estimated mean value and improve the anti-interference ability of the established model. The formula of the pooling operation can be expressed as

$$\mathbf{A}_k^l(i,j) = \left[\sum_{x=1}^{f}\sum_{y=1}^{f}\left[\mathbf{A}_k^l(i,j)(s_0 i + x, s_0 j + y)^p\right]\right]^{\frac{1}{p}}. \quad (2)$$

In the above formula, s_0 represents the pooling step size, when $p = 1$ is the average pooling; when p tends to infinity, it is the maximum pooling.

2.3. Fully Connected Layer. The fully connected layer is composed of several neurons "holding hands" with each other, and each neuron in the latter part is interconnected with the neuron in the former part. This is equivalent to the aforementioned forward feedback network, as the end layer of the convolution model. It is the integration of the features provided by the previous convolution and pooling operations. The feature map is expanded into a one-dimensional column vector and used as the input corresponding to the fully connected layer. It is easy to realize the nonlinearity of the input feature by using the activation function and finally

extract feature vectors with more expressive ability and then realize feature classification.

All in all, the structure corresponding to a typical convolutional neural network used to process image tasks can usually be expressed as follows.

Because convolutional neural networks have certain advantages in processing images, they are generally regarded as a commonly used method in the field of image recognition. In the traditional network structure, there are many classic structures, such as AlexNet network, GoogleNet network, VGGNet network, and ResNet network.

2.3.1. Gradient Descent. The standard gradient descent can be described as

$$\theta = \theta - \eta \cdot \nabla_\theta J(\theta). \tag{3}$$

The standard gradient descent refers to the replacement of the secondary parameters of the gradient of the overall calculation example. This standard gradient descent method has certain drawbacks such as relatively slow calculation speed and certain application limitations.

2.3.2. Stochastic Gradient Descent (SGD). Different from calculating the gradient after calculating the loss of all samples in the standard gradient descent case, SGD calculates the gradient once for each sample and updates the parameters. It can be described as

$$\theta = \theta - \eta \cdot \nabla_\theta J\left(\theta; x^{(i)}; y^{(i)}\right). \tag{4}$$

2.3.3. Minibatch Gradient Descent (MBGD). This method is a relative compromise between batch stochastic gradient descent and gradient descent. Among them, the main idea of the minibatch gradient descent method is as follows: based on a dataset of n training samples, update the corresponding parameters in real time and select a minibatch data sample of size m ($m < \mathrm{n}$) to calculate its corresponding. The gradient of the formula is as follows:

$$\theta = \theta - \eta \cdot \nabla_\theta J\left(\theta; x^{(i;i+n)}; y^{(i;i+n)}\right). \tag{5}$$

AdaGrad is a method of adaptive learning rate, which implements a high learning rate for low-frequency parameters and a low learning rate for high-frequency parameters. This feature makes it more suitable for processing sparse data. Standard parameters are a very small constant, generally 10e-8, which represents the global learning rate, and usually, the final gradient accumulation variable r.

$$\begin{aligned} g &\leftarrow \nabla_\theta J(\theta), \\ r &\leftarrow r + g^2, \\ \Delta\theta &\leftarrow \frac{\delta}{\sqrt{r+\epsilon}} \cdot g. \end{aligned} \tag{6}$$

3. Action Feature Model Algorithm Based on Convolutional Neural Network

In order to simplify the processing of video sequences and make the computer's judgment and recognition of the types of human actions more efficient, a method of compressing the depth of human actions is proposed, shown in Figure 4. This method considers the spatial and temporal characteristics of the video at the same time, which can save the video action characteristics and reduce the information redundancy. The algorithm includes the following steps: remove the background interference in the depth video, leaving only the shape of the human body. After removing the background interference from a k-frame depth video sequence, it is expressed as

$$\langle d_1, d_2, ..., d_t, d_k \rangle, \tag{7}$$

where dt represents the average value of the depth features of all frames as of t time. The deep depth image is shown in Figure 5.

Rank pooling processing for depth videos: At each time t, define a score value, and the score value must meet the conditions; the more the current frame number, the greater the score value.

Perform rank pooling on deep video sequences to extract features. The process of rank pooling is to find an optimal solution that satisfies the following objective function:

$$\begin{aligned} &\arg\min \frac{1}{2}\|\omega\|^2 + \lambda \sum_{i>j} \xi_{ij}, \\ &s.t. \omega^T\left(d_i - d_j\right) \geq 1 - \xi_{ij}, \xi_{ij} \geq 0. \end{aligned} \tag{8}$$

In the depth image, each joint point is used as the center point for extension, and the depth image is cropped using a frame of size q∗p to obtain the image block of the joint point. In order to make the extracted image blocks have the same scale feature, we find the maximum range of motion for the same unit in a video frame sequence and define it as mask. For each unit, use 0 to fill the unit to the mask size. In this way, the relative scale between the units can be maintained, which is very necessary for the preservation of spatial information, as shown in Figure 6.

Convolutional neural network has an excellent effect on image and speech processing. By using a single neuron to respond to the coverage and surrounding pixels, convolutional neural networks have unique advantages for large-scale image processing. The structural design of the weight sharing of the convolutional neural network is more like the principle of the human neural network. Through the pooling layer, the weight sharing fully linked layer design greatly reduces the parameters of the entire neural network. The design uses multiple convolutional templates for feature extraction, which is far better than the existing machine learning methods in terms of image classification and recognition. Deep learning through convolutional neural networks can automatically extract useful features from images and use these features to

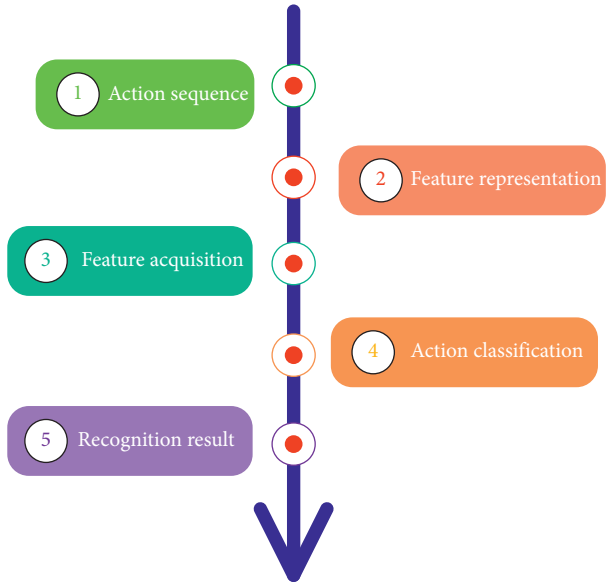


Figure 4: Action recognition framework flowchart.

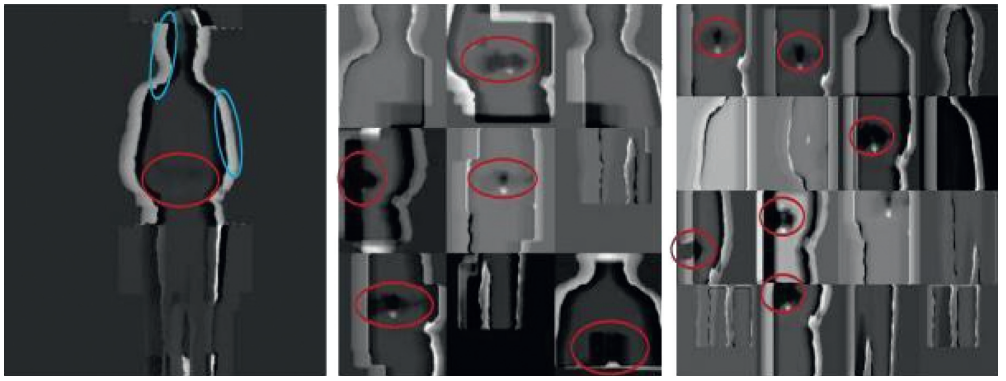


Figure 5: Deep depth image.

classify images. The level of perfection of learning features exceeds many existing methods of manually specifying features.

In order to achieve a good classification effect, we adopted the AlexNet network structure, which has achieved remarkable results on the ImageNet dataset, as the neural network we use to classify gestures. The network structure has 5 convolutional layers and 3 full link layers.

There is the following relationship between every two convolutional layers of the network structure, as shown in Figure 7.

Among the three fully connected layers in AlexNet, each fully connected layer contains 4096 neurons. Such a network maximizes the multiclass logistic regression objective; that is, it maximizes the average log probability of the correct label in the training sample under the prediction distribution, thereby making the classification more accurate. In order to make the convolutional neural network get good results faster, the model trained on the ImageNet of the AlexNet network is used in this article to initialize the network parameters.

3.1. Hardware Platform. The workstation for training in the experiment is equipped with an Intel E5-2300 CPU and 16 GB DDR3 memory using GPUNvidia TitanX to accelerate the training process of the neural network. The predicted result is plotted in Figure 8.

3.2. Software Platform. The deep learning platform used in the experiment is Caffe (Convolutional Architecture for Fast Feature Embedding), which is a general framework for deep learning algorithms. The framework uses many libraries that can perform fast calculations, models for fast data storage, and function templates that can be directly called, allowing developers to quickly implement the network structure envisioned. The architecture abstracts many common operations of convolutional neural networks, and they are implemented with CPU and GPU, respectively. The entire calculation can be seamlessly switched between CPU and GPU. Caffe allows users to implement convolutional neural networks only by specifying the network structure in the configuration file.

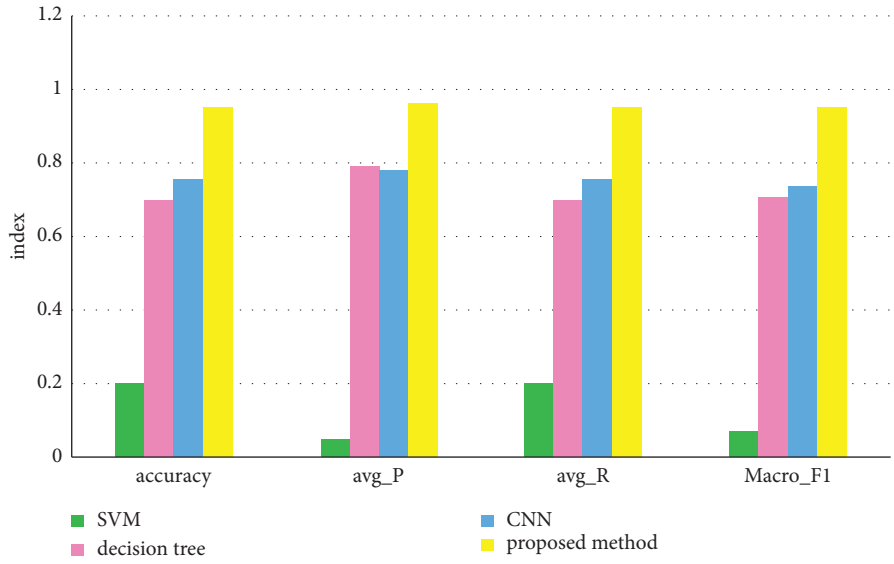


Figure 6: Statistical table of evaluation indexes for different algorithms.

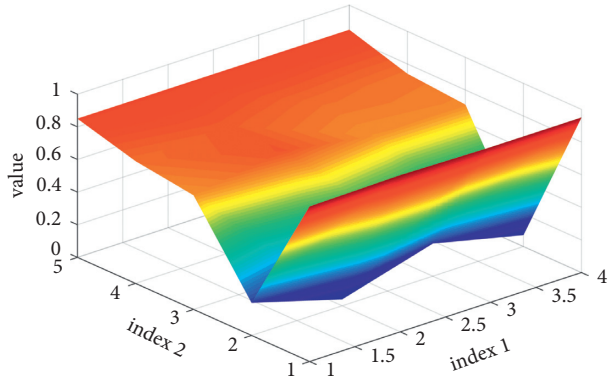


Figure 7: Value vs. index.

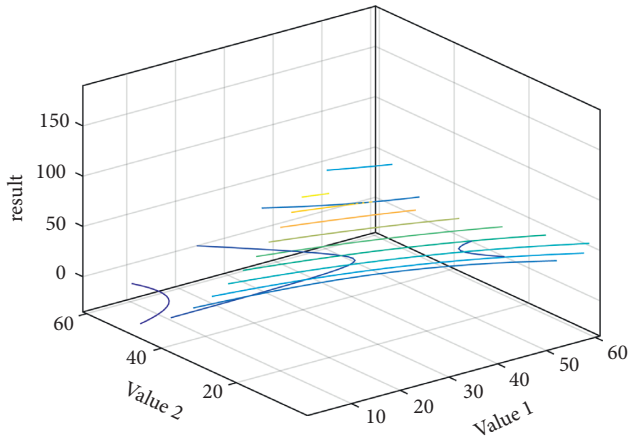


Figure 8: Predicted result.

The parameters of the network set in the experiment are as follows: The network uses 256 training images (batch size) for one iteration. According to the size of the training dataset, a total of about 90 cycles (epoch) are trained. In these 90 cycles, the initial learning rate is set to 0.001 (base learning rate), which drops to 0.0001 after 60 cycles. For the rest of the training parameters, we refer to the method proposed by A. Krizhevsky et al.

In the verification of the algorithm in this article, a total of five human motion datasets are used. They are MSR Activity 3D dataset, G3D dataset, MSR Daily Activity dataset, SYSU 3D HOI dataset, and UTD-MHAD dataset. These datasets cover most types of actions, including single actions, in-game actions, daily action actions, human and object interaction actions, and fine-grained actions. Therefore, these datasets can show that the algorithm proposed in this paper is universal.

3.2.1. MSR Action 3D Dataset. MSR Action 3D dataset contains 20 simple actions performed by 10 people facing the camera, and each person performs each action 2 to 3 times. The experiment adopts the method of cross-validation. The human movement data labeled 1,3,5,7, and 9 are trained, and the movement data labeled 2,4,6,8, and 10 are tested.

3.2.2. G3D Dataset. G3D is a game action type dataset, which is a series of action data shots for scenes in the game, including 20 game actions displayed by 10 people. The experimental verification method is to keep the first five labeled people as the training data and the last five labeled people as the test data.

3.2.3. MSR Daily Activity 3D Dataset. The MSR Daily Activity 3D dataset contains 16 actions displayed by 10 people, and each person performs each action twice, once in a standing position and once in a sitting position. Most of this dataset contains interactive actions of characters. The experiment adopts a cross-validation method, and people's movement data with labels of 1, 3, 5, 7, 9, 11, 13, and 15 are used for training, and those with labels of 2, 4, 6, 8, 10, 12, 14, and 16 action data are tested.

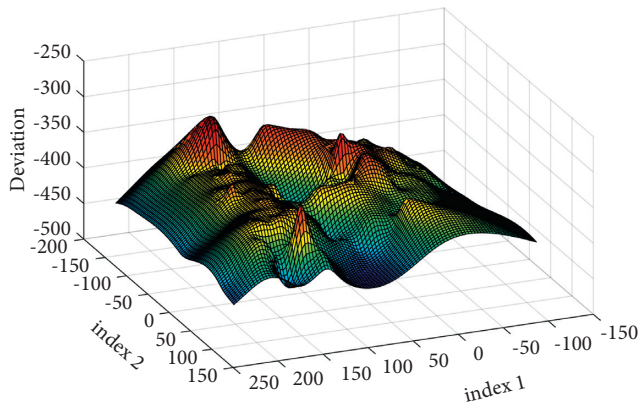


Figure 9: Deviation.

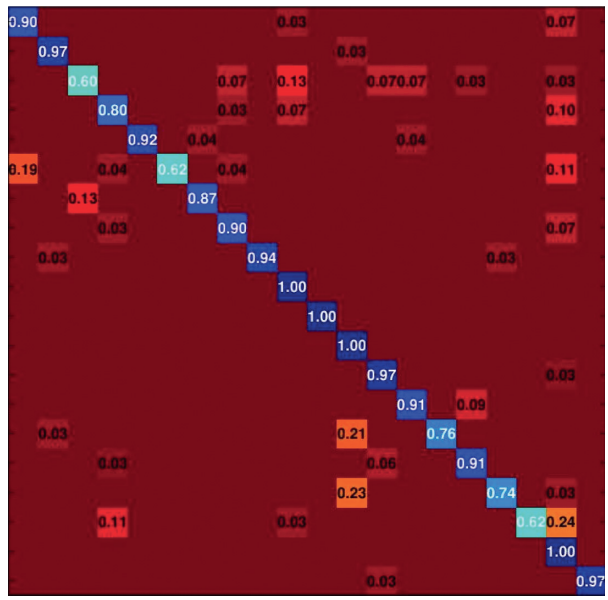


Figure 10: Matrix of MSR Action 3D.

3.2.4. SYSU 3D HOI Dataset. SYSU 3D HOI dataset was designed and collected by Sun Yat-Sen University, and it is aimed at the action types of character interaction. This dataset consists of 40 people showing 12 actions. For each action, each person interacts with the following six items: phone, chair, school bag, wallet, mop, and broom. The experiment adopts the method of cross-validation, the action data with odd numbers of people is trained, and the action data with even numbers is tested.

3.2.5. UTD-MHAD Dataset. UTD-MHAD dataset contains 27 actions displayed by 8 people, 8 people contain 4 men and 4 women, and each person performs each action 4 times. The experiment adopts the method of cross-validation. The human movement data labeled 1, 3, 5, and 7 are trained, and the movement data labeled 2, 4, 6, and 8 are tested. The deviation is plotted in Figures 9 and 10.

4. Conclusions

This paper analyzes the framework structure of action recognition and designs a new framework from two directions: feature extraction and feature classification. In terms of feature extraction, in order to retain more information about the original video, this article starts from the time-space structure and uses the hierarchical sorting and pooling dynamic depth map DDI to obtain three types of dynamic depth maps, namely, the whole body DDI, part DDI, and joint DDI; in the feature classification stage, this

article uses the AlexNet convolutional neural network provided by the deep learning platform Caffe to train the model; multiplicative fusion of the final result can get a better action recognition rate.

4.1. Bidirectional Pooling Sorting Algorithm. The pooling algorithm is based on the premise of a forward and backward time assumption. It extracts and sorts video features and compresses the sorting information on a picture, which is verified on RGB color images. It is a video compression method with excellent performance. Two-way pooling reverses this assumption of forward and backward time and trains them separately. This is quite effective for actions that are sensitive to time information and can obtain more comprehensive information. On the other hand, more data can make the model training more adequate.

4.2. Hierarchical Image Segmentation Combination. This paper proposes a simple and effective way of extracting video spatial information. The human body is divided according to joint points and combined according to the whole body, parts, and joint points. In the three types of pictures obtained in this way, the whole body DDI can provide body contour information, and the partial and joint DDI can provide detailed information, which plays a complementary role and is of great significance for the next step of recognition. And this way of segmentation is easy to understand, and it is very convenient for other researchers to reverify.

4.3. The Challenge of Human Motion Recognition. The difference between within and between classes is the same action, and the performance of different people may vary greatly. For example, because of the differences between individuals, people with different simple actions of running have different speeds and step lengths. A robust action recognition method should have good generalization; the environment here can be divided into the complex background where the action is executed and the camera environment. The environment in which the action occurs is an important differentiating factor. In a complex and chaotic background, it is difficult to accurately track and locate feature points of interest. Moreover, important parts of the human body are likely to be occluded by objects and other people. At the same time, the lighting conditions will further affect the appearance of the person's contours, which will greatly interfere with the recognition of human movements. The problem of occlusion can be alleviated by using multiple cameras to observe movement from different perspectives, but this will face the problem of synchronization and cannot achieve real time. In addition, a moving camera will increase the difficulty of positioning and tracking the human body and will also cause the ability to change the scale of the human body, which is not affected by changes between classes. At the same time, it should be able to distinguish the differences between different categories. However, as the types of actions increase, there is a certain overlap between actions, which makes the recognition more difficult.

Generally, research assumes that actions are easily segmented in the time dimension. Although this assumption reduces the segmentation burden in the recognition task, another segmentation process must be added in advance, which makes the real-time performance of the entire recognition worse. The speed of people performing actions varies greatly, and it is difficult to determine the starting point of the action, which has the greatest impact when extracting features from the video to represent the action. Therefore, a robust human action recognition method should be invariant to the execution speed of the action.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by Science and Technology Projects of Henan Science and Technology Department, Assessment of action in sports video based on multi-feature fusion, 182102310041, Science and Technology Projects of Henan Science and Technology Department, Assessment of Taichi Action Based on Vision Transformer, 222102320016, and Sanmenxia Vocational and Technical College Project, Assessment of action in low-quality large-displacement sports videos, SZYGCCRC-2020-005.

References

[1] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, Las Vegas, NV, USA, 2016.

[2] M. Yun, J. Zhao, J. Zhao, X. Weng, and X. Yang, "Impact of in-vehicle navigation information on lane-change behavior in urban expressway diverge segments," *Accident Analysis & Prevention*, vol. 106, no. 1, pp. 53–66, 2017.

[3] S. Kumar Dwivedi, R. Amin, V. Satyanarayana, and R. Chaudhry, "Blockchain-based secured event-information sharing protocol in internet of vehicles for smart cities," *Computers & Electrical Engineering*, vol. 86, no. 1, pp. 1–9, 2020.

[4] Z. Khan and S. Amin, "Bottleneck model with heterogeneous information," *Transportation Research Part B: Methodological*, vol. 112, no. 1, pp. 157–190, 2018.

[5] J. M. Cairney, K. Rajan, D. Haley et al., "Mining information from atom probe data," *Ultramicroscopy*, vol. 159, no. 1, pp. 324–337, 2020.

[6] J. Yu and P. Lu, "Learning traffic signal phase and timing information from low-sampling rate taxi GPS trajectories," *Knowledge-Based Systems*, vol. 110, no. 1, pp. 275–292, 2016.

[7] K. P. Wijayaratna, V. V Dixit, L. Denant-Boemont, and S. Travis Waller, "An experimental study of the Online Information Paradox: does en-route information improve road network performance?" *PLoS ONE*, vol. 12, no. 9, pp. 184–191, 2017.

[8] Z. Wang, H. Ren, Q. Shen, W. Sui, and X. Zhang, "Seismic performance evaluation of a steel tubular bridge pier in a five-span continuous girder bridge system," *Structures*, vol. 31, no. 1, pp. 909–920, 2021.

[9] S. Nakayama and J. Takayama, "Traffic network equilibrium model for uncertain demands," in *Proceedings of the 82nd Transportation Research Board Annual Meeting*, Washington, DC, USA, 2003.

[10] H. Shao, W. H. K. Lam, and M. L. Tam, "A reliability-based stochastic traffic assignment model for network with multiple user classes under uncertainty in demand," *Networks and Spatial Economics*, vol. 6, no. 3, pp. 173–204, 2019.

[11] A. Chen, J. Kim, S. Lee, and Y. Kim, "Stochastic multi-objective models for network design problem," *Expert Systems with Applications*, vol. 37, no. 2, pp. 1608–1619, 2020.

[12] H. Wang, W. H. K. Lam, X. Zhang, and H. Shao, "Sustainable transportation network design with stochastic demands and chance constraints," *International Journal of Sustainable Transportation*, vol. 9, no. 2, pp. 126–144, 2015.

[13] S.-M. Hosseininasab and S.-N. Shetab-Boushehri, "Integration of selecting and scheduling urban road construction projects as a time-dependent discrete network design problem," *European Journal of Operational Research*, vol. 246, no. 3, pp. 762–771, 2015.

[14] S.-M. Hosseininasab, S.-N. Shetab-Boushehri, S. R. Hejazi, and H. Karimi, "A multi-objective integrated model for selecting, scheduling, and budgeting road construction projects," *European Journal of Operational Research*, vol. 271, no. 1, pp. 262–277, 2018.

[15] M. Johnson, M. Schuster, Q. Le et al., "Google's multilingual neural machine translation system: enabling zero-shot translation," *Transactions of the Association for Computational Linguistics*, vol. 5, no. 1, pp. 339–351, 2017.

[16] M. D. Moreno, "Translation quality gained through the implementation of the iso en 17100:2015 and the usage of the blockchain," *Babel*, vol. 1, no. 2, pp. 1–9, 2020.

[17] X. Wang, X. Yu, L. Guo, F. Liu, and L. Xu, "Student performance prediction with short-term sequential campus behaviors," *Information*.vol. 11, no. 4, p. 101, 2020.

[18] Q. Guo, Z. Zhu, Q. Lu, D. Zhang, and W. Wu, "A dynamic emotional session generation model based on Seq2Seq and a dictionary-based attention mechanism," *Applied Sciences*, vol. 10, no. 6, pp. 1–10, 2020.

[19] H. Ren, Xi Mao, W. Ma, J. Wang, and L. Wang, "An English-Chinese machine translation and evaluation method for geographical names," *ISPRS International Journal of Geo-Information*, vol. 9, no. 3, pp. 193–201, 2020.

[20] J. Arús-Pous, T. Blaschke, S. Ulander, J.-L. Reymond, H. Chen, and O. Engkvist, "Exploring the GDB-13 chemical space using deep generative models," *Journal of Cheminformatics*, vol. 11, no. 1, pp. 20–29, 2019.

[21] T. Moon, T. In Ahn, and E. S. Jung, "Long short-term memory for a model-free estimation of macronutrient ion concentrations of root-zone in closed-loop soilless cultures," *Plant Methods*, vol. 15, no. 1, pp. 1–12, 2019.

[22] N. Pourdamghani and K. Knight, "Neighbors helping the poor: improving low-resource machine translation using related languages," *Machine Translation*, vol. 33, no. 3, pp. 239–258, 2019.

[23] L. Bote-Curiel, S. Muñoz-Romero, A. Gerrero-Curieses, and J. L. Rojo-Álvarez, "Deep learning and big data in healthcare: a double review for critical beginners," *Applied Sciences*, vol. 9, no. 11, pp. 1–11, 2019.

[24] J. Zhang and T. Matsumoto, "Corpus augmentation for neural machine translation with Chinese-Japanese parallel corpora," *Applied Sciences*, vol. 9, no. 10, pp. 1–12, 2019.

[25] Y. Chen, Y. Ma, X. Mao, and Q. Li, "Multi-task learning for abstractive and extractive summarization," *Data Science and Engineering*, vol. 4, no. 1, pp. 14–23, 2019.

[26] P. Zhou and Z. Jiang, "Self-organizing map neural network (SOM) downscaling method to simulate daily precipitation in the Yangtze and Huaihe River Basin," *Climatic and Environmental Research*, vol. 21, no. 5, pp. 512–524, 2016.

[27] X. Xiao, "Analysis on the employment psychological problems and adjustment of retired athletes in the process of career transformation," *Modern Vocational Education*, vol. 5, no. 12, pp. 216-217, 2018.

[28] S. Sahoo and M. K. Jha, "Pattern recognition in lithology classification: modeling using neural networks, self-organizing maps and genetic algorithms," *Hydrogeology Journal*, vol. 25, no. 2, pp. 311–330, 2016.

[29] Y. Zhou and B. Yang, "Sports video athlete detection using convolutional neural network," *Journal of Natural Science of Xiangtan University*, vol. 39, no. 1, pp. 95–98, 2017.

[30] J. Pang, "Research on the evaluation model of sports training adaptation based on self-organizing neural network," *Journal of Nanjing Institute of Physical Education*, vol. 16, no. 1, pp. 74–77, 2017.

[31] G. Querzola, C. Lovati, C. Mariani, and L. Pantoni, "A semi-quantitative sport-specific assessment of recurrent traumatic brain injury: the TraQ questionnaire and its application in American football," *Neurological Sciences*, vol. 40, no. 9, pp. 1909–1915, 2019.

[32] J. Wang, X. Luo, and H. Yan, "Correlation analysis between injuries and functional movement screening for athletes of the National Shooting Team," *Journal of Capital Institute of Physical Education*, vol. 5, no. 4, pp. 352–355, 2016.

[33] G. Ma, "Research on the design of juvenile football players' sports injury prediction model," *Automation Technology and Application*, vol. 277, no. 7, pp. 141–144, 2018.

Hindawi
Scientific Programming
Volume 2022, Article ID 7102267, 8 pages
https://doi.org/10.1155/2022/7102267

Research Article

Evaluation and Analysis of the Impact of Airport Delays

Fen Zhou, Guosong Jiang, Zhengwu Lu, and Qingdong Wang

School of Computer Science and Technology, Huanggang Normal University, HuangGang Hubei 438000, China

Correspondence should be addressed to Qingdong Wang; wangqingdong@hgnu.edu.cn

Received 18 December 2021; Revised 12 January 2022; Accepted 23 January 2022; Published 24 February 2022

Academic Editor: Baiyuan Ding

To characterize the real dynamic process of delay propagation between airports and to understand the mechanism of delay propagation from a global perspective, this paper establishes a network of delay propagation relationships between airports based on causal analysis and conducts an example analysis. First, the flight operating status data processing is used to obtain the arrival delay time series of each airport. Then, the delay propagation relationship between airports is analyzed in pairs through the causal analysis method to obtain the delay propagation relationship network between all airports. Finally, the complex network theory and related indicators are used. The network is analyzed, and these analysis results can provide theoretical support for the formulation of delay propagation mitigation measures.

1. Introduction

With the rapid development of today's air transport industry [1], the number of flights operated by airlines has increased significantly. The problem of declining punctuality has followed, and flight delays have gradually emerged. As a result, the air transport industry will suffer substantial economic losses every year [2] and even endanger the safety of passengers, aircraft, and airports [3]. Therefore, flight delays have become a significant challenge facing the air transportation system [4]. The upstream and downstream flights share aviation resources, such as aircraft, crew, and airport ground resources. Therefore, when the upstream flight is delayed, the downstream flight will also be delayed with significant probability. This phenomenon is called delay propagation [5, 6]. If there is no reasonable method to control the propagation of delays, the spread of delays will continue to expand [7, 8]. At the same time, due to the current development of the air transportation industry, the aviation operation scene has become very complicated, and flights will be delayed due to varying degrees of influence from different sources at the same time. Therefore, the research on the propagation mechanism of flight delays is essential and challenging.

So far, there has been a lot of research in flight delay propagation. For example, Beatty et al. [9] put forward the concept of delay multiplier to quantify delay propagation by analyzing an airline's flight status table. The delay propagation analysis of multiple flights at a single airport is carried out. These methods mainly study the local dynamics of some flights, a single airport, or an airline [10–14]. Only a small number of people research the air transportation system level from the perspective of the entire aviation network. Therefore, in the field of flight delay propagation, there is still a lack of away from the perspective of the whole aviation network that can accurately dig out the causal relationship of flight delays from the current highly complex data: the delay propagation relationship between airports. Obtaining the flight delay propagation relationship is helpful to understand the delay propagation mode and law between airports, provide theoretical support for formulating control delay propagation measures, and have guiding significance for the improvement of flight delays.

2. Related Work

The academic literature on flight delays can be divided into three categories:

(1) Statistical models that explore the impact of various components of travel time

(2) Econometric models that analyze the economic drivers of flight delays

(3) Operational management models

Among them, the impact of air transportation delays on operations is investigated. Due to the highly random nature of air transportation, different aspects of flight scheduling problems have been explored in the past. Some researchers have developed statistical models for predicting the other components of air travel time. In the econometric model, the influence of various factors on the onset and progress of propagation delay is quantified. Either way, it is necessary to analyze the delay propagation process from a broader and network-based perspective because the flight schedules operated by airlines and airports are oriented towards network performance optimization. Although progress has been made in understanding the spread of flight delays [15–20], few studies have investigated the space of delays by considering the interdependence of delays. With the continuous development of time-series analysis, much progress has been made in various fields. The time-series dependency analysis [21] and medical applications of complex human body systems (such as brain-computer interface [22–26], epilepsy [27, 28], sleep staging [29–34]). Since the air transportation system is also typically large and complex, the mechanism of delay propagation has not been fully understood, especially the interdependence between different airports. Therefore, understanding the spread of flight delays is a complex problem, and few studies have studied the space of delays by considering the interdependence of delay time series. Based on the dependence of delays and using network graph theory similar to [35–40], we evaluate and analyze the airport flight delay network.

3. Network Analysis Evaluation Index

We use a flight delay propagation relationship discovery algorithm based on transfer entropy. This algorithm is used to discover the delay propagation relationship between two airports. Due to the large number of airports and complex interactions, the characteristics of delayed propagation cannot be understood from only the information at the level of a single airport. The complex network theory and related indicators provide a suitable method for studying air transportation systems. Therefore, network-level analysis is used to capture the global structure of functional interaction.

$A_{ij=}(a_{ij})_{N*N}$ is the adjacency matrix of the sample network. $a_{ij} = 1$ if and only if there is an edge from one node i to another node j; otherwise, $a_{ij} = 0$. In this sample network, the total number of airports is N, and the total number of edges is $M = \sum_{i,j=1}^{N} a_{ij} = 12$. Later, some practical topological structures were introduced to help analyze the delay propagation.

3.1. Degree. The degree of an airport reflects the number of airports with which it has a delayed propagation relationship. In the directed network, the in-degree $k_i^{in} = \sum_{j=1}^{N} a_{ji}$ and the out-degree $k_i^{\text{out}} = \sum_{j=1}^{N} a_{ij}$ of airport i, respectively, represent the delay from airport j propagate to airport i and delays propagate from airport i to airport j. The total degree of airport i is $k_i = k_i^{\text{in}} + k_i^{\text{out}}$.

3.2. Reciprocity Parameter. The reciprocity parameter reflects the two-way nature of delay propagation between airport pairs. Reciprocity means that airport i affects airport j, and airport j also affects airport i ($a_{ij} = a_{ji} = 1$). The parameter R is used to evaluate the symmetry of the directed network. It is defined as follows:

$$R = \frac{\sum_{i\neq j}^{N}\left(a_{ij} - \overline{a}\right)\left(a_{ji} - \overline{a}\right)}{\sum_{i\neq j}^{N}\left(a_{ij} - \overline{a}\right)^2}. \tag{1}$$

Here, $\overline{a} = \sum_{i\neq j}^{N} a_{ij}/N(N-1)$. The maximum R is 1, which means that the links between all airport pairs are bidirectional. The larger the R value, the more symmetrical the network.

3.3. Link Density. Link density reflects the proportion of active links in the total number of potential links, that is, the proportion of airport pairs that influence all airport pairs. The calculation formula is as follows:

$$l_{\text{d}} = \frac{\sum_{i\neq j}^{N} a_{ij}}{N^2} = \frac{M}{N^2}. \tag{2}$$

The higher the l_{d}, the tighter the network connection, the easier it is for the delay to spread through it and the less likely it is to stop the delay from spreading.

3.4. Transitivity. Transitivity measures the existence of triangles in the network. Mathematically defined as the relationship between the number of triangles (3 nodes, each pair of nodes have edges) N_Δ and the number of connected triples N_3 in the network, the formula is as follows:

$$T = \frac{3N_{\Delta}}{N_3}. \tag{3}$$

Here, $N_{\Delta} = \sum_{k>i>j} a_{ij}a_{ik}a_{jk}$, $N_3 = \sum_{k>i>j}(a_{ij}a_{ik} + a_{ji}a_{jk} + a_{ki}a_{kj})$. The existence of a large number of triangles indicates that the groups of three airports are closely connected, so the delay generated in any one of them can easily spread to other airports.

3.5. Assortativity. The coordination coefficient reflects whether nodes with similar degree values tend to be connected. It can be expressed by the conditional probability, that is, the probability of an airport with degree k and an airport with the degree of influence. The formula is as follows:

$$\frac{1}{M}\sum_{j>i}\frac{1}{2}\left(k_i + k_j\right)a_{ij}. \tag{4}$$

A positive co-match coefficient value indicates that airports with large degrees, in general, tend to airports with high connectivity, and the network is said to be co-matched;

a negative co-match coefficient value indicates that airports with large degrees, in general, tend to airports with low connectivity, which is said that the network is heterogeneous.

3.6. Efficiency. The efficiency of the network reflects how easy the delay is to spread between the two airports, that is, how many intermediate nodes the delay has to pass from a node to reach the target node. The formula is as follows:

$$E = \frac{1}{n(n-1)} \sum_{i \neq j} \frac{1}{d_{ij}}. \quad (5)$$

Here, d_{ij} is the distance (or the number of hops) between nodes i and j.

3.7. Clustering Coefficient. The clustering coefficient reflects the inherent clustering trend of airports. The aggregation coefficient of an airport is the proportion of adjacent airports (airports with delayed propagation links with other airports) that have direct delay propagation links (that is, the number of triangles in the network). For the network, the overall clustering coefficient is calculated as follows:

$$C^{D} = \frac{1}{n} \sum_{i=1}^{n} \frac{(1/2)\sum_j \sum_h \left(a_{ij} + a_{ji}\right)\left(a_{ih} + a_{hi}\right)\left(a_{jh} + a_{hj}\right)}{\left[d_i^{\text{dot}}\left(d_i^{\text{dot}} - 1\right) - 2d_i^{\leftrightarrow}\right]}. \quad (6)$$

Here, $d_i^{\text{dot}} = \sum_{j \neq i} a_{ji} + \sum_{j \neq i} a_{ij}$ and $d_i^{\leftrightarrow} = \sum_{j \neq i} a_{ij} a_{ji}$.

3.8. Largest Connected Cluster. The largest connected cluster reflects the degree of delayed propagation. The largest related group is through a set of interconnected airports. To represent the area of delayed propagation, we set a baseline for the connected cluster. Only if an airport affects enough airports will it be added to the connected group; it is greater than a certain threshold.

To measure the similarity of the largest connected clusters of different networks, the Jaccard index is introduced, which is defined as $J = |A \cap B|/|A \cup B|$, where A and B are both finite sets composed of airport members in the largest connected cluster. If the airport sets are the same, then $J = 1$; if they are completely different, then $J = 0$.

3.9. Community. The community assesses whether airport delay propagation can be divided into multiple subregions. Each subregion has dense delay propagation links inside the airport and sparse delay propagation links with airports outside the subregion. In addition, the modular Q_d is used to measure the strength of dividing the network into multiple communities. The calculation formula of Q_d is as follows:

$$Q_d = \frac{1}{M} \sum_{ij} \left(A_{ij} - \frac{k_i^{\text{out}} k_j^{\text{in}}}{M} \right) \delta\left(c_i, c_j\right). \quad (7)$$

Here, if airport i and airport j are in the same community, the δ function outputs 1; otherwise, it outputs 0. M represents the total number of edges.

3.10. Network Motifs. The network motif reflects the local relationship pattern between any three airports. The three airport clusters have similar relationship patterns. An essential tool for evaluating the significance level of the motif is Z-score, which is defined as follows:

$$Z(G') = \frac{F_G(G') - \mu_R(G')}{\sigma_R(G')}, \quad (8)$$

where $F_G(G')$ represents the frequency of G' in the network G. $\mu_R(G')$ and $\sigma_R(G')$, respectively, represent the frequency average and standard deviation of N random networks.

3.11. Network Randomization. Network randomization is used to generate a random network for network comparison. During the randomization process, self-connection and duplication of edges are prohibited.

4. Data Description and Processing

4.1. Data Description. This dataset contains information on the operation status of all flights departing and arriving in China from December 1st to 31st 2000. The calculation method of the flight arrival delay in this article is the difference between the actual arrival time and the planned arrival time, representing the actual delay perceived by the passengers, rather than starting to calculate the arrival delay more than 15 minutes after the scheduled arrival time. We obtained the average daily arrival delay time for all flights in December 2000 across the country based on this calculation method. The results showed that the delay time on December 4, 8, 20, and 21 was relatively high, while on the 11th and 17th, the delays on the day and the 31st are relatively low as shown in Figure 1.

4.2. Data Preprocessing

4.2.1. Data Cleaning. Specific strong values need to be calculated based on the original data, such as flight arrival delay, departure delay, ground transit, and air flight. At the same time, there is a certain percentage of dirty data in the original dataset, which needs to be predicted. The detailed steps of preprocessing are as follows:

Step 1: delete 11 pieces of data whose planned arrival times are empty.

Step 2: 1649 entries.

Step 3: the original data only have the planned departure date and the actual departure date but not the planned arrival date and actual arrival date. To facilitate subsequent processing, the scheduled arrival date and actual arrival date need to be added to the data.

Step 4: for the convenience of subsequent processing, splice the two fields of date and time into one area.

Step 5: the original data are not the data of the flight segment level, so it is necessary to set whether each data are the tag of the flight segment, which is seg_count. It

is not a flight segment level. For example, there is an aircraft flying from A to B and from B to C, but there is another piece of data in the data that will pass from A to C. Therefore, the seg_count field corresponding to the data flying from A to C is set to 2. The seg_count flying from A to B and from B to C are both set to 1. After the final setting is completed, there are 54891 data with seg_count>1.

Step 6: delete the data of seg_count>1, these data are not of the flight segment level, so they are useless.

Step 7: the flight registration number is the same, the departure airport is the same, the planned departure time is the same, and the arrival airport may have the same or different data. Only one of these data need to be deleted, but which one do you choose to delete? or delete one randomly (currently select the one with the smallest id), about 180.

Step 8: calculate the planned flight time and actual flight time (flight time = arrival time-departure time) in minutes.

Step 9: calculate the planned passing time and the actual passing time. This month, this task needs to find the flight chain based on the aircraft registration number. The transit time = the departure time of the next flight in the flight chain-the departure time of the previous flight in the flight chain, in minutes. Although this calculation does not consider the midway rest of some aircraft, there may be some very long transit times, but this does not affect the subsequent operations, so there is no need to worry.

Step 10: calculate take-off delay and arrival delay. Departure delay = actual departure time-planned departure time, and arrival delay = actual arrival time-planned arrival time, in minutes.

Step 11: delete the data with the larger absolute value of the difference between the actual flight time and the planned flight time (currently take the data with the total value greater than 5 hours). From practical considerations, the flight time between the two places should not fluctuate too much, so the data that fluctuate more than 5 hours are deleted, 108 pieces. Of course, it is not necessarily correct if it is less than 5 hours, but there is no way.

Step 12: delete the data whose planned and actual transit time and flight time are less than zero, 3810 items.

Step 13: complete data preprocessing.

From the above 13 steps, all the fundamental values that need to be used can be obtained, the dirty data can be removed, and finally, 256758 flight operating status data can be obtained.

4.2.2. Construction of Airport Delay Data. The input for transferring entropy needs to be time-series data, so after data calculation and data cleaning, they still need further processing.

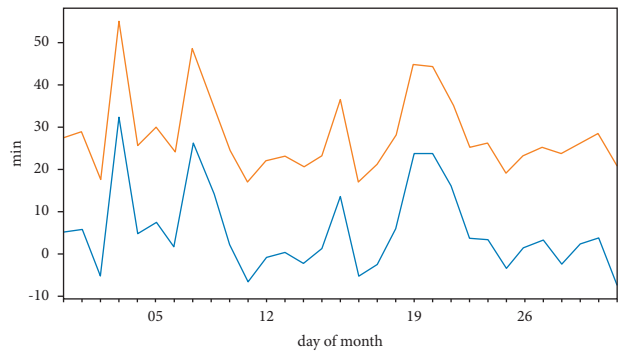


Figure 1: National average daily delay time.

Step 1. Aggregate the arrival delays of all flights at all airports at the hourly level to obtain the arrival delay time series of all airports.

Precisely, the average arrival delay time of all flights is calculated whose planned arrival time is within the h hour on day d at the i-th airport.

$$D^i(d,h) = \frac{\sum(T_{act} - T_{sch\ d})}{N^i(d,h)}, i \in \{1,2\}, \tag{9}$$

where T_{act} represents the actual arrival time of all flights whose scheduled arrival time is within the h hour on day d of the i-th airport and T_{schd} represents the actual arrival time of all flights whose scheduled arrival time is within the h hour on day d of the i-th airport. Planned arrival time, $N^i(d,h)$, represents the number of all flights whose planned arrival time is within the h hour on the d day of the i-th airport.

Then, the i-th airport arrival delay time-series set can be expressed as

$$S^i = \left\{D^i(1,1), D^i(1,2), \ldots, D^i(2,1), D^i(2,2), \ldots, D^i(d,h)\right\}. \tag{10}$$

Here, the length of the airport arrival delay time series is $n = d * h$.

Step 2. Use Z-score standardization to process the arrival delay time series, and the processing method is as follows:

$$D^{i\prime}(d,h) = \frac{D^i(d,h) - < D^i(\cdot,h) >}{\sigma\left(D^i(\cdot,h)\right)}, i \in \{1,2\}. \tag{11}$$

where $D^{i\prime}(d,h)$ represents the normalized average arrival delay time of all planned arrival times at the hth hour at the dth day of the i-th airport, $D^i(d,h)$ represents the original average arrival delay time of all planned arrival times at the hth hour of the i-th airport on the d day, $< D^i(\cdot,h) >$ represents the average value of all sample points at the h-th hour of the i-th airport, and $\sigma(D^i(\cdot,H))$ represents the standard deviation of all sample points at the h-th hour of the i-th airport.

Then, the preprocessed set of the i-th airport arrival delay time series can be expressed as

$$S^{i\prime} = \left\{ D^{i\prime}(1,1), D^{i\prime}(1,2), \ldots, D^{i\prime}(2,1), D^{i\prime}(2,2), \ldots, D^{i\prime}(d,h) \right\}. \tag{12}$$

Here, the length of the preprocessed airport arrival delay time series is $n = d * h$.

From the above two steps, the final required transfer entropy input time-series data can be obtained. The processing of the original data is completed.

5. Experiment Analysis

5.1. Global Characteristics of the Airport Delay Propagation Relationship Network. To analyze the delay propagation relationship, we need to transfer entropy between the paired computer field delay time series and build a network of delay propagation. We aggregate flight data hourly so that each airport corresponds to a time series every day. Then, we build a network of delay propagation relationships every day, and we get a total of 31 networks. December 4th is the day with the most serious flight delays. We use the communication network on this day to analyze the characteristics of delayed propagation. After removing the unconnected airports, it is found that there are 169 airport nodes and 1381 edges, as shown in Figure 2, which means that about three-quarters of the airports have a delayed propagation relationship with other airports that day. This delayed propagation relationship network includes 17 4F-class airports, 35 4E-class airports, 38 4D-class airports, 71 4C airports, and eight 3C airports. It can be seen that larger airports such as 4F, 4E, and 4D are almost all caught in the spread of delays. Although small airports such as 4C and 3C occupy a large proportion of the entire civil aviation network, only half of them are involved in the spread of flight delays.

For this communication network, we will use network analysis tools to answer the following questions about delayed communication:

How many airports are affected or affected by each airport? For a certain airport i in the network, k_i^{out} is the number of airports affected by airport i, and k_i^{in} is the number of airports affected by airport i. In this network, $<k^{in}> = <k^{out}> = 8.17$, which means that each airport affects about 8 airports and is affected by about 8 airports.

Is the delayed propagation connection between airport pairs bidirectional? The reciprocity parameter of this network is $R = 0.77$. Randomly, 1000 networks with the same number of nodes and edges are generated, and their average reciprocity coefficient is found to be 0.049, which is much lower than 0.77. Therefore, the network is more symmetrical. One possible reason is that delays caused by two-way flights between airports are prone to two-way propagation.

How big is the aggregation trend between airports? The average aggregation coefficient of the entire network is 0.42, which is much larger than the average aggregation coefficient of the random network (=0.048), indicating that the airport network has a strong aggregation tendency.

Can the delay propagation between airports be divided into multiple subareas? The community detection algorithm is used to analyze it. Modularity is used to evaluate the strength of the network divided into communities. The greater the modularity value, the more obvious the community structure. The modularity value of this network is 0.168, and the average modulus value of 1000 random networks is 0.205. Therefore, there is no evidence that the delayed propagation on December 4 can clearly delineate the subregions.

What is the possible range of delay propagation? We use the largest connected cluster to represent the range of possible delay propagation. The member airports of the largest connected cluster are selected by the k^{out} threshold so that the largest connected cluster contains a set of airports that have more influence on other airports. We define the threshold k^{out} to be greater than the average out-degree of the network $k^{out} = 8.17$, which means that the impact of the airport is greater than or equal to the other five airports. The largest connected cluster of the network contains the number of airports $M_{\text{d}} = 50$, indicating that the delay may spread among 50 airports.

5.2. Time Characteristics of the Delay Propagation Relationship Network. As time changes, the daily delay situation is different; how will the delay propagation network change? To answer this question, we evaluated the Pearson correlation coefficient between the daily network topology indicators of the delay propagation relationship network and the average daily arrival delays of all flights in December. It can be seen from Table 1 that the number of edges, average clustering coefficient, average degree, link density, and comatch coefficient are all highly correlated with the average daily arrival delay. At the same time, the correlations of other indicators are relatively low. According to the formula of average degree and link density, it can be said that these two indicators are directly proportional to the number of sides, so the number of sides, average degree, and link density can be analyzed together. This result means that the longer the national average daily delay to Hong Kong, the more edges in the delay propagation network, and vice versa.

The higher the clustering coefficient of the node, it means that the neighboring nodes of the node are more likely to be connected. It is easier to form a local area with the node as the starting point so that the delay will spread in it; on the contrary, if the clustering coefficient of the node is lower, the node's neighbor nodes are more inclined to connect to other nodes except their neighbors. The formed network is closer to a tree so that the delay will spread quickly in a larger area. Therefore, due to the large degree and the low aggregation coefficient, the large airport will produce a more extensive range of delay propagation. In contrast, in the small airport, the delay generated can only spread in a local area due to the low degree and the high aggregation coefficient.

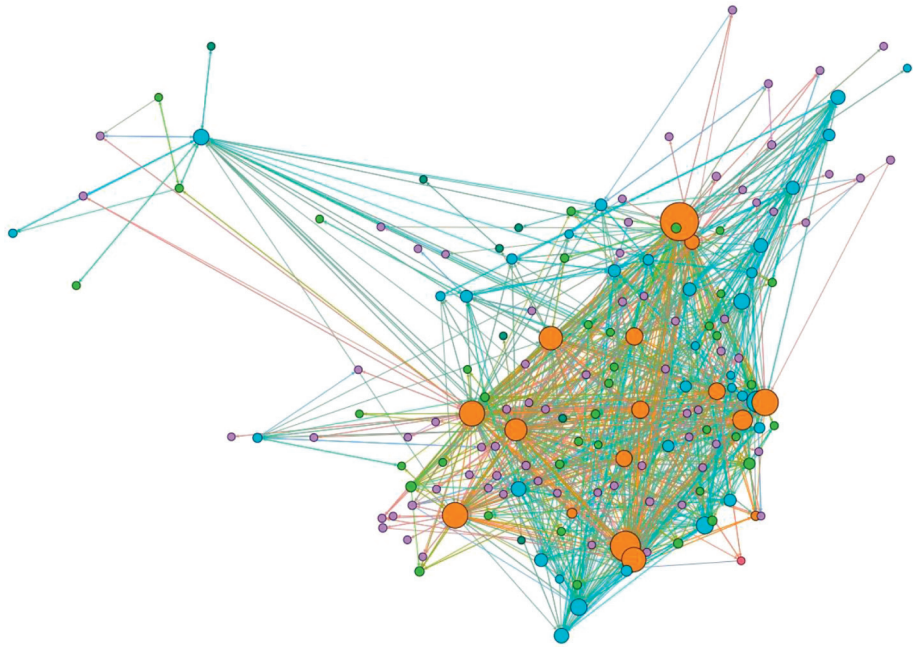


Figure 2: Delayed dissemination network on December 4.

Table 1: Correlation analysis between network topology indicators and average daily arrival delays.

Network topology indicators	Pearson correlation coefficient
Number of nodes	0.396
Number of sides	0.634
Reciprocity coefficient	0.550
Average clustering coefficient	0.652
Average degree	0.627
Link density	0.704
Transitivity	0.522
Coefficient of coordination	−0.632
Largest connected cluster	0.327
Modularity	−0.448
Efficient	0.457

The co-match coefficient has an inverse relationship with the average daily arrival delay time, which means that the more serious the delay, the more nodes with higher degrees tend to nodes with lower degrees of connectivity, because the degree is proportional to the number of flights, and the number of flights is larger. The daily operation of the airport has approached saturation, so the more serious the delay, the greater the tendency of large airports to spread the delay to various small airports and try to avoid affecting other larger airports and reduce losses.

6. Discussion and Conclusion

From the perspective of the entire aviation network, this study established a delay propagation relationship network based on the time-series relationship of each airport arrival delay. It used network analysis and evaluation indicators to reveal the macro performance of delay propagation. Specifically, we have constructed 31 flight delay propagation relationship networks using the information on the operation status of all flights departing and arriving in China. Selecting the day with the most serious delays in the 31 days as the analysis object, we found that the larger airports of 4F, 4E, and 4D are almost all caught in the spread of delays, although small airports such as 4C and 3C occupy a large proportion of the entire civil aviation network. But only half of them are involved in the spread of flight delays. The average degree shows that, on average, each airport affects about eight airports and is also affected by about eight airports. The very high reciprocity parameter indicates that the two-way delay spread between the airport pairs is more serious. It may be because a plane performs round-trip flights between the two airports a day, but there is no suitable buffer absorption mechanism between the airports to cause the delay. The high clustering coefficient of the airport indicates that the network has a strong trend of clustering, and the airport nodes tend to form a denser network cluster. We have also found the largest connected cluster in the network, where each airport node has more influence. The ingress and egress of airport nodes are directly proportional to the number of flights at the airport. The intensive flight of large airports has made the airport capacity close to saturation. Once an accident occurs, the delay will happen and spread, but because large airports generally have better-delayed handling measures may reduce some of the delays.

Next, we analyzed the time characteristics of the delay propagation relationship network, compared 31 indicators of the delay propagation relationship network, and calculated the Pearson correlation coefficient between each indicator and the average daily arrival delay of all flights in December. The results obtained show that the aggregation coefficient strongly correlates positively with the average daily arrival delay. This explains that large airports have a

larger degree and a low aggregation coefficient, which will cause a larger range of delay propagation, while small airports have a higher degree of delay. Small and high aggregation coefficients and the resulting delay can only spread in a local area. The co-match coefficient has an inverse relationship with the average daily arrival delay time. This means that the more serious the delay, the more nodes with higher degrees tend to nodes with lower degrees of connectivity. Therefore, the more serious the delay, the greater the tendency for large airports to spread the delay. Delays can be dispersed to various small airports to avoid affecting other larger airports and reduce losses. The proposed method is a general causal analysis method. In the future, we will use this method on other types of time-series data, such as physiological time series and financial time series.

Data Availability

The dataset can be accessed upon request to the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by the University industry university research innovation fund of science and technology development center of the Ministry of education of China (2020ITA05022), the Natural Science Foundation of Hubei Province (2021CFB316), the Preliminary Support Project of Hubei Social Science Foundation (21ZD137), and the Hundreds of Schools Unite with Hundreds of Counties-University Serving Rural Revitalization Science and Technology Support Action Plan (BXLBX0847).

References

[1] John Wiley & Sons, *The Global Airline Industry*, John Wiley & Sons, Hoboken, NJ, USA, 2015.

[2] M. Ball, C. Barnhart, M. Dresner et al., *Total Delay Impact study, NEXTOR Research Symposium*, Washington DC, USA, 2010, http://www.nextor.org.2010.

[3] S. AhmadBeygi, A. Cohn, Y. Guan, and P. Belobaba, "Analysis of the potential for delay propagation in passenger airline networks," *Journal of Air Transport Management*, vol. 14, no. 5, pp. 221–236, 2008.

[4] A. I. Czerny, "Airport congestion management under uncertainty," *Transportation Research Part B: Methodological*, vol. 44, no. 3, pp. 371–380, 2010.

[5] N. Kafle and B. Zou, "Modeling flight delay propagation: a new analytical-econometric approach," *Transportation Research Part B: Methodological*, vol. 93, pp. 520–542, 2016.

[6] S. Lan, J.-P. Clarke, and C. Barnhart, "Planning for robust airline operations: optimizing aircraft routings and flight departure times to minimize passenger disruptions," *Transportation Science*, vol. 40, no. 1, pp. 15–28, 2006.

[7] L. Daqing, J. Yinan, K. Rui, and S. Havlin, "Spatial correlation analysis of cascading failures: congestions and Blackouts," *Scientific Reports*, vol. 4, no. 1, p. 5381, 2014.

[8] L. Meng and X. Zhou, "Robust single-track train dispatching model under a dynamic and stochastic environment: a scenario-based rolling horizon solution approach," *Transportation Research Part B: Methodological*, vol. 45, no. 7, pp. 1080–1102, 2011.

[9] R. Beatty, R. Hsu, L. Berry, and J. Rome, "Preliminary evaluation of flight delay propagation through an airline schedule," *Air Traffic Control Quarterly*, vol. 7, no. 4, pp. 259–270, 1999.

[10] L. W. Fan, F. Wu, and P. Zhou, "Efficiency measurement of Chinese airports with flight delays by directional distance function," *Journal of Air Transport Management*, vol. 34, pp. 140–145, 2014.

[11] W.-B. Du, M.-Y. Zhang, Y. Zhang, X.-B. Cao, and J. Zhang, "Delay causality network in air transport systems," *Transportation Research Part E: Logistics and Transportation Review*, vol. 118, pp. 466–476, 2018.

[12] M. Maziarz, "A review of the Granger-causality fallacy," *The Journal of Philosophical Economics: Reflections on economic and social issues*, vol. 8, no. 2, pp. 86–105, 2015.

[13] G. Ver Steeg and A. Galstyan, "Information transfer in social media," in *Proceedings of the 21st international conference on World Wide Web*, pp. 509–518, New York, United States, 16 April 2012.

[14] M. Lungarella, K. Ishiguro, Y. Kuniyoshi, and N. Otsu, "Methods for quantifying the causal structure of bivariate time series," *International journal of bifurcation and chaos*, vol. 17, no. 3, pp. 903–921, 2007.

[15] N. Pyrgiotis, K. M. Malone, and A. Odoni, "Modelling delay propagation within an airport network," *Transportation Research Part C: Emerging Technologies*, vol. 27, pp. 60–75.

[16] N. Nayak and Y. Zhang, "Estimation and comparison of impact of single airport delay on national airspace system with multivariate simultaneous models," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2206, no. 1, pp. 52–60, 2011.

[17] L. Hao, M. Hansen, Y. Zhang, and J. Post, "New York, New York: two ways of estimating the delay impact of New York airports," *Transportation Research Part E: Logistics and Transportation Review*, vol. 70, pp. 245–260, 2014.

[18] Z. Jia, Y. Lin, Y. Liu, Z Jiao, and J Wang, "Refined nonuniform embedding for coupling detection in multivariate time series," *Physical Review*, vol. 101, no. 6, Article ID 062113, 2020.

[19] P. Fleurquin, J. J. Ramasco, and V. M. Eguiluz, "Systemic delay propagation in the US airport network," *Scientific Reports*, vol. 3, no. 1, p. 1159, 2013.

[20] B. Campanelli, P. Fleurquin, A. Arranz et al., "Comparing the modeling of delay propagation in the US and European air traffic networks," *Journal of Air Transport Management*, vol. 56, pp. 12–18, 2016.

[21] Z. Jia, Y. Lin, Z. Jiao, Y. Ma, and J. Wang, "Detecting causality in multivariate time series via non-uniform embedding," *Entropy*, vol. 21, no. 12, p. 1233, 2019.

[22] Z. Jia, Y. Lin, J. Wang, Z. Feng, X. Xie, and C. Chen, "HetEmotionNet: two-stream heterogeneous graph recurrent neural network for multi-modal emotion recognition," 2021, http://arXiv.org/abs/2108.03354.

[23] Z. Li, J. Wang, Z. Jia, and Y. Lin, "Learning space-time-frequency representation with two-stream attention based 3D network for motor imagery classification," in *Proceedings of the 2020 IEEE International Conference on Data Mining (ICDM)*, pp. 1124–1129, IEEE, Sorrento, Italy, 17-20 Nov. 2020.

[24] Z. Jia, Y. Lin, T. Liu, K. Yang, X. Zhang, and W. Jing, "Motor imagery classification based on multiscale feature extraction and squeeze-excitation model," *Journal of Computer Research and Development*, vol. 57, no. 12, p. 2481.

[25] Z. Jia, Y. Lin, X. Cai, H. Chen, H. Gou, and J. Wang, "Sst-emotionnet: spatial-spectral-temporal based attention 3d dense network for eeg emotion recognition," in *Proceedings of the 28th ACM International Conference on Multimedia*, pp. 2909–2917, WA, Seattle, USA, 12 October 2020.

[26] Z. Jia, Y. Lin, J. Wang, K. Yang, T. Liu, and X. Zhang, "A multi-branch multi-scale convolutional neural network for motor imagery classification," in *Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 736–751, Springer, Ghent, Belgium, 14-18 September.

[27] Y. Liu, Y. Lin, Z. Jia, J. Wang, and Y. Ma, "A new dissimilarity measure based on ordinal pattern for analyzing physiological signals," *Physica A: Statistical Mechanics and Its Applications*, vol. 574, Article ID 125997, 2021.

[28] Y. Liu, Y. Lin, Z. Jia, Y. Ma, and J. Wang, "Representation based on ordinal patterns for seizure detection in EEG signals," *Computers in Biology and Medicine*, vol. 126, Article ID 104033, 2020.

[29] Z. Jia, Y. Lin, J. Wang, R. Zhou, and X. Ning, "GraphSleepNet: adaptive spatial-temporal graph convolutional networks for sleep stage classification," *IJCAI*, pp. 1324–1330, 2020.

[30] Z. Jia, Y. Lin, J. Wang, X. Wang, P. Xie, and Y. Zhang, "SalientSleepNet: multimodal salient wave detection network for sleep staging," arXiv preprint: http://arXiv.org/abs/2105.13864, 2021.

[31] X. Cai, Z. Jia, M. Tang, and G. Zheng, "Brainsleepnet: learning multivariate eeg representation for automatic sleep staging," in *Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine*, pp. 976–979, Seoul, Korea, 16-19 Dec. 2020.

[32] Z. Jia, X. Cai, G. Zheng, J. Wang, and Y. Lin, "SleepPrintNet: a multivariate multimodal neural network based on physiological time-series for automatic sleep staging," *IEEE Transactions on Artificial Intelligence*, vol. 1, no. 3, pp. 248–257, 2020.

[33] Z. Jia, Y. Lin, H. Zhang, and J. Wang, "Sleep stage classification model based ondeep convolutional neural network," *Journal of Zhejiang University (Science Edition)*, vol. 54, no. 10, pp. 1899–1905, 2020.

[34] Z. Jia, Y. Lin, J. Wang et al., "Multi-view spatial-temporal graph convolutional networks with domain generalization for sleep stage classification," *IEEE Transactions on. Neural Systems and Rehabilitation Engineering*, vol. 29, pp. 970–979, 2021.

[35] W. Li, Q. A. Wang, L. Nivanen, and A. Le Méhauté, "How to fit the degree distribution of the air network?" *Physica A: Statistical Mechanics and Its Applications*, vol. 368, no. 1, pp. 262–272, 2006.

[36] G. Bagler, "Analysis of the airport network of India as a complex weighted network," *Physica A: Statistical Mechanics and Its Applications*, vol. 387, no. 12, pp. 2972–2980, 2008.

[37] R. Guimera, S. Mossa, A. Turtschi, and L. A. N. Amaral, "The worldwide air transportation network: anomalous centrality, community structure, and cities' global roles," *Proceedings of the National Academy of Sciences*, vol. 102, no. 22, pp. 7794–7799, 2005.

[38] Z.-C. Li, W. H. K. Lam, S. C. Wong, and X. Fu, "Optimal route allocation in a liberalizing airline market," *Transportation Research Part B: Methodological*, vol. 44, no. 7, pp. 886–902, 2010.

[39] H. E. Silva, E. T. Verhoef, and V. A. C. Van den Berg, "Airline route structure competition and network policy," *Transportation Research Part B: Methodological*, vol. 67, pp. 320–343, 2014.

[40] D. Gillen, H. Hasheminia, and C. Jiang, "Strategic considerations behind the network-regional airline tie ups-a theoretical and empirical study," *Transportation Research Part B: Methodological*, vol. 72, pp. 93–111, 2015.

Hindawi
Scientific Programming
Volume 2022, Article ID 1002517, 10 pages
https://doi.org/10.1155/2022/1002517

Research Article

Analysis of Persuasive Design Mechanism Based on Unconscious Calculation

Hongtao Zheng and Shuo Li

School of Design, Shanghai Jiao Tong University, Shanghai 200240, China

Correspondence should be addressed to Hongtao Zheng; zhenghongtao@sjtu.edu.cn

Received 9 December 2021; Revised 22 December 2021; Accepted 17 January 2022; Published 22 February 2022

Academic Editor: Baiyuan Ding

In order to make users' behavior more standardized, a persuasive design mechanism analysis method based on unconscious calculation is proposed. Taking the concept of persuasion and the goal of persuading function as the theoretical basis, the user behavior data is obtained in the form of correlation calculation, and the ant colony algorithm is used to classify the behavior data. According to the results of data processing, analyze unconscious behavior and its characteristics, and obtain the persuasion mechanism through unconscious calculation, persuasion model, and persuasion model design. The experimental results show that the method in this paper has a higher acceptance of behavioral persuasion and higher satisfaction of the persuaded, indicating that the method has strong practical applicability.

1. Introduction

Guide the user's behavior to be more in line with the norms, improve the user's motivation and ability to complete the behavior, and enable them to develop good behavior habits so as to provide users with better services. Today, with the increasing popularity of Internet technology, persuasion technology can be used to change the behavior and attitude of users and to guide and persuade users' behavior [1–3]. At present, the application research of persuasion technology mainly involves the field of health, and its persuasion strategy is also proposed for the field of health. The research is more one-sided and only takes meeting the needs of users as the research goal, which cannot achieve behavior persuasion in the real sense. Therefore, we should induce and persuade users to develop or change their behavior through design intervention so as to achieve a purpose other than the design itself [4–6]. Based on this, how to intervene with the target users, persuade them, and improve the execution of the users is a topic that needs to be studied.

The application of persuasive technology in various fields needs to be realized through design. As a mode of thinking, persuasive design has been widely used in the cross-research of design and other disciplines, including education, medical care, health, sports, games, advertising, e-commerce, and other fields. Persuasion is to achieve psychological and behavioral guidance through nonmandatory means. There are various ways of persuasion, mainly including direct diarrhea persuasion, impact persuasion, and retrograde persuasion. Direct persuasion is mainly for the purpose of informing. There is no specific persuasion object and no targeted behavior to persuade. It is only a popular persuasion method to understand things from the perspective of cognition, which is not persuasive. Impulse persuasion is a special persuasion method. It has specific persuasion objects and targeted target behaviors. Finally, it is necessary to clearly change the deep-rooted views and opinions of the persuasion objects. Retrograde persuasion refers to stimulating the change of behavior or attitude from the opposite point of view. Although the above persuasion methods can achieve behavior persuasion to a certain extent, in practical application, the acceptance of behavior persuasion is not high, there is a certain gap between the persuasion effect and the expected effect, and the persuader's satisfaction with the method is not high.

In view of the problems existing in the above persuasion methods, this paper proposes an analysis method of

persuasion design mechanism based on unconscious computing. Unconsciousness refers to a kind of consciousness that is unnoticed and unconscious. Unconscious thinking refers to the thinking process below the level of consciousness. Unconscious thinking plays a positive role in problem solving. Therefore, in the design of persuasion design mechanism, unconscious thinking calculation is introduced to further improve the persuasion effect.

2. Persuasion Concept and Functional Objective Analysis

2.1. Persuasion Concept. Persuasion is a concept in psychology, which refers to allowing the persuaded to accept content that is purposeful under noncompulsory circumstances. Persuasive design refers to the use of persuasive psychological methods in the design to allow users to change their attitudes toward the product or their behavior in using the product. Its main purpose is to guide users to perform purposeful operations. Based on the theory of persuasive design, a persuasive design model is proposed. This model contains three elements, namely, motivation, ability, and motivation point.

2.1.1. Motivation. Motivation refers to the user's internal reasons when performing operations or using behaviors. It can be divided into three categories: fun and pain, hope and fear, and social identification and social rejection. Among them, fun and pain are derived from human instinct, hope and fear are a result of human behavior, and social identification and social rejection are feedback level content after the behavior is over.

2.1.2. Ability. Capability refers to the ability of a user to complete a certain behavior. In the persuasive design theoretical model, the most important principle is simplicity. For example, in product design, the higher the ease of use of the product, the lower the requirements for the user's ability to complete the use of the product and the higher the user's sense of pleasure. Therefore, designers should try their best to reduce the requirements of products to users' ability in interactive design and make products more useable and easy to use. This principle is also applicable in other research fields.

2.1.3. Promoting Point. The promotion point refers to the clue provided to the user or a metaphor so that the user can complete the persuasive operation behavior. The promotion point can be related to motivation and promotion of the motivation elements of users. The promotion point can also be related to the ability to complete a certain behavior under the existing abilities of the user. Finally, the promotion point can also be a reminder behavior point to remind users of some operation behaviors.

In the process of persuasive design, according to the three elements of the persuasive design model, the guidance of user behavior is realized through the control of user motivation, ability and promotion point. Only skillfully balancing the relationship between users' motivation, users' ability, and promotion point can design an effective persuasion mechanism.

2.2. Persuasive Functional Goals. Based on the description of the concept of persuasion, the persuasive design model, and its elements, the persuasion technology uses information as the external trigger element of behavior change to enhance the user's intrinsic motivation and behavior ability so as to achieve the purpose of persuading users to change their behavior. This process can be abstracted and become the basis for the realization of the persuasion mechanism. Based on this functional goal, through relevant technology to perceive, collect user behavior state data and environmental data included in the user persuasion target behavior context, and convert it into persuasion information that can be perceived by the user and generate behavioral ability and behavioral motivation. The purpose of persuading the target behavior can be achieved by timely selecting the appropriate carrier to convey information to user.

3. User Behavior Data Processing

According to the goal analysis of persuasion function, persuasion technology takes information as the external trigger element of behavior change. The information here specifically refers to user behavior data. Therefore, before the design of the persuasion mechanism, first process the user behavior data and obtain the user behavior data processing results through two steps of data mining [7, 8] and data classification [9, 10], and provide the necessary trigger elements for the design of persuasion mechanism.

3.1. Behavioral Data Mining. Assuming that the user behavior data is in a data area, the data in the entire area is described through the undirected traversal graph $H = (A, B, C)$, where A represents the node set, B represents the link set, and C represents the number of users in the area. Assuming that the data in the regional environment is composed of K regions, it is represented by a set form, specifically $K = (k_1, k_2, \ldots, k_n)$, where n represents the number of data regions, and then the data flow density of the data region [11] can be expressed by formula (1):

$$\rho = 1 - \frac{d_c}{\alpha_c (x - 1)^2 + x}, \tag{1}$$

where α_c represents the useful information in the data stream; d_c represents the community to which the data stream belongs; x represents the node where the data stream is located.

If W_i represents the amount of behavior data of a user i in the data area and G_w is the corresponding feature set, where w represents the user behavior feature, then the expression for the degree of association between user i and behavior feature w is

$$D_{iw} = \sum_{i=1}^{N} w_i \left\{ \log_{10} \left[\frac{w_i}{\beta_i} + D\left(\lambda_i \| \varphi_i\right) \right] \right\}, \quad (2)$$

where β_i represents the correlation factor, which is related to the amount of data in the data area; N represents the number of users; λ_i represents the salient features of user i; φ_i represents the insignificant features of user i [12, 13].

Since the persuasion mechanism is designed to persuade users' behaviors, persuasion is mainly oriented to the salient features of user behaviors, and the insignificant features that have a little impact can be ignored. Considering the above factors, if μ_i^l represents the payload length of the salient features of user behavior in the data area environment and D represents the total length, formula (2) is optimized to obtain an improved correlation calculation expression:

$$\mu_i^l = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} D_i\left(\mu_{ij}\right)}{\sum_{i=1}^{N} \sum_{l=1}^{N} D_l\left(\mu_{ij} + \varpi_{ij}\right)}, \quad (3)$$

where D_i represents the set of salient features of user i; μ_{ij} and ϖ_{ij} represent the payload length and total length of salient features of user behavior, respectively. Combining formula (2) and formula (3), we can obtain the data of salient features of user behavior, that is, to achieve user behavior data mining.

3.2. Classification of Behavioral Data. Based on the results of user behavior data mining, in order to avoid behavior deviations in the persuasion process and improve the persuasion effect, further classification of behavior data can not only reduce behavior deviations but also improve the efficiency of persuasion. The traditional method mainly uses the support vector machine (SVM) method to classify data. This method is mainly suitable for static data. There are certain limitations to the dynamic data of user behavior data [14–16]; therefore, this paper adopts the ant colony algorithm [17–19] to optimize it.

Ant colony algorithm is an intelligent bionic algorithm through which the traditional SVM method is optimized and applied to user behavior data classification to achieve the purpose of improving the accuracy of persuasion results [20, 21]. The specific operation process is given as follows.

Step 1. Initialize the position and pheromone of the ant colony.

Determine the initial pheromone size of ant z through the SVM parameter range; the calculation formula is

$$E_z = e\left(z_i \| z_j\right) + \left(z_i' \| z_j'\right), \quad (4)$$

where z_i represents the initial pheromone concentration; z_j represents the initial search speed; z_i' and z_j' both represent the direction guidance vector.

In order to prevent the ant colony from accelerating the convergence, a fitness function $X(t)$ [22] is set, and the fitness function $X(t)$ is modified; then

$$X(t) = \max_{q=1,2,\ldots,Q} \rho\left(S_q, S_p\right), \quad (5)$$

where S_q and S_p both represent genetic operators.

Step 2. Ant colony transfer.

Select the maximum pheromone concentration in the individual ant colony as the target individual, denoted by Y_{uf}.

$$Y_{uf} = \begin{cases} Y_{\text{best}}, & y \le y', \\ Y', & \text{otherwise}, \end{cases} \quad (6)$$

where Y_{best} represents the optimal solution obtained in the iterative process, that is, the maximum value of the pheromone concentration [23, 24].

Determine the moving direction of ant z 's position by formula (7):

$$R_z = \sum_{z=1}^{N} \left[p\left(\left|\theta_z\right|\right) \times Y_{uf} \right]^t, \quad (7)$$

where θ_z represents the expected moving direction of ant z.

To perform a local search for the ant in the dominant position in the data field, there are

$$H_z(A, B, C) = \sqrt{H_z(A)^2 + H_z(B)^2 + H_z(C)^2}, \quad (8)$$

where $H_z(A)^2$, $H_z(B)^2$, and $H_z(C)^2$ represent the relevant pheromone of nodes, links, and users in the data area.

Step 3. Pheromone update [25, 26].

After completing Step 1 and Step 2, update different pheromones. The specific update rules are as follows:

$$H_{abc}^z = \sum_{c=1}^{N} p_z\left(\left|\theta_{\omega_z \cap \alpha}\right|\right), \quad (9)$$

where α represents the volatilization coefficient of the pheromone.

Through the above steps, it can be seen that the use of the ant colony algorithm to optimize the data classification effect of the traditional support vector machine method can obtain more accurate data classification results [27, 28] and provide a user behavior data basis for the persuasion mechanism design.

4. Analysis Method of Persuasive Design Mechanism Based on Unconscious Calculation

Through the above analysis, the results of user behavior data processing are obtained, and the design and design of the persuasion mechanism will be analyzed in detail in the following. Analyze unconscious behavior and its characteristics, and give the application case of the persuasion mechanism. On this basis, give the persuasion design model, design the persuasion model, and complete the design of the persuasion mechanism.

4.1. Unconscious Calculation Analysis

4.1.1. Analysis of Unconscious Behavior and Its Characteristics. Unconscious behavior is an instinctive behavior made without subjective analysis and judgment, such as reflection and stress response, which can be designed persuasively by using the characteristics of unconscious behavior [29, 30]. For example, when passengers enter the security inspection device and leave the security inspection device, setting the two links to move continuously will make passengers unconsciously avoid light due to stress response, so as to achieve the effect of driving passengers, so that passengers can quickly pass the security inspection device. After the security check, leave quickly, thereby improving the efficiency of passengers' security check passage. Through the analysis of people's unconscious behavior, it can be found that unconscious behavior has the characteristics of universality, richness, and concealment. Almost everyone has unconscious behaviors in their behaviors. Unconscious behaviors are common in daily life and gradually integrated into their daily habits.

Unconscious behaviors are widely used in the design. Integrating unconscious behaviors into related designs can bring new design concepts to design and provide users with a natural user experience. A large number of unconscious behaviors run through people's increasingly frequent operation of mobile terminals [31, 32]. Introducing unconscious design into the persuasion mechanism design can enhance the user experience.

4.1.2. Unconscious Calculation. Based on unconscious behavior and its characteristics, unconscious computing has been studied. Unconscious computing has developed several times and has formed a variety of specific methods. Among these methods, the processing separation program (PDP) is still the best method. It separates conscious extraction and automatic extraction in simple recognition tasks by including tests and elimination tests. This paper intends to use the process separation program to separate the implicit and explicit components of unconscious behavior so as to realize unconscious computing [33, 34].

In order to explore the complex relationship between consciousness and unconsciousness, 5 (age: elderly, middle-aged, college students, junior high school, and junior high school) were adopted × 2 (contribution source: consciousness and unconsciousness). The contribution of consciousness and unconsciousness was calculated through the inclusion and exclusion test of the processing separation program (PDP). The specific process is as follows.

The subjects were divided into five age groups with 23 people in each group. All subjects were in good health and had a normal corrected vision. The age of the elderly group was 60–71 years. The age of the middle-aged group was between 30 and 55 years. The subjects in the university group were between 18 and 25 years old. The age of the subjects in the third group of junior middle school is between 14 and 15 years. The average age of the subjects in the high primary group was 11 years.

In the experiment, 50 specific pictures in the study of the best age of recognition ability were used as experimental materials. The whole experiment was divided into two stages: learning and testing. All subjects were tested separately. Ten of the 50 pictures were randomly selected as learning materials. In the learning phase, the learning time of each material is 2 seconds. In the test stage, each learned material and 4 unlearned materials are grouped into a group, and then the 10 groups of materials are subjected to inclusion test and exclusion test successively. In the inclusion test, the subjects were told to recognize the materials just presented in the five pictures, and if they cannot recognize the pictures just presented, pick out the first picture they think of. In the exclusion test, the subjects were told to recognize the material just presented in the five pictures, but not this one, but another possible picture.

According to the results of the inclusion test and exclusion test in the processing separation program (PDP), the conscious and unconscious contributions of each age group to recognizing specific pictures are calculated by using the formula (see Table 1).

The results show that (1) there is a developmental separation between conscious and unconscious contributions; (2) the unconscious contribution of the elderly group was higher than that of the conscious contribution but did not reach a significant level. The other four groups showed that the level of conscious contribution was extremely significantly higher than that of the unconscious contribution.

According to the above analysis, this paper uses the processing separation program to separate the implicit and explicit components of unconscious behavior and realizes unconscious computing.

4.2. Persuade Design Patterns.

Based on the unconscious computing theory, this paper determines five persuasion behavior components, constructs the persuasion behavior process mechanism, grasps the persuasion behavior psychological mechanism, analyzes the change mechanism of the five components and their overall impact on users' attitude and behavior, and takes it as a breakthrough in persuasion design. The structure of the persuasion behavior component is shown in Figure 1.

(1) Clue reminder: in order to increase the degree of user substitution, design familiar, clear, and attractive scene themes and report on user behavior in real time

(2) Behavior plan: in order to improve the user's execution ability, introduce environmental variables to stimulate behavior, design a reasonable plan, and provide a heuristic path [35, 36]

(3) Execution plan: in order to maintain the persistence of behavior, create a clear task process and give positive encouragement and guidance

(4) Social relevance: in order to enhance social recognition and improve behavior motivation, social sharing, self-expression, and peer comparison are carried out

Table 1: Conscious and unconscious contributions of each age group.

	High primary	Third junior	University	Middle age	Elderly
Consciousness contribution	0.569	0.625	0.584	0.387	0.273
Unconscious contribution	0.202	0.190	0.241	0.253	0.305

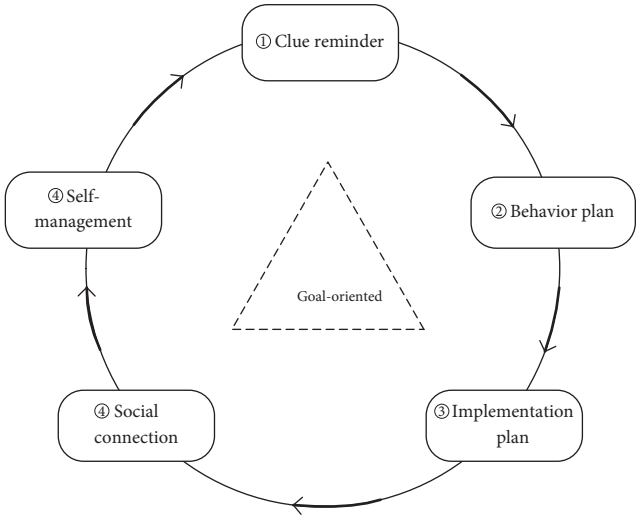


Figure 1: Persuasion behavior component structure.

(5) Self-management: in order to conduct self-management scientifically and rationally, the user behavior execution process is managed based on monitoring, comparison and evaluation

In actual persuasion, the brain will take intuitive response, active psychological tendency, and mental cooperation evaluation to make decisions according to similar or unfamiliar scenes. The behavior threshold and trigger conditions of persuasion are different. Therefore, according to the size of motivation and ability, this paper combines them into four quadrants and divides persuasion behavior into three types of modes. These three types of persuasion modes all have the above five persuasion design components, but each has its own emphasis. Class A represents repetitive habitual behavior with high motivation and ability, which refers to the formed behavior. It focuses on the cues and stimuli of creating a behavior environment, including explicit things or scene cues (interface cues), implicit habits, or experience connections (cognitive cues). Class B represents the assisted autonomous behavior with a low factor in motivation and ability, which refers to changing and transforming behavior habits. It focuses on activating behavior parameters and reducing obstacles to user behavior with the help of visible and measurable goals and feasible path plans. Class C represents the heuristic induced behavior with low motivation and ability. It refers to creating or using the situation shared by some persuasion behavior with the help of narrative, metaphor, empathy, and other rhetorical communication means such as cognitive therapy, design behavior route in line with the user's cognition, and gradually inducing the user.

4.3. Persuade Model Design. The persuasion context in the persuasion model includes intention (initiator of the intention to change behavior and attitude), event (clear the use context of persuasion technology, user context, and technical context), and strategy (information content, form, and dissemination path are accurately targeted to target users to achieve persuasion) aspects. Figure 2 is a schematic diagram of the persuasion model.

4.3.1. The Internal Function Stage of the Persuasion Model. Combining with the behavioral data processing and unconscious calculations mentioned above, the functional phase process inside the persuasion model is obtained. From data information input to persuasion information output, the entire functional stage is divided into three functional modules that affect the persuasion function: data information acquisition, data information transformation, and persuasion information transmission. The data information acquisition module is used to perceive the original data information and pass this information into the persuasion product, and then the data information conversion module will process the original data information into persuasive information that can be used to improve the user's behavior motivation and behavior ability. Finally, the persuasion information transmission module transmits the converted persuasion information to the user. Figure 3 shows the internal functional phase process of the whole persuasion model.

Through the establishment of the functional modules of the above persuasion model, it can be seen that the persuasion model is divided into three functional modules: data information acquisition, data information transformation, and persuasion information transmission.

4.3.2. Persuasion Mechanism Design. The application of persuasion mechanism in various fields needs to be realized through design. According to the above analysis, the specific process of persuasion mechanism design is as follows:

(1) Determine the target behavior. First, choose a simple and specific target behavior, and develop it into a large series of target behaviors step by step in a step-by-step manner.

(2) Determine the user object. Must first choose a willing target user group, and then step by step to expand those who are not willing or have the willingness to oppose the boycott.

(3) Analyze the reasons why users do not adopt the target behavior. Find out whether the reason for preventing user behavior is the lack of motivation or

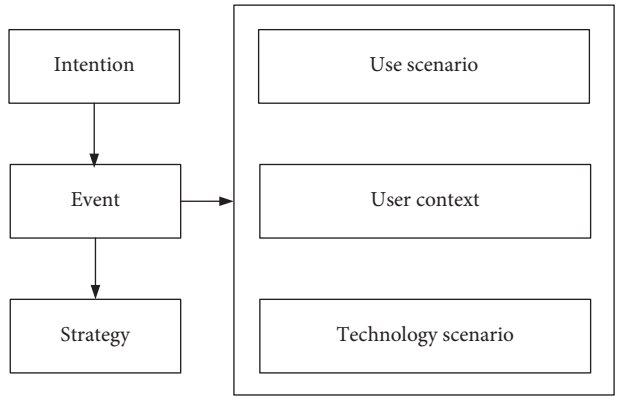


Figure 2: Schematic diagram of persuasion model.

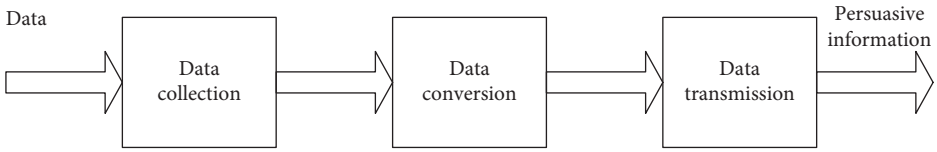


Figure 3: The internal function stage process of the persuasion model.

lack of ability. If both lack of motivation and lack of ability, they need to go back to the previous two steps to consider whether the previously determined target behavior and target user are appropriate.

(4) Choose a persuasion mechanism or persuasion strategy that meets the application conditions. The choice of persuasion mechanism is mainly considered from three aspects: the target behavior, the target user, and the reasons for preventing the behavior from occurring.

(5) Investigate the application cases of the persuasion mechanism. To determine whether the persuasion mechanism is appropriate, we need to find three cases with similar target behaviors, three cases with similar target users, and three cases with the same persuasion mechanism.

(6) Follow successful cases. According to previous studies on the application cases of the persuasion mechanism, we can find more successful cases from similar cases to imitate. This imitation is not just a hard copy, but to discover the essence of its persuasive effect from successful cases.

(7) Rapid iteration of prototype design. In persuasive design, the persuasion mechanism must be iterated quickly. With the development of related technologies, the persuasion mechanism must also be continuously developed. This is more important than deep thinking about how to persuade with the best mechanism.

(8) Expand the trial scope to verify the effectiveness of the function. As mentioned earlier, a small and easy-to-achieve goal should be selected to determine the goal. With the success of the small goal, the scale needs to be expanded. After the effectiveness of the persuasion mechanism is verified, a new target behavior can be established to achieve iterative and gradual development.

5. Application Case Analysis

In order to verify the effectiveness of the method in this paper, the following takes the field of shared electric vehicles as an example to further verify the feasibility of the method in this paper.

5.1. Scenario Design and Research Purpose. This paper attempts to introduce a persuasion mechanism to solve the problem of users' bad behavior. Therefore, after a systematic understanding of the persuasion mechanism, it analyzes the efforts of persuasion technology in the field of shared electric vehicles. Refining the application methods and modes of persuasive technology in the shared electric vehicle service system provides theoretical hypotheses and clearer research goals for further empirical research.

In recent years, the sharing economy of rent for sale mode has become more and more mature. The shared electric vehicle combined with the two has become a hot spot in the sharing field. Shared electric vehicles not only fill the gap in the field of public transportation but also make daily transportation more convenient and efficient and avoid environmental problems caused by exhaust emissions. In the long run, it provides a good scheme for the improvement of the energy crisis, air pollution, traffic congestion, and other problems. Even though electric shared vehicles have many advantages, they still encounter many obstacles in the early development. The main problem is that the vehicle body is damaged, the environment inside the vehicle becomes worse, the vehicle is parked disorderly, and even serious traffic accidents will

be caused by users' nonstandard operation. In order to make the healthy and sustainable development of shared electric vehicles, it is necessary to guide users' behavior to be more standardized, improve users' motivation and ability to complete behavior, and make them form good behavior habits so as to make shared electric vehicles better serve users. Therefore, in the case analysis, this paper uses persuasion technology to change users' behavior and attitude so as to guide and persuade users' behavior habits.

Through the investigation of users and their behaviors using shared electric vehicles, they can intuitively discover the users' bad behavior problems in the process of using shared electric vehicles and analyze the influencing factors of users' bad behaviors and attitudes. Accurately analyze the user's travel behavior path and key scenarios so as to obtain an opportunity for behavior persuasion. At the same time, on the basis of the original persuasive design, investigate whether the user's behavior is affected or changed and the user's recognition and trust in such behavior persuasion.

5.2. Basic Information of Survey Objects

5.2.1. Target User Attributes. Users of shared electric vehicles are the main research objects. It is necessary to study their age, gender, occupation, driving age, purpose, and other pieces of information, analyze the attributes of users in the field of shared electric vehicles, understand the background factors of users, and more accurately find their behavior path so as to guide the establishment of key scenes. Table 2 shows the relevant contents applied in the collection of basic information of survey objects.

The first part is the basic situation of shared electric vehicle users, collecting information on the user's age, gender, occupation, driving age, and whether there is a private car. Quantitatively study the characteristics of target users in the shared electric vehicle field, and collect data for the output of user models.

The second part is a study on the travel situation of shared electric vehicle users, which mainly includes the selected shared travel brand, the reason and purpose of choosing shared travel, and the frequency and duration of shared travel. The purpose of this part is to supplement the user model on the one hand and to make the behavior path for the subsequent on-site observation of users and on the other hand to analyze some of the reasons for the generation of user behavior.

The third part is a study on the behavioral cognition of users of shared electric vehicles, which mainly includes users' understanding and views on shared electric vehicles, the status of shared travel in the minds of users, overall feelings and difficulties encountered during use, and uncivilized and irregular behaviors that users have learned about themselves. On the one hand, this part studies the psychological attitude of users so as to see the implementation of behavior persuasion in sharing electric vehicles. On the other hand, they can also find out which factors may affect the behavior of shared electric vehicle users and the importance of these factors so as to conduct in-depth research on these factors in the next stage.

5.2.2. Travel Situation of Users Using Shared Electric Vehicles. The above is mainly to study the attributes of target users and then need to study and analyze the use of shared electric vehicles for such users to travel, which can provide directions for behavioral persuasion design. The content of the travel situation study includes the average number of users using a car per month and the average length of time each time the car is used. See Figures 4 and 5 for details.

Based on the above survey results, the effectiveness of this method is verified from the two aspects of behavior persuasion acceptance and persuasion satisfaction. In the verification, in order to highlight the advantages of this method, direct persuasion and impact persuasion are used as comparative methods to analyze the application effects of different persuasion methods.

5.3. Analysis of Experimental Results

5.3.1. Satisfaction of the Persuaded. Taking the satisfaction of the persuaded as the experimental index, the direct persuasion method, the impact persuasion method, and the method in this paper are compared. In the comparison process, 10 users were randomly selected, and the comparison was achieved through the form of scoring. The score interval is [0, 100]. The larger the value, the higher the satisfaction level. The results are shown in Table 3.

By analyzing the data in Table 3, it can be seen that, in the scores given by 10 users, the satisfaction of this method is higher than that of the direct persuasion method and impact persuasion method. Among them, user 4 gave the highest score of 97.1 for the method of this paper, while the highest scores for the direct persuasion method and impact persuasion method were 81.7 and 85.1, respectively, which were significantly lower than this method. It shows that users can better accept the persuasion mechanism designed by this method, and satisfaction can reflect users' executive power. The higher the satisfaction, the stronger the users' executive power and the better the persuasion effect. Therefore, it can be seen that the application effect of this method is better.

5.3.2. Acceptance of Behavioral Persuasion. Taking the acceptance of behavioral persuasion as the experimental index, the direct persuasion method, the impact persuasion method, and the method in this paper are compared, and the results are shown in Figure 6. Among them, the acceptance of behavior persuasion is represented by data, specifically 0.1–1.0. The higher the value, the higher the acceptance.

Analysis of Figure 6 shows that, with the increase in the number of iterations, the acceptance of behavioral persuasion of different methods has shown a continuous downward trend. The change trend is more obvious than the method in this paper. At the same time, analyzing the results of the acceptance of behavioral persuasion shows that the method in this paper is significantly greater than the direct persuasion method and the impact persuasion method, which shows that the method of this paper is more effective.

Table 2: Basic information collection form of survey objects.

Basic situation	Age, gender, occupation, driving experience, and ownership of a private car
Travel situation	The selected shared travel brand The reason and purpose for choosing shared travel Shared travel frequency and duration
Behavioral awareness	Degree of connection and perception of shared electric vehicles Share the status that appears in the minds of users General feelings and difficulties encountered during use Uncivilized and irregular behavior that users have learned about themselves

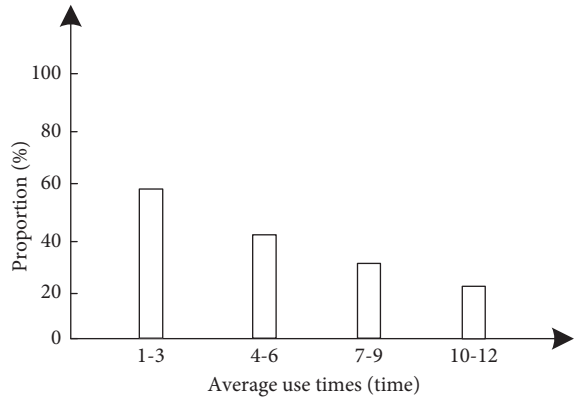


Figure 4: Average number of users using a car per month.

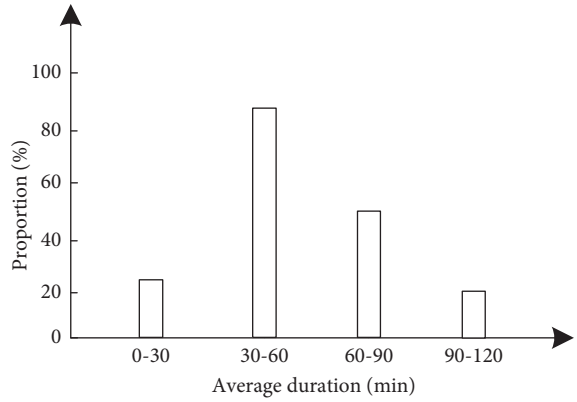


Figure 5: The average car usage time each time.

Table 3: Comparison results of persuaded persons' satisfaction.

User	Method of this paper	Direct persuasion	Impact persuasion
1	94.3	78.4	82.5
2	95.6	77.1	83.6
3	92.0	75.3	85.1
4	97.1	76.2	84.7
5	93.3	75.0	79.8
6	96.0	79.4	78.9
7	94.2	80.1	76.3
8	95.7	81.7	78.4
9	91.7	80.9	78.0
10	93.4	79.9	79.0

5.3.3. Persuasion Duration. Taking the persuasion duration as the experimental index, the direct diarrhea persuasion method, the impact persuasion method, and the method in this paper are compared. The results are shown in Figure 7. Among them, the persuasion time is recorded through the clock. The higher the value, the longer the persuasion time.

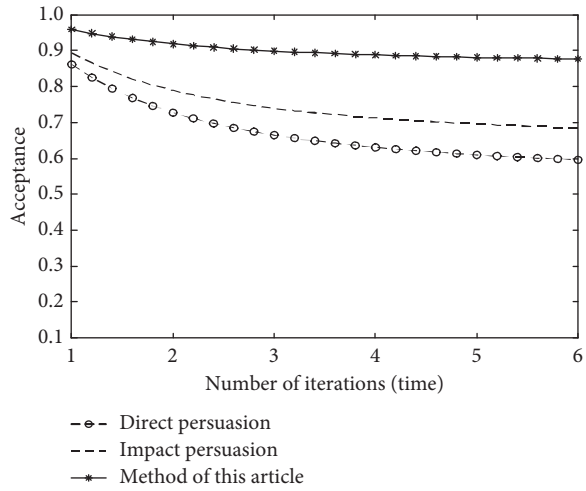


Figure 6: Comparison results of acceptance of behavioral persuasion.

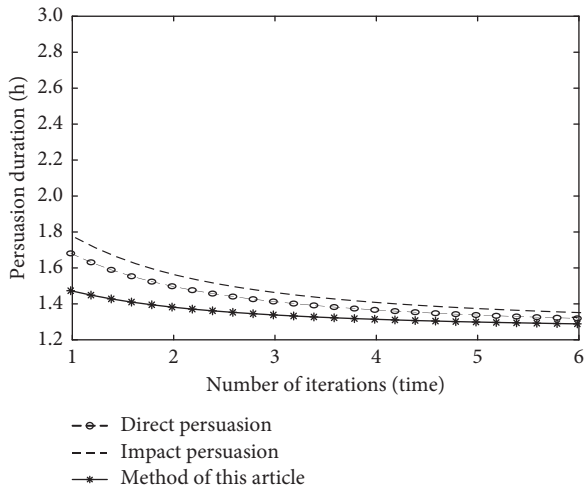


Figure 7: Comparison results of persuasion duration.

According to the analysis of Figure 7, the persuasion duration of different methods shows a continuous downward trend with the increase of the number of iterations. Among them, the persuasion duration of this method is less than 1.5 h, while the persuasion duration of the direct persuasion method and impact persuasion method is higher than that of this method, up to 1.8 h and 1.7 h. At the same time, by analyzing the persuasion time, it can be seen that the persuasion efficiency of this method is significantly higher than that of the direct diarrhea persuasion method and impact persuasion method, which shows that this method is more effective.

6. Conclusion

For the purpose of improving the acceptance of behavioral persuasion, an analysis method of persuasion design mechanism based on unconscious calculation is proposed. Taking the concept of persuasion and the goal of persuasion as the theoretical basis, the analysis of the experimental results shows that the method in this paper has a higher acceptance of behavioral persuasion and can obtain a higher satisfaction from the persuaded, which fully validates the advantages of the method in this paper.

Data Availability

The raw data supporting the conclusions of this article will be made available by the authors without undue reservation.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding this work.

References

[1] H. Zhang, M. Wang, L. Yang, and H. Zhu, "A novel user behavior analysis and prediction algorithm based on mobile social environment," *Wireless Networks*, vol. 25, no. 2, pp. 791–803, 2019.

[2] D. Shi, L. Wang, and Z. Wang, "What affects individual energy conservation behavior: p," *Energy Policy*, vol. 128, no. 5, pp. 150–161, 2019.

[3] E. Garr and A. R. Delamater, "Exploring the relationship between actions, habits, and automaticity in an action sequence task," *Learning & Memory*, vol. 26, no. 4, pp. 128–132, 2019.

[4] B. Yin, T. Teng, L. Tong et al., "Efficacy and acceptability of parent-only group cognitive behavioral intervention for treatment of anxiety disorder in children and adolescents: a meta-analysis of randomized controlled trials," *BMC Psychiatry*, vol. 21, no. 1, pp. 1–12, 2021.

[5] A. Sankar, D. Goldman, L. Colic, J. Kim, and H. Blumberg, "Targeting emotion regulation circuitry with a psychobehavioral intervention to reduce hopelessness in bipolar disorder," *Biological Psychiatry*, vol. 87, no. 9, p. 116, 2020.

[6] L. C. Ozoemena, A. B. Ikechukwu-Ilomuanya, L. C. Ezema, E. E. Offordile, and A. J. Ihebuzoaju, "Impacts of cognitive-behavioral intervention on anxiety and depression among social science education students: a randomized controlled trial," *Medicine*, vol. 98, no. 15, Article ID 14935, 2019.

[7] J. Tinoco, M. D. Granrut, D. Dias, T. Miranda, and A. G. Simon, "Piezometric level prediction based on data mining techniques," *Neural Computing & Applications*, vol. 32, no. 1, pp. 4009–4024, 2020.

[8] J. D. Campo-vila, A. Takilalte, A. Bifet, and L. Mora-López, "Binding data mining and expert knowledge for one-day-ahead prediction of hourly global solar radiation," *Expert Systems with Applications*, vol. 167, no. 8, Article ID 114147, 2020.

[9] A. Vanarse, J. I. Espinosa-Ramos, A. Osseiran, A. Rassau, and N. Kasabov, "Application of a brain-inspired spiking neural network architecture to o data classification," *Sensors*, vol. 20, no. 10, p. 2756, 2020.

[10] D. Griffiths and J. Boehm, "A review on deep learning techniques for 3D sensed data classification," *Remote Sensing*, vol. 11, no. 12, p. 1499, 2019.

[11] L. F. Maldaner, J. P. Molin, and M. Spekken, "Methodology to filter out outliers in high spatial density data to improve maps reliability," *Scientia Agricola*, vol. 79, no. 1, p. 2022, 2021.

[12] M. Shaeri and A. M. Sodagar, "A framework for on-implant spike sorting based on salient feature selection," *Nature Communications*, vol. 11, no. 1, p. 3278, 2020.

[13] A. Sak, B. Tha, C. Mik, and B. Aa, "Salient features of dufour and soret effect in radiative MHD flow of viscous fluid by a rotating cone with entropy generation," *International Journal of Hydrogen Energy*, vol. 45, no. 28, pp. 14552–14564, 2020.

[14] F. Borges, A. Pinto, D. Ribeiro, T. Barbosa, and D. Ferreira, "An unsupervised method based on support vector machines and higher-order statistics for mechanical faults detection," *IEEE Latin America Transactions*, vol. 18, no. 6, pp. 1093–1101, 2020.

[15] B. Koo, S. La, N. W. Cho, and Y. Yu, "Using support vector machines to classify building elements for checking the semantic integrity of building information models," *Automation in Construction*, vol. 98, no. 2, pp. 183–194, 2019.

[16] W. da Silva Cotrim, L. B. Felix, V. P. R. Minim, R. C. Campos, and L. A. Minim, "Development of a hybrid system based on convolutional neural networks and support vector machines for recognition and tracking color changes in food during thermal processing," *Chemical Engineering Science*, vol. 240, no. 4, Article ID 116679, 2021.

[17] Z. Khan, S. Fang, A. Koubaa, P. Fan, F. Abbas, and H. Farman, "Street-centric routing scheme using ant colony optimization-based clustering for bus-based vehicular ad-hoc network," *Computers & Electrical Engineering*, vol. 86, no. 1, Article ID 106736, 2020.

[18] S. L. G. de Oliveira and L. M. Silva, "Evolving reordering algorithms using an ant colony hyperheuristic approach for accelerating the convergence of the ICCG method," *Engineering with Computers*, vol. 36, no. 4, pp. 1857–1873, 2020.

[19] P. Stodola, K. Michenka, J. Nohel, and M. Rybanský, "Hybrid algorithm based on ant colony optimization and simulated annealing applied to the dynamic traveling salesman problem," *Entropy*, vol. 22, no. 8, p. 884, 2020.

[20] A. S. Sharma and D. S. Kim, "Energy efficient multipath ant colony based routing algorithm for mobile ad hoc networks," *Ad Hoc Networks*, vol. 113, no. 8, Article ID 102396, 2021.

[21] A. Akinyelu and E. S. Ezugwu, "Ant colony optimization edge selection for support vector machine speed optimization," *Neural Computing & Applications*, vol. 13, no. 3, pp. 1–33, 2020.

[22] K. Chatra, V. Kuppili, and D. R. Edla, "Texture image classification using deep neural network and binary dragon fly optimization with a novel fitness function," *Wireless Personal Communications*, vol. 108, no. 3, pp. 1513–1528, 2019.

[23] M. Ghosh, R. Guha, R. Sarkar, and A. Abraham, "A wrapper-filter feature selection technique based on ant colony optimization," *Neural Computing & Applications*, vol. 32, no. 12, pp. 7839–7857, 2020.

[24] S. K. Lakshmanaprabu, K. Shankar, S. Sheeba Rani et al., "An effect of big data technology with ant colony optimization based routing in vehicular ad hoc networks: towards smart cities," *Journal of Cleaner Production*, vol. 217, no. 20, pp. 584–593, 2019.

[25] M. Y. Shabir, A. Ullah, and Z. Mahmood, "ANT-colony based disjoint set assortment in wireless sensor networks," *Wireless Networks*, vol. 25, no. 8, pp. 5137–5150, 2019.

[26] A. F. Tuani, E. Keedwell, and M. Collett, "Heterogenous adaptive ant colony optimization with 3-opt local search for the travelling salesman problem," *Applied Soft Computing*, vol. 97, no. 4, p. 106720, 2020.

[27] M. G. H. Omran and S. Al-Sharhan, "Improved continuous ant colony optimization algorithms for real-world engineering optimization problems," *Engineering Applications of Artificial Intelligence*, vol. 85, no. 10, pp. 818–829, 2019.

[28] V. Sangeetha, R. Krishankumar, K. S. Ravichandran, and S. Kar, "Energy-efficient green ant colony optimization for path planning in dynamic 3D environments," *Soft Computing*, vol. 25, no. 15, pp. 1–21, 2021.

[29] M. C. Fadus, K. R. Ginsburg, K. Sobowale et al., "Unconscious bddbehavior dahy," *Academic Psychiatry*, vol. 44, no. 1, pp. 95–102, 2020.

[30] H. Song and N. Moon, "Eye-tracking and social behavior preference-based recommendation system," *The Journal of Supercomputing*, vol. 75, no. 4, pp. 1990–2006, 2019.

[31] L. Yan, H. Fang, and S. M. Lin, "Design of intelligent interactive financial management product for children with persuasive characteristics," *Packaging Engineering*, vol. 41, no. 22, pp. 108–113, 2020.

[32] Y. H. Du and M. S. Gong, "Research on the design of professional cognitive education based on behavior persuasion," *Design*, vol. 34, no. 4, pp. 118–121, 2021.

[33] Q. Y. Chen and N. N. Sun, "Research on interactive design strategies in online courses based on persuasive design," *Design*, vol. 32, no. 11, pp. 28–30, 2019.

[34] J. J. Li, "Research on healthy breakfast based on behavioral design theory," *Design*, vol. 33, no. 1, pp. 84–87, 2020.

[35] N. Guo, P. Li, and R. K. Wei, "Simulation of user behavior extraction based on generative tree outlier detection method," *Computer Simulation*, vol. 37, no. 6, pp. 257–261, 2020.

[36] J. L. Wang and W. L. Luo, "Prediction of user information behavior intention in mobile terminal message push process," *Computer Simulation*, vol. 36, no. 3, pp. 440–443, 2019.

Hindawi
Scientific Programming
Volume 2022, Article ID 3800169, 7 pages
https://doi.org/10.1155/2022/3800169

Research Article

The Mental Health Evaluation System of College Students Based on Data Mining

Peng Li

Faculty of Marxism, Lishui University, LiShui Zhejiang 323000, China

Correspondence should be addressed to Peng Li; lsxymkszyxylp@lsu.edu.cn

Received 14 December 2021; Revised 7 January 2022; Accepted 12 January 2022; Published 21 February 2022

Academic Editor: Baiyuan Ding

In order to solve the problems of high misevaluation rate and low work efficiency in the current mental health evaluation process of college students, a mental health evaluation system based on data mining algorithms is proposed. First, analyze the research status of college students' mental health evaluation and data mining algorithms and build a mental health evaluation system framework; then, collect college students' mental health questionnaire data, use the Apriori algorithm based on a three-dimensional matrix to analyze and classify, traverse each attribute of each transaction in the two-dimensional matrix, and directly obtain the frequent item set, frequent binomial set, and frequent three-item set by reading the three-dimensional attribute matrix and the mental health evaluation data to obtain mental health intelligent evaluation results. Finally, specific simulation experiments are used to analyze the feasibility and superiority of the mental health intelligent evaluation system. The results show that the system in the article overcomes the shortcomings of the current mental health intelligent evaluation system, improves the accuracy of mental health intelligent evaluation, improves the efficiency of mental health intelligent evaluation, and the system is more stable, which can meet the actual requirements of current college students' mental health evaluation.

1. Introduction

The psychological quality education of college students is one of the important contents of higher education. Through the psychological quality education, it is expected to cultivate the good character and will quality of college students, enhance the psychological adaptability and endurance ability, and promote the development of college students' physical and mental health [1–4]. With the effective support of relevant national policies, the psychological quality education of colleges and universities has gradually developed, and the professional team of psychological health quality education in colleges and universities has begun to take shape. Considering the coordination between student health management and career development, the related issues of psychological counseling have attracted more and more attention from school administrators [5–7]. However, in real life, whether facing college groups or other social groups, due to their own obstructive reasons, they are often unwilling to seek help from the counseling center so that they cannot get professional help in time when they have unhealthy manifestations such as psychological depression. Online psychological counseling can avoid this problem well, help people answer mental health problems more widely, provide various online communication and knowledge popularization, and obtain psychological support by browsing other people's comments, thereby effectively improving the efficiency of mental health management consulting, to promote the development of students' mental health and improve their psychological quality [7–11].

Mental health education for college students in China started in the 1990s and initially established the leading role of college students' mental health education courses in educational work, but it still has the problem of poor curricula education effects. Many scholars have conducted research on the issue of mental health consultation [12–14]. Literature [15] pointed out that the psychological problems of college students have become an important factor in integrating into the society. In order to effectively alleviate the psychological pressure of students, it designed a student

mental health consulting service system based on the Web and designed the system technology according to actual needs, but the safety of the system needs to be improved. Literature [9] designed a multisource mental health questionnaire, visualized the data obtained from the survey, studied the attribute correlation between various types of data, focused on analyzing the uncertain mental health data, and comprehensively judged students. For psychological problems, this method works well, but the data quality is not high. Literature [16] designed a mental health service system for the elderly based on Internet technology and analyzed it in combination with application cases. The practicality is good, but it takes a long time. Literature [17] introduced artificial intelligence, designed a knowledge system with a multilevel framework by building a knowledge base, using Bayesian and certainty factor methods to reason about data, and designed a psychological consultation expert system, but the system is not safe enough. Literature [18], with the support of Android technology, designed a mobile-end student mental health consultation system based on the C/S architecture, analyzed the client functional modules in detail, and designed the system architecture by combining the complete communication between the client and the server. The system function is clarified, but it takes a long time. In the 21st century, the country has gradually begun to pay more attention to the mental health of college students. The State Council (2004) No. 16 "Opinions on Further Strengthening and Improving the Ideological and Political Education of College Students" requires major colleges and universities across the country to attach importance to the mental health education of college students. Jiao Sizheng Hall (2011) No. 1 "Basic Construction Standards for Students' Mental Health Education in General Colleges and Universities (Trial)" requires colleges and universities to incorporate college students' mental health education into the school's talent training system and provide psychological counseling services through multiple channels. At the same time, more and more colleges and universities have begun to gradually establish and improve the psychological files of college students, paying attention to the monitoring and prevention of students' psychological crises. At present, many colleges and universities have to conduct a special mental health assessment (SCL-90) for freshmen when they first enter the school, to understand the psychological status of the students, to find out the psychological crisis in the students in time, to promptly guide and intervene according to the psychological status, and to deal with the students. Provide targeted psychological counseling. At the same time, it has gradually begun to consciously save the relevant data on the mental health of college students in a certain way to lay the foundation for the formation of psychological files.

With the development of science and technology, computer applications have gradually entered the operation and management work of all walks of life. Many colleges and universities have adopted psychological evaluation systems to evaluate the mental health of students. Through the psychological evaluation system, college psychological counseling staff can quickly collect students' psychological conditions and make judgments on the results, improve work efficiency, reduce work intensity, and provide great help to a certain extent. At the same time, the database also accumulates a large amount of psychological data [19–22]. However, most colleges and universities currently conduct psychological evaluations based on the access, query, statistics, and backup of basic psychological information of students. They have not conducted in-depth analysis of large amounts of psychological data, and have not fully utilized these data to dig out the hidden information. Some of the information knowledge grasp the trend of students' psychological development, so it cannot effectively provide decision-making assistance to the psychological counseling work [23]. Data mining is to dig out the hidden information knowledge from massive data. The development and application of data mining technology in the retail industry, financial industry, telecommunications industry, and other industries have been fruitful. If data mining technology is applied to the development of mental health management systems, a large number of fuzzy records in student mental health files will be removed through related algorithms. The random data information is processed, and the potential information is unearthed. To a certain extent, it will help the psychological counseling teacher to more scientifically and quickly conduct student psychological judgment and prevention, timely conduct psychological counseling and intervention, improve work efficiency, and reduce psychological incidents which happened [24].

In order to find a more effective psychological consultation system, this paper designs a mental health education consultation management system based on the Apriori algorithm on the basis of existing research results. By introducing data mining technology, it can dig out what is needed in massive data. On the basis of ensuring the quality of psychological quality teaching, improve the effect of psychological quality education for college students.

2. The Overall Structure of the Mental Health Evaluation System for College Students

In order to enable people to have an accurate understanding of their own mental health and, at the same time, to promote the scientific and informatization of mental health guidance, this paper constructs an intelligent evaluation system for mental health based on data mining, using scientific mental health evaluation tools. To comprehensively reflect the user's mental health level, the overall structure of the system is shown in Figure 1.

The mental health intelligent evaluation system based on data mining mainly includes four modules: basic information management module, evaluation test question management module, evaluation result analysis module, and personal psychological evaluation module.

(1) The basic information management module is the basic function in the text system, and its main function is to manage the basic information in the text system. This module has the function of personal information synchronization and evaluation type management. The former can synchronize the

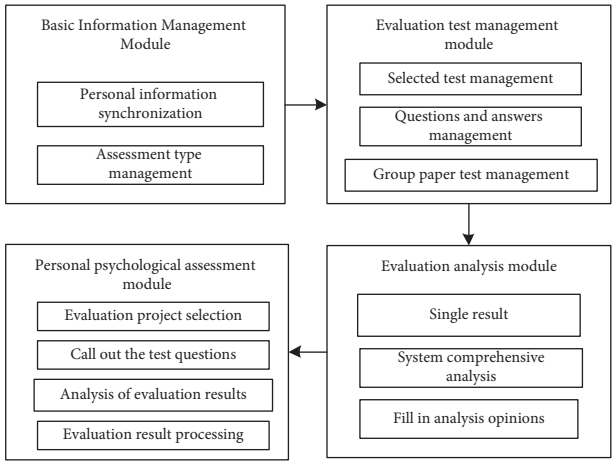


Figure 1: Overall diagram of the system structure.

personal information filled in by the user, accurately track the user's real-time personal situation, and ensure the accuracy of the follow-up evaluation of this system. The latter can manage the system in this paper. The latter can manage the assessment types in the system in a unified manner, and add or delete assessment types according to user needs. In the process of using the system, users have many types of choices, high versatility, and powerful overall system functions and can be widely used in the crowd.

(2) The evaluation test question management module is an important part of this system. The quality of the evaluation test questions seriously affects the correctness of the evaluation results. This module mainly includes three types of test questions: choice test questions, question answering test questions, and test paper grouping test questions. The form of multiple choice questions can be single-choice or multiple choice questions. There is a strong correlation between the test questions. Answers are made in the evaluation test questions through the selection method, and the answers are selected according to the user's choice. With different jumps to different question types to complete the answer, the answer points of each question are different. The question form of the essay questions is a question and answer, which is designed according to different types of assessment. The test question format of the test paper combines the choice test and the essay test. The above two test question types are combined in different evaluation types, and the jump to the multiple choice or essay is based on the user's answer, and the jump to the essay is terminated after the jump to the essay. This system can effectively reduce the user's psychological precautions through the skipping of thinking in the choice of test questions, increase their interest in testing, make the test easier to be accepted by them, and obtain a deeper understanding of the user's psychological state through the question and answer test.

(3) The evaluation result analysis module is responsible for statistically analyzing the psychological evaluation results of users, discovering the psychological problems of users in the first time, grasping their psychological conditions, and providing help to users. This module can provide users with the analysis results of a single evaluation, users with the comprehensive analysis results of multiple evaluations, and provides users with the function of filling in analysis opinions, so as to effectively and timely obtain user feedback information on the evaluation results.

(4) The personal psychological evaluation module is the main module in the system of this paper, which highlights the significance of the system of this paper to serve users. The evaluation item selection part provides users with all evaluation items of this system. Users can select evaluation items according to their own needs. The evaluation test question recall function can call out the corresponding evaluation test questions to the user for evaluation according to the evaluation item selected by the user and automatically save the user's evaluation. Answer the result, give the corresponding score, and finally process the user's evaluation result through the evaluation result processing function and feedback the processing result.

3. Design of Mental Health Evaluation System Based on Data Mining

The client software of the system in this paper uses the method of psychological evaluation data mining to realize the analysis of evaluation results, processes the obtained user evaluation result data and establishes a database, analyzes the evaluation results through the decision tree algorithm, and obtains an evaluation. Psychological assessment data mining includes extracting data from the database, cleaning the data, selecting the mining mode, and outputting the results. The complete data mining process is shown in Figure 2.

In this paper, an initial dataset is established by using the answering results of the college student survey questionnaire. After preprocessing of data integration, data extraction, data cleaning, and data conversion, a mental health assessment dataset to be mined is obtained, and the Apriori algorithm is used to mine the data. Association rules: Apriori algorithm completes data mining through continuous scanning, candidate item set, and minimum support comparison. Apriori algorithm needs to face a large amount of data when running, which is not conducive to efficient mining. For this reason, the system designed first divides the user psychological counseling database, divides it into multiple equivalent parts, and performs mining operations one by one. Finally, the data mining is completed in a unified manner, and the classification of the data is obtained through the mining results,so as to obtain the specific classification of

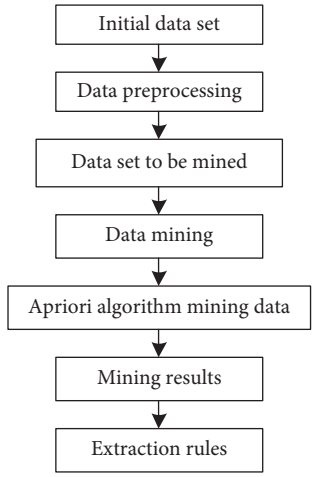


Figure 2: Data mining flowchart of mental health assessment.

the mental health status of college students. The operation steps are as follows:

(1) Divide the university student mental health questionnaire database into n equivalent parts.

(2) Perform frequent item set mining operations for a certain part, obtain data candidate item sets through scanning, compare it with the given minimum support to obtain frequent item sets, and then calculate new items based on frequent item sets and support set of candidates. In the same way, the frequent item set is obtained by comparing the candidate item set with the minimum support, and the operation is repeated continuously according to this step to obtain the final frequent item set.

(3) Use the previous step to process each part of the college student mental health questionnaire database to generate corresponding frequent item sets.

(4) Summarize the frequent item sets of each part and then generate the global frequent item sets. Based on the above, to complete the server-side information processing, we provide an important foundation for the efficient and safe operation of the system.

3.1. Algorithm Description. The Apriori algorithm is used to mine the association rules between the data and find the datasets that often appear in the data values, that is, the frequent item sets; by finding the connections in these sets, it is helpful for decision-making.

Definition 1. The support in association rules refers to the ratio of the number of transactions that contain both X and Y to all transactions in database D, provided that $X{=>}Y$ is established in database D:

$$\text{Support}(X \Rightarrow Y) = P(X \cup Y). \quad (1)$$

Definition 2. The confidence in the association rule refers to the ratio of the number of transactions that contain both X and Y to the number of transactions that only contain X in database D, namely,

$$\text{Confidence}(X \Rightarrow Y) = \frac{\text{Support}(X \Rightarrow Y)}{\text{Support}(X)} = P(Y/X). \quad (2)$$

The two attributes of confidence and support in association rules are obtained through experience. If they are satisfied,

$$\text{Sup}(X \Rightarrow Y) \geq \min\_\text{sup\&conf}(X \Rightarrow Y) \geq \min\_\text{conf}. \quad (3)$$

Then, X and Y are strong association rules and vice versa are weak association rules.

3.2. Apriori Algorithm Process Based on Three-Dimensional Matrix. The traditional Apriori algorithm will generate a large number of candidate item sets when the amount of data is large and the analysis categories are many, especially when generating binomial sets and tri-items' sets. And every time a higher level of frequent item sets are generated, the database needs to be rescanned, which will generate a lot of computational redundancy and low efficiency. On the basis of the Apriori algorithm, improvements are made to address these shortcomings of the traditional algorithm. The improved algorithm idea is as follows.

(1) Scan the database first and abstract the database into a two-dimensional matrix based on the attributes contained in all its transactions, which are used to store all the information in the database.

(2) Traverse each attribute of each transaction in the two-dimensional matrix. By reading two different attributes in the same transaction each time without repeating the reading, the three-dimensional upper triangular attribute matrix Matrix (i, j, k) is established, and the coordinates are established according to the corresponding attributes. The coordinate intervals in the three dimensions are all [1, N] (N is the largest attribute type). During the scanning process, each time the coordinates are repeated, the corresponding weight is increased by one, and the matrix can be expressed as

$$\text{Matrix}(i, j, k) = \text{Matrix}(i, j, k) + 1. \quad (4)$$

(3) Secondly, by reading the three-dimensional attribute matrix, we can directly obtain the frequent item set, frequent binomial set, and frequent three-item set. The space diagonal of the first hexagram limit of the three-dimensional matrix is the support of frequent item sets, and the coordinates (i,j,j) on the corresponding plane are the support of frequent binomial sets and coordinates (i,j,k). It is the support degree of the corresponding three-item set.

(4) In the transaction with the number of attributes less than k, there must be no possibility of containing k itemsets. Therefore, after getting the frequent three-item set, scan the database. Delete the affairs that

contain no more than four attributes, and simplify the database.

(5) Use the standard Apriori algorithm for subsequent calculations through the frequent three-item sets that have been obtained. The specific algorithm flow is shown in Figure 3.

4. Experiment Analysis

Randomly select 1,000 college students from a certain university as the experimental subjects, and compare the mental states of employees measured by this system and professional psychologists. 1 represents neuroticism, 2 represents compulsive paranoia, 3 represents emotional instability, 4 represents emotional weakness, and 5 represents communication sensitivity. The result is shown in Figure 4.

According to Figure 4, it can be seen that the psychological evaluation results of the system in this paper are basically consistent with those of professional psychologists, which proves that not only the system in this paper is effective but also its evaluation results are also very good. At the same time, it can be seen that the number of employees with neurasthenia accounted for the largest proportion. The number of employees with forced paranoia and emotional instability were 27% and 22%, respectively. The number of employees with neuroticism and communication sensitivity was relatively small. The reason was the high work pressure and competition of the unit. The pressure is strong. These employees can effectively reduce the proportion of various psychological problems after a period of emotional resolution based on the results of the processing opinions given by the system of this paper, which proves that the system of this paper is practical and can provide users with effective solutions to psychological problems.

In order to verify the data processing efficiency of the system in this paper, the Internet-based mental health service system for the elderly and the online forum user mental health automatic evaluation system based on multifeature fusion are used as the comparison system. The time result is shown in Figure 5.

Analyzing Figure 5, we can see that the data processing time of the system in this paper has been lower than that of the other two systems. As the amount of data continues to increase, the data processing time of the three systems has changed. The data processing time of the comparison system 1 and the comparison system 2 has increased sharply with the increase in the amount of information, and the fluctuation range is large and the stability is poor. The increase in processing time is small, the curve is smoother, and the stability is good. After the data volume reaches 5×103 GB, the data processing time gradually stabilizes, which proves that the system in this paper has high data processing efficiency and strong stability.

Different noises were added to the three systems to evaluate the mental state of employees, and the accuracy of the three systems' mental health evaluation was tested. The results are shown in Table 1.

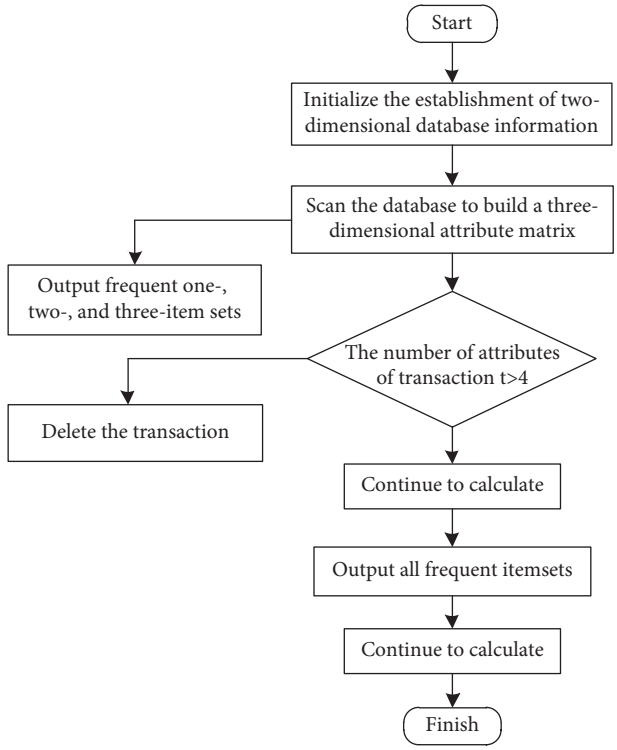


Figure 3: Apriori algorithm flow.

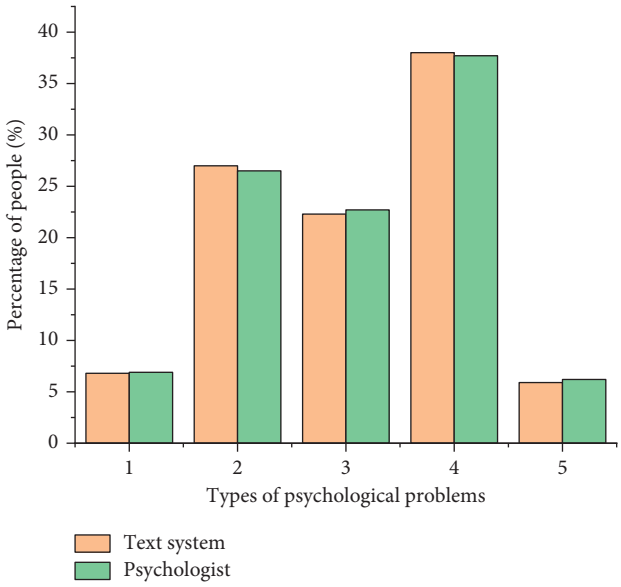


Figure 4: Comparison results.

It can be seen from Table 1 that with the continuous increase of noise, the mental health evaluation accuracy of the three systems has decreased, but the mental health evaluation accuracy of the system in this paper has been above 90%, and the accuracy rate has not decreased much. It is relatively stable, while the mental health evaluation accuracy of the other two systems is less than 88%, indicating that the different noises in the three systems have the least impact on the mental health evaluation accuracy of the

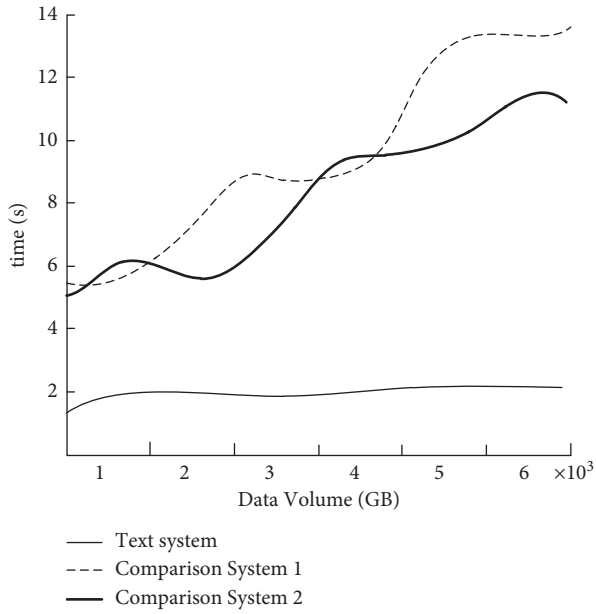


Figure 5: Data processing time of the three systems.

Table 1: Mental health evaluation accuracy.

Noise (dB)	20 (%)	40 (%)	60 (%)	80 (%)
Text system	93.33	92.08	92.77	91.11
Comparison system 1	88.12	87.64	87.63	85.88
Comparison system 2	81.68	81.62	81.11	80.33

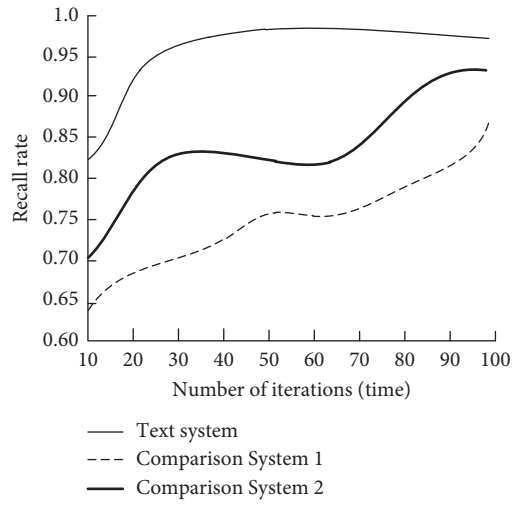


Figure 6: Comparison of system recall rates.

system in this paper. The antinoise performance of the system in this paper is better. The accuracy is high.

The information recall rate can be used to verify the system's data scheduling capability. The higher the recall rate, the more stable it is, and the better the system's data scheduling capability. Test the information recall rate of the three systems in the psychological evaluation, and the results are shown in Figure 6.

It can be seen from Figure 6 that the recall rates of the three systems all increase with the increasing number of iterations. The recall rate of the system in this paper is the highest, and it is stable after the number of iterations reaches 20. It has good recall and adaptive scheduling performance. The recall rate of comparison system 1 and comparison system 2 fluctuates greatly, and the recall rate is lower than that of the system in this paper.

5. Conclusion

In order to improve the effect of college students' mental health evaluation, this paper constructs a mental health intelligent evaluation system based on the Apriori algorithm, which can fully help college students understand their own mental health, solve their psychological problems, and enhance their self-awareness. By collecting college students' mental health questionnaire data and using the Apriori algorithm based on the three-dimensional matrix, we analyze and classify the mental health evaluation data and obtain the mental health intelligent evaluation results; finally, the feasibility and superiority of the mental health intelligence evaluation system are analyzed by specific simulation experiments. The results show that the system in the study overcomes the shortcomings of the current mental health intelligent evaluation system, improves the accuracy of mental health intelligent evaluation, improves the efficiency of mental health intelligent evaluation, and has better system stability, which can meet the actual requirements of current college students' mental health evaluation. In order to make the system of this paper more functional and have better development, it is also necessary to improve the analysis of system evaluation results, improve the situation of system data loss during power failure, increase the system's data backup and recording functions, solve the problem of waste of storage space, and make the function of this text system more perfect.

Data Availability

The dataset used to support the findings of the study can be obtained from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] X. Li, Y. Fei, and Y. Fan, "A review of research on occupational mental health of construction workers," *China Safety Science Journal*, vol. 30, no. 9, pp. 202–210, 2020.

[2] M. Wang, F. Wang, and J. Wang, "Reliability and validity analysis of the Chinese version of the psychological flexibility comprehensive questionnaire in financial institution staff," *Chinese Journal of Behavioral Medicine and Brain Science*, vol. 27, no. 1, pp. 72–77, 2018.

[3] S. Li and Q. Huang, "A review of researches on the intervention efficacy of gardening activities based on quantitative measurement of elderly physical and mental health

indicators," *Journal of Northwest University*, vol. 50, no. 6, pp. 852–866, 2020.

[4] M. Wang, F. Han, and L. Jia, "Meta-analysis of the detection rate of depressive symptoms and related factors in college students," *Chinese Mental Health Journal*, vol. 34, no. 12, pp. 1041–1047, 2020.

[5] L. Wen, N. Zhang, and Z. Yu, "Meta-analysis of the relationship between emotion regulation and mental health of children and adolescents," *Chinese Journal of Clinical Psychology*, vol. 28, no. 5, pp. 1002–1008, 2020.

[6] S. Zhang, T. Liu, and F. Xia, "Social network analysis of interpersonal relationship and mental health of college students," *Chinese Mental Health Journal*, vol. 34, no. 10, pp. 855–859, 2020.

[7] J. Liang and H. Luo, "The application of big data mining methods in the psychological early warning system of college students," *Chinese School Health*, vol. 39, no. 12, pp. 1821–1824, 2018.

[8] D. Liu, X. Xia, and C. Wan, "Automatic evaluation of online forum users' mental health based on multi-feature fusion," *Chinese Journal of Computers*, vol. 42, no. 7, pp. 1553–1569, 2019.

[9] X. Chen, M. Tong, and S. Chen, "Visual analysis of multi-source mental health questionnaire data for college students," *Journal of Computer Aided Design and Graphics*, vol. 32, no. 2, pp. 181–193, 2020.

[10] Z. Liu and X. Song, "A multivariate decision tree algorithm that can be used for categorized attribute data," *Journal of Northeastern University*, vol. 41, no. 11, pp. 1521–1527, 2020.

[11] J. Liang and H. Luo, "The application of big data mining methods in the psychological early warning system of college students," *Chinese School Health*, vol. 39, no. 12, pp. 1821–1824, 2018.

[12] H. Gong, "Construction of a mental health management platform for college students based on new media technology," *Library Work and Research*, vol. 1, no. 10, pp. 124–128, 2017.

[13] K. Jia, H. Li, and Y. Yuan, "Application of data mining based on Apriori algorithm in mobile medical system," *Journal of Beijing University of Technology*, vol. 43, no. 3, pp. 394–401, 2017.

[14] F. Zhao and B. Liu, "Association analysis of college students' performance based on improved Apriori algorithm," *Journal of Qiqihar University (Natural Science Edition)*, vol. 34, no. 1, pp. 11–15, 2018.

[15] H. Wang, "Design and implementation of a web-based mental health consultation service system for college students," *Journal of Xi'an University of Arts and Science (Natural Science Edition)*, vol. 20, no. 1, pp. 72–76, 2017.

[16] Y. Qian, D. Chen, and M. Zhu, "Design of mental health service system for the elderly based on Internet technology," *Packaging Engineering*, vol. 38, no. 22, pp. 53–59, 2017.

[17] L. Yang, "Psychological consultation expert system reasoning model design," *Computer and Digital Engineering*, vol. 46, no. 6, pp. 1145–1148, 2018.

[18] Y. Yin, *Design and Implementation of a mobile College Student Mental Health Consulting Service System Based on C/S Architecture [D]*, Jilin University, Changchun, 2018.

[19] L. Deng, J. Liang, and B. Li, "The status quo of mental health education in primary and secondary schools: the different perspectives of psychological teachers and school administrators," *Teacher Education Research*, vol. 30, no. 4, pp. 58–64, 2018.

[20] R. Su, *The Design and Implementation of an Android-Based Mental Health Management System for College Students*, Shandong University, Jinan, 2017.

[21] Y. Zhang, C. Hu, and S. Huang, "Data mining and analysis method of secondary equipment defects based on Apriori algorithm," *Automation of Electric Power Systems*, vol. 41, no. 19, pp. 147–151, 2017.

[22] Z. Xie and P. Wang, "Parallel matrix Apriori algorithm based on MapReduce architecture," *Computer Application Research*, vol. 34, no. 2, pp. 401–404, 2017.

[23] R. Zhong and H. Wang, "Specific data query technology in the university cloud computing management system based on data mining," *Modern Electronic Technology*, vol. 41, no. 2, pp. 130–132, 2018.

[24] M. Luo, "Research on students' mental health based on data mining algorithms," *J Healthc Eng*, vol. 20211382559 pages, 2021, PMID: 34733450; PMCID: PMC8560244, Article ID 1382559.

Hindawi
Scientific Programming
Volume 2022, Article ID 1838468, 10 pages
https://doi.org/10.1155/2022/1838468

Research Article

Taekwondo Action Recognition Method Based on Partial Perception Structure Graph Convolution Framework

Jianqiao Liang[1] and Guocai Zuo[1,2]

[1]*Shenyang Institute of Urban Construction, Shenyang, Liaoning 110167, China*
[2]*Hunan Software Vocational and Technical University, Hunan, Xiangtan 411100, China*

Correspondence should be addressed to Guocai Zuo; zuoguocai@hnsoftedu.com

Received 22 December 2021; Revised 4 January 2022; Accepted 13 January 2022; Published 21 February 2022

Academic Editor: Baiyuan Ding

Action recognition in Taekwondo competitions and training is an important task, which can provide a very valuable reference factor for technicians, athletes, and coaches. We propose a graph convolution framework with part of the perception structure to recognize, decompose, and analyze Taekwondo actions. Taking advantage of the long short-term memory of a part of the perception structure, the recognized Taekwondo actions are marked in time series, and then features are extracted from the graph convolution level to obtain the spatial and temporal associations between joints. Predict the action category and perform score matching based on the manual tag database. Finally, it is verified on our self-made Taekwondo competition data set. Our method has an average accuracy of 90% in action recognition, and an average action score matching rate of 74.6%. The accuracy of action recognition is high, which provides great assistance to Taekwondo e training and competitions.

1. Introduction

With the rapid development of science and technology, the sports industry urgently needs the intervention of artificial intelligence technology. This field has gradually attracted a lot of research. Previously, sports training was done in a more traditional manner. When the coach explains the training essentials to the athletes, the manner they play the video is not obvious, lacking in sensibility and interaction for the athletes [1]. With the upgrading of computer vision technology, human motion recognition began to be applied to the sports training industry. The development of these technologies has directly promoted the interpretation, prediction, and interaction of training actions in the sports industry. They have become a factor in the successful application of techniques and tactics in the process of sports training [2]. Taekwondo competitions have been updated day by day, and new technical systems have been introduced for training and practical exercises, which have a direct connection to Taekwondo athletes to obtain better results. Different from the development of the scoring system, the model of athletes adapting to technical and tactical methods to achieve victory has gradually become standard [3]. It is important to understand the technical actions of players in taekwondo competitions. The number of actions that occur in taekwondo battles is very high [4], but only a few achieve the main goal, which is to score [5, 6]. The main focus of understanding the scoring actions and thus promoting victory is not unique to Taekwondo. It has also been the research object of other fighting sports such as karate [7], boxing [8], fencing [9], and judo [10].

In order to be able to recognize taekwondo actions, in-depth analysis of the scoring points of each action is required, to provide athletes with skills during the training process. This paper uses an action recognition algorithm to recognize and decompose the actions of athletes in the Taekwondo competition. However, action recognition has higher requirements for scenes and characters, so in most cases, the tasks faced are more complicated. Therefore, most scholars have begun to study motion recognition methods based on bone joint points. The bone joint point–based action recognition

method and the image-based action recognition method are very different in principle, and the recognition effect is also very different [11]. Bone-based methods mainly rely on human skeleton data, and image recognition relies on a variety of pattern information, which is more susceptible to nonobjective factors. With the development of convolutional neural networks, human bone keypoint detection algorithm technology has made great breakthroughs in accuracy and precision. Skeleton data, as the mainstream input for action recognition, provide a more solid foundation for action recognition algorithms. Most researchers prefer to use bone data instead of RGB data, because bone data are compact and low in cost, and are not affected by nonobjective factors.

In the previous neural network action recognition tasks, the spatial relationship between joints has not been used, and features can only be extracted from the time level. Some researchers started to focus on this problem and validated various neural network models through experiments and finally found that long short-term memory (LSTM) neural network can obtain the spatial relationship between joints. The LSTM neural network can divide the bones and joints into independent parts, then implement partial feature perception for each part, so as to extract the features of the part, and finally obtain the spatial relationship between joints through the association between the partial features. Each part corresponds to an independent LSTM network, and finally, all the outputs are combined. The proposal of this method improves the use of space. However, the network is subject to manual network architecture predetermined rules, resulting in poor overall network robustness. The graph convolutional network (GCN) is used to recognize actions [12], which enriches the feature capture at the time and space level. Although GCN is often used in the analysis and classification of social networks, it is extraordinary in processing arbitrary graph structures [13]. The application of CNN in action recognition only stays at the 2D and 3D levels. The GCN can learn feature information from adjacent joints in the skeleton data. Some researchers have focused on designing a GCN architecture suitable for bones. For example, Li proposed a spatial temporal graph convolutional network (ST-GCN), which first modeled the bone joint point data, performed spatiotemporal convolution on all joint points, and obtained the spatial correlation and time correlation, and then used filters for feature capture. Gao et al. [14] used GCN to obtain spatiotemporal features from the bone joint point, which gave the model strong expressive ability and good generalization ability.

Based on the previous investigations and experiments, the purpose of this research is to identify and decompose the scoring techniques in the Taekwondo competition. The analysis results will provide valuable references for technicians, coaches, and athletes; help improve Taekwondo training methods; and optimize skills and tactics systems [15]. Therefore, to extract the important features of bone joint point-based Taekwondo action, a new approach combining LSTM and ST-GCN is provided. According to the temporal and spatial relationship between human bones, each frame is coded in sequence. Through the coded bone joint point sequence, the dynamic changes can be mapped in real time. The LSTM algorithm recognizes the network topology features as an input.

2. Related Work

In the study of motion recognition, compared with RGB data, bone data are not affected by a nonobjective environment and have high robustness, so it is widely used. Skeleton diagrams belong to non-Euclidean space and are fundamentally different from grid data. The skeletal data are arranged in a grid to use the structure of the neural network to obtain powerful feature learning capabilities [16, 17]. However, Monti et al. [18] raised a problem that in the vertex domain, meaningful operators cannot be fully expressed. Therefore, most current research points to GCNs because the operators defined by GCN networks are not in the vertex domain, but in the non-Euclidean space. Yan et al. proposed a bone-based action recognition method for the first time [19]. Gao et al. proposed a GCN network based on sparse regression to take advantage of the dependence between joints [14]. Shi et al. proposed a dual-stream method that uses the GCN architecture to capture the second-order information between the joints while acquiring the joint information, and then uses the classification strategy to filter the features [20]. Our method still uses graph convolution, as graph topology matrix can help us better traverse the skeleton task.

NAS (neural architecture search) [21] is a crucial component of automated machine learning, which automatically builds neural networks. At present, there have been a large number of literature focusing on NAS research, such as the reinforcement learning and black-box optimization researched by Zoph; the evolutionary search method researched by Real et al. [22]; the gradient-based processing proposed by Liu et al. [23]. In addition, Liu has unique insights in search space network design; in the realm of semantic segmentation and picture automated classification architecture design, Saxena et al. have more in-depth research. Although the method based on RGB data is not robust, Peng et al. still have a very high right to speak in the field of RGB data processing [24]. At present, Pham et al. proposed to optimize the graph neural network of ENAS to achieve inductive learning and citation, and the effect is also quite good [25]. Compared with our tasks, there is a certain gap, because its goal is to find a network with only two to three layers of conversion, propagation, and aggregation functions so that the post-decomposition work of Taekwondo can be effectively completed.

Graph convolutional neural networks have great advantages in irregular data and biological data processing, and they have gradually been applied to social networks and have begun to show their prowess. At the beginning of the definition of GCN, Defferrard proposed a spectral domain method, which aims to decompose the Fourier domain model at the time level [26]. Monti proposed a node domain method to realize the free switching of operators between graph nodes and leader nodes. However, the above studies are difficult to simulate the global structure of GCN. In order to further optimize GCN, Veli introduces an attention mechanism, which realizes automatic selection of key information [27] in all outputs. Since then, Velickovic and others have continued to explore the road of optimizing

GCN and proposed a node classification method of attention mechanism graph with better results. Sankar also researched the attention mechanism. He said that the attention mechanism was introduced into the dual dimensions of time and space, and in the process of achieving self-attention, he predicted better characteristics. Nevertheless, our approach is unique. We need to construct a dynamic display of joint points through the correlation characteristics between Taekwondo actions. The other is the problem of calculating the weights of different representations or different frames.

3. Method

3.1. Taekwondo Action and Scoring. Taekwondo is a martial art mainly based on legwork. In actual combat, the use of footwork can ensure that the power of the legs is fully utilized, which is of great significance for achieving victory in actual combat. Taekwondo mainly uses the hind legs as the core strength, so the footwork of Taekwondo has distinct characteristics. The athlete's center of gravity needs to fall between the feet or the front foot, and most of the body posture is to stand sideways to protect the body and the vital parts below so that the back legs can be turned by twisting the waist and turning the hips to increase the strength and speed of hitting. In our research, we mainly analyze the scoring points of eleven taekwondo actions, which are a front kick, push kick, cross kick, downward kick, side kick, hook kick, back kick, backspin kick, single leg kick, double leg kick, and double flying kick. According to relevant research, taekwondo is a sport mainly based on leg actions. Therefore, this paper, as the first stage of Taekwondo action decomposition research, will take leg action as the main research object. In the later research, the action decomposition research of other body parts such as the hand will be gradually increased. The main Taekwondo leg actions are shown in Figure 1.

In the scoring system of Taekwondo, a precise and powerful hit to the effective scoring area using the allowed techniques is required to score in the competition. 3 points are awarded for headbutts, 2 points for spin kicks and back kicks, 1 point for other techniques, and no additional points for the referee's reading. The maximum score for a technique is 3 points, and the valid scoring areas include the abdomen and both ribs, and the areas of the face that are allowed to be attacked. If the permitted technique is used to hit a nonvalid scoring area protected by protective gear, the knockdown will be scored. The specific scoring rules are shown in Table 1.

3.2. Graph Convolutional Network. We will go through the search-based GCN in detail in this part. To keep the article self-contained, we go over how to utilize GCN to represent spatial maps briefly.

Consider the undirected graph $G = \{V, E, A\}$ composed of $n = |V|$. The nodes connected by $|E|$ are encoded in the adjacency matrix $A \in R^{n\times n}$. Let $X \in R^n$ X be the input representation of G, and $\{x_i, \forall i \in V\}$ be its n elements. Then, in order to model the representation of G, the graph is Fourier transformed, so that the transformed signal, such as in Euclidean space, can then process basic calculation formulas, such as filtering. For this reason, the graph Laplacian L whose normalization is defined as $L = I_n - D^{-1/2}AD^{-1/2}$, and $D_{\text{ii}} = \sum_j A_{\text{ij}}$ is used for Fourier transform. Then, the graph filtered by the operator g_θ, parameterized by θ, can be expressed as

$$Y = g_\theta(L)X = Ug_\theta(\Lambda)U^T X, \tag{1}$$

where Y is the extracted graphic feature. U is the Fourier basis, which is a set of orthogonal eigenvectors of L, so $L = U\Lambda U^T$, where Λ is the corresponding eigenvector. However, multiplying with the eigenvector matrix is expensive. The computational burden of this nonparametric filter is $O(n^2)$. As suggested by Hammond et al. [28], the filter g_θ can be a good approximation of the Chebyshev polynomial of order R.

$$Y = \sum_{\text{r}=0}^{R} \theta_r' T_r(\widehat{L})X \tag{2}$$

where θ_r' represents the Chebyshev coefficient. Chebyshev polynomial $T_r(\widehat{L})$ is recursively defined as

$$T_r(\widehat{L}) = 2\widehat{L}T_{r-1}(\widehat{L}) - T_{r-2}(\widehat{L}), \tag{3}$$

$T_0 = 1$ and $T_1 = \widehat{L}$. Here $\widehat{L} = 2L/\lambda_{\max} - I_n$ is normalized to [−1, 1]. For equation (2), the work in Kipf's research [29] sets $R = 1$, $\lambda_{\max} = 2$ and adapts the network to this change. In this way, the first-order approximation of the spectrogram convolution is formed. So,

$$Y = \theta_0' X + \theta_1'(L - I_n)X = \theta_0' X + \theta_1'\left(D^{-1/2}AD^{-1/2}\right)X. \tag{4}$$

Similarly, θ_r' can also be approximated by a uniform parameter θ, that is, $\theta = \theta_0 = -\theta_1$, so that the training process can adapt to the approximation error, then

$$Y = \theta\left(I_n + D^{-1/2}AD^{-1/2}\right)X. \tag{5}$$

The computational cost is $O(|E|)$. Multiple GCN layers can be stacked to obtain advanced graph features. For simplicity, in the following sections, we set $L = I_n + D^{-1/2}AD^{-1/2}$. In general, $X \in R^{n\times C}$ and multichannel. Therefore

$$Y = LX\theta. \tag{6}$$

Yan proposed the ST-GCN method in 2018. This method uses the bone joint matrix as input to obtain the spatiotemporal features between the joint points. The method we propose takes the function of the module as the demarcation point, separately obtains the spatiotemporal features before the skeleton node, and maps into a dynamic network, abandoning the predefined graph.

3.3. Inception Structure. A huge proportion of classic scholars have demonstrated that the convolutional neural network will cause the receptive field to have the characteristics of scaling invariance [30]. Therefore, in the application process, the network should be optimized, and multi-

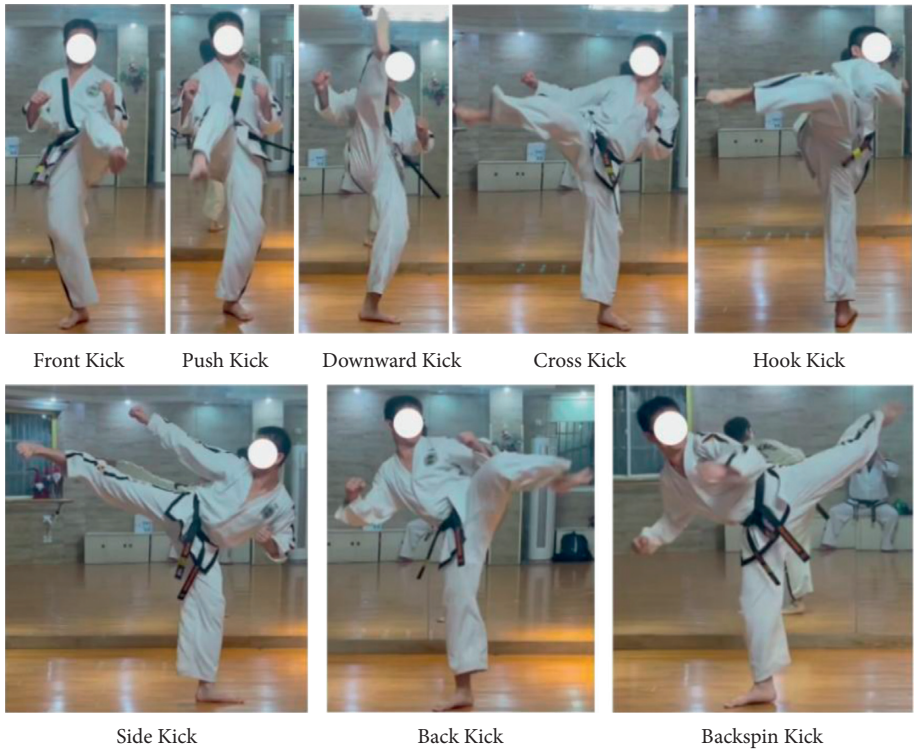


Figure 1: Regular taekwondo action.

Table 1: Taekwondo action scoring and requirements.

Score	Requirement
1	Attack with hands to the middle or high section; attack with feet to the middle section; perfect defense
2	Attack with the feet to the high section; jump into the air and hit the high section with hands (foot off the ground); jump with the feet to attack the middle section
3	Jump up and attack to the high section; jump up to 180 degrees and turn and kick to the middle section; jump up to 180 degrees or more and turn around and attack to the high section
4	Jump up 180-degree turnaround leg kick to high section; jump up 360 degrees or more turnaround leg kick to the middle section
5	Jump 360 degrees or more, turn and kick to the high section

scale filters should be appropriately combined selectively to achieve good generalization performance.

Szegedi et al. [31] found a structure, which they called inception, which can fuse convolution kernels of different scales, as shown in Figure 2. When the previous layer's output becomes the input, the extracted feature map will be shunted, and the convolution kernels of different sizes will obtain feature information of different scales. However, this method also has a disadvantage, that is, too many kernels can easily cause gradient explosion. To deal with the issue, Lin et al. [32] add one-dimensional convolution (conv1 × 1) before the original structure, followed by the activation function ReLU. In this way, the one-dimensional convolution adds the receptive fields corresponding to each other, greatly reducing the amount of calculation, and also realizing the feature fusion between different channels.

Considering that the trained data are all single-channel data, they are represented as 2D matrix in the neural network. And, each part of the human bone data that we take

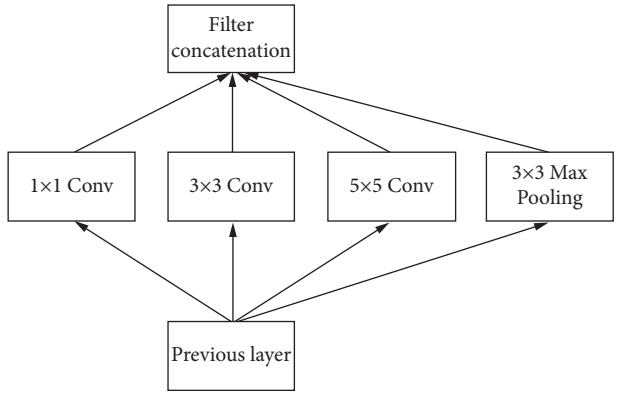


Figure 2: Inception structure.

contains three adjacent joints, forming a three-dimensional data with a column number of 3. A time series of joint coordinates is represented by every column of the matrix.

A convolution kernel with a scale of $m \times n$ has a span of scale m in the time dimension, and arranges n of each scale in a 3D sequence. Consider it a single-joint temporal convolution. From the perspective of the $m \times n$ convolution kernel, m is the time window spanning the entire time sequence, and has the following representation:

$$h^{m\times n} = w_0 p_t^k + w_1 p_{t+1}^k + \cdots + w_{m-1} p_{t+m-1}^k. \tag{7}$$

When the $m \times n$ convolution kernel is applied to extract spatial and temporal features from the action labels, the following output forms can be adopted:

$$h^{m\times n} = \sum_{i=0}^{m-1} \sum_{j=0}^{m-1} w_{ij} p_{t+i}^{k+j}. \tag{8}$$

The pixels in the image structure have continuity, but the graph structure does not have this feature. Therefore, pooling operations cannot be performed between different joint point coordinates. Therefore, in the inception structure, the pooling operation only appears in the temporal convolution dimension, and the size of the pooling operation is $(s \times 1)$. For a single joint, its trajectory can be regarded as continuous. In this case, in the temporal dimension, the pooling process can be carried out.

3.4. Partially Aware Network. Long short-term memory neural network (LSTM) was proposed by Hochreiter and Schmidhuber [33] in 1997.

LSTM is a derivative of Recurrent Neural Network (RNN). Since 2010, it has been proven that RNN has been successfully applied to speech recognition [34], language modeling [35], and text generation [36]. However, the disappearance of gradients and explosions makes RNN difficult to apply to long-term dynamics research. As an improved network of RNN, LSTM can handle this problem well. LSTM gives the network a lot of freedom, so that the network memory unit has an adaptive solution to learn and update information, which greatly improves the performance of some perception networks.

Assume that $X = (x_1, x_2, \ldots, x_n)$ represents an input sentence composed of word representations of n words. In every position t, the RNN produces a hidden layer h in the middle denoted as y_t, and the hidden state h_t uses a nonlinear activation function to update the previously hidden layers h_{t-1} and the input x_t, as follows:

$$\begin{aligned} y_t &= \sigma\left(W_y h_t + b_y\right), \\ h_t &= f\left(h_{t-1}, x_t\right), \end{aligned} \tag{9}$$

where W_y and b_y are the parameter matrices and vectors learned during the training process, and σ represents the element-wise softmax function.

The LSTM unit includes an input gate i_t, a forget gate f_t, an output gate o_t, and a memory unit c_t to update the hidden state h_t, as follows:

$$\begin{aligned} i_t &= \sigma\left(W_i x_t + V_i h_{t-1} + b_i\right), \\ f_t &= \sigma\left(W_f x_t + V_f h_{t-1} + b_f\right), \\ o_t &= \sigma\left(W_o x_t + V_o h_{t-1} + b_o\right), \\ c_t &= f_t \odot c_{t-1} + i_t \odot \tanh\left(W_c x_t + V_c h_{t-1} + b_c\right), \\ h_t &= o_t \odot \tanh\left(c_t\right), \end{aligned} \tag{10}$$

where $\odot$ is a kind of function which is similar to the multiplication operate, V represents a matrix related to weight, and b represents the learning vector. To increase the model's performance, morpheme training was carried out on two LSTMs. The first one is a morpheme that begins on the left and works its way to the right; the next one is a reverse duplicate of a character. Before passing to the next layer, the outputs of the forward and reverse passes are combined in series. Finally, the prediction value is observed using the activation function.

After understanding the partial perception algorithm LSTM, I was inspired by it, because in the human body recognition process, the human skeleton will be divided into multiple parts. Each part is an interconnected joint. These parts composed of joints are made by hand, for the graph convolution to be able to explore the relationship between these parts and extract the corresponding spatial features of the joint points. To obtain the information of a point in GCN, it is necessary to start from the field of that point. According to the adjacency matrix in the field, the skeleton data are automatically segmented, and then all the feedback information is input to the next joint point to complete the capture of the feature points of the entire human skeleton. Through this operation, the defects of manual design features are avoided, and the spatial features on the time series are obtained.

If an ordinary convolutional neural network is used, all parts will be merged into a whole for feature extraction of convolution operations. Partial perception networks can divide joints into different departments and capture individual features for each part. Separately extracting features in this way helps to explore the connection between parts, that is, the spatial temporal relationship between joints, as shown in Figure 3.

3.5. Network Structure. Our Taekwondo action recognition network structure is shown in Figure 4. When the athlete's bone features are input to the Part Dividing layer, all joints are grouped and merged, and each 3 joints share one part. A channel matrix is formed when all of the parts are put together, and then the three-channel spatiotemporal features are extracted from these channel matrices through the graph convolutional layer. Finally, all the extracted spatial temporal features are input into the fully connected layer, as shown in Figure 5.

To obtain the actions of the Taekwondo athletes, we stacked 7 GCN units. Each GCN unit is made up of a partial perceptual GCN layer, a BatchNorm layer, and a ReLU activation layer. The convolutional layer in each unit is filled

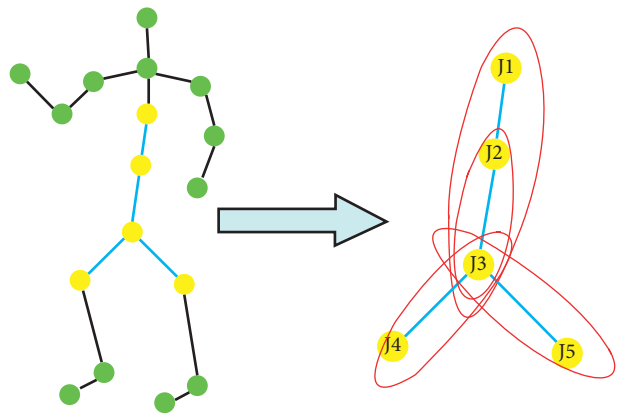


Figure 3: Spatial relationship of human skeleton joints.

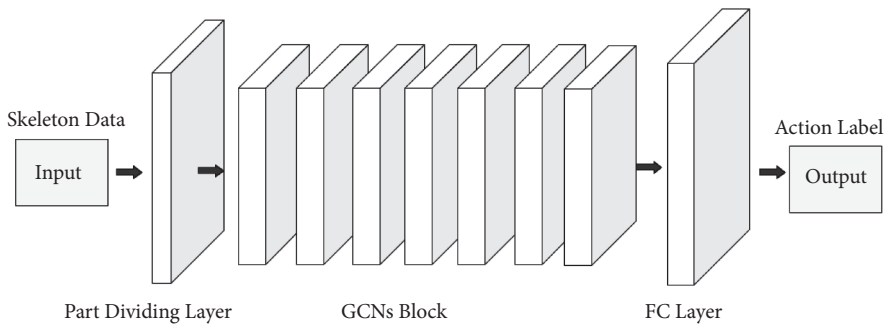


Figure 4: Global network structure.

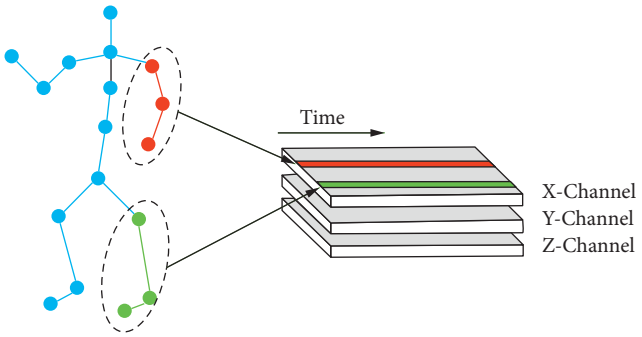


Figure 5: Part dividing layer structure.

according to requirements, with a step size of 1, as shown in Figure 6. Therefore, the size of the bone data will not change from input to output through these convolutional layers.

According to the preliminary test, the 7, 5, and 3 layers are supplemented with the initial structure. Then, set the first two layers to (3×1) filters. The third layer is the Inception structure, supplemented by (1×1), (3×1), and (3×3) filters. The fourth and sixth layers are supplemented by (3×1) filters. The fifth and seventh layers are consistent with the first layer structure. The global network structure is shown in Figure 4.

4. Experiments

4.1. Data set. Taekwondo is a kind of sports, and there is no dedicated data set for taekwondo action recognition in the world. To verify the performance of our method, we applied to the relevant departments for the data of the Chinese Taekwondo competition. In data processing, we use Openpose to preprocess the action video data. After the preprocessed data, we manually proofread the Taekwondo action, and then perform a series of tasks such as label production and action score matching. The following experiments are all based on this data set. A total of 45,000 training sets and 5,000 test sets are prepared.

4.2. Training. All experiments in this paper are executed on the Ubuntu framework, and configure python 3.7 as the language environment version. The experimental hardware environment uses RTX 2080 GPU, Intel i7-7700 CPU, 50 GB memory, and Pycharm Community 2021 as development tools. The Max-Margin criterion is used during training. It concentrates on the model's decision boundary's robustness and proposes a probabilistic-based estimating approach as an alternative. To avoid the difficulty of over-fitting during the training of the deep network, the dropout method was used in the training. The relevant training parameters are shown in Table 2:

4.3. Experimental Result. This paper proposes a method of convolutional time-space diagram to recognize Taekwondo athletes' actions. It mainly involves two parts of innovation, namely, the GCN network and part of the perception structure. In order to verify their respective effects, ablation

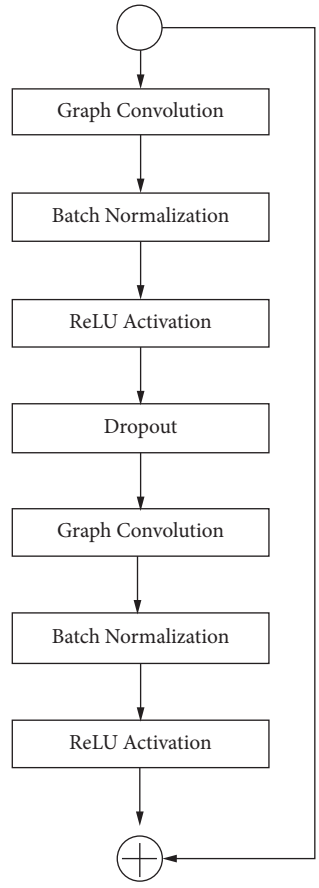


Figure 6: GCN block internal structure.

Table 2: Training parameter settings.

Parameter	Value
Epoch	70
Regularization	0.0001
Dropout rate	0.5
Initial learning rate	0.05
Hidden unit number	200
Weight attenuation coefficient	0.0005
Momentum	0.9

experiments were carried out. First, replace the network structure with a pure GCN network, named after the letter *G*, and build a GCN module performance verification experimental group. Secondly, replace the GCN module in the network with an LSTM structure and name it with the letter *L* to construct an experimental group to verify the performance of the LSTM module. Our method will be verified on a self-made Taekwondo competition data set, and the original GCN network will be compared with our method. From the level of joint recognition accuracy (Joint), bone recognition accuracy (Bone), accuracy (Acc), and parameter number (Param), the experimental results are shown in Table 3.

Table 2 shows that the GCN approach is effective, the overall accuracy is increased by 3.7%, and the number of parameters is also lowered accordingly. Through the partial perception structure method, the overall accuracy is increased by 7.5%, and the portion of parameters is lowered by half. The experimental results show that part of the perceptual structure contributes more to the overall performance improvement. Although the overall performance improvement effect of GCN is not as good as the part of the perceptual structure, it is indispensable at the level of capturing global information. The two methods contrast with each other. The experimental results demonstrate that our proposed GCN method based on partial perception structure is effective.

In order to further verify the effectiveness of this method, this paper compares four different types of bone-based action recognition methods, namely, Dynamic Skeleton [37], P-LSTM [16], TCN [38], and ST-GCN [12]. Among them, Dynamic represents hand-made action recognition methods; TCN represents CNN-based action recognition methods; P-LSTM represents RNN-based action recognition methods; and ST-GCN represents GCN-based action recognition methods. The above 4 methods and the IST-GCN method in this paper are verified on the self-made Taekwondo competition data set. The experimental results are shown in Table 4.

Table 3 shows that the verification experiment of the self-made Taekwondo competition data set yielded positive results, the GCN-based action recognition method is better than other types of action recognition methods, which proves that the graph convolutional network has a huge

Table 3: Results of ablation experiments.

Method	Joint (%)	Bone (%)	Acc (%)	Param (M)
ST-GCN	81.1	81.8	82.1	3.14
G	84.9	85.2	85.8	2.36
L	88.8	89.1	89.6	1.61
Ours	90.9	91.5	92.2	1.36

Table 4: Comparison of different types of action recognition methods.

Method	Action recognition accuracy (%)	Accurate scoring accuracy (%)
Dynamic skeleton	61.2	46.2
P-LSTM	63.9	51.3
TCN	75.3	64.1
ST-GCN	82.5	69.3
Ours	90.7	74.6

Table 5: Taekwondo competition data recognition and analysis results.

Taekwondo action	Usage frequency	Score frequency	Total score	Action recognition accuracy (%)	Score match correct rate (%)
Front kick	1845	25	53	86	69
Cross kick	866	94	186	90	74
Push kick	1566	25	49	86	71
Downward kick	180	9	12	91	77
Double flying kick	102	15	31	92	65
Single leg kick	131	299	33	93	69
Hook kick	21	2	8	88	73
Back kick	55	6	16	86	76
Backspin sick	16	0	0	95	69
Side kick	23	0	0	96	72

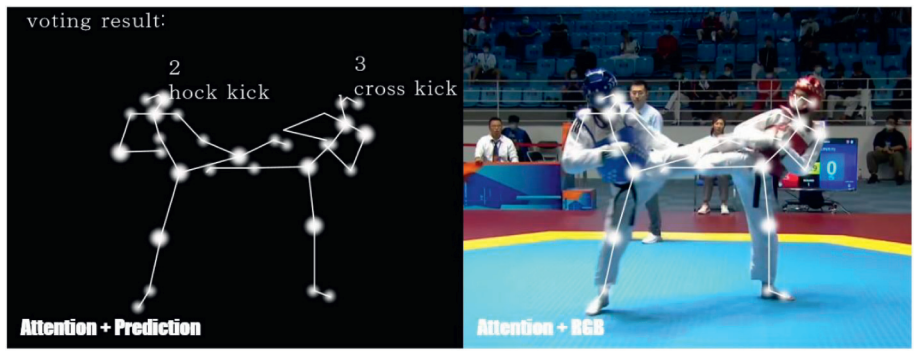


Figure 7: Action recognition and scoring effect.

advantage in the realm of action recognition. The effect of taekwondo action recognition by this method is shown in Figure 7, and the action category and corresponding score are, respectively, marked.

Through our method, all data of taekwondo competitions are identified and decomposed. The most frequently used technical actions in the game include: forward kick, side kick, push kick, down split, back kick, back spin kick, whirlwind kick, side kick, and other techniques. Perform statistical analysis on the identification data and accuracy of each technology, and match the scores. The results are shown in Table 5.

The results of taekwondo action recognition by our method are shown in Table 2. From the above data and analysis, it can be seen that in terms of the technical use of high-level Taekwondo athletes in simulated competitions in the team, the front kick and cross kick techniques are the most commonly used and best skills for athletes, whether it is the technique usage rate or the technical score rate. All of them are among the best. At the same time, rotating technology is more difficult to use and has higher requirements for specific physical fitness, so the utilization rate is generally low. The accuracy of Taekwondo technical action recognition is maintained at 80%, and the score matching

rate is relatively low. Because there are a lot of nonhuman factors in the Taekwondo action, there is a certain error in the score matching rate. In the next study, we will focus on the accuracy of the score matching rate.

5. Conclusion

In this paper, we examine the current state of work and research in the realm of action recognition before moving on to research on Taekwondo action recognition. Taekwondo is a sport, and its action recognition is an important task, which can provide a very valuable reference factor for technicians, athletes, and coaches. But action decomposition and score matching are a difficult point in this technique. This paper proposes a graph convolution framework of part of the perception structure to recognize, decompose, and analyze Taekwondo actions. Taking advantage of the long short-term memory of part of the perception structure, the recognized Taekwondo actions are marked in time series, and then features are extracted from the graph convolution level to obtain the spatial and temporal associations between joints. Predict the action category and perform score matching based on the manual tag database. In general, our method has an average accuracy of 90% in action recognition, and an average action score matching rate of 74.6%.

Because of the low matching rate of Taekwondo scores in this paper, in the future, we hope to use a dual-stream graph convolutional network to process actions, so that the secondary feature sequence of the action can be obtained, to get the second match between the action and the score, and further improve the score matching rate. Deeper picture feature extraction is also a key direction for our future research.

Data Availability

The data set can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Natural Science Foundation of Hunan Province (no. 2020JJ7007).

References

[1] A. Hernandez-Mendo, J. Castellano, O. Camerino, G. Jonsson, Á. Blanco-Villaseñor, and M. T. Antonio Lopes, "Anguera. observational software, data quality control and data analysis," *Revista de Psicología del deporte*, vol. 23, no. 1, pp. 111–121, 2014.

[2] D. G. Liebermann, L. Katz, M. D. Hughes, R. M. Bartlett, J. McClements, and I. M. Franks, "Advances in the application of information technology to sport performance," *Journal of Sports Sciences*, vol. 20, no. 10, pp. 755–769, 2002.

[3] U. Moenig, "La evolución de las técnicas de patada en taekwondo," *Revista de Artes Marciales Asiáticas*, vol. 6, no. 1, pp. 117–140, 2012.

[4] C. González, X. Iglesias, and M. Anguera, "Tactical moves in top level competition taekwondo combat. A descriptive study," in *Proceedings of the Scientific Congress on Martial Arts and Combat Sports*, pp. 48-49, Associaçao para o desenvolvimiento e investigaçao de Viseu Viseu, Portugal, Europe, 2011.

[5] M. Kazemi, C. Casella, and G. Perri, "2004 Olympic tae kwon do athlete profile," *Journal of the Canadian Chiropractic Association*, vol. 53, no. 2, p. 144, 2009.

[6] M. Kazemi, J. Waalen, C. Morgan, and A. R. White, "A profile of Olympic taekwondo competitors," *Journal of Sports Science & Medicine*, vol. 5, p. 114, 2006.

[7] A. Riveiro-Bozada, O. García-García, V. Serrano-Gómez, J. Antonio, L. Lopez, and A. Hernández-Mendo, "Influence the level of competition in technical actions performed point female Shiai Kumite-karate," *Cuadernos de Psicología del Deporte*, vol. 16, no. 1, pp. 51–67, 2016.

[8] M. Pic and G. K. Jonsson, "Professional boxing analysis with T-Patterns," *Physiology & Behavior*, vol. 232, Article ID 113329, 2021.

[9] R. Tarragó, X. Iglesias, D. Lapresa, M. Teresa Anguera, L. Ruiz-Sanchis, and Javier Arana, "Analysis of diachronic relationships in successful and unsuccessful behaviors by world fencing champions using three complementary techniques," *Anales de Psicología*, vol. 33, no. 3, p. 471, 2017.

[10] K. Ito, N. Hirose, M. Nakamura, N. Maekawa, and M. Tamura, "Judo kumi-te pattern and technique effectiveness shifts after the 2013 international judo federation rule revision," *Archives of Budo*, vol. 10, no. 1, 2014.

[11] C. Chen, R. Jafari, and N. Kehtarnavaz, "UTD-MHAD: a multimodal dataset for human action recognition utilizing a depth camera and a wearable inertial sensor," in *Proceedings of the 2015 IEEE International conference on image processing (ICIP)*, pp. 168–172, IEEE, Quebec City, Canada, 2015.

[12] L. Chaolong, C. Zhen, Z. Wenming, C. Xu, and J. Yang, "Spatio-temporal graph convolution for skeleton based action recognition," in *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.

[13] Y. Li, D. Tarlow, M. Brockschmidt, and R. Zemel, "Gated graph sequence neural networks," 2015, https://arxiv.org/abs/1511.05493.

[14] X. Gao, W. Hu, J. Tang, J. Liu, and Z. Guo, "Optimized skeleton-based action recognition via sparsified graph regression," in *Proceedings of the 27th ACM International Conference on Multimedia*, pp. 601–610, 2019.

[15] L. Quevedo Junyent, A. Padrós Blázquez, J. Solé i Fortó, and G. Cardona Torradeflot, "Entrenament perceptivocognitiu amb el Neurotracker 3D-MOT per potenciar el rendiment en tres modalitats esportives," *Apunts Educació Física i Esports*, vol. 119, no. 119, pp. 97–108, 2015.

[16] A. Shahroudy, J. Liu, T. T. Ng, and G. Wang, "Ntu rgb+ d: a large scale dataset for 3d human activity analysis[C]," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, New York, NY, USA, 2016.

[17] S. Song, C. Lan, J. Xing, W. Zeng, and J. Liu, "An end-to-end spatio-temporal attention model for human action recognition from skeleton data[C]," in *Proceedings of the AAAI Conference on Artificial Intelligence, AAAI Press*, Palo Alto, CA, USA, 2017.

[18] F. Monti, D. Boscaini, J. Masci, and E. Rodolà, "Geometric deep learning on graphs and manifolds using mixture model cnns[C]," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, New York, NY, USA, 2017.

[19] S. Yan, Y. Xiong, and D. Lin, "Spatial temporal graph convolutional networks for skeleton-based action recognition," in

Proceedings of the AAAI Conference on Artificial Intelligence, AAAI Press, Palo Alto, CA, USA, 2018.

[20] L. Shi, Y. Zhang, J. Cheng, and H. Lu, "Two-stream adaptive graph convolutional networks for skeleton-based action recognition," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 12026–12035, 2019.

[21] B. Zoph and Q. V. Le, "Neural architecture search with reinforcement learning," 2016, https://arxiv.org/abs/1611.01578.

[22] E. Real, A. Aggarwal, Y. Huang, and Q. V. Le, "Regularized evolution for image classifier architecture search," *Proceedings of the aaai conference on artificial intelligence*, vol. 33, no. 1, pp. 4780–4789, 2019.

[23] H. Liu, K. Simonyan, and Y. Yang, "Darts: differentiable architecture search," 2018, https://arxiv.org/abs/1806.09055.

[24] W. Peng, X. Hong, and G. Zhao, "Video action recognition via neural architecture searching," in *Proceedings of the 2019 IEEE International Conference on Image Processing (ICIP)*, pp. 11–15, IEEE, 2019.

[25] H. Pham, M. Guan, B. Zoph, Q. Le, and J. Dean, "Efficient neural architecture search via parameters sharing," in *Proceedings of the International Conference on Machine Learning*, pp. 4095–4104, PMLR, 2018.

[26] M. Defferrard, X. Bresson, and P. Vandergheynst, "Convolutional neural networks on graphs with fast localized spectral filtering," *Advances in Neural Information Processing Systems*, vol. 29, pp. 3844–3852, 2016.

[27] P. Veličković, G. Cucurull, A. Casanova, A. Romero, L. Pietro, and Y. Bengio, "Graph attention networks," arXiv preprint arXiv:1710.10903, 2017.

[28] D. K. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 129–150, 2011.

[29] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks[J]," arXiv preprint arXiv: 1609.02907, 2016.

[30] J. P. Jones and L. A. Palmer, "An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex," *Journal of Neurophysiology*, vol. 58, no. 6, pp. 1233–1258, 1987.

[31] C. Szegedy, W. Liu, Y. Jia et al., "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1–9, 2015.

[32] M. Lin, Q. Chen, and S. Yan, "Network in network," arXiv preprint arXiv:1312.4400, 2013, 2013.

[33] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[34] O. Vinyals, S. V. Ravuri, and D. Povey, "Revisiting recurrent neural networks for robust ASR[C]," in *Proceedings of the 2012 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pp. 4085–4088, IEEE, 2012.

[35] T. Mikolov, M. Karafiát, L. Burget, J. Cernocký, and S. Khudanpur, "Recurrent neural network based language model," in *Proceedings of the INTERSPEECH 2010, 11th Annual Conference of the International Speech Communication Association*, vol. 2, no. 3, pp. 1045–1048, Makuhari, Japan, 2010.

[36] I. Sutskever, J. Martens, and G. E. Hinton, "Generating text with recurrent neural networks," in *Proceedings of the 28th International Conference on Machine Learning, ICML 2011*, Bellevue, Washington, USA, 2011.

[37] J. F. Hu, W. S. Zheng, J. Lai, and J. Zhang, "Jointly learning heterogeneous features for RGB-D activity recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 5344–5352, 2015.

[38] T. S. Kim and A. Reiter, "Interpretable 3d human action analysis with temporal convolutional networks," in *Proceedings of the 2017 IEEE conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1623–1631, IEEE, Honolulu, HI, USA, 2017.

Hindawi
Scientific Programming
Volume 2022, Article ID 4973875, 10 pages
https://doi.org/10.1155/2022/4973875

Research Article

Research on Innovative Design of Product Packaging Based on Big Data Technology

Fang Gan,[1,2] **Nurul Hanim Romainoor**,[1] **and ZhiMin Guo**[3]

[1]*School of the Arts, Universiti Sains Malaysia, Penang 11800, Malaysia*
[2]*School of the Arts, Guangxi Normal University for Nationalities, Chongzuo, Guangxi 532200, China*
[3]*Nortel Instrument Co., Ltd, Shenzhen 515100, China*

Correspondence should be addressed to Fang Gan; ganfang@gxnun.edu.cn

Received 24 November 2021; Revised 3 December 2021; Accepted 13 December 2021; Published 17 February 2022

Academic Editor: Baiyuan Ding

Traditional product packaging design mainly relies on the designer's personal experience and intuition, but there are problems such as uncontrollable content and lack of knowledge and guidance in the field of product design. With the development of big data technology, product packaging design in the era of big data is carried out under the support of a large amount of real data, which has high predictability, success rate, and short development cycle. Compared with structured data, unstructured data such as text, images, and audio has a higher value. Among them, text big data and image big data have good application prospects in the field of product packaging design due to their easy data acquisition, mature processing technology, and simple operation. This paper proposes a combination of big data technology and neural style transfer model and proposes an innovative design method for product packaging that can generate high-quality images and controllable content. First, the perceptual engineering theory is used to obtain user needs, to build a mapping model between product modeling elements and product semantics, and to guide the selection of product semantics and style maps; second, use neural style transfer models to reconstruct and combine the color features of style maps. After the integration, it migrated to product packaging design, based on big data product packaging innovation design methods, and developed a product innovation design auxiliary prototype system based on actual needs to improve the company's R&D and innovation capabilities, shorten product development cycles, and reduce R&D costs. Improve product success rate and user satisfaction.

1. Introduction

With the advent of the 21st century, the design has truly been integrated into all aspects of our lives. People are carrying out design activities all the time to obtain a more comfortable living environment. Standardized and mechanized production, on the one hand, enables the market citizens to purchase aesthetic products with lower expenditures and increases the quality of life through the quantification of products; on the other hand, due to the impact of economic globalization, the design is consistent. The international style lacks a clear personality. The design of "form follows function" and "equal emphasis on form and function" brings people a wealth of material enjoyment, but it neither bring people's spiritual warmth and considerate comfort nor can it satisfy consumers' demand for individualization and uniqueness. Because of demand, more and more people are buying "works" instead of "commodities." [1–4]. In the context of modern design, the design not only satisfies functions and people's surface visual experience but also needs to convey a certain emotional experience to consumers through design language.

Traditional product appearance innovation design mainly relies on the designer's personal experience and intuition and has defects such as strong subjectivity, weak interpretability, low predictability, long development cycle, slow response to user needs, and low product success rate. The advent of the era of big data has made it possible to improve product design capabilities. Big data contain a lot of valuable information such as product feedback, user preferences, market demand, and visual display. This information plays an important role in guiding product

appearance innovation and design. Use the valuable information in big data to quickly, accurately, and fully capture. Responding to user needs has become the focus and hotspot of domestic and foreign research. This article tries to use big data technology to realize the design of personalized packaging system.

2. Related Work

With the rapid development of the Internet, the Internet of Things, and communication technologies, the amount of global data has exploded, and mankind has officially entered the era of big data. Big data technology describes a new generation of technology and architecture system, through high-speed collection, discovery, or analysis, to extract the economic value of various large amounts of data. The Gartner organization summarized big data as a massive, high growth rate, and diversified information asset that requires new processing models to enhance decision-making, insight, and process optimization capabilities. Academician Xu Zongben defined big data as "incapable of centralized storage, and it is difficult to analyze and process within an acceptable time, in which individual or part of the data presents a low value, while the data as a whole present a massive and complex dataset with a high value." Although the above definitions have different perspectives and focuses, they all reflect that big data are a dataset with the characteristics of large data volume, multiple data types, high data authenticity, fast data generation speed, high value, and low value [5–10].

At present, with the development of artificial intelligence technologies such as machine learning, deep learning, natural language processing, and computer vision, big data have been widely used, such as transportation, health care, fault diagnosis, tourism, and product design, and achieved certain results. Achievement: in the product design, traditional product design mainly relies on the designer's personal experience and intuition. Product design in the era of big data is carried out under the support of a large amount of real data, which has high predictability, success rate, and short development cycle. Compared with structured data, unstructured data such as text, images, and audio have a higher value. Among them, text big data and image big data have good application prospects in the field of product design due to their easy data acquisition, mature processing technology, and simple operation. Specifically, in terms of data research, text big data can make up for the limitations of traditional research methods in the number of surveys, geographical scope, time interval, and timeliness of data, ensuring the authenticity, reliability, and timeliness of survey information; in products, in terms of design plan visualization, image big data can directly generate product images, providing users and designers with intuitive visual displays. The current product design is developing in the direction of intelligence, systematization, and short development cycle. Making full use of the value contained in big data and introducing cutting-edge technologies such as machine learning and deep learning are the hotspots and difficulties of current product innovation design.

3. Related Theoretical Methods

3.1. Big Data Technology. With the development of social economy, the levels of user demand for products will become more diversified and individualized. Researchers must dare to break through the shackles of inherent concepts and break conventions to find new ideas and new ways of the product design. Big data contain a lot of valuable information, which is of great significance to product innovation and design. Figure 1 shows the life cycle data of the entire big data product [11–14].

3.2. Neural Style Transfer. Batch normalization (BN), IN, CIN, and AdaIN are commonly used methods in neural style transfer. BN calculates the characteristic data (mean and variance) of each channel of a batch of samples, while IN calculates the characteristic data of each channel and sample independently. CIN performs style transfer by panning and zooming the normalized result. Each style needs to learn two parameters that control panning and zooming. AdaIN is obtained by improving CIN, and AdaIN calculates the control parameters adaptively from the feature representation of the image. Figure 2 shows the neural style transfer structure, using the "encoder-AdaIN-decoder" architecture, which mainly includes three parts: encoder, AdaIN, and decoder [15–17].

The encoder part adopts the VGG-19 network pretrained on the ImageNet dataset, which can detect the features in the image. The AdaIN layer normalizes the content map and matches the mean and variance of each channel feature map in the style map by aligning the mean and variance of each channel feature map in the content map:

$$\begin{aligned} t &= \mathrm{AdaIN}(fc, fs) \\ &= \sigma(fs)\left[\frac{fc - \mu fc}{\sigma(f(c))}\right] + \mu(f(s)), \end{aligned} \tag{1}$$

where f represents the encoder, c and s represent the content map and the style map, respectively, and μ and σ represent the mean and variance, respectively. The decoder part converts the feature space to the image space to generate a stylized image $T(c, s)$:

$$T(c, s) = \mathrm{g}(t), \tag{2}$$

where $g$represents the decoder.

The decoder adopts a network structure symmetrical to the encoder, and what needs to be trained in the entire network is the weight parameter of the decoder network. The loss function of the network is mainly composed of content loss and style loss. The training goal is to minimize the loss function [18–20]:

$$L = L_c + L_s, \tag{3}$$

where L is the loss function and L_c and L_s are content loss and style loss, respectively. The content loss is the Euclidean distance (L_2 norm) between the target feature and the input image feature:

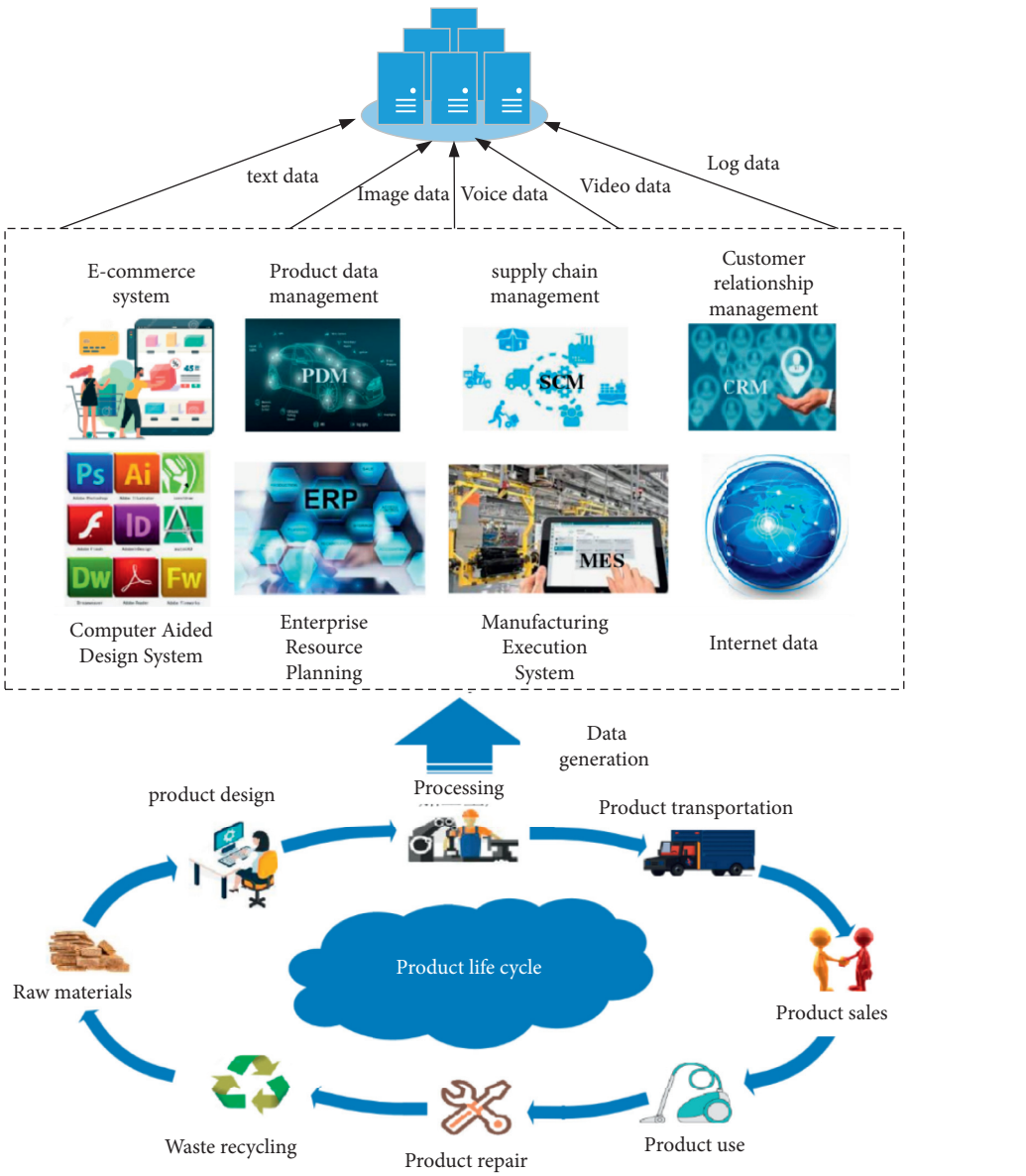


Figure 1: Product life cycle data.

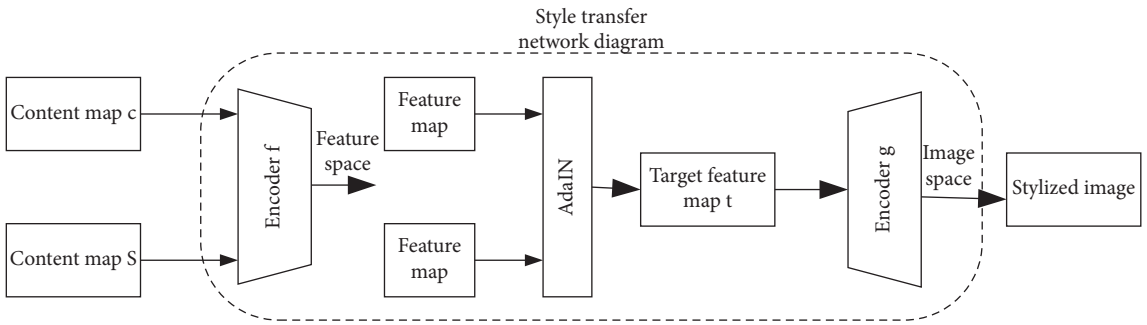


Figure 2: Neural style transfer.

$$L_c = fg(t) - t_2. \quad (4)$$

Since the AdaIN layer transmits the mean and variance of the style features, the style loss can be obtained by matching the mean and variance:

$$L_s = \sum_{i=1}^{K} \mu(\varnothing_i g(t)) - \mu(\varnothing_i s)_2 + \sum_{i=1}^{K} \sigma(\varnothing_i g(t)) - \mu(\varnothing_i s)_2, \quad (5)$$

where ϕ represents the pretrained VGG-19 model, K represents the number of convolutional layers of the model, and $\phi_i(s)$ represents the activation value of the style map in the ith layer of the model ϕ.

3.3. Semantics of Product Packaging Design Style Diagram. Sensibility is the feeling that people produce after being stimulated by the outside world. Human sensibility has the attributes of vague and polysemous, including vision, hearing, touch, smell, and other aspects. It is a kind of recognition and comprehensive judgment of the object by the person himself. Perceptual engineering is a technology that combines perceptual and engineering. It aims to link people's emotional response with the design characteristics of products and attempts to transform emotions into measurable design elements. It is a user-centered design method. The process of perceptual engineering applied to product design is shown in Figure 3. The basic idea is to describe products from two different perspectives, product semantics, and product attributes and establish a mapping model between the two in the synthesis stage. Kansei engineering mainly includes four steps: product field selection, product semantic space establishment, product attribute space establishment, and mapping model construction [21].

Color perception is the psychological perception that light passes through the lens to stimulate the retina and is transmitted to the visual center of the brain through nerve fibers. It reflects the emotion and preference of people for color. Studies have shown that, in people's psychological reactions to objects, the influence of color in the first 5 minutes accounts for 50–80%, and after 5 minutes, it remains unchanged at 50%. It can be seen that color has a greater impact on human sensory characteristics. In the field of image color research, Ishihara et al. tried to use self-organizing neural networks to construct a perceptual engineering expert system for monochrome images. Hsiao applies fuzzy set theory to product color planning, but only considers the evaluation of monochromatic images. Tsai and Chou proposed a two-color emotional model based on gray theory. However, when the product contains multiple components of different colors, predicting the evaluation of the entire image is much more complicated. Chen Lili et al. used perceptual engineering methods to analyze the color and volume of different wavelengths ("reduced-enlarged"), distance ("highly cold-close"), soft and hard feeling ("soft-hard"), sense of lightness ("light-heavy"), sense of warmth ("passionate-high cold"), mood ("relaxed-depressive"), personality ("exaggerated-low-key"), and texture ("Gorgeous-Plain") The correlation of the seven color semantic dimensions, as well as the relationship between sensory characteristics and stimulus response time, and the results are shown in Table 1. The results in Table 1 show that the long-wavelength colors give people a strong sense of intimacy, publicity, and gorgeousness, while the short-wavelength colors are high-cold, low-key, and simple. This section uses the above research conclusions to evaluate the semantics of product packaging design style diagrams. In the selection process of product packaging style images, images composed of colors with similar wavelengths are consciously selected as the style map [22].

4. Construction of Aided Design System for Product Personalized Packaging Driven by Big Data

4.1. Data Processing. Big data have rich data types, including voice data, video data, image data, text data, geographic location data, and numerical data. These data can be divided into three types: structured, semistructured, and unstructured data. Among them, structured data are composed of defined data types, usually stored in relational databases (RDBMS), and have the characteristics of easy retrieval; the structure and content of semistructured data are mixed storage, which makes it strong flexibility. Common semistructured data include webpages and reports; unstructured data have the characteristics of large volume, high growth rate, and high value and are difficult to retrieve. Common unstructured data include text, images, and audio. Compared with the first two types of data, unstructured data processing is more difficult, but it contains a high value [23].

The purpose of data preprocessing is to make the original data meet the minimum specifications and standards for data analysis. The data generated during the product packaging design life cycle contain a large amount of data, which will bring difficulties and deviations to data analysis and require preprocessing. Data preprocessing mainly deals with missing data, abnormal data, duplicate data, logical error data, and inconsistent data; data integration is to merge and store data obtained from multiple data sources, and it is necessary to resolve multiple entity name recognition, redundancy, and data conflicts. Data transformation is to normalize, discretize, and sparse data to make the data suitable for data analysis. Normalization is to solve the dimensional difference between different features; data specification is a way to maintain the integrity of the original data. The operation of reducing the data scale under the premise can improve the efficiency of data analysis.

Big data have the characteristics of low value density and high commercial value. How to dig out valuable information from large-scale data is the core part of the big data processing process. The process of using analytical techniques to reveal valuable rules and results from data is called data analysis. Commonly used big data analysis techniques include machine learning, deep learning, natural language processing, and statistical analysis. Many researchers have integrated big data analysis technology to build a data analysis work platform. As shown in Table 2, commonly used working platforms include WEKA, Mahout, RapidMiner, and KNIME.

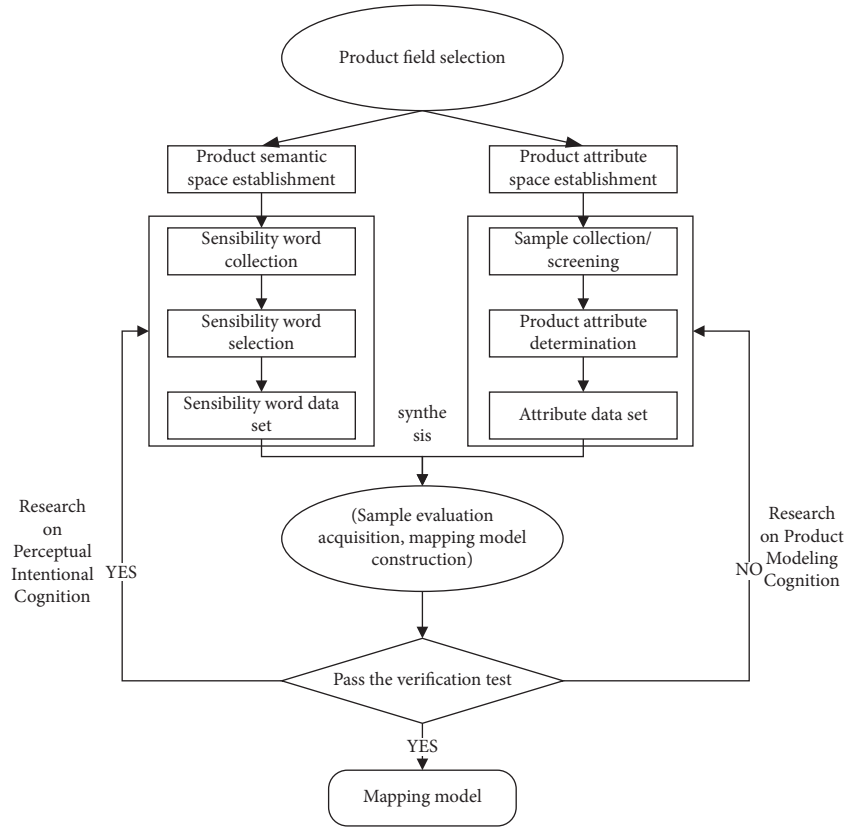


Figure 3: Kansei engineering process.

Table 1: Color perception grade.

	Volume sense reduced-enlarged	Sense of distance cold-close	Soft and hard soft-stiff	Soft and hard soft-stiff	Soft and hard soft-stiff	Mood relaxed-depressive	Character public-low key	Texture gorgeous-plain
Red	2.167	1.778	1.111	−2.611	1.306	1.139	−2.583	−2.417
Orange	1.806	1.944	−1.000	−0.889	−2.028	−0.639	−2.083	−1.056
Yellow	2.028	2.000	−1.361	−1.500	−1.500	−0.194	−2.194	−1.000
Green	−1.194	−0.778	0.556	0.306	0.722	−1.500	0.889	0.278
Blue	0.056	−0.750	−1.639	−2.194	0.444	−2.333	0.861	0.833
Blue	−1.028	−1.500	−0.833	-0.306	1.167	−1.389	−0.333	−0.139
Purple	−1.056	−1.389	1.194	1.056	1.083	1.056	−0.306	−2.000
Pink	0.861	2.278	−2.361	−2.056	−0.611	−1.083	−1.056	−0.028
White	1.722	0.528	−1.222	−2.556	0.944	−1.556	1.222	1.778
Black	−2.139	−2.333	2.778	2.889	2.278	2.611	2.833	2.472
Gray	−1.306	−1.417	0.472	0.861	2.028	1.056	2.194	1.806

4.2. System Design

4.2.1. Architecture Design. (1) System architecture selection: typical distributed system architectures include client/server (Client/Server, C/S) and browser/server (Browser/Server, B/S), two types; the structure of the two is shown in Figure 4.

The biggest difference between the C/S structure and the B/S structure is whether the application logic layer is independent. The application logic layer contains the algorithms and calculation processes required by the system functions. The C/S structure includes both the presentation layer and the application logic layer on the client side, while the B/S structure adds a WEB server, which is independent of the application logic layer. In the B/S structure, all users can share the WEB server, and the client is only responsible for receiving user requests and displaying the response results, which reduces the client's requirements for computer configuration [24].

The C/S structure and the B/S structure are also quite different on the client. The C/S structure requires special software to be installed on the client, while the B/S structure can be accessed using a standard browser. The differences in the client side make the C/S structure more complicated in terms of system upgrades, installation, maintenance, and

Table 2: Data analysis work platform.

Work platform	Advantage	Shortcoming
WEKA	① The code is open source and can be used for secondary development ② Good effect of low-dimensional data	① The effect of high-dimensional data is poor ② Limited algorithm library
Mahout	① Can be integrated with other big data platforms ② Based on Hadoop implementation, suitable for processing large amounts of data	① No graphical operation mode ② Use interface development
RapidMiner	① Fast integration ② Support multiple data resource access methods	① Poor parallelism
KNIME	① Simple operation process and intuitive results ② Complete data analysis methods	① The integration is difficult ② Poor support for statistical models

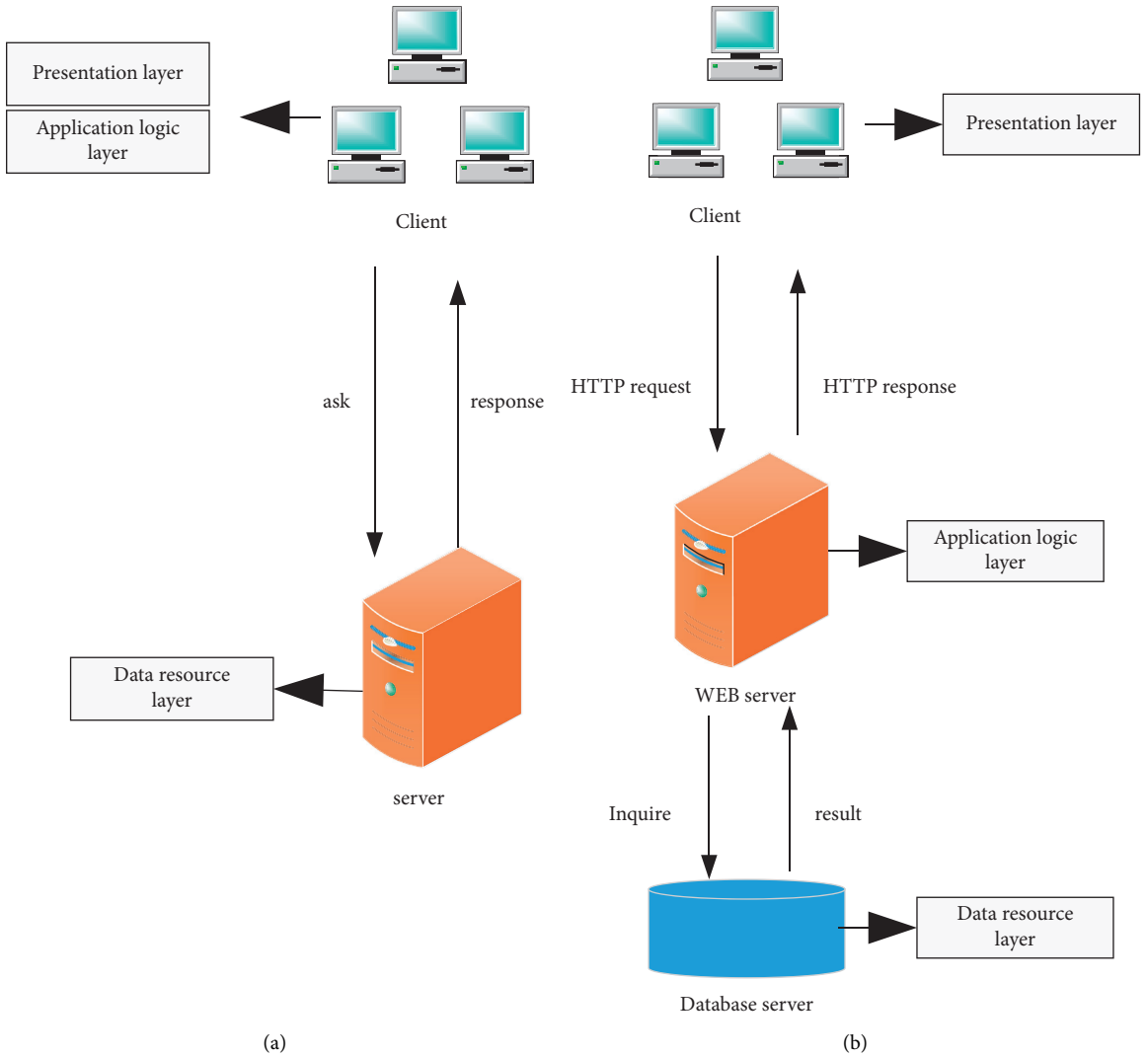


Figure 4: C/S and B/S structure.

user training. Considering that the prototype system developed in this section faces a large number of users, and the functions of the system will be further enriched as the development progresses, a B/S structure with low technical requirements for users, low client computer configuration requirements, and easy update and maintenance is selected.

4.2.2. Database Design. *(1)* Choice of database: a database is a collection of data stored together in a certain way, which has the characteristics of low redundancy, easy expansion, high sharing, and strong independence. Users can add, delete, modify, and query data in the database. Databases can be divided into relational databases and nonrelational

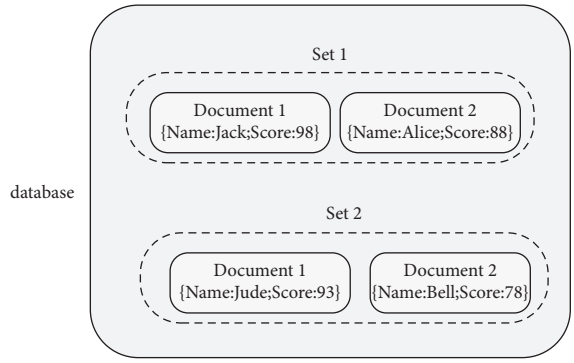


Figure 5: MongoDB database.

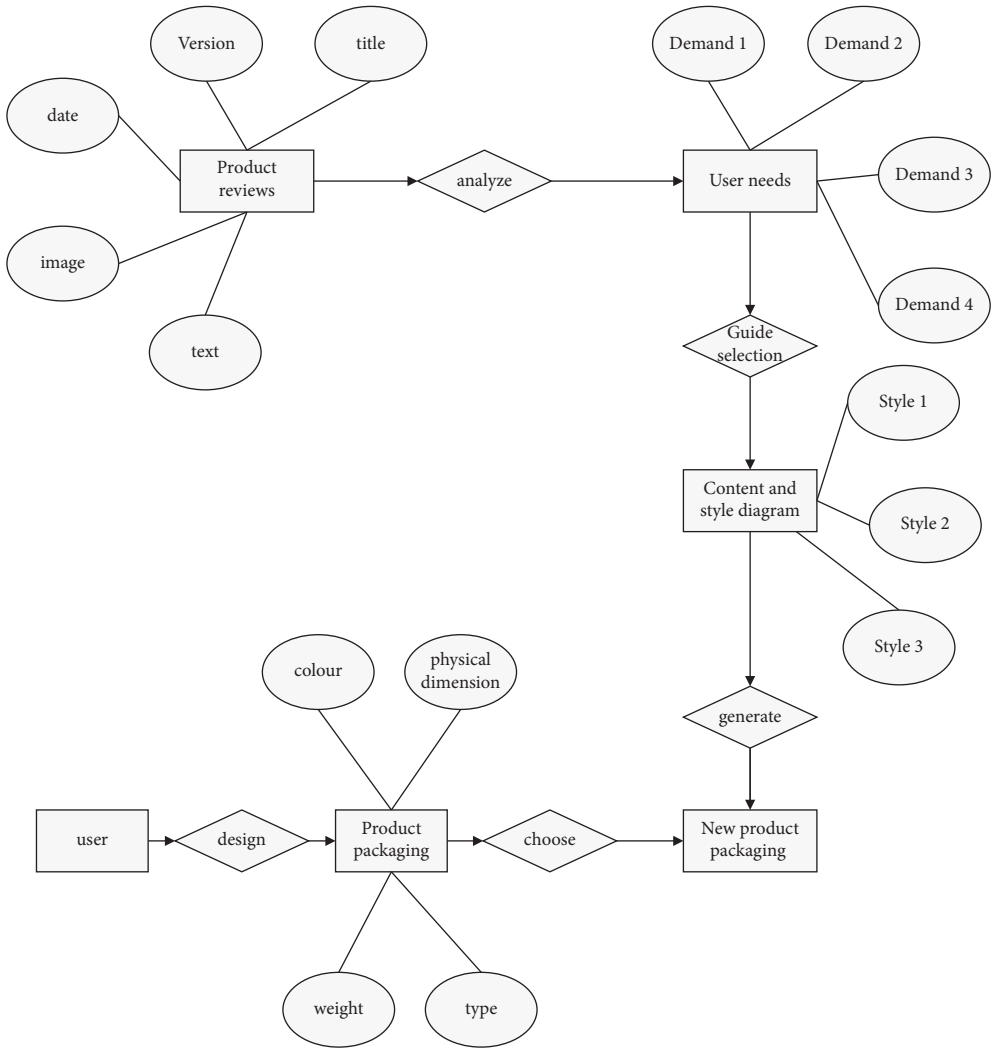


Figure 6: Big data-driven product innovation design auxiliary prototype system.

databases. Relational databases use a two-dimensional table model to organize data. They have the characteristics of consistent data format and easy management and operation. The mainstream relational databases include DB2, Microsoft SQL Server, Microsoft Access, and MySQL. Nonrelational database is an extension and extension of relational data. Its storage format is very flexible, including document format, key-value format, and image format. Due to the complex data types involved in the prototype system to be developed, this section uses a nonrelational database for storage. MongoDB is a commonly used nonrelational database with high performance, easy deployment, easy use, and scalability. As shown in Figure 5, MongoDB consists of three parts: document, collection, and database. Key-value pairs

form a set of documents; multiple records composed of multiple sets of documents are called collections, and multiple collections form a database.

(2) Data structure design: according to the functional requirements of the big data-driven product innovation design auxiliary prototype system, the database has designed product online comment information (Comment_information), product attribute information (Attributes_information), content image information (Content_image), style image information (Style_image) and objective weight values,iInformation (Objective_weight), and other collections. The entity-relation diagram (ER) of the prototype system database is shown in Figure 6.

Comment_information is used to store product comment information, including product category, product image, product version, comment date, and comment text. The MongoDB database has two methods for storing image data. The first is to convert the image into a binary format and store it as a dictionary key-value pair; the second is to use the MongoDB submodule GridFS for storage. This section uses the first method to store image data. The detailed field descriptions are shown in Table 3.

Style_image and Content_image are used to store style image and content image, respectively. The data structure of the two is the same, including image number and image, as well as the score value of the image in the system's predefined style type. The detailed fields are shown in Table 4.

Objective_weight is used to store the objective weight of the product. The subjective weight of the product is input by the user through the system interface, and the two together form a comprehensive weight value through the game theory method to rank the alternatives. Objective_weight includes the product category and the weight value of the product on several criteria predefined by the system. The detailed fields are shown in Table 5.

4.2.3. Development and Operating Environment. The overall operating framework of the entire big data-driven product innovation design auxiliary prototype system adopts the B/S mode, uses the *Python* language to complete the development, and uses the MongoDB database to store data. The software configuration in the system development process is shown in Table 6.

4.2.4. System Implementation. (1) System login: users access the big data-driven product innovation design auxiliary prototype system through a standard browser. Tourist users need to register before they can log in and access the system. The login interface of the system is shown in Figure 7. The registered user enters the system homepage interface, as shown in Figure 8, after logging in. In addition to the core functions of the system, system announcements and some application cases are also displayed on the system homepage. System announcements facilitate users to keep abreast of system update functions and notifications of projects that users participate in; the application case part can help users better understand system functions.

Table 3: Comment_information data structure.

Serial	Number data name	Data type	Data description
1	Product_category	String	Product name
2	ID_id	Number	Product ID
3	Product_image	Binary date	Product image
4	Title	String	Title
5	Product_version	String	Product version
6	Comment_data	Date	Comment date
7	Comment_user	String	Comment user
8	Comment_test	String	Comment text

Table 4: Style_image data structure.

Serial	Number data name	Data type	Data description
1	Image_number	Number	Image number
2	Image	Binary date	Image
3	Style1_score	Number	Style 1 point
4	Style2_score	Number	Style 2 point
5	Style3_score	Number	Style 3 point
6	Style4_score	Number	Style 4 point
7	Style5_score	Number	Style 5 point
8	Style6_score	Number	Style 6 point

Table 5: Objective_Weight data structure.

Serial	Number data name	Data type	Data description
1	Product_category	String	Product category
2	Criterion1_weight	Number	Criterion 1 weight value
3	Criterion2_weight	Number	Criterion 2 weight value
4	Criterion3_weight	Number	Criterion 3 weight value
5	Criterion4_weight	Number	Criterion 4 weight value
6	Criterion5_weight	Number	Criterion 5 weight value
7	Criterion6_weight	Number	Criterion 6 weight value

Table 6: Development environment software configuration.

Name	Tool
Operating system	Ubuntu 18.04
Development environment	P *y* charm professional version:2018.1
Development language	*Python*,XML
Database	MongoDB
WEB server	Tomcat, nginx
Access mode	B/S

4.3. Analysis of Test Results. Through the construction of the personalized packaging system established in this section of this article and the trained style transfer model, the content and style features in the content graph and style graph can be extracted and integrated separately through the style transfer model. The image generation module of the new product includes three parts: content map selection, style map selection, and image generation. As shown in Figure 8, the "Content Map" and "Style Map" columns, respectively, display the content map and style map selected by the user. After clicking the "Generate" button, the generation interface will automatically generate a new product image.

Figure 7: System login interface.

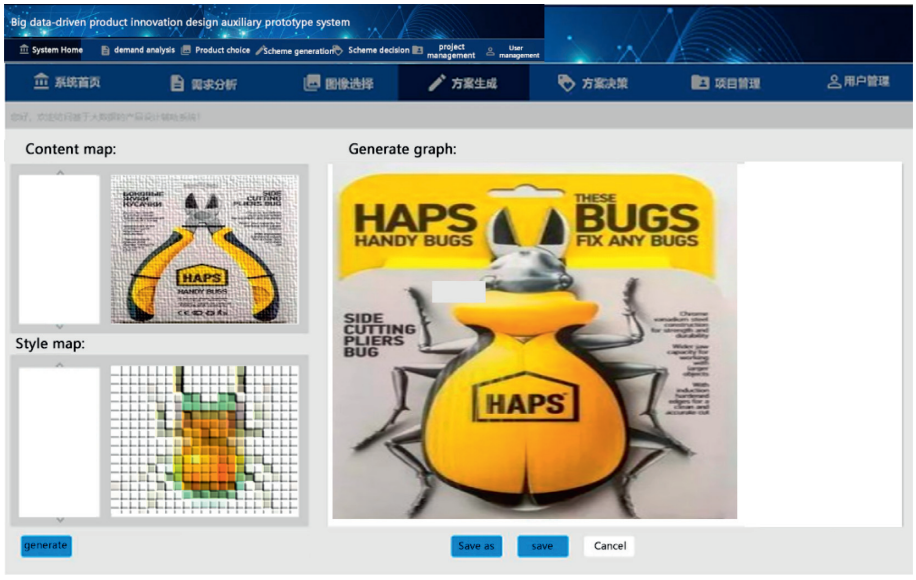


Figure 8: New product packaging design.

The generated new product image has both the product shape in the content map and the color characteristics of the style map. In addition, the plan generation interface also allows all users to upload local images for image generation, but the generated images cannot be stored in the system and can only be exported as jpg format and saved locally.

5. Conclusion

This study explores how to obtain information related to the product design from the big data generated in the product life cycle and uses machine learning, deep learning, and natural language processing technologies to conduct in-depth analysis to further form intelligent and systematic products' innovative packaging design methods; the main research work is as follows. A set of big data-driven product innovation packaging design auxiliary prototype system is developed. First, it analyzes the design-related data generated in the product life cycle and the functional requirements of the system; secondly, the technical framework is designed according to the functional requirements; finally, the two core functions of the system, such as new product generation and product design plan decision-making, are shown in interface.

The innovative design of product packaging in the new era involves the intersection of multiple disciplines, including industrial design, computer, and psychology. This study introduces theories of big data, machine learning, deep learning, and natural language processing technology into the field of product design and proposes intelligence. The

innovative design method of the product has been verified by experiments. The following aspects will be perfected and improved in the next research work. This study only studies two kinds of big data, image and text, and the behavior big data generated by operations such as clicking and browsing is also of great significance for obtaining user needs. In the follow-up research, we will explore the application of more different types of big data in product innovation design.

Data Availability

The data used to support the findings of the study are available from the corresponding upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] C. R. Wilkinson and A. De Angeli, "Applying user centred and participatory design approaches to commercial product development," *Design Studies*, vol. 35, no. 6, pp. 614–631, 2014.

[2] A. Burnap, Y. Liu, Y. Pan, H. Lee, R. Gonzalez, and P. Y. Papalambros, "Estimating and exploring the product form design space using deep generative models," in *Proceedings of the ASME 2016 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, vol. 1–13, Charlotte, NC, USA, August 2016.

[3] D. V. Horn, A. Olewnik, and K. Lewis, "Design analytics: capturing, understanding, and meeting customer needs using big data," in *Proceedings of the Asme International Design Technical Conferences*, pp. 863–875, Chicago, IL, USA, August 2012.

[4] M. Saifullah, McC. Rebecca, McC. Adam, and V. V. Quan, "Effect of drying techniques and operating conditions on the retention of color, phenolics, and antioxidant properties in dried lemon scented tea tree (Leptospermum petersonii) leaves," *Journal of Food Processing and Preservation*, vol. 45, no. 3, 2021.

[5] R. Ireland and A. Liu, "Application of data analytics for product design: s," *CIRP Journal of Manufacturing Science and Technology*, vol. 23, pp. 128–144, 2018.

[6] H. Ah-Reum, B. Nam, B. R. Kim et al., "Phytochemical composition and antioxidant activities of two different color Chrysanthemum flower teas," *Molecules*, vol. 24, no. 2, 2019.

[7] S. Vinodh and K. J. Manjunatheshwara, "Application of fuzzy QFD for environmentally conscious design of mobile phones," *Green and Lean Management*, Springers, Berlin, Germanny, pp. 149–160, 2017.

[8] K. Chakraborty, S. Mondal, and K. Mukherjee, "Analysis of product design characteristics for remanufacturing using Fuzzy AHP and Axiomatic Design," *Journal of Engineering Design*, vol. 28, no. 5, pp. 338–368, 2017.

[9] S. Tripathi, L. L. Henrekin, C. D. Read, and K. F. Welke, "Identification of critical to quality elements for intensive care rounds by kano analysis," *Pediatric Quality And Safety*, vol. 2, no. 4, 2017.

[10] R. M. Wong and O. O. Adesope, "Meta-analysis of emotional designs in multimedia learning: a replication and extension study," *Educational Psychology Review*, vol. 33, pp. 1–29, 2020.

[11] A. G. María, V. Pardo, R. Lucía, and M. N. Daniel, "Do products respond to user desires? A case study. Errors and successes in the design Process,under the umbrella of emotional design," *Symmetry*, vol. 12, no. 8, 2020.

[12] M. Antoine, A. Marie-France, and B. Éric, "Ordinary users, precursory users and experts in the anticipation of future needs: evaluation of their contribution in the elaboration of new needs in energy for housing," *Applied Ergonomics*, vol. 94, 2021.

[13] G. Büyüközkan and F. Göçer, "Application of a new combined intuitionistic fuzzy MCDM approach based on axiomatic design methodology for the supplier selection problem," *Applied Soft Computing*, vol. 52, pp. 1222–1238, 2017.

[14] D. Mourtzis, E. Vlachou, and N. Milas, "Industrial big data as a result of IoT adoption in manufacturing," *Procedia Cirp*, vol. 55, pp. 290–295, 2016.

[15] M. Kopel, "Analyzing music metadata on artist influence," in *Proceedings of the Intelligent Information & Database Systems: Asian Conference*, Bali, Indonesia, March 2015.

[16] H. Kim, "Towards a sales assistant using a product knowledge graph," *Web Semantics Science Services and Agents on the World Wide Web*, vol. 46, pp. 14–19, 2017.

[17] J. Ahn, J.-M. Park, L. Won-Ho, and G.-Y. Noh, "Website interactivity and processing: menu customization and sense of agency are keys to better interaction design," *International Journal of Human-Computer Studies*, vol. 147, 2021.

[18] D. Mourtzis, M. Doukas, and C. Vandera, "An ensemble deep convolutional neural network model with improved DS evidence fusion for bearing fault diagnosis," *Sensors*, vol. 17, no. 8, p. 1729, 2017.

[19] R. Estrada and I. Ruiz, *Big Data SMACK: A Guide to Apache Spark, Mesos, Akka, Cassandra, and Kafka*, Apress, New York, NY, USA, 2020.

[20] A. Kangale, S. K. Kumar, M. A. Naeem, M. Williams, and M. K. Tiwari, "Mining consumer reviews to generate ratings of different product attributes while producing feature-based review-summary," *International Journal of Systems Science*, vol. 47, no. 13, pp. 3272–3286, 2016.

[21] Pew Research Center, "Online reviews," 2019.

[22] E. Ilbahar and S. Cebi, "Classification of design parameters for E-commerce websites: a novel fuzzy Kano approach," *Telematics and Informatics*, vol. 34, no. 8, pp. 1814–1825, 2017.

[23] J. Vieira, J. M. A. Osório, S. Mouta et al., "Kansei engineering as a tool for the design of in-vehicle rubber keypads," *Applied Ergonomics*, vol. 61, pp. 1–11, 2017.

[24] M. Nagamachi, "Successful points of kansei product development," in *Proceedings of the KEER2018, Go Green with Emotion. 7th International Conference on Kansei Engineering & Emotion Research*, vol. 19-22, no. 146, pp. 177–187, Kuching, Malaysia, March 2018.

Hindawi
Scientific Programming
Volume 2022, Article ID 8750580, 12 pages
https://doi.org/10.1155/2022/8750580

Research Article

Research on Fresh Product Logistics Transportation Scheduling Based on Deep Reinforcement Learning

Hongshen Yu

Changchun University of Finance and Economics, Changchun 130177, Jilin, China

Correspondence should be addressed to Hongshen Yu; yuhongs301@ccufe.edu.cn

Received 24 November 2021; Revised 16 December 2021; Accepted 23 December 2021; Published 14 February 2022

Academic Editor: Baiyuan Ding

With the improvement of the economic level, people's quality of life continues to improve, the demand for fresh food is increasing, and the logistics of fresh products is also developing rapidly. We effectively balance the relationship between transportation costs and service levels in fresh product logistics and transportation businesses, improve the transportation capacity and efficiency of logistics transportation businesses, and improve the resource utilization of businesses. It is important for the development of the logistics transportation scheduling industry for fresh products. *Significance*. Based on this, this paper proposes a DNQ algorithm based on pointer network, which solves the single fresh product distribution service center-regional efficient logistics scheduling problem, and a feasible logistics transportation scheduling scheme can be obtained through simulation experiments. The simulation results show that the algorithm is superior to other common intelligent algorithms in terms of accuracy and stability, which proves that the algorithm is effective and feasible (the research results cannot be directly shown in the abstract and need to be supplemented) At the same time, it further explored the DNQ algorithm to improve the correction network, which can solve the complex problem of multiple fresh product distribution service centers-regional efficient logistics scheduling. It is a successful attempt to improve the solution algorithm. Complex logistics and transportation scheduling problems provide ideas and have good guidance and reference significance.

1. Introduction

With the further improvement of residents' living standards, the scale of online and offline demand in the fresh product market has gradually expanded. How to achieve accurate delivery of customer needs has put forward higher requirements for the efficient transportation of fresh product logistics. The logistics and transportation of fresh products need to be comprehensively considered from the aspects of precooling and refrigeration technology, fresh product logistics links, etc.; for the huge fresh market demand, the continuous improvement of the logistics distribution optimization problem of fresh products becomes the problem especially important. According to the latest research, the transportation cost of fresh product logistics is 40%–60% higher than the logistics of ordinary goods, and the cost incurred is on the rise. How to plan vehicle scheduling scientifically and reasonably plays a very important role in reducing distribution costs and total logistics costs, as well as providing better services to consumers. For the optimization of the logistics and distribution vehicle scheduling of agricultural products and fresh products, it is necessary to apply modern information technology to carry out real-time and accurate positioning of agricultural products cold-chain transportation vehicles, driving data collection, customer information collection and analysis, agricultural product transportation volume analysis, and distribution route planning to achieve distribution. Intelligent and efficient vehicle scheduling can meet consumer demand at the lowest cost and guarantee the quality of agricultural products. This article summarizes the life scenarios of residents, the distribution network of fresh products is scattered, the production and sales are separated, the distribution path is too long, the transportation capacity is wasted, the vehicle loading rate is too low, and the transportation efficiency is low. Research on the efficient logistics transportation

scheduling problem of the distribution center has certain guidance and reference significance for solving the complex logistics transportation scheduling problem of fresh products.

2. Related Work

The research content of fresh product logistics and transportation mainly focuses on the development status and countermeasures of fresh products, safety management, route optimization, precooling and cold storage technology, and optimization of fresh product logistics links. A brief introduction to the research contents of fresh product logistics and transportation in recent years is now given.

In terms of qualitative research, Macheka et al. [1] pointed out that the core technology and core equipment of fresh product logistics are backward, the cold chain supply chain is severely disconnected, and smooth operations cannot be formed. Fresh product logistics lacks systematic laws and regulations and authoritative logistics standards and other internal development contradictions. Musavi and Bozorgi-Amiri [2] put forward that my country's fresh food logistics is facing problems such as reprocessing and packaging and last-mile delivery. Neves-Moreira et al. [3] used the fuzzy analytic hierarchy process to analyze the weights of IS indicators and summarized three indicators that have the greatest impact on the development of agricultural and fresh products logistics. Rahimi et al. [4] found out the restrictive factors restricting the development of my country's agricultural and fresh products logistics from the five perspectives of policy, personnel, hardware, software, and management. Lei [5] analyzed the development conditions of my country's fresh agricultural products and fresh product logistics from economic factors and social factors. Samuel Mercier [6] summarized and analyzed the current situation of the food cold chain in Canada and believed that the main problems of the cold chain include seasonal outbreaks of food waste and food poisoning, cold chain transportation in summer, and products produced in severe cold areas, freezing damage, and other issues. The research of Xin et al. [7–9] focused on the technical improvement of the fresh-keeping link in the logistics of fresh food products. New Zealand scholar James K. Carson [10] focused on how to optimize the preservation process to extend the life cycle of products on the shelf. In the case study, he introduced how to improve the carton packaging and stacking methods of kiwifruit to extend the life of the product. Marlies de Keizer [11] believes that the different spoilage rates of different types of products will have an impact on the effect of the fresh agricultural product logistics network. He believes that the product delivery cycle and the different spoilage rates of products should be considered as factors in the logistics network model.

In terms of quantitative research, the research on the logistics of fresh agricultural products and fresh products has a wide range of research. Sun Zhidan and others used cost-saving methods and human ant colony algorithms to solve the path optimization problem of fresh agricultural products and fresh products logistics distribution [12–14]. Ge Changfei et al. [15] optimized the existing logistics and transportation model of fresh food products from the perspective of product quality and order quantity. Huang Chunhui's research [16] focuses on the emergency complementation of intersupplier inventory in the food cold chain and the optimization of the entire fresh product logistics and transportation network. Zhang Wenfeng [17] established an optimization model for the layout of fresh product logistics outlets with the goal of outlet operating costs and construction costs and used a particle swarm algorithm to solve the model. Xiunian Zhan [18] established a value-based decision model. In order to improve the efficiency of the cold chain of fresh products, Soto-Silva et al. [19] designed three models to improve the efficiency of the cold chain of fresh products: fresh product purchasing model, fresh product storage model, and the combination of purchasing, transportation, and warehousing. The model was finally validated with a Chilean dried apple processing plant. Ghezavati et al. [20–22] established an optimization model for the supply chain distribution network of fresh agricultural products from the place of production to the customer with the maximum benefit as the objective function and verified the effectiveness of the model. Marco Bortolini and others created a three-objective distribution planning model to optimize the logistics distribution path of fresh products with the minimum cost as the goal [22, 23].In solving the scheduling problem model, Zhang Jianfeng and others used a deep recurrent neural network model embedded with a pointer network to solve the job shop scheduling problem.

3. Related Theories

In reinforcement learning, modeling strategy $\pi \vee (a \vee s)$ and value function $V^{\pi}(s)$ and $Q^{\pi}(s, a)$ are generally required. Early reinforcement learning algorithms mainly focused on the discrete and limited problems of states and actions, and tables can be used to record these probabilities. But in many practical problems, the number of states and actions of some tasks is very large. In order to effectively solve these problems, in order to be able to design a stronger strategy function, so that the agent can deal with complex environments, learn better strategies, and have better generalization capabilities, the pointer network model is used.

3.1. Pointer Network Model. Compared with the traditional local search algorithm, the pointer network model has two obvious advantages. First, compared with the traditional local search algorithm, the pointer network model responds faster when entering new data. In the traditional heuristic algorithm, when a new case is entered, it needs to be recalculated without any experience. Second, the output of the pointer network model is related to the length of the input but has nothing to do with the dictionary. The versatility of the model is greatly improved. The same model is used in the same type of problem, saving the trouble of training the model in each case.

Pointer Networks (Ptr-Nets) are composed of an encoder (Encoder), a decoder (Decoder), and a shallow neural network. The encoder is a two-way long and short-term memory network, and the decoder is a single Xiang's long and short-term memory network [24]. The structure of the pointer network is shown in Figure 1. [24].

Sequence to sequence is a conditional sequence generation problem. Given a sequence $x_{1:s}$, generate another sequence $y_{1:T}$. The length S of the input sequence and the length of the output sequence can be different. The goal of the sequence-to-sequence model is to estimate the conditional probability as

$$p_\theta(y_{1:T} \vee x_{1:S}) = \prod_{t=1}^{T}\left(y\middle|t \vee y_{1:(t-1),x_{1:S}}\right), \tag{1}$$

where is an element in the set V. Given a set of training data, the maximum likelihood estimation can be used to train the model parameters.

$$\max_\theta \sum_{n=1}^{N} \log p_\theta\left(y_{1:T_n} \vee x_{1:S_n}\right). \tag{2}$$

After the training is completed, the model can generate the most likely target sequence based on an input sequence X,

$$\hat{y} = \arg\max_y p_\theta(y \vee x). \tag{3}$$

The specific generation process can be completed by the greedy method or beam search.

Similar to the general sequence generation model, conditional probability can be implemented using various neural networks. The most direct way to achieve sequence to sequence is to use two recurrent neural networks to encode and decode, respectively, which is also known as the Encoder-Decoder model.

3.1.1. Encoder. First, use a recurrent neural network Ren. To encode the input sequence $x\_(1{:}s)$ to obtain a fixed-dimensional vector u, u is generally the hidden state of the encoding recurrent neural network at the last moment.

$$\begin{aligned} h_t^e &= f_{enc}\left(h_{t-1}^e, e_{x_{t-1}}, \theta_{enc}\right), \forall t \in [1{:}S], \\ u &= h_S^e. \end{aligned} \tag{4}$$

Among them, $f_{\text{enc}}(\cdot)$ is the coded recurrent neural network, its parameter is θ_{enc}, and e_x is the vector of x.

3.1.2. Decoder. When generating the target sequence, another recurrent neural network R_{dec} is used to decode. In the t-th step of the decoding process, the generated prefix sequence is $y_{1:t-1}$. Let $\overset{t}{h}$ denote the hidden state in the network R_{dec}; $o_t \in (0,1)^{V\vee}$ is the posterior probability of all words in the vocabulary; then,

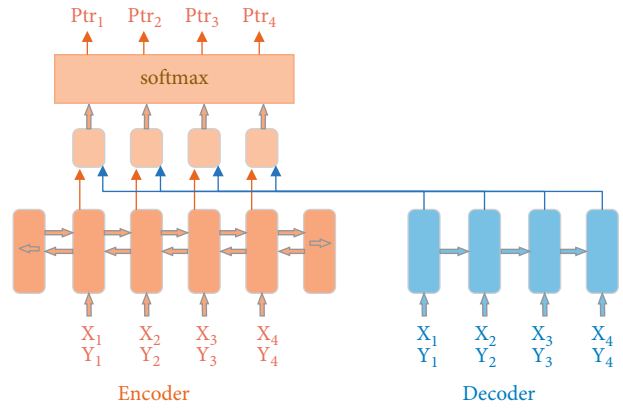


Figure 1: Schematic diagram of pointer network structure.

$$\begin{aligned} h_0^d &= u, \\ h_t^d &= f_{\text{dec}}\left(h_{t-1}^d, e_{y_{t-1}}, \theta_{\text{dec}}\right), \\ o_t &= g\left(h_t^d, \theta_0\right), \end{aligned} \tag{5}$$

where $f_{\text{enc}}(\cdot)$ is the Decoder recurrent neural network, $g(\cdot)$ is the last layer of the feedforward neural network with the softmax function, θ_{dec} and θ_0 are the parameters of the neural network, e_y is the vector of y, and y_θ is a special symbol, such as "$."

Although the sequence-to-sequence model based on the recurrent neural network is relatively easy to implement in theory, there are some shortcomings: (1) the length of the input sequence is difficult to determine, so the capacity of the encoding vector cannot be determined, and the information of the input sequence cannot be saved; (2) length dependence is an unavoidable problem of cyclic neural networks. If the input sequence is too long, it is easy to lose information.

In order to obtain more abundant input sequence information, the attention mechanism can be used to select useful information from the input sequence in each step. The hidden layers of Encoder and Decoder are represented by $(e_1, \ldots, e_s)$ and $(d_1, \ldots, d_{t(T)})$; then,

$$\begin{aligned} h_0^d &= u, \\ h_t^d &= f_{\text{dec}}\left(h_{t-1}^d, e_{y_{t-1}}, \theta_{\text{dec}}\right), \\ o_t &= g\left(h_t^d, \theta_0\right). \end{aligned} \tag{6}$$

In the t-th step of the decoding process, the model takes the node e_j of the hidden layer of input j and the current state d_t as simple input through the network, and the output obtained is the corresponding input j at this time, and all the inputs are calculated. After that, the normalized processing is performed by softmax, and then the processing result is weighted and summed with the node e_j, which is the probability distribution of the next input.

The attention mechanism can be used alone, but more often as a component in a neural network. The attention mechanism is mainly used for information screening, selecting relevant information from the input information. The attention mechanism can be divided into two steps: one is to calculate the attention distribution α, and the other is to calculate the weighted average of the input information according to α [25].

The input of the pointer network [26] is a vector sequence $X = x_1, \ldots, x_{n\_}$ n of length n, and the output is a subscript sequence $c_{1:m} = c_1, c_2, \ldots, c_m, c_i \in [1, n], \forall i$. Unlike general sequence-to-sequence tasks, the output sequence here is the subscript (index) of the input sequence.

The calculation formula of the conditional probability $p(c_{1:m} \vee x_{1:n})$ is

$$\begin{aligned} p\left(c_{1:m} \vee x_{1:n}\right) &= \prod_{i=1}^{m} p\left(c_i \vee c_{1:m}, x_{1:n}\right) \\ &\approx \prod_{i=1}^{m} p\left(c_i \vee c_1, c_2, \cdots, c_m\right). \end{aligned} \tag{7}$$

In formula (7), the conditional probability $p(c_i \vee x_{c_1}, \ldots, x_{c_{i-1}}, x_{c_{1:n}})$ is calculated using the attention distribution. For $x_{c_1}, \ldots, x_{c_{i-1}}, x_{c_{1:n}}$, if a recurrent neural network is used to encode the vector h_i, then

$$p\left(c_{1:m} \vee x_{1:n}\right) = \operatorname{softmax}\left(s_{i,j}\right). \tag{8}$$

In formula (8) $s_{i,j}$ is the unnormalized attention distribution in each input vector when decoding at the i-th step:

$$s_{i,j} = v^T \tanh\left(W x_j + U h_i\right), \quad \forall j \in [1, n]. \tag{9}$$

In formula (14), v, W, and U represent learnable parameters.

3.2. Deep Reinforcement Learning Technology. Agents can perceive their environment through sensors and act on anything in the environment with the help of actuators. Reinforcement learning can obtain the optimal strategy for sequential decision-making by maximizing the cumulative reward value that the agent obtains from the environment, which is more suitable for exploration of effective strategies to solve problems, with strong decision-making ability but lack of perception ability.

Deep learning is another important research field of machine learning, and it has made considerable progress in recent years. DL mainly combines low-level features of things through multilayer neural networks and nonlinear transformations to form abstract and easily distinguishable high-level feature representations to realize effective perception and expression of things. Although deep learning has a strong perception ability, its decision-making ability is insufficient.

Deep reinforcement learning integrates deep learning and reinforcement learning. It can give full play to their respective advantages and integrate perception and decision-making capabilities. Specifically, deep learning methods are used to obtain abstractions of large-scale input data. Representation is used as the environmental observation value for reinforcement learning, thereby obtaining the optimization of the problem-solving strategy. The basic process of deep reinforcement learning is shown in Figure 2.

The basic iterative process of deep reinforcement learning is as follows.

In the first step, the agent obtains the observations about the environment state by interacting with the environment and uses the deep learning method to realize the perception of the environment based on the obtained environment state information and determine the characteristics of the environment state. The second step is to use reinforcement learning methods to map the current state features to corresponding actions through a certain strategy.

In the third step, the environment gives feedback to the action taken by the agent and forms the next state, returning to the first step.

Deep Q-Learning (DQN, Deep Q Network) is a typical deep reinforcement learning algorithm to solve the problem that the Q-Learning algorithm cannot be applied when the state space is too high or the action space is continuous.

The basic idea of DQN is to use a neural network instead of the action value function, use a neural network to receive state-action pair as input, and output the corresponding Q-value. DQN algorithm converts the original Q-Learning problem into the corresponding deep learning problem; namely, the problem of updating the Q-table is transformed into a problem of approximating the Q function using neural network fitting, as shown in Figure 3.

DQN introduces an experience playback mechanism and dual network mechanism in the training process.

The experience playback mechanism mainly relies on the storage unit of the experience pool. At each time step, the trajectory transfer sample $e_t = (s_t, a_t, r_t, s_{t+1})$ obtained from the interaction between the agent and the environment is stored in the experience pool $P = (e_1, \ldots, e_t)$, as shown in Figure 4. When training the DQN network, each time a small batch of trajectory transfer samples is randomly taken from P, this approach effectively reduces the correlation between training samples.

In addition to using a neural network to approximate the current state-action value function, DQN also uses another network alone to generate the target Q-value. Specifically, $f(s, a \vee \theta)$ represents the output of the current Q-value network; $f(s, a \vee \theta')$ represents the output of the target Q-value network; usually, $Y = r + \gamma \max_{a'} f'(s, a \vee \theta')$ approximately represents the goal of value function optimization, namely, the target value Q_{target}. During each iteration, the parameters θ of the current Q-value network will be updated. After several iterations, the parameters θ of the current Q-value network will be used to replace the target Q-value network parameters θ', so that the target will be within a period of time. The Q-value remains unchanged, thereby reducing the correlation between the current Q-value and the target Q-value, as shown in Figure 5.

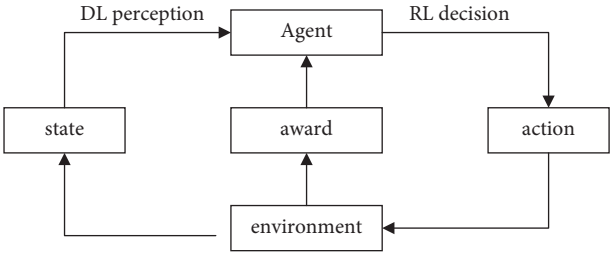


Figure 2: The basic process of deep reinforcement learning.

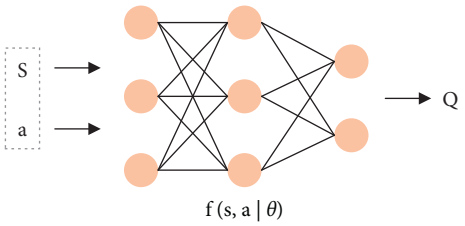


Figure 3: Schematic diagram of neural network fitting Q function.

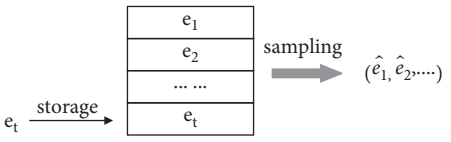


Figure 4: Schematic diagram of experience playback mechanism in DQN.

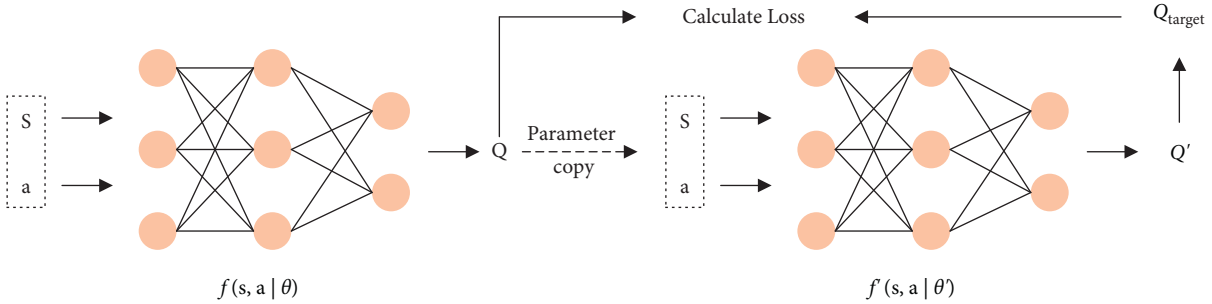


Figure 5: Schematic diagram of DQN dual network.

3.3. DQN Algorithm. In Q-Learning, let $Q(s,a)$ directly estimate the optimal state value function $^{Q}(s,a)$. The estimation method of the Q function is

$$Q(s,a) \longleftarrow Q(s,a) + \alpha\left(r + \gamma \max_{a'} Q\left(s',a'\right) - Q(s,a)\right). \tag{10}$$

When the Q-Learning algorithm is used in practical problems, the problem of dimensional explosion often occurs; that is, the state space or action space is often very large, and Q-table storage cannot be used, because the hardware cost of storage and the time cost of querying are very high. In the continuous reading state and action space, the dimensions are even greater, and obviously, storage is impossible. So we need a function $Q_{\varnothing}(s,a)$ fitted value function $Q^{\pi}(s,a)$.

$$Q_{\varnothing}(s,a) \approx Q^{\pi}(s,a). \tag{11}$$

The function $Q_{\varnothing}(s,a)$ is usually a function with a parameter of $\varnothing$, usually a neural network, and the output is a real number. If the action is a finite discrete m action "$a_1, \ldots, a_m$," we can make the Q network output a dimensional vector, where each dimension is represented by $Q_{\varnothing}(s,a_i)$, and the corresponding value function $Q(s,a_i)$ approximate value.

$$Q_{\varnothing}(s) = \begin{matrix} Q_{\varnothing}(s,a_1) \\ \vdots \\ Q_{\varnothing}(s,a_m) \end{matrix} \approx \begin{matrix} Q^{\pi}(s,a_1) \\ \vdots \\ Q^{\pi}(s,a_m) \end{matrix}. \tag{12}$$

It is necessary to learn a function $Q_{\varnothing}(s,a)$ containing the parameter $\varnothing$ to approximate the value function $Q^{\pi}(s,a)$. By fitting the mapping relationship between Q-value and s-a, the limited Q-value is replaced. To obtain good reinforcement learning results, a better fitting function is needed. In order to obtain the required fitting function, the pointer network introduced in Section 3.1 will be used. Driven by this kind of thinking, the DeepMind team proposed the Deep Q-Learning Network (Deep Q-Learning Network, DQN). According to the knowledge of the Q-Learning algorithm in the previous section, the core of each iteration of the algorithm is to seek

$$Q(s,a) = E_{s',a'}\left[r + \max_{a'} Q(s',a')\right]. \tag{13}$$

Introduce a deep neural network with parameter $\varnothing$ to fit $Q_{\varnothing}(s,a) \approx {}^{Q}(s,a)$, and train the network by gradient descent at each iteration, and the sample mean square error is caused by the loss function instead:

$$\begin{aligned} L_i(\varnothing_i) &= E_{s,a}\left[\left(y_i - Q_{\varnothing_i}(s,a)\right)^2\right], \\ y_i &= E_{s',a'}\left[r + \max_{a'} Q_{\varnothing_{i-1}}(s',a')\right]. \end{aligned} \tag{14}$$

The gradient of the loss function on $\varnothing$ is

$$\begin{aligned} \nabla_{\varnothing_i} L_i(\varnothing_i) &= E_{s,a,s'} \\ &\cdot \left[\left(r + \max_{a'} Q_{\varnothing_{i-1}}(s',a') - Q_{\varnothing_i}(s,a)\right)\nabla_{\varnothing_i} Q_{\varnothing_i}(s,a)\right]. \end{aligned} \tag{15}$$

In order to make the training of deep neural networks more efficient and convergent, the deep Q network algorithm also introduces an experience playback mechanism and a target network mechanism [27].

4. Optimization Algorithm for Logistics Transportation of Fresh Products Based on Deep Reinforcement Learning

4.1. Algorithm Structure Design

4.1.1. Design of DNQ Algorithm Structure Based on Pointer Network. For tasks that require the perception of high-dimensional raw input and output decision-making control at the same time, deep reinforcement learning algorithms have made huge and substantial progress. Because the convolutional neural network has natural advantages in image processing, the deep Q network has a level of competence with humans when it solves the complex problems related to image processing but is close to the real environment. However, the problems faced by deep reinforcement learning often have a strong time dependence. If there is a delay in the rewards in the environment, the agent needs a long number of steps to optimize the strategy. Faced with such problems, the deep Q network is not good. Performance and recurrent neural networks are suitable for handling problems related to time series [28]. Based on this, this chapter uses the pointer network model introduced in Section 3.1 to replace the convolutional neural network in the traditional deep Q network model. This modified deep Q network is called the deep Q network based on the pointer network (Deep Q Network) (DQN-PN).

In the structure diagram of the pointer network model shown in Figure 1, the encoder of the model is a two-way long and short-term memory network, and the decoder is a one-way long and short-term memory network. Assuming that a single fresh product logistics service center has 4 customer points in the deterministic logistics transportation scheduling problem, the location information of the 4 customer points is input into the encoder, and the output of the encoder each time focuses on the customer point information and other 4 points. Concerning the information of a customer point, such output is called the node of the customer point. We input the information of the customer point where the agent is at this time into the decoder and input the node containing the previous path information and the 4 customer points obtained before into a shallow neural network, and 4 outputs can be obtained. The output is normalized by softmax, and what is obtained is the location information of the next client point that the agent will visit. The size of the output dictionary is the number of client points input to the encoder.

4.1.2. Design of DNQ Algorithm Structure Based on Improved Pointer Network. When using the value function and strategy function of the pointer network model-fitting algorithm proposed in the previous section, it is found that for the efficient logistics transportation scheduling problem of multiple fresh product service centers, after the distribution service center is added, the number of transportation vehicles will increase or decrease at the same time. The input becomes very complicated; therefore, the encoder input becomes inefficient, which affects the performance of the algorithm. In order to improve the efficiency of input, this section improves the pointer network model, simplifies its structure, and makes it suitable for solving the efficient logistics transportation scheduling problem of multiple fresh product distribution service centers.

The RNN encoder adds additional complexity to the encoder, but it is not actually necessary, and the method can be made more versatile by omitting it. The encoder RNN is necessary only when the input transfers order information, and when there is no meaningful order in the input set, it is not necessary to use it for combinatorial optimization problems in the encoder. Therefore, in the model established by et al., the encoder RNN is simply omitted, and the embedded input is used directly instead of the hidden state of the RNN. With this modification, much of the computational complexity disappears without reducing the efficiency of the model.

As shown in Figure 6, the model consists of two main components. The first is a set of embeddings, which maps the input to a D-dimensional vector space. There may be multiple embeddings corresponding to different elements of the input, but they are shared between the inputs. The

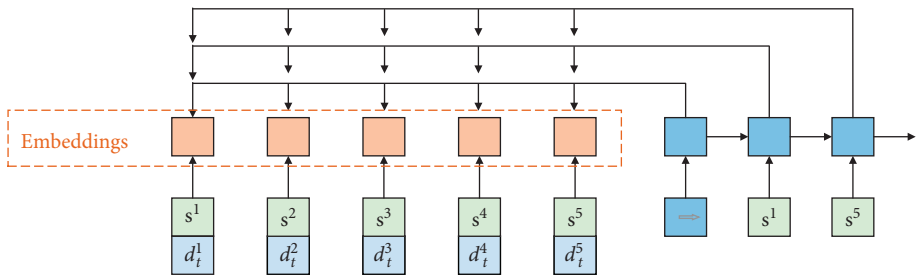


Figure 6: Improved pointer network model.

second component of the model is the decoder, which points to the input of each decoding step. Use RNN to simulate the decoder network, and provide static elements as input to the decoder network, and dynamic elements can also be the input of the decoder.

In the model in Figure 6, the embedding layer on the left maps the input to a high-dimensional vector space. On the right, the RNN decoder stores the information of the decoded sequence. Then, the hidden state and embedded input of the RNN use the attention mechanism to generate a probability distribution on the next input.

4.2. Algorithm Steps. Through the introduction of the pointer network and the deep Q network in the previous chapter, the specific algorithm steps of the deep Q network based on the pointer network are as follows: d_t^1

4.3. Algorithm Environment Design. From the steps of Algorithm 1 in the previous section, it can be seen that it essentially uses the reinforcement learning of the deep learning fitting value function. Therefore, the process of the deep Q network algorithm based on the pointer network is also the continuous interaction between the agent and the environment. If we want to use this algorithm, we need to design a single fresh product efficient logistics and transportation scheduling problem, including the state space S design, the action space A design, and the action reward r design. The discount rate γ, the learning rate α, and the parameter update interval C are relatively simple inputs, and their values will be given in the experimental part, and no design will be done in this section.

4.3.1. State Space S Design. One of the inputs of the algorithm is the state space S, so the state space S is designed first. According to the algorithm structure design ideas in Section 4.1, the fresh product logistics transportation scheduling problem model is determined. At time t, the state s_t of the environment at this time is represented by a set of vectors $(p, q_t, j_1, j_2, \ldots, j_N)$. The state of the environment is constantly changing over time. Although time is a continuous variable, according to the Markov property of reinforcement learning, the state change trajectory of the environment can be divided into countless discrete states in a certain short period of time. Here, the unit is separated by the length of time the delivery vehicle moves from one customer point to another customer point. Assuming there are T states, in the fresh product logistics transportation scheduling problem, $T = N + 2m$, m is the number of times to return to the distribution center. All states form a state set, that is, the state space S, $S = \{s_t: 0, \ldots, T\}$.

4.3.2. Action Space A Design. Another input of the algorithm is the action space A. The output of the deep Q network algorithm based on the pointer network is the mapping relationship function between state and action. Therefore, the design of the action space is indispensable for environmental design. In the logistics and transportation scheduling problem of fresh products, in each step of the iteration, the agent chooses the next customer point to be visited which is the action. The agent can only choose one action to execute. In addition, the time discretization of the action space is consistent with the state space of the environment. In the problem to be solved, the transportation vehicle needs to depart from the distribution service center and eventually return to the distribution service center, so the default " $a_0 = a_T = p_0$, that is, the first action and the last action select the distribution center. At other times, it has never chosen one of the customer points visited. If the remaining load capacity of the vehicle cannot meet the needs of the customer point, the strategy is to return the vehicle to the distribution service center to unload the goods and then continue to provide services.

4.3.3. Action Reward r Design. Action reward is the key to the algorithm because it determines the learning direction and efficiency of the deep Q network algorithm of the pointer network. In each step of the iterative process, the environment will give a reward value according to the current state and the action chosen to be executed, and the evaluation and improvement of the strategy will be carried out according to the reward r. In the fresh product logistics transportation scheduling problem, the goal is to make the total visit distance the smallest, and the reward for each step can be represented by the distance between two customer points.

5. Simulation Experiment and Result Analysis

5.1. Experimental Environment and Parameter Settings. All the experimental algorithms in this article are implemented on the Linux system based on the Tensorflow

platform using Python. The computer's CPU is Intel Core i7 8700, the CPU frequency is 3.2 GHz, the graphics card is NVIDIA GeForce GTX 1080Ti, and the RAM is 16 GB.

The parameter settings are shown in Table 1.

5.2. Simulation Experiment Results and Analysis. Simulation Experiment 1: solve a single fresh product distribution service center-regional efficient logistics scheduling problem.

Since there is no public cvRPSD calculation example set, in order to verify the solution effect of Deep Q Network based on pointer network (DQN-PN), in this section, the experimental data will be obtained from the benchmark example of the deterministic CVRP example set (http://www.branchandcut.org//data/), and then through the improvement of the deterministic CVRP example set, the experiments needed in this section will be obtained. CVRPSD calculation example set is as follows. Compared with the CVRP calculation example, other information is not changed except for the needs of the customer. Assume that the random demand of the customer point in this experiment obeys a discrete distribution, and the expected value of the random demand is equal to the demand of the corresponding customer point in the CVRP calculation example. The random distribution function of the demand of customer point i is shown in Table 2. The random distribution function of the demand of other customer points obeys the same formula as that of customer point i, and the difference lies in the different distribution probabilities. Table 2 is only used to show the form of the data and does not have the data details of the examples. The space is limited, and the data details of each customer point are no longer listed. At the same time, since no result set of using deep reinforcement learning to solve such problems was found, the results of the frequently used simulated annealing algorithm (simulated annealing, SA) were used as a reference.

In Table 3, BKS represents the best-known value of the calculation example; BS represents the best value obtained in the algorithm experiment; AS is the average value obtained in the algorithm experiment; CT represents the average time consumed in the algorithm experiment and the unit of time consumption for seconds.

Combining Table 3 and Figure 7 shows that for different types and scales of calculation examples, the deep Q network algorithm based on the pointer network is effective and stable for solving the single fresh product distribution service center-regional efficient logistics scheduling problem. There is a higher solution accuracy. When the solution time is not much different, most of the results are better than the solution results of the simulated annealing algorithm. This shows that the method designed in the article can effectively solve the problem of single fresh product distribution service center-regional efficient logistics scheduling. The main reason is that the pointer network model network with better approximation performance is adopted.

Simulation Experiment 2: solve the problem of multiple fresh product distribution service centers-regional efficient logistics scheduling.

In order to verify the effectiveness of the improvement based on the improved DNQ algorithm of pointer network (I-DNQ-NP) in solving multiple fresh product distribution service centers-regional efficient logistics scheduling problems, this section designs and the previous one contrast experiment of section algorithm. Since there is no public MDCVRPSDTW example set, and it is not meaningful to directly use the result data of simulation experiment 1, this section changes the examples Q1 and Q2 in the related literature [30] to change the definite requirements of each customer point used after being a random demand variable.

Calculation example Q1 is to solve the logistics scheduling problem of multiple fresh product distribution service centers with 3 distribution service centers and 15 customer points. The specific information of the distribution service center is shown in Table 4, and the specific information of the customer points is shown in Table 5.

Example Q2 is the logistics scheduling problem of multiple fresh product distribution service centers with 4 distribution service centers and 25 customer points. The specific information of the distribution service center is shown in Table 6, and the specific information of customer points is shown in Table 7.

In Tables 4–7, the X and Y coordinates represent the coordinates of the customer point and the coordinate position of the freight yard; K represents the number of vehicles available in the distribution center; Q represents the carrying capacity of each vehicle in the distribution center; I represents the basic salary of the driver for each vehicle startup in the distribution service center; H represents the hourly salary of the driver of the distribution center; E is the expected value of the customer's random demand; ET is the time when the service can be started; LT is the time when the customer can be serviced.

The comparison of the experimental results is shown in Table 8. In the table, TO in BS represents the timeout time of the optimal solution, DIS in BS represents the path length of the optimal solution, TO in AS represents the average timeout time of the solution obtained by solving the calculation example 20 times, and DIS in AS represents the average path length of the solution obtained by solving the calculation example 20 times, F is the driver's salary, CT is the calculation time, and the unit of time is seconds.

Through simulation test two, it can be known that under the conditions of the above experimental parameters, both DNQ-NP and I-DNQ-NP can obtain feasible solutions that do not exceed the carrying capacity of the transport vehicle and meet the customer's point time. Analyzing the data in Table 8 shows that under the same parameter conditions, I-DNQ-NP is equivalent to DNQ-NP in terms of solution quality and accuracy (no difference between the optimal value and the average value). However, it has been drastically reduced. The experimental results prove that the modification of the pointer network is effective. The RNN encoder does add additional complexity to the encoder, simply omitting the encoder RNN, and directly using embedded input instead of RNN hidden state. Through this modification, a lot of computational complexity disappears, which greatly improves the efficiency of the model.

Input: State space S, action space A, discount rate γ, learning rate α, parameter update interval C
(1) Initialize experience pool D, the capacity is N_D;
(2) Randomly initialize the parameters of the Q network $\varnothing$;
(3) Randomly initialize the parameters of the target Q network $\widehat{\varnothing} = \varnothing$;
(4) repeat
(5) Initialize the initial state s;
(6) repeat
(7) In state S, choose action $a = \pi^{\varepsilon}$
(8) Perform action a, observe the environment, get an instant reward r and a new state s';
(9) Put s, a, r, s' into D;
(10) Sample ss,aa,rr,ss' from D;
(11) $y = \begin{cases} rr, ss' \text{Terminal state,} \\ rr + \gamma \max\limits_{a'} Q_{\varnothing}(ss', a') \text{other.} \end{cases}$;
(12) Use $(y - Q_{\varnothing}(ss, aa))^2$ as the loss function to train the Q network;
(13) $s \longleftarrow s'$;
(14) Every C steps, $\widehat{\varnothing} = \varnothing$;
(15) Until s is the termination state;
(16) Until $\forall s, a$, $Q_{\varnothing}(s, a)$ converges;
output : Q network $Q_{\varnothing}(s, a)$

Algorithm 1: Deep Q network based on pointer network

Table 1: Parameter setting [29].

Parameter	Parameter description	Value
Enc-net	Encoder parameters	$(N+2) \times 256 + (N+2) \times 256^2$
Dec-net	Decoder parameters	$(N+2) \times 256 + (N+2) \times 256^2$
γ	Discount rate	1.0
α	Learning rate	0.001
β	Learning rate	0.002
ε	Parameters of ε − greedy method	[0.1, 1.0]
C	Parameter update interval	0
N_D	Capacity of experience pool D	5000

Table 2: The random distribution function table of the demand of customer i.

Demand	ζ_1	ζ_2	⋯	ζ_k
Probability distributions	$p_{i,1}$	$p_{i,2}$	⋯	$p_{i,k}$

Table 3: Comparison of result sets of CVRPSD calculation examples.

Examples	BKS	SA		DNQ-PN		
		BS	CT	BS	AS	CT
E-n22-k4	375	411.57	104.1	406	409	84.6
P-n23-k8	529	619.53	21.1	580	642	17.2
A-n32-k5	784	853.6	199.8	822	868	138.4
A-n33-k4	661	704.2	178.2	695	730	196.5
E-n33-k4	835	850.27	371.7	813	855	239.2
A-n34-k5	778	826.87	236.4	788	823	240.7
A-n36-k5	799	858.71	276.1	815	823	235.2
A-n39-k5	822	869.18	257.6	754	809	243.6
P-n40-k5	458	472.5	367.2	430	495	338.1
A-n44-k6	937	1025.48	281.4	1019	1136	253.0
A-n45-k7	1146	1264.99	216.0	1213	1285	248.9
P-n50-k10	696	760.94	150.6	732	764	157.6
E-n51-k5	521	552.26	586.0	594	621	445.8
A-n55-k9	1073	1179.11	265.5	1149	1269	248.5
A-n60-k9	1354	1529.82	393.7	1477	1504	354.8

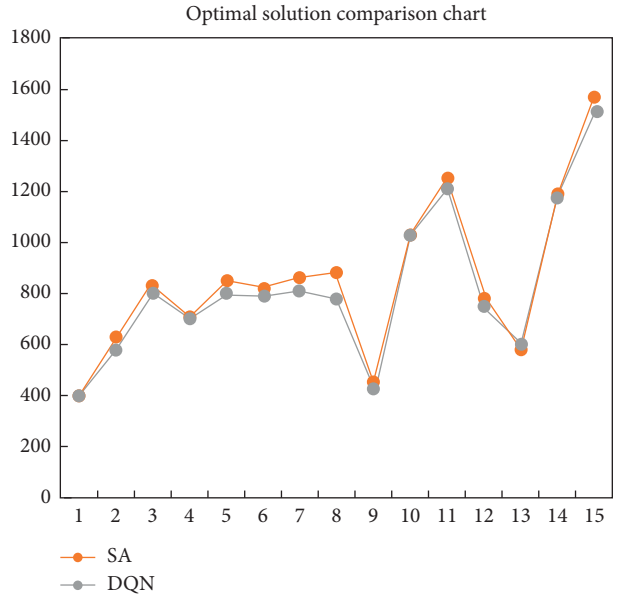


Figure 7: Comparison of optimal solutions.

Table 4: Calculation example Q1 distribution service center information table.

Distribution Service center	X coordinate	Y coordinate	K	Q	I	H
1	60	50	2	100	1000	10
2	10	80	2	100	1200	12
3	100	80	2	100	1500	15

Table 5: Example Q1 customer point information table.

customer	X coordinate	Y coordinate	$E(j)$	ET	LT
1	9	115	10	140	510
2	80	86	30	100	260
3	90	112	10	65	146
4	40	112	10	410	640
5	72	65	10	15	67
6	60	99	20	265	710
7	25	66	20	50	300
8	10	60	20	10	70
9	38	70	10	334	705
10	35	66	10	100	350
11	35	89	10	260	450
12	25	105	20	152	321
13	52	75	30	30	135
14	65	85	10	167	620
15	20	100	40	384	529

Table 6: Calculation example Q2 distribution service center information table.

Distribution Service center	X coordinate	Y coordinate	K	Q	I	H
1	30	80	3	150	1200	12
2	22	20	3	150	1200	12
3	80	90	3	150	1500	15
4	75	40	3	150	1800	18

Table 7: Example Q2 customer point information table.

Customer	X coordinate	Y coordinate	$E(j)$	ET	LT
1	65	40	10	129	450
2	100	86	30	100	170
3	32	66	10	65	146
4	85	90	10	10	70
5	75	65	10	115	367
6	90	69	20	150	210
7	80	46	20	170	225
8	88	98	20	50	95
9	45	70	10	234	605
10	80	66	10	200	450
11	65	29	10	448	505
12	25	85	20	0	100
13	22	75	30	30	92
14	52	85	10	367	467
15	50	80	40	184	429
16	20	95	40	155	548
17	58	65	20	99	148
18	35	95	20	179	254
19	45	90	10	258	345
20	30	50	10	10	73
21	39	42	20	310	532
22	28	36	20	12	433
23	18	41	10	102	397
24	25	20	10	0	60
25	5	14	40	250	300

Table 8: Comparison of experimental results.

			Q1	Q2
DNQ-NP	BS	TO	0	0
		DIS	4450	6267
		F	10055	16632
	AS	TO	0	0
		DIS	4521	6748.5
		F	10127	17684
	CT		194	383
I-DNQ-NP	BS	TO	0	0
		DIS	4461	6194
		F	10169	15568
	AS	TO	0	0
		DIS	4487	6446.8
		F	10207	17006
	CT		161	328

6. Conclusion

The language expression of the conclusion analysis part is more colloquial. It is recommended to clarify the logical relationship and make a rigorous expression. In today's highly developed material civilization, the daily fresh product needs of residents' lives are more personalized, and at the same time, the effectiveness of customer needs is gradually becoming higher. This further brings challenges to the logistics and transportation scheduling of the fresh product market (how to achieve it). The precise delivery of customer needs puts forward higher requirements for the efficient transportation of fresh product logistics. Based on this, this paper proposes a DNQ algorithm based on a pointer network to solve the single fresh product distribution service center-regional efficient logistics scheduling problem, and a feasible logistics transportation scheduling plan can be obtained through simulation experiments. The simulation results show that the algorithm is superior to other common intelligent algorithms in terms of accuracy and stability, which proves that the algorithm is effective and feasible. Based on this, we further explored the improved DNQ algorithm of the correction network, which can solve multiple fresh product distribution service centers-regional efficient logistics scheduling problems. The improved algorithm reflects its efficient model solving efficiency on complex problems, and through simulation, the analysis can further obtain multiple fresh product distribution service centers-regional efficient logistics scheduling schemes. It is a successful attempt to improve the solution algorithm. It is of great significance for further research on solving the complex logistics transportation scheduling problems of fresh products. Through this research, we can see that the DNQ algorithm of the pointer network has an important position and practical value in solving the actual problem of fresh product logistics transportation scheduling, but the transportation scheduling model in this article is a very abstract model. When the follow-up is further combined with the classic algorithm research, it needs to be further improved in practical applications.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

Research and Practice of Applied Undergraduate Talent Training Mode from the perspective of AACSB International Certification No. zd20054.

References

[1] L. Macheka, E. Spelt, J. G. A. J. van der Vorst, and P. A. Luning, "Exploration of logistics and quality control activities in view of context characteristics and postharvest losses in fresh produce chains: a case study for tomatoes," *Food Control*, vol. 77, pp. 221–234, 2017.

[2] M. Musavi and A. Bozorgi-Amiri, "A multi-objective sustainable hub location-scheduling problem for perishable food supply chain," *Computers & Industrial Engineering*, vol. 113, pp. 766–778, 2017.

[3] F. Neves-Moreira, D. Pereira da Silva, L. Guimarães, P. Amorim, and B. Almada-Lobo, "The time window assignment vehicle routing problem with product dependent deliveries," *Transportation Research Part E: Logistics and Transportation Review*, vol. 116, pp. 163–183, 2018.

[4] M. Rahimi, A. Baboli, and Y. Rekik, "Multi-objective inventory routing problem: a stochastic model to consider profit, service level and green criteria," *Transportation*

Research Part E: Logistics and Transportation Review, vol. 101, pp. 59–83, 2017.

[5] N. Lei, "Intelligent logistics scheduling model and algorithm based on Internet of Things technology," *Alexandria Engineering Journal*, vol. 61, no. 1, pp. 893–903, 2022.

[6] qingX. Xin, Z. Fu, Z. Zhu, and X. Zhang, "Improved preservation process for table grapes cleaner production in cold chain," *Journal of Cleaner Production*, vol. 221, pp. 1171–1179, 2019.

[7] H. Zhao, S. Liu, C. Tian, G. Yan, and D. Wang, "An overview of current status of cold chain in China," *International Journal of Refrigeration*, vol. 88, no. 88, pp. 483–495, 2018.

[8] X. Zhang and G. Li, "Study on the time-space optimization for cold-chain logistics of fresh agricultural products," in *Proceedings of the 2010 International Conference on Future Information Technology and Management Engineering*, no. 1, pp. 331–333, Changzhou, China, October 2010.

[9] J. K. Carson and A. R. East, "The cold chain in New Zealand-a review," *International Journal of Refrigeration*, vol. 87, no. 87, pp. 185–192, 2018.

[10] M. de Keizer, R. Akkerman, M. Grunow, J. M. Bloemhof, R. Haijema, and J. G. A. V. D. Vorst, "Logistics network design for perishable products with heterogeneous quality decay," *European Journal of Operational Research*, vol. 262, pp. 535–549, 2017.

[11] M. Grazia Speranza, "Trends in transportation and logistics," *European Journal of Operational Research*, vol. 264, no. 3, pp. 830–836, 2018.

[12] H. Shi, L. Sun, Y. Teng, and X. Hu, "An online intelligent vehicle routing and scheduling approach for B2C e-commerce urban logistics distribution," *Procedia Computer Science*, vol. 159, pp. 2533–2542, 2019.

[13] L. Liu, H. Wang, and S. Xing, "Optimization of distribution planning for agricultural products in logistics based on degree of maturity," *Computers and Electronics in Agriculture*, vol. 160, pp. 1–7, 2019.

[14] Z. Lu, Z. Zhuang, Z. Huang, and W. Qin, "A framework of multi-agent based intelligent production logistics system," *Procedia CIRP*, vol. 83, pp. 557–562, 2019.

[15] J. Cheng, B. Yang, M. Gen, Y. J. Jang, and C.-J. Liang, "Machine learning based evolutionary algorithms and optimization for transportation and logistics," *Computers & Industrial Engineering*, vol. 143, Article ID 106372, 2020.

[16] A. V. Barenji, W. M. Wang, Z. Li, and D. A. Guerra-Zubiaga, "Intelligent E-commerce logistics platform using hybrid agent based approach," *Transportation Research Part E: Logistics and Transportation Review*, vol. 126, pp. 15–31, 2019.

[17] C. Qi and L. Hu, "Optimization of vehicle routing problem for emergency cold chain logistics based on minimum loss," *Physical Communication*, vol. 40, Article ID 101085, 2020.

[18] X. Zhang, J. S. L. Lam, Shipping mode choice in cold chain from a value-based management perspective," *Transportation Research Part E: Logistics and Transportation Review*, vol. 110, pp. 147–167, 2018.

[19] W. E. Soto-Silva, M. C. González-Araya, M. A. Oliva-Fernández, and L. M. Plà-Aragonés, "Optimizing fresh food logistics for processing: application for a large Chilean apple supply chain," *Computers and Electronics in Agriculture*, vol. 136, pp. 42–57, 2017.

[20] V. R. Ghezavati, S. Hooshyar, and R. Tavakkoli Moghaddam, "A Benders' decomposition algorithm for optimizing distribution of perishable products considering postharvest biological behavior in agri-food supply chain:a case study of tomato," *CEJOR*, vol. 25, pp. 29–54, 2017.

[21] J. Chai, "Study on route optimization of cold chain logistics of fresh food," *Carpathian Journal of Food.Science and Technology*, vol. 8, no. 2, pp. 113–121, 2016.

[22] M. Bortolini, M. Faccio, E. Ferrari, M. Gamberi, and F. Pilati, "Fresh food sustainable distribution: cost, delivery time and carbon footprint three-objective optimization," *Journal of Food Engineering*, vol. 174, no. 1, pp. 56–67, 2016.

[23] S. Y. Wang, Y. Shi, F. Tao, and H. Wen, "Optimization of vehicle routing problem with time windows for cold chain logistics based on carbon tax," *Sustainability*, vol. 9, no. 694, pp. 1–23, 2017.

[24] X. Dai, M. Chen, and Y. Zhou, "Optimal logistics transportation and route planning based on fpga processor real-time system and machine learning," *Microprocessors and Microsystems*, vol. 80, Article ID 103621, 2021.

[25] O. Stopka, K. Jeřábek, and M. Stopková, "Using the operations research methods to address distribution tasks at a city logistics scale," *Transportation Research Procedia*, vol. 44, pp. 348–355, 2020.

[26] O. Vinyals, M. Fortunato, and N. Jaitly, *Pointer Networks*, Computer Science, 2015, https://arxiv.org/abs/1506.03134.

[27] S. R. Cardoso, A. P. F. D. Barbosa-Póvoa, and S. Relvas, "Design and planning of supply chains with integration of reverse logistics activities under demand uncertainty," *European Journal of Operational Research*, vol. 226, no. 3, pp. 436–451, 2013.

[28] Y. Wang, J. Zhang, X. Guan, M. Xu, Z. Wang, and H. Wang, "Collaborative multiple centers fresh logistics distribution network optimization with resource sharing and temperature control constraints," *Expert Systems with Applications*, vol. 165, Article ID 113838, 2021.

[29] B. Bai, K. Zhao, and X. Li, "Application research of nano-storage materials in cold chain logistics of e-commerce fresh agricultural products," *Results in Physics*, vol. 13, Article ID 102049, 2019.

[30] H. M. Stellingwerf, L. H. C. Groeneveld, G. Laporte, A. Kanellopoulos, J. M. Bloemhof, and B. Behdani, "The quality-driven vehicle routing problem: model and application to a case of cooperative logistics," *International Journal of Production Economics*, vol. 231, Article ID 107849, 2021.

Hindawi
Scientific Programming
Volume 2022, Article ID 2802573, 13 pages
https://doi.org/10.1155/2022/2802573

Research Article

Multimodal Music Emotion Recognition Method Based on the Combination of Knowledge Distillation and Transfer Learning

Guiying Tong

College of Music, Huizhou University, Huizhou 516000, Guangdong, China

Correspondence should be addressed to Guiying Tong; tgy@hzu.edu.cn

Received 11 December 2021; Revised 4 January 2022; Accepted 7 January 2022; Published 11 February 2022

Academic Editor: Baiyuan Ding

The main difficulty of music emotion recognition is the lack of sufficient labeled data. Only the labeled data with unbalanced categories are used to train the emotion recognition model. Not only is accurate labeling of emotion categories costly and time-consuming, but it also requires extensive musical background for labelers At the same time, the emotion of music is often affected by many factors. Singing methods, music styles, arrangement methods, lyrics, and other factors will affect the expression of music emotions. This paper proposes a multimodal method based on the combination of knowledge distillation and music style transfer learning and verifies the effectiveness of the method on 20,000 songs. Experiments show that compared with traditional methods, such as single audio, single lyric, and single audio with multimodal lyric methods, the method proposed in this paper has significantly improved the accuracy of emotion recognition, and the generalization ability has been significantly improved.

1. Introduction

Music works contain rich human emotions, and emotions play an indispensable role in the transmission of musical emotions and understanding and appreciation of music [1–3]. With the current development of Internet technology and artificial intelligence, the amount of digital music is growing rapidly. With a large number of music works, how to recommend suitable music according to different environments and different moods of users has become a hot topic of research in recent years. In this context, the automatic recognition technology of music emotion has attracted more and more attention from the industry. In recent years, deep learning technology has replaced traditional statistical algorithms as the mainstream technology in the field of automatic music emotion recognition [4–6]. The main content of music includes digital audio and lyrics text, and the current research on music emotion recognition mainly focuses on these two aspects.

Existing studies have applied artificial intelligence and Internet of Things technology to the teaching of art courses and the work of art exhibition. In traditional statistical methods, models need to design complex elements and features for repeated understanding and recognition manually, which requires a lot of time and learning costs. Unlike traditional methods, deep learning algorithms can automatically identify and extract the most appropriate representative feature elements from the data. At the same time, deep learning has shown outstanding contributions in many fields such as speech recognition and machine translation. Some scholars have achieved excellent results in the music emotion classification competition task through the in-depth application of deep learning technology (Music Information Retrieval Evaluation Exchange), combined with the emotion recognition method of music audio [7–9] in Figure 1.

Some scholars also proposed an end-to-end multimodal emotion recognition binary classification method based on music audio and lyrics, and the superiority of this method was verified through specific experiments. Compared with the emotion recognition method of single music audio and single lyric music, this method has a significant improvement in recognition accuracy [10–12]. Some scholars have conducted research on the emotion recognition of multimodal music and carried out detailed comparative experiments on the methods of fusion between different music in

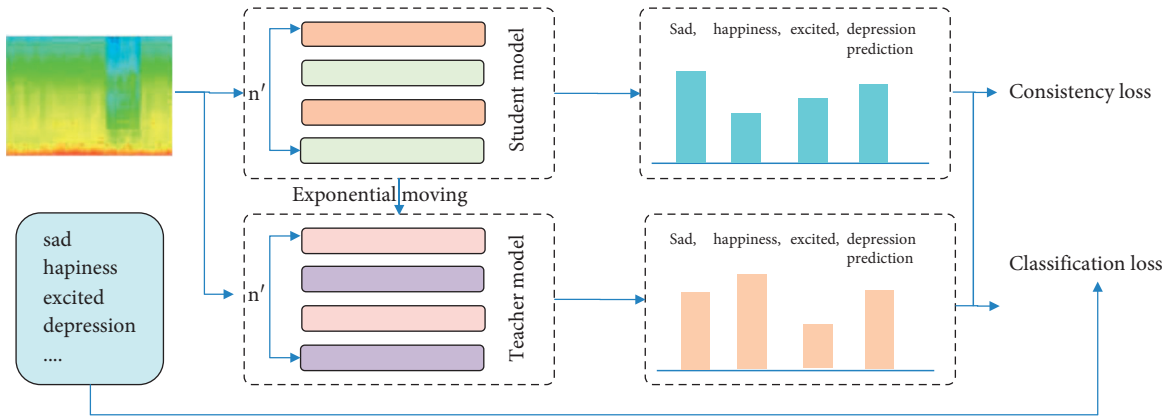


Figure 1: The structure of the student-teacher model.

different periods. The results show that the middle fusion method of music is better than the front fusion and the post-fusion. In addition, some researches have carried out comparative analysis and research on several classic music emotion data sets that are currently available in the industry, e.g., CAL500, MIREX 2007 AMC, and MediaEval Emotion in Music. It is found that labeling music emotion data sets has high cost and is relatively time-consuming. In the research on the related content of the music public data set, it is found that the data quantity and quality of the public data set in this respect are not ideal, and many data sets have problems such as the unbalance of emotion categories [13–15].

Based on the multimodal music emotion recognition method, combining knowledge distillation and transfer learning, this article attempts to improve the accuracy of music emotion recognition effectively when the amount of labeled data is too small or the emotion categories are not balanced. The method based on knowledge distillation specifically analyzes the application scenarios of the teacher-student model in the image field. On the one hand, the model shows relatively excellent processing performance [16–18]. On the other hand, it also builds a multimodal neural network structure based on audio and lyrics recognition methods. This network structure can quickly identify and efficiently classify different music styles, which is conducive to the rapid application of music recommendation [19–23] models in Figure 2.

2. Data Preprocessing

2.1. Audio Signal Expression and Preprocessing Technology. Mel spectrum, as a signal representation method generally applicable to common audio classification tasks, is better than other high-level audio signal representations. This method relatively completely retains the characteristic information of a series of signals in the music work and truly achieves the lossless preservation of the music information [24–27]. At the same time, Mel's frequency spectrum is more in line with the auditory characteristics of the field of ergonomics. Based on the above multiple factors, this paper selects the Mel spectrum method as the input data for music audio analysis. In addition, this article adopts the Voice Activity Detection (VAD) method to specifically detect whether there is a silent frame in the music signal. Considering the audio part of the silent frame will affect the overall recognition of the specific audio and the flow of the expression and preprocessing of the music audio signal in this paper shown in Figure 3.

A piece of music mainly includes audio information and lyrics text information, and audio information includes time domain features, frequency domain features, and cestrum features. Time domain features usually refer to time domain parameters calculated in fixed-length music, which mainly include short-term energy, short-term average amplitude, and average amplitude difference functions. Frequency domain characteristics refer to the characteristic parameters obtained by converting fixed-length music from time domain signals into frequency domain signals through Fourier transform, which mainly include spectrum centroid, spectrum roll-off, and spectrum flux. The cestrum feature refers to the use of human ears to have different perception effects on the different loudness, pitch, and timbre of the sound. Therefore, it is used to classify different music emotions. The specific operation methods are as follows: (1) the model divides the music into specific lengths and uses Fourier conversion to frequency domain signals; (2) the model calculates the logarithm of the frequency domain signals and then performs the inverse Fourier transform [28–31].

2.2. Word Vector Pre-Representation of Lyrics Text. Taking into account the types of English songs used in this article, the lyrics of the songs include English and other different types of languages. In order to ensure the compatibility of the system with multiple languages, this article cuts the words of each character in English songs according to special characters, which guarantees the versatility and fluency of the word-slicing algorithm [32–35]. The word vector expresses the adaptive expression method of dynamic lyrics recognition and segmentation. At the same time, this paper performs word frequency statistics and word vector initialization processing on the pretrained lyrics collection. In the specific

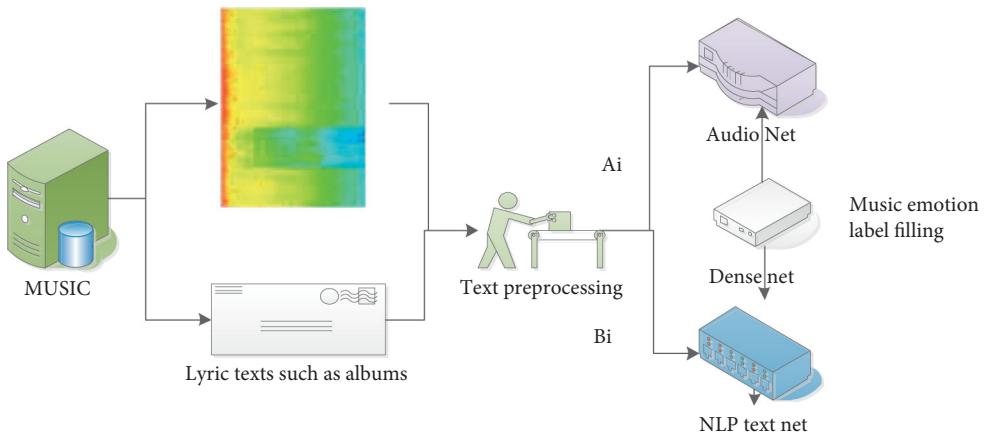


Figure 2: Multimodal neural network structure based on audio and lyrics text.

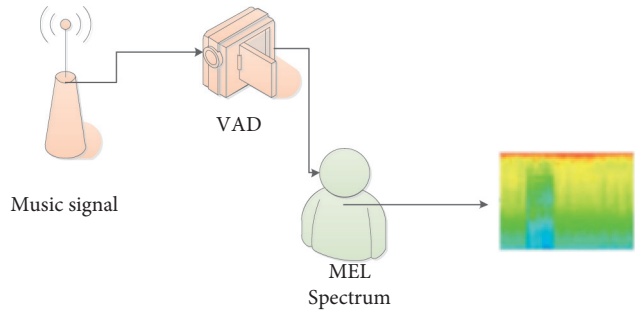


Figure 3: System pretreatment process.

model training, for the convenience of adjusting the word vector, part of the algorithm optimization work is carried out [36–38]. In addition, this article sets the word vector word_dim to 138, and the maximum length parameter of lyrics to 200. In this paper, the preprocessing of the word vector in the lyric text is shown in Figure 4.

3. Student-Teacher Model and Transfer Learning Methods

3.1. Student-Teacher Model. Considering that the data was less labeled and relatively unevenly distributed, this paper uses the process model of Figure 1 to express the student-teacher model, thereby improving the accuracy of the overall music emotion recognition. The teacher-student model specifically uses the existed music genre network recognition architecture, and we then use different genre recognition systems to learn the network parameters in stages.

At the same time, considering that the network reasoning ability and performance of the teacher system are generally higher than those of the student group, this paper specifically optimizes the network parameters of the music style selected in the teacher model, and its performance is generally better than the network parameters of the students. On this basis, the characteristics, advantages, and disadvantages of this type of network structure were further elaborated.

The machine learning method used in the text mainly uses the thinking mode of transfer learning and mainly uses two methods for knowledge transfer. The first mode is to use the trained song style network structure as a feature extraction tool for emotion recognition and use it for the recognition of part of the music emotion network after feature training. The second mode mainly combines the already trained song style network structure with the newly added network into the system for training. This article mainly chooses the second training mode. Considering that the teacher network model does not participate in the backpropagation of the neural network system, this paper compares the teacher model parameter W with the student model parameter w through Exponential Moving Average (EMA) analysis; the specific parameter expression results of the teacher model at time t are as follows:

The Wt expression of the most commonly used function is

$$\begin{aligned} Wt &= \alpha \times Wt - 1 + (1-\alpha) \times \omega t, \\ \text{Floss1} &= Hw(\alpha) + \lambda Kw * Lw(\alpha), \\ \text{Floss2} &= Hw(\alpha) + \lambda w * \text{KL}(\alpha) + (1-\alpha) \times \omega t, \\ \text{Floss3} &= Hw(\alpha) + \lambda w * \text{KL}(\alpha) + (1-\alpha) \times \omega t + \alpha \\ &\quad \times Wt - 1. \end{aligned} \tag{1}$$

Among them, α represents the attenuation-smoothing rate, Wt represents the parameters of the student model at

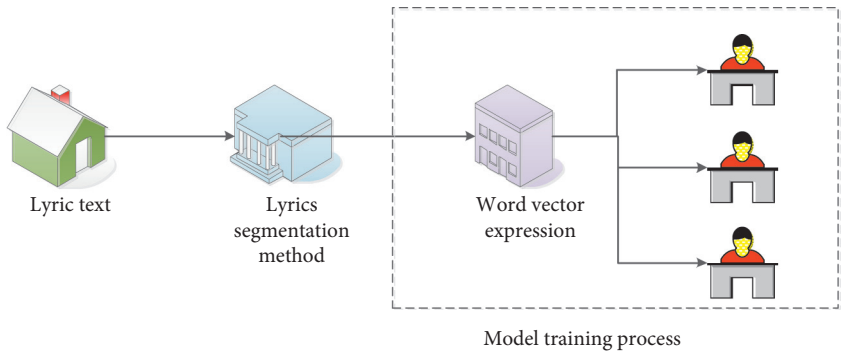


Figure 4: Word vector pre-expression of lyrics.

time t. The F-loss function of the teacher-student model is composed of relative entropy (Kullback–Leibler divergence) $\mathrm{KL}(\alpha)$ and classification cross entropy $\mathrm{Hw}(\alpha)$ (control of the importance of two parts of CE). For the same sample X, first get the output s through the student model, then get the output t through the teacher model, and use the data t and s to calculate the $\mathrm{KL}(\alpha)$ loss. On this basis, the result data of the gradient analysis calculation is passed back to the student model, and the CE loss value in the s model is further calculated. Finally, according to the two parts of the loss function, the specific parameters of the student model are updated; the student model parameters are further analyzed by exponential smoothing.

At the same time, in the process of model training, iterative analysis of model parameters is required. Input the sample data set X into the model. The data set contains both emotionally labeled samples and unlabeled samples. In addition, these sample sets are enhanced through the system model, which further improves the generalized understanding of emotional data of the system machine learning model.

3.2. Data Enhancement

3.2.1. Gaussian Noise Analysis. English songs are often accompanied by various types of noise during the specific recording and dissemination process. In the process of preprocessing the audio, this paper can further use Gaussian noise processing for all audio content and perform noise enhancement processing on audio effects.

3.2.2. Audio Cutting Process. Before analyzing the audio, perform audio tailoring according to different audio types and styles. According to different types of song fragments, the machine learning model system simulates human emotions when listening to music. The average length of the audio clip collection of the model is set to 30 seconds to ensure the stability and fluency of the audio recognition process of the system.

3.2.3. Audio Mixing Analysis. This article randomly mixes different types of audio. The model specifically assumes that a pair of audio samples (Xa, Xb) are processed, and the resulting mixed sample is set to *Xcom*. It should be noted that the generated *Xcom* is only used to calculate the loss of the model *KL* and cannot be used for multidimensional systematic mixing of different types of audio for emotional tags. Among them, γ is the sample-mixing coefficient.

$$\begin{aligned} \mathrm{Xcom} &= (1-\gamma) * \mathrm{Xa} + \gamma * \mathrm{Xb}, \\ \mathrm{Ycom} &= (1-\gamma) * \mathrm{Ya} + \gamma * \mathrm{Yb}. \end{aligned} \tag{2}$$

Forward propagation of working signal: take a sample N_m from the sample set and set θ_i as the threshold of the i neuron of the hidden layer:

$$\begin{aligned} Y_m &= f\left(\sum_{i=1}^{m} w_{ih} X_m - \theta_i\right), \\ O_j &= g\left(\sum_{j=1}^{m} w_{hj} Y_m - \theta_j\right). \end{aligned} \tag{3}$$

In addition to audio information in music, lyrics text information also implies certain emotions. In order to avoid discretization of emotional expression due to the sparse distribution of words in the lyrics, short sentences, Y_m, and repetition, it is necessary to filter the lyrics. First, create a vocabulary of word emotion recognition. The common words in the lyrics are classified according to the emotions and strengths they express. For example, the term "sun" is usually used to express hopeful and positive emotions, while the term "car" expresses emotions that are rather vague. Therefore, in terms of emotion recognition, "sun " is more efficient than "car." When creating a vocabulary of word emotion recognition, vocabulary like "sun" should be kept, and vocabulary like "car" should be avoided at the same time. After the word emotional recognition vocabulary is created, Oj, a certain vocabulary, needs to be identified and quantified. The specific formula is as follows. The total network error is

$$e_{(n)} = \frac{1}{2} \sum_{j=1}^{m} \left(O_j - d_j\right)^2. \tag{4}$$

Because the data collected in the research are not in the same order of magnitude, the prediction error of neural network will become larger, and even the network output value will fall into a certain interval, it is impossible for e(n) to carry out normal output and error feedback Therefore,

this paper normalizes the original data $f(x)$ to eliminate the data dimension. In this paper, map-minimax (max_x–min_x) is used to process the original data into data in [−1, 1] interval:

$$y = (\max y - \min y)\frac{x - \min x}{\max x - \min x} + \min y,$$

$$x = (\max x - \min x)\frac{y - \min y}{\max y - \min y} + \min x, \quad (5)$$

$$f(x) = \left(1 + e^{-ax}\right)^{-1},$$

$$f(x) = \frac{1 - e^{-x}}{1 + e^{-x}}, \ (-1 < f(x) < 1).$$

4. Multimodal Network Structure

The teacher-student model proposed in this article specifically uses a multimodal neural network structure based on text information such as audio and lyrics. Specifically, the input audio will be cut into N segments according to a fixed length. The model outputs the final sentiment label as N segments and averagely obtains its analysis results.

4.1. Audio Model. This article specifically uses the Mel spectrum information used in the previous article as the model data parameters. The Mel spectrum model is a two-dimensional matrix data with time dimension and characteristic length. Specifically, the time length feature_len = 1,024, and the feature length mel_dim = 128, so the number of input samples for the overall model is $1{,}024 \times 128$. In view of the relatively little emotional information in the lyrics, this paper compares and analyzes the model with the existing network analysis models such as ResNet, GoogLeNet, and VGGNet. Specifically, this paper designs a relatively suitable lightweight convolutional neural network system and then sorts out the network organization structure of the audio model, as shown in Figure 5. The convolutional neural network selected in this article has a certain feasibility, considering that many objects in life have local correlations. In particular, each frame of signal in a music piece is not isolated but is linked and interacts with others, so as to effectively convey the music information to the listener.

In order to prevent words with strong emotional inclinations and less frequent occurrences, they are filtered due to the low results of the above formula. Therefore, improvements are needed for such words. This article uses synonyms to prevent such words from being eliminated. The specific idea is to find the similar words to the words in the vocabulary and add them to the vocabulary after creating a vocabulary of word emotion recognition. The process flow of multifeature fusion of audio information and lyric text is shown in the Figure 5. Preprocessing is required before feature extraction. The preprocessing of audio information mainly includes signal filtering. The preprocessing of lyrics text is the operation of word emotion recognition vocabulary and supplementing similar words. After preprocessing, feature extraction is performed separately; that is, feature fusion and output are performed. Aiming at the feature vector data of music audio and lyrics, this paper improves the traditional forward neural network. Using specific Chebyshev orthogonal polynomial clusters, a forward neural network model with a specific structure is constructed. In this model, the forward neural network uses a single hidden layer to reduce the complexity of the overall model.

The activation function of each neuron in the hidden layer uses each function in Chebyshev's orthogonal polynomial cluster, while the activation functions of the neurons in other layers of the model all use linear activation functions. The weight of each neuron in the hidden layer and the output layer is w. The specific structure of the model is shown in the figure below. Emotion recognition classification can effectively improve the quality of music retrieval services and has important practical significance. In this paper, the Chebyshev polynomial cluster is incorporated into the forward neural network model, and the recognition of higher music emotions is achieved under the supervision and training of the gradient descent learning algorithm. It is noted that the accuracy of music sentiment classification is not only related to the number of layers of the model and the number of neurons, but also related to the number of dimensions of training samples.

4.2. Text Model. This paper focuses on screening the song's lyrics information, song title information, album style, and other characteristic information. Specifically, the maximum text length max_lengh is set to 200, the word vector dimension parameter word_dim is set to 128, and the word vector matrix-embedding matrix of the lyrics text is set to 200×128 matrix data. Specifically, the initial value of embedding matrix refers to the above, and the matrix data of the word vector is further incorporated into the model to train the stability and recognition accuracy of the model. The network architecture in this model is shown in Figure 6.

4.3. Multimodal Model. The text uses a multimodal model to combine the text model and the individual audio and then remove the respective full link parts. On this basis, the text model and the audio model are combined and merged, and then the emotional label prediction results are divided into N equal parts. The specific structure mode of the model is shown in Figure 6.

5. Experimental Results and Analysis

Considering that the previous article has analyzed the existing methods of multimodal emotion recognition, compared with traditional single audio or lyrics methods, the multimodal emotion model used in this article has significant advantages. In addition, the experiments in this article do not pay attention to the actual comparison of these two types of methods. The model in this paper mainly verifies the effect of knowledge distillation and transfer learning algorithms on existing networks under the

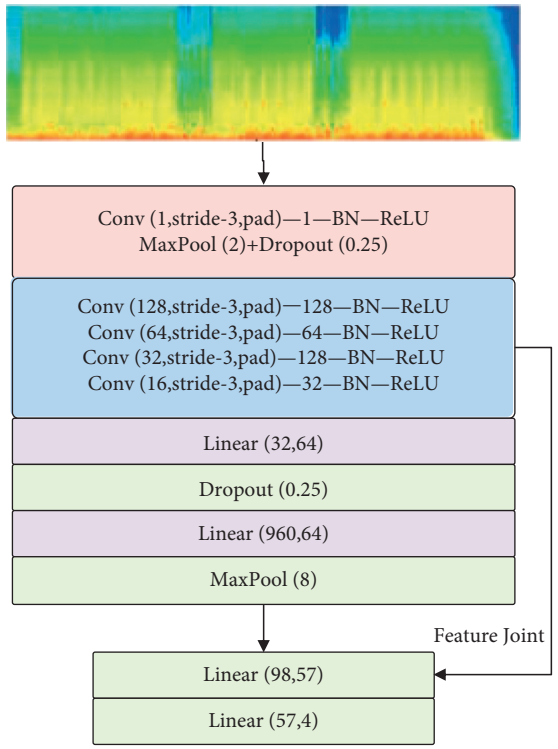


Figure 5: Audio model network.

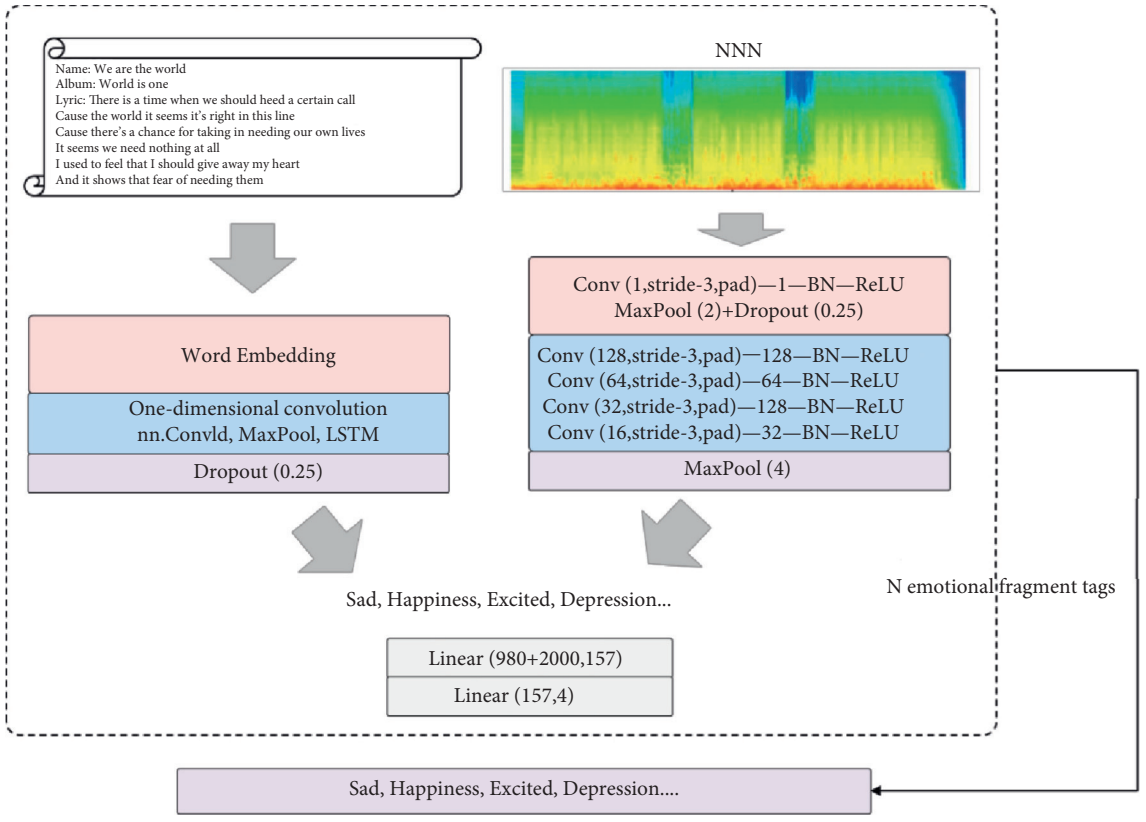


Figure 6: Multimodal network structure.

condition of missing or unbalanced sentiment labeled data. The data set used in the experiment contains about 2,000 English sentiments with or without annotated emotions.

5.1. Introduction to Related Data Sets. The data set used in this experiment contains about 2,000 English songs. In the repertoire, there are about 87% of English songs and about 13% of songs in other words. In addition, the data set comprehensively selects 15,000 songs as the specific training set and 5,000 songs as the test data. The emotion of the song is subdivided into four types: sad, lyrical, happy, and not sad or not happy. The music emotion used in this article expresses a relatively discrete model, and each category is mainly labeled with mutually exclusive discrete states. The number of songs in each category in the training set is 6,000 (2,000 labeled; 4,000 unlabeled), 7,000 (4,000 labeled; 3,000 unlabeled), and 2,000 (800 labeled; 1,200 unlabeled). Among them, the audio uses a unified WAV format, the sampling frequency is 2,205 Hz, and the monophonic channel is sampled. The music data was labeled by 5 people with more than 5 years of music education experience, and the samples of more than 3 people were included in the sample database. The data labeling of the training set is shown in Figure 7.

5.2. Comparison of Model Methods. Experiment 1, experiment 2, and experiment 3 compare and analyze the fitting results of different models. Specifically, this paper compares the results produced by two experimental methods based on knowledge distillation and nonknowledge distillation Considering that in different experimental environments, the number of labeled data items in the knowledge distillation method is small and unbalanced, the recognition accuracy of its training set and test set has shown a significant improvement in different models. The experimental results are shown in the figure below. The results indicate that using the teacher-student model combined with the knowledge distillation method for analysis, the teacher network can guide the student network to learn the answers consistent with the teacher network from the marked and unmarked emotional data. At the same time, the teacher network model is relearned according to the model parameters completed in different stages of the student network training, and the best model parameters are obtained after continuous iteration. The research results show that the accuracy of the multimodal method combined with the knowledge distillation method is greatly improved compared with the uncombined multimodal method in Figures 8 and 9.

In experiment 2, a comparative analysis of the knowledge distillation multimodal method combined with genre transfer learning is conducted.

Furthermore, the multimodal method combining only knowledge distillation and the model method proposed are compared and analyzed in this paper. Overall, this paper adopts the knowledge distillation multimodal method for song transfer learning and compares the model learning results with traditional multimodal learning methods. The experimental results are shown in Figure 10. The results show that when the initial values of the parameters of the model are selected, we further substitute the data into the model to analyze whether the local optimal solution can be obtained. In the traditional case, the model may not be able to analyze the global optimal solution. Existing experimental analysis further shows that the use of this paper's genre transfer algorithm can identify the key characteristics of music emotions and perform transfer learning. The model has high generalization adaptability. The initial values of the model parameters generally affect whether the model falls into a local optimal solution and even cause the model to be unable to solve the global optimal solution.

Through experimental analysis, it can be found that the genre characteristics of songs can be well identified in the model. Further transfer learning in the process of music emotion model recognition can show the model's good generalization ability. At the same time, the model achieves the global optimal model effect very well and converges to a higher accuracy rate in the case of fewer iterations in Figures 11 and 12.

Experiment 3 includes a comparison between the lightweight convolutional neural network and the classic network in the audio model selected in this article.

This paper avoids the use of words with strong emotional tendencies, so as to avoid biased analysis results. improvements are needed for such words. This article uses synonyms to prevent such words from being eliminated. The specific idea is to find the similar words to the words in the vocabulary and add them to the vocabulary after creating a vocabulary of word emotion recognition. The process flow of multifeature fusion of audio information and lyric text is shown in the figure below. Preprocessing is required before feature extraction. The preprocessing of audio information mainly includes signal filtering. The preprocessing of lyrics text is the operation of word emotion recognition vocabulary and supplementing similar words. After preprocessing, feature extraction is performed separately; that is, feature fusion and output are performed. Aiming at the feature vector data of music audio and lyrics, this paper improves the traditional forward neural network. Using specific Chebyshev orthogonal polynomial clusters, a forward neural network model with a specific structure is constructed. In this model, the forward neural network uses a single hidden layer to reduce the complexity of the overall model.

This paper compares the selected lightweight convolutional neural network with the classic network, mainly to illustrate that under the actual data set, choosing a suitable design of the lightweight convolutional neural network is better than the generalization ability brought by the classic network promotion. This article compares the two without combining knowledge distillation and transfer learning to ensure the independence of experimental comparison. The experimental results are shown in Figure 13. The method of using the classic ResNet network on the training set shows a higher accuracy rate, but on the test set its accuracy rate drops more. According to this conclusion, the model shows that there is an overfitting phenomenon under the ResNet network model, and it cannot have the generalization ability on the entire data set in Figures 14 and 15.

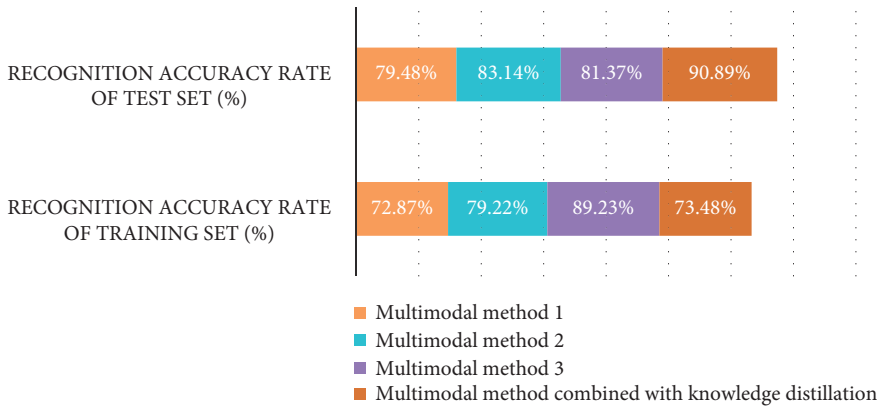


Figure 7: Statistical table of data labeling of control group set.

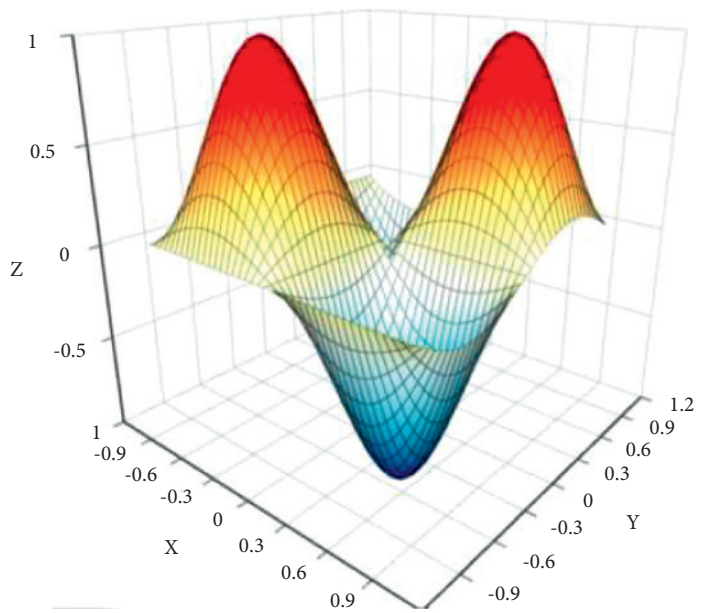


Figure 8: Result statistics of experiment 1.

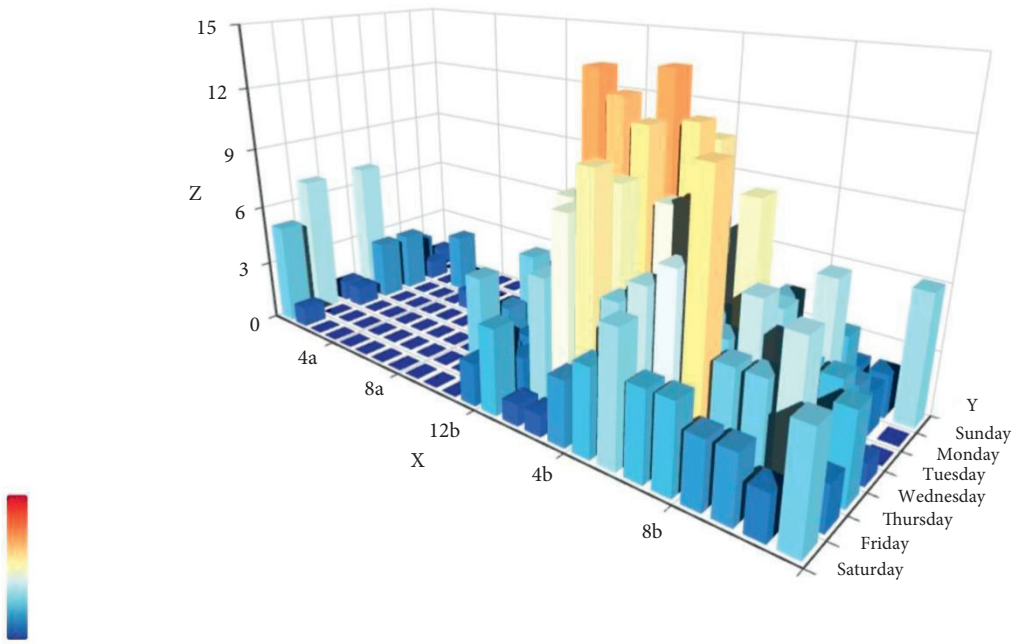


Figure 9: Result statistics of multimodal music emotion recognition in experiment 2 in a week.

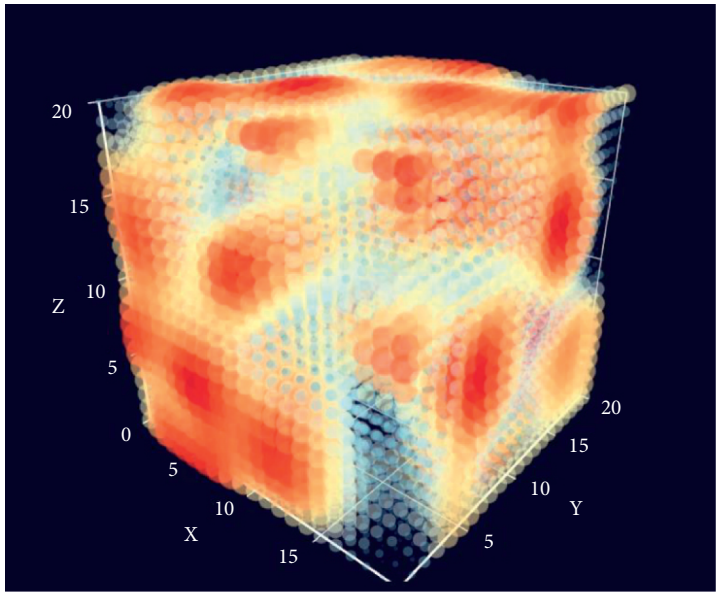


Figure 10: Result statistics of multimodal music emotion recognition in experiment 3.

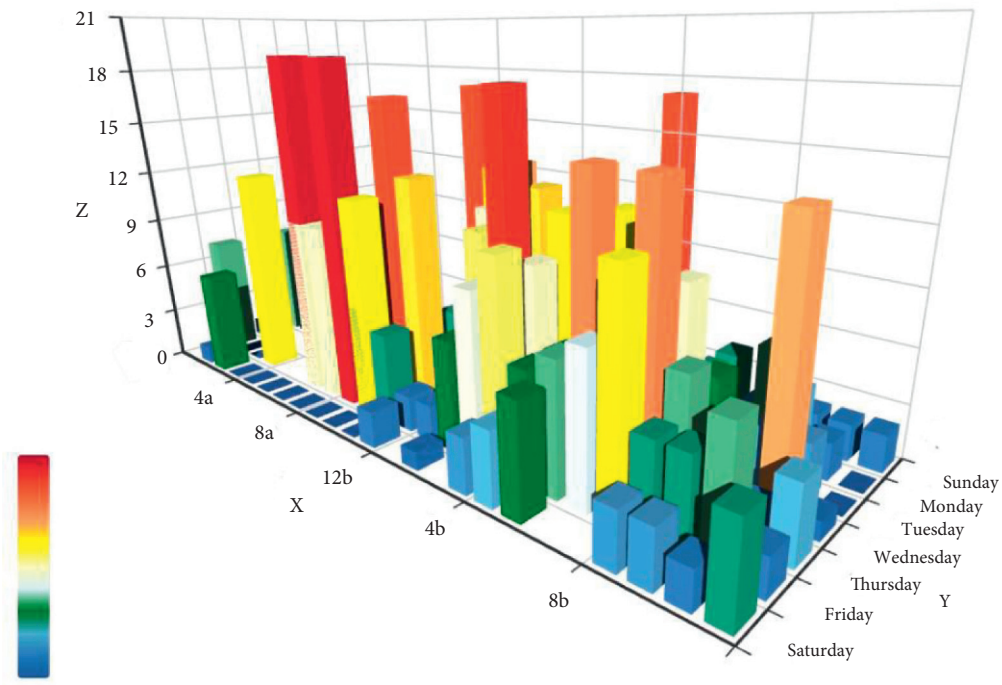


Figure 11: Result statistics of multimodal music emotion recognition in experiment 4 in a week.

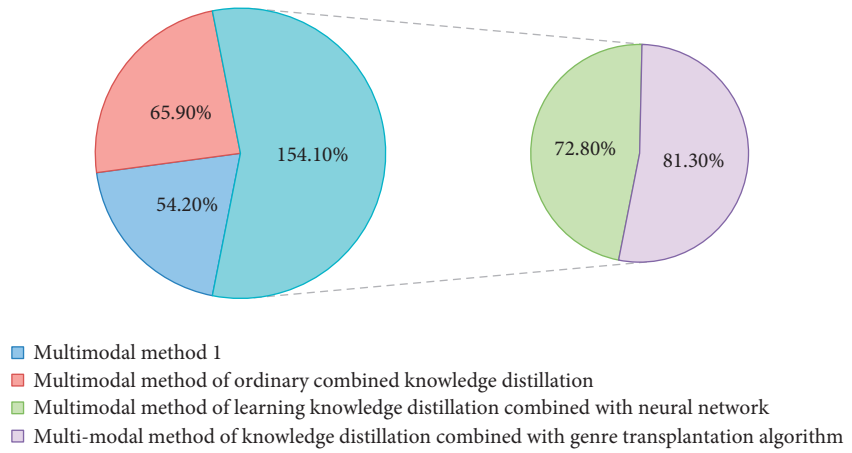


Figure 12: Comparison of conclusions of different data models.

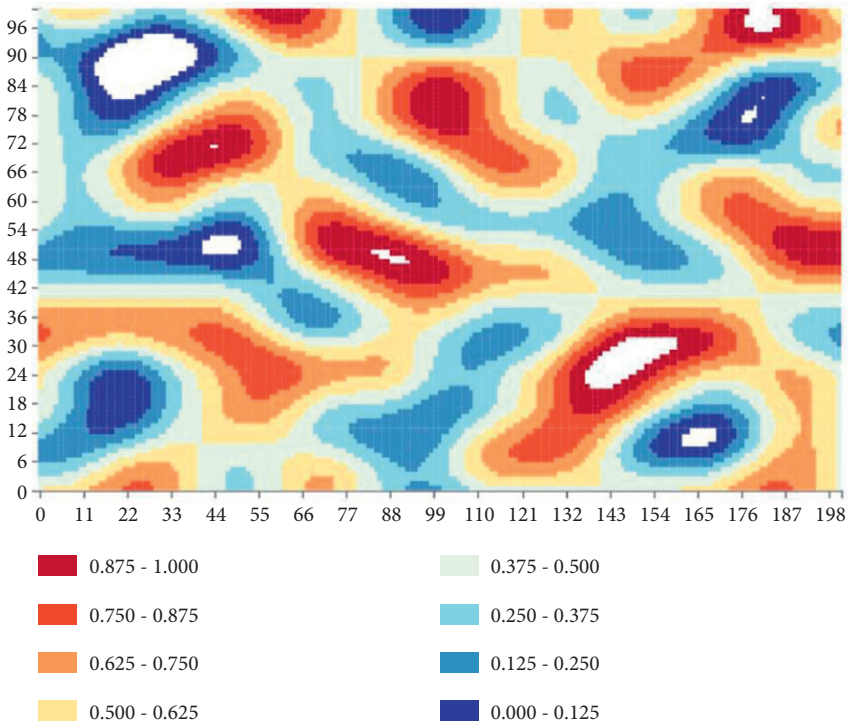


Figure 13: Experimental comparison results using different convolutional networks.

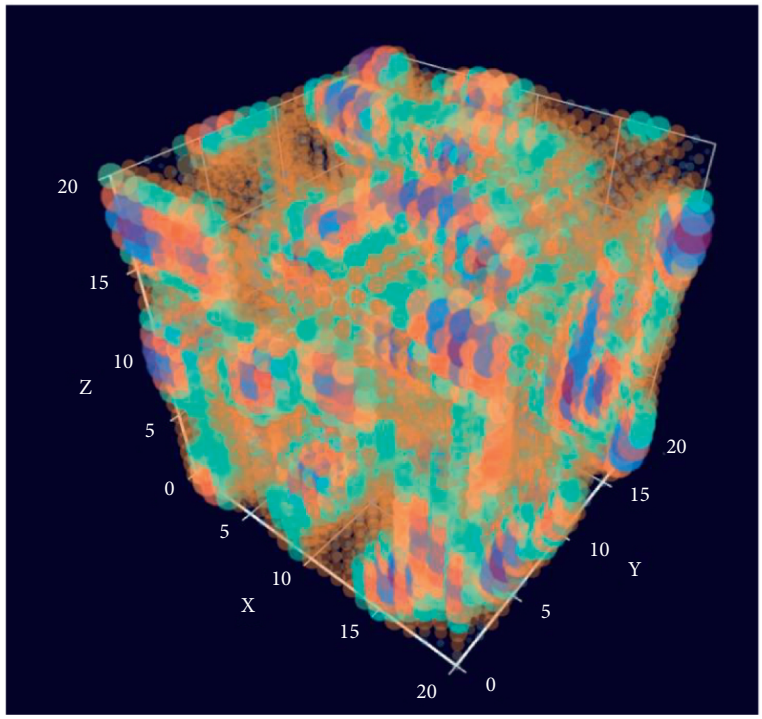


Figure 14: Result statistics of multimodal music emotion recognition in experiment 4.

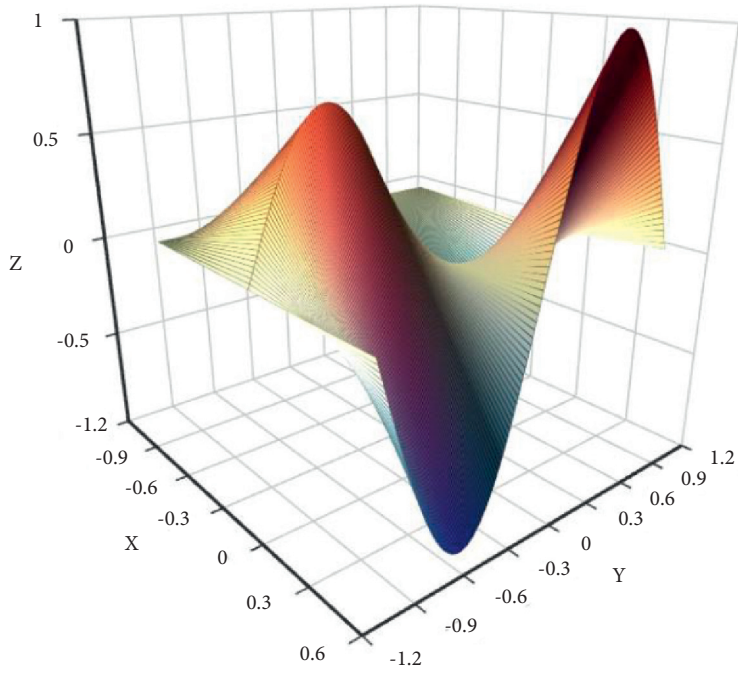


Figure 15: The results based on the combination of knowledge distillation.

6. Conclusion

This paper proposes a multimodal method based on the combination of knowledge distillation and music style transfer learning. This research uses the knowledge distillation teacher-student model to unearth the consistent relationship between the unlabeled and labeled emotional data between the teacher model and the student model. In addition, this study improves the accuracy of the model when there are less sentiment labeled data and unbalanced categories. At the same time, this article uses big data and machine learning analysis methods to make the model learn the characteristics of song genre, which characterizes the transfer learning process in music emotion recognition. In

the case of a small number of iterations, the overall accuracy is higher, and the convergence speed in the training process is improved. The research results prove that the relevant features in the audio field have good effects in emotion recognition tasks. Combining the needs of the research conclusions of this article, two issues need to be explored in the next step of the research. The first is to analyze the ambiguity of sentiment labels in different subjects and analyze the instability of training and testing caused by the ambiguity. The second is to study the influence of big data migration learning for different tasks on the effect of emotion recognition, such as the influence of visual tasks on the effect of emotion recognition.

Data Availability

The data set can be accessed upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This work was supported by the research project on 2021 teaching quality and reform of Huizhou University, name: Research on Guzheng Teaching Mode Reform Based on Application-Oriented Talent Cultivation, no. X-JYJG2021070.

References

[1] M. Doyran, A. Schimmel, P. Baki et al., "Mumbai: multi-person, multimodal board game affect and interaction analysis dataset," *Journal on Multimodal User Interfaces*, vol. 15, no. 4, pp. 373–391, 2021.

[2] M. C. Camacho, E. M. Williams, K. Ding, and S. B. Perlman, "Multimodal examination of emotion processing systems associated with negative affectivity across early childhood," *Developmental Cognitive Neuroscience*, vol. 48, no. 1, pp. 915–917, 2021.

[3] D. Liu, L. Chen, Z. Wang, and G. Diao, "Speech expression multimodal emotion recognition based on deep belief network," *Journal of Grid Computing*, vol. 19, no. 2, pp. 59–72, 2021.

[4] T. Chen, H. Yin, X. Yuan, Y. Gu, F. Ren, and X. Sun, "Emotion recognition based on fusion of long short-term memory networks and SVMs," *Digital Signal Processing*, vol. 117, no. 1, pp. 137–153, 2021.

[5] D. S. Cortes, C. Tornberg, T. Bnziger, H. A. Elfenbein, H. Fischer, and P. Laukka, "Effects of aging on emotion recognition from dynamic multimodal expressions and vocalizations," *Scientific Reports*, vol. 11, no. 1, 2021.

[6] M. Dahmane, J. Alam, P.-L. St-Charles, M. Lalonde, K. Heffner, and S. Foucher, "A multimodal non-intrusive stress monitoring from the pleasure-arousal emotional dimensions," *IEEE Transactions on Affective Computing*, vol. 99, p. 1, 2020.

[7] L. Schoneveld, A. Othmani, and H. Abdelkawy, "Leveraging recent advances in deep learning for audio-Visual emotion recognition," *Pattern Recognition Letters*, vol. 146, pp. 1–7, 2021.

[8] S. Hans-Eckhardt, "Music-evoked emotions—current studies," *Frontiers in Neuroscience*, vol. 11, p. 600, 2017.

[9] M. Tzelepi, N. Passalis, and A. Tefas, "Online subclass knowledge distillation," *Expert Systems with Applications*, vol. 181, pp. 115–132, 2021.

[10] G. Valle-Pérez, H. Enter G E, J. Beskow, A. Holzapfel, P. Y. Oudeyer, and S. Alexanderson, "Transflower: probabilistic autoregressive dance generation with multimodal attention," *ACM Transactions on Graphics*, vol. 40, no. 6, pp. 1–14, 2021.

[11] A. Del Ponte, P. Canofari, and A. De Dominicis, "Financial and trade relationships between the Eurozone and China in the age of resilience," *Asia Europe Journal*, vol. 19, no. 4, pp. 489–506, 2021.

[12] J. Dong, L. Yin, X. Liu, M. Hu, X. Li, and L. Liu, "Impact of internet finance on the performance of commercial banks in China," *International Review of Financial Analysis*, vol. 72, no. 12, pp. 1–12, 2020.

[13] B. Zhao, K. Kenjegalieva, J. Wood, and A. Glass, "A spatial production analysis of Chinese regional banks: case of urban commercial banks," *International Transactions in Operational Research*, vol. 27, no. 4, pp. 2021–2044, 2019.

[14] C. Zheng, S. Chen, and Z. Dong, "Economic fluctuation, local government bond risk and risk-taking of city commercial banks," *Sustainability*, vol. 13, no. 17, pp. 1–26, 2021.

[15] A. G. Assaf, A. N. Berger, R. A. Roman, and M. G. Tsionas, "Does efficiency help banks survive and thrive during financial crises?" *Journal of Banking & Finance*, vol. 106, no. 7, pp. 445–470, 2019.

[16] O. Badunenko and S. C. Kumbhakar, "Economies of scale, technical change and persistent and time-varying cost efficiency in Indian banking: do ownership, regulation and heterogeneity matter?" *European Journal of Operational Research*, vol. 260, no. 2, pp. 789–803, 2017.

[17] S. Kumar Dwivedi, R. Amin, S. Vollala, and R. Chaudhry, "Blockchain-based secured event-information sharing protocol in internet of vehicles for smart cities," *Computers & Electrical Engineering*, vol. 86, no. 1, pp. 1–9, 2020.

[18] Z. Khan and S. Amin, "Bottleneck model with heterogeneous information," *Transportation Research Part B: Methodological*, vol. 112, no. 1, pp. 157–190, 2018.

[19] J. M. Cairney, K. Rajan, D. Haley et al., "Mining information from atom probe data," *Ultramicroscopy*, vol. 159, no. 1, pp. 324–337, 2020.

[20] Z. Wang, H. Ren, Q. Shen, W. Sui, and X. Zhang, "Seismic performance evaluation of a steel tubular bridge pier in a five-span continuous girder bridge system," *Structures*, vol. 31, no. 1, pp. 909–920, 2021.

[21] S. Nakayama and J. Takayama, "Traffic network equilibrium model for uncertain demands," in *Proceedings of the 82nd Transportation Research Board Annual Meeting*, Washington, DC, USA, January 2021.

[22] H. Shao, W. H. K. Lam, and M. L. Tam, "A reliability-based stochastic traffic assignment model for network with multiple user classes under uncertainty in demand," *Networks and Spatial Economics*, vol. 6, no. 3, pp. 173–204, 2019.

[23] A. Chen, J. Kim, S. Lee, and Y. Kim, "Stochastic multi-objective models for network design problem," *Expert Systems with Applications*, vol. 37, no. 2, pp. 1608–1619, 2020.

[24] S.-M. Hosseininasab, S.-N. Shetab-Boushehri, S. R. Hejazi, and H. Karimi, "A multi-objective integrated model for selecting, scheduling, and budgeting road construction projects," *European Journal of Operational Research*, vol. 271, no. 1, pp. 262–277, 2018.

[25] M. Johnson, M. Schuster, Le et al., "Google's multilingual neural machine translation system: enabling zero-shot translation," *Transactions of the Association for Computational Linguistics*, vol. 5, no. 1, pp. 339–351, 2017.
[26] M. D. Moreno, "Translation quality gained through the implementation of the iso en 17100:2015 and the usage of the blockchain," *Babel*, vol. 1, no. 2, pp. 1–9, 2020.
[27] X. Wang, X. Yu, L. Guo, F. Liu, and L. Xu, "Student performance prediction with short-term sequential campus behaviors," *Information*, vol. 11, no. 4, p. 101, 2020.
[28] Q. Guo, Z. Zhu, Q. Lu, D. Zhang, and W. Wu, "A dynamic emotional session generation model based on Seq2Seq and a dictionary-based attention mechanism," *Applied Sciences*, vol. 10, no. 6, pp. 1–10, 2020.
[29] H. Ren, X. Mao, W. Ma, J. Wang, and L. Wang, "An English-Chinese machine translation and evaluation method for geographical names," *ISPRS International Journal of Geo-Information*, vol. 9, no. 3, pp. 193–201, 2020.
[30] J. Arús-Pous, T. Blaschke, S. Ulander, J.-L. Reymond, H. Chen, and O. Engkvist, "Exploring the GDB-13 chemical space using deep generative models," *Journal of Cheminformatics*, vol. 11, no. 1, pp. 20–29, 2019.
[31] T. Moon, T. In Ahn, and E. Son, "Long short-term memory for a model-free estimation of macronutrient ion concentrations of root-zone in closed-loop soilless cultures," *Plant Methods*, vol. 15, no. 1, pp. 1–12, 2019.
[32] N. Pourdamghani and K. Knight, "Neighbors helping the poor: improving low-resource machine translation using related languages," *Machine Translation*, vol. 33, no. 3, pp. 239–258, 2019.
[33] L. Bote-Curiel, S. Muñoz-Romero, A. Gerrero-Curieses, and J. Luis Rojo-Álvarez, "Deep learning and big data in healthcare: a double review for critical beginners," *Applied Sciences*, vol. 9, no. 11, pp. 1–11, 2019.
[34] J. Zhang and T. Matsumoto, "Corpus augmentation for neural machine translation with Chinese-Japanese parallel corpora," *Applied Sciences*, vol. 9, no. 10, pp. 1–12, 2019.
[35] Y. Chen, Y. Ma, X. Mao, and Q. Li, "Multi-task learning for abstractive and extractive summarization," *Data Science and Engineering*, vol. 4, no. 1, pp. 14–23, 2019.
[36] X. Xiao, "Analysis on the employment psychological problems and adjustment of retired athletes in the process of career transformation," *Modern Vocational Education*, vol. 5, no. 12, pp. 216-217, 2018.
[37] Y. Zhou and B. Yang, "Sports video athlete detection using convolutional neural network," *Journal of Natural Science of Xiangtan University*, vol. 39, no. 1, pp. 95–98, 2017.
[38] J. Pang, "Research on the evaluation model of sports training adaptation based on self-organizing neural network," *Journal of Nanjing Institute of Physical Education*, vol. 16, no. 1, pp. 74–77, 2017.

Hindawi
Scientific Programming
Volume 2022, Article ID 1395265, 11 pages
https://doi.org/10.1155/2022/1395265

Research Article

Research on Toy Design for Special Children Based on Sensory Integration Training, D-S Theory, and Extenics: Taking Physical Toys for ADHD Children as an Example

Shilin Wu

School of Fine Arts & Design, Guangzhou University, Guangzhou, Guangdong 511400, China

Correspondence should be addressed to Shilin Wu; wslw@gzhu.edu.cn

Received 13 December 2021; Revised 31 December 2021; Accepted 8 January 2022; Published 11 February 2022

Academic Editor: Baiyuan Ding

Based on sensory integration training, D-S theory and extenics, this paper discusses the innovation and development of ADHD children's toys, so as to make the design scheme more real, scientific, and comprehensive and meet the internal needs of users. Considering the advantages of extenics and the D-S theory in dealing with uncertain and incompatible decision-making problems, an evaluation method of toy design scheme for special children based on extension analysis and the D-S theory is proposed. Through the combination of the D-S theory and extenics evaluation method and sensory integration training theory, the designed toy scheme has a certain auxiliary therapeutic effect on the rehabilitation of ADHD children. Through the design case of physical fitness for ADHD children, it is proved that the combination of the D-S theory and extenics based on the sensory integration training theory has a certain reference value for the generation of product scheme and makes the product design process more comprehensive and objective.

1. Introduction

With the progress of bioengineering, medical technology, electronic equipment, and other industries, more and more diseases have been found and paid attention to. Attention deficit hyperactivity disorder (ADHD) is one of the most common mental diseases in children [1–4]. According to domestic media reports, the incidence rate of ADHD in childhood is 2.59%~7.25%. Children with ADHD have lower basic abilities than normal children, such as learning ability, executive ability, activity ability, and social communication ability. Childhood is the main period of people's learning. If they are not treated in time, it will affect their adulthood and even their whole life [5, 6].

Physical toy design itself has the characteristics of diversity, complexity, and experience, which makes the evaluation of physical toy design scheme have obvious uncertainty and incompatibility. Specifically, the evaluation of physical toy design scheme is essentially a multicriteria decision-making (MCDM) problem [7–11], which needs to comprehensively consider many factors. The importance of each evaluation criterion has certain uncertainty, and the process of determining the importance of each evaluation criterion by expert evaluation method has obvious subjectivity and fuzziness, Therefore, the evaluation process of physical toy design scheme is actually an uncertain reasoning process with the characteristics of imprecision, fuzziness, and subjectivity. At the same time, there is a certain degree of incompatibility between multiple criteria for the evaluation of physical toy design scheme. For example, the selection of high-quality materials may not only significantly improve the quality of toys but also lead to the increase of cost and the reduction of economy. When the functional design is more complex, it meets the functional requirements, but may lead to the decline of security.

Extenics [12] is a new subject spanning philosophy, mathematics, and engineering. As the key application method of extenics, the extension analysis method takes the matter-element theory and the extension set theory as the theoretical framework [13–15]. Its basic idea is to establish the classical domain, node domain, and evaluation level of things, calculate the correlation function of the matter-

element to be evaluated about the evaluation level according to the actual data, and quantitatively and objectively describe the degree to which the matter-element to be evaluated belongs to a certain evaluation level through the correlation function, The levels of different matter elements to be evaluated can be divided according to the large cell of the correlation function. It can be seen that extension analysis can deal with the incompatibility in evaluation and provides a new way for thing classification and pattern recognition. The Dempster Shafer theory (D-S theory) [16–18] can directly express uncertainty and provides an effective method for the expression and synthesis of uncertain information, but its evidence is mainly obtained through expert knowledge and experience, which is highly subjective. Extension analysis can effectively transform the contradictory problems existing in things into compatible ones, so as to reduce the conflict between various evidences and optimize the results of evidence fusion to the greatest extent.

At present, the treatment of ADHD children mainly focuses on behavioral intervention, rehabilitation training, and drugs, but the repeatability of behavioral intervention and the side effects of drugs lead to poor treatment effect, which promotes the gradual development of physical training methods of sensory integration [19]. Because of its advantages of safety, simplicity, and low cost, the sensory integration training is accepted and welcomed by many parents. In this paper, according to the characteristics of the sensory integration training theory, the correlation function matrix of the matter element to be evaluated is obtained by using the idea of extension analysis; after normalization, the discrimination framework of the D-S theory is established, and the basic probability assignment (BPA) function on the discrimination framework of the D-S theory is obtained. The critical method considering the contrast intensity and conflict is used to calculate the importance of each feature. Considering the importance of features, the evidence is fused. Finally, according to the nature of BPA function, the grading results of multiple physical toy design schemes for ADHD children and the ranking relationship of the same level schemes are obtained, so as to provide reference for the physical toy design of ADHD children.

2. Basic Concepts

2.1. Sensory Integration Training. Sensory integration (SI) is a process of unifying neuropsychology after connecting and selecting sensory information from various senses such as human vision, hearing, touch, proprioception, and vestibular perception. It is the basis of human life, learning, and work. Sensory integration training (SIT) refers to planned training activities to reduce sensory integration disorders and their negative impact on individual learning and life and to improve individual sensory integration ability [19–21]. SIT is divided into three areas: the tactile training, the vestibular function training, the and proprioception function training. The tactile training includes thermal training, pressure training, and perception training; the vestibular training includes balance ability, spatial perception abilityv and concentration ability; and the proprioception training includes joint static perception, joint dynamic perception, and muscle response [22].

2.2. Physical Toy Design for ADHD Children. Toys are closely related to children's healthy growth. They are a good spiritual partner in children's life and will accompany children through the purest time of their life. For ADHD children, toys are not only their friendly playmates but also can be treated during the game. ADHD children mainly have several characteristics: large emotional changes; inattention; large amount of activities; inflexible large action, or precision action [23]. According to these characteristics, children can complete sensory integration training with the help of corresponding toys. At present, the main types of common integration training toys include tactile ball, tactile plate, balance cap, and other toys. Such toys are proprietary equipment for sensory integration, but they will inevitably feel boring after long-term use, especially for ADHD children with attention deficit symptoms. This single toy is difficult to attract the attention of ADHD children for a long time. It is necessary to design suitable children's physical toys according to the inherent characteristics of ADHD children, so that they can exercise in play and recover in training [23].

3. Toy Evaluation Model for ADHD Children Based on Analytic Hierarchy Process

3.1. Construction of Evaluation Hierarchy of Toy Design Indexes for ADHD Children. Through the research on the design of toys for children with ADHD, we can understand the way and purpose of sensory integration training for children with ADHD. After consulting experts and experienced toy designers and reviewing relevant literature, starting from the tactile training, preauditory training, and proprioception training of sensory integration training for ADHD children, the criterion layer of ADHD children's toys is constructed into three aspects: the vestibular training B1, the proprioception training B2, and the tactile training B3.

For B1, B2, and B3, nine subcriteria indicators are subdivided, including balance ability training C11, spatial perception training C12, focus ability training C13, joint static perception ability C14, joint dynamic perception ability training C15, muscle response ability C16, heat perception training C17, pressure perception training C18, and perceived weight training C19.

After the criteria indicators and subcriteria indicators are established, three criteria indicators are compared each time, and then nine subcriteria indicators are compared with each other. Finally, the analytic hierarchy process model of toy evaluation indicators for ADHD children is obtained, as shown in Figure 1.

3.2. Extension Analysis of Physical Toy Design Schemes for ADHD Children. According to the extension analysis theory [12–15], the matter elements of physical toy design scheme are defined as ordered triples $R = (N, I, \nu)$ as follows:

$$R = \begin{bmatrix} N & I_1 & \nu_1 \\ & I_2 & \nu_2 \\ & \vdots & \vdots \\ & I_n & \nu_n \end{bmatrix}, \quad (1)$$

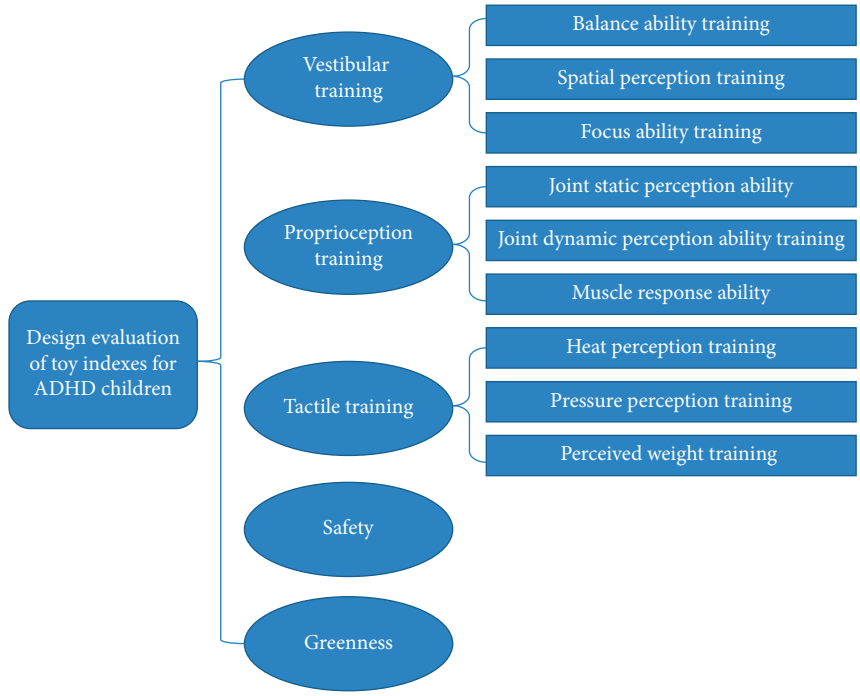


Figure 1: Analytic hierarchy process model of toy evaluation indicators for ADHD children.

where N stands for physical toy design scheme, $I = \{I_1, I_2, \ldots, I_n\}$ is the characteristic of physical toy design scheme, and $v = \{v_1, v_2, \ldots, v_n\}$ is the corresponding characteristic value.

The characteristics $I_1, I_2, \ldots, I_n$ of physical toy design scheme are divided into l evaluation levels, so that

$$R_j = \begin{bmatrix} N_j & I_1 & v_{j,1} \\ & I_2 & v_{j,2} \\ & \vdots & \vdots \\ & I_n & v_{j,n} \end{bmatrix} = \begin{bmatrix} N_j & I_1 & [a_{j,1}, b_{j,1}] \\ & I_2 & [a_{j,2}, b_{j,2}] \\ & \vdots & \vdots \\ & I_n & [a_{j,n}, b_{j,n}] \end{bmatrix}, \quad (2)$$

where R_j is the classical domain matter element of N, N_j represents the j-th ($j = 1, 2, \ldots, l$) evaluation level of N, and $v_{j,i} = [a_{j,i}, b_{j,i}]$ is the value interval of N_j corresponding to the characteristic I_i ($i = 1, 2, \ldots, n$), that is, the classical domain.

It is assumed that

$$R_p = \begin{bmatrix} N_p & I_1 & v_{p,1} \\ & I_2 & v_{p,2} \\ & \vdots & \vdots \\ & I_n & v_{p,n} \end{bmatrix}, = \begin{bmatrix} N_p & I_1 & [a_{p,1}, b_{p,1}] \\ & I_2 & [a_{p,2}, b_{p,2}] \\ & \vdots & \vdots \\ & I_n & [a_{p,n}, b_{p,n}] \end{bmatrix}, \quad (3)$$

where R_p is the joint domain matter element of N, N_p represents all evaluation level of N, and $v_{p,i} = [a_{p,i}, b_{p,i}]$ is the value interval of N_p corresponding to the characteristic I_i ($i = 1, 2, \ldots, n$), that is, the joint domain.

It is assumed that

$$R_0 = (N_0, I, x) = \begin{bmatrix} N_0 & I_1 & x_1 \\ & I_2 & x_2 \\ & \vdots & \vdots \\ & I_n & x_n \end{bmatrix}, \quad (4)$$

where R_0 is the matter element to be evaluated, N_0 is the unknown evaluation level of the matter element to be evaluated, and the value of R_0 corresponding to the characteristic I_i ($i = 1, 2, \ldots, n$) is x_i, $x = \{x_1, x_2, \ldots, x_n\}$ is the value vector.

After determining the classical domain matter element, joint domain matter element, and the matter element to be evaluated of physical toy design scheme, the relationship between classical domain matter element, joint domain matter element, and matter element to be evaluated can be quantified by using the correlation function in extenics.

Specifically, correlation function $k_j(x_i)$ describes the quantitative relationship between the value x_i in matter element to be evaluated R_0 and two intervals (classical domain $v_{j,i} = [a_{j,i}, b_{j,i}]$, and joint domain $v_{p,i} = [a_{p,i}, b_{p,i}]$) represents the degree to which R_0 belongs to the evaluation level N_j ($j = 1, 2, \ldots, l$) on feature I_i, and its calculation formula is

$$k_j(x_i) = \begin{cases} \dfrac{-D(x_i, v_{j,i})}{b_{j,i} - a_{j,i}}, & x_i \in v_{j,i}, \\ \dfrac{D(x_i, v_{j,i})}{D(x_i, v_{p,i}) - D(x_i, v_{j,i})}, & x_i \notin v_{j,i}, \end{cases} \tag{5}$$

where $D(x_i, v_{p,i}) = |x_i - (a_{p,i} + b_{p,i})/2| - (b_{p,i} - a_{p,i})/2$ and $D(x_i, v_{j,i}) = |x_i - (a_{j,i} + b_{j,i})/2| - (b_{j,i} - a_{j,i})/2$ represent the distance between x_i and $v_{j,i} = [a_{j,i}, b_{j,i}]$ and the distance between x_i and $v_{p,i} = [a_{p,i}, b_{p,i}]$, respectively.

Thus, the correlation function matrix of R_0 is obtained as $K_0 = (K_{ji})_{l\times n}$, here $K_{ji} = k_j(x_i)$.

3.3. Evaluation of Physical Toy Design Schemes for ADHD Children Based on the D-S Theory

3.3.1. Solution of BPA Function Based on Correlation Function. The concept of correlation function in the extension analysis theory extends the logical value from {0, 1} to (−∞, +∞). According to the size of the correlation function, the membership relationship between elements and sets can be judged, so that the qualitative description of either or between elements and sets can be extended to quantitative description, which can more comprehensively and accurately characterize the relationship between elements in the set. When the correlation function is greater than 0, it indicates that the element has a certain property, and the larger the value is, the closer it is to the property; when the correlation function is less than 0, it means that the element does not have this property, and the smaller the value, the farther away from this property; when the correlation function is equal to 0, it indicates that the element may or may not have this property, which is a critical element. Therefore, the correlation function can be extended to the BPA function in the D-S theory, that is, the larger the correlation function is, the larger the value of BPA converted is, on the contrary, the smaller the value of BPA converted. In addition, the range of BPA assignment in the D-S theory is [0, 1], so normalization is required.

It can be seen that the matter-element concept of extension analysis theory quantitatively and objectively describes the degree of elements belonging to a certain property through the correlation function and can distinguish different levels of elements in the same domain according to the large cell of the correlation function, which provides a new method for the classification and pattern recognition of things. Using the classification idea of extension analysis, this paper establishes the D-S theoretical discrimination framework of physical toy design scheme evaluation and obtains the BPA function on the discrimination framework through normalization. The details are as follows:

First, a D-S theoretical discrimination framework [16–18] for physical toy design scheme evaluation is established as $\Theta = \{\theta_1, \theta_2, \ldots, \theta_l\}$, where $\theta_1, \theta_2, \ldots, \theta_l$ indicate the evaluation levels of physical toy design scheme in turn. The BPA function on discrimination framework Θ is represented by $m_i(\theta_j)$, where $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, l$.

Second, correlation function matrix $K_0 = (K_{ji})_{l\times n}$ of matter element to be evaluated R_0 is transformed into BPA function matrix $m = (m_{ji})_{l\times n}$ on the discrimination framework, that is,

$$m_{ji} = m_i(\theta_j) = \frac{e^{k_j(x_i)}}{\sum_{j=1}^{l} e^{k_j(x_i)}}. \tag{6}$$

As can be seen from equation (6), the value of $e^{k_j(x_i)}$ increases with the increase of $k_j(x_i)$, $0 \le m_i(\theta_j) \le 1$, $\sum_{j=1}^{l} m_i(\theta_j) = 1$. When $k_j(x_i) \longrightarrow +\infty$, $m_i(\theta_j) = 1$; when $k_j(x_i) \longrightarrow -\infty$, $m_i(\theta_j) = 0$. It can be seen that equation (6) can realize the transformation between the correlation function of extension analysis and the BPA function of the D-S theory, so as to solve the BPA function on discrimination framework Θ.

3.3.2. Calculation of Feature Importance Based on CRITIC Method. In multifeature evaluation and decision-making problems, the common calculation methods of feature importance include entropy weight method, standard deviation method, CRITIC method, and so on. Compared with the entropy weight method and the standard deviation method, the CRITIC method comprehensively considers the contrast strength and conflict between features and can more completely reflect the competitive relationship between features [24–26]. Here, the CRITIC method is selected to calculate the importance of features.

For n features of the physical toy design scheme , each feature has l different evaluation states (i.e. l evaluation levels). The contrast intensity in CRITIC method is expressed in the form of standard deviation , so the standard deviation of feature i ($i = 1, 2, \ldots, l$) is as follows :

$$\sigma_i = \sqrt{\frac{1}{l-1} \sum_{j=1}^{l} \left(m_{ji} - \frac{1}{l} \sum_{j=1}^{l} m_{ji} \right)^2}. \tag{7}$$

The conflict is based on the correlation between the two features. If there is a strong positive correlation, the conflict between the two features is low. For two features i and h (i, $h = 1, 2, \ldots, n$ and $i \ne h$), their correlation coefficient is

$$\psi_{ih} = \frac{\sum_{j=1}^{l} \left(m_{ji} - 1/l \sum_{j=1}^{l} m_{ji} \right) \left(m_{jh} - 1/l \sum_{j=1}^{l} m_{jh} \right)}{\sqrt{\sum_{j=1}^{l} \left(m_{ji} - 1/l \sum_{j=1}^{l} m_{ji} \right)^2} \sqrt{\sum_{j=1}^{l} \left(m_{jh} - 1/l \sum_{j=1}^{l} m_{jh} \right)^2}}. \tag{8}$$

Therefore, the conflict between feature i and other features can be expressed as

$$\xi_i = \sum_{h=1, h\ne i}^{n} (1 - \psi_{ih}). \tag{9}$$

At last, the importance of feature i is

$$\omega_i = \frac{\delta_i}{\sum_{i=1}^{n} \delta_i}, \tag{10}$$

where $\delta_i = \sigma_i \cdot \xi_i$ represents the amount of information contained in feature i.

Thus, the importance vector of n features is obtained as $\omega = [\omega_1, \omega_2, \ldots, \omega_n]$.

3.3.3. Evidence Fusion considering Feature Importance. The classical D-S theory holds that each evidence is equally important in evidence fusion, but in practice, each evidence has different importance. In other words, in the evaluation of physical toy design scheme, the importance of each feature is different, so the role of participating in evidence fusion is also different. Therefore, this paper introduces feature importance to make the result of evidence fusion more reasonable.

From the above, there are n features in the physical toy design scheme evaluation, and each feature has l states. The proposition elements in framework $\Theta = \{\theta_1, \theta_2, \ldots, \theta_l\}$ represent different evaluation states of the feature, then the BPA function on Θ is m: $m_1, m_2, \ldots, m_n$.

According to feature importance vector $\omega = [\omega_1, \omega_2, \ldots, \omega_n]$, the relative importance of feature i ($i = 1, 2, \ldots, n$) is

$$\mu_i = \frac{\omega_i}{\max\{\omega_1, \omega_2, \ldots, \omega_n\}}. \tag{11}$$

Then, after considering the feature importance, the BPA function equation (6) on discrimination framework Θ of physical toy design scheme evaluation is improved as: when $\theta_j \neq \Theta$, $\widetilde{m}_i(\theta_j) = \mu_i m_i(\theta_j)$; when $\theta_j = \Theta$, $\widetilde{m}_i(\theta_j) = 1 - \sum_{j=1}^{l} \widetilde{m}_i(\theta_j)$.

According to the idea of the D-S theory, for the discrimination framework $\Theta = \{\theta_1, \theta_2, \ldots, \theta_l\}$ of physical toy design scheme evaluation, all possible set in Θ are represented by power set 2^{Θ} . Obviously, there are l elements in Θ, and each element is incompatible with each other, so the number of elements in the power set 2^{Θ} is 2^l.

Let A be the evaluation conclusion of physical toy design scheme. A can be one of $\theta_1, \theta_2, \ldots, \theta_l$ (indicating that the physical toy design scheme to be evaluated belongs to this level) or several of them (indicating that the physical toy design scheme to be evaluated belongs to these levels). For $\forall A \subseteq \Theta$, the Dempster fusion rule for the finite BPA functions $\widetilde{m_1}, \widetilde{m_2}, \ldots, \widetilde{m_n}$ on the framework Θ is

$$M(A) = \frac{1}{K} \sum_{A_1 \cap A_2 \cap \ldots \cap A_n = A} \widetilde{m_1}(A_1) \cdot \widetilde{m_2}(A_2) \cdot \ldots \cdot \widetilde{m_n}(A_n), \tag{12}$$

where M is the BPA function after fusion and $M = \widetilde{m_1} \oplus \widetilde{m_2} \oplus \ldots \oplus \widetilde{m_n}$, here $\oplus$ 为is the evidence fusion symbol, K is the normalized constant and $K = \sum_{A_1 \cap A_2 \cap \ldots \cap A_n \neq \varnothing} \widetilde{m_1}(A_1) \cdot \widetilde{m_2}(A_2) \cdot \cdots \cdot \widetilde{m_n}(A_n)$.

According to the rules of equation (12), the n features of physical toy design scheme are fused to realize the evaluation of physical toy design scheme.

4. Case Study

4.1. Scheme Design. According to the actual needs of ADHD children's treatment and the analysis of existing products in the market, it is found that the toy equipment in the form of children's balance scooter has three characteristics that can exercise spatial perception ability, concentration ability, and balance ability. Then, taking children's scooter as the research object, this paper designs the physical toy equipment for ADHD children.

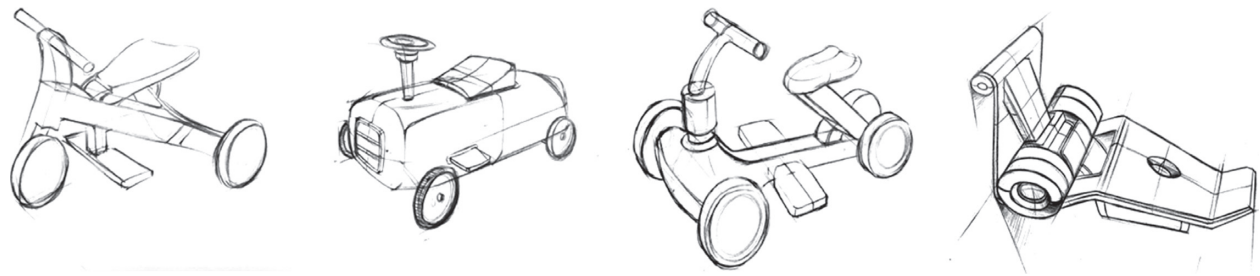
Through the investigation and visit to the families of 10 ADHD children, 10 ADHD children and their parents were exchanged and interviewed to obtain their cognition of the shape of children's balance scooters. The survey results show that for ADHD children and their parents, the most important characteristics of balance scooters are interest, visual impact, and sense of security. Summarizing the design expectations, three preliminary schemes are obtained from the perspectives of comfortable handling, interesting modeling, and lightness and flexibility, as shown in Figure 2.

The toy car is driven forward by children squatting and reciprocating on the balance scooter, which plays the role of exercising children's proprioception and vestibular perception, and can train children's spatial perception ability, concentration ability, perceived weight, joint static perception ability, and balance ability. The finger pressing plate structure is used at the pedal to exercise children's touch. The whole design can enable ADHD children to achieve the comprehensive training effect of sensory integration in the whole process of using the balance scooter.

Scheme 1 has a simple and lovely shape, and the seat part is relatively wide, which can fully protect children's hips and make them more comfortable when riding. The wheel is designed to be fully enclosed without spokes, so as to prevent children's feet from being involved in spokes when playing and reduce potential safety hazards. Scheme 2 is generous in shape and imitates retro cars, which is interesting and easy to attract children's attention. At the same time, wide curve seats are also used to protect children's hips and non spoke wheels to prevent feet from being drawn into spokes. Scheme 3 and scheme 4 have simple and light shape and higher flexibility. The seat is suspended, which greatly improves the shock absorption effect. The chassis is in an equilateral triangle, which has the effect of stabilizing and preventing side fall, and improves the safety of the toy car.

4.2. Fuzzy Comprehensive Evaluation. The vestibular training, the proprioception training, the tactile training, safety, and greenness are selected as the features of physical toy design scheme evaluation, which are represented by $I_1, I_2, \ldots, I_5$ in turn. Through expert evaluation, the feature values of the four physical toy design schemes on $I_1, I_2, \ldots, I_5$ are shown in Table 1.

For evaluation level, $l = 4$, that is: acceptable (N1), fair (N2), good (N3), and excellent (N4). Referring to the feature value of each physical toy design scheme to be evaluated, the classical domain matter elements of four evaluation levels



Scheme 1 Scheme 2 Scheme 3 Scheme 4

Figure 2: Preliminary scheme set.

Table 1: The feature values of the four physical toy design schemes on $I_1, I_2, \ldots, I_5$.

Physical toy design scheme	I_1	I_2	I_3	I_4	I_5
1	0.9333	1.0000	0.8902	0.2500	0.2500
2	0.8615	0.6667	0.8295	0.2500	0.7500
3	1.0000	1.0000	1.0000	0.2500	0.2500
4	0.8000	0.6667	0.9125	1.0000	1.0000

during the evaluation of physical toy design scheme are obtained according to equation (2), which are, respectively:

$$R_1 = \begin{bmatrix} N_1 & I_1 & [0\cdot 7667, 0\cdot 8333] \\ & I_2 & [0\cdot 6111, 0\cdot 7222] \\ & I_3 & [0\cdot 8011, 0\cdot 8579] \\ & I_4 & [0\cdot 1250, 0\cdot 3750] \\ & I_5 & [0\cdot 1250, 0\cdot 3750] \end{bmatrix},$$

$$R_2 = \begin{bmatrix} N_2 & I_1 & [0\cdot 8333, 0\cdot 9000] \\ & I_2 & [0\cdot 7222, 0\cdot 8333] \\ & I_3 & [0\cdot 8579, 0\cdot 9147] \\ & I_4 & [0\cdot 3750, 0\cdot 6250] \\ & I_5 & [0\cdot 3750, 0\cdot 6250] \end{bmatrix},$$

$$R_3 = \begin{bmatrix} N_3 & I_1 & [0\cdot 9000, 0\cdot 9667] \\ & I_2 & [0\cdot 8333, 0\cdot 9444] \\ & I_3 & [0\cdot 9147, 0\cdot 9716] \\ & I_4 & [0\cdot 6250, 0\cdot 8750] \\ & I_5 & [0\cdot 6250, 0\cdot 8750] \end{bmatrix}, \quad (13)$$

$$R_4 = \begin{bmatrix} N_4 & I_1 & [0\cdot 9667, 1\cdot 0333] \\ & I_2 & [0\cdot 9444, 1\cdot 0556] \\ & I_3 & [0\cdot 9716, 1\cdot 0284] \\ & I_4 & [0\cdot 8750, 1\cdot 1250] \\ & I_5 & [0\cdot 8750, 1\cdot 1250] \end{bmatrix}.$$

According to equation (3), the joint domain matter element during physical toy design scheme evaluation is obtained as follows:

$$R_p = \begin{bmatrix} N_p & I_1 & [0\cdot 7667, 1\cdot 0333] \\ & I_2 & [0\cdot 6111, 1\cdot 0556] \\ & I_3 & [0\cdot 8011, 1\cdot 0284] \\ & I_4 & [0\cdot 1250, 1\cdot 1250] \\ & I_5 & [0\cdot 1250, 1\cdot 1250] \end{bmatrix}. \quad (14)$$

First physical toy design scheme 1 is evaluated, and then the matter element to be evaluated is

$$R_0 = \begin{bmatrix} N_0 & I_1 & 0\cdot 9333 \\ & I_2 & 1\cdot 0000 \\ & I_3 & 0\cdot 8902 \\ & I_4 & 0\cdot 2500 \\ & I_5 & 0\cdot 2500 \end{bmatrix}. \quad (15)$$

By substituting into equation (5), the correlation function matrix obtained is

$$K_0 = \begin{bmatrix} -0\cdot 5000 & -0\cdot 8332 & -0\cdot 2661 & 0\cdot 5000 & 0\cdot 5000 \\ -0\cdot 2498 & -0\cdot 7499 & 0\cdot 4313 & -0\cdot 5000 & -0\cdot 5000 \\ 0\cdot 4993 & -0\cdot 5000 & -0\cdot 2157 & -0\cdot 7500 & -0\cdot 7500 \\ -0\cdot 2504 & 0\cdot 5000 & -0\cdot 4774 & -0\cdot 8333 & -0\cdot 8333 \end{bmatrix}. \quad (16)$$

The discrimination framework of physical toy design scheme evaluation is established as $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$, where $\theta_1, \theta_2, \theta_3, \theta_4$ indicate the evaluation levels in turn: acceptable (N1), fair (N2), good (N3), and excellent (N4).

The BPA function matrix on the discrimination framework is obtained from equation (6):

$$m = \begin{bmatrix} 0\cdot 1591 & 0\cdot 1374 & 0\cdot 2054 & 0\cdot 5214 & 0\cdot 5214 \\ 0\cdot 2044 & 0\cdot 1494 & 0\cdot 4125 & 0\cdot 1918 & 0\cdot 1918 \\ 0\cdot 4322 & 0\cdot 1918 & 0\cdot 2160 & 0\cdot 1494 & 0\cdot 1494 \\ 0\cdot 2043 & 0\cdot 5214 & 0\cdot 1662 & 0\cdot 1374 & 0\cdot 1374 \end{bmatrix}. \quad (17)$$

The importance vector of features $I_1, I_2, \ldots, I_5$ is further obtained from equations (7)–(10):

$$\omega = [0\cdot 1745\ 0\cdot 2745\ 0\cdot 1461\ 0\cdot 2024\ 0\cdot 2024]. \quad (18)$$

Therefore, the relative importance vector of features $I_1, I_2, \ldots, I_5$ is

$$\mu = [0 \cdot 6358\ 1 \cdot 0000\ 0 \cdot 5320\ 0 \cdot 7375\ 0 \cdot 7375]. \quad (19)$$

The BPA function is improved according to the relative importance vector. The value of the improved BPA function is shown in Table 2.

The evidence fusion for the five features is conducted according to (12). For physical toy design scheme 1, the BPA function after fusion is $M = \widetilde{m_1} \oplus \widetilde{m_2} \oplus \widetilde{m_3} \oplus \widetilde{m_4} \oplus \widetilde{m_5}$. Therefore, $M(\theta_1) = 0.4773$, $M(\theta_2) = 0.1705$, $M(\theta_3) = 0.1701$, $M(\theta_4) = 0.1821$.

Similarly,for physical toy design scheme 2, $M(\theta_1) = 0.7198$, $M(\theta_2) = 0.1503$, $M(\theta_3) = 0.0983$, $M(\theta_4) = 0.0316$; for physical toy design scheme 3, $M(\theta_1) = 0.3905$, $M(\theta_2) = 0.0629$, $M(\theta_3) = 0.0645$, $M(\theta_4) = 0.4821$; for physical toy design scheme 4, $M(\theta_1) = 0.4238$, $M(\theta_2) = 0.0724$, $M(\theta_3) = 0.0722$, $M(\theta_4) = 0.4315$.

According to the D-S theory, the BPA function value after fusion actually represents the degree of support for a proposition after evidence fusion, as shown in Figure 3.

Taking physical toy design scheme1 as an example, the degree of support for the four propositions of "scheme 1 belongs to N1, N2, N3, and N4" after evidence fusion is 0.4773, 0.1705, 0.1701, and 0.1821, respectively, of which the degree of support for "scheme 1 belongs to N1" is the highest, so it is considered that scheme 1 belongs to N1 level. Similarly, scheme 2 belongs to N1 and schemes 3 and 4 belong to N4.

For schemes 1 and 2 belonging to N1 level, the total support for "scheme 1 belongs to higher level (i.e. N2, N3 and N4)" after evidence fusion is 0.5179, while the total support for "scheme 2 belongs to higher level (i.e. N2, N3 and N4)" after evidence fusion is 0.2802. It can be seen that compared with scheme 2, the degree of support for "scheme 1 belongs to a higher level" after evidence fusion is higher, so scheme 1 is better than scheme 2.

Similarly, for schemes 3 and 4 belonging to N4, it can be seen that after evidence fusion, the total support for "scheme 3 belongs to lower level (i.e. N1, N2 and N3)" is 0.5227, while after evidence fusion, the total support for "scheme 4 belongs to lower level (i.e. N1, N2 and N3)" is 0.5684. It can be seen that compared with scheme 3, the degree of support for "scheme 4 belongs to lower level" after evidence fusion is higher, so scheme 3 is better than scheme 4.

Therefore, for the scheme at the same level, the total support degree for the scheme at a higher level and a lower level after evidence fusion can be calculated, respectively, according to the connotation of BPA function, and the ranking relationship of the scheme at the same level can be obtained through further comparative analysis.

Finally, schemes 1 and 2 belong to N1 level, schemes 3 and 4 belong to N4 level, scheme 1 is better than scheme 2, and scheme 3 is better than scheme 4. The total ranking relationship of the four physical toy design schemes to be evaluated is 3>4>1>2, where $\succ$ means "better than," and the optimal scheme is 3.

Extension analysis has been applied in many fields to verify that it can deal with the incompatibility in decision-making problems. In this paper, extension analysis is used to deal with the incompatibility between the characteristics of product physical toy design scheme under the background of green manufacturing, which has a strong theoretical basis. The evidence theory is an uncertain reasoning method with strong theoretical basis. It can reason without a priori probability and conditional probability and continuously reduce the hypothesis set by relying on evidence accumulation. It has also been widely used. Therefore, this paper combines the two methods to establish a physical toy design scheme evaluation framework. This method is easy to apply. It can not only solve the total ranking relationship of all schemes to be evaluated but also get the specific classification of each scheme to be evaluated. In addition, the classical domain matter elements of each evaluation level can be set according to the needs of decision makers, and different classification results can be obtained when the total ranking relationship to be evaluated remains unchanged.

Table 2: Improved BPA function value.

	$\widetilde{m_1}(\theta_j)$	$\widetilde{m_2}(\theta_j)$	$\widetilde{m_3}(\theta_j)$	$\widetilde{m_4}(\theta_j)$	$\widetilde{m_5}(\theta_j)$
θ_1	0.1012	0.1374	0.1093	0.3845	0.3845
θ_2	0.1299	0.1494	0.2194	0.1414	0.1414
θ_3	0.2748	0.1918	0.1149	0.1102	0.1102
θ_4	0.1299	0.5214	0.0884	0.1014	0.1014
Θ	0.3642	0	0.4680	0.2625	0.2625

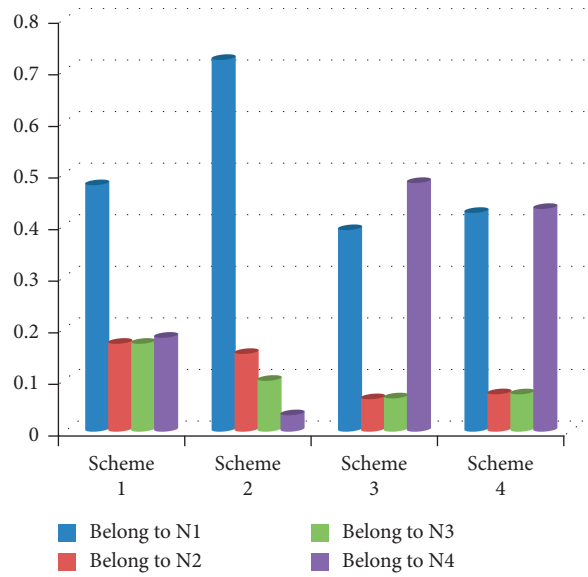


Figure 3: The support degree of evidence fusion to four process planning schemes at a level.

According to the requirements of sensory integration training at the criterion level, scheme 3 is deeply and carefully designed, as shown in Figure 4. The children's balance scooter provides power for children by squatting and reciprocating on the car, which can effectively exercise children's spatial perception ability, concentration ability, balance ability, perceived weight, and balance ability; Placing a finger pressure plate at the pedal can achieve the effects of pressure training and weight perception training. At the same time, it has a simple and lovely appearance, which meets the aesthetic needs of children.

Figure 4: Design of physical toy for children with ADHD.

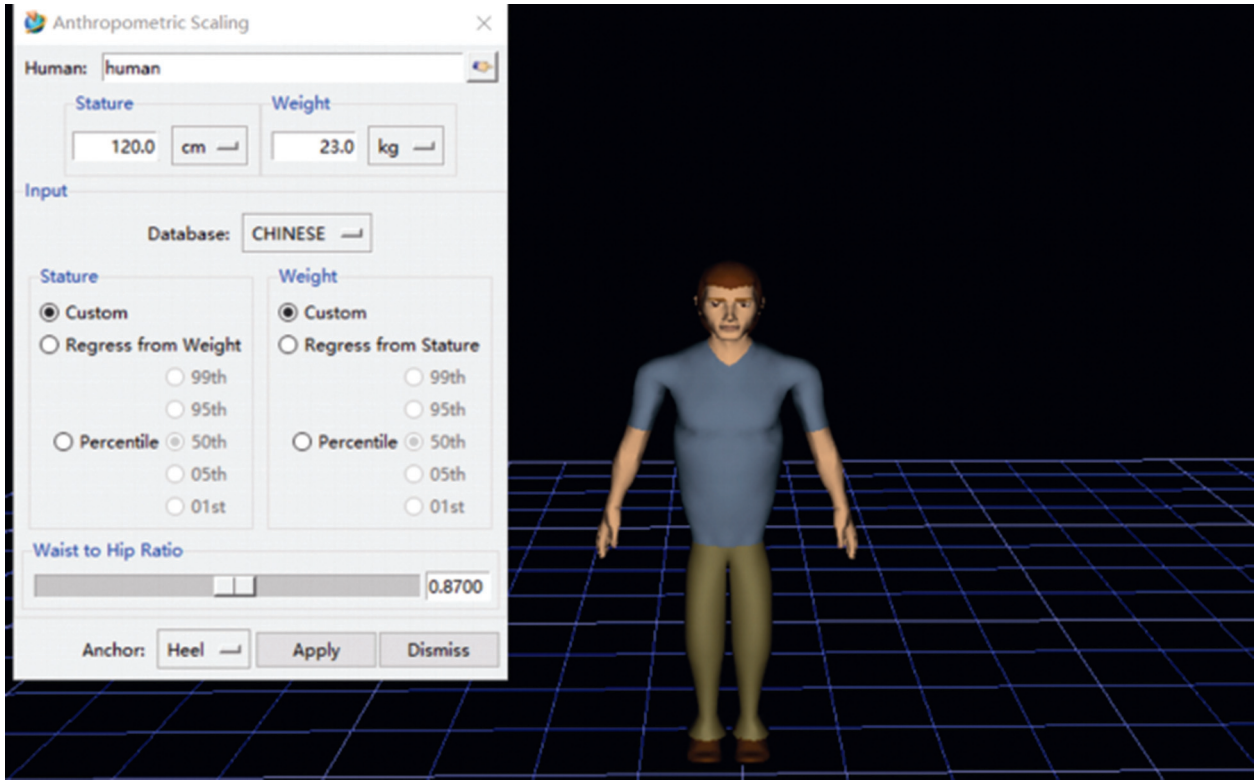


Figure 5: Establishment of virtual human body model size.

4.3. Product Testing. In the product test, Jack software of Siemens company is used for man-machine simulation of the product. Through the analysis of simulation data to verify whether the design scheme is safe and reasonable.

Input the parameters of the character model in Jack software and select the average height and weight of a normal 6-year-old Chinese boy of 120 cm and 23 kg, as shown in Figure 5.

Import the 3D digital model of the balance scooter into Jack software. Set the virtual human's posture through [human control], adjust the virtual human's hand to hold the handle through the [behaviors] command, place the virtual human's foot on the pedal, and seat the virtual human on the seat through the [attach to] command, as shown in Figure 6.

Use Jack software to simulate the loading process of the balance scooter and select [lower back analysis] and [NIOSH] functions to obtain the chart, as shown in Figure 7.

From the test and analysis results, the lower back analysis and lifting analysis values are within a reasonable range, so the loading process of the balance scooter is safe and reasonable. ADHD children can safely operate the product. By doing squat reciprocating motion and driving and playing, they can achieve the stimulation of rise and fall vibration, sudden start and emergency stop, and reflex adjustment. These stimulation forms can enable ADHD children to achieve the effect of rehabilitation training.

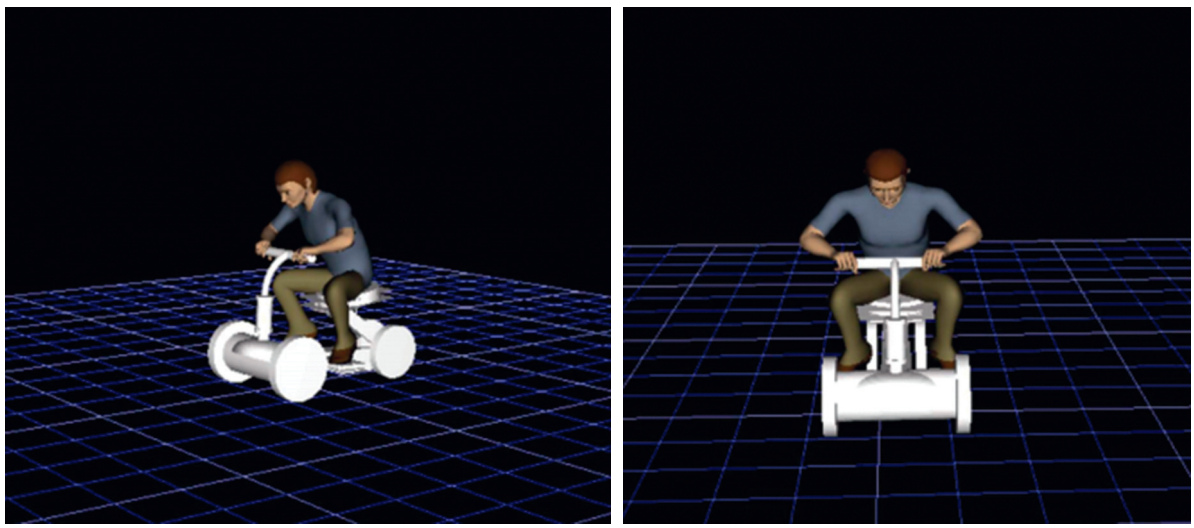


Figure 6: Simulation of balance sliding step vehicle.

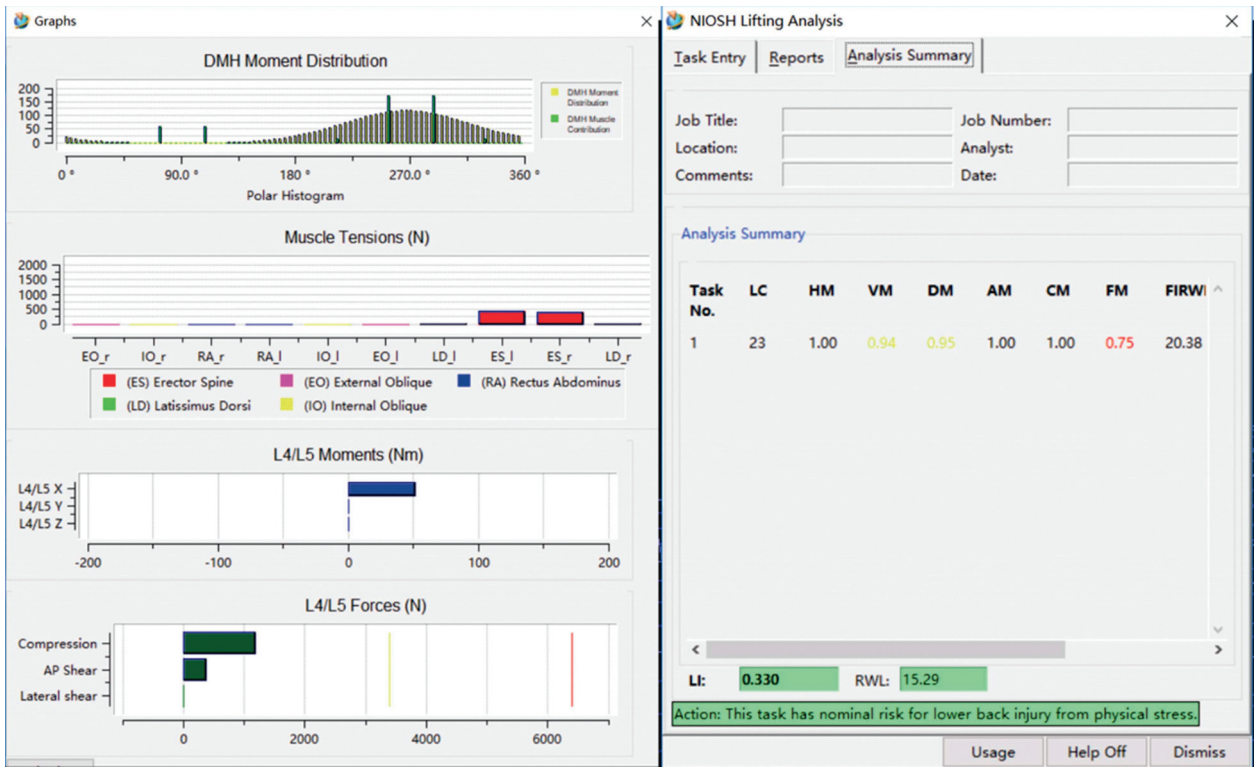


Figure 7: Data chart of lower back analysis and NIOSH.

5. Conclusions

At present, the treatment of ADHD children is mainly behavioral intervention, and its repeatability and drug side effects lead to poor treatment effect. The physical training method of sensory integration has the advantages of safety, simplicity, and low cost, which is accepted and welcomed by many parents. Extenics and the D-S theory combined with sensory integration training were used to evaluate the toy design of ADHD children. Integrating the rehabilitation criteria of sensory integration training, nine secondary indicators are extended from the three criteria level indicators of tactile training, the vestibular training, and the proprioceptive training to construct the evaluation system. According to the weight ranking of the indicators, clarify the toy type and toy function, design four primary schemes, and then score the three primary schemes by extenics and the D-S theory evaluation to obtain the best design scheme and refine it. In the product test, Jack software of Siemens is used to verify whether the

final product is safe and reasonable through virtual simulation. This study takes children with psychological problems as the research object, provides new ideas for the design of children's toys, and improves the scientificity and reliability of the physical toy design for children with ADHD.

Data Availability

The data set can be accessed upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This work was supported by the Ministry of Education of Humanities and Social Science Project 2021 (21YJA760091 and 21YJC760042).

References

[1] E. G. Willcutt, A. E. Doyle, J. T. Nigg, S. V. Faraone, and B. F. Pennington, "Validity of the executive function theory of attention-deficit/hyperactivity disorder: a meta-analytic review," *Biological Psychiatry*, vol. 57, no. 11, pp. 1336–1346, 2005.

[2] Y. Tang, X. Li, Y. Chen, Y. Zhong, A. Jiang, and C. Wang, "High-accuracy classification of attention deficit hyperactivity disorder with l2,1-norm linear discriminant analysis and binary hypothesis testing," *IEEE Access*, vol. 8, no. 99, Article ID 56228, 2020.

[3] L. Pérez-Crespo, J. Canals-Sans, E. Suades-González, and M. Guxens, "Temporal trends and geographical variability of the prevalence and incidence of attention deficit/hyperactivity disorder diagnoses among children in Catalonia, Spain," *Scientific Reports*, vol. 10, no. 1, p. 6397, 2020.

[4] D. Clark, K. E. Seymour, R. L. Findling, and S. H. Mostofsky, "Subtle motor signs as a biomarker for mindful movement intervention in children with attention-deficit/hyperactivity disorder," *Journal of Developmental and Behavioral Pediatrics*, vol. 41, no. 5, pp. 349–358, 2020.

[5] A. Ronald, N. D. Bodec, and T. Polderman, "Systematic review: how the attention-deficit/hyperactivity disorder polygenic risk score adds to our understanding of ADHD and associated traits," *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 60, no. 10, pp. 1234–1277, 2021.

[6] S. L. Li, K. W. Kam, A. Chee et al., "The association between attention-deficit/hyperactivity disorder and retinal nerve fiber/ganglion cell layer thickness measured by optical coherence tomography: a systematic review and meta-analysis," *International Ophthalmology*, vol. 41, no. 9, pp. 3211–3221, 2021.

[7] L. Li, C. Mao, H. Sun, Y. Yuan, and B. Lei, "Digital twin driven green performance evaluation methodology of intelligent manufacturing: hybrid model based on fuzzy rough-sets AHP, multistage weight synthesis, and PROMETHEE II," *Complexity*, vol. 2020, no. 6, 24 pages, Article ID 3853925, 2020.

[8] L. Jian, J. Q. Wang, and J. H. Hu, "Multi-criteria Decision-Making Method Based on Dominance Degree and BWM with Probabilistic Hesitant Fuzzy information," *International Journal of Machine Learning & Cybernetics*, vol. 10, no. 7, 2019.

[9] L. Li, J. Hang, H. Sun, and L. Wang, "A conjunctive multiple-criteria decision-making approach for cloud service supplier selection of manufacturing enterprise," *Advances in Mechanical Engineering*, vol. 9, no. 3, Article ID 168781401668626, 2017.

[10] M. M. Mousavi and J. Lin, "The application of PROMETHEE multi-criteria decision aid in financial decision making: case of distress prediction models evaluation," *Expert Systems with Applications*, vol. 159, Article ID 113438, 2020.

[11] L.-h. Li, J.-c. Hang, Y. Gao, and C.-y. Mu, "Using an integrated group decision method based on SVM, TFN-RS-AHP, and TOPSIS-CD for cloud service supplier selection," *Mathematical Problems in Engineering*, vol. 2017, Article ID 3143502, 14 pages, 2017.

[12] S. A. Churchill, J. Inekwe, K. Ivanovski, and R. Smyth, "Stationarity Properties of Per Capita CO2 Emissions in the OECD in the Very Long-Run: A Replication and Extension analysis," *Energy Economics*, vol. 90, Article ID 104868, 2020.

[13] W. Cai and Y. Shi, "Extenics: its significance in science and prospects in application," *Journal of Harbin Institute of Technology*, vol. 38, no. 7, pp. 1079–1086, 2006.

[14] G. Zheng, Y. Jing, H. Huang, X. Zhang, and Y. Gao, "Application of Life Cycle Assessment (LCA) and extenics theory for building energy conservation assessment," *Energy*, vol. 34, no. 11, pp. 1870–1879, 2009.

[15] Z. G. Tao, D. D. Zhao, X. J. Yang, J. M. Wang, and Y. Shu, "Evaluation of open-pit mine security risk based on FAHP-extenics matter-element model," *Geotechnical & Geological Engineering*, vol. 38, no. 2, pp. 1653–1667, 2020.

[16] L. Li and H. Wang, "A parts supplier selection framework of mechanical manufacturing enterprise based on D-S evidence theory," *Journal of Algorithms & Computational Technology*, vol. 12, no. 4, pp. 333–341, 2018.

[17] G. Zhao, A. Chen, G. Lu, and W. Liu, "Data fusion algorithm based on fuzzy sets and D-S theory of evidence," *Tsinghua Science and Technology*, vol. 25, no. 1, pp. 12–19, 2020.

[18] L. Li, T. Qu, Y. Liu et al., "Sustainability assessment of intelligent manufacturing supported by digital twin," *IEEE Access*, vol. 8, Article ID 174988, 2020.

[19] M. Gandolfi, D. Munari, C. Geroin et al., "Sensory Integration Balance Training in Patients with Multiple Sclerosis: A Randomized, Controlled trial," *Multiple Sclerosis*, vol. 21, no. 11, 2015.

[20] G. Marialuisa, G. Christian, P. Alessandro et al., "Robot-assisted vs. sensory integration training in treating gait and balance dysfunctions in patients with multiple sclerosis: a randomized controlled trial," *Frontiers in Human Neuroscience*, vol. 8, p. 318, 2014.

[21] K. H. Li, S. J. Lou, H. Y. Tsai, and R. C. Shih, "The effects of applying game-based learning to webcam motion sensor games for autistic students' sensory integration training," *Turkish Online Journal of Educational Technology - TOJET*, vol. 11, no. 4, pp. 451–459, 2012.

[22] A. Setti, J. Stapleton, D. Leahy, C. Walsh, R. A. Kenny, and F. N. Newell, "Improving the efficiency of multisensory integration in older adults: audio-visual temporal discrimination training reduces susceptibility to the sound-induced flash illusion," *Neuropsychologia*, vol. 61, pp. 259–268, 2014.

[23] Y.-Y. Liao, Y.-R. Yang, Y.-R. Wu, and R.-Y. Wang, "Virtual reality-based wii fit training in improving muscle strength,

sensory integration ability, and walking abilities in patients with Parkinson's disease: a randomized control trial," *International Journal of Gerontology*, vol. 9, no. 4, pp. 190–195, 2015.

[24] L. J. Deng and M. A. Ai-Ling, "Discuss on Water-Saving Irrigation Schemes Optimization Based on TOPSIS Model and CRITIC Weights method," *Water Ences and Engineering Technology*, vol. 21, 2020.

[25] L. Li and C. Mao, "Big data supported PSS evaluation decision in service-oriented manufacturing," *IEEE Access*, vol. 8, no. 99, Article ID 154663, 2020.

[26] X. Ping, "Application of CRITIC Method in Medical Quality Assessment," *Value Engineering*, vol. 30, 2011.

Hindawi
Scientific Programming
Volume 2022, Article ID 1383520, 9 pages
https://doi.org/10.1155/2022/1383520

Research Article

Intangible Cultural Heritage Management Using Machine Learning Model: A Case Study of Northwest Folk Song Huaer

Liusuo Huang and Yan Song

Software College, Henan Finance University, Zhengzhou 450046, China

Correspondence should be addressed to Liusuo Huang; huangliusuo@hafu.edu.cn

Received 23 December 2021; Revised 9 January 2022; Accepted 18 January 2022; Published 9 February 2022

Academic Editor: Baiyuan Ding

Machine learning is the core field of artificial intelligence. It provides a guarantee for the digital intangible cultural heritage management of artificial intelligence. However, the existing theoretical and practical research points out that there are still many gaps in this field. Huaer is a folk song popular in Qinghai, Gansu, Ningxia Hui Autonomous Region, and individual regions of Xinjiang. It is known as the soul of the northwest. It is a national human intangible cultural heritage. It was listed as human intangible cultural heritage by the United Nations in September 2009. With the rapid development of network technology and machine learning, it is very important to manage the network communication and deep mining of Huaer information. In this regard, use of machine learning natural language processing to mine the information of Huaer lyrics is proposed. By constructing the Huaer model of recurrent neural network (RNN), data mining of Huaer lyrics is carried out, and the built-in language module in Python is interconnected with dynamic Web pages. Four Huaer image segmentation methods and five deep learning algorithms are proposed, and the steps of image segmentation algorithm and BP neural network algorithm based on block technology are introduced. The results can provide new ideas for the protection and inheritance of music intangible cultural heritage and provide effective and high-quality information for Huaer art researchers and lovers.

1. Introduction

Artificial intelligence is the greatest scientific and technological innovation in the 21st century [1–5]. As the core field of artificial intelligence, the goal of machine learning is to let computers learn by themselves. The machine learning algorithm enables it to identify the laws in the observation data, build a model to explain the world, and predict things without clear preprogramming rules and models [6–9]. Therefore, using machine learning to protect intangible cultural heritage can help people better inherit intangible cultural heritage.

Huaer refers to the traditional folk songs popular in Northwest China such as Gansu, Ningxia, and Qinghai. As a world intangible cultural heritage, it is the most bloody, magnificent, and heroic logo in Chinese folk songs. It is also the most complete and representative model in the Chinese folk song music system. Huaer is born with the people's working life. The first research has carried out extensive and in-depth research on the musical attributes and cultural forms of Huaer. The research content covers the musical characteristics of Huaer, aesthetic identity, humanistic spirit, inheritance, and protection. The research objects include Gansu Huaer, Xinjiang Huaer, Qinghai Huaer, and Ningxia Huaer. Research perspectives include musicology, anthropology, folklore, and ecology [10–13].

With the continuous development of computer technology, foreign and domestic Internet users obtain information more rapidly, and people's growing spiritual and cultural needs are higher and higher. Huaer folk lovers and Huaer researchers urgently need to obtain effective, comprehensive, and large-scale Huaer information [11]. At present, the world culture is in a period of great development, various ideological and cultural exchanges and blending are more frequent, and the role and status of culture in the competition of comprehensive national

strength are more prominent. As a part of the traditional culture of China, the development and dissemination of Huaer are also indispensable at present [13]. In this regard, the author uses the advantages of computer technology to analyze the characteristic culture of Huaer in this region. By studying the development of Huaer, the author comes to the conclusion that the collection of Huaer information on the Internet is very small, and the content is incomplete and difficult to collect. As a unique art in this region, Huaer has great development space and receives resource constraints in the process of inheritance and development. Therefore, the research on Huaer information is more necessary and urgent at present.

Unfortunately, the existing theoretical and practical research points out that there are many gaps in this field. How to identify intangible cultural heritage with the help of machine learning, build a model to explain the characteristics of Huaer music, and then inherit intangible cultural heritage with artificial intelligence is an urgent practical problem in front of the academic interface. From the perspective of machine learning, this study examines and discusses the identification of Huaer, the world intangible cultural heritage.

2. Related Works

The current research also discusses the identification of Huaer, but mostly focuses on distinguishing various characteristics of Huaer, such as musical structure, mode and tonality, singing style, and poetic characteristics. On the whole, the research on these characteristics focuses on language description rather than on strategy extraction, so it is not good for the guidance of using artificial intelligence to improve the effect of Huaer recognition. How to adopt machine learning strategies to effectively identify Huaer is one of the few related studies.

Except for a few studies that have developed the emotion recognition system of music audio, most studies focus on audio pitch recognition and note recognition, while there are few empirical studies on Huaer case base based on machine learning, only the regional pattern classification of Chinese folk songs [14–17]. Therefore, the domestic research on Huaer recognition needs to be strengthened in terms of the comprehensiveness of research methods, the innovation of research perspectives, and the diversity of research contents.

In terms of machine learning theory, the research results of foreign folk song recognition research in research methods and perspectives, online recognition, and music emotion [18] are far more than those of domestic similar research. In addition, it is rare to analyze and discuss folk song recognition through a deep neural network algorithm at home and abroad. Most of them are fuzzy systems based on sound-level contour and harmony analysis. However, the recognition theoretical system of other scholars still has the deficiency of emphasizing content and neglecting category.

To sum up, in addition to the lack and deficiency of theoretical research, practical and objective factors also increase the difficulty of identifying folk song species. For example, the question of whether Longnan folk songs belong to Huaer in Longzhong or still belong to folk songs is becoming more and more intense in the academic circle, and the relevant empirical research needs to be followed up. There are also deficiencies in the research of traditional music recognition at home and abroad, focusing on case description, interpretation, and strategy refinement, neglecting strategy classification, and combining theoretical system construction. It still needs to be further improved in the systematic and interdisciplinary vision of the research. No one has set foot in the field of introducing machine learning into Huaer recognition research. Therefore, from the perspective of machine learning, this study introduces musicology, computer science, and technology to explore the construction of Huaer resource database, to summarize various algorithms and strategies of Huaer recognition.

3. Basic Concepts

3.1. Recurrent Neural Network. A recurrent neural network (RNN) [19–21] is a kind of neural network specially used to process sequences. RNN can be extended to longer sequences, and most RNN can also handle variable length sequences. It shares parameters in different ways. Each output item is a function of the previous item, and each output item is generated by applying the same update rules to the previous output. RNN can be applied to spatial data across two dimensions. When an application involves time data and the whole sequence can be observed before providing the whole sequence data to the network, RNN can have a backward connection about time. The basic structure of the mainstream recurrent neural network model is shown in Figure 1.

The left half of Figure 1 is the basic structure diagram of RNN model not expanded by time (the black box in the figure represents the delay of a single time step), and the right half is the diagram expanded by time. In the structure diagram, each time step t is generally represented as follows:

(1) $x^{(t)}$ represents the input of training samples at time step t, and $x^{(t-1)}$ and $x^{(t+1)}$ represent the input of training samples in time steps $t-1$ and $t+1$, respectively.

(2) $h^{(t)}$ represents the activation function of the hidden layer at time step t. $h^{(t)}$ is determined by $x^{(t)}$ and $h^{(t-1)}$. The sigmoid function is used for general dichotomy problem, and softmax function is used for K-category classification problem.

(3) $o^{(t)}$ represents the output of the model at time step t. $o^{(t)}$ is only determined by the current hidden state $h^{(t)}$ of the model.

(4) $L^{(t)}$ represents the loss function of the model in time step t, and $L^{(t)}$ represents the length of the output value $o^{(t)}$ and the corresponding training target $y^{(t)}$.

(5) $y^{(t)}$ represents the target output of the training sample sequence at time step t. The connections from the input layer to the hidden layer, the hidden layer to the output layer, and the hidden layer to the

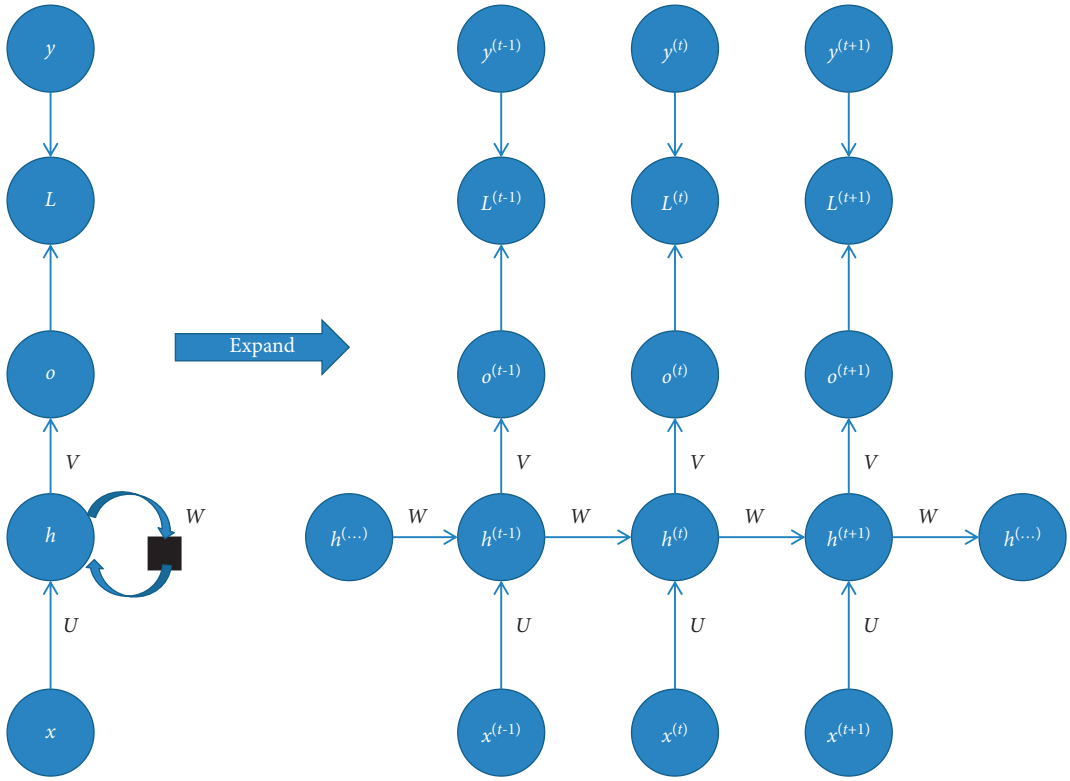


Figure 1: Basic structure of the RNN model.

hidden layer are parameterized by weight matrices U, V, and W, respectively.

3.2. Hypertext Preprocessor. Hypertext preprocessor (PHP) is a general purpose programming language originally designed for dynamic Web page development [22, 23]. PHP absorbs the characteristics of C, Java, and Perl languages. The syntax is simple and easy to learn. It is widely used in Internet dynamic Web page development technology and is mainly applicable to the field of Web page development. PHP's unique syntax combines C, Java, Perl, and PHP's own syntax. It can execute dynamic Web pages faster than common gateway interface (CGI) or Perl. Compared with other programming languages, PHP embeds the program into hypertext markup language (HTML) for execution, and the execution efficiency is much higher than CGI that completely generates HTML tags [4]. PHP can also execute compiled code, which can encrypt and optimize code operation to make the code run faster.

3.3. Python Built-In Module and Dynamic Web Page Interconnection. Python is an interpretative programming language, so it has the operation mechanism of interpretative language. So far, Python has many advantages over other languages because of its scalability, cross-platform, and other characteristics. Python's scalability is reflected in its modules, and its powerful class library provides effective help for computer cutting-edge disciplines such as machine learning [24, 25]. The Python modules involved in this experiment include (1) Gensim module applied to automatically extract semantic topics from text, and (2) sys, os, time, json, process, and network module socket in Python standard library.

Gensim is an open-source third-party Python module, which is used in unsupervised learning of the topic vector expression of the text hidden layer from the original unstructured text. It supports a variety of topic model algorithms including term frequency-inverse document frequency (TF-IDF), latent semantic analysis (LSA), latent Dirichlet allocation (LDA), and word2vec, supports flow training, and provides application program interface (API) for some common tasks such as similarity calculation and information retrieval. In this experiment, we use word2vec, a tool launched by Google to obtain word vectors. The internal algorithm of the tool realizes the transformation from word to vector through deep learning. The word vector output from word2vec model can be used to do a lot of related work of natural language processing, such as clustering, finding synonyms, part of speech analysis, and prediction.

In this experiment, Python, an open-source project of Google, is used to realize the interconnection between Python and dynamic Web pages. It complements the advantages of the two languages. The combination of Python program and PHP program can be understood as the technology of the combination of Python language and PHP language, and the popular technology can be understood as the mixed programming technology of Python language and PHP language. Python and PHP languages have their own internally defined data types. When PHP data are sent to Python or Python data are sent to PHP, transcoding is required in traditional technology, while Python technology

can directly send data by serializing different data types of Python and PHP without transcoding, which greatly improves the development speed.

Because of its global interpreter lock (GIL) feature, Python language has low multithreading efficiency. In Python based on the mixed programming mechanism of Python program and PHP program, the Python end can be deployed in a multi-process manner, to improve the overall working efficiency of the Python program. This technology improves the multithreading efficiency of Python. The basic principle of Python technology is socket communication, so network module socket support is required. Socket is a basic component in network programming. Socket is basically an information channel, and there are programs at both ends [8]. These programs may be located on different computers (connected through the network) and send information to each other through socket. The main principle of network communication in Python is shown in Figure 2.

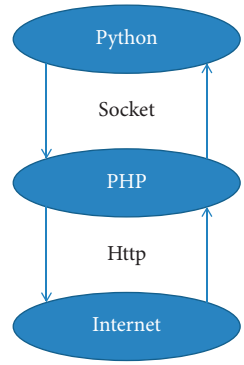


Figure 2: Communication schematic diagram of Python socket.

4. Construction of the Huaer Model Based on RNN

4.1. Python Crawler Building. The construction of cyclic neural network model has high requirements for the amount of data, so the Huaer lyric information used in this experiment is obtained by Python Web crawler technology. Using crawler technology, we can quickly and accurately obtain Huaer lyric information from Web applications and provide data support for subsequent experiments. Python Web crawlers are built to simulate computer network connections. That is, the computer makes a request to the server (with request header and message body), and the server responds to the computer's request (with HTML file). The crawler simulates the computer to initiate a request to the server, accepts the server's response content, and parses and extracts the required information.

By analyzing the Huaer information source, we found that the existing Huaer information is mainly distributed in Web application NetEase cloud music and QQ music. According to the structure of Web pages on different platforms, the crawler program is designed accordingly. Through analysis, the NetEase cloud music Web page structure is a multipage Web page structure. The process of this type of web crawler is as follows:

(1) Manually turn the page, observe the universal resource locator (URL) composition characteristics of a Web page, construct the URLs of all pages, and store them in the list

(2) Loop out the URLs according to the URL list

(3) Define crawler functions

(4) Call the crawler function in a loop to store data

(5) After the cycle, end the crawler

The flow chart of multipage crawler is shown in Figure 3.

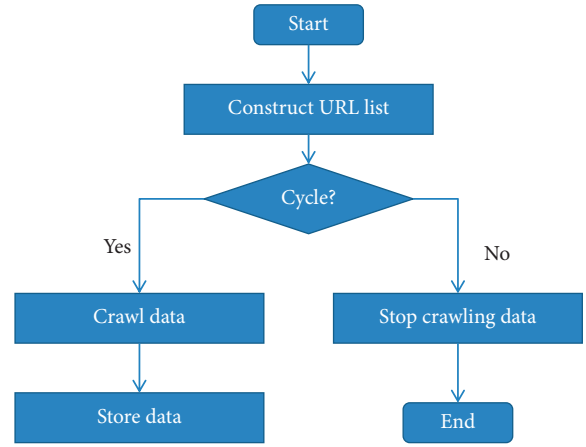


Figure 3: Multipage crawler flow chart.

The QQ music Web page structure is a cross-page Web page structure, and the cross-page crawler process is as follows:

(1) Define the crawl function to crawl the URLs of all albums on the list page

(2) Save the album URL to the list

(3) Define the function to crawl detailed page data

(4) Enter the album detailed page and crawl the detailed page data

(5) Store data, end the cycle, and end the crawler program

The flow chart of cross-page crawler is shown in Figure 4.

All the Huaer lyric data obtained in this experiment are saved in the text file for subsequent processing.

4.2. Lyric Splitting. Compared with English, English takes space as a very obvious separator, and an English word can be divided horizontally by letters. However, Chinese has no separator between words because it inherits the tradition of ancient Chinese and is divided into 8 kinds according to the "eight methods of Chinese character Yong," namely point, horizontal, vertical, apostrophe, pressing, bending, bending, and hook. In ancient Chinese, except for people's names, place names, and continuous words, words were usually single Chinese characters, so there was no need to write word segmentation at that time. In modern Chinese, there are

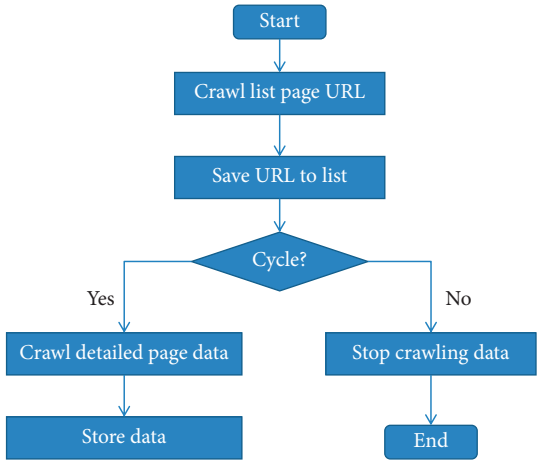


Figure 4: Cross-page crawler flow chart.

many double-word or multi-word words, and a word is no longer equivalent to a word. Therefore, it brings difficulty to Chinese word segmentation. To get more effective experimental data and more rigorous analysis of experimental results, Jieba, the mainstream word segmentation tool, is used for Huaer lyrics. We use the word segmentation tool Jieba to denoise the crawled Huaer lyrics and eliminate the special characters. After word segmentation, we can get the prediction database that can be used for training and then train the RNN model.

4.3. System Modeling. The larger the corpus, the better the training results of the model, while for the smaller corpus, the opposite is true. The model training requires the Python NLP Gensim package. Gensim needs to be installed first, but Gensim has requirements for the versions of science and technology libraries such as NumPy and SciPy. It needs to pay attention to the versions of NumPy and SciPy. When the algorithm program does not make an error during import, it will succeed. In Gensim, the parameters related to the training algorithm are in Gensim.models.word2vec.Word2vec. The parameters to be noted are as follows:

(1) sentences: this parameter sets the corpus to be analyzed, which can be sequence and character files. In this experiment, file traversal is used to read.

(2) size: this parameter sets the dimension of word vector. The default value is 100. The value of this dimension is generally related to the size of the corpus currently used. If the corpus is very small, such as text corpus less than 100 m, the default value is used. If the corpus is large, the dimension is increased.

(3) window: this parameter is the maximum distance from the word vector context. The larger the window is, the words farther away from a word will also have a context relationship. The default value is 5. In practical use, the size of this parameter can be dynamically adjusted according to the actual needs. If it is a small language material, this value can be set smaller.

(4) sg: this parameter is the selection of two models of word2vec. If it is 1, it is the skip-gram model. If it is 0, it is the continuous bag of word (CBOW) model. The default parameter value is 0.

(5) hs: this parameter is the choice of two solutions for word2vec. If the parameter value is 1 and the number of negative samples is greater than 0, it is hierarchical softmax. If the parameter value is 0, it is negative sampling. The default parameter value is 0.

(6) negative: this parameter is the number of negative samples when using negative sampling. The default is 5. It is recommended to be between [3, 10].

(7) cbow_mean: this parameter is used for CBOW projection. If it is 1, it is the average value of the word vector of the context. If it is 0, xw in the algorithm is the sum of the word vectors of the context. In this study, the average value is used to represent xw, and the default value is also 1. It is not recommended to modify the default value.

(8) min_count: this parameter is the minimum word frequency of the word vector to be calculated. This parameter is added to remove rare low-frequency words. The default is 5. If the corpus is too small, this value is lowered.

(9) iter: this parameter is the maximum number of iterations in the random gradient descent method. The default value is 5. For smaller corpora, you can reduce the parameter value, and for larger corpora, you can increase the parameter value.

(10) alpha: this parameter is the initial step of iteration in the random gradient descent method. The default is 0.025.

(11) min_alpha: because the algorithm supports gradually reducing step size in the iterative process, min_alpha gives the minimum iteration step value.

*Invoking Word2vec.*Word2vec() algorithm is enough for the training of the model. For corpora of different sizes, the algorithm parameters need to be adjusted to achieve better training results. When the model training is completed, saving the model for reuse is needed. In word2vec, there are two ways to save the model. One is to save the model directly, and the other is to store it in a form that can be parsed by C language. For this, it can be saved according to needs. In addition, the speed of model training is affected by the running environment of the training program and the size of the corpus. When the corpus is very large, the computer with better performance can train the model faster.

4.4. Identification Strategy. Machine learning recognition strategy is a set of systematic and operable Huaer recognition algorithms, which can provide theoretical support and practical guidance for Huaer recognition model and retrieval field. The folk song category artificial intelligence (AI) is

developed according to the recognition strategy, which can accurately distinguish the categories of folk songs according to the model. Before recognizing the characteristics of Huaer music, it is necessary to establish a certain scale database and sample size.

4.4.1. Huaer Database Construction. A database is a warehouse that organizes, stores, and manages data according to data structure. It is a collection of organized, shareable, and uniformly managed large amounts of data stored in computers for a long time. Huaer database is a collection of Huaer music resource data that are stored together in a certain way, can be shared with multiple users, has as little redundancy as possible, and is independent of the application. It can be regarded as an electronic file cabinet—the place where electronic files are stored. Users can add, query, update, delete, and other operations on Huaer music data in the file.

As Huaer is popular in Gansu, Qinghai, and Ningxia provinces (regions) in the northwest, it is a popular folk song among local Han, Hui, Tibetan, Dongxiang, Baoan, Sala, Tu, Yugu, and other nationalities. It covers a wide range of areas and objects, so it is difficult to collect materials. Therefore, the construction of Huaer database can apply the method of typical example extraction to collect representative Huaer music orders such as Gansu Huaer, Qinghai Huaer, and Ningxia Huaer. The collected genres and forms include Huaer's literary theory, pictures, music scores, and audio and video resources and focus on the audio data of all kinds of music cards sung by Northwest Huaer masters (Zhonglu Zhu in Gansu, Jun Ma in Qinghai, Dexian Wang in Ningxia, etc.). After the collection, the resources are classified, coded, labeled, and analyzed to realize the real-time dynamic update and expansion of the resource database. After completion, it will be uploaded to the existing local official social networking website platform to provide online sharing services for the whole society free of charge. The database is shown in Figure 5.

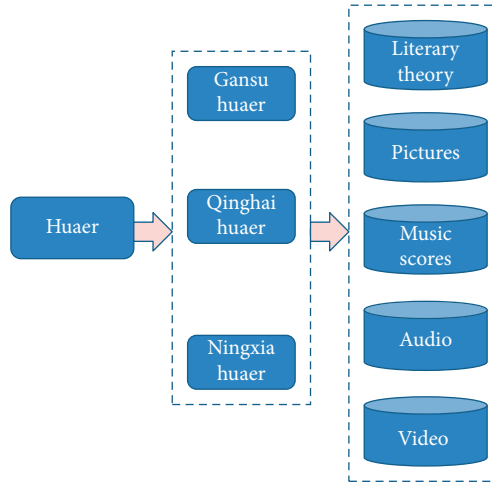


Figure 5: Huaer resource database.

4.4.2. Huaer Recognition Path. Recognition refers to the development of a human-level recognition system based on in-depth learning. Machine learning is a computer algorithm that predicts the image (music score), sound (tone), and other data of Huaer through multilayer nonlinear feature learning and hierarchical feature extraction. Its core is neural network algorithm.

(1) Image Recognition of the Huaer Music Score. According to the writing order, block and line features of Huaer notation and staff notation, image segmentation methods such as DFS-based image marker segmentation method, bounding box algorithm, adaptive iterative projection smoothing algorithm, and score image segmentation method based on cluster analysis are proposed, to lay a foundation for folk song information classification and recognition.

An image segmentation algorithm based on block technology is proposed in this study, and the main steps of image segmentation algorithm to extract Huaer score image are introduced as follows.

Firstly, a music score image is divided into independent small images, and each small image is divided initially; then, the initial division of each small image is fused, and then, a point is randomly selected in each division of the fusion, and the selected points constitute the characteristic data of the original score image; finally, the spectral clustering method is used to cluster these feature data, and the category of each pixel in the original image is determined according to the clustering results of the feature data. The proposed algorithm is mainly composed of image segmentation, initial presegmentation, clustering of feature data, and determining the category of pixels in the original image.

(1) Image segmentation

The purpose of image segmentation is to transfer a pair of large music score images into memory for processing, enhance the applicability of the algorithm, and avoid the problem that too large music score images cannot be processed. After the large-scale score image is divided into blocks, the amount of data and calculation pressure of each operation are reduced.

(2) Initial presegmentation

Each independent small image after blocking is initially divided, and then, one division from the initial division of each small image is selected as the reference. Without losing generality, the initial division of the first small image is taken as the reference, and the division of the remaining small images is compared with the reference division: if the division of the small image has the same grayscale as the reference division, then the pixels corresponding to the division of the small image are classified into the reference division; if the category of pixels in the small image division does not appear in the reference division, a new division is generated and added to the reference division. Through this method, an initial division of the whole image will be formed as follows:

$$c = [c_1, c_2, \ldots, c_n]. \quad (1)$$

(3) Clustering of characteristic data

A pixel i is randomly selected from each c_i, where $i = 1, 2, \ldots, n$, and the gray value x_i of the original music score image where the point is located is taken as an input object in the spectrum clustering. Assume that the data (called feature data) obtained by this method are as follows:

$$X = [X_1, X_2, \ldots, X_n]. \quad (2)$$

The feature data x are clustered using the spectral clustering algorithm, and the clustering result as v is set, where each element in v is the class mark of the input vector x_i, and the vector composed of these class marks is as follows:

$$v = [v_1, v_2, \ldots, v_n], \quad (3)$$

where $v_i \in \{1, 2, \ldots, k\}$, in which k is the number of categories to be divided.

(4) Determination of the category of pixels in the original music score image

It is assumed that the category of x_j is v_i and the division of x_j is c_j. All pixels in c_j are classified into the same category as x_j, to obtain the segmentation result of the original score image.

(2) Feature Recognition of Huaer Music. There are Bayes, nearest neighbor method, BP neural network, decision tree, and support vector machine algorithms to support Huaer music feature recognition [26–28]. Four recognition modules, audio processing, feature extraction, pattern classification, and semantic representation, are proposed according to the pitch, melody, rhythm, tone, and beat of Huaer melody. BP neural network method [29–33] is proposed for feature recognition of Huaer music in this study, and the main steps of extracting Huaer music features by BP neural network algorithm are introduced as follows.

BP neural network is a multilayer feedforward neural network. The main characteristics of the network are signal forward transmission and error back propagation. In forward transmission, the input signal is processed layer by layer from the input layer through the hidden layer to the output layer. The neuron state of each layer only affects the neuron state of the next layer. If the output layer cannot get the expected output, it will turn to back propagation and adjust the network weight and threshold according to the prediction error, so as to make the prediction output of BP neural network approach the expected output continuously. The topology of BP neural network can be regarded as a nonlinear function. The network input value and prediction value are the independent variable and dependent variable of the function, respectively.

When the number of input nodes is n and the number of output nodes is m, BP neural network expresses the functional mapping relationship from n independent variables to m dependent variables. Before BP neural network prediction, the network must be trained first. Through training, the network has the ability of associative memory and prediction. The training process of BP neural network includes the following steps:

Step 1: network initialization.

According to the system input and output sequence, the number of network input layer nodes, hidden layer nodes, and output layer nodes is determined. The connection weights between input layer, hidden layer, and output layer neurons are initialized. Then, the threshold of initialization hidden layer and the threshold of output layer are initialized. The learning rate and neuron excitation function are given.

Step 2: hidden layer output calculation.

According to the connection weight between input layer and hidden layer and the hidden layer threshold, the hidden layer output is calculated.

Step 3: output layer output calculation.

According to the hidden layer output, the connection weight, and the threshold, the prediction output o of BP neural network is calculated.

Step 4: error calculation.

The network prediction error is calculated according to the network prediction output and the expected output.

Step 5: weight update.

The network connection weight is updated according to the network prediction error.

Step 6: threshold update.

The network node threshold is updated according to the network prediction error.

Step 7: whether the algorithm iteration is over is judged. If not, return to Step 2.

5. Case Study

5.1. Mining Results of Huaer Lyrics. The training of this experimental model belongs to unsupervised learning. There are not many objective evaluation methods similar to supervised learning, and it depends more on end-to-end application. RNN is used to process Huaer's information to realize clustering, synonym finding, part of speech analysis, and prediction of Huaer's lyrics. Word2vec language model has high performance, but it has high requirements for the amount of data. The experimental results are as shown in Tables 1–3.

5.2. Analysis of Mining Results. In this experiment, the circular neural network is used, which is an excellent neural network designed to comprehensively use the forward information and reverse information in the historical information. In terms of data collection, due to the limitations of the historical attribute and cultural background of Huaer, the information collected by the author is limited. In addition, the neural network has high requirements for the size

Table 1: Most approximate words to "A Ge."

Word	Proximity
Guo Er	0.8006
Er Qing	0.7920
Ge Jiao	0.7892
Shi Lun	0.7865
Xuan Shi	0.7807
Shan Tou	0.7788
Da Sao	0.7787
Shi Yi	0.7731
Lao Mao	0.7714
Wan Nan	0.7713

Table 2: Most approximate words to "Da Mai."

Word	Proximity
Chi Yao	0.9893
Tian Tang	0.9886
Fa Zi	0.9877
Xian Hua	0.9876
Tou Dai	0.9873
Zhi Jia	0.9869
Tai Yuan	0.9866
Hou Tou	0.9864
Guo Yuan	0.9862
Xue Shan	0.9860

Table 3: Similarity between two approximate words.

Approximate words	Similarity
A Ge and Lao Xiang	−0.7582
A Ge and La Ba	−0.7815
A Ge and Min Ge	−0.7643
Lao Xiang and La Ba	0.9800
Lao Xiang and Min Ge	0.9859
La Ba and Min Ge	0.9713

Figure 6: Word cloud diagram.

of the corpus, and the correctness of the training results needs to be improved. It is more ideal to use speech recognition technology to convert the existing Huaer video and audio speech signals into corresponding text or commands to obtain a large amount of Huaer data.

As the traditional culture of Huaer, the key content of Huaer's lyrics is extracted to form a word cloud. After sorting the words, only the first 300 high-frequency words are intercepted due to too many words. The word cloud is shown in Figure 6.

By observing the word cloud picture, we can see that the word "Huaer" (the center of Figure 6) has the highest word frequency, and other words are successively expanded from the origin and distributed at each point according to the word frequency.

6. Conclusions

The 21st century is an era of intelligence. The research results of artificial intelligence have entered many fields such as science and engineering, humanities, and social sciences, as well as intangible cultural heritage protection and people's daily life. With the rapid development of science and technology, a large number of Huaer samples are identified by means of machine learning, which not only provides a strong sample library and gene library for the survival of Huaer as intangible cultural heritage, but also provides fertile soil for subsequent intelligent machine creation and inheritance of Huaer. In the face of the realistic dilemma of lack of successors and generation division in inheriting music intangible cultural heritage, this research may provide new ideas for the protection and inheritance of music intangible cultural heritage.

In this study, deep learning natural language processing is applied to Huaer information mining through cyclic neural network sequence modeling. Many interesting features are found in the experiment. The principle of analyzing and mining Huaer text information can also be applied to many aspects. The development of artificial intelligence will bring us more opportunities and challenges. Reasonable and effective use of computer technology can bring convenience to all aspects of human beings, such as speech recognition, deoxyribonucleic acid (DNA) sequence analysis, emotion classification, machine translation, and named body recognition. In the experiment, there are some difficulties in the process of Huaer information collection. It has certain limitations for practical applications, but the ideas provided have laid a theoretical foundation for subsequent practical application research. A larger corpus and optimizing the RNN internal network architecture can more effectively analyze Huaer information. How to obtain a larger data corpus with the help of speech recognition and video behavior recognition technology and optimize the internal network architecture is the next research problem.

Data Availability

The dataset can be accessed upon request to the corresponding author.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

[1] W. Yang, G. Li, Y. Xu, and J. Hu, "An algorithm for mining of association rules for the information communication network alarms based on swarm intelligence," *Mathematical Problems in Engineering*, vol. 2014, no. 1, 14 pages, Article ID 894205, 2014.

[2] M. K. Alomar, M. M. Hameed, N. Al-Ansari, and M. A. AlSaadi, "Data-driven model for the prediction of total dissolved gas: robust artificial intelligence approach," *Advances in Civil Engineering*, vol. 2020, Article ID 6618842, 2020.

[3] J. S. Suri, A. Puvvula, M. Majhail et al., "Integration of cardiovascular risk assessment with COVID-19 using artificial intelligence," *Reviews in Cardiovascular Medicine*, vol. 21, no. 4, pp. 541–560, 2020.

[4] C. K. Sahu, C. Young, and R. Rai, "Artificial intelligence (AI) in augmented reality (AR)-assisted manufacturing applications: a review[," *International Journal of Production Research*, vol. 59, no. 1, pp. 1–57, 2020.

[5] C. Shi, X. Feng, and Z. Jin, "Sustainable Development of China's Smart Energy Industry Based on Artificial Intelligence and Low-carbon economy," *Energy Science & Engineering*, vol. 10, 2020.

[6] F. Pedregosa, G. Varoquaux, A. Gramfort et al., "Scikit-learn: Machine Learning in Python," *The Journal of machine Learning research*, vol. 12, 2012.

[7] Y. Xin, L. Kong, Z. Liu et al., "Machine learning and deep learning methods for cybersecurity," *IEEE Access*, vol. 6, Article ID 35365, 2018.

[8] A. Alexandre, P. Fabian, E. Michael et al., "Machine learning for neuroimaging with scikit-learn," *Frontiers in Neuroinformatics*, vol. 8, p. 14, 2013.

[9] L. Li, H. Ruan, C. Liu et al., "Machine-learning reprogrammable metasurface imager," *Nature Communications*, vol. 10, no. 1, 2019.

[10] J.-X Zhang, "Research on Chinese traditional folk music from the perspective of music semiotices——a case study of Qinghai huaer," *Journal of Qinghai Nationalities Institute (Social Sciences)*, vol. 47, no. 1, pp. 117–122, 2021.

[11] Y. Yang and G. Welch, "Pedagogical challenges in folk music teaching in higher education: a case study of Hua'er music in China," *British Journal of Music Education*, vol. 33, no. 01, pp. 61–79, 2016.

[12] Y. Yang, J. Sundberg, G. Welch, and E. Himonides, "The tuning and vocal formant features of Chinese folk song singing: a case study of Hua'er music," *Journal of the Acoustical Society of America*, vol. 131, no. 4, p. 3376, 2012.

[13] Y. Yang and G. Welch, "Contemporary challenges in learning and teaching folk music in a higher education context: a case study of Hua'er music," *Music Education Research*, vol. 16, no. 2, pp. 193–219, 2014.

[14] S. Sheykhivand, Z. Mousavi, A. Farzamnia, and T. Yousefi Rezaii, "Recognizing emotions evoked by music using CNN-lstm networks on EEG signals," *IEEE ACCESS*, vol. 8, 2020.

[15] M. Soleymani, M. Pantic, and T. Pun, "Multimodal emotion recognition in response to videos," *IEEE TRANSACTIONS ON AFFECTIVE COMPUTING*, vol. 3, no. 2, pp. 211–223, 2012.

[16] J. A. Miranda-Correa, M. K. Abadi, I. Patras, and N. Sebe, "AMIGOS: a dataset for affect, personality and mood research on individuals and groups," *IEEE TRANSACTIONS ON AFFECTIVE COMPUTING*, vol. 12, no. 2, pp. 479–493, 2021.

[17] M. L. R. Menezes, A. Samara, R. Bond et al., "Towards emotion recognition for virtual environments: an evaluation of eeg features on benchmark dataset," *Personal and Ubiquitous Computing*, vol. 21, no. 6, pp. 1003–1013, 2017.

[18] M. Punkanen, T. Eerola, and J. Erkkil, "Biased emotional recognition in depression: perception of emotions in music by depressed patients - ScienceDirect," *Journal of Affective Disorders*, vol. 130, no. 1–2, pp. 118–126, 2011.

[19] L. Su and L. Zhou, "Exponential synchronization of memristor-based recurrent neural networks with multi-proportional delays," *Neural Computing & Applications*, vol. 31, no. 5, 2019.

[20] Y. Zhang, W. Kong, Z. Y. Dong et al., "Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network," *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 841–851, 2019.

[21] L. I. Peng, Y. Yang, X. Gao et al., "A Study of Chinese Speech Recognition Based on Bidirectional Recurrent Neural network," *Journal of Applied Acoustics*, vol. 39, no. 3, 2020.

[22] S. U. Xiao-Jian, C. L. Zhu, Q. G. Yang, J. Yan-bao, and G. Ling-wang, "Hypertext Preprocessor Based Multimedia System Assisting Ticks and Mites identification," *Chinese Journal of Frontier Health and Quarantine*, vol. 36, 2013.

[23] M. Juvane, P. Spiesberger, K. Pinter, R. Vallon, and T. Grechenig, "Assessing Mozambique's Software Industry to Foster Local Universities-Industry Collaboration," in *Proceedings of the 2020 IEEE Global Engineering Education Conference (EDUCON)*, April 2020.

[24] R. Cimrman, V. Luke, and E. Rohan, "Multiscale finite element calculations in Python using SfePy," *Advances in Computational Mathematics*, vol. 45, pp. 1897–1921, 2019.

[25] B. D. Palma, M. Erba, L. Mantovani, and N. Mosco, "A Python Program for the Implementation of the Γ," *Computer Physics Communications*, vol. 234, 2019.

[26] Z. Ning and S. University, "Research on Fountain Control System Based on Music Features Recognition," *Microcomputer Applications*, vol. 35, 2019.

[27] M. Chmulik, R. Jarina, M. Kuba, and E. Lieskovska, "Continuous Music Emotion Recognition Using Selected Audio Features," in *Proceedings of the 2019 42nd International Conference on Telecommunications and Signal Processing (TSP)*, Budapest, Hungary, July 2019.

[28] F. Yan, "Music recognition algorithm based on T-S cognitive neural network," *Translational Neuroscience*, vol. 10, 2019.

[29] G. Huo and Z. Ning, "Strength Prediction of Laminates Containing Embedded Fiber Wrinkles Using BP Neural Networks," *Journal of Nanjing University of Aeronautics & Astronautics*, vol. 52, 2020.

[30] Z. Wang, Y. Zhang, Z. Ren, C. Koh, and O. A. Mohammed, "Modeling of anisotropic magnetostriction under DC bias based on an optimized BP neural network," *IEEE Transactions on Magnetics*, vol. 56, no. 99, p. 1, 2020.

[31] X. L. Zhu, "A face recognition system using ACO-BPNN model for optimizing the teaching management system," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 5194044, 2021.

[32] F. Y. Guo, Y. C. Zhang, Y. Wang et al., "Fault detection of reciprocating compressor valve based on one-dimensional convolutional neural network," *Mathematical Problems in Engineering*, vol. 2020, Article ID 8058723, 2020.

[33] D. Wu and Y. J. Shen, "English feature recognition based on GA-BP neural network algorithm and data mining," *Computational Intelligence and Neuroscience*, Article ID 1890120, 2021.

Hindawi
Scientific Programming
Volume 2023, Article ID 9803492, 1 page
https://doi.org/10.1155/2023/9803492

Retraction

Retracted: Dynamic Early-Warning Model of College Students' Psychological Crisis Based on Characteristic Attribute

Scientific Programming

Received 26 September 2023; Accepted 26 September 2023; Published 27 September 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] X. Chen, "Dynamic Early-Warning Model of College Students' Psychological Crisis Based on Characteristic Attribute," *Scientific Programming*, vol. 2022, Article ID 6026248, 11 pages, 2022.

Hindawi
Scientific Programming
Volume 2022, Article ID 6026248, 11 pages
https://doi.org/10.1155/2022/6026248

Research Article

Dynamic Early-Warning Model of College Students' Psychological Crisis Based on Characteristic Attribute

Xiaojing Chen

School of Education, Xinyang University, Xinyang, Henan 464000, China

Correspondence should be addressed to Xiaojing Chen; 160301137@stu.cuz.edu.cn

Received 10 December 2021; Revised 30 December 2021; Accepted 18 January 2022; Published 9 February 2022

Academic Editor: Baiyuan Ding

At present, the existing dynamic early-warning model of college students' psychological crisis cannot obtain psychological characteristics, and the error of psychological crisis prediction is large. The early-warning result of psychological crisis appears big deviation. A new dynamic early-warning model of college students' psychological crisis is designed based on characteristic attributes. In the process of modeling negative psychological crisis of college students, the fusion constraint fuzzy theory obtains the essence of college students' psychology from subjective and objective reasons, gives the different stages of college students' psychology, obtains the corresponding information entropy of the behavior hindrance in each stage of college students' psychology, and calculates the threshold value of the behavior tendency of college students' psychology. Based on the psychological state information of college students obtained by perception, the fuzzy comprehensive early-warning data model of feature attributes is constructed. According to the fuzzy comprehensive early-warning method and the early-warning index system, the early-warning factor set of college students' psychological crisis is constructed, and the dynamic early-warning model of college students' psychological crisis is constructed. Experimental results show that the mining results are consistent with the actual situation, and the grey relational degree, small error probability, and mean square ratio index are better than the present model.

1. Introduction

With the development of the economy, the introduction of social competition mechanism and the enlargement of the enrollment scale of colleges and universities, the conflicts, and pressures in the fields of ethics, values, behavior patterns, interpersonal relationships, employment, and job hunting of college students are increasing, and the probability of psychological frustration is increasing [1]. College students are just coming of age and are about to enter the transitional period of society. They are often faced with pressure and confusion in all aspects of study, employment, love, and family. College students are more likely to have various psychological problems and have low self-healing ability. The obsessive-compulsive symptoms, anxiety, phobia, and psychosis of Chinese college students are significantly higher than the national norm. Therefore, in the process of education and teaching, colleges and universities should not only pay attention to the cultivation of college students' professional learning skills and the improvement of their ideological and moral qualities but should also pay more attention to the mental health of college students. Survey shows that college students have become a high incidence group of mental diseases. College students account for 20% of the patients with mental diseases and 50% of the dropouts due to mental diseases. Moreover, there is an upward trend [2, 3]. Mental diseases generally have four stages of development: psychological crisis ⟶ psychological barriers ⟶ mental illness ⟶ psychological decline. A psychological crisis is an early symptom of mental illness. When a person is confronted with a difficult situation, and his previous way of dealing with the problem and his usual supporting model is not enough to deal with the present situation, that is, the difficult situation he has to face exceeds his ability to deal with the problem, the person will have temporary psychological distress. This temporary psychological imbalance is the psychological crises [4–6]. A psychological crisis is a sign that one is experiencing upheaval

and turmoil in one's life. It temporarily interferes with or disrupts one's habitual pattern of life, characterized by high levels of tension, accompanied by anxiety, frustration, and confusion. The psychological crisis of college students can lead to behavioral or emotional disorders and even lead self-injury or injury and other serious consequences [7]. But people's psychological problems from the emergence to the degree of serious need to go through a gradual development process; if the problem in this process is resolved or alleviated, then the tragedy may not occur [8]. At present, psychological crises caused by psychological problems on university campuses are on the rise, and it is not uncommon to see all kinds of out-of-control behaviors from stress, depression to self-injury, and other injuries. It is very urgent to study the methods of predicting the psychological crisis of college students.

College students' psychological crisis early-warning model refers to a kind of psychological crisis management model by means of a set of scientific early-warning index system and crisis evaluation model, through analyzing and comparing the collected early-warning information, timely discovering and identifying the potential or actual crisis factors, giving out the crisis alarm, immediately activating the emergency plan for crisis, preventing the outbreak of crisis, and reducing the loss of crisis [9, 10]. With the rapid development of science and technology, the quickening pace of life, and the complexity of interpersonal relationship, the pressure and conflict of college students in the fields of values, ethics, behavior, employment, and so on are increasing, the psychological crisis is escalating, and malignant incidents emerge one after another, and psychological problems of college students have gradually become the focus of the whole society. The control and response of college students' psychological crisis are based on accurate identification and early warning [11]. The correlation and redundancy among the indicators of the early-warning model of psychological crisis is one of the effective methods to identify crisis signals and reduce crisis components [12].

Reference [13] established a correlation prediction model between adolescent blind obedience psychology and criminal behavior using the multiobjective evolutionary algorithm. Firstly, the risk factors and protective factors of teenagers' social growth are screened from the perspective of individual, family, school, and society; secondly, the support degree of fuzzy item set of each factor is defined; thirdly, the weight vector of the characteristic index of criminal psychological manifestation of teenagers' blind obedience is calculated; finally, the grey correlation method is used to extract the subjective and objective factors of teenagers' criminal psychology. An accurate correlation prediction model between adolescent blind obedience psychology and criminal behavior is established. Reference [14] takes the learning performance of knowledge and skills of crisis intervention as the intermediary variable to explore the influence process of emotional intelligence on the performance of crisis intervention behavior. Based on the sequence pattern analysis and comprehensive evaluation of learning performance in the previous research, the emotional intelligence of 104 mental health service personnel was collected through the emotional intelligence scale, and the intermediary effect model was used to analyze the relationship between emotional intelligence, learning performance, and behavior performance.

However, the above dynamic early-warning model of college students' psychological crisis cannot obtain the attribute of psychological characteristics, and the prediction error of psychological crisis is large. Most of these studies stay at the level of mental health education, and there is no effective means to predict whether college students have a psychological crisis. People's psychological activities are extremely complex, and the process of change is not visible. The psychological crisis prediction can detect problems as early as possible, and timely intervention measures can be taken to effectively avoid the consequences of the formation and deterioration of the psychological crisis. For this reason, a dynamic early-warning model of college students' psychological crisis based on characteristic attributes is designed. This paper introduces the grey Verhulst forecasting model to analyze the characteristic attributes of college students' psychological crisis. Combining the theory of constraint fuzzy theory from subjective and objective reasons, the essence of college students' psychology is obtained. According to the corresponding information entropy of the behavioral obstacles in the various stages of the college students' psychology, the threshold of the behavior tendency of the college students to produce psychology is calculated. Early-warning indicators is set, and the construction of a dynamic early-warning model of college students' psychological crisis is completed.

2. Dynamic Early-Warning Model of College Students' Psychological Crisis Based on Characteristics

2.1. Research on the Characteristics of College Students' Psychological Crisis. The proposed method uses the grey Verhulst prediction model [15, 16] to study the characteristic attributes of college students' psychological crisis. The grey Verhulst prediction model combines the advantages of the Verhulst model. This model is suitable for non-monotonic swing development sequence or S-shaped sequence with saturation. The characteristics of psychological crisis of college students belong to the category of non-monotonous swinging development, so the grey Verhulst predictive model is used to analyze the attributes of psychological crisis.

The initial value $X^{(1)}(0) = X^{(0)}(1)$ is set, the parameter column $Y = [x, y]^T$ is brought into the whitening equation, the differential equation is solved, and the whitening equation solution corresponding to the grey Verhulst model is obtained as follows:

$$X^{(0)}(t) = e^{at}\left[\frac{(1-e^{-2})}{X^{(1)}(0)} - \frac{y}{x}\right]^{-1}. \quad (1)$$

Let $X^{(0)} = [x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)]$ represent the data sequence, $X^{(0)}(n) \geq 0, n = 1, 2, \cdots, n$ and $X^{(1)}$ represent the first-order accumulation generation sequence of the original data sequence $X^{(0)}$, and its expression is as follows:

$$X^{(1)} = \left[x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n)\right]. \quad (2)$$

Let $\mathfrak{F}^{(1)}$ represent the immediate mean generation sequence corresponding to the original data sequence $X^{(0)}$, and its expression is as follows:

$$\begin{cases} \mathfrak{F}^{(1)} = \left[\mathfrak{F}^{(1)}(1), \mathfrak{F}^{(1)}(2), \cdots, \mathfrak{F}^{(1)}(n)\right] \\ \mathfrak{F}^{(1)}(k) = \dfrac{x^{(1)}(n) + x^{(1)}(n-1)}{2} \end{cases}, \quad (3)$$

where $n = 2, 3, \cdots, n$.

Let Y represent the least square estimation of the grey Verhulst model, and its calculation formula is as follows:

$$Y = [x, y]^T = \frac{B^T}{B}. \quad (4)$$

The effect term of the Verhulst-like model is introduced into the grey Verhulst model to obtain the whitening equation corresponding to the grey GM (1,1) power model and GM (1,1) power model as follows:

$$\begin{cases} X^{(0)}(N) + X^{(1)}(N+1) \geq 0, \\ \dfrac{dX^{(1)}}{dt} + x^{(1)} = y\left(x^{(1)}\right)^{\delta}, \end{cases} \quad (5)$$

When parameter δ has a value of 2, the grey Verhulst model and its corresponding whitening equation are obtained as follows:

$$\begin{cases} x^{(0)}(k) + bz^{(1)}(k) = b\left(z^{(1)}(k)\right)^2, \\ \dfrac{dx^{(1)}}{dt} + bx^{(1)} = b\left(x^{(1)}\right)^2, \end{cases} \quad (6)$$

In the above equation, b represents the ash action and δ represents the development coefficient.

The grey Verhulst model obtains the time response formula by solving the whitening equation:

$$\frac{\hat{x}^{(1)}(N+1) = x^{(1)}(N)}{\left[x^{(1)}(N) + e^{at}\left(X - Yx^{(1)}(N)\right)\right]}, \quad (7)$$

where $N = 1, 2, \cdots, n-1$.

Through the above process, the prediction equation of the Grey Verhulst Model is obtained, and the characteristics of the psychological crisis of college students are studied:

$$X^{(0)}(N+1) = X^{(1)}(N+1) - X^{(1)}(N). \quad (8)$$

2.2. Calculation of the Threshold of College Students' Psychological Behavior Tendency. This section is based on the analysis of the characteristics of college students' psychological crisis and the calculation of the threshold of behavior tendency. In the process of modeling the negative psychological crisis of college students, the fusion of constraint fuzzy theory obtains the essential characteristics of college students' psychology from subjective and objective reasons. Considering the different stages of college students' psychology, the corresponding information entropy of the behavioral obstacles is obtained in each stage of college students' psychology, and the threshold of the behavior tendency of college students is calculated to produce psychology [17, 18]. The process is as follows:

The s et $\{Q(t), q = 1, 2, \cdots, n\}$ represents the mental time series of different stages of college students. Because of the obvious differences in forms of college students' psychology, it is necessary to reconstruct the psychological process in phase space to obtain a matrix as follows:

$$Q = \begin{bmatrix} q(1) & q(1+\tau)\cdots q(1+(m-1)\tau) \\ q(2) & q(2+\tau)\cdots q(2+(m-1)\tau) \\ & \cdots \\ q(n) & q(n+\tau)\cdots q(n+(m-1)\tau) \end{bmatrix}, \quad (9)$$

where ξ represents the psychological complexity of college students, and τ represents the psychological duration, meeting the following requirements:

$$N = (\xi - 1) \cdot \tau. \quad (10)$$

Each row in the matrix is regarded as a variety of influencing factors of psychological changes, and the total number is N. The influencing factors are ranked according to the degree of conflict psychology, and the index of the affected categories of the affected factors in the initial psychological performance state is obtained to obtain the psychological performance state of the affected categories [19–21]. For r dimension college students, the total number of mental state sequences corresponding to different mental states is u, assuming that the probability of the occurrence of different mental state sequences is $P_1, P_2, \cdots, P_k$; then, the probability of mental state sequences is ranked according to the form of Shannon entropy. The amount of information of a piece of information is directly related to its uncertainty. For example, if we want to figure out something very uncertain or something we do not know, we need to know a lot of information. Shannon entropy can effectively solve the problem of quantitative measurement of information. And the permutation entropy is defined by the following formula:

$$D = -\sum_{j=1}^{k}\left(P_j \cdot InP_j\right). \quad (11)$$

The state sequence is integrated given by formula (11), and formula (12) is used to express the behavioral emotional characteristics corresponding to the psychological state at this stage:

$$\frac{0 \leq D = D}{In(\xi) \leq 1}. \quad (12)$$

Suppose that $\{x(t), t = 1, 2, \cdots, N\}$ represents the time series set of behavioral emotional characteristics of college students' psychological state. Due to the complexity of college students' behavioral tendency dominated by

psychology, it is necessary to establish a relationship with psychological duration τ, select the duration series $x(t+\tau)$ to form a new duration stage point column $y(t)$, and determine the psychological duration according to the correlation between $x(t)$ and $y(t)$.

For the time series X and Y of behavioral propensity dominated by subjective will and objective events at two different stages, according to the performance of university students' psychology in work, life, and study [22], the average information amount of psychological behavior hindered by subjective factors and objective factors is selected from the two different stages, and the threshold of behavioral propensity formed by different reasons is calculated respectively by using formula (13) and formula (14):

$$G(X) = -\sum_{i=1}^{n} U(x_i)\lg U(x_i), \tag{13}$$

$$G(Y) = -\sum_{i=1}^{n} U(y_i)\lg U(y_i). \tag{14}$$

Among them, $U(x_i)$ represents the set X of time series of behavioral tendencies dominated by the psychology of college students formed by subjective will, x_i represents the set of characteristic indicators of subjective factors such as emotional fluctuation, weak will, psychological conflict, and ideological extremism extracted, and $U(y_i)$ represents the set Y of time series of behavioral tendencies dominated by the psychology of college students formed by objective events, which is characterized by such objective factors as individualism, pursuit of fame and interests, abuse of power for personal gain, bribery, and parents' emphasis on intellectual education.

2.3. Dynamic Early-Warning Model of College Students' Psychological Negative Crisis. The fuzzy early-warning method based on feature attributes is used to conduct the early warning of psychological crisis of college students [23]. Psychological adaptability, frustration endurance, emotional stability, safe inclination, hedging ability, and temperament are six indicators of college students' early warning of psychological crisis. Based on these six warning indicators, the information of warning indicators is constructed, and the set of warning factors is constructed according to the fuzzy comprehensive warning method and the warning indicator system, which is represented by S [24]. Each index used in college students' early warning of psychological crisis is divided into 5 grades, and the evaluation results of constructing the comment set are divided into excellent, good, medium, poor, and very bad.

In the process of constructing the fuzzy comprehensive evaluation model, weight distribution is a key link. The analytic hierarchy process or the Delphi method are generally selected. The weight distribution can be obtained by the analytic hierarchy process, as shown in formula (4):

$$L = \{L_1, L_2, L_3, L_4, L_5\} = \{0.36, 0.09, 0.14, 0.13, 0.12\}. \tag{15}$$

The membership vectors of all factors are summarized to construct the fuzzy early-warning matrix of Si (i.e., membership matrix), which can be expressed by E_i. The fuzzy comprehensive evaluation vector of each early-warning index is obtained by using the compound operation process of fuzzy matrix:

$$B_i \cdot B_i = A_i \cdot E_i = (b_{i1}, b_{i2}, b_{i3}, b_{i4}, b_{i5}). \tag{16}$$

It is necessary to construct a scientific subordinate function in the early warning of college students' psychological tolerance. Symptom self-rating scale SCL-90 is the most commonly used mental health test in the world. The national norm of the symptom index is 1.43 ± 0.51. The mean 1.43 is chosen as the midpoint, and the standard deviation is taken as the interval to construct the corresponding relationship between the early-warning score of college students' mental endurance and the comment set. Comprehensive early warning is to use the index to judge abnormal state of psychology. Based on the correspondence between the mental health standards and the commentary set, the comprehensive early-warning levels are set as excellent (0–0.4), good (0.4–0.92), medium (0.92–1.43), poor (1.43–1.96), and bad (1.96–2.88). At the same time, the measurement scale vector is set to $L = (0.36, 0.09, 0.14, 0.13, 0.12)$. Thus, the mathematical calculation model of college students' psychological crisis index can be determined as follows:

$$\aleph = E \cdot L. \tag{17}$$

Based on the psychological crisis index of college students, the set of subjective and objective psychological factors formed by college students is obtained, the objective function of the relationship between college students' psychology and their behavioral obstacles is given, and the problem model of college students' psychological negative crisis is established.

The set of all factors is set for college students to have the tendency of rebellious behavior dominated by psychology as follows:

$$\Phi = \{\varphi_1, \varphi_2, \cdots, \varphi_n\}. \tag{18}$$

The objective function of the relationship between college students' psychology and their behavioral obstacles is obtained as follows:

$$Z = \Phi \cdot \sum_{k=1}^{n}\sum_{i=1}^{n} \varphi_n \cdot d_{ik}^2, \tag{19}$$

where $d_{ik} = \|x_k - v_i\|$ represents the correlation strength between each resistance factor x_k and the category v_i of rebellious behavior tendency. Based on the evaluation function φ_n of psychological negative crisis relationship obtained by the above formula, the global optimal solution is solved by the following formula:

$$F = \frac{\sum_{n=1}^{k} (\varphi_n \cdot x_n)^m}{\sum_{n=1}^{k} \varphi_n}. \tag{20}$$

2.4. Model Solving Based on Ant Colony Algorithm. The ant theory is used to solve the model in the above section, thus effectively solving the negative psychological crisis model of college students. The ant colony algorithm is a probabilistic algorithm used to find optimal paths. The algorithm uses a positive feedback mechanism to make the search process converge continuously and finally approach the optimal solution. Therefore, this algorithm was selected in the model solving. The specific process is as follows: suppose that λ represents the intensity of hindrance behavior tendency of college students in t period, Δv represents the pheromone of negative psychological crisis of college students searched by ant k, and $f\,(0 \le f \le 1)$ represents the influence of negative psychological crisis, and the intensity of hindrance behavior tendency of college students in the next period searched is calculated by using the formula below:

$$\lambda\,(t+1) = f \cdot \lambda\,(t) + \sum \Delta\lambda_{ij}\,(t). \tag{21}$$

Assuming that d_k represents the path length traveled by the k ant in this cycle, there are

$$\frac{\Delta\lambda\,(t) = d}{\lambda\,(t+1)}. \tag{22}$$

Assuming that η_{ij} represents the visibility of path (i, j), usually taken as $1/d_{ij}$, set d_{ij} represents the length of path (i, j), the importance corresponding to path visibility is $\beta\,(\beta \ge 0)$, the relative importance of path trajectory is $\alpha\,(\alpha \ge 0)$, U represents the feasible point set, and the transition probability of ant k in t-time domain is $p_{ij,k}\,(t)$, then $p_{ij,k}\,(t)$ is defined by the following formula:

$$p_{ij,k}\,(t) = \begin{cases} \dfrac{[\Delta\lambda(t)]\left[\eta_{ij}\right]^{\beta}}{\sum_{l\in\eta_{ij}} {}^{[\Delta\lambda(t)]}l\left[\eta_{ij}\right]^{\beta}}, & j \in \eta_{ij}, \\ 0, & j \notin \eta_{ij}. \end{cases} \tag{23}$$

According to the above description, the optimization function of college students' psychological negative crisis model is calculated by using formula (20):

$$\min Z = f\,(x) \cdot x \in [a, b]. \tag{24}$$

Suppose m ant is at the equal division of interval m in the initial time domain, λ_{ij} is replaced in formula (19) with λ_j, which is called the neighborhood attraction intensity of ant j, and η_{ij} is defined as $f_i - f_j$, which represents the difference value of objective function and parameter $\alpha, \beta \in [1, 5]$ of college students' psychological negative crisis model. The transfer probability of ants is calculated by the following formula:

$$p_{ij} = \left(\left[\lambda_j\right]\left[\eta_{ij}\right]^{\beta}\right) / \left[\eta_{ik}\right]^{\beta}. \tag{25}$$

The negative psychological crisis model of college students is established by using the following formula:

$$\lambda_j\,(t+1) = \frac{\lambda_j\,(t) + \sum_j \Delta\lambda_j}{\rho}, \tag{26}$$

where $\Delta\lambda_j$ represents the increase of the regional attraction intensity of the j-th ant in this cycle, L_j represents the change of $f\,(X)$ in this cycle, which is defined as $f\,(X + r) - f\,(X)$. The optimization of the function is carried out with the help of the continuous movement of m ants. When $\eta_{ij} \ge 0$ is satisfied, ant i transfers from its nearest neighborhood i to the neighborhood of ant j according to probability p_{ij}; when $\eta_{ij} \le 0$ is satisfied, ant i searches the nearest neighborhood. The search radius is r, that is, each ant either moves to the location of other ants or carries out near neighborhood search and gradually converges to the global optimal solution of the problem.

3. Pyramid Relationship Analysis of College Students' Psychological Crisis Prevention and Intervention

People are products of relationships. The psychological crisis of college students inevitably affects a series of interpersonal relationships. Generally speaking, in the psychological crisis incidents of university students, the individuality of psychological crisis is the student himself, and there are many closely related individuals of psychological crisis based on the individual of psychological crisis, including parents and relatives (dependants, guardians), school teachers and students, staff of government departments, staff of social institutions, and concerned people [25–27]. These related individuals of psychological crisis constitute a huge network of psychological crisis relations. The tight pyramid model of the relationships between and with these crisis-associated individuals is shown in Figure 1, where schools, families, government, and society all have relationships of equal importance to the crisis-associated individuals [28, 29]. However, in the prevention and intervention of college students' psychological crisis in the past, people often pay too much attention to the relationship between schools and individuals in psychological crisis. It can get twice the result with half the effort when we intervene in college students' psychological crisis according to the tight relationship and straighten out all the relationship networks [30–32].

It is the common pursuit and goal of the government, the university, the society, the family, and even the individual students who meet the psychological crisis to make full use of the pyramid model to prevent and intervene the psychological crisis of college students [33]. Although the pyramid model of psychological crisis of college students can achieve the characteristics of abstraction, prominence, and systematic linkage [34–36]; in order to reduce losses, avoid risks, and improve efficiency, it is necessary to carry out advance education and rehearsal of psychological crisis of college students for key groups closely related to psychological crisis.

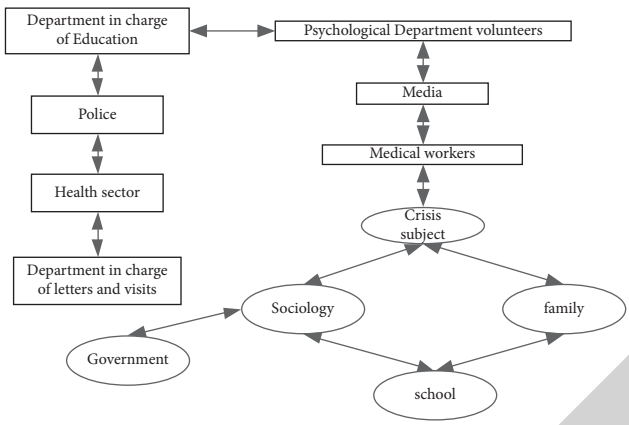


Figure 1: Pyramid relationship model of college students' psychological crisis intervention.

4. Experimental Design and Result Analysis

In order to verify the effect of the dynamic early-warning model of college students' psychological crisis based on perceptual data, this paper selects 4,200 college students as the object of study and uses this model to test the dynamic early-warning of college students' psychological crisis.

4.1. Cluster Mining of Abnormal Mental State. This model is used to mine clusters with an abnormal psychological state in all the selected research objects. Figure 2 shows the mining results of some research objects.

From Figure 2, the results of mining the mental state of some research objects using this model are completely consistent with the actual mental state of the research objects. This shows that the model in this paper can accurately dig out clusters of college students with abnormal mental states from all the research objects. This article uses the Verhulst predictive model to analyze the characteristics of college students' psychological crisis. In the process of research object mining, it can be accurately classified according to characteristic attributes. Early warning of psychological crisis for the excavated clusters can improve the early-warning efficiency of this model.

4.2. Psychological Crisis Early-Warning Results. Based on the abnormal psychological state of college students' cluster discovered in the previous experiment, this paper uses the method of six indicators such as psychological adaptability, frustration tolerance, emotional stability, security tendency, hedging ability, and temperament to warn the subjects of the abnormal psychological state cluster of a mental crisis and determines the trend of the overall psychological state of the subjects according to the summary of various indicators. Some results are shown in Tables 1–3.

Analysis Tables 1–3 can be obtained; most of the selected study subjects are at the medium level of psychological crisis ability; a very few of the study subjects are poor or very poor in psychological crisis ability, so it is necessary to adopt different measures through the relevant teachers to carry out psychological health education and psychological counseling. The method in this paper obtains the essential causes of college students' psychological crisis according to the theory of constraint fuzzy theory and can accurately classify the psychological state of students according to the index of psychological crisis of students. According to the results of the threshold value of students' mental behavior tendency calculated in this paper, the emotional stability and frustration tolerance of students can be accurately identified. Based on the above data, the analysis of the overall mental crisis ability of college students can help colleges and universities to grasp the direction of mental health education.

4.3. Comparative Test. Taking 4,200 students in a university as the research object, the comparative test data comes from about 4,500 original records of student mental health assessment in the past two years. Then, the original data are preprocessed according to the threshold of psychological behavior tendency, and the result of psychological crisis early warning is taken as a reference. The grey correlation degree, small error probability, and mean square deviation of the psychological crisis early-warning results of the method of this paper, the method of reference [4], and the method of reference [5] are calculated through the test platform.

4.3.1. Comparative Test of Grey Correlation Degree, Small Error Probability, and Mean Square Deviation Ratio. In order to verify the overall effectiveness of the proposed method, the proposed method is tested in the TTE network platform developed by visual c + +. Taking grey correlation degree γ, small error probability P, and mean square deviation ratio C as test indexes, the proposed method, literature [4] method, and literature [5] method are tested. The calculation formulas of test indexes are as follows:

$$\gamma = \frac{1}{n}\sum_{k=1}^{n}\delta(k),$$

$$P = P\left\{\left|\Delta^{(0)}(k) - \overline{\Delta}^{(0)}\right| < 0.6745S_1\right\}, \tag{27}$$

$$C = \frac{S_2}{S_1},$$

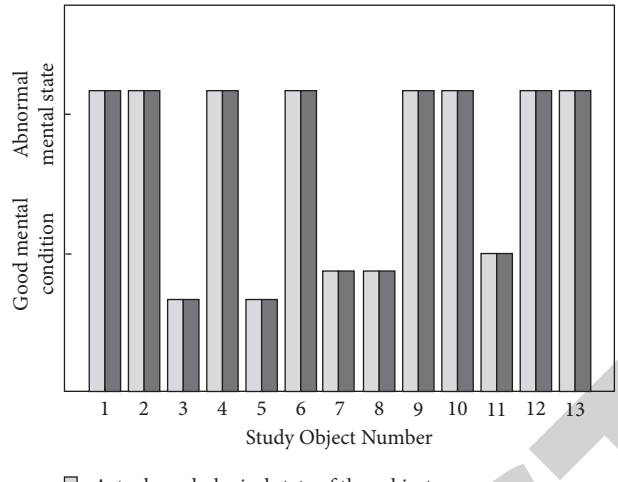


Figure 2: Mining results of some research objects.

Table 1: Test results of six indicators of college students.

Index psychological adaptability	Student number			
	1	2	3	4
Frustration tolerance	90	95	74	64
Emotional stability	92	86	83	98
Safety tendency	95	98	86	83
Hedging ability	97	98	92	77
Index	88	90	98	62

Table 2: Emotional stability test results.

Gender	Male	Female	Total
Excellent (number)	146	56	180
Good (number)	198	80	258
Medium (number)	1,189	558	1,790
Poor (number of people)	116	69	182
Poor (number of people)	5	56	11

Table 3: Test results of frustration tolerance.

Gender	Male	Female	Total
Excellent (number)	245	58	170
Good (number)	88	82	181
Medium (number)	1,546	492	1,888
Poor (number of people)	125	115	219
Poor (number of people)	25	20	49

where $\delta(k)$ represents the correlation coefficient; $\Delta^{(0)}(k)$ represents absolute error sequence; $\overline{\Delta}^{(0)}$ is the mean value corresponding to $\Delta^{(0)}(k)$; S_2 is the variance corresponding to $\Delta^{(0)}(k)$.

Figure 3 shows the grey correlation test results of the proposed method, literature [13] model, and literature [14] model. The higher the grey correlation, the more accurate the change of college students' psychological crisis characteristic attributes analyzed by the method. By analyzing the data in Figure 3, it can be seen that the grey correlation of the score analysis results obtained by the proposed method in multiple iterations is higher than that of the other two methods.

Figure 4 shows the small error probability of different methods. The smaller the small error probability, the greater the error of the result of the characteristic attribute analysis

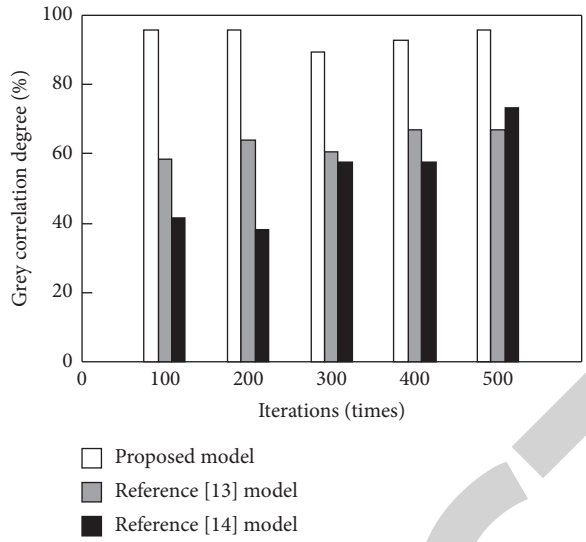


Figure 3: Grey correlation degree of different models.

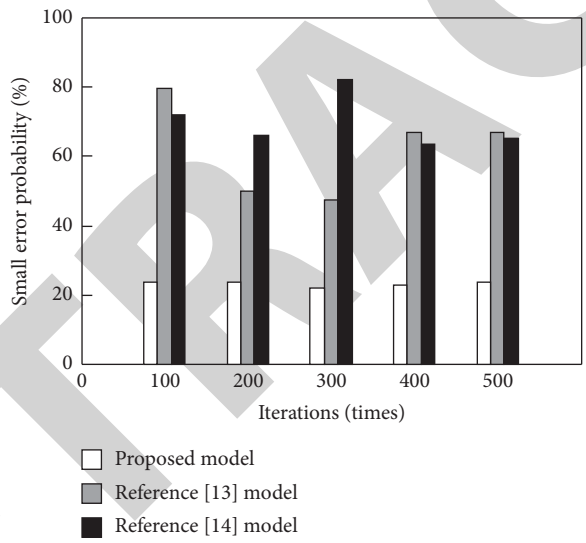


Figure 4: Small error probability of different models.

of the method in the evolution process of college students' psychological crisis. It can be seen from Figure 4 that the small error probability of the analysis results obtained by the literature [13] model and the literature [14] model is lower than the small error probability of the analysis results of the proposed method.

Figure 5 shows the mean square deviation ratio obtained by different methods in multiple iterations. The larger the mean square deviation ratio, the greater the error of the analysis results. The mean square deviation ratio obtained by the proposed method in multiple iterations is within 0.02, which is far lower than the mean square deviation ratio of the methods in literature [4] and literature [5]. The method in this paper constructs a fuzzy comprehensive early-warning model of characteristic attributes after obtaining the information of the mental state of college students. The determination of early-warning indicators helps to predict the psychological crisis of college students in time. Therefore, the method in this paper has great advantages in the test of grey correlation degree, probability of small error and mean square error, and other indicators.

4.3.2. Comparison of Early-Warning Accuracy of Psychological Crisis Model. The early-warning accuracy of this model and the two comparison models is compared, and the results are shown in Figure 6.

According to the analysis of Figure 6, among the early-warning results of the three early-warning models on the psychological crisis of the research object, the early-warning results of this model completely coincide with the fitting curve of the actual situation. Because this article sets early-warning indicators, it obtains information on the early-warning indicators of college students' psychological crisis. And according to the fuzzy comprehensive early-warning method and the early-warning index system, the early-

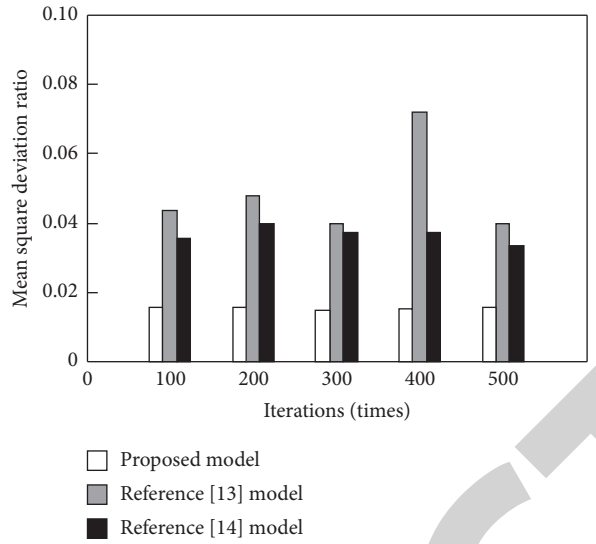


Figure 5: Mean square deviation ratio of different models.

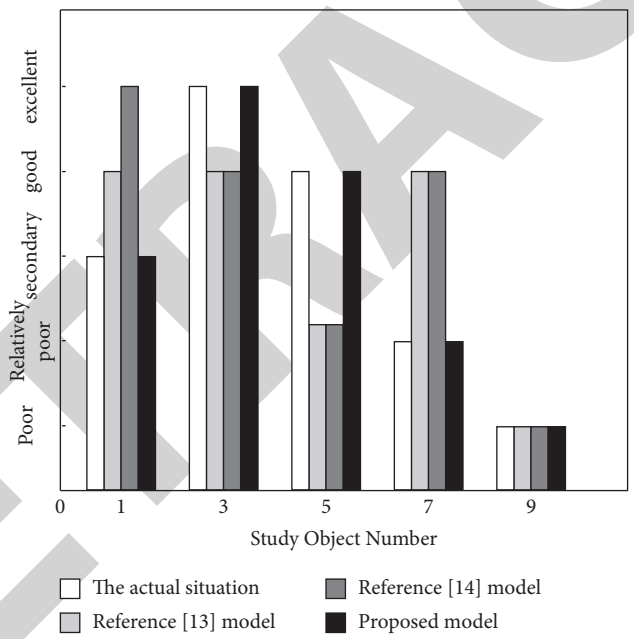


Figure 6: Comparison of early-warning results of different models.

warning factor set of college students' psychological crisis is constructed, thereby improving the accuracy of psychological crisis early warning of the method in this paper, which shows that the early-warning accuracy of this model is significantly higher than that of the other two early-warning models.

4.3.3. Complexity Comparison of Psychological Crisis Early-Warning Models. Early-warning efficiency is an important test direction of the college students' psychological crisis early-warning model. Figure 7 shows the comparison results of early-warning efficiency between this model and the comparison model, in which the early-warning efficiency is reflected by the index complexity in early warning.

It can be seen from Figure 7 that among the three early-warning models, the index complexity of the model in this paper is controlled below 20%, which is significantly reduced compared with the two comparative models. This is because before constructing the dynamic early-warning model of psychological crisis, this article analyzes the psychology of students at different stages in advance and calculates the threshold of psychological behavior tendency based on the behavior information entropy. These processes process psychological data in advance and simplify the process of dynamic early warning of psychological crises. Therefore, it

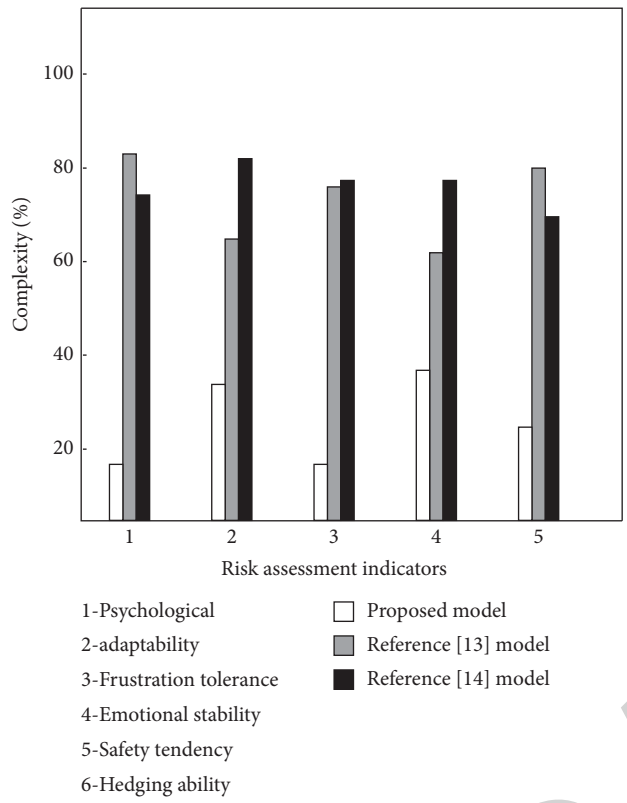


Figure 7: Comparison of early-warning results of different models.

shows that the early-warning indicators used in this model are relatively simple and conducive to the improvement of early-warning efficiency.

5. Conclusion

Psychological crisis intervention is the most important link to relieve the psychological pressure of college students and avoid mental illness. Therefore, it is important to understand the psychological state of college students in time and predict the psychological crisis as early as possible. In order to solve the problem of large error in the traditional early-warning model of college students' psychological crisis, this paper introduces the grey Verhulst method to analyze the characteristics of college students' psychological crisis, constructs the dynamic early-warning model of college students' psychological crisis, calculates the threshold value of psychological behavior tendency, perceives the information of college students' psychological state, sets the warning index, and realizes the dynamic warning of college students' psychological crisis. And the early-warning accuracy of the method in this paper is higher than the comparison of the two methods, which is consistent with the actual situation. Because the method in this paper preprocesses the psychological data and optimizes the calculation process, the computational complexity of the model is also low. The above experimental results prove that the constructed model has better performance in early warning of college students' psychological crisis. This article has achieved the purpose of predicting the psychological crisis of students more accurately. The data selected this time are conventional student data information, and the actual types of psychological crisis attributes are far more than these. In future research, we consider adding data on student internships, work, and other attributes to improve the prediction range of the model.

Data Availability

The raw data used to support the conclusions of this article are available from the corresponding author upon request without undue reservation.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] A. R. G. Mónica and B. A. S. Gutiérrez, "Influence of diversity in lectures on the students' learning process and on their perspectives about renewable energies in an international context - the students' view," *Solar Energy*, vol. 173, no. 5, pp. 268–271, 2018.

[2] Y.-H. Wang, Q.-Y. Luo, and Z.-T. Shi, "A study on the relationship between sexual morality education and mental health level of Chinese female college students," *Medicine*, vol. 98, no. 19, p. e15576, 2019.

[3] A. Fossas, "Psychological maturity predicts different forms of happiness," *Journal of Happiness Studies*, vol. 20, no. 6, pp. 1933–1952, 2019.

[4] Y. Jia, Z. J. Hou, and D. Wang, "Calling and career commitment among Chinese college students: career locus of control as a moderator," *International Journal for Educational and Vocational Guidance*, vol. 21, no. 2, pp. 211–230, 2020.

[5] N. Kupper, M. Jankovic, and W. J. Kop, "Individual differences in cross-system physiological activity at rest and in response to acute social stress," *Psychosomatic Medicine*, vol. 83, no. 2, pp. 138–148, 2020.

[6] N. Jiang and W. S. Xu, "Analysis of students' Campus consumption behavior based on data mining technology," *Popular Science and Technology*, vol. 1, no. 1, pp. 26–36, 2019.

[7] B. W. Cowan and Z. Hao, "Medicaid expansion and the mental health of college students," *Health Economics*, vol. 30, no. 6, pp. 1306–1327, 2021.

[8] G. Sn and H. Meena, "Socio-demographic factors and locus of control on mental health among college students," *International Journal of Scientific Research*, vol. 15, no. 1, pp. 53–55, 2021.

[9] M. A. Loftis, T. Michael, and C. Luke, "College student suicide risk: the relationship between alexithymia, impulsivity, and internal locus of control," *International Journal of Educational Psychology*, vol. 8, no. 3, pp. 246–257, 2019.

[10] W. Lu, G. L. Xu, Y. H. Cheng, K. Y. Li, and B. Wang, "Soft classification and recognition of human behavior based on local spatiotemporal features," *Computer and Modernization*, vol. 41, no. 3, pp. 94–98, 2019.

[11] X. D. Gao and X. K. Hao, "Student psychological intervention and behavior monitoring system based on machine vision," *Modern electronic technology*, vol. 43, no. 24, pp. 71–74, 2020.

[12] S. E. Williams, D. M. Sarno, J. E. Lewis, M. K. Shoss, M. B. Neider, and C. J. Bohil, "The psychological interaction of spam email features," *Ergonomics*, vol. 62, no. 8, pp. 983–994, 2019.

Hindawi
Scientific Programming
Volume 2022, Article ID 6538069, 13 pages
https://doi.org/10.1155/2022/6538069

Research Article

Research on Optimization and Application of University Student Development and Management Strategy Driven by Multidimensional Big Data

Zhimei Lv

Engineering Faculty of Economics, Henan Institute Economic and Trade, Zhengzhou, Henan 450018, China

Correspondence should be addressed to Zhimei Lv; lvzhimei@henetc.edu.cn

Received 17 December 2021; Accepted 17 January 2022; Published 8 February 2022

Academic Editor: Baiyuan Ding

The purpose of education is to enable students to develop fully, freely, comprehensively, and harmoniously. The development of college students is a social problem that must be focused on in higher education under the background of China's powerful human resources. It is an important subject that higher education must attach great importance to and solve to find out the way and method to solve and avoid the development of college students to help them face risks correctly, meet challenges, and grow healthily. Under the background in the era of big data, Internet and mobile intelligent terminal has great popularity in colleges and universities, especially in the digital library, the official platform construction, and is widely used in Internet multimedia technology in college classroom, although for college students and the teacher provides vast amounts of information data, exploring the college students' horizons, enriching the students' knowledge structure. But the explosive growth of data and information security threat to the development of college students management work still bring the severe test of this; education workers in colleges and universities should comply with the development of The Times, the big data organic blend in the current education system, and the specific practice of education, for modern education practice and college students' all-round development's age characteristic new development mode and effective way. In order to adapt to the new form of the development of The Times, the management of colleges and universities should keep pace with The Times, and integrate the concept of big data and advanced technology into the education management work, so as to realize the national deployment of the education system innovation development strategy, explore teaching rules, students' growth and development tasks, and other realistic needs of the times. In addition, making big data the most powerful internal driving force for educational development helps colleges and universities to realize the frontier, timeliness, interactivity, and individuation of educational management. Colleges and universities should actively apply the idea of big data, improve the data resource integration and use ability of the staff in the field of education management through diversified means, improve the management level in multiple dimensions, and promote the updating and upgrading of the management structure of higher education and teaching. By analyzing the background of big data era, this paper analyzes the challenges of college students' management in the era of big data and puts forward the strategies of college students' management in the era of big data, so as to improve the quality and efficiency of college students' management.

1. Introduction

'College students' development' is a relatively broad concept. At present, the academic circles at home and abroad have not formed a consensus on its meaning. Overall, due to focus on topics or scholars in the emphasis of study on the definition of the elements or dimensions "college students" development in scattered, in summary mainly two levels, namely, the development of cognitive ability (IQ) and the development of cognitive ability (quotient and), are receiving higher education. College students' knowledge, ability, social relations, quality, and personality can be developed in an all-round way. The development of college students cannot be separated from specific historical conditions, and the theme of the era often becomes the theme of college students' development [1–3]. College students'

development refers to the students at the university of this particular time and space scope is not only the development of specific tasks to complete the university stage, in order to achieve the basic curriculum standards of graduating from college; what is more important to within limited time and space form can make college students after graduation life still can let themselves get comprehensive and ability of sustainable development [4, 5]. In other words, the development of college students is an internal and gradual growth process of college students with development tasks as the goal orientation.

To accurately grasp the real situation development is the premise and the study of college students' development. Environment is the external influencing factor of college students' development, and the change of environment often leads to profound changes in college students' behavior and lifestyle. Scholars mainly start from the social environment and educational environment of college students' development, through studying the change of the environment and its impact on college students' development, discuss how college students choose the best, and remove the bad in the environmental transformation to promote their own development [6]. The decrease of the social status of the students, the ideological and moral quality of ordinary questioned, learning stress, the growth of employment prospects uncertain, psychological problems and psychological barriers, mental disease frequently occurred, etc., are the outline of the growth of contemporary college students; this group development needs to cause the social, especially the higher education.

There is no doubt that higher and newer requirements are put forward for the education quality of college students. The traditional teaching mode is relatively rigid, education system is lack of necessary, consistent with the facts, and the timeliness of data information support makes the current college education practice in the ideological content and scientific shortcomings in one way or the other difficult to be fundamental for college students to provide an influential modern education content of keeping the pace [7]. Therefore, college education should be inclined or give consideration to digitalization and networking to meet the needs of the development. From the macro level, relying on big data can make educators timely obtain the real dynamic information related to college students' thoughts and behaviors, and can analyze and track it and predict its prospects, so that the importance of data information to know the behavior dynamics of college students is self-evident; from the micro level, according to the big data to analyze the ideological situation, behavior habits, and values of college students, explore their internal roots and influencing factors, so as to develop practical and highly targeted education programs. Therefore, through the analysis of the collected, screening, and college students' information data can be found in time and accurate prediction for the college students during the period of the thought and behavior development and can also find their aptitude or the circumstances of education method; education under the big data is more scientific, precision, accuracy, effectiveness and flexibility [8].

Big data provide massive or all data, which are more conducive to quantitative analysis and precise guidance; it reflects the general direction of the dynamic development of things, which is more conducive to grasp and control on the macro level [9]. It gives the correlation between things, which is more conducive to decision-making and targeted problem-solving. Therefore, big data are of more and more attention and aroused people's attention, both academic, business, enterprise, and government; almost all industries or different use big data acquisition, collection of the platform selection, and quantitative analysis of huge amounts of information data, greatly improving the work efficiency, and promote the rapid development.

Many universities in the teaching, learning, research, general services, library, and other business departments have to establish and perfect the relevant data of information system and accumulated a lot of college students information data, such as the recruitment of students employment, course selection, test scores, poor students archives, consumer spending, library access records, etc., and all of these data contain a multitude of complicated correlations, from consumer spending to the amount of poverty subsidies to course choice to grades [10]. As an organic whole, the data and information resources of various departments should be reasonably shared or seamlessly connected, but the fact is not so. First, there are problems such as weak connection, small horizontal data flow, and insufficient information resource integration among various business departments. Second, the data information standards of each business system information data platform are different, with various formats and poor accuracy, making it difficult to share resources among databases. In a word, the phenomenon of "information data isolation" among various departments in most colleges and universities is quite serious, which has become the bottleneck problem that information data resources of various departments in schools cannot be shared.

2. Student Development and Management Strategies

As globalization penetrates from the economic field to the educational field, the development of college students is confronted with the value conflicts of tradition and openness, eastern culture and western culture, and social interests and personal interests. At the same time, most families in China lack rational thinking about independent education for their children, and the educational management of colleges and universities is also largely deficient in this aspect, which causes a series of social problems. Universities should optimize the content of educational management. To cultivate the students' independent ability and self-education ability in important position, cultivate their good psychological quality, improve their social adaptation ability and the ability to withstand setbacks, to strengthen their emotional education and moral education, promote comprehensive coordinated development, and cultivate talents with high comprehensive quality and spiritual character. The development process of individual social attribute has regularity, objectivity, universality, and particularity

[11–13]. A comprehensive survey of the basic situation in colleges and universities shows that educators do not pay enough attention to it. First, there is a general lack of overall and overall top-level design and the lack of guiding specific educational practice with the concept of big data. Second, many higher education workers do not know much about the concept of big data, let alone apply it to educational practice. Third, the lack of grasp of the character and behavior dynamics of college students unable to actively use big data to collect relevant information data and naturally unable to find a more willing form of communication; the traditional teaching habit of reading from a book is hard to recover [14].

From the perspective of college student management, students are the "absolute" subject and the primary goal and fundamental starting point of education. College students mainly refer to the cultivation of students with excellent ideological and moral quality, reasonable knowledge structure, excellent ability and quality, profound scientific and humanistic quality, positive personality psychology, healthy body, and appropriate degree of socialization in the development process and finally achieve comprehensive development. Teachers should strengthen the political theory on "thought" learning, unswervingly adhere to the student's main body status, give the students more "humanistic care," close to the student's learning, life, and thoughts, focus on the students' inner feelings and reality need, understand the autonomy and initiative of students, really respect students, concern for students, to cultivate students, take the overall development of students as the fundamental, value realization as the goal, and do "everything for students." The action should strengthen their own vocational skills training, at the same time adhere to the subject status of students unshaken, and change the past cramming, exam-oriented, didactic teaching way, to achieve personalized, humanized, humane teaching [15]. The management of college students must face all students and be committed to the comprehensive development of all students. College teachers must remove colored glasses and respect, love, and care for every student. In terms of role positioning, teachers must realize the importance of equality and reposition themselves as guides and cooperators. At the same time, students should increase their sense of identity, timely eliminate their inner "subordinate" understanding, good at expression, and the courage to act. In the system construction, we should strengthen the participation of students, optimize the structure of student organizations, expand the function of student organizations, respect and protect the rights of students through the system construction, and wake up the sense of ownership of students. Process diagram of knowledge discovery is shown in Figure 1.

The growth and development of college students is a complex system development process, which is not only the physical nature of individual growth and maturity, but also the construction of complex social relations. Since entering the 21st century, big data can increasingly reflect the specific laws of the development of things. With the gradual maturity of data mining technology, compared with personal experience, decisions made based on big data are more objective and effective. On college students' education policy formulation, the traditional meaning of the research type and decision questioned by many, but in the new century is based on the context of the dynamic attention by the people, is a decentralized, constantly updated, multidimensional, can store, classification and analysis of different regions and education subject and the object of information data [16, 17]. Large-scale educational data can also be obtained from the websites of the Ministry of Education and local education administrative departments and college students' education. The basic data obtained through these means can be used to formulate sound education policies. Thus, for college students' education, data collection will make education more convincing and credible. Big data can make educational activities achieve maximum effect and adjust educational methods and means at any time. Big data are real time, dynamic, and forward-looking. In order to provide corresponding feedback and guidance, with the deepening of learning activities, students' learning database and background information system, past assessment, and learning data will be constantly updated [18].

In the context of big data, the teaching data can be subdivided to an incredible extent through computer intelligent teaching and evaluation system, and the massive data generated by students inside and outside the classroom can be mined to record the data generated by students in the classroom. Therefore, big data analysis can fully and truly reflect the quality of education, making education personalized, diversified, and accurate. From the aspect of students, we can dig out their education level, cultural habits, and so on; at the national level, data such as education input and relevant curriculum can be collected through tests and questionnaires. For example, it can understand the grasp of classroom knowledge and the accuracy of notes, the notes learned in class [19]. It can understand students' learning rules and predict students' learning behavior. And in order to make students present as holographic data, big data can pay attention to students' health data, social communication data, etc., can collect multidimensional data about students' learning, so that educational institutions and educators pay attention to what students are good at, and guide students' sustainable development.

By tracking students' online learning trajectory and analyzing each forum post that students read, we can find the best teaching methods and determine which forum posts are most suitable for students to read. For example, when students watch teaching videos, analyze which section of content can be directly skipped, which section can be replayed, and so on. The analysis of these massive data can fully give college students with autonomy and choice in learning and can completely change the current backward teaching methods. Online education providers can formulate the following learning courses, progress, and difficulty according to students' learning conditions, so as to avoid the situation of losing one to the other. For example, due to different regions and educational levels, students' Chinese foundation is too different, which leads to the poor teaching quality of college Chinese in public

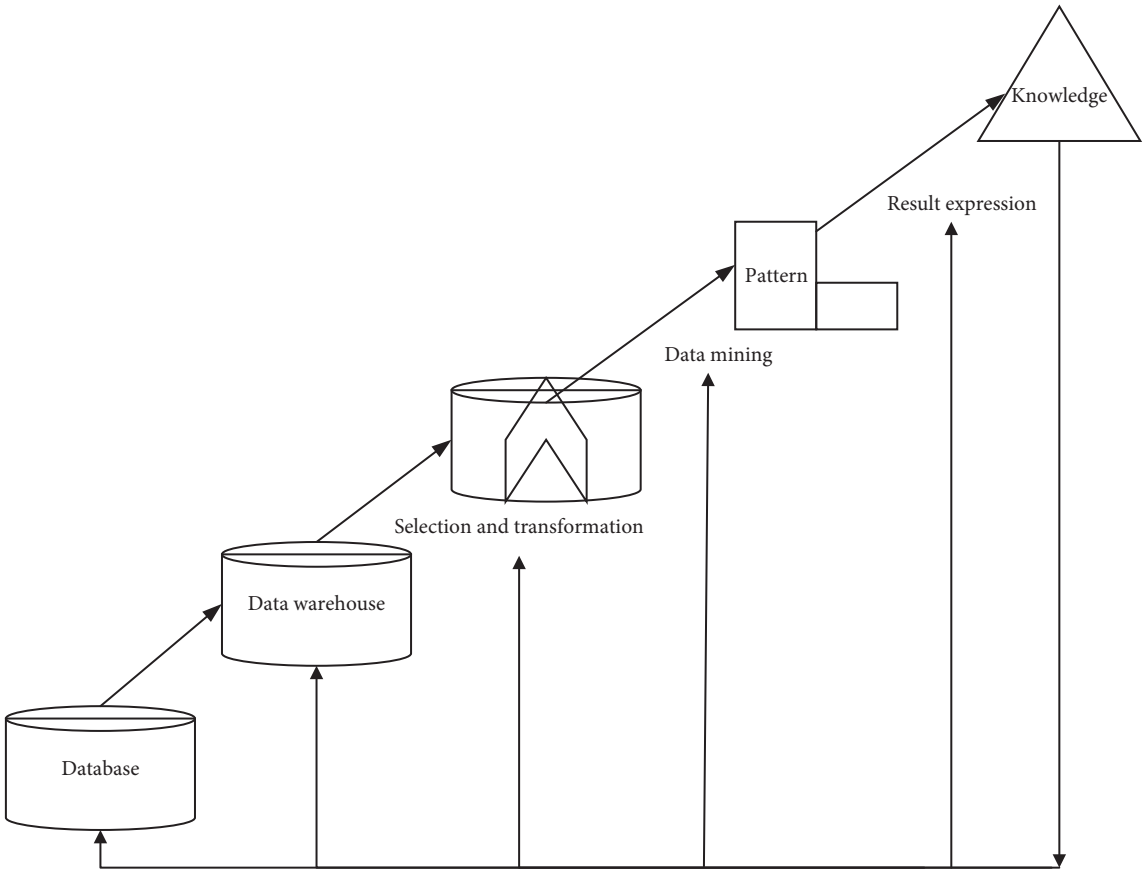


Figure 1: Process diagram of knowledge discovery.

courses. The results obtained through the analysis of students' learning situation through big data can maximize the classroom effect, but also enable teachers to adjust the content [20].

In short, people are still accustomed to the traditional education mode of light driving, and the awareness of big data utilization needs to be improved. The sample size of big data is large and complicated, and the structure is diverse, and the distribution is extremely extensive. The information data generated by college students every day in the library, classroom, dormitory, and so on are enormous and exist in the form of chaotic fragments. The collection and storage of big data is the primary problem that people are faced with. Big data types and sources of information are widely multifarious, generate and update frequency is high, how to carry on the scientific and reasonable collection, storage, and thus have a complete information data, how to select the timeliness of information data sorting, finally get scientific analysis results, and it is also higher educators are facing a big problem. Renderings of two algorithms are shown in Figure 2.

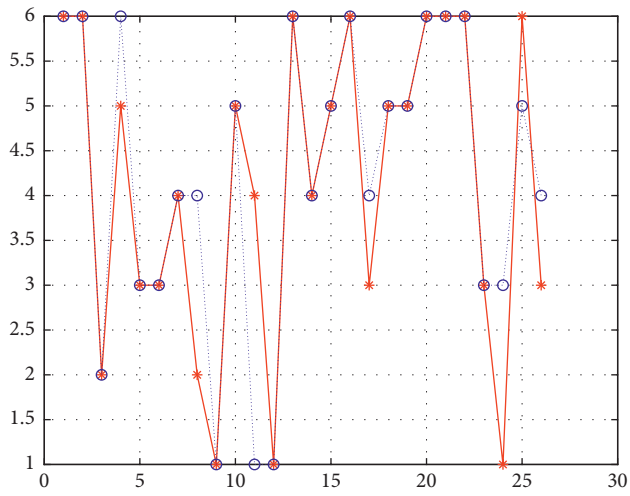


Figure 2: Renderings of two algorithms.

3. Multidimensional Big Data-Driven Approach

The first category is predicting the unknown. The second category is automatic partitioning of vast amounts of unlabeled data and automatic clustering of similar things. Classification is mainly the classification of new things. At present, there are many very successful classifiers, such as decision tree, naive Bayes, artificial neural network, support vector machine, etc. All of these algorithms can be classified, and some algorithms can build good structural models, which can explain the whole data set well [21, 22]. For example, psychology uses clustering to study the causes of depression, biology uses automatic classification of

organisms, and information retrieval uses clustering to group similar web pages together. The following uses C4.5 as an example to describe the process of building a decision tree.

The calculation formula of information entropy of each node is as follows:

$$\begin{aligned} \text{gain} &= \text{info}(T) - \sum_{i=1}^{s} \frac{|T_i|}{T} \times \text{info}(T_i), \\ \text{info}(T) &= -\sum_{j=1}^{N\text{Class}} \frac{\text{freq}(C_j, T)}{|T|} \times \log_2\left(\frac{\text{freq}(C_j, T)}{|T|}\right). \end{aligned} \tag{1}$$

where, T is the data set owned by the node and $\text{freq}(C_i, T)$ is the frequency of each $i \in [1, N\text{class}]$.

The calculation principle of local consistency is relatively simple, mainly using the Kendal Concorde coefficient, and the specific calculation formula is as follows:

$$W = \frac{12\sum (R_i)^2 - n(\overline{R})^2}{\mathrm{K}^2\left(n^3 - n\right)}, \tag{2}$$

where K is the number of evaluation, n is the number of individuals, R_i is the sum of the rank of the i-th individual, and $\overline{R}$ is the average value of all individuals.

The calculation formula of single-sample statistics is as follows:

$$t = \frac{\overline{X} - \mu}{\delta_x / \sqrt{n-1}}, \tag{3}$$

where t is the statistics of the sample mean and the population mean, $\overline{X}$ is the sample mean, μ is the population mean, δ_x is the standard deviation of the sample, and n is the number of samples.

The calculation of the statistics of the hypothesis of the single-body sample is as follows:

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2 / n_1 + n_2 - 2\left(1/n_1 - 1/n_2\right)}}, \tag{4}$$

where $\overline{X}_1$ and $\overline{X}_2$ are the mean of the sample, S_1^2 and S_2^2 are the mean of the sample, and n_1 and n_2 are the number of samples.

Set two random sequences X and Y, Pearson correlation coefficient between the two sequences is r, and then

$$r = \frac{\text{cov}(X, Y)}{\sqrt{\sigma_x^2}\sqrt{\sigma_y^2}} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 (y_i - \overline{y})^2}}, \tag{5}$$

where σ_x^2 and σ_y^2 is the variance of random sequence X and Y, $\text{cov}(X, Y)$ is the covariance of random sequence X and Y, $\overline{x}$ and $\overline{y}$ are the mean value of variables, and n is the length of sequence X and Y.

In order to ensure the accuracy of the results, this paper uses two evaluation indexes, mean absolute error, and root mean square error, to evaluate the optimization effect of the model. The specific calculation formulas are as follows:

$$\begin{aligned} \text{MAE} &= \frac{1}{s}\sum_{i=1}^{s}\left|\widehat{y}_i - y_i\right|, \\ \text{RMSE} &= \sqrt{\frac{1}{s}\sum_{i=1}^{s}\left[\widehat{y}_i - y_i\right]^2}, \end{aligned} \tag{6}$$

where y_i is the actual value of the variable, $\widehat{y}_i$ is the predicted value of the variable, and s is the number of predicted samples.

There are many ways to normalize data, but one common goal is to make data dimensionless between data. This paper summarizes the following data standardization methods based on several literature.

Range transformation formula of cost index is as follows:

$$y_{ij} = \frac{x_{ij} - \min\limits_{i\in\{1,2,\ldots,n\}} x_{ij}}{\max\limits_{i\in\{1,2,\ldots,n\}} x_{ij} - \min\limits_{i\in\{1,2,\ldots,n\}} x_{ij}}, \tag{7}$$

where x_{ij} is data sample.

The transformation formula of fixed indicators is as follows:

$$y_{ij} = \begin{cases} \dfrac{\min\left|x_{ij} - a_j\right|}{\left|x_{ij} - a_j\right|}, & x_{ij} \neq a_j, \\ 1, & x_{ij} = a_j, \end{cases} \tag{8}$$

where a_j is the optimal stable value of the j-th indicator f_j.

Interval transformation formula is as follows:

$$y_{ij} = \begin{cases} 1 - \dfrac{q_1^j - x_{ij}}{\max\left\{q_1^j - \min\limits_{i\in\{1,2,\cdots,n\}} x_{ij}, \max\limits_{i\in\{1,2,\cdots,n\}} x_{ij} - q_2^j\right\}}, & x_{ij} < q_1^j, \\ 1, & x_{ij} \in \left[q_1^j, q_2^j\right], \\ 1 - \dfrac{x_{ij} - q_1^j}{\max\left\{q_1^j - \min\limits_{i\in\{1,2,\cdots,n\}} x_{ij}, \max\limits_{i\in\{1,2,\cdots,n\}} x_{ij} - q_2^j\right\}}, & x_{ij} > q_2^j, \end{cases} \tag{9}$$

where $[q_1^j, q_2^j]$ is the optimal stability interval of f_j.

And based on the current social development, focus on shaping the concept of big data, gradually improve and optimize the campus information management platform, create a first-class professional data faculty, and accelerate the implementation of campus big data innovation management work at multiple levels for college students to create the campus order conducive to their deep learning [23]. Colleges and universities can build a convenient

platform and innovate the mode for education management through big data, which can also help universities achieve the goal of information-based supervision and facilitate the construction of campus internal information management. Colleges will create an orderly, harmonious, and humanized management atmosphere in the process of properly planning and building the educational management structure. In order to cater to the trend of social development and meet the demands of current educational management, universities should focus on the integration into educational management and carry out educational management upgrading centering on information innovation.

3.1. Shaping "Big Data View." In order to ensure the routinization and systematization of educational management, school leaders should lead in-service teachers to shape correct "big data view" and strengthen the effectiveness of educational management. First, when implementing educational management, the school should take data as the core basis, establish the management concept of "handling affairs with data and managing with data," and speed up the transformation of information educational management. Collect, integrate, analyze, and distinguish the existing educational resources in the campus, the teachers and staff groups, the number and scale of students, the latest educational requirements, the basis of education management, and other aspects of data, strengthen the management of campus data resources, and ensure that it matches with the teaching management, to ensure the effectiveness of management; Second, when building the big data platform for education management, the school should organize education managers to hold relevant meetings on the confidentiality of data resources to emphasize the risk of information resource disclosure, so as to maintain the stability and security of the operation of the data platform. Three inside the campus is big data concept expansion training activities; leaders at all levels and teacher encouraged to actively participate, by learning the latest management idea, instead of the old, do not cut time management concept, forming large data thinking, actively explore the education management channel, increase management of modern elements, and ensure that the new educational management concept meets students' current psychological expectations [24].

3.2. Information Platform Construction. As the core of multidirectional application of university data resources, the construction of university's own information platform is extremely important. At present, local colleges and universities in China often encounter many obstacles in the establishment of information platform, such as lack of capital and human resources, resulting in the platform's single function, poor experience, lack of personalized services, and other problems. In order to completely solve the above problems and complete the task of constructing big data platform, colleges and universities can adopt the following methods to solve them: first, integrate data. The detailed information data of the campus are analyzed in a centralized manner, and the data with distinctive characteristics are analyzed to explore its causes and unified input of sorted data resources into the data platform, reflecting the characteristics of open platform; second, pay attention to the financial capital reserve data. Carry out a rigorously situation of colleges and universities, and make a detailed budget for the funds used in platform construction. If there is still surplus funds after the platform is perfected, innovation sections can be continuously added to improve the comprehensiveness of the platform when it is put into use in the future. If the cost is insufficient, colleges and universities can seek investment help from enterprises. The university can determine the quality of its products by itself and obtain platform investment through internal publicity on campus [25].

3.3. Data Team Construction. In-service teachers are the key group in the educational management. In order to make their comprehensive ability and quality match the cutting-edge information education management, universities should pay attention to the professional training of teachers. According to their job functions, college teachers can be divided into two categories: one is the teachers who participate, and the other is the administrative teachers who carry out management work. First, cultivate the big data concept of full-time educational management teachers. The school should be regularly invited big data information industry pundits, campus at all levels of line management experts and scholars such as teachers, information platform maintenance upgrade, big data shaping skills such as content, information platform launched special training lectures, strengthen management teachers' ability to use big data technology, and guarantee the education management effect. Second, encourage teachers management in the free time to learn the advanced education management through multiple channels, information platform, big data technology, and related theory knowledge, such as the Internet, radio and television, books, newspapers and magazines, etc., make its armed mind, and strengthen the big data effectiveness of education and teaching management, to ensure smooth implementation of education management work, and finally, the landing of collaborative innovation [26].

3.4. Formulating a Management System. In this regard, during the formal planning and construction of the system, the school should highlight the following points to enhance innovation efficiency: first, the collection system. Much data will be generated in the daily management, and a large data storage database will be established over time. In the absence of specific detailed management, data resources can not be properly utilized, but also cause problems such as data accumulation disorder and loss [27].

The overall normative teaching standards, teaching language, teaching methods, educational concepts, and other aspects are set as the content of data collection, and unified standards of data entry are set to avoid the risk of invalid data. In this way, teachers and students can timely

harvest the real-time refinement of teaching and learning and carry out targeted optimization and improvement on the hidden problems, so as to strengthen the effectiveness of teaching and management [28]. The second is the application system. To improve the construction of big data platform as the focus, colleges and universities need to add many applications one by one, such as capital and finance, campus security, resource allocation, teaching management, teacher-student interaction, teaching support, events, major decisions, and other core sections, to avoid the problem of single data platform.

To sum up, integrating big data into the field of education management can become a basic strategy in the future education system innovation. Based on the current social development, focus on shaping the concept of big data, gradually improve and optimize the campus information management platform, create a first-class professional data faculty, and accelerate the implementation of campus big data innovation management work at multiple levels [29]. For college students to create the campus order conducive to their deep learning, build a solid foundation for their long-term development, for the country to cultivate a large-scale, higher quality of college students [26]. Data table structure diagram is shown in Figure 3.

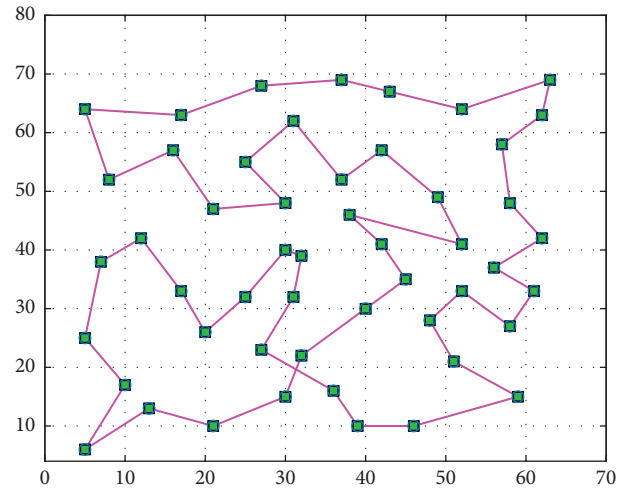


Figure 3: Data table structure diagram.

4. Analysis and Discussion of Calculation Results

At present, the idea of humanistic quality education in universities has become the theme of education and teaching. Under the guidance of this concept, to promote the quality development of college students requires that the education and management of the university should be constantly innovative, in the work to be targeted, not only to encourage outstanding on the basis of qualified, standard, but also to develop personality on the basis of standard requirements. The construction of practical guidance system from campus to real life, closely combined with social reality and the characteristics of college students' development, combined with university education and students' development ideal, to help college students set specific tasks in each stage. Second, in the process of practice, college students' development plan is established on the basis of the practice of the comprehensive evaluation system, the learning attitude and conduct performance, social practice, and so on into the appraisal scope, in integrating theory with practice, comprehensive evaluation, help students further understanding society, self-understanding, grasp myself, constantly updated knowledge and ability structure, and strive to achieve unity of knowledge and action to achieve optimal development. Quality measures under different parameters are shown in Figure 4.

Scatter plots with different parameters are shown in Figure 5. As can be seen from Figure 4, compared with the mean score of 3.49, the mean score of 3.65, and the mean score of expectation is the highest, which indicates that as follows: first, at present, college students in Y University have a strong desire to and they are eager to have the right to participate in school management. Second, college students have a high understanding of their own behavior ability and believe that they have certain management ability and can exercise the right to participate in school management. The reason for this result may be that serving as a student cadre is an important means to train college students to effectively exercise their right to participate in school management. By serving as student cadres, students can exercise their own behavior ability related to school management, so that they can exercise the right to participate in school management given by law more effectively.

Cloud image of clustering results is shown in Figure 6. It can be seen from Figure 5 that at present, college students in Y University are mainly involved in teaching management and student management, while students are less involved in logistics management and administrative management. The proportion of college students choosing not to participate is relatively high in administrative management and logistics management, especially in administrative management matters. For example, 78.7% of college students choose not to participate in financial supervision activities, and 74.9% choose not to participate in institutional setting and personnel management reform. In terms of student management, only 10.4% of students do not participate in the democratic evaluation of the election of class and community leaders, and only 8.3% of the expenses such as class expenses form a sharp contrast. It can be seen that the overall situation of Y University students is better in terms of teaching management and student management than in administrative management and logistics management. In teaching management and student management, student knowledge, to present an opinion, participation in supervision and participation in decision-making are relatively high, the proportion of the students can not only be in the process of participating in details and provide opinions and suggestions, and in the administrative and logistics management students to participate in the activities of school management is more aware, but lack of voice, supervision, and decision-making power.

The class characteristic curve is shown in Figure 7. The data analysis of this study shows that there is not a simple

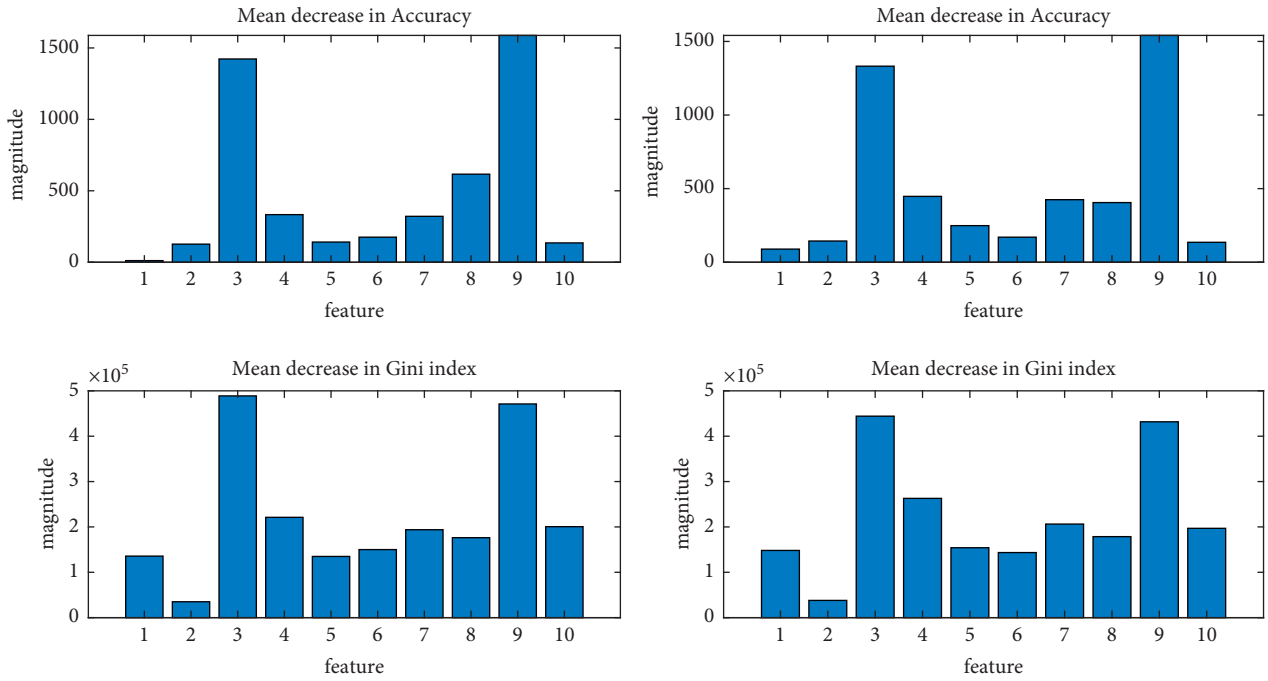


Figure 4: Quality measures under different parameters.

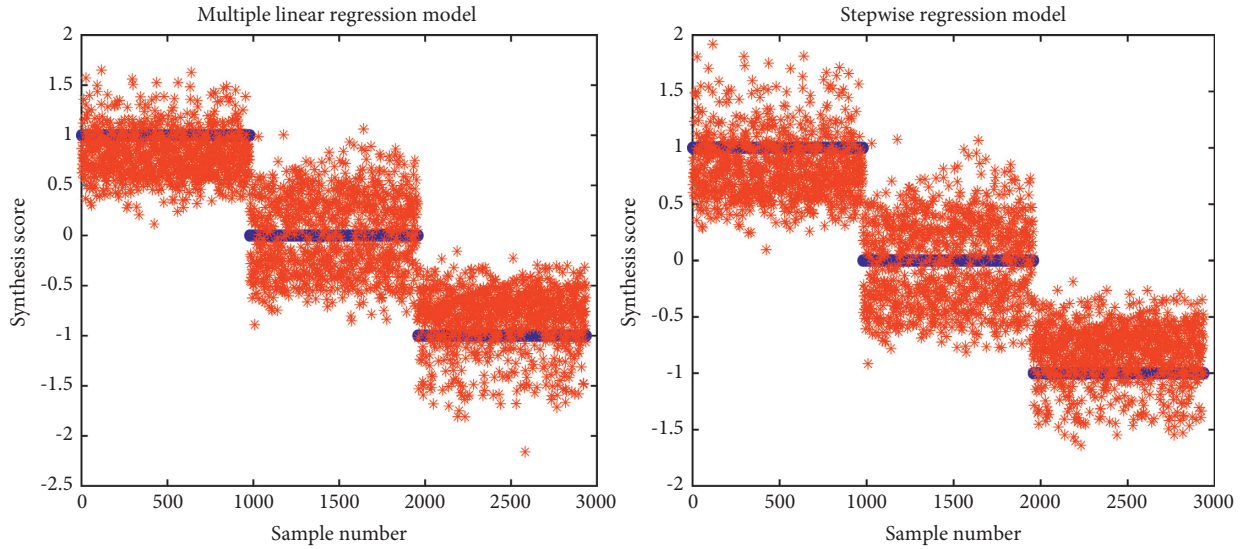


Figure 5: Scatter plots with different parameters. (a) Multiple linear regression model. (b) Stepwise regression model.

linear relationship between the cognitive formation path of data literacy individuals, the formation path mechanism of auxiliary skills, and the formation path mechanism of core skills. The influence of the three is mutual, and the formation path mechanism of auxiliary skills has the most significant effect. Therefore, the education of individual cognition, auxiliary skills, and core skills of data literacy should not be independent of each other, and the interaction of the three should be strengthened in the course setting so as to promote the formation of individual cognition, auxiliary skills, and core skills of data literacy to influence and promote each other. And the auxiliary skill formation path mechanism on individual cognitive data path mechanisms and core skills greatly influenced by the mechanism of forming path, therefore, in the curriculum should appropriately increase the proportion of secondary skills education courses and strengthen the education of the auxiliary skill curriculum and data quality personal interaction of cognitive and core skills course. In addition, different teaching practice contents have different influences on individual cognition, auxiliary skills, and core skills of data literacy, so different emphasis should be put on corresponding curriculum setting.

With the unique characteristics of big data technology, information will be continuously pushed according to the

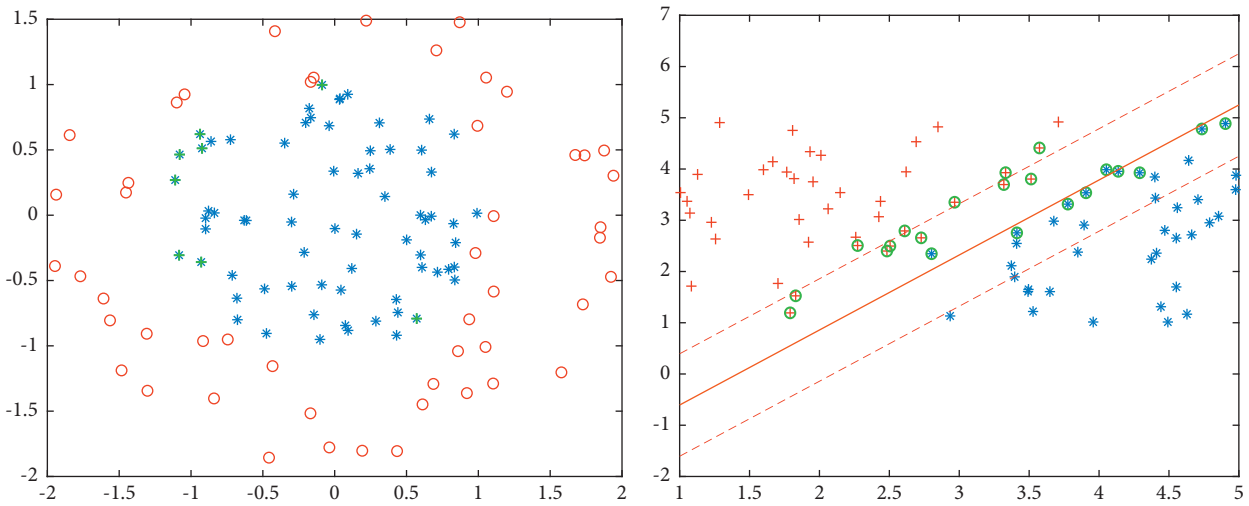


Figure 6: Cloud image of clustering results.

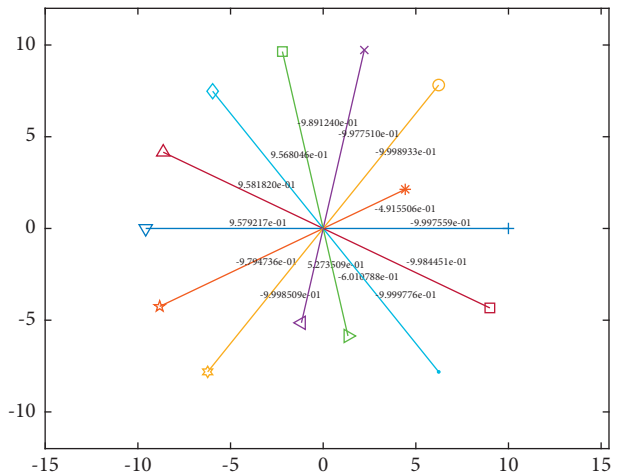


Figure 7: Class characteristic curve.

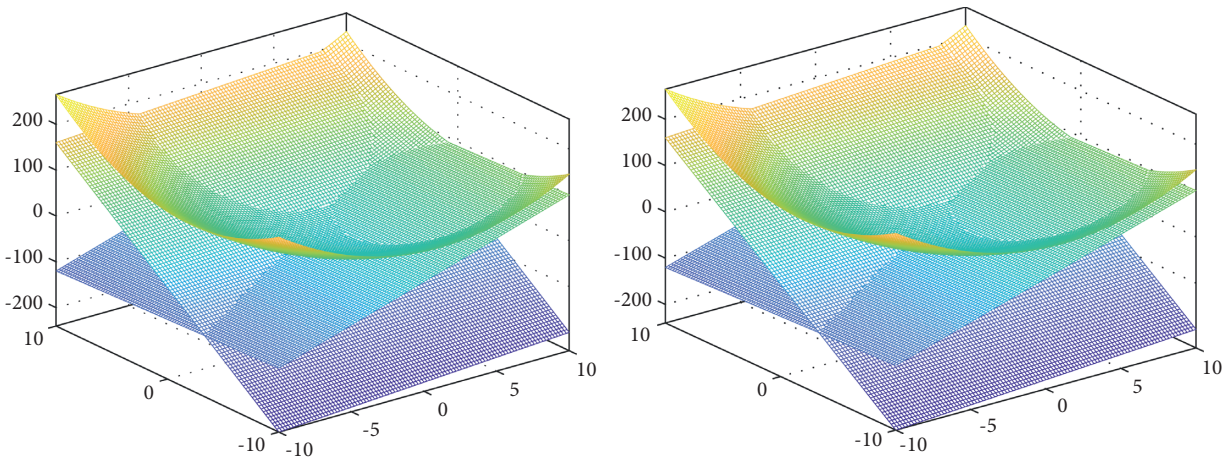


Figure 8: Distribution of eigenvalues of the observation matrix.

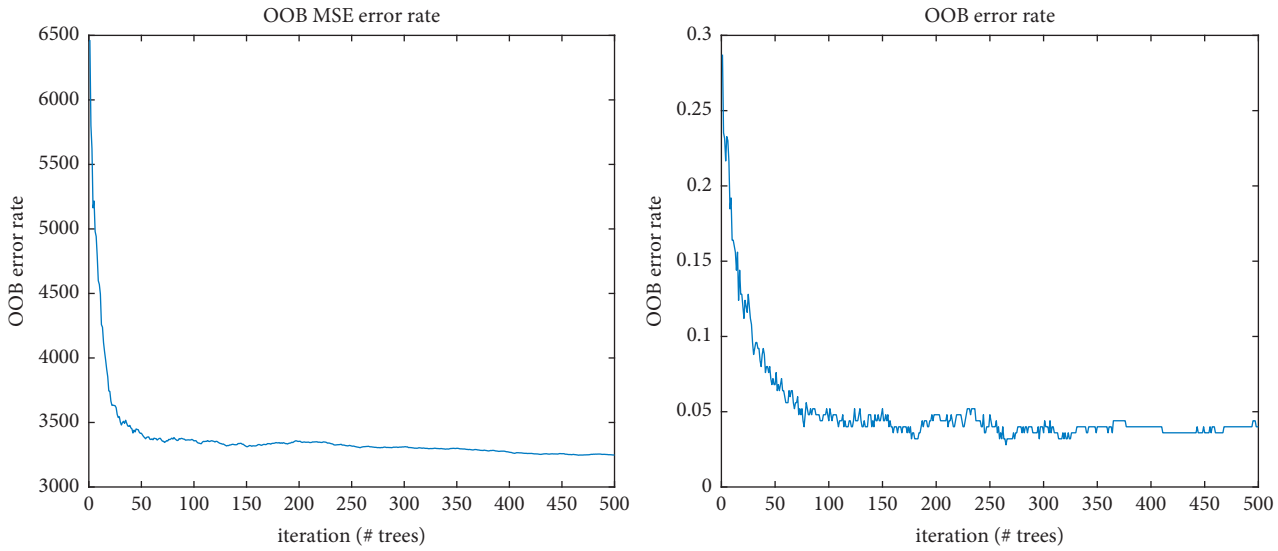


Figure 9: Information content and standard error distribution. (a) OOB MSE error rate. (b) OOB error rate.

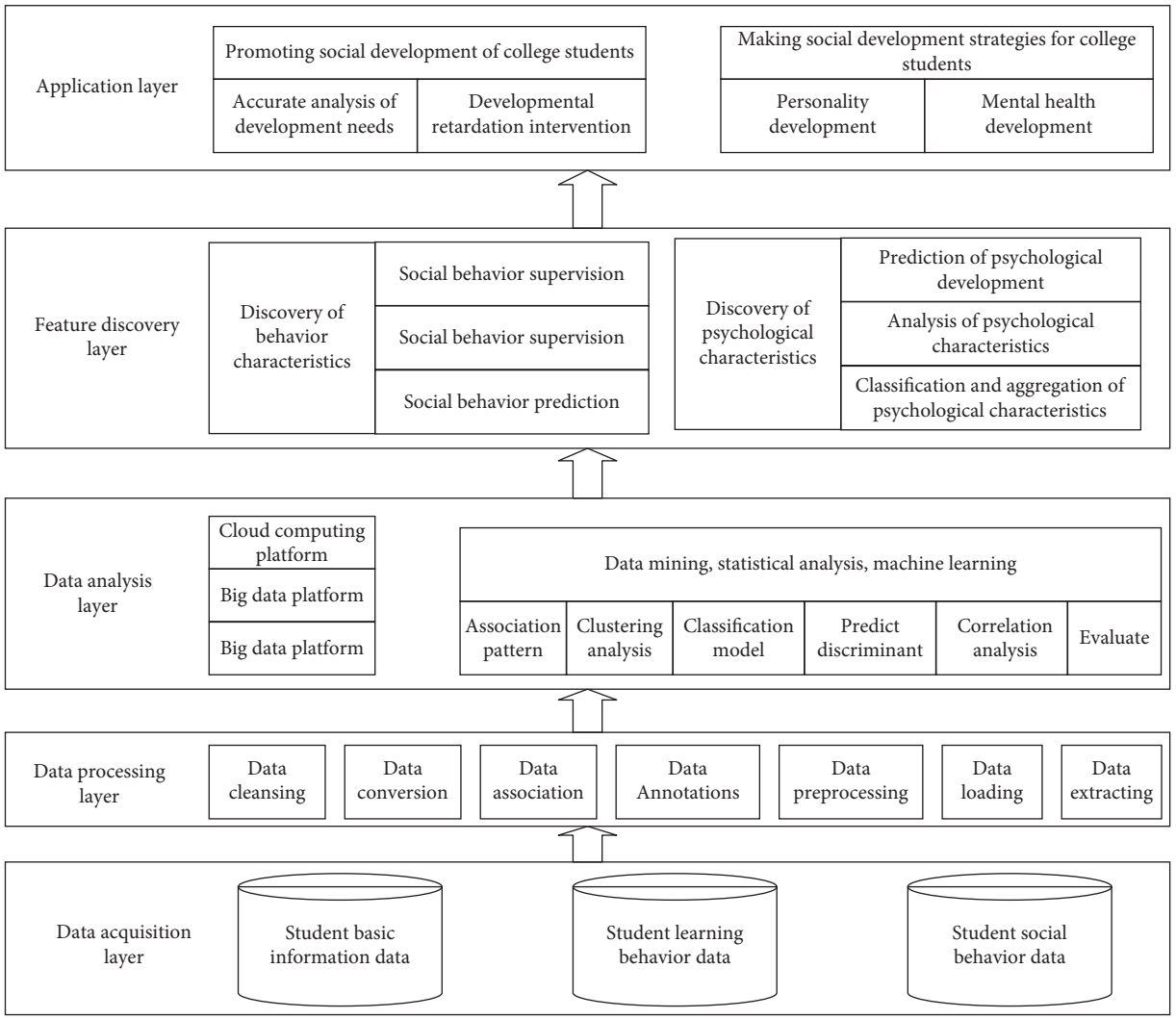


Figure 10: College student development model driven by multidimensional big data.

public's interest points. First is the promotion of "participatory" teaching methods in teaching. This teaching method is characterized by question-based teaching and open content, with no standard answers to questions and few or no assignments and papers, which can give students enough time and space to think freely. Use network technology and computer technology to collect relevant information to answer questions, through the process of answering questions to complete the learning process. As information, identify the key to assimilate, process, and produce their opinions, and transfer their ideas to others. In this process, students not only master the ability to answer various questions with the help of the Internet but also finally learn the knowledge related to "questions." For different students, basic and comprehensive research promotion and three levels of learning and training and teaching, according to the characteristics of students to establish appropriate training objectives, design and formulate strict student learning plans, as far as possible to get good development. Distribution of eigenvalues of the observation matrix is shown in Figure 8.

At present, most universities implement the credit system, which is the management mode formed in the era of planned economy, with insufficient flexibility, too strong rigidity, and too much common binding force. In today's big data environment, we advocate a more personalized model of student management. The teacher management system takes the student as the center, the student as the leading, and the teacher as the auxiliary and establishes the student service center. The specific operations are as follows: first, the establishment of psychological consultation, first aid, work research, learning guidance mechanism, and the establishment of the corresponding community management department; second, based on student dormitories, classes are cancelled and 8–15 students and teachers form a whole. Third, graduate students or senior outstanding students to help manage students, to provide guidance for students. This management mode can realize students' self-education, self-management, and self-service. Information content and standard error distribution is shown in Figure 9. The college student development model driven by multidimensional big data is shown in Figure 10.

5. Conclusion

Big data have been widely used in many industries and play an important role in promoting the development of human society in the new era. Admittedly, big data have not been widely applied to some extent in the educational practice of colleges and universities. Therefore, the workers on the front of college education must have firm confidence in reform, overcome all difficulties and obstacles, and make solid efforts to do big data homework. We must attach importance to improving resource allocation and optimizing data information technology management platform. Enhance the data awareness of teachers and students and improve the efficiency of data processing. Construct the authority organization and maintain the authority of university management work. Establish early warning mechanism, ensure the data security of teachers and students, and optimize and update the education management plan according to the feedback results of big data; ensure that the management of college students can keep pace with the development of the times, efficiency, and orderly direction. To this end, the paper proposes the following multidimensional big data-driven development and management strategies for college students:

(1) Improve resource allocation and optimize technology platform. The function of big data technology depends on the construction of Internet information technology to a great extent. Universities should clear the importance of big data information management, attach great importance to improve the allocation of resources, increase the investment of the construction of information platform, complete the necessary hardware and software facilities construction, optimizing the management information technology platform, efficient use of the popularity of the Internet, hyperlink information extension function, and mode of online education and the database platform sharing, and to ensure the standard input of personal data information of teachers and students, so as to form a systematic and perfect big data sharing platform, to ensure that college teachers, parents, and students themselves can understand the trends of students' thinking dynamics, behavioral preferences, values, and other trends through the sharing platform. Through rational use of the interactive transmission and sharing functions of multidirectional information, highlight the advantages of big data information management, form a comprehensive pattern of education, and constantly realize intelligent management.

(2) Enhance data awareness and improve processing efficiency. Information technology era, the vast amounts of data and information structure, and the information change speed have been accelerated, and with efficiency of college students management work, teachers need to strengthen the consciousness of the data; all relevant data should be brought into the scope of data collection, give prominence to the big data integration, and ensure the timeliness of data information. Make full use of big data processing technology, set up a professional big data management team, emphasize the compatibility of data between different databases, and ensure the dynamic budget of college students' ideological dynamics, behavioral preferences, and other accurateness, so as to realize the scientific warning of college students' study and behavior. Actively use the sharing of modern network technology, with the help of big data information collection and processing ability, to achieve the information collection, analysis, and prediction of all the students in colleges and universities, to summarize the common problems, and put forward effective management measures, and constantly improve the management efficiency of college students.

(3) Colleges and universities should combine educational objectives and talent training mode effectively, emphasize the integrated management of data sources, and coordinate teaching system, behavior system, and early warning system. In order to ensure the effectiveness of college student management, colleges should pay attention to strengthening matrix construction and actively build authoritative institutions. In the process of daily life and learning management of college students, strengthen the management and punishment of college students, in strict accordance with the relevant provisions of the punishment of students who violate discipline, in order to fundamentally improve the importance of college students to the school management system, so that they can consciously manage and standardize their own behavior.

(4) Establish an early warning mechanism. The student administration departments of colleges and universities should strengthen the publicity and education and ensure that all teachers, students, and administrators can correctly understand big data technology. Attach importance to the introduction of high-quality data management personnel and ensure that they can extract sensitive data from the mass of data information. Strictly regulate the input operation of teachers and students' information and increase data education, in order to ensure the accuracy of background data processing and analysis and reduce the difficulty of data processing management teachers. Formulate data management plans and programs and optimization of firewall technology, determine the level of database information disclosure, clear different identities to view the level and authority settings, and do a good job of confidentiality provisions and agreement signing, to effectively protect the data security, to avoid the occurrence of network violence and infringement.

Data Availability

The data set can be accessed upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

[1] J. Fan, "Research on individualized service of ideological and political education in universities in big data era," *Journal of Changchun University*, vol. 27, no. 6, pp. 47–50, 2017.

[2] T. A. Hart and M. Sharfman, "Assessing the concurrent validity of the revised kinder, lydenberg, and domini corporate social performance indicators," *Business & Society*, vol. 54, no. 5, pp. 575–598, 2015.

[3] G. D. A. Pederzini, "Strategic management cultures: historical connections with science," *Journal of Management History*, vol. 22, no. 2, pp. 214–235, 2016.

[4] A. Amran, H. Fauzi, and Y. Purwanto, "Social responsibility disclosure in Islamic banks: a comparative study of Indonesia and Malaysia," *Journal of Financial Reporting & Accounting*, vol. 15, no. 1, pp. 99–115, 2017.

[5] A. Lauder, R. F. Sari, and N. Suwartha, "Critical review of a global campus sustainability ranking: GreenMetric," *Journal of Cleaner Production*, vol. 108, pp. 852–863, 2015.

[6] F. U. Gómez, C. Sáez-Navarrete, S. R. Lioi, and V. I. Marzuca, "Adaptable model for assessing sustainability in higher education," *Journal of Cleaner Production*, vol. 107, pp. 475–485, 2015.

[7] J. Dlouhá, L. Henderson, and D. Kapitulčinová, "Sustainability-oriented higher education networks: characteristics and achievements in the context of the UN DESD," *Journal of Cleaner Production*, vol. 172, pp. 4263–4276, 2018.

[8] L. Liu, Y. S. Wang, and T. J. Wu, "Student satisfaction scale development and application for sport management in China," *Eurasia Journal of Mathematics, Science and Technology Education*, vol. 13, no. 5, pp. 1429–1444, 2017.

[9] I. N. Emelyanova, O. A. Teplyakova, and O. A. Teplyakova, "Mobility of Russian university students as a phenomenon and a management problem," *University Management: Practice and Analysis*, vol. 24, no. 2, pp. 131–144, 2020.

[10] M. Turner, C. Scott-Young, and S. Holdsworth, "Developing the resilient project professional: examining the student experience," *International Journal of Managing Projects in Business*, vol. 12, no. 3, pp. 716–729, 2019.

[11] K. H. Yang, "A study on the development of university student's learning competency scales," *The Journal of Lifelong Education and HRD*, vol. 12, no. 1, pp. 29–64, 2016.

[12] M. Jakubiak, M. Cholewa-Wiktor, and A. Sitko-Lutek, "Sustainable development and pro-ecological education in the opinion of students of management," in *Proceedings of the 11th International Conference on Education and New Learning Technologies (EDULEARN)*, pp. 700–705, Palma de Mallorca, (Spain), July 2019.

[13] F. Y. Wei, "Research on the students' rights guarantee in the management of colleges and universities," in *Proceedings of the International Conference on Innovations in Economic Management and Social Science (IEMSS)*, vol. 29, pp. 1582–1587, Hangzhou, China, April 2017.

[14] M. Chankseliani, "The politics of student mobility: links between outbound student flows and the democratic development of post-Soviet Eurasia," *International Journal of Educational Development*, vol. 62, pp. 281–288, 2018.

[15] R. Brooks, "Higher education mobilities: a cross-national European comparison," *Geoforum*, vol. 93, pp. 87–96, 2018.

[16] M. N. D. Tran, K. Moore, and M. C. Shone, "Interactive mobilities: conceptualising VFR tourism of international students," *Journal of Hospitality and Tourism Management*, vol. 35, pp. 85–91, 2018.

[17] A. Cuzzocrea and E. Damiani, "Making the pedigree to your big data repository: innovative methods, solutions, and algorithms for supporting big data privacy in distributed settings via data-driven paradigms," in *Proceedings of the 43rd IEEE-Computer-Society Annual International Computers, Software and Applications Conference (COMPSAC)*, vol. 2, pp. 508–516, Milwaukee, WI, USA, July 2019.

[18] P. Senellart, "Provenance and probabilities in relational databases: from theory to practice," *Sigmod Record*, vol. 46, no. 4, pp. 5–15, 2017.

[19] S. Kumar, S. Madria, and M. Linderman, "M-Grid: a distributed framework for multidimensional indexing and

querying of location based data," *Distributed and Parallel Databases*, vol. 35, no. 1, pp. 55–81, 2017.

[20] X. Yan, X. Feng, C. Song, and X. Hu, "Bidirectional feedback dynamic particle filter with big data for the particle degeneracy problem," *Tsinghua Science and Technology*, vol. 23, no. 4, pp. 463–478, 2018.

[21] L. Cai, X. W. Wang, J. N. Cao, K. Li, and C. Hui, "Adaptive caching strategy based on big data learning in ICN," *Journal of Internet Technology*, vol. 19, no. 6, pp. 1677–1689, 2018.

[22] S. Akter, M. A. Hossain, Q. Lu, and S. M. R. Shams, "Big data-driven strategic orientation in international marketing," *International Marketing Review*, vol. 38, no. 5, pp. 927–947, 2021.

[23] J. R. Huie, C. A. Almeida, and A. R. Ferguson, "Neurotrauma as a big-data problem," *Current Opinion in Neurology*, vol. 31, no. 6, pp. 702–708, 2018.

[24] J. Ding, V. Nathan, M. Alizadeh, and T. Kraska, "Tsunami: a learned multi-dimensional index for correlated data and skewed workloads," *Proceedings of the Vldb Endowment*, vol. 14, no. 2, pp. 74–86, 2020.

[25] J. Z. Qi, G. L. Liu, L. Kulik, and C. S. Jensen, "Effectively learning spatial indices," *Proceedings of the Vldb Endowment*, vol. 13, no. 11, pp. 2341–2354, 2020.

[26] P. Ferragina and G. Vinciguerra, "The PGM-index: a full-y-dynamic compressed learned index with provable worst-case bounds," *Proceedings of the Vldb Endowment*, vol. 13, no. 8, pp. 1162–1175, 2020.

[27] T. Kraska, A. Beutel, N. Polyzotis, H. Chi, and J. Dean, "The case for learned index structures," in *Proceedings of the 44th ACM SIGMOD International Conference on Management of Data*, pp. 489–504, Houston TX USA, May 2018.

[28] L. Chen, Y. Gao, and X. Li, "Efficient metric indexing for similarity search and similarity joins," *IEEE Transactions on Knowledge and Data Engineering*, vol. 29, no. 3, pp. 556–571, 2017.

[29] J. Yu and M. Sarwat, "Two birds, one stone: a fast, yet lightweight, indexing scheme for modern database systems," *Proceedings of the Vldb Endowment*, vol. 10, no. 4, pp. 385–396, 2016.

Hindawi
Scientific Programming
Volume 2023, Article ID 9829285, 1 page
https://doi.org/10.1155/2023/9829285

Retraction

Retracted: Classroom Visualization Based on the Relationship between User Roles and Scene Content

Scientific Programming

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Y. Li, D. Yu, K. Yu, and X. Chen, "Classroom Visualization Based on the Relationship between User Roles and Scene Content," *Scientific Programming*, vol. 2022, Article ID 5300840, 10 pages, 2022.

Hindawi
Scientific Programming
Volume 2022, Article ID 5300840, 10 pages
https://doi.org/10.1155/2022/5300840

Research Article

Classroom Visualization Based on the Relationship between User Roles and Scene Content

Yingzhi Li,[1] Di Yu,[2] Kan Yu,[1] and Xi Chen[3]

[1]*Wenhua College, Wuhan 430074, Hubei, China*
[2]*Hubei University of Economics, Wuhan 430205, Hubei, China*
[3]*Shanghai Ocean University, Shanghai 201306, China*

Correspondence should be addressed to Di Yu; tranqull@hbue.edu.cn

Received 15 November 2021; Accepted 17 January 2022; Published 7 February 2022

Academic Editor: Baiyuan Ding

With the increasing complexity of the internal structure of university classrooms and the increasing amount of time humans spend indoors, indoor classroom scenes have become an important part of people's daily lives. Compared with open outdoor space, indoor environments are more complex in terms of 3D models, spatial layouts, feature types, and connectivity relationships, especially for large indoor buildings that often contain a large amount of information. This paper introduces the development tools and common rendering techniques of Unity3D game engine, explores the visualization model of a single object, designs a regional comparison algorithm to calculate the visualization intensity, and establishes a 3D visualization intensity mapping table. The classroom is equipped with visualization effect, which makes students get comfortable and at ease in classroom learning, and the comfort level of this paper's solution is improved by 8% compared with other IoT-based solutions.

1. Introduction

With the increasing complexity of the internal structure of university classrooms and the increasing amount of time spent indoors, indoor maps have become an important part of people's daily lives. Compared with the outdoor open space, the indoor environment of classrooms is more complex in terms of three-dimensional models, spatial layout, feature types, and connectivity, especially for large indoor buildings, which usually contain a lot of information [1, 2]. Classroom interior space has three outstanding features: (1) the internal pattern is complex. The interior space of university classrooms is clearly divided, usually with multiple floors, and the internal space is intricately connected and divided. (2) The view is highly obscured, with a large number of vertical and horizontal walls and indoor artificial facilities not only blocking people's view, but also hindering people's understanding of the overall spatial structure of the room and limiting their access to local information. (3) The access conditions are special [3–5]. There is no obvious road in the indoor space. However, access through open space enables people to locate themselves quickly and accurately.

Due to the complexity of classroom indoor environments, unadorned visualization of stacked features will not provide users with valuable information. Therefore, indoor map visualization that can be adapted to different application contexts, different user purposes, and different scene patterns has application value [6]. Such user-centered maps, also called personalized maps and adaptive maps, are important applications of indoor maps and location services. By establishing the mapping relationship between users and maps, the map expression content and representation method are dynamically changed in real time, and the personalized service method of the map is changed from the operation interaction between users and maps to autonomous push [7, 8]. Therefore, it is a very meaningful research direction to make full use of the visual advantages of 3D maps, to integrate user cognitive features and specific needs with map display interaction, to dynamically and effectively visualize the indoor scenes of classrooms, and to realize user personalized 3D visualization of indoor scenes.

Traditional map representation is the design process of map content, symbol system, representation method, preparation principle, and drawing mechanism under the premise of clear map usage, expression content, output form, etc. The content and manner of map representation are completed before cartography, and the map effect is closely related to the level of designers [9, 10]. However, in the adaptive visualization map, except for describing the basic information of indoor space such as the layout of fixed facilities and the structure of university classrooms, all other information is triggered and drawn in real time; the process of extracting the representation content, selecting the representation method, and symbolizing and drawing the map is done dynamically; and the effect of using the map in the user's presence in the mobile state should be taken into account. Some studies show that in indoor map services, map design, especially the design of dynamic map representation, is the most important factor affecting the effectiveness of indoor map services and is an urgent problem for indoor map services [11].

In view of the above urgent needs and existing conditions, the research in this paper revolves around user roles and scene contents, constructs the association relationship between them, and mines valuable information about the current environment based on ontological semantic reasoning. Secondly, we define the visualization model of classroom indoor scenes, combine user and environment information, and dynamically present 3D maps to realize user-centered visual representation of classroom indoor scenes. This indoor visualization method that associates user roles with scene contents is important for providing personalized and intelligent services in classroom indoor environments.

2. Related Work

This paper proposes the indoor visualization method of user role and scene content association, which establishes the mapping relationship between user and indoor map from both user role and scene content, so that the map expression content and representation method change dynamically in real time, and therefore is an adaptive map visualization, i.e., the theory and method of organically organizing spatial data and establishing user-centered indoor adaptive map [12]. At the same time, the method proposed in this paper is applied to the visualization of the indoor environment, and it is necessary to explore the indoor map representation related to it. Therefore, this paper will explore the current status of domestic and international research on both adaptive map visualization and indoor map representation.

2.1. Adaptive Map Visualization. The proposal of adaptive map visualization research can be traced back to as early as 1998. Subsequently, a large number of studies have extended and expanded the framework, and some highly valuable research results have emerged. For example, [13] provides a more comprehensive discussion on the user model of adaptive maps, adaptive spatial database design, and adaptive symbol design and establishes an adaptive strategy framework. However, in the field of 3D visualization, the research scope is mainly on urban scale scenes. Reference [14] studied a multi-detail level management method based on 3D R-tree indexing and proved through experimental analysis that its defined parameters can quantitatively adjust the scene complexity and thus adaptively control the detail level of 3D model visualization. For example, [15] proposes a novel application of focus + context scaling technique that can be used to zoom in on navigation paths and their related features in 3D urban environments as a way to effectively reveal the focus area; [16] shows a technique to automatically generate 3D virtual city models and dynamically zoom in on landmark objects. In addition, there are many related studies at home and abroad, but in terms of their research contents and progress, most of them focus on the adaptive strategies of electronic maps or the efficiency of big data display of 3D scenes, and there are fewer studies on 3D visualization for indoor environments [17].

User model, as one of the core elements of adaptive map visualization, is a collection of rules for users in terms of interface, data amount, map representation, etc., generated based on user background information, historical behavior records, and initial adaptive rules, and is the core of the system constituting adaptive map visualization. In the study of user models for adaptive maps, [18] earlier proposed user modeling applicable to adaptive maps and applied it initially to urban tourism. A large amount of literature was then devoted to the study of user models, covering the classification of user models, representation methods, example applications, etc. Reference [19] designed a preliminary mechanism for adaptive user interfaces for map visualization systems and explored ways to make the system's user interface automatically adapt to user features. Reference [20] proposed a map representation and visualization model, user model construction, and matching algorithm in adaptive map visualization system. Reference [21] established an adaptive representation model of navigation electronics maps content by analyzing the user's background information, behavioral habits, and other information and used the plain Bayesian algorithm for model matching. Reference [22] studied the characteristics of context-aware mobile device based on adaptive user interface and established a context-driven mobile device based adaptive user interface model.

2.2. Indoor Map Representation. In the area of indoor map representation, many research results have also been achieved in recent years. Reference [23] proposed a method for displaying and navigating indoor location maps by using a mobile terminal with a camera to superimpose path information on a paper map. Reference [24] proposed an indoor access representation method that can represent the topological relationship of cross-floor access from the perspective of underground pipeline maps. In [25], the form of map expression suitable for spatial cognitive results was studied, and three visualization techniques for road network

framing diagrams, oriented to road network cognitive expression and suitable for mental image map expressive features, were proposed. The research team of [26] conducted special theoretical and technical methods from indoor maps, location maps, and mobile maps and achieved various results; [27] focused on the indoor map representation methods under single-floor, multi-floor, and indoor-outdoor switching conditions; and [28] studied the dynamic mapping process and mechanism of mobile maps based on the cognitive semantic theory and established the dynamic mapping of mobile maps model.

Indoor spatial scene modeling, as a research part of indoor scenes, is important for indoor map representation. Indoor GML spatial standard describes the basic semantic information of indoor space, lacking complex semantic relationships such as opposite and upstairs, without user-centered information expression [29]. The above model standards focus on the construction, storage, and display of 3D models, ignoring the role of the user in the 3D scene, who is more like a "visitor" than a "participant" who can change the scene. On this basis, [30] proposes indoor ontology modeling for holographic location maps, which defines the semantic concepts, attributes, and relationships of indoor space based on the association of people, things, and objects, and improves the indoor space ontology modeling method. In this paper, we hope to use the existing results of the above model to propose the association model of user model and indoor scene in indoor environment.

3. Introduction to Related Concepts and Theories

3.1. User Roles and Scenario Content. User personas are a way of classifying users to build a user model, which divides user groups in the form of personas. User model refers to the description, classification, and identification of user background information, user behavior information, user scenario information, and user cognitive rules. User background information describes the basic characteristics of users in terms of education, culture, society, economy, and nature; user behavior information is the record of all operations of users on data, interfaces, and functions, which is the basis of user model analysis; user scene information includes user purpose, current location, speed, time, brightness, and direction [15]; user cognitive rules include those influenced by users' visual saturation, color contrast, symbolic cognitive ability, and information acquisition ability, as well as the guidelines that need to be followed when designing the map. The role concept is the role played by a user in the context. Roles can reflect both the user's interests (e.g., the user's professional characteristics imply his or her range of interests) and the user's behavior (e.g., tourists and other role types with behavioral purpose). At the same time, roles are temporal in nature; a user may have multiple roles at the same time to describe multiple aspects of his or her characteristics, and a user can have different roles at different times.

3.2. Theoretical Approach to Ontology Modeling. An ontology is a shared conceptualization of knowledge in a particular domain. Originally ontology is a philosophical concept, an abstract and systematic explanation of the nature of objective existence. In recent years, information abstraction and knowledge description tools have been widely used in the field of computing. Many people have given different levels of understanding about ontology in their research history, among which Gruber's definition has received much recognition, which considers ontology as an explicit normative description of a conceptual model. Reference [3] also elaborates on the definition of ontology, which is a formal normative description of a shared conceptual model.

At present, the more mature, commonly used, and well-known ontology construction tools mainly include Apollo [14], OilEd [16], OntoEdit [18], Prot6g6 [19], and WebODE [19]. Prot6g6, ontology editing and knowledge acquisition software, is open source software developed by Stanford University School of Medicine Bioinformatics Research Center based on Java language, mainly used for the construction of ontologies in the semantic web, and is a common development tool for ontology modeling. It provides methods for the construction of ontology concepts, relationships, attributes, and instances and hides the specific ontology description language. The ontology structure of Prot6g6 is represented by a tree hierarchy, and users can add or edit classes, subclasses, instances, etc. by clicking the corresponding items. Therefore, users do not need to master the specific ontology representation language to use it, which makes it a relatively easy to learn and master ontology development tool. In this paper, we use Prot6g6 software and OWL modeling language for ontology modeling of user characteristics and indoor scenes according to specific research needs.

4. A 3D Visualization Method Based on Associative Reasoning

4.1. Rendering Techniques for 3D Scenes. This study explores the 3D visual display model of a university classroom and implements 3D rendering on the Unity3D game engine, so this section will briefly introduce the Unity3D game engine.

Unity3D game engine is 3D development software that supports cross-platform and has a high A rate in the current 3D engine market. It has a complete graphics rendering subsystem, network subsystem, physics subsystem, audio and video subsystem, editorial system, shading system, and GUI system. Unity3D has good support for mainstream dimensional modeling tools, an excellent design environment, perishable design flow, and fast and easy to operate scene editor. The final program can not only be embedded in the browser environment to run directly, but also support multiplatform applications; these advantages make it the engine of choice for 3D games and virtual simulation projects. In 3D games, virtual reality, Web3D, and other fields have a wide range of applications.

(1) Unity3D components: the properties of each object in the Unity3D scene are composed of individual

pieces, which are the behavior and care of the objects in the scene, and they are the functional modules of the objects. There are many kinds of Unity3D components; the common ones are scripting class, particle class, physics class, sound class, rendering class, etc. [20].

(2) Unity3D scripting: Unity3D scripting is essentially a custom functional component. It is a code segment that implements specific functions by calling functional components of Unity3D or prepackaged real-time runtime classes, according to the data input and output situation, business process, result display, and other requirements [8].

(3) Development tools: MonoDevelop is a cross-platform open source integrated development environment, integrating many Eclipse and Microsoft Visual Studio features and currently supporting languages such as C#, Java, Boo, and Visual Basic. The default is to use MonoDevelop to program and implement Unity3D scripting programs.

Colorful rendering effects can be programmed and controlled by the program. A shade is a program that can be manipulated by objects and executed by the GPU. Programmers can use these programs to obtain most of the desired dimensional graphics effects. Shaders are divided into Vertex Shader and Pixel Shader. Shaders replace the traditional fixed rendering pipeline and make it possible to compute dimensional graphics. Due to the editable nature of shades, a wide variety of image effects can be achieved without being limited by the fixed rendering pipeline of the graphics card, thus greatly improving image quality [21].

This paper presents examples of two rendering techniques applied in this study, namely, glow effects and epi-glow effects. Glow and halos of light sources are phenomena found everywhere in nature, and they provide strong visual information about brightness and atmosphere. When viewing computer graphics, film, and print, the intensity of light reaching the eye is limited, so the only way to identify the intensity of light sources is through their surrounding glow and halo using modern graphics hardware; you can reproduce this effect through some simple rendering operations. When we trivialize a scene rendered in real time with bright and interesting objects, the objects will look more realistic or fantastic. To achieve the glow effect, you must first isolate and separate the growing parts of the scene or model from the non-glowing parts by some method. The scene is rendered as normal without the glow, and a texture map is created using the glow source information, which is black everywhere except for the glow source. This rendered texture map can be used as a normal texture in the later rendering. For the simple ensemble, a two-step image convolution operation is performed with multiple samples at each pixel, so that the points of the glow source are painted outward as a blurred, large glow mass, and finally the blurred glow is applied to the top of the normal rendering using additional alpha blending. In this way, using hardware rendering and texture mapping, the flow source is expanded into a convincing glow atmosphere. The glow effect before and after rendering is shown in Figures 1 and 2.

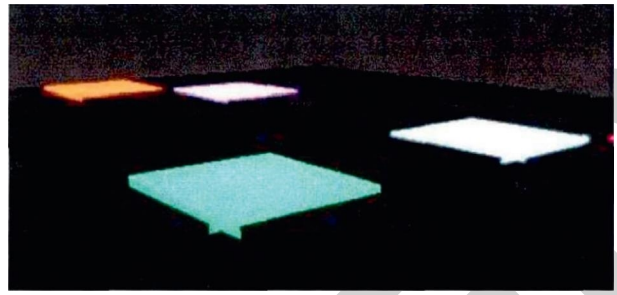


Figure 1: Glow effect (not rendered).

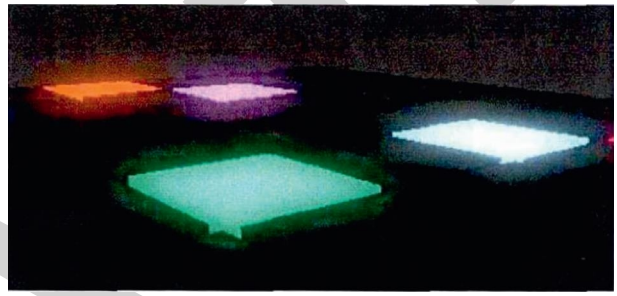


Figure 2: Glow effect (rendered).

The outer glow effect is like an extra layer on the outside of the object; the imaginary layer is filled with a slightly larger range than the above and is a blend of screen transparency and object mode, resulting in a "glow" effect on the outer edge of the layer. This effect can be used to show the outline of the object as well as to blur the object and even vary the glow in a variety of ways. To see the outer glow effect in the background, you must first detect the edge of the object, and the angle between surfaces can be determined by using incident light and perpendicular to the outer surface of the object. The outer glow layer is like an extra layer on the F-side of the layer, and the screen transparency will affect the blending effect of the display of the object. In addition to expanding, the outer light has a gradient speed between the area of color and completely transparent area; the color and intensity of the outer and inner light can be adjusted to achieve a more desirable effect. A variety of external light effects are displayed as showing in Figure 3.

5. Three-Dimensional Visualization Performance Intensity Mapping Table

5.1. Single-Object Visual Representation Method. The visualization information of indoor scenes is colorful and difficult to be accurately classified and summarized; however, complex indoor scenes are composed of independent objects that can be divided into different types of objects, so it is important to explore the visualization information of a single object sufficient to simplify complex scenes. In this paper, based on the visual perception of human eyes, the visualization information of a single object is divided into

Figure 3: Display of various external light-emitting effects.

four subtypes of information: light information, color information, motion information, and text information, and several basic visualization methods are listed for each subtype; see Table 1.

(1) Light information refers to the light of the object in the three-dimensional display. This mode determines the vast majority of the object display form. According to the sensitivity of the human comedian to light, this paper divides the light mode into five visualization expressions: hidden, transparent, normal, glowing, and twinkling. Figure 4 illustrates the performance of the above five visualization methods. On the basis of this, colorful color rendering and visual effects can be added to make the 3D visualization of objects more diverse and hierarchical.

(2) Color information refers to the color of the object in the three-dimensional display. Distinguish objects according to the ability of the human eye to distinguish colors. This paper takes red and green as a typical representative example, and encodes the color according to the same method.

(3) Motion mode refers to the form of motion of the object in the three-dimensional display. The human eye is more sensitive to moving objects than stationary objects, so moving objects are more likely to attract the user's attention. Jumping, rotation, and other motion modes added to the display mode of the object can significantly enhance the visual effect of the object.

(4) Text information refers to whether the three-dimensional display of the object appears when people are immersed in the world of virtual roaming; text display can attract attention, indicating to the user that this is an important scene object. In the above coding example, 90 kinds of object visualization information are obtained by free combination.

5.2. Indoor Scene Performance Intensity Mapping Table. This paper provides an area comparison method to evaluate the visual effect of 3D visualization performance. As shown in Figure 5, the four directional axes of the coordinate system represent the four subtypes of information, forming a visual area, representing the intensity of object visualization performance, and the area is calculated and compared to evaluate and grade the visualization effect of 90 objects. The specific formula is as follows:

$$\text{reascore} = (L + C)(M + T), \tag{1}$$

where L, C, M, and T represent the light information score, color information score, motion information score, and text information score.

For example, the area score of object visualization mode 1 [L1C1M1T1] is 4.4, which consists of 2 points for hidden light information, 3.3 points for normal color information, 3.3 points for normal motion information, and 5 points for no text prompt information. The area score of object visualization mode 2 [L5C3M1T1] is 400, which consists of 10 points for twinkling light information, 10 points for green information, 10 points for rotational motion information, and 10 points for text prompt information. The area of object visualization mode 2 is significantly larger than that of object visualization mode 1, thus leaving more impressions on users. Therefore, object visualization mode is 2 [L5C3M1T1] was judged to have a higher score [5]. The area comparison method is a relatively simple scoring calculation, and in order to judge the visualization effect more accurately, it is necessary to make some manual selection after the preprocessing of this method to screen out some visualization performance modes that are obviously out of the ordinary. The final result is shown in Table 2, which is called the 3D visualization intensity mapping table, and the area score is the score of the area comparison method. 3D visualization strength mapping table will now give examples of visualization with reasonable visual effects, graded from weak to strong according to the strength of the effect, to correspond to the list of important levels of reasoning results in the previous section. The specific use will be explained in detail in the next section on the visual representation algorithm for associative reasoning.

6. Visual Representation Algorithm for Associative Reasoning

6.1. Visual Representation Algorithm Flow. This section proposes a visual representation algorithm based on associative reasoning, which maps the reasoning results to the representation intensity mapping table with priority and sets the visualization mode of each category of objects within the scene to perform the hierarchical rendering of the whole scene to achieve an effective scene visualization effect. The specific flow of the algorithm is shown in Figure 6.

(1) Rule reasoning. Using the expression rule base and rule priorities in Section 3, we filter out the indoor

Table 1: Single object visualization information.

Information	Information subtype	Visualization mode	Code	Score
Object visualization information	Optical information	Hidden	1.1	2
		Transparent	1.2	4
		Normal	1.3	6
		Glowing	1.4	87
		Twinkling	1.5	10
	Color information	Normal	C1	3.3
		Gules	C2	6.7
		Green	C3	10
	Motion information	Normal	M1	3.3
		Rotation	M2	6.7
		Jumping	M3	10
	Text message	No text	T1	5
		Having text	T2	10

space or things associated with the user role by query function and get the result priority list sorted from

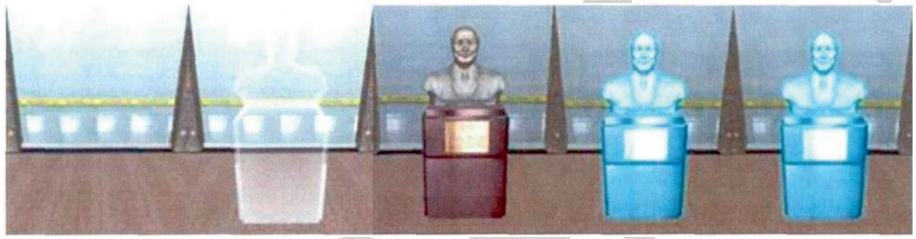


Figure 4: Optical information visualization.

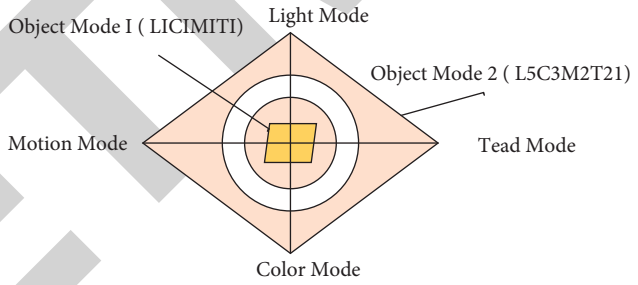


Figure 5: Example of a regional scoring map.

Table 2: 3D visualization performance intensity mapping table.

Object visualization	Area score	Normalization	Optical information	Color information	Motion information	Text message
L1	Default	0	Hidden	—	—	—
L2CIM1T1	61	0.15	Transparent	Normal	Normal	No text
L3CIM1T1	77	0.19	Normal	Normal	Normal	No text
L3CIM1T2	124	0.31	Twinkling	Gules	Normal	Having text
L4C2M1T1	131	0.33	Twinkling	Gules	Normal	No text
L3C2M1T1	139	0.35	Glowing	Green	Normal	No text
L4C3M1T1	149	0.37	Twinkling	Green	Rotation	No text
L5C3M1T1	170	0.43	Glowing	Gules	Rotation	No text

strong to weak, which provides a preliminary inference result table for the visual mapping in the later steps.

(2) Focus calculation. We design role attributes to describe the focus of the role instance on the scene, added in the role instantiation. The focus of character instances with focus value greater than 1 is more bifurcated, and the things that need to be highlighted are more prominent, in order to strengthen the visual central visual. This step further provides a perfect inference result table for the visualization mapping later.

(3) Visual representation mapping table on the list of results of the mapping. According to the visual representation of a single object in the visual representation mapping table, the visual information pattern of each object is determined in order from strong to weak according to the existing inference result table. In the process of rendering indoor scenes, the weight in the rule inference process is adjusted, and the reasonable hierarchical visualization rendering will make the scene present a more distinctive hierarchical effect.

(4) Visualization performance updates. When the user information and scene information change, the rule inference and mapping process are reworked to dynamically adjust the visualization of objects in the scene. For example, as the user's position keeps changing, the objects in the scene change, and the accessibility of the user to the objects also changes, so when the user's position is detected to have changed, the indoor scene is rerendered to achieve the visualization effect update in the user's view.

6.2. Visual Representation Algorithm Application Examples. This section shows an example of a feasible visual rendering strategy for a typical indoor scene. A simplified indoor scene containing two adjacent regular rooms and a walkway with walls, doors, windows, and a small number of concrete objects in each space is used, as shown in Table 3. This case study completes the inference process, visualization mapping, and indoor scene rendering with the user role visitor at the center to illustrate the specific application of the visual representation algorithm.

Following the flow of the visual representation algorithm, the preliminary list of results is obtained using the correspondence rules of behavioral targets, the association rules of target things, and the association rules of topological relations in turn. Then, the association rules of semantic location are used and the weights are recalculated. Finally, the weights are calculated again using focus = 1.5 for visitor role instances, and the results of the weight calculation are shown in Table 4.

7. Simulation Effect

Unlike the traditional way of managing resources in the classification system, LCS uses semantic web-based ontology technology to organize all kinds of learning resources in the platform. The platform allows users to collaboratively edit learning content and use group intelligence to promote the growth of learning resources. Since the platform allows any user to edit the learning content, in order to ensure that the resources can fully absorb the group wisdom, ensure that the absorbed content is meaningful to the growth of the resources, and avoid the growth of resources in a disorganized manner, the platform adopts a complete content evolution intelligent control technology combined with human review technology to realize the orderly control of the content evolution of the learning resources. Learners can view the visualized resource evolution path (as shown in Figure 7) to understand the evolution of resources as a whole and can

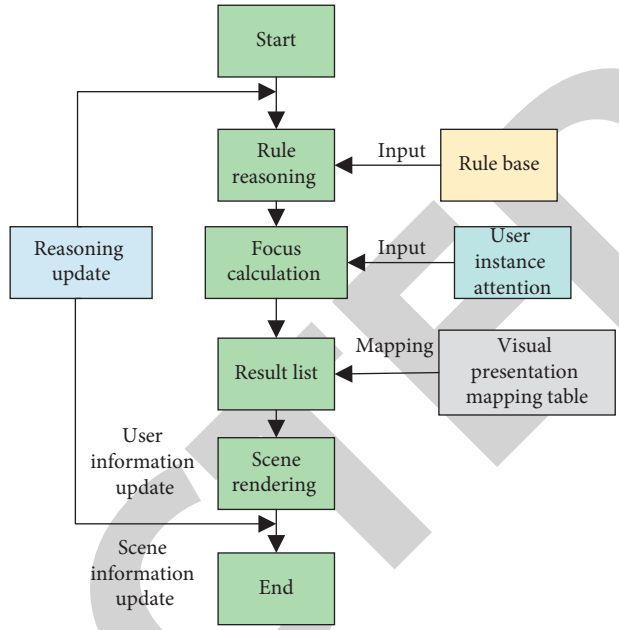


Figure 6: Flow chart of visual representation algorithm.

Table 3: Simplified indoor scene examples.

Interior space	Space type	Object	Number
Room 201	Room	Wall	Wall 201
		Window	Window 201
	General space	Door	Door 201
		Touring statue	Touring statue 001
		Sofa	Sofa 201
Room 203	Room	Wall	Wall 203
		Window	Window 203
		Door	Door 203
		Exhibition board	Exhibition board 001
		Desk	Desk 203
Corridor 001	Corridor	Wall	Wall 001
		Window	Window 001
	Passages	Exhibition board	Exhibition board 002
	Transfer space		

Table 4: List of weight calculation results.

Rule reasoning	Result item	Weight	Semantic location weight	Focus calculated weight	Weight reduction
Professional interest rules of social roles	Touring statue 001	5	5	7.5	1.00
Target object	Exhibition board 001	4	1.6	2.4	0.32
Association rules	Exhibition board 002	4	2	3	0.40
Topological relation	Wall 201	0.5	0.5	0.33	0.04
	Window 201	0.5	0.5	0.33	0.04
	Door 201	5	5	7.5	1.00
	Wall 203	0.5	0.2	0.13	0.02
Association rules	Window 203	0.5	0.2	0.13	0.02
	Door 203	5	2	3	0.40
	Wall 201	0.5	0.25	0.17	0.02
	Window 001	0.5	0.25	0.17	0.02
Semantic position	Sofa 201	1	1	1.5	0.20
Association rules	Desk 203	1	0.4	0.27	0.04

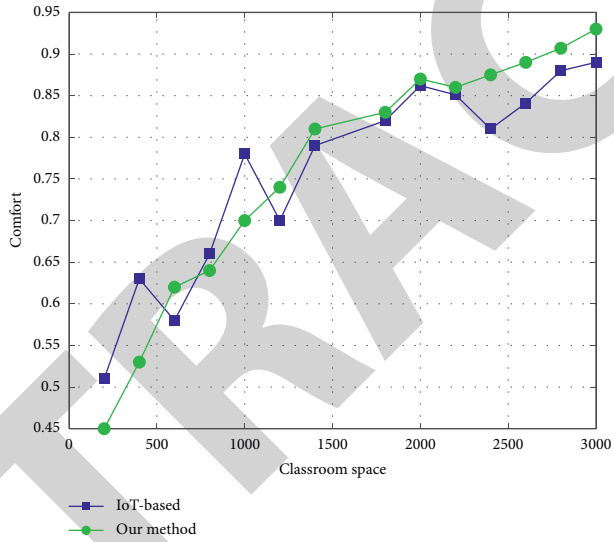


Figure 7: Classroom space comfort.

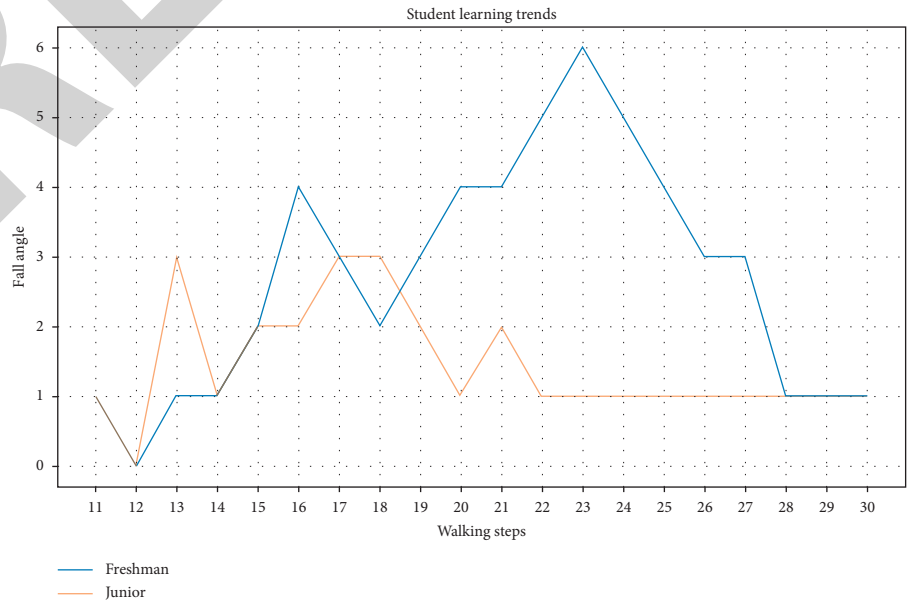


Figure 8: Comfort level of students in different grades.

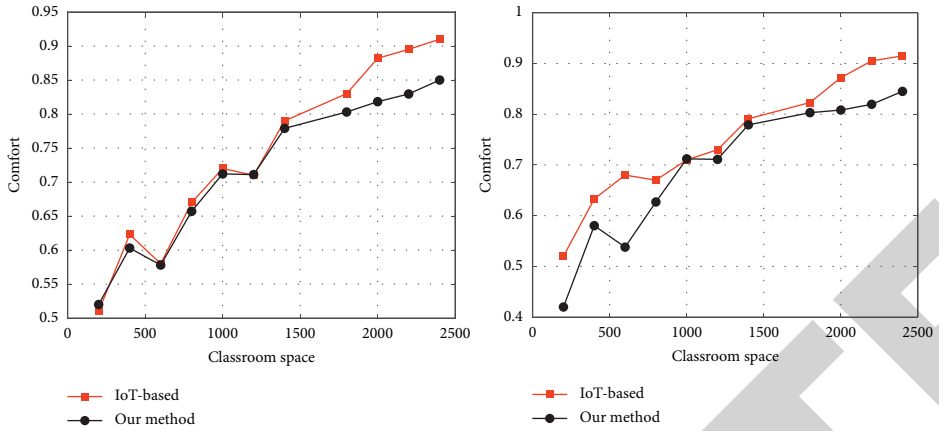


Figure 9: Design of personalized evaluation program based on process information.

also discover the differences between any two content versions through the version comparison function.

Interaction in ubiquitous learning is not only the interaction between learners and physical learning resources, but also the process of learning, drawing on the wisdom of others, and establishing dynamic connections between learners and teachers through learning resources, so that learners can acquire new knowledge from their peers and receive help in learning. This trend has led to the inclusion of "people" as an important resource for learning.

Learning resources in LCS not only include learning contents and learning activities, but also attach social cognitive network properties to the learning contents as shown in Figure 8. Learners learning the same or similar topics can also realize social cognitive networks through learning resources, which is consistent with the value of "connection and recreation" advocated by the constructivist view of learning. As learners continue to interact with each other, a cognitive network with the same learning interests and hobbies and frequent interactions will gradually form. Each learner is an entity node in the cognitive network space and can establish learning connection with different learner nodes through learning resources. The strength of connection between nodes is represented by a multifactor composite cognitive model, and as learners continue to learn and interact, the status and connections of nodes in the learning community network will be continuously updated, so as to obtain knowledge and wisdom from each other, thus promoting the learners' learning.

In addition, ubiquitous learners differ from traditional learners in that the former have clear learning purpose and need, and their learning is generally goal-driven; they simply want to understand some knowledge and wish to master other in depth. Therefore, the effectiveness of learners' learning cannot be measured by a uniform evaluation standard. Therefore, with different learning targets and different learning goals, ubiquitous learning environments need to provide learners with personalized evaluation criteria to measure the effectiveness of different learning targets in achieving their different learning goals. Ubiquitous learning requires the provision of evaluation based on process information, which is recorded as the main basis for evaluating the effectiveness of learners' learning (as shown in Figure 9).

8. Conclusions

The LCS records all kinds of process information generated by learners during the learning process and classifies them into five categories: learning attitudes, learning activities, content interactions, resource tools, and evaluation feedback. The evaluator (usually the creator of the resource, assuming the role of the teacher) selects appropriate information according to different learners and different learning objectives and presets a number of personalized evaluation schemes. The system selects the appropriate solution as the evaluation criteria based on the learner's learning objectives and knowledge mastery, then collects data based on the evaluation criteria, and calculates the evaluation results using a simple and easy to understand weighting method. To ensure the accuracy of the evaluation, the evaluator is allowed to manually modify the evaluation results according to the learners' specific performance. Both the evaluator and the learner can view the current evaluation results in real time.

In future research, the data in this paper needs to be expanded, for example, increasing the map area, adding types of IoT devices, and using artificial intelligence for research.

Data Availability

The dataset used in this paper is available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding this work.

Acknowledgments

This work was supported by the Department of Science and Technology of Hubei Province, "Research on Semantic Fusion Method of Medical Data Based on Ontology," under Grant no. 2020CFB675.

Hindawi
Scientific Programming
Volume 2022, Article ID 3828478, 11 pages
https://doi.org/10.1155/2022/3828478

Research Article

Video Face Detection Technology and Its Application in Health Information Management System

Yuxuan Liao,[1] Zhenyu Tang,[2] Jiehong Lei,[1] Jiajia Chen,[1] and Zhong Tang[3]

[1]*School of Information and Management, Guangxi Medical University, Nanning, Gangxi 530021, China*
[2]*Department of Computer Engineering, University of CA, Irvine, CA 92697, USA*
[3]*School of Humanities and Social Sciences, Guangxi Medical University, Nanning, Gangxi 530021, China*

Correspondence should be addressed to Zhong Tang; tangzhong@stu.gxmu.edu.cn

Received 9 December 2021; Revised 2 January 2022; Accepted 10 January 2022; Published 4 February 2022

Academic Editor: Baiyuan Ding

Computer face detection, as an early step and prerequisite for applications such as face recognition and face analysis, has attracted people's attention for a long time. With the popularization of computer applications, the improvement of performance, and the gradual maturity of research in the field of image processing and pattern recognition, face-related applications have become more and more a reality, so the research on face detection and positioning is also receiving more and more attention. Face detection and positioning are an important part of face analysis technology. Its goal is to search for the location of facial features (such as eyes, nose, mouth, and ears) in images or image sequences. It can be widely used in the fields of face tracking, face recognition, gesture recognition, facial expression recognition, head image compression and reconstruction, facial animation, etc. Based on the health information management system, this study mainly discusses the application of face recognition technology in video systems. Compared with other biological characteristics, such as fingerprints and eye masks, human faces are easier to obtain. In research and exploration, stable and effective face detection and face recognition algorithms have been proposed, which can achieve good recognition results even in real-time video surveillance. Aiming at the automatic face recognition technology in video surveillance, this study introduces in detail the video face detection technology in the health information management system of video image collection, image preprocessing, face detection, and face recognition. The prototype system of hygiene management is recognized.

1. Introduction

Face recognition is a kind of biometric identification technology based on human facial feature information. The research on face recognition (as shown in Figure 1) can be traced back to a study published by Galton in "Nature" magazine in 1888, which proposed the use of face information for personal identification, but it does not involve the problem of timely recognition of faces. In 1965, Bledsoe_ proposed a semiautomatic face recognition system model, which is the earliest research on face detection and recognition. The model is mainly based on the geometric features of the face, such as facial features and skin color. Through the standard two-dimensional pattern recognition method, the face image is expressed as geometric feature parameters. At the end of the 1970s, a real face recognition system began to appear [1–10]. In 1973, Dr. Kanade used the method of integration to design a backtracking recognition system based on a certain amount of artificial assistance so that the face can achieve a good matching recognition. Parke has established a high-quality face grayscale model through a computer, but the detection and recognition of human faces require prior knowledge of these human faces and cannot achieve the function of automatic recognition. By the 1990s, on the basis of a large number of scientific research results, face recognition has also received more and more attention from people around the world. Many classic algorithms are produced, among which K-L transformation achieves the purpose of reducing the dimensionality of facial image features, which becomes the basis of the later principal component analysis method. Subsequently, the eigenface method based on the

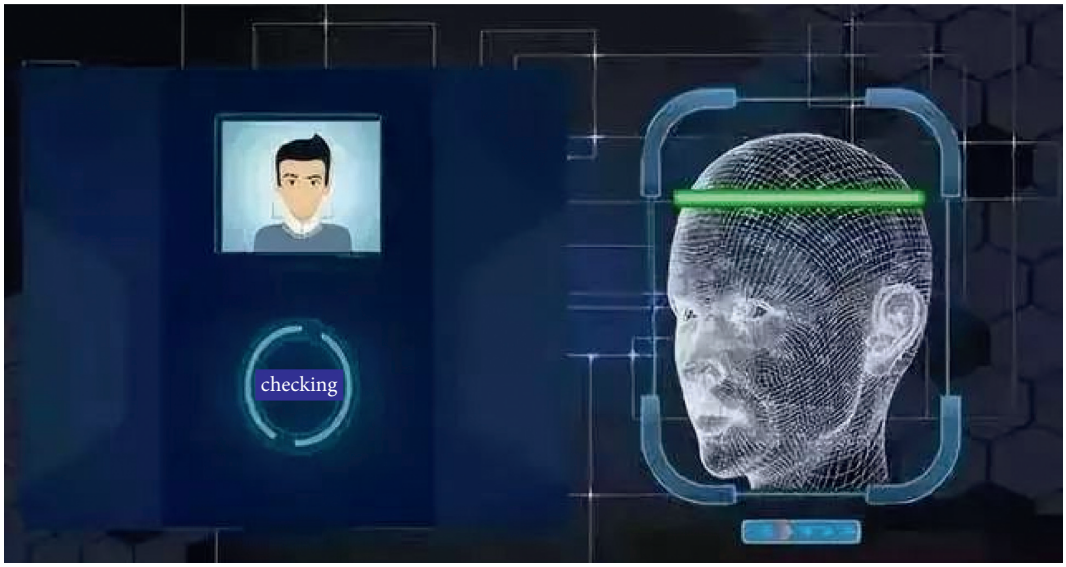


Figure 1: Face recognition.

template has appeared one after another, which has played a very good effect in the actual system application. The method based on the template has a better effect on face detection and recognition than the method based on geometric feature information. Fisher's linear discriminant analysis method developed on the basis of principal components uses linear judgment analysis to select the most advantageous feature points, and it is still widely used. After 1998, the system based on face detection and recognition began to enter practical research. The entire system was guaranteed to be automatically completed without human involvement. The focus of the research was also on solving the detection rate caused by illumination and complex background. On the low issue, support-vector machines and three-dimensional illumination models are applied to face detection, making the system closer to the thinking of real people, and the researched system is also more commercialized and has achieved good market response [11–15].

The main methods of face recognition technology include recognition methods based on geometric features, recognition methods based on algebraic features, recognition methods based on connection mechanisms, face recognition technologies based on three-dimensional data, and hybrid recognition methods of several technologies. A series of related technologies use cameras or cameras to capture images or video streams containing human faces, automatically detect and track human faces in the images, and then perform face recognition on the detected faces.

With the recognition method based on geometric features, the geometric feature vector represents the face, and the feature vector is constructed by extracting some feature points on the front contour line from the face profile and is used to recognize the face. Recognition methods based on algebraic features are also called eigenface recognition, including principal component analysis method, independent component analysis, and methods based on hidden Markov models. The principal component analysis party performs orthogonal transformation according to the statistical characteristics of the image and obtains the eigenvectors with successively decreasing eigenvalues, that is, eigenfaces. The high-dimensional space representation of the image can be converted to the low-dimensional space through orthogonal transformation. The independent component analysis method performs some linear decomposition of the observed data. Hidden Markov models used for face recognition are one-dimensional hidden Markov models and two-dimensional hidden Markov models. The one-dimensional orthogonal model first transforms the two-dimensional image signal into a one-dimensional observation sequence and divides the face image into strip regions from top to bottom, and each region represents a state in the continuous hidden Markov model. Recognition methods based on connection mechanism mainly include artificial neural networks and elastic matching methods. The neural network is a complex system composed of a large number of simple processing units (neurons) interconnected to solve recognition problems. Commonly used neural networks are BP neural network, self-organizing neural network, convolutional neural network, radial basis function network, and fuzzy neural network. The elastic matching method is a method based on a dynamic link structure. Face recognition based on three-dimensional data mainly includes a method based on curvature and a method based on model synthesis. Since the curvature is the most natural and basic local feature that expresses the surface, the curvature of the face surface was first used to deal with the 3D face recognition problem [16–22].

The above various face recognition methods have achieved certain success, but each has advantages and disadvantages. The processing of obstructions such as glasses is more difficult, and the robustness to large expression changes or posture changes is also relatively poor. The recognition method based on algebraic features extracts principal components through

different transformation methods [23–27]. However, this method has insurmountable defects, especially when the ambient light changes, and the recognition effect will sharply drop, which cannot meet the needs of the actual system. The advantage of the recognition method based on the connection mechanism is that it saves the material information in the face image while avoiding the more complicated feature extraction work. However, due to the huge amount of original image data and the time-consuming process, when the number of samples greatly increases, its performance will sharply drop. Face recognition methods based on three-dimensional data are characterized by the use of three-dimensional data, which opens up new ideas for face recognition, but it is difficult to obtain information sources, and the amount of data storage and calculation is also huge [28–33].

2. Image Preprocessing

The traditional face recognition technology is mainly based on the face recognition of visible light images, which is also a familiar recognition method with more than 30 years of research and development history. Face image preprocessing is an important link in the research and development of face recognition in video images, as shown in Figure 2. The video image quality is poor, the background is complex, the face is unpredictable, and the system must be guaranteed under a large amount of video image data. Those are the real-time requirements. Therefore, the quality of image preprocessing before face detection and recognition will directly affect the performance of the video system. The process is shown in Figure 3. Due to the changeable environment of video image capture, affected by the equipment itself, the captured video images often have various image quality problems. Through the continuous efforts of scientists, image preprocessing technology has also been greatly developed. A system has been formed for static digital image preprocessing, and it can achieve good results. Image enhancement, image grayscale transformation, and image normalization process the image, which will be separately described below. In face image preprocessing, image preprocessing for faces is a process of processing images based on the results of face detection and finally serving feature extraction.

2.1. Eliminate Image Noise. There is no specific definition of image noise, and there are different classifications in different fields. If it is defined according to the shape of the amplitude distribution generated by the noise, the amplitude distribution is called Gaussian noise according to the Gaussian distribution, and its probability density function is as follows:

$$p(z) = \frac{1}{\sqrt{2\pi}\delta} e^{-\left((z-u)^2/2\delta^2\right)}. \quad (1)$$

Z represents the gray value of the image pixel, Z represents the gray value of the image pixel, μ represents the average or expected value of Z, and δ represents the standard deviation of Z. The square of the standard deviation δ^2 is called the variance of Z. Salt and pepper noise is usually generated in the image sensor, transmission channel, decoding process, etc. The human eye can observe the bright and dark spots in black and white. The original image obtained by the system cannot be directly used due to various limitations and random interference, and it must be preprocessed such as grayscale correction and noise filtering in the early stage of image processing.

A large number of experimental studies have found that the digital images acquired by video surveillance cameras are more affected by salt and pepper noise and zero-mean Gaussian noise. Therefore, this study will also make the corresponding image preprocessing for these two types of noise. The salt and pepper noise uses median filtering, and the Gaussian noise uses wavelet analysis for denoising.

The median filter is a nonsmooth linear filter, which can eliminate noise while preventing the image from becoming blurred, protecting the details and edge information of the image. The median filter method has a good effect on filtering out salt and pepper noise.

Median filtering is to take the value of each pixel in the image in its square area and sort the pixel values in the area from small to large. If there are an odd number of pixel values, the middle one is taken as the pixel value of the corresponding position in the new image. If it contains an even number of pixel values, the average value of the middle two is taken as the pixel value of the corresponding position of the new image. This can eliminate isolated noise points. Time vs normalized frequency is shown in Figure 4.

In fact, the median filter is to select the closest value of the square pixel value of the area as the gray value of the pixel so that special points, namely, noise points, can be filtered out and that isolated noise points, namely, salt and pepper noise, can be eliminated by this method. Compared with the mean filter, the median filter does not change the value of the pixel point in the field but selects a representative point, and the resulting ambiguity is smaller. The solutions to the lighting problem include 3D image face recognition and thermal imaging face recognition. However, these two technologies are far from mature, and the recognition effect is not satisfactory.

Wavelet transform is an image analysis method based on image resolution. Through the multiscale refinement analysis of image signals, it can extract information on the tiny details of the image. It has been attracting attention in recent years. When denoising the image, it is necessary to analyze the affinity between the noise and the image information. When the overlap between the noise signal and the image signal is small, the traditional Fourier transform can be used for the denoising process, and the noise can effectively remove pollution, but when the noise signal and the image signal overlap a large part, the removal of noise by using this method will affect the image quality and damage the detailed information of the image. While transforming the noise signal, it will also transform the image signal. The feature extraction effect of the image is affected. Using wavelet transform to remove image noise can solve this problem. Wavelet analysis can remove image noise in the high-frequency region while retaining the image detail information in the high-frequency region. Power vs frequency is shown in Figure 5.

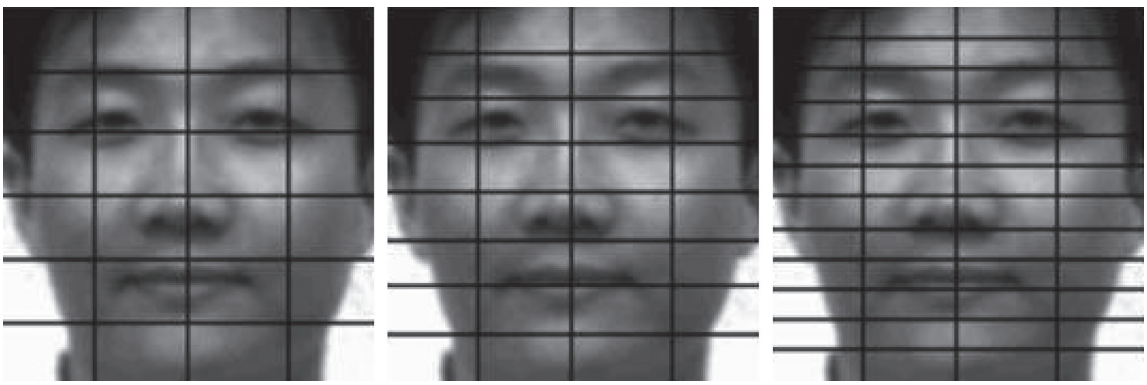


Figure 2: Image preprocessing.

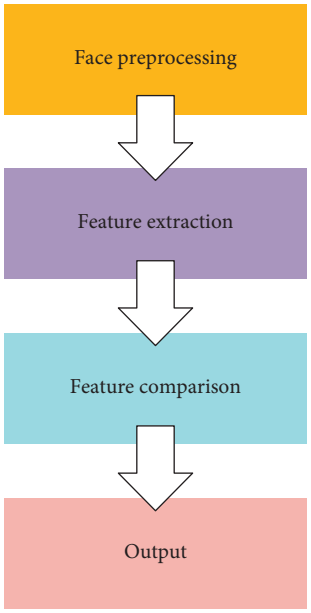


Figure 3: Process.

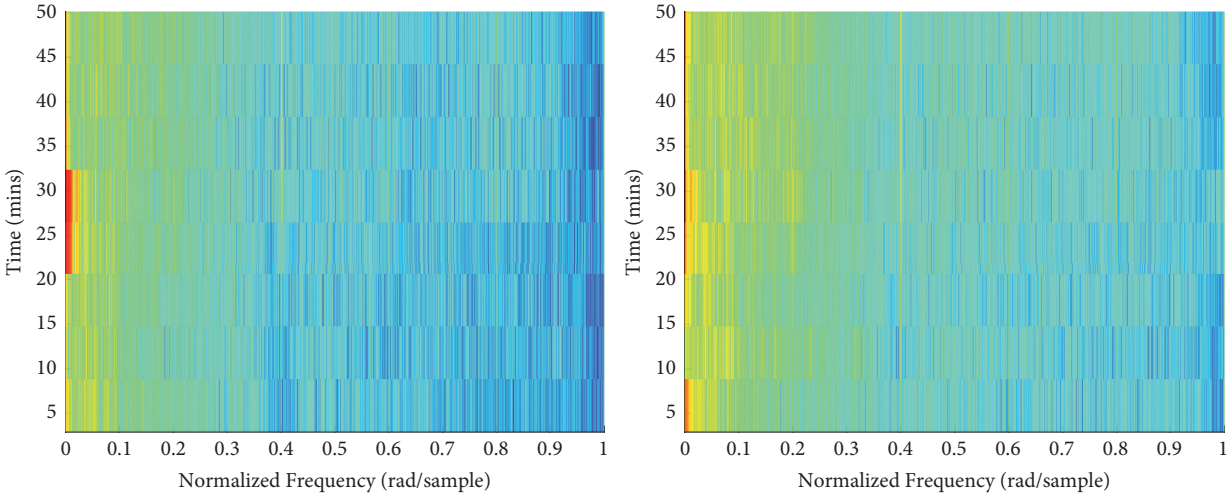


Figure 4: Time vs normalized frequency.

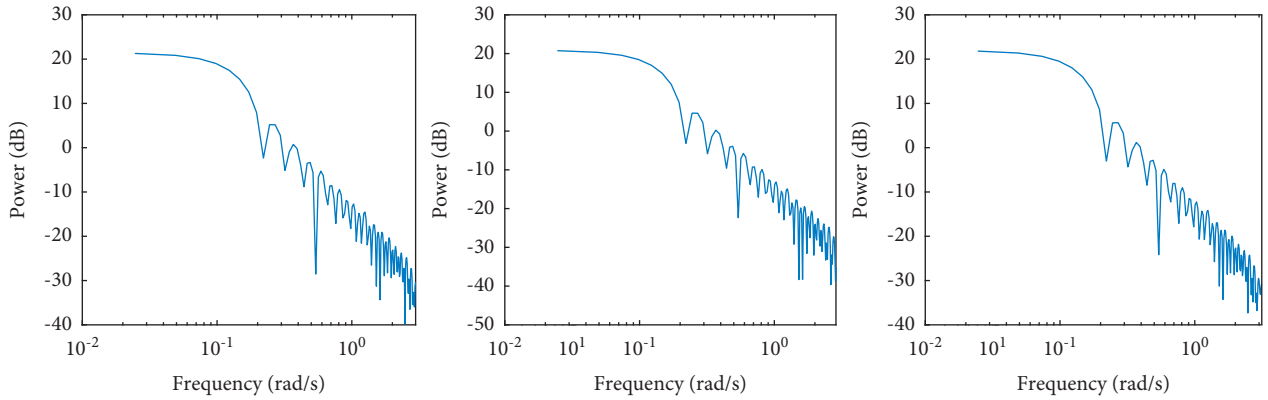


Figure 5: Power vs frequency.

The basic idea of wavelet analysis is as follows: a basic wavelet function is defined, operations such as scaling or translation are performed on this function, and multiscale refinement is performed to obtain a series of function clusters, which are called wavelet bases. When using wavelet transform to process an image, a set of wavelet bases is used to approximate an image signal, and then, the signal is projected onto the wavelet base to realize wavelet transform.

The basic idea of image denoising using wavelet analysis is as follows: through wavelet transform, most useful signals are compressed while dispersing noisy signals. Wavelet decomposition of image information is to project the image signal to each orthogonal basis and then obtain the relationship coefficient between the signal and a series of wavelet bases, that is, the wavelet transform value, to achieve the purpose of the wavelet decomposition. The wavelet decomposition of the image makes the image key information and the image noise information have different decomposition results. The wavelet transform value of the image's own signal only has a large value on some scales, while the noise signal after wavelet analysis is on most scales. Both have larger wavelet transform values. In this way, through the decomposition and reconstruction of the image signal, the purpose of denoising can be achieved.

The local singularity of the image signal can be described by the Lipschitz index. A function f(x) is defined with the following:

$$\left|f(x) - f(x_0)\right| \le K\left|x - x_0\right|^{\alpha}. \tag{2}$$

There is a unique constant k that satisfies formula (2). If formula (2) holds for all x and x0, then f(x) is consistent with Lipschitz a in the interval (a, b). If the function has a discontinuity or a certain order of derivative discontinuity at a point, then this point can be described as a singular point of the function. Then,

$$W_s f(x) \le As^{a}. \tag{3}$$

Among them, Wsf(x) is the wavelet transform of f(x) on the scale s, and then, equation (3) becomes

$$W_{2^j} f(x_k) \le A2^{ja}. \tag{4}$$

The logarithm of both sides is as follows:

$$\log_2\left|W_{2^j} f(x_k)\right| \le \log_2 A + ja. \tag{5}$$

It can be seen that the coefficients of the wavelet transform are affected by the Lipschitz exponent a of the function f(X). When a>0, the coefficients of the wavelet transform are proportional to the scale. If a<0 of the function f(x), the wavelet coefficient is inversely proportional to the scale. The Lipschitz exponent of a function reflects the singularity of the function at a certain point. The larger the scale, the larger the wavelet transform amplitude corresponding to the signal, but the smaller the wavelet transform amplitude corresponding to image noise. Therefore, different thresholds can be set on different decomposition scales, the wavelet transform smaller than the given threshold is regarded as the wavelet transform of noise, the pixel value is set to 0, and the wavelet transform larger than the threshold is regarded as the wavelet transform. The wavelet transform of the signal preserves the pixel. By reconstructing the image after wavelet transformation, the purpose of noise removal can be achieved. The prediction is shown in Figure 6.

2.2. Fractional Differentiation Realizes Image Enhancement. The Cauchy definition of the fractional differential is as follows:

The Cauchy definition form is based on the traditional definition form of integer order calculus, which is an extension of integer order calculus:

$$D^{\gamma} f(t) = \frac{\Gamma(\gamma + 1)}{2\pi j} \int_c \frac{f(\tau)}{(\tau - t)^{\gamma+1}} \mathrm{d}\tau, \tag{6}$$

where c is the smooth curve surrounding f(t) single value and analytic domain, it is the Gamma function, and its integral transformation form is as follows:

$$\Gamma(x) = \int_0^{\infty} y^{x-1} e^{-y} \mathrm{d}y. \tag{7}$$

The Grunwald-Letnikov definition of the fractional differential is as follows:

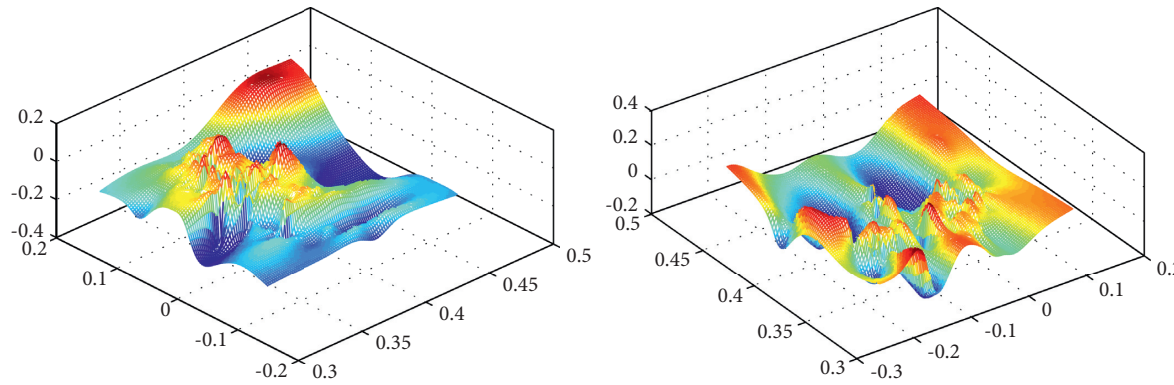


Figure 6: Predicted results.

$$\lim_{h\longrightarrow 0} s_h^{(v)}(t) \approx h^{-v} \sum_{r=0}^{[(t-a)/h]} w_r^{(v)} s(t-rh), \tag{8}$$

where

$$w_r^{(v)} = (-1)^r \binom{v}{r}. \tag{9}$$

In a complete digital image, if the human eye can see a clear image effect, it means that there must be a certain correlation between adjacent pixels in this image, and the image signal is highly self-similar. The similarity is usually expressed in complex texture details. Fractional differentiation can be used to enhance the complex texture details in the image signal.

2.3. Image Normalization Processing Based on Face. The key point of a video surveillance system based on face recognition is to obtain as good a face image as possible and to accurately obtain the feature points of the face image when it is convenient to use. After the images captured from the video are processed by denoising, image enhancement, etc., a series of issues such as illumination and complex background need to be considered when performing face detection so that the obtained face images are as close to the same standard as possible. The image feature points will better reflect the facial features. The following will introduce how to normalize the face image. The data variation is shown in Figure 7.

When forming a standard face database, it is necessary to normalize the size of the face image and adjust the image to a certain size to facilitate feature extraction later and calculate the recognition rate. There are many methods for image enlargement and reduction. In fact, image enlargement can be regarded as excessive, and reduction can be regarded as undersight, but enlargement and reduction are applicable to digital images. Image scaling is not to simply obtain the pixels of the source image, and it is often necessary to find new pixels to replace, which can be approximated by the method of difference. The following describes two commonly used methods for calculating interpolation: nearest-neighbor interpolation and linear interpolation.

2.3.1. Nearest Neighbor Interpolation. The nearest-neighbor interpolation is to select an input pixel value closest to the desired pixel as the pixel gray value of the point. It is a simple interpolation method with intuitive thinking, but the quality of the image obtained is not high. When there are slight changes in the pixel gray level in the image, the processed image obtained by this method will be unnatural and appear as artificial traces.

2.3.2. Bilinear Interpolation. Bilinear interpolation is a method of equally dividing the width and height of the input image according to the width and height of the output image to determine the gray value of the output image. The interpolation function is as follows:

$$\begin{aligned} f(R_1) &\approx \frac{x_2 - x}{x_2 - x_1} f(p_{11}) + \frac{x - x_1}{x_2 - x_1} f(p_{21}), \\ f(R_2) &\approx \frac{x_2 - x}{x_2 - x_1} f(p_{12}) + \frac{x - x_1}{x_2 - x_1} f(p_{22}). \end{aligned} \tag{10}$$

Then, linear interpolation is performed in the y direction to get

$$f(P) \approx \frac{y_2 - y}{y_2 - y_1} f(R_1) + \frac{y - y_1}{y_2 - y_1} f(R_2). \tag{11}$$

The result of linear interpolation has nothing to do with the order of interpolation. Regardless of whether the interpolation in the y direction is performed first, or the interpolation in the x direction is performed first, the result obtained will not change.

3. Face Recognition

Face image feature extraction is as follows: the features that can be used by a face recognition system are usually divided into visual features, pixel statistical features, face image transformation coefficient features, and face image algebraic features. With the continuous development of science and technology in recent years, face recognition technology has also achieved good developmental effects in different fields, and many practical technical methods have emerged. Under different lighting conditions and complex backgrounds, face

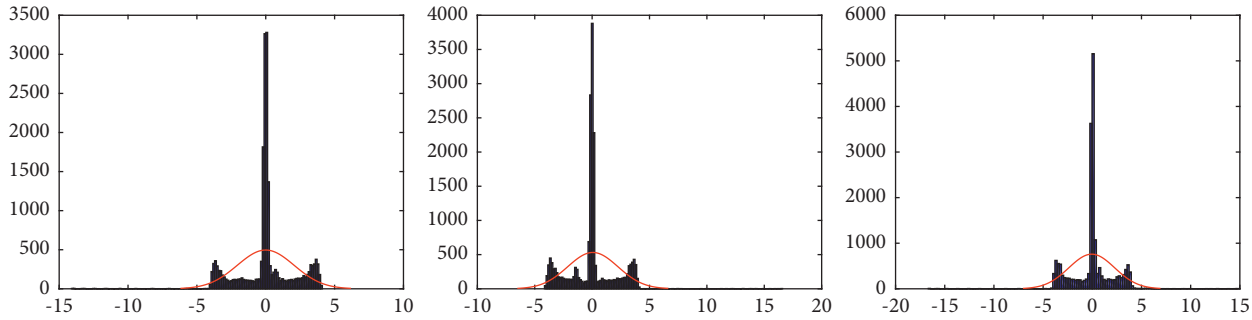


Figure 7: Data variation.

recognition technology has made great progress. It can achieve the expected effect in a specific application environment, but there is still a certain distance from the recognition ability of the human eye itself. Obtaining it is easier than fingerprints, eye masks, etc., but it also leads to a variety of facial features, which are more unpredictable, and it is difficult for computers to express all possible conditions through numbers. The prediction is compared in Figure 8.

At present, the common basic algorithms for face recognition can be divided into several categories: methods based on geometric features, subspace methods, neural network methods, and hidden Markov methods. It involves computer vision, pattern recognition, image processing, neural network, and other related discipline theories.

The method based on geometric features is one of the earliest face recognition methods. In fact, a face image can be regarded as composed of eyes, nose, mouth, etc. The difference between a face and a face lies in these face components and is not the same as the difference between color, size, shape, and so on. The method based on geometric features is to perform face recognition through the description of geometric features such as eyes, nose, mouth, and face shape, to obtain the geometric relationship between these geometric features, which is represented by feature vectors, angles, and curvatures., and then match these features. The similarity between vectors is usually calculated by Euclidean distance. The classic algorithms based on geometric features include the active contour model proposed by Huang to extract the contours of face parts and the deformable template model proposed by Grenander. The face recognition algorithm based on geometric features is simple and easy to understand, adapts to certain lighting changes, and has a small storage space; however, a unified feature extraction standard is not formed, the obtained features are not stable enough, and they cannot achieve good results in the case of complex expressions and side faces, such as face recognition effect.

The main idea of face recognition based on subspace is to map an image to a low-dimensional subspace through a certain spatial transformation, thereby reducing the complexity of classifying face images in high-level spaces and improving computational efficiency. The main subspace methods at this stage include principal component analysis (PCA), linear discriminant analysis (LDA), independent component analysis (ICA), etc. Turk first used the PCA method for face recognition and developed it into the later eigenface method. Its main idea is to obtain eigenvalues and corresponding eigenvectors by performing KL transformation on the feature space and arranging them in descending order, before selecting N feature vectors are used as principal components to obtain a low-dimensional face vector space to achieve the purpose of dimensionality reduction. The main idea of the linear discriminant analysis (LDA) method is to find a linear transformation so that the distance between the classes is the largest and the distance within the classes is the smallest after the training sample is projected to the linear transformation. When there are many training samples, this method is better than the PCA algorithm. Face feature extraction is carried out for certain features of the face. Facial feature extraction, also known as face representation, is the process of modeling features of human faces.

The main idea of the independent component analysis (ICA) method is to find a set of mutually independent components from the training sample through linear transformation and use this to describe the sample data containing high-order statistical information. The ICA algorithm is a generalization of the PCA algorithm. It is widely used in the field of face recognition. The neural network-based face recognition method mainly uses neural network theory for facial feature extraction and recognition. The neural network simulates human thinking through computers and strives to be closer to the learning process of the human brain.

Completing complex tasks by coordinating a large number of simple arithmetic processing (neurons) has certain advantages for face processing with complex structures. The advantage of the neural network is that there is no fixed process, but it is constantly adjusted according to the changes in conditions and environment to obtain the inherent laws and characteristics of the face image. The disadvantage is that the number of neurons is large and the training time is long, which is difficult to implement in practice.

The hidden Markov model (HMM) for face recognition was first introduced by Samaria and then widely used in the field of face recognition. The hidden Markov model is based on statistical analysis and is a form of Markov chain. It consists of a Markov chain with specific state values and a set of random functions to form a bidirectional random process,

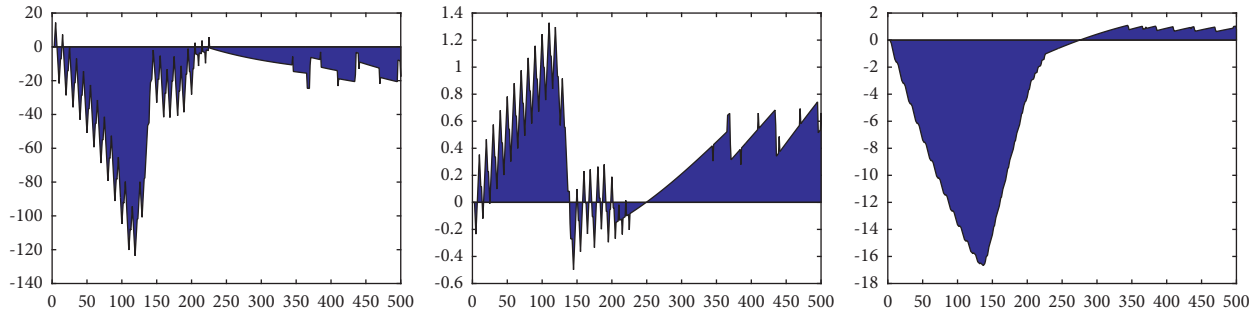


Figure 8: Prediction comparison.

which can extract the face image. Interrelationships between various organs are as follows. Face recognition based on the HMM model also has better recognition results when the facial expressions are complex and the face scale greatly changes, but the feature extraction and model training operations are large.

In addition, there are methods for face recognition based on elastic graph matching, support-vector machines, and other methods. In actual applications, it is often difficult to meet the needs by using only one method. Therefore, it is necessary to use the advantages and disadvantages of various methods and the scope of application. Different methods are combined to achieve better recognition results.

Face recognition has gradually entered the research of three-dimensional face recognition from the initial recognition of frontal two-dimensional images to the subsequent recognition in multiple poses and complex environments. With the deepening of research, the performance of face recognition systems has also continued. However, due to the complex environment people are in and the influence of age, face occlusion, etc., the robustness of the face recognition system still needs to be improved. The face features extracted by the SIFT algorithm have strong robustness and rich information and have good feature extraction effects under the interference of image scale transformation, noise, brightness, etc., and have been widely used in images in recent years. such as retrieval, image matching, and image tracking, and this method has been proven to have higher matching performance than other local features of the same type and has high research value and practical significance in the field of face recognition. The "eigenface" method represented by PCA is a benchmark algorithm for detecting the performance of face recognition systems. This study will analyze the specific application of the PCA algorithm and SIFT algorithm in face recognition and compare the advantages and disadvantages of the two through simulation experiments. The recognized portrait is shown in Figure 9.

This part introduces the specific implementation process of face recognition using the PCA algorithm on the ORL face database.

(1) The face database is read, the ORL database contains a total of 40×10 images, and each image is 92×112 in size. Each image is converted into a row vector of 92×112 columns, and 6 images of each person are selected as training samples, that is, 240 images. The image forms a matrix of 240 rows and 92×112 columns. The formula is to calculate the average image vector of the training sample to get the average face, then calculate the interpolation between each image and the average face, and then build the covariance matrix from the formula.

(2) The singular value decomposition method is used to calculate the eigenvalues and eigenvectors of the covariance matrix, arrange the eigenvalues and their corresponding eigenvectors in descending order, and select the top P largest eigenvalues and their corresponding eigenvectors. These P feature vectors (principal components) are constructed into a feature face space.

(3) Six images of each person are selected as training samples, and the remaining 4 images are selected as test samples, and the difference between the training sample face and the average face to the eigenface space is mapped so that the matrix can be used as the basis for face recognition. Each image of the test sample is mapped to the eigenface space, and the obtained vector can be used as a comparison with the eigenface space of the training sample.

(4) There are many ways to compare the test sample and the training sample. To find out which sample the face image to be recognized belongs to, you can first simplify the 6 images of the training sample to form a feature that can approximately represent this class vector, and the simplest way is to do the arithmetic average of 6 vectors, compare the test sample with 40 feature vectors representing 40 classes, and divide it into the closest class; you can also compare the sample to be identified with the training sample, compare each feature vector of, and classify it into the category of the image closest to it. The disadvantage is that if there is a special situation, it happens that there is an image of another category that is closest to it, which will cause and avoid misclassification. In this case, you can find the closest few pictures and divide them into the category with the largest number of similarities. In this experiment, the calculation of the distance between the vector and the vector uses the Euclidean distance metric.

The PCA algorithm and the SIFT algorithm are used to perform face recognition on the ORL database and the

Figure 9: Recognized portrait.

laboratory database under the condition of selecting different training samples and test samples. The overall face recognition efficiency is not as high as the recognition rate on the ORL database; as the number of samples increases, the recognition rate on both databases accordingly increases. The overall quality of the face images of the laboratory database is not as high as that of the ORL database and is affected by illumination and posture, so the recognition rate is not as high as the recognition rate on ORL.

The PCA algorithm is a characteristic face method. The entire recognition process is for the entire face image, not for the characteristics of the face. Therefore, the recognition result is greatly affected by the light, facial expressions, etc., and the difference of the characteristic face is obtained under different brightness of the big image. The PCA algorithm is only able to describe the difference between images not human faces. Different from the PCA algorithm, the SIFT algorithm realizes face recognition by extracting local feature points of the image, especially the extracted feature points are concentrated around the eyes, mouth, nose, etc. The algorithm has good invariance under different conditions such as rotation, scale scaling, and illumination changes. The feature information extracted by the SIFT algorithm is rich, and a large number of feature vectors can be extracted from a small amount of image samples. Mikolajczyk obtained through experiments that the matching performance of SIFT descriptors is higher than that of other local feature descriptors.

Through experiments, it is found that, compared to the PCA algorithm, the face recognition based on the SIFT algorithm can extract feature points for recognition without image normalization; for most algorithms, it is necessary to establish a training library first and then recognize the test samples after training. The SIFT algorithm does not need to perform sample training in advance, as long as the video image is collected, and the extracted facial features are matched with the sample features of the sample library. In summary, the SIFT algorithm has certain robustness for face recognition in the case of poses, expressions, and simple illumination changes. Although the feature vector extracted by the SIFT algorithm has 128 dimensions, it takes up a lot of space in the matching operations of a large number of databases, which affects the speed of face recognition, but is constantly improving and expanding, such as the SURF algorithm that reduces the dimensionality of the feature vector, the PCA-SIFT algorithm that combines with the PCA algorithm, and so on.

4. Conclusions

Based on the health information management system, this study mainly discusses the application of face recognition technology in video systems. Compared with other biological characteristics, such as fingerprints and eye masks, human faces are easier to obtain. In research and exploration, stable and effective face detection and face recognition algorithms have been proposed, which can achieve good recognition results even in real-time video surveillance. Aiming at the automatic face recognition technology in video surveillance, this study introduces in detail the video face detection technology in the health information management system of video image collection, image preprocessing, face detection, and face recognition. The prototype system of hygiene management is recognized.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Guangxi Health and Economic and Social Development Research Center Project (Research on the Practice Mechanism of Social Work in Public Health Emergencies in Guangxi; no. 2021RWB03) and 2021 Guangxi Philosophy and Social Science Planning Research Project (Research on the Construction of Urban Smart Home-Based Elderly Care Service System under the Healthy Guangxi Strategy; no. 21FSH017).

References

[1] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference On Computer Vision And Pattern Recognition*, pp. 770–778, Las Vegas, NV, USA, June 2016.

[2] M. Yun, J. Zhao, J. Zhao, X. Weng, and X. Yang, "Impact of in-vehicle navigation information on lane-change behavior in urban expressway diverge segments," *Accident Analysis & Prevention*, vol. 106, no. 1, pp. 53–66, 2017.

[3] S. Kumar Dwivedi, R. Amin, V. Satyanarayana, and R. Chaudhry, "Blockchain-based secured event-information sharing protocol in internet of vehicles for smart cities," *Computers & Electrical Engineering*, vol. 86, no. 1, pp. 1–9, 2020.

[4] Z. Khan and S. Amin, "Bottleneck model with heterogeneous information," *Transportation Research Part B: Methodological*, vol. 112, no. 1, pp. 157–190, 2018.

[5] J. M. Cairney, K. Rajan, D. Haley, and B. Gault, "Mining information from atom probe data," *Ultramicroscopy*, vol. 159, no. 1, pp. 324–337, 2020.

[6] J. Yu and P. Lu, "Learning traffic signal phase and timing information from low-sampling rate taxi GPS trajectories," *Knowledge-Based Systems*, vol. 110, no. 1, pp. 275–292, 2016.

[7] K. P. Wijayaratna, V. V Dixit, L. Denant-Boemont, and S. T. Waller, "An experimental study of the online information paradox: does en-route information improve road network performance?" *PLoS One*, vol. 12, no. 9, pp. 184–191, 2017.

[8] Z. Wang, H. Ren, Q. Shen, W. Sui, and X. Zhang, "Seismic performance evaluation of a steel tubular bridge pier in a five-span continuous girder bridge system," *Structures*, vol. 31, no. 1, pp. 909–920, 2021.

[9] S. Nakayama and J. Takayama, "Traffic network equilibrium model for uncertain demands," in *Proceedings of the 82nd Transportation Research Board Annual Meeting*, Washington, DC, USA, 2021.

[10] H. Shao, W. H. K. Lam, and M. L. Tam, "A reliability-based stochastic traffic assignment model for network with multiple user classes under uncertainty in demand," *Networks and Spatial Economics*, vol. 6, no. 3, pp. 173–204, 2019.

[11] A. Chen, J. Kim, S. Lee, and Y. Kim, "Stochastic multi-objective models for network design problem," *Expert Systems with Applications*, vol. 37, no. 2, pp. 1608–1619, 2020.

[12] H. Wang, W. H. K. Lam, X. Zhang, and H. Shao, "Sustainable transportation network design with stochastic demands and chance constraints," *International Journal of Sustainable Transportation*, vol. 9, no. 2, pp. 126–144, 2015.

[13] S.-M. Hosseininasab and S.-N. Shetab-Boushehri, "Integration of selecting and scheduling urban road construction projects as a time-dependent discrete network design problem," *European Journal of Operational Research*, vol. 246, no. 3, pp. 762–771, 2015.

[14] S.-M. Hosseininasab, S.-N. Shetab-Boushehri, S. R. Hejazi, and H. Karimi, "A multi-objective integrated model for selecting, scheduling, and budgeting road construction projects," *European Journal of Operational Research*, vol. 271, no. 1, pp. 262–277, 2018.

[15] M. Johnson, M. Schuster, Le et al., "Google's multilingual neural machine translation system: enabling zero-shot translation," *Transactions of the Association for Computational Linguistics*, vol. 5, no. 1, pp. 339–351, 2017.

[16] M. D. Moreno, "Translation quality gained through the implementation of the iso en 17100:2015 and the usage of the blockchain," *Babel*, vol. 1, no. 2, pp. 1–9, 2020.

[17] X. Wang, X. Yu, L. Guo, F. Liu, and L. Xu, "Student performance prediction with short-term sequential campus behaviors," *Information*.vol. 11, no. 4, p. 101, 2020.

[18] Q. Guo, Z. Zhu, Q. Lu, D. Zhang, and W. Wu, "A dynamic emotional session generation model based on Seq2Seq and a dictionary-based attention mechanism," *Applied Sciences*, vol. 10, no. 6, pp. 1–10, 2020.

[19] H. Ren, Xi Mao, W. Ma, J. Wang, and L. Wang, "An English-Chinese machine translation and evaluation method for geographical names," *ISPRS International Journal of Geo-Information*, vol. 9, no. 3, pp. 193–201, 2020.

[20] J. Arús-Pous, T. Blaschke, S. Ulander, J.-L. Reymond, H. Chen, and O. Engkvist, "Exploring the GDB-13 chemical space using deep generative models," *Journal of Cheminformatics*, vol. 11, no. 1, pp. 20–29, 2019.

[21] T. Moon, T. In Ahn, and J. E. Son, "Long short-term memory for a model-free estimation of macronutrient ion concentrations of root-zone in closed-loop soilless cultures," *Plant Methods*, vol. 15, no. 1, pp. 1–12, 2019.

[22] N. Pourdamghani and K. Knight, "Neighbors helping the poor: improving low-resource machine translation using related languages," *Machine Translation*, vol. 33, no. 3, pp. 239–258, 2019.

[23] L. Bote-Curiel, S. Muñoz-Romero, A. Gerrero-Curieses, and J. L. Rojo-Álvarez, "Deep learning and big data in healthcare: a double review for critical beginners," *Applied Sciences*, vol. 9, no. 11, pp. 1–11, 2019.

[24] J. Zhang and T. Matsumoto, "Corpus augmentation for neural machine translation with Chinese-Japanese parallel corpora," *Applied Sciences*, vol. 9, no. 10, pp. 1–12, 2019.

[25] Y. Chen, Y. Ma, X. Mao, and Q. Li, "Multi-task learning for abstractive and extractive summarization," *Data Science and Engineering*, vol. 4, no. 1, pp. 14–23, 2019.

[26] P. Zhou and Z. Jiang, "Self-organizing map neural network (SOM) downscaling method to simulate daily precipitation in the Yangtze and Huaihe River Basin," *Climatic and Environmental Research*, vol. 21, no. 5, pp. 512–524, 2016.

[27] X. Xiao, "Analysis on the employment psychological problems and adjustment of retired athletes in the process of career transformation," *Modern Vocational Education*, vol. 5, no. 12, pp. 216-217, 2018.

[28] S. Sahoo and M. K. Jha, "Pattern recognition in lithology classification: modeling using neural networks, self-organizing maps and genetic algorithms," *Hydrogeology Journal*, vol. 25, no. 2, pp. 311–330, 2016.

[29] Y. Zhou and B. Yang, "Sports video athlete detection using convolutional neural network," *Journal of Natural Science of Xiangtan University*, vol. 39, no. 1, pp. 95–98, 2017.

[30] J. Pang, "Research on the evaluation model of sports training adaptation based on self-organizing neural network," *Journal of Nanjing Institute of Physical Education*, vol. 16, no. 1, pp. 74–77, 2017.

[31] G. Querzola, C. Lovati, C. Mariani, and L. Pantoni, "A semi-quantitative sport-specific assessment of recurrent traumatic brain injury: the TraQ questionnaire and its application in

American football," *Neurological Sciences*, vol. 40, no. 9, pp. 1909–1915, 2019.

[32] J. Wang, X. Luo, and H. Yan, "Correlation analysis between injuries and functional movement screening for athletes of the National Shooting Team," *Journal of Capital Institute of Physical Education*, vol. 5, no. 4, pp. 352–355, 2016.

[33] G. Ma, "Research on the design of juvenile football players' sports injury prediction model," *Automation Technology and Application*, vol. 277, no. 7, pp. 141–144, 2018.

Hindawi
Scientific Programming
Volume 2022, Article ID 9415286, 10 pages
https://doi.org/10.1155/2022/9415286

Research Article

Design of Athlete's Running Information Capture System in Space-Time Domain Based on Virtual Reality

Chun Zhu

Wuhu Institute of Technology, Wuhu 241006, China

Correspondence should be addressed to Chun Zhu; zhunchun@whit.ah.cn

Received 10 November 2021; Revised 29 November 2021; Accepted 3 December 2021; Published 3 February 2022

Academic Editor: Baiyuan Ding

In order to improve the training effect of athletes, aiming at the problems of inaccurate information capture results, poor real-time performance of sports information capture, and inability to effectively suppress noise interference in traditional methods, an athlete space-time running information capture system based on virtual reality is designed. Establish the athlete's human skeleton, obtain the athlete's sports joint points, design the overall architecture of the athlete's space-time running information capture system, and realize the whole link design of sports information capture through RF chip, infrared camera, sports data acquisition module, data transmission module, and human-computer interaction module. Based on virtual reality technology, a virtual reality environment is built to obtain the characteristic parameters of athletes' sports posture in space-time domain, and the median filtering algorithm is used to filter the original signal to eliminate the impact of noise signal on athletes' sports information capture. Finally, the activity of the motion region is detected, and the motion information is captured combined with the Gaussian mixture model. The experimental results show that the system designed in this paper has high accuracy and anti-interference and can realize the real-time capture of motion information.

1. Introduction

According to the training characteristics of athletes, an athlete sports information collection system is designed, which can collect relevant data according to the athlete's training data by formulating a reliable communication protocol, using special instruments and transmitting them to the computer through IC card, display various data parameters, and automatically generate accurate training data from the original data by software [1–3]. Through the application of the system, it can provide auxiliary decision-making means for athletes to exercise scientifically and reasonably. Nowadays, mature human motion capture systems in the market include optical, electromagnetic, and ultrasonic [4–6]. These systems have disadvantages in practical application. They have high requirements for the site and inconvenient installation. The most important thing is that the price is expensive, which is difficult for most athletes to afford, making it difficult for athletes' motion capture systems to be popularized [7].

In view of the above problems, relevant scholars are committed to studying an effective athlete sports information capture system. Reference [8] proposes a human knee motion capture system based on MATLAB. MATLAB software is used to analyze and process the motion of human knee. Firstly, a program is written to process the coordinated data of human lower limb marker points measured by motion capture and interpolate them. Secondly, according to the obtained coordinate data of bone marker points in the process of human lower limb movement, the local coordinate system of human femur and tibia is established by using bone marker points. Finally, through coordinate transformation, the flexion, extension, adduction, abduction, and internal and external rotation of knee femoral tibial joint in human lower limb movement are calculated. This method can quickly process the motion data of human knee joint and obtain the kinematic characteristics of femoral tibial joint of human knee joint, which provides a reference for athlete training. Reference [9] designed a human motion trajectory tracking system based on three-frame difference method.

The system conducted in-depth research on the underwater environment, set up initialization layer, motion detection layer, and human tracking layer in the hardware architecture, and constructed three-dimensional human model, motion feature extraction module, and motion feature segmentation module so as to determine the boundary features and gray features of human body, contour features, and skin color features. The tracking trajectory of the system is highly similar to the actual running trajectory. Reference [10] proposes a three-dimensional human posture tracking method based on dual Kinect sensors, which represents the human motion posture by the method of human joint degree of freedom vector, uses the traceless Kalman filter method to track the human posture, and constructs a set of human posture tracking system based on dual Kinect sensors. Compared with the traditional human motion capture system, the system can realize three-dimensional human posture tracking and reflect the special properties of the motion process under complex actions.

Although the above systems have effectively solved the shortcomings of traditional methods, some new problems have been found in the application, such as the inaccurate information capture results of some joints in the motion state, the poor real-time performance of motion information capture, and the inability to effectively suppress noise interference. Therefore, this paper designs a running information capture system based on virtual reality. With the help of computers, virtual reality can bring people realistic feelings such as seeing and listening, and with the help of specific instruments and equipment, people can interact in the virtual environment created by computer.

2. Hardware Design of Athletes' Running Information Capture System in Space-Time Domain

2.1. Establishment of Athletes' Human Skeleton. In order to accurately describe the movement state of athletes, it is necessary to establish a skeleton model for the human body. As shown in Figure 1, some human joint nodes are given. There are 21 joint points in the human body, including 38 joint degrees of freedom and 6 degrees of freedom. The motion of the human body can be regarded as the motion of these joint points. When capturing the motion of these joint points, the motion of the skeleton can be determined.

It can be seen from the skeleton model that the rotational motion degrees of freedom of each marked point are not the same. According to the rotational motion degrees of freedom of the marked points, they can be divided into four categories:

(1) Fixed, nonrotatable marker points, can be used to measure the displacement of the whole skeleton model in space, such as hips.

(2) Hinge shaped, the marking point connects two parts, with a degree of freedom of rotation, such as the marking point at the elbow.

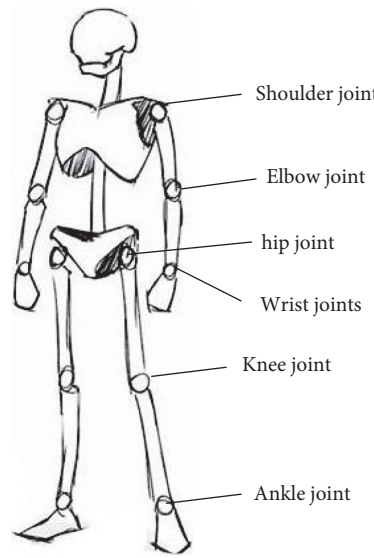


Figure 1: Athletes' skeleton model.

(3) Spherical, the marking point has three rotational degrees of freedom, such as the marking point of shoulder joint.

(4) Generally, this marker point has two rotational degrees of freedom, such as ankle marker point.

Such classification is helpful to reduce the time complexity of the algorithm and improve the accuracy of tracking in motion capture. In the specific operation process, too many marked points will increase the amount of calculation, and too few marked points will make the action impossible to be accurately restored, resulting in abnormal action. Therefore, the number of marked points should be determined according to the task to be completed.

2.2. Overall System Architecture and Hardware Module Design

2.2.1. Overall System Architecture. In the overall design of the athlete's time-space running information capture system, the signal input can be electrical signals, physical signals, and so on. In the forward channel, various signals are converted into electrical signals through sensors, and then the signals are converted into digital signals through ADC. PCI, VXI, ISA buses, and so on are inserted into the computer to work to realize the simulation of the athlete's running posture. Combined with the skeleton model of the human body, the sensor nodes are arranged to realize the full information perception of the athlete's movement posture.

This article comprehensively considers system technical indicators, system performance, development difficulty and development cost, and so on, chooses 5409A as the main processor of the system, and chooses PCI bus and PCI9054 bridge chip as the system data communication interface. The overall design model structure of the system is as shown in Figure 2.

Since the athlete's motion capture system requires a total of 21 joint data information, 17 sensors must be used to collect the attitude information of the carrier. These 21 sensors are

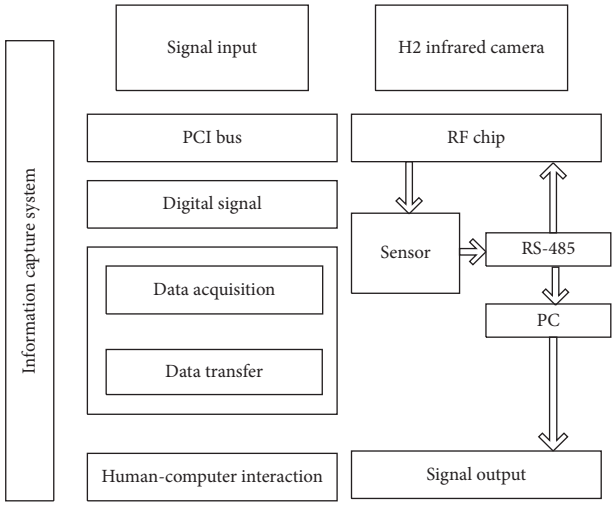


Figure 2: The overall architecture diagram of the athlete's running information capture system in the temporal and spatial domains.

divided into 3 channels, and each channel uses an RS-485 bus to connect 7 sensors and a signal catcher is then connected to the transmitter through the signal catcher and transmitted to the receiver on the PC side through the wireless network, and then the receiver is connected to the PC.

2.2.2. Hardware Module Design

(1) RF Chip Selection. In order to meet the requirements for real-time changing athletes' running posture recognition, this system adopts an ultra-high-frequency RFID system. Through comparative analysis, it is found that the nRF24LE1 radio frequency chip produced by Nordic has the characteristics of ultra-low-power consumption, small size, long range, and low cost. nRF24LE1 is a highly integrated, ultra-low-power consumption, system-on-chip technology 2.4 GHz radio frequency chip [11]. It integrates a radio frequency transceiver, an enhanced 51 FLASH high-speed microcontroller, and a wealth of peripherals and interfaces in a very small package. The radio frequency transceiver works in the ISM frequency band of 2.400 GHz to 2.483 GHz and adopts GFSK modulation and demodulation with strong anti-interference ability, and the transmission rate can reach 2 Mbps. The peak current provided by RX/TX is less than 14 mA, there is a power saving mode below μA, advanced power management, and the power supply range is 1.9 to 3.6 V. nRF24LE1 truly achieves ultra-low-power consumption performance. What is more worth mentioning is that nRF24LE1 provides an Enhanced ShockBurst^{TM} transceiver mode. In this mode, data is sent in at a low speed, but it can be transmitted at a high speed (2 Mbps), which not only saves energy, but also improves anti-interference. The embedded fast 8-bit MCU executes the traditional 8051 instruction set, but the speed is 8 times that of the traditional 8051 microcontroller. It has high execution capability and powerful functions, and high-speed data transmission through the SPI interface with the radio frequency transceiver [12].

In addition, the RFID system inevitably has to consider the signal collision problem of multiple tags entering the reading area at the same time, and the nRF24LE1 integrates the carrier monitoring function, which can accurately monitor whether the current working channel has interference. Therefore, the system chooses nRF24LE1 as the core chip, which can not only meet the requirements of low cost, low power consumption, high integration, and high performance, but also enhance the system's ability to prevent collisions and identify moving targets [13, 14].

(2) Infrared Camera. In order to collect high-definition images of athletes' running posture, the H2 model infrared camera was selected, which has a CAD infrared light bluetooth camera aimed at 60 meters. Using this camera, the photoelectric conversion main board and internal and external dual optical path correction optics can be used as compensation to reduce signal interference, so it can make the signal more stable and ensure that the measurement data are also in a stable state [15]. Choose a high-quality 2-million-pixel camera, equipped with a large-memory high-speed multimedia platform; the picture quality is clear and smooth, and it is easier to capture the athlete's movement posture. 1000 sets of data are stored and exported, which can be read directly in the machine, or exported using a USB connection to a computer [16]. The specific parameters of the H2 model infrared camera are shown in Table 1.

It can be seen from Table 1 that the H2 model infrared camera is a camera that integrates a camera, goggles, infrared camera, and heat sink. The basic principle is to use ordinary CCD black-and-white cameras to perceive the spectral characteristics of infrared light and combine infrared lamps as the "light source" for night vision imaging. The power, angle, camera configuration of infrared lamps, and infrared lenses with a certain focal length can realize rapid imaging of athletes' movement posture in a dynamic environment.

(3) Movement Data Acquisition Module. The generation and collection of motion posture data is the basis of the entire information capture system, which controls the dynamic gain code of the analog signal preprocessor and performs spectrum analysis on the signal [17, 18]. In summary, the data acquisition model of the motion posture data capture is obtained. The workflow of the motion data acquisition model is as shown in Figure 3.

In Figure 3, sensor nodes are arranged at each joint bone node of the human body to read the posture data of the human body movement and build a human body behavior database based on the read data information.

(4) Data Transmission Module. In the above motion data acquisition module, ARM Cortex-M3 series chips are used to collect data and send it to the transmitter. The transmitter also adopts ARM chip. The difference is that the device also includes a transmitting chip with 2.4 G wireless network, which uses NRF24L01 [19]. Through this chip, the data is sent out wirelessly. The receiver is also a processor with ARM chip and 2.4 G wireless receiving device with NRF24L01

Table 1: H2 model infrared camera parameters.

Parameter	Specification
Measuring range	>50m
Precision	±2%
Infrared camera resolution	256∗192
Visible light resolution	1920∗1080
Size	54∗25∗79 mm
Frame rate	25 Hz

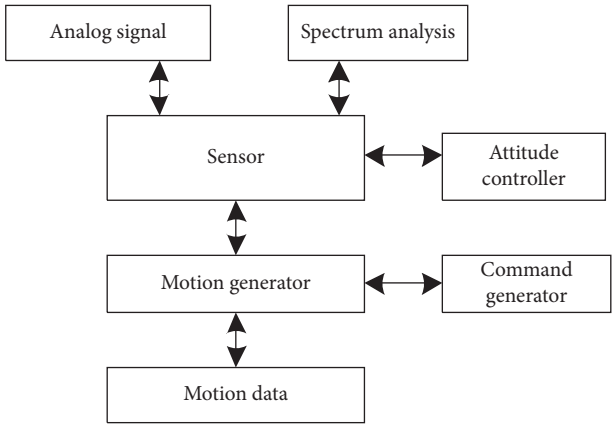


Figure 3: Workflow chart of the motion data acquisition module.

chip. After receiving the data from the transmitter through the receiver, it is sent to the PC through the serial port.

NRF24L01 is a single-chip wireless transceiver chip produced by NORDIC that works in the ISM frequency band of 2.4 GHz–2.5 GHz. The wireless transceiver includes frequency generator, enhanced "ShockBurst" mode controller, power amplifier, crystal oscillator, modulator, and demodulator [20]. It can almost be connected to various single-chip computers. Even if the SPI interface is not available, the SPI interface can be simulated through the ordinary IO port, and finally the data transmission is completed. Its main performance parameters are as shown in Table 2.

(5) Human-Computer Interaction Module. The ROS stack of the human-computer interaction module obtains the athlete's motion information through human body gesture recognition and voice recognition and displays the information on the system's human-computer interaction interface. The remote ROS stack is responsible for solving the received posture information and, at the same time, recognizing the athlete's current posture and feeding it back to the human-computer interaction control terminal [21, 22]. The real-time data synchronization and interaction between different ROS stacks are carried out through the wireless network and displayed on the main display interface.

The ROS stack of the human-computer interaction module includes two parts: the athlete's motion gesture recognition package and the adaptive speech recognition package. The human-computer interaction module obtains the athlete's movement information through the athlete's movement posture recognition and voice recognition and at the same time receives the movement data collected from the remote movement data acquisition module and displays the athlete's movement posture information on the human-computer interaction interface [23].

Table 2: Specific parameters of the data transmission module.

Parameter	Specification
Voltage working range	1.9V–3.6 V
Temperature range	−40°C–80 °C
Working frequency	2.4GHz–2.525 GHz
Transmit power	−6 dBm, −12 dBm, −18 dBm
Receive working current	12.3 mA
Launch operating current	11.3 mA
Data transfer rate	1 Mbps
Packet received length	1–32 bytes

Through the above analysis, it can be seen that the athletes' running information capture system designed in this paper realizes the whole link design of athletes' sports information collection, transmission, and interaction through infrared camera, sports data acquisition module, data transmission module, and human-computer interaction module so as to provide reliable hardware support for athletes' sports information capture.

3. Software Design of Running Information Capture System for Athletes in Time and Space

3.1. Construction of Virtual Reality Environment. Virtual reality is based on computer technology as the core, combined with related science and technology, to generate a digital environment that is highly similar to a certain range of real environments in terms of sight, hearing, and touch. Users use necessary equipment to interact and interact with objects in the digital environment and can experience the feeling of adapting to the real environment. Virtual reality is created by human beings in the process of exploring and understanding nature and gradually formed a scientific method and technology for understanding nature, simulating nature, and better adapting and using nature.

The construction of the virtual reality environment is the basis of the software design of the running information capture system of the entire athlete's time and space domain, which mainly includes three-dimensional visual modeling and three-dimensional auditory modeling [24, 25]. Among them, visual modeling mainly includes geometric modeling, motion modeling, physical modeling, object behavior modeling and model segmentation, and so on [26, 27]. Auditory modeling usually just adds interactive sound responses to the activities of users and objects. The modeling cycle of the virtual reality environment is as shown in Figure 4.

Geometric modeling is an abstract model used to describe the inherent geometric properties of an object. The content expressed includes the shape of the object and the appearance of the object. The shape of the object can be represented by points, straight lines, polygonal figures, curve or surface equations, or even images. And so on, it depends on the comprehensive consideration of storage and computing overhead. Abstractly representing the outline and

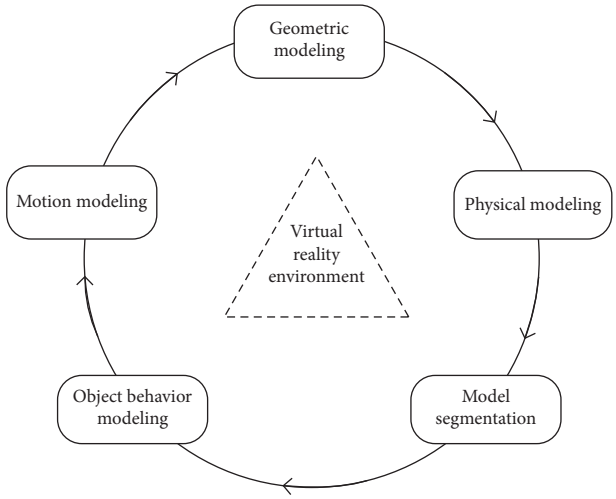


Figure 4: Virtual reality environment modeling cycle.

shape of primitives in an object is conducive to storage but needs to be recalculated when used. Specifically, the calculation time can be saved during generation, but the time and space overhead required for storage and access to the storage is relatively large. The content of the object's appearance description includes surface texture, surface light intensity, and surface color.

After setting the shape and appearance of the object, the next step is motion modeling [28]. In general, 4×4 homogeneous transformation matrix is used to describe the motion modeling so as to determine the position of the three-dimensional object in the coordinate system. The general form of homogeneous transformation matrix is

$$W_{x \leftarrow y} = \begin{bmatrix} P_{3\times3} & Q_{3\times1} \\ 0\,0 & 0\,1 \end{bmatrix}. \tag{1}$$

Among them, $P_{3\times3}$ represents the rotation submatrix of the coordinate system x relative to the coordinate system y; $Q_{3\times1}$ represents the position vector of the origin of the coordinate system x relative to the origin of the coordinate system y. The method of using homogeneous transformation matrix to orthogonalize two coordinate systems can save the amount of calculation. Rotation and translation can be defined in the same way, and they can also be combined. The inversion is relatively easy and suitable for complex modeling.

$$\begin{aligned} W_{x \leftarrow y} &= W_{y \leftarrow x} \\ &= \begin{bmatrix} P^W & P^W H \\ 0\,0 & 0\,1 \end{bmatrix}. \end{aligned} \tag{2}$$

When the object is scaled, it is only necessary to multiply the scale factor about the three axes of the world coordinate system in the diagonal elements of the $P_{3\times3}$ matrix; the rotation of the object only needs to be transformed in the rotation submatrix $P_{3\times3}$.

3.2. Obtaining the Athlete's Movement Posture Characteristic Parameters in the Time and Space Domain. Fully consider the time and space domain of the athlete's motion information so that real-time motion information can be obtained from the simulation image in virtual reality, and real-time motion simulation is realized. The captured motion information is simulated by the forward kinematics theory. In using virtual reality to simulate athletes' movements, each joint chain is set with at least 3 nodes, and at the same time, its own local coordinate system is established for the joint chain nodes [29, 30].

Set the coordinates of node u in the world coordinate system to (x_u, y_u), and set the coordinate position of node v to (x_v, y_v), and the position information of node u and node v is obtained as follows:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -x_v \\ 0 & 1 & -y_v \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_u \\ y_u \\ 1 \end{bmatrix}. \tag{3}$$

For multinode links, in fact, the conversion matrix between the local coordinate systems of u to v can be obtained by the mutual combination of adjacent parent nodes; namely,

$$D_{u \longrightarrow v} = d_{u \longrightarrow u+1} \cdot H_{v \longrightarrow v+1} \cdots H_{u \longrightarrow u+n} \cdot H_{v \longrightarrow v+m}, \tag{4}$$

where $D_{u \longrightarrow v}$ represents the combined transformation matrix from node u to node v. Through formula (4), all nodes in the coordinate system can be represented as a complete joint chain. Expand formula (4) and find that the coordinate position of the node can be expressed as a continuous function:

$$G = (g_1, g_2 \cdots g_N). \tag{5}$$

Among them, N represents the number of rotation angles between nodes. Through the acquisition of the characteristic parameters of the athlete's movement posture in the time and space domain, the joint movement of the athlete in the sense of the time and space domain is determined.

3.3. Athlete Motion Signal Filtering. In order to eliminate the influence of noise signals on athletes' motion information capture, it is also necessary to preprocess the collected data. This paper uses the median filter algorithm to filter the original signal. The median filter technology has a good filtering effect on impulse noise, especially while filtering the noise, it can still retain the characteristics of the original sensor data. The comparison effect of the collected sensor's original motion signal and the median filtered motion signal is as shown in Figure 5.

The motion signal waveform after median filtering is smoother, which eliminates waveform burrs and jitter caused by noise signals and at the same time corrects samples with large deviations for further processing.

3.4. Regional Activity Detection. When using virtual reality technology to analyze the behavior of athletes, areas of intense movement can provide more discriminative

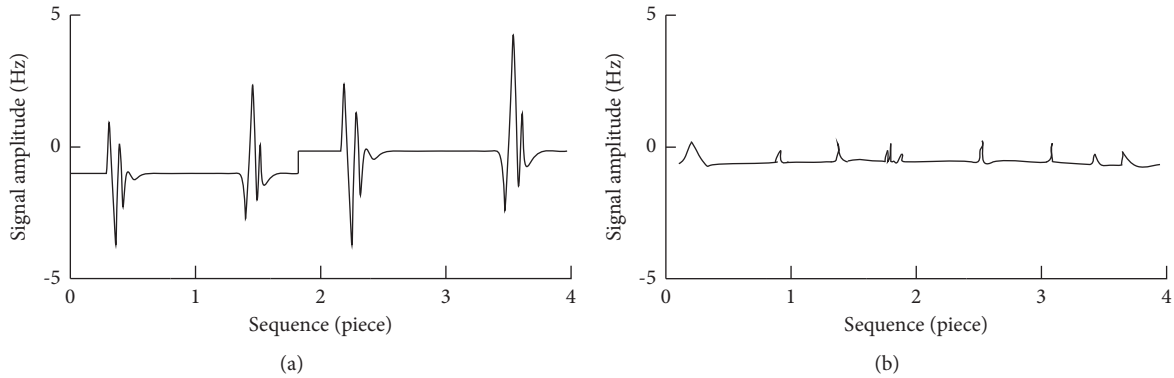


Figure 5: Comparison of motion signals before and after median filter processing.

information. Calculating the Lucas–Kanade optical flow characteristics of the area, as an energy function, can evaluate the degree of activity in the area. The calculation method is shown in the following formula:

$$F(k) = \left[1 - \frac{\left(f^t_{\max} - f^t_{\min}\right)}{f^t}\right] \exp\left(k_u k_v - \theta(k)\right). \tag{6}$$

Among them, $F(k)$ represents the optical flow vector of the kth pixel in the area; $\theta(k)$ represents the weight value of the pixel in the area, usually $\theta(k)$ takes a constant; t represents the number of pixels in the area; and f^t represents the magnitude of the optical flow vector value. The calculation method is as shown as follows:

$$f^t = f^t_{\text{best}} - r\left(f^t_{\max} - f^t_{\min}\right). \tag{7}$$

Among them, r represents the energy function in the area. The larger the value, the more active the area is.

3.5. Motion Information Capture Algorithm Based on Gaussian Mixture Model. The Gaussian mixture model refers to the use of Z Gaussian distributions to describe the feature distribution of multiple samples in a type of attitude capture result, that is, the weighted sum of Z Gaussian distribution functions to approximate the distribution function of each type of attitude observation [31, 32].

Suppose that the athlete's typical motion posture is ∂_z; after classification, it contains a total of M posture samples; namely, $S = \{s_1, s_2, \ldots, s_M\}$. For a single sample s_i, the density function of the Gaussian mixture distribution is

$$\mu(z) = \frac{1}{S_i} \sum_{i,j=1}^{M} f_{ij}(x, y). \tag{8}$$

Among them, f_{ij} represents the mixing coefficient, that is, the weight, which satisfies

$$\sum_{i,j=1}^{Z} f_{ij} = 1. \tag{9}$$

Among them, $Z = \{z_1, z_2, \ldots, z_n\}$. In this way, the athlete's motion posture can be described by a submodel containing Z Gaussian distribution functions.

In fact, the result of the activity detection in the previous area will affect the number of sub-Gaussian models and the mean and variance of each submodel. The number of Gaussian models can be estimated using Akaike information criterion and Bayes information criterion, or can be selected based on experience; that is, when the number of samples contained in a typical pose category is small (less than 20 samples), the value of Z is general choose 1 to 2; otherwise, when the number of samples is too small, the estimation of the mean and variance in the sub-Gaussian model will be out of generality [33, 34].

The calculation of the mean and variance in the sub-Gaussian model is actually a parameter estimation problem in a small sample space. There are many methods for parameter estimation. In comparison, EM is a regression analysis algorithm for incomplete data sets based on the principle of Maximum Likelihood Estimator (MLE). When the data is incomplete, the iterative algorithm for solving the maximum likelihood estimation of the parameters of the distribution density function, which greatly reduces the computational complexity of the maximum likelihood estimation, is conducive to improving the accuracy of the athlete's motion information capture [35, 36].

4. System Performance Test Experiment

In order to verify the effectiveness of the designed athlete space-time running information capture system based on virtual reality, the system designed in this paper is used to verify the capture effect of runner running information, and the human knee motion capture system based on MATLAB and the human motion trajectory tracking system based on three frame difference method are compared with the system designed in this paper.

4.1. Experimental Conditions Setting. According to the human skeleton model established above, it was finally decided to select 5 representative running actions for recognition, as shown in Table 3.

A capture experiment was performed on the movement information generated in the five running states shown in Table 3. In the experiment, MATLAB software was used for data processing to generate simulation experiment

Table 3: Running state classification table.

Serial number	Sports posture
1	Start
2	Upper limb swing
3	Leg swing
4	Stable trunk
5	Hit the line

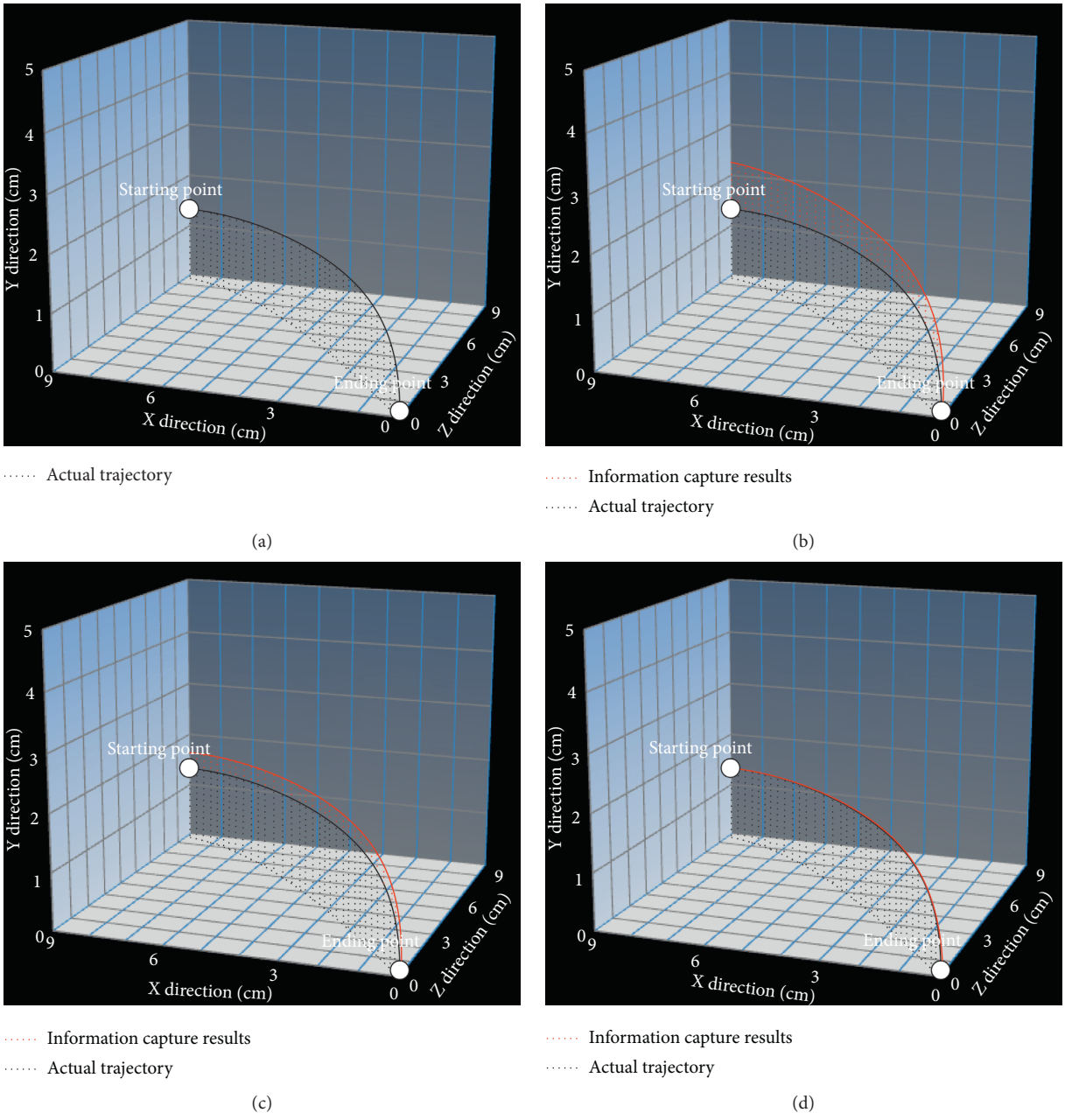


Figure 6: Comparison results of the accuracy of joint position information capture.

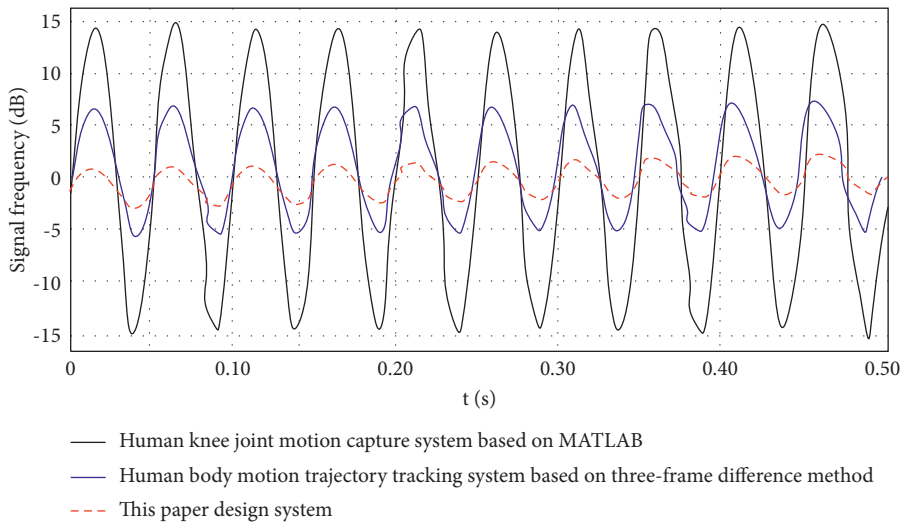


Figure 7: Comparison results of information capture signal components.

diagrams, and related conclusions were drawn based on the analysis of the experimental results.

4.2. Analysis of Experimental Results. Based on the above-mentioned experimental environment, a motion information capture experiment is carried out, and the application performance of the system is displayed through three indicators: the accuracy of joint position information capture, the effect of human motion signal processing, and the system processing time.

4.2.1. Accuracy of Capturing Joint Position Information. When the athlete is running, the ankle joint is taken as the target, and the accuracy of capturing ankle joint motion information by different systems is analyzed. The result is shown in Figure 6.

According to Figure 6, when three systems are used to capture the motion information of the ankle, the accuracy of information capture shows inconsistent characteristics. Among them, the information capture results of the human knee motion capture system based on MATLAB are obviously different from the actual motion trajectory of the ankle, indicating that the accuracy of information capture is low. Although the information capture effect of human knee motion capture system based on MATLAB is better than that of human knee motion capture system based on MATLAB, there is still a certain gap between its capture results and the system designed in this paper. The coincidence between the information capture results of the lower ankle joint and the actual motion trajectory of the system designed in this paper is high, which shows that the information capture accuracy of the system is high and can provide a reliable data basis for athletes' training. Because the system design in this paper filters the athlete's motion signal and uses the motion information capture algorithm based on the Gaussian mixture model to capture the information, the accuracy of the captured information is improved.

4.2.2. Human Motion Signal Processing Effect. Carrying out the comparison of human motion signal processing effects, the results of information capturing signal components of different systems are shown in Figure 7.

Analyzing Figure 7 shows that in the 0.5s capture time, the signal frequency of human knee joint motion capture system based on MATLAB is between 14 dB and 15 dB. The signal frequency of human body motion trajectory tracking system based on three-frame difference method is between 5 dB and 7 dB. The signal frequency of this paper design system is between 1 dB and 3 dB. When the system designed in this paper is used to capture athletes' running information in time and space, the signal output has good stability and a small range of fluctuations up and down, while the signal fluctuation range of the traditional system is large, indicating that the system designed in this paper has better anti-interference ability. This is because the system designed in this paper uses GFSK with strong anti-interference ability for modulation and demodulation in the hardware design, which not only saves energy, but also improves anti-interference.

4.2.3. System Processing Time. In order to further verify the effectiveness of the system designed in this paper, the application performance of different systems is further compared with the system processing time as an experimental indicator. The results are shown in Figure 8.

Analyzing Figure 8 shows that when different systems are used to capture motion information, the processing time of the system has shown a rapid increase in the initial stage of the experiment. As the experiment progresses, the area gradually stabilizes. The comparison shows that the processing time of the system designed in this paper is significantly lower than that of the two traditional systems, which shows that the processing time of the system designed in this paper is shorter and the processing efficiency is higher. It can realize real-time capture of athletes' running information in time and space. Because the system

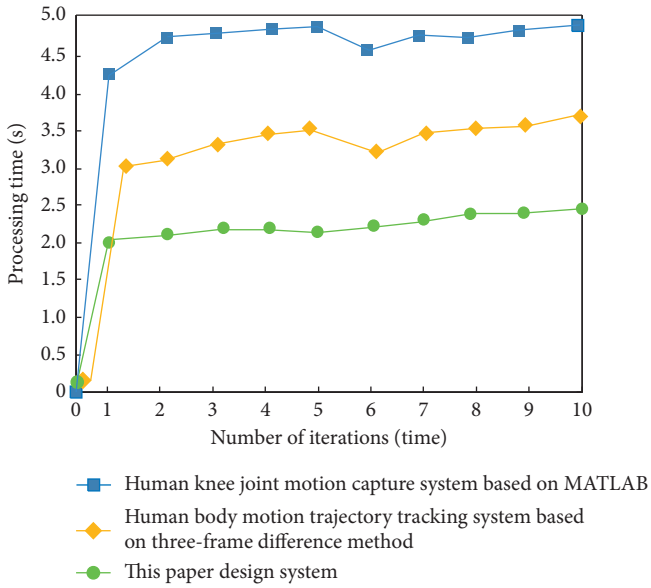


Figure 8: System processing time comparison results.

designed the movement data acquisition module and the data transmission module in this paper, the processing time can be shortened and the processing efficiency can be improved.

5. Conclusion

This paper designs an athlete's time-space running information capture system based on virtual reality. In the process of modern sports training, the information capture system based on virtual reality technology is innovative. It can help athletes better master relevant knowledge and technical essentials and greatly improve the training efficiency of athletes. The experimental results show that the system designed in this paper has better capture effect, can achieve accurate and real-time acquisition of athletes' sports information, and has better application effect. In future research, wireless transmission may be adopted. The sensor is attached to the human body, and the data is directly transmitted to the PC for motion tracking through the motion of the human body so as to realize the automatic collection of motion information. The fidelity of the human body model needs to be improved. The pure joints and bones lack a sense of image. In the future, real human models could be made, including skin and hair, which can track and simulate people's subtle movements and even smiles.

Data Availability

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding this work.

References

[1] M. Jiang, "Research on athlete training behavior based on improved support vector algorithm and target image detection," *Journal of Intelligent and Fuzzy Systems*, vol. 39, no. 4, pp. 5725–5736, 2020.

[2] Y. Wang, "Real-time collection method of athletes' abnormal training data based on machine learning," *Mobile Information Systems*, vol. 2021, no. 3, 11 pages, Article ID 9938605, 2021.

[3] C. Li and J. Cui, "Intelligent sports training system based on artificial intelligence and big data," *Mobile Information Systems*, vol. 2021, no. 1, 11 pages, Article ID 9929650, 2021.

[4] S. Zhao, "Research on scientific sports training of students majoring in physical education," *Revista Brasileira de Medicina do Esporte*, vol. 27, no. 5, pp. 460–463, 2021.

[5] Y. Liu and Y. Ji, "Target recognition of sport athletes based on deep learning and convolutional neural network," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 2, pp. 2253–2263, 2021.

[6] L. Tang, C. Zhu, and H. Luo, "Training prediction and athlete heart rate measurement based on multi-channel PPG signal and SVM algorithm," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 4, pp. 1–12, 2020.

[7] V. Isheyskiy, E. Martinyskin, S. Smirnov, A. Vasilyev, and K. Knyazev, "Specifics of MWD data collection and verification during formation of training datasets," *Minerals*, vol. 11, no. 8, p. 798, 2021.

[8] J. P. Wang, X. Zhao, H. Hu, L. Y. Wu, D. Guo, and Y. Li, "Motion capture measurement and analysis of human knee joint based on MATLAB," *Journal of Henan Polytechnic University(Natural Science)*, vol. 39, no. 3, pp. 86–93, 2020.

[9] X. J. Lin, "Design of human body underwater motion trajectory tracking system based on three-frame difference method," *Modern Electronics Technique*, vol. 42, no. 13, pp. 51–55, 2019.

[10] Q. Li, X. D. Wang, and H. Li, "3D human pose tracking approach based on double Kinect sensors," *Journal of System Simulation*, vol. 32, no. 8, pp. 1446–1454, 2020.

[11] J. Altet, E. Barajas, D. Mateo, A. Billong, and F. Reverter, "BPF-based thermal sensor circuit for on-chip testing of RF circuits," *Sensors*, vol. 21, no. 3, p. 805, 2021.

[12] H. Gevorgyan, A. Khilo, Y. Ehrlichman, and M. A. Popovi, "Triply resonant coupled-cavity electro-optic modulators for RF to optical signal conversion," *Optics Express*, vol. 28, no. 1, pp. 788–815, 2020.

[13] F. Adrion, M. Keller, G. B. Bozzolini, and C. Umstatter, "Setup, test and validation of a UHF RFID system for monitoring feeding behaviour of dairy cows," *Sensors*, vol. 20, no. 7035, p. 1, 2020.

[14] F. K. Matheus, P. M. Frank, F. B. Marcelo, and G. R. C. Thomas, "RFID wireless system for detection of water in the annulus of a flexible pipe," *Marine Structures*, vol. 72, no. 3, Article ID 102776, 2020.

[15] J. Flores, I. Garmendia, and I. C. Axpe, "Thermal monitoring and control by infrared camera in the manufacture of parts with laser metal deposition," *Dyna*, vol. 95, no. 1, pp. 360–364, 2020.

[16] C. Bowen, N. Reeve, T. Pettinger, and J. Gurnell, "An evaluation of thermal infrared cameras for surveying hedgehogs in parkland habitats," *Mammalia*, vol. 84, no. 4, pp. 354–356, 2020.

[17] A. Krtali, M. Baji, T. Ivelja, and I. Racetin, "The AIDSS module for data acquisition in crisis situations and environmental protection," *Sensors*, vol. 20, no. 5, p. 1267, 2020.

[18] F. Wen, H. Xiang, T. Zhang, Y. Wang, and X. Gao, "Upgrade of the data acquisition and control system of microwave

reflectometry on the experimental advanced superconducting tokamak," *IEEE Transactions on Nuclear ence*, vol. 66, no. 99, pp. 1340–1345, 2019.

[19] H. Aljuaid and S. A. Parah, "Secure patient data transfer using information embedding and hyperchaos," *Sensors*, vol. 21, no. 1, p. 282, 2021.

[20] M. Helwig, S. Zimmer, P. Lucas, A. Winkler, and N. Modler, "Multiphysics investigation of an ultrathin vehicular wireless power transfer module for electric vehicles," *Sustainability*, vol. 13, no. 17, p. 9785, 2021.

[21] T. Zhu and F. Zhang, "Design of marine two-way voice communication system based on human-computer interaction," *Journal of Coastal Research*, vol. 95, no. 1, p. 1389, 2020.

[22] Y. Jing, "Research on fuzzy English automatic recognition and human-computer interaction based on machine learning," *Journal of Intelligent and Fuzzy Systems*, vol. 39, no. 4, pp. 5809–5819, 2020.

[23] M. Tu, "Gesture detection and recognition based on pyramid frequency feature fusion module and multiscale Attention in human-computer interaction," *Mathematical Problems in Engineering*, vol. 2021, no. 7, 10 pages, Article ID 6043152, 2021.

[24] G. S. Heidner, P. M. Rider, J. C. Mizelle, C. M. O'Connell, and Z. J. Domire, "Anterior–posterior balance perturbation protocol using lifelike virtual reality environment," *Journal of Applied Biomechanics*, vol. 36, no. 4, pp. 1–5, 2020.

[25] M. Yamada, K. A. Nikita, J. A. Diekfuss, and L. D. Raisbeck, "The effect of attentional focus on movement accuracy in an immersive and interactive virtual reality environment," *Neuroscience Letters*, vol. 752, no. 1, Article ID 135814, 2021.

[26] S. Zhang, H. Yu, T. Wang, and J. Dong, "Augmented visual feature modeling for matching in low-visibility based on cycle-labeling of Superpixel Flow," *Knowledge-Based Systems*, vol. 195, no. 3, Article ID 105699, 2020.

[27] H. Hu, M. Cheng, F. Gao, Y. Sheng, and R. Zheng, "Driver's preview modeling based on visual characteristics through actual vehicle tests," *Sensors*, vol. 20, no. 21, p. 6237, 2020.

[28] X. Cheng, G. Li, R. Skulstad et al., "Data-driven uncertainty and sensitivity analysis for ship motion modeling in offshore operations," *Ocean Engineering*, vol. 179, no. 1, pp. 261–272, 2019.

[29] J. Y. Zhao and D. F. Zhang, "Simulation of human motion information capture in time-space domain based on virtual reality," *Computer Simulation*, vol. 38, no. 8, pp. 391–395, 2021.

[30] L. Li and Q. H. Zhuang, "Prediction and simulation of human behavior continuity based on time domain segmentation," *Computer Simulation*, vol. 38, no. 5, pp. 339–343, 2021.

[31] I. Shahin and A. B. Nassif, "Novel cascaded Gaussian mixture model-deep neural network classifier for speaker identification in emotional talking environments," *Neural Computing & Applications*, vol. 32, no. 7, pp. 2575–2587, 2020.

[32] C. Ananth and D. Brabin, "Enhancing segmentation approaches from Gaussian mixture model and expected maximization to super pixel division algorithm," *Sylwan*, vol. 164, no. 4, pp. 15–32, 2020.

[33] Z. Hui, S. Jin, D. Li, Y. Z. Yao, and B. Liu, "Individual tree extraction from terrestrial LiDAR point clouds based on transfer learning and Gaussian mixture model separation," *Remote Sensing*, vol. 13, no. 2, pp. 1–32, 2021.

[34] Q. Xu, S. Yuan, and T. Huang, "Multi-dimensional uniform initialization Gaussian mixture model for spar crack quantification under uncertainty," *Sensors*, vol. 21, no. 4, p. 1283, 2021.

[35] S. L. Yoo and J. C. Jeong, "Safe navigation distance between marine routes and aquaculture farms in South Korea using Gaussian mixture model," *Sensors*, vol. 20, no. 5, p. 1246, 2020.

[36] Y. Wang and M. Wen, "Simulation of tennis match scene classification algorithm based on adaptive Gaussian mixture model parameter estimation," *Complexity*, vol. 2021, no. 1, 12 pages, Article ID 3563077, 2021.

Hindawi
Scientific Programming
Volume 2022, Article ID 7905992, 10 pages
https://doi.org/10.1155/2022/7905992

Research Article

Vocal Music Recognition Based on Deep Convolution Neural Network

Zhuo He

Zhengzhou Normal University, Zhengzhou, Henan 450044, China

Correspondence should be addressed to Zhuo He; hezhuo@zznu.edu.cn

Received 24 November 2021; Revised 18 December 2021; Accepted 31 December 2021; Published 2 February 2022

Academic Editor: Baiyuan Ding

In order to achieve fast and accurate music technique recognition and enhancement for vocal music teaching, the paper proposed a music recognition method based on a combination of migration learning and CNN (convolutional neural network). Firstly, the most standard timbre vocal music is preprocessed by panning, flipping, rotating, and scaling and then manually classified by vocal technique features such as breathing method, articulation method, pronunciation method, and pitch region training. Then, based on the migration learning method, the weight parameters obtained from the convolutional model trained on the sound dataset CNN are migrated to the sound recognition, and the convolutional and pooling layers of the convolutional model are used as feature extraction layers, while the top layer is redesigned as a global average pooling layer and a Softmax output layer, and some of the convolutional layers are frozen during training. The experimental results show that the average test accuracy of the model is 86%, the training time is about 1/2 of the original model, and the model size is only 74.2 M. The F_1 values of the model are 0.88, 0.80, 0.83, and 0.85 in four aspects, such as breathing method, exhaling method, articulation method, and phonetic region training, etc. The experimental results show that the method is efficient for voice and vocal music teaching recognition. The experimental results show that the method is efficient, effective, and transferable for voice and vocal music teaching research.

1. Introduction

The concept of vocal music, also known as artistic singing, is a musical performance art that uses the combination of artistic language (singing dream) and scientific singing voice (artistic voice) to create a vivid and pleasant auditory image, singing voice, to express the highly condensed lyrics (poems or words) and typical, emotional melodic tones (good learning song) to learn vocal music, to express thoughts and feelings, and to create a second degree [1–3]. In short, vocal music is music with language sung and the human voice. Vocal music includes American singing, Gregorian chant, folk singing, and popular singing, as shown in Figure 1.

With the continuous development of economy, people's needs for material and spiritual aspects become more comprehensive and high level. With the continuous reform and development of education and the integration of various arts into people's daily life, people's appreciation level of vocal art has gradually increased, which has put forward higher requirements for the vocal art itself. High-quality vocal music appreciation and tasting are given better requirements. In order to improve the technical level of vocal teaching, this requires better technical development skills and error correction. The improvement of vocal technique is directly related to the content of vocal art. Therefore, the exploration of the status and role of vocal technique in vocal art can help the further development of vocal art by providing a better understanding of the current situation and future direction of vocal art. However, there are still many problems in vocal music teaching, such as low learning efficiency and ineffectiveness. Therefore, a new method needs to be found.

Convolutional neural network method is considered as a good learning and training method [4–10]. Deep learning theory was first proposed by Hinton et al. [11–14] as an effective method to simulate the sound learning process of brain recognition; it shows big advantages in vocal music processing and pattern recognition. For example, Janssens

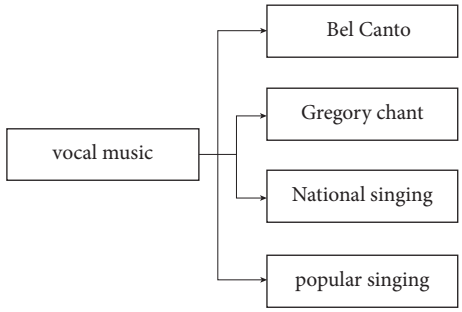


Figure 1: Vocal classification.

et al. [15, 16] proposed an end-to-end convolutional neural network based on the frequency domain representation of vibration signals to achieve fault classification of bearings with an accuracy of 93.6%. 97.58% accuracy was achieved by Qiao et al. using a convolutional neural network model to process SAR type images for classification. Qin et al. [17] gave a deep dual convolutional neural network in order to extract multidomain information of vibration signal and achieved 97.02% accuracy for engine misfire diagnosis. Zhang Kang et al. [18] proposed a deep convolutional neural network with random discard and batch normalization to accurately identify engine misfire faults based on the original cylinder head diagnostic signal. From the image processing and pattern recognition mentioned above, it can be seen that CNN models have powerful automatic feature extraction capability to perform deep feature extraction on signals [19–26] with stronger robustness and better generalization capability. Meantime, a number of parameters are greatly reduced by weight sharing and pooling operations, which reduces the training cost. In terms of pattern recognition of voice, most scholars use vibration signal for pattern recognition of voice, which has an impact on the online diagnosis of educated voice due to the high sampling rate of vibration signal and thus the relatively large scale of the constructed network. Considering that the frequency signal is easy to collect and stable, and the sampling rate is relatively low, the constructed network is easier to realize the online diagnosis. Therefore, this paper proposes the method of using CNN to automatically obtain the vocal features of the sound signal and then compare it with the standard music, write the algorithm into the STM32 microcontroller, and analyze it with the standard vocal music recorded in it to quickly give the vocal defects of the singer.

2. CNN Contrast Diagnosis Algorithm

A CNN is a feedforward neural network with convolutional operations. Compared to fully CNN of the same size, CNN has local connectivity, weight sharing, and downsampling [27] and therefore requires fewer parameters and memory for network training. The composition of neural network includes input layer, hidden layer, and output layer. The loss of forward propagation is transmitted forward by the input layer, and then the activation function is defined. On this basis, the class of the two-layer neural network is defined, the activation function is initialized, and the weights are used to realize the forward propagation of the core training logic of the neural network and calculate the loss and mutual propagation update. The final prediction process is a process of calculating the output value forward. The above is the core process of the whole calculation.

2.1. Two-Dimensional Convolution. Convolution is one of the core mathematical operations in CNN, mainly used to extracts more abstract feature from the data. Convolution layers convolve the local area of the input signal with the convolution kernel and then add the corresponding bias to the convolution output and perform a nonlinear transformation by the activation function to get the corresponding feature map. For two-dimensional linear non-shift systems, the output sequence $y(m, n)$ is equal to the convolution sum of the input sequence, $X(m, n)$, and the unit impulse response sequence $H(m, w)$. The one-dimensional wave signal is decomposed by wavelet packet to obtain multiple groups of wavelet coefficients, and then these wavelet coefficients are arranged into a two-dimensional matrix as the input of depth learning algorithm. The main calculation procedure is shown as follows:

$$a^l = f\left(\sum_{iem} x_i^{l-1} \otimes w_i^l + b_i^l\right), \tag{1}$$

where x_i^{l-1} is the element in the convolution region of the i ck (convolution kernel) of the l-1st layers; w_i^l is a weight of the i ck of the l layers; $\otimes$ is the convolution operation; b_i^l is the bias of the i-th ck of the l-th layers; M is the convolution region; and a^l is the output of the l-th layers convolution after the action of the activation function $f(-)$.

In order to avoid accelerate the learning speed of the network, the activation function adopted in this paper is Rectified Linear Unit (ReLU (x)), whose formula is shown as follows:

$$\text{relu}(x) = \max(0, x)\begin{Bmatrix} 0, & x \le 0 \\ x, & x > 0 \end{Bmatrix}. \tag{2}$$

x indicates the input value.

2.2. Maximum Pooling. The main purpose of max pooling is to compress the output features of the convolutional layer and extract the main features, thus reducing the sizes of the input vector and the parameters of the convolutional NN, reducing the training time and controlling overfitting. The maximum value of the upper left region is 9, the maximum element value of the upper right region is 2, the maximum value of the lower left region is 6, and the maximum value of the lower right region is 3. In order to calculate the values of the four elements on the right, we need to calculate the 2 of the input matrix × 2. Do the maximum operation in the area. It is like applying a filter with a scale of 2, because we chose 2 × Zone 2, stride 2; these are the hyperparameters for maximum pooling. Because the filter we use is 2 × 2, the final output is 9. Then move 2 steps to the right to calculate the

maximum value of 2. Then, in the second line, move down 2 steps to get the maximum value of 6. Finally, move 3 steps to the right to get the maximum value of 3. This is a 2×2 matrix, i.e., $f = 2$; step length is 2, i.e., $s = 2$. This is an intuitive understanding of the maximum pooling function. You can put this 4×4 region as regarded as the set of some features, that is, the set of inactive values of a certain layer in neural network. In this paper, we adopt the maximum pooling method as shown in Figure 2. K-maxpooling means that the original max pooling over time only takes the strongest value from a series of eigenvalues of the revolution layer, so our idea can be expanded. K-MAX pooling can take the value scored in the top-K of all eigenvalues and retain the original order of these eigenvalues (Figure 3 is the schematic diagram of 2-max pooling). That is, more feature information is reserved for subsequent stages. You can implement a minpooling layer and embed it into any network to see how the effect is. Computer science is based on practice. Why is there no minpooling? Suppose that the minpooling layer is used to replace the maxpooling layer. Generally, the maxpooling layer is an active layer before it, such as the ReLU layer, which will make the value of convolution characteristics greater than or equal to 0. If you use minpooling again, the resulting activation graph is the minimum value in the neighborhood of 0 or close to 0. When you go through minpooling several times, you will find that all activation values are 0, so your network cannot train, because any useful information is gone. Maxpooling is now commonly used to reduce the dimension of features and retain the maximum response of low-level features such as edges and textures after the first convolution layer and convolution block of convolution network. The pooling process can be expressed as shown in Figure 2.

2.3. Fully Connected Layer after Convolution and Pooling. Take vocal music for example. The size of a piece of sound is width × height × channel number (generally three-color channel). Assuming that the size of a group of sounds is n, if we use the traditional neural network to deal with this vocal music, the input layer needs n neurons, and if the fully connected structure is adopted, there will be many weight parameters, which is very difficult and time-consuming for the training of the network. Moreover, if the image is very complex, it is impossible to capture more advanced sound features by increasing the number of hidden layers, because neural networks with too many hidden layers may have problems in gradient back propagation, such as gradient explosion and gradient disappearance. Therefore, the traditional neural network is not suitable for dealing with vocal music tasks. In contrast, the convolution neural network adopts the design idea of local connection, weight sharing (that is, the convolution kernel is only connected with one window, and the convolution kernel can be shared by multiple windows), and pooling. The superposition of the three strategies greatly reduces many, many unnecessary weight parameters in the network, making network training easier. It should be pointed out that CNN using gap instead of FC usually has better prediction performance. The final classification of the network is achieved at network by the fully connected layer. The results of the fully connected layer are calculated as

$$z^l = f\left(w^l x^{l-1} + b^l\right). \tag{3}$$

In equation (3), w^l is a weight of the lth layers; x^{l-1} is the output value of the lth layers; b^l is the bias of the lth layers; $f(-)$ is a activation function, and a ReLU activation function is used; and z^l is the output value of the lth layers. The whole connection layers process can be expressed as follows.

2.4. Classification Evaluation and Loss Function for the Classification of Vocal Patterns. Scoring function is widely used in structure based computational aided drug design. It provides a theoretical basis for drug efficacy evaluation in drug research and development by quantitatively evaluating drug target interaction [1–5] and improves the efficiency of screening active compounds. Quantitative evaluation of the interaction between drugs and target proteins is usually divided into two steps. One step is docking process, which mainly refers to conformation search to find out potential binding poses; the other step is scoring process, which usually refers to scoring to predict drug target binding force. Most scoring functions are not approximated based on the complete physical model, so they often do not strictly follow the multi-body expansion theory, conservation law, symmetry invariance, etc. Even the expressions of some knowledge-based scoring functions do not contain physical meaning at all. In fact, as a tool applied in the scenario of high-throughput drug screening, most scoring functions focus on efficiency and pursue the balance between accuracy and efficiency by approximate means.

When using machine learning model to solve problems, there are two important concepts when it comes to model construction and model evaluation.

Loss function: most machine learning algorithms need to maximize or minimize a function, namely, "objective function." Generally, the minimization function is called "loss function."

Loss function is used in model construction (some simple model construction does not need loss function, such as KNN), so it is used to guide model generation. The damage function can be classified as follows (Figure 4).

Evaluation index: evaluate the machine learning algorithm model. In some problems, the loss function can be directly used as the evaluation index (for example, in regression problems, the mean square error (MSE) can be used not only to guide the model construction, but also to evaluate the model performance after the model is completed).

The evaluation index is used after the model is built, so it is used to evaluate the performance of the model. Common evaluation indicators of classification types can be expressed as follows.

Classification type common evaluation indicators: confusion matrix, accuracy, precision, recall, ROC-AUC, P-R curve. The Softmax evaluation function is used as the probabilistic output of the final classification layers of the CNN.

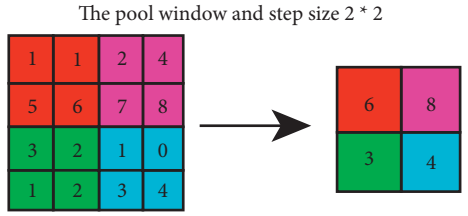


Figure 2: Schematic diagram of the largest set.

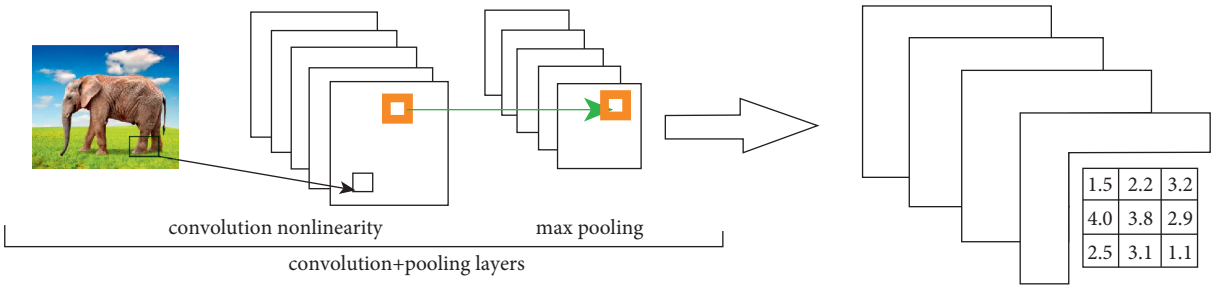


Figure 3: Process diagram of the whole connection layer.

$$p(z_j) = \text{softmax}(z_j) = \frac{e^{z_j}}{\sum_{k=1}^{N} e^{z_k}}. \tag{4}$$

In equation (4), z_j is the activation value of the jth neuron in the output layer; N is the number of vocal skill classifications; and $p(z_j)$ is the probabilistic output of each neuron.

The loss function used in this paper is the classification cross-entropy function, which is obtained by calculating the cross-entropy in the output vector after Softmax and the actual labels of the samples. The loss function is calculated as shown in the following:

$$\text{Loss} = -\sum_{i=1}^{N} y_i' \log(y_i). \tag{5}$$

In equation (5), y_i' is the i-th value of the actual labels; y_i is the i-th value of the output layers; and N is a number of vocal skill classifications. In the backpropagation stage, Adam algorithm is chosen in this paper to effectively update the weights and bias values of the network, and Adam algorithm uses first-order moment estimation and second-order moment estimation to dynamically adjust the learning rate of updating each parameter, so as to update the some weights to find the optimal solution [4, 28–30]. The standard deviation σ in data evaluation is shown as follows:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (x_i - x)^2}{n}}. \tag{6}$$

x is the standard value and x_i is the data sample.

The more common loss functions currently in use are shown as follows:

(1) 0-1 loss:

$$L(Y, f(X)) = \begin{cases} 1, & Y \neq f(X), \\ 0, & Y = f(X). \end{cases} \tag{7}$$

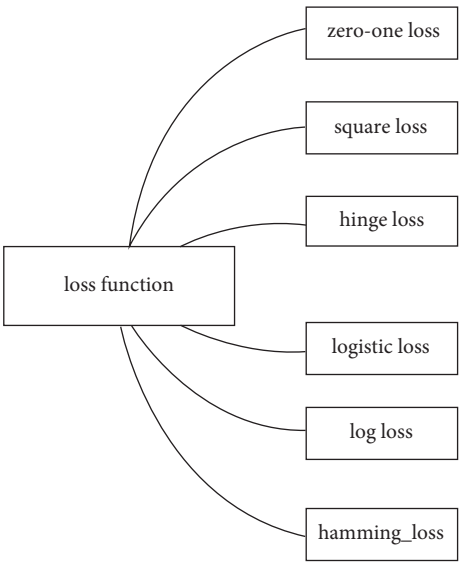


Figure 4: Classification of loss function.

0-1 loss means that the predicted value and the target value are not equal to 1; otherwise they are 0. 0-1 loss function directly corresponds to the number of classification errors, but it is a non-convex function, which is not very applicable.

(2) Squared loss:

$$L(Y, f(X)) = (Y - f(X))^2. \tag{8}$$

This represents frequent application and regression problems.

(3) Absolute loss:

$$L(Y, f(X)) = |Y - f(X)|. \tag{9}$$

It is very sensitive to outliers and noise. It is often used in AdaBoost algorithm.

(4) Logarithmic loss:

$$L(Y, P(Y|X)) = -\log P(Y|X). \tag{10}$$

(a) Log loss function can characterize the probability distribution very well, in many scenarios especially multi-categorization; if you need to know the confidence that the result belongs to each category, then it is very suitable.
(b) Not very robust, more sensitive to noise than hinge loss.
(c) The loss function of logistic regression is the log-log loss function.

(5) Hinge loss:

$$L(Y, f(X)) = \max(0, 1 - f(X))^2. \tag{11}$$

(a) Hinge loss function means that if it is correctly classified, the loss is 0; otherwise the loss is [Formula]. SVM uses this loss function.
(b) The general [Formula] is the predicted value. Between −1 and 1, [Formula] is the target value (−1 or 1). It means that the value of [Formula] is between −1 and +1. It does not encourage [Formula]; that is, it does not encourage the classifier to be overconfident. There will be no reward for making a correctly classified sample more than 1 from the division line, so that the classifier can focus more on the overall error.
(c) It is relatively robust and insensitive to outliers and noise, but it does not have a good probability interpretation.

(6) Loss function of LR:

$$L(Y, \pi(X)) = -Y \log \pi(X) - (1 - Y)\log(1 - \pi(X)). \tag{12}$$

2.5. Network Structure of the Convolutional Neural Network. CNN is a variant of multilayer perceptron (MLP). It was developed by biologists Huber and Wiesel's early research on cat visual cortex. The cells in the visual cortex have a complex structure. These cells are very sensitive to subareas of visual input space, which we call receptive fields, and tile the whole field of vision in this way. These cells can be divided into two basic types, simple cells and complex cells. Simple cells respond to the marginal stimulation pattern in the receptive field to the greatest extent. Complex cells have larger acceptance domains, which is locally invariant to stimuli from the exact location, from NN to convolutional neural network, as shown in Figure 5.

This tight relationship between interlayer connections and null domain information in CNNs makes them suitable for vocal processing and understanding. Moreover, they have also shown superior performance in automatically extracting salient features of the voice. In an example, GF (Gabor filters) have been used in initialization preprocessing step to achieve simulating the response of the visual system to visual stimuli. In most of the current work, researchers have applied CNNs to a variety of machine learning problems, including sound recognition, document analysis, and language detection. For the purpose of finding frame-to-frame coherence in sound, CNNs are currently trained by a temporal coherence, but this is not specific to CNNs.

In fact, according to the training results in the document, you may get a lot of such conclusions. For example, the neuron on the first floor is the simplest classifier. What it does is to detail whether there are green, yellow, and oblique stripes. The second layer is more complicated than this. According to the output of the first layer, it can see that the straight line and horizontal line are part of the window frame, the brown grain is wood grain, and the diagonal stripe + gray may be many things (part of the tire, etc.). According to the output of the second hidden layers, the third hidden layer will do more complex things. But the problem now is that when we directly use the fully connect feed forward network for image processing, we often need too many parameters. For example, suppose this is a $100 * 100$ color map (a small image), you pull this into a vector (how many pixels does it have), and it has $100 * 100\,3$ pixels. If it is a color graph, each pixel needs three values to describe it, that is, 30,000 dimensions. If the input vector is 30,000 dimensions, assuming that the hidden layer has 1000 neurons, the parameters of the hidden layer are $30{,}000 * 1000$, which is too much. So what CNN does is simplify the architecture of the neural network.

We know from human knowledge and from our images that some weights are not useful, and we filter them out at the beginning. Instead of using fully connect feed forward network, it uses relatively few parameters for image processing, so CNN is simpler than ordinary DNN.

After we finish our talk, we will find that you may think the operation of CNN is very complex, but in fact, its model is simpler than DNN. We use power knowledge to remove some parameters from the original fully connect layer and become CNN. Suppose that the input of our network is a $6 * 6$ image. If it is black and white, a pixel only needs a value to describe it. 1 means that ink is applied, and 0 means that ink is not applied. In the revolution layer, it consists of a group of filters (each filter is actually equivalent to a neuron in the fully connect layer), each filter is actually a matrix $(3 * 3)$, and the parameters in each filter (each element value in the matrix) are network parameters (these parameters need to be learned and not designed by people). If each filter detects $3 * 3$, it means that it detects another $3 * 3$ pattern (see a range of $3 * 3$). When detecting a pattern, you can decide whether a pattern appears by looking at only a range of $3 * 3$ without looking at the whole image. This is the first property we consider, Stripe = 1, which can ensure that every part of the image can be convoluted to $6 - 3 + 1 = 4$. The filter will tell you that the maximum value ($3 * 3$ matrix inner product) at the top left and bottom left represents that the pattern to be detected by the filter appears in the upper left corner and lower left corner of the image. This matter is

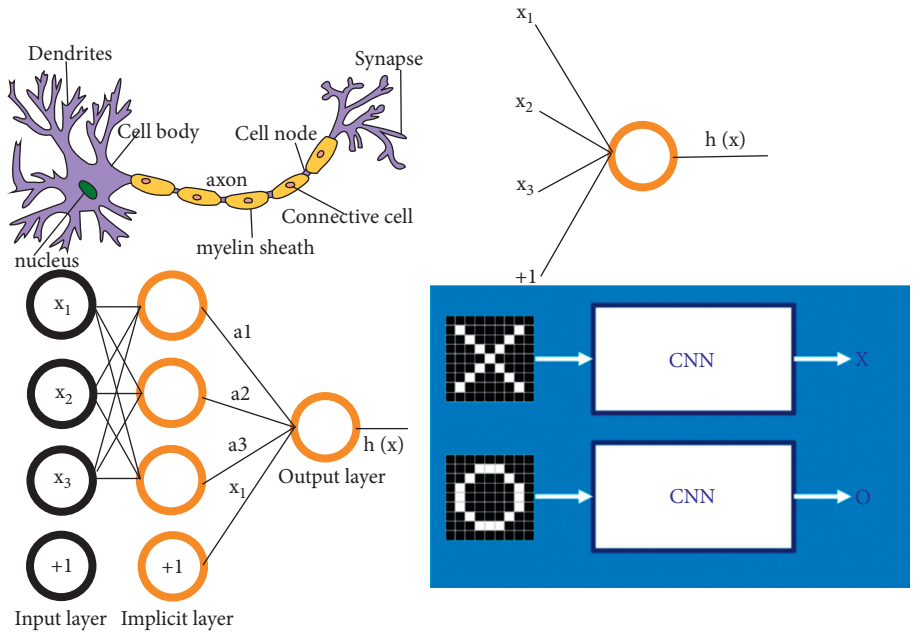


Figure 5: Neural network to convolution neural network process diagram.

considered as property 2. If the same pattern appears in the upper left corner and the lower left corner, we can use filter 1 to detect it. We do not need different filters to do this. Revolution is fully connected to remove some weight. Set the $6 * 6$ image flag as vector (36) and weight as filter (marked with different colors), such as 1, 2, 3, 7, 8, 9, 13, 14, 15 (9 each time, originally 36 each time). Different neurons share one weight (shared weight). Reducing weight and sharing weight can reduce parameter max pool down sampling pool: reduce the amount of calculation, but lose information. Here is a question: for the first time, there are 25 filters to get 25 feature maps. For the second, there are also 25 filters. Do you want to get a $25 * 25$ feature map? It is not like that! Suppose there are two filters in the first layer, and the filter in the second layer will consider the depth when considering this input. Instead of considering each channel separately, all channels are considered at once. Therefore, the output has as many filters as the revolution has (the revolution has 25 filters and the output has 25 filters; however, these 25 filters are a cube).

This greatly reduces the parameter scale of NN architecture. The designed CNN has 8 layers, including input layers, two sets of alternately connected convolutional and pooling layers, tiling layer, fully connected layers, and output layers, as shown in Figure 6.

The specific signal processing is shown in Figure 7.

The features are extracted through the convolution layer, and then the convolution results are mapped nonlinearly using the ReLU activations function. The pooling layers are then processed to eliminate the redundancy of information and reduce the number of model parameters. Finally, feature classification is performed by a fully connected layer and a Softmax output layer. The loss function used in the network is the classification, and the weights of the network are updated using the Adam optimizer. The structural parameters of the CNN model after iterative optimization are shown in Table 1.

Vocal voice recognition is often carried out after time-frequency analysis to obtain the speech spectrum. Among them, the speech sequence spectrum is characterized by a sequence of waves. In order to increase the effect of voice identification, it is necessary to overcome various characteristics of voice signal, including vocal breathing method, enunciation method, pronunciation method, and sound area training. CNN provides convolution in time and space. The idea of CNN is applied to the music modeling of speech recognition. Convolution can be used to overcome the diversity of speech signals. Signal is regarded as a wave and is widely used in deep convolution network recognition in sound. Vocal music and speech recognition are similar to the specific process in Figure 8.

The main algorithm flow is as follows:

(1) Standard vocal preprocessing. The collected high quality vocal music is preprocessed by panning, flipping, rotating, scaling, etc. to realize the expansion of the dataset. And the vocal music is uniformly adjusted to different frequencies and peaks.

(2) Input vocal music samples. Three vocal music samples are randomly selected from educators' vocal music styles as training samples input.

(3) Construct vocal music education recognition model. Based on the overall architecture of the pretrained VGG16 model, the original Softmax classification layer is replaced with a Softmax classifier with 6 neurons and the remaining fully connected layers are replaced in a global average pooling layer.

(4) Parameter migration and fine-tuning. Initialize the parameters of the pretrained VGG16 model by

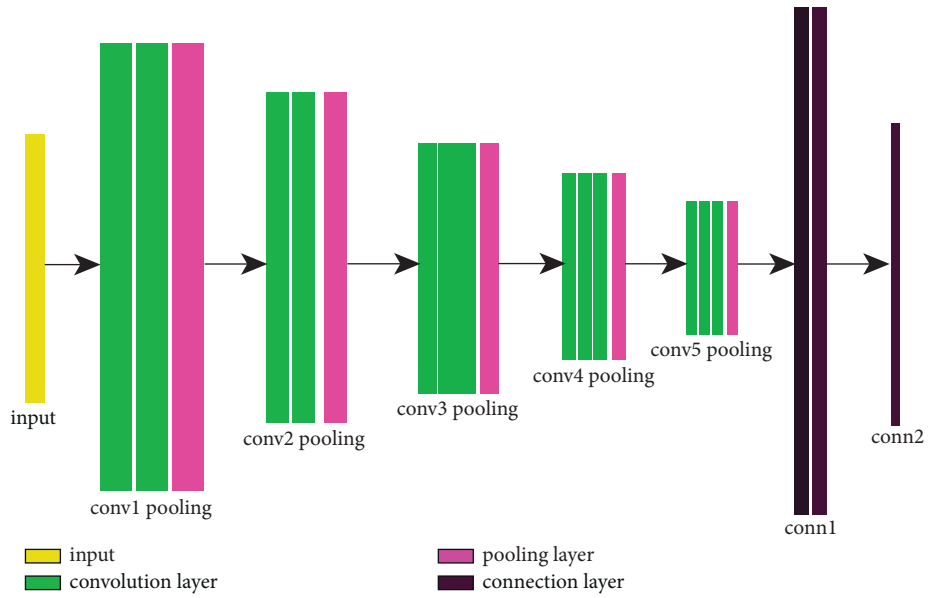


Figure 6: Convolutional neural network structure diagram.

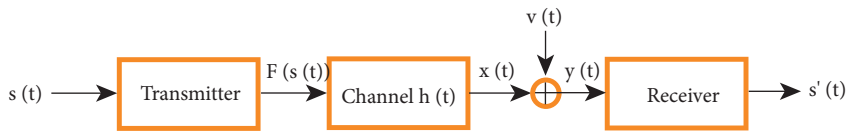


Figure 7: Spectrogram as observed signal frequency spectrum.

Table 1: Structural parameter of CNN.

Serial number	Neural network layer	Nuclear size	Number of nuclear	Step length	Activation function	Make-up zeroes	Output size
1	Input layer	—	—	—	—	—	[116, 8, 1]
2	Convolutional layer C1	6 * 2	16	1 * 1	ReLU	SAME	[116, 8, 1]
3	Pooling layer S1	3 * 2	16	3 * 2	—	SAME	[39, 4, 16]
4	Convolutional layer C2	6 * 2	32	1 * 1	ReLU	SAME	[39, 4, 36]
5	Pooling layer S2	2 * 2	32	2 * 2	—	SAME	[20, 2, 32]
6	Flat layer	—	—	—	—	—	1280
7	Fully connected layer	1	1	—	ReLU	—	60
8	Output layer	1	1	—	Sigmoid	—	7

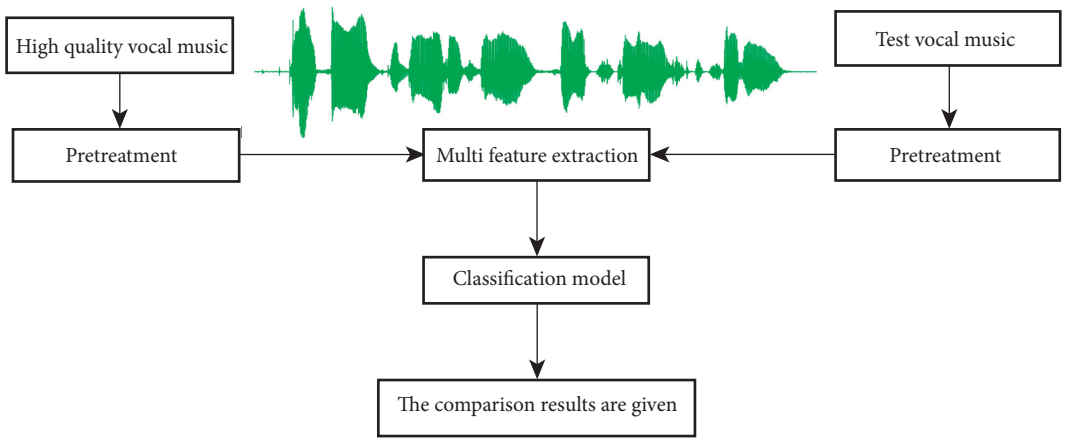


Figure 8: Convolutional neural network audio signal processing technique.

parameter migration, set the training parameters such as learning rate, momentum factor, and batch size, freeze the parameters of the pooling layer and some convolutional layers by updating the loss function iteratively, and update and optimize the parameters of some convolutional layers and Softmax classification layers.

(5) Testing of the model. The model is tested by taking performance pieces from the educated human dataset as test samples to verify the accuracies of the model.

3. Results and Analysis

There are two ways of migration learning: freezing all the convolutional layers and fine-tuning some of the network layers. Freezing all the convolutional layers means that the top layer of the model is allowed to participate in the training and all the convolutional layers are not involved. Freezing all the convolutional layers means that the top layer of the model is allowed to participate in the training, and none of the convolutional layers are involved. In this experiment, three scenarios are designed to analyze the effectiveness of migration learning by freezing all the convolutional layers, depending on the structure of the custom top layer: (1) dense layers with 64 neurons and Softmax output layers; (2) dense layers and Softmax output layers; (3) using the original top layer structure of VGG16; (4) pooling layers and Softmax output layers. The experimental results obtained for the three samples are shown in Table 2.

The relationship between the variation of the vocal value and the effective value at each time point during the experiment is shown in Figure 9.

It is obvious to see that different vocal test results have certain errors from the valid values, and different test samples can get the errors and losses from the valid values, so that the effect of each vocal sample can be analyzed and the problems in vocal music can be fed back, mainly through the four characteristics of vocal music (breathing method, articulation method, pronunciation method, and pitch range training) methods to improve. The graph below shows the errors values between the different characteristics and the valid standard samples, as shown in Figure 10.

The purpose of our vocal music teaching research is to quickly improve the vocal music level of the trainees. The difference between the trainees and the standard value or standard vocal music shown in Figure 10 will help to quickly make up for or improve the shortcomings in vocal music teaching. This shows that there are still great differences in our trainees. In ordinary education and teaching, we should further understand and train for this difference.

The stagger analysis of the training data is shown in Figure 11:

Through the errors in the training, we can determine the key points and special projects in the real vocal music training, which will help us make different training plans for each trainer and ensure the reasonable learning of vocal music technology, as shown in Figure 12.

Through the error change, we can find that the difference between the sample value and the standard value is changing

Table 2: Testing and accuracy.

Experiment number	Accuracy (%)	Test time (s)
1	78	30
2	90	30
3	86	30

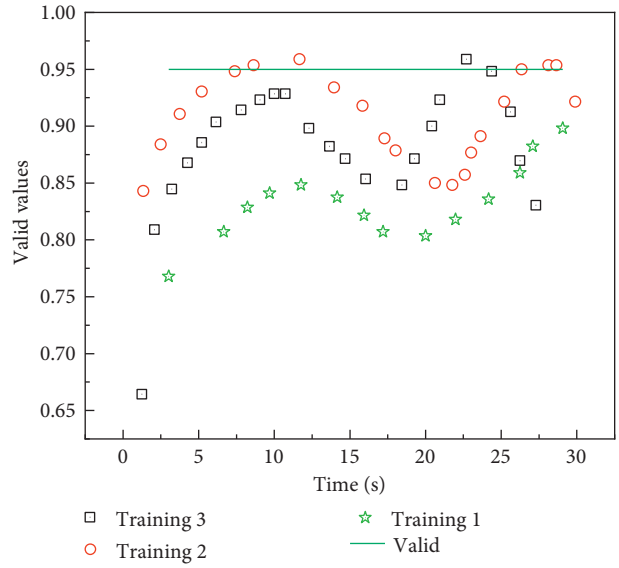


Figure 9: Standard errors for different test samples.

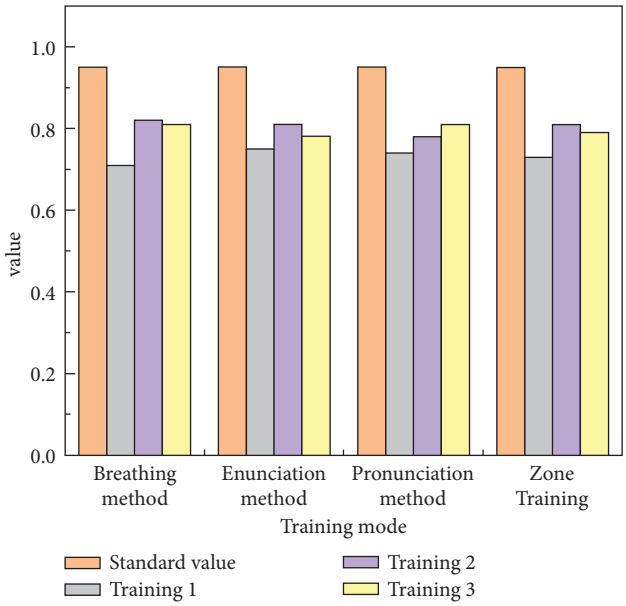


Figure 10: Error situation under different characteristics.

in the whole process, which is what we need to improve in the error learning. Only through the continuous error change to find the error can we continuously improve the ability and skills of vocal music learners in vocal music teaching, so as to fundamentally improve the level of vocal music teaching.

The comparison between the test sample and the standard sample can be evaluated for each segment and the gaps that exist

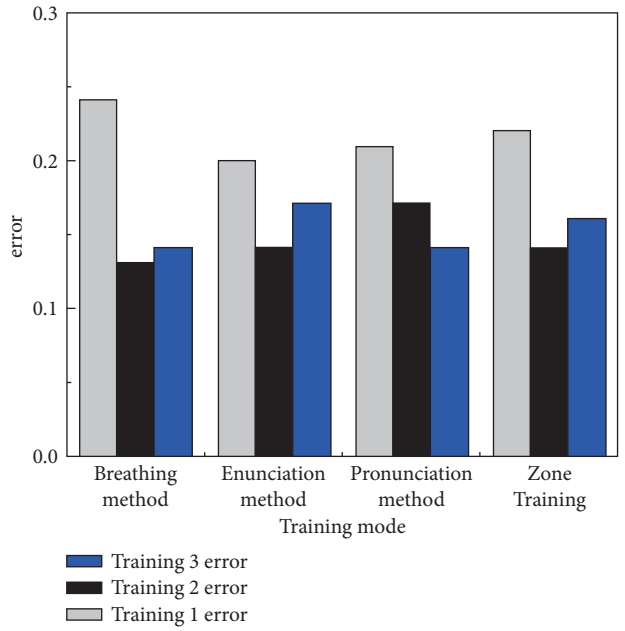


Figure 11: Error training chart.

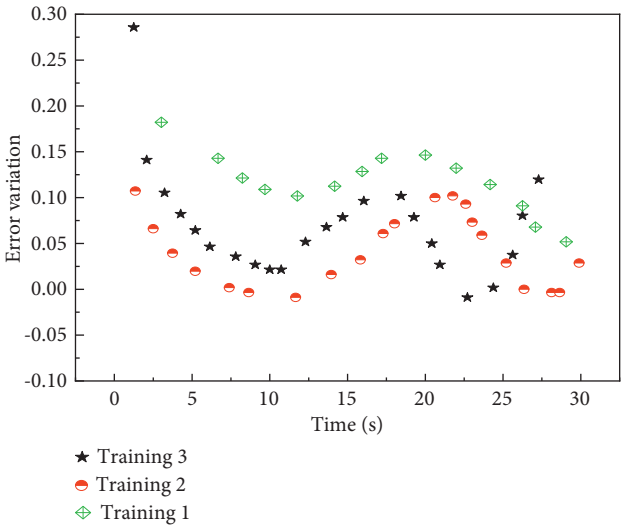


Figure 12: Error variation curve.

in the four characteristics of vocal music can be analyzed. This allows for timely targeted training tests for the gaps that exist in order to achieve rapid improvement, which is the application of the neural network method in vocal music teaching. It is also an important method for research in vocal music teaching.

In summary, this result shows that the standard sample convolutional neural network for uninvolved training samples for the control obtained the corresponding accuracy. And the accuracy of different characteristics was analyzed.

4. Conclusion

For the traditional recognition methods, there are problems such as poor recognition effect and low efficiency. In this paper, we introduced migration learning, redesigned the top layer of the model, compared the recognition effect of different sampling schemes and whether to freeze all convolutional layers during training, and obtained the following conclusions. In this study, the human voice features were analyzed and modeled by CNN method, and the accuracy and error under four features were predicted and combined with the development of the change curve of the error (0.1–0.3). The results show that the deep learning method can obtain better model results giving a quick evaluation with what is not known and giving relevant training suggestions for the trainer. It can quickly improve the vocal skills of learners and the vocal characteristics of vocal educators and helps learners to target their deficiencies with special training.

Data Availability

The dataset can be accessed upon request to the author.

Conflicts of Interest

The author declares no conflicts of interest.

References

[1] J. Sun, "Research on resource allocation of vocal music teaching system based on mobile edge computing," *Computer Communications*, vol. 160, no. 2, 2020.

[2] J. M. P. Wilbiks, D. T. Vuvan, P. Y. Girard, I. Peretz, and F. A. Russo, "Neurocase. Effects of vocal training in a musicophile with congenital amusia," *Neuron*, vol. 33, no. 2, pp. 1–191, 2020.

[3] Y. L. Huang, S. Y. Meng, X. S. Li, and W. Y. Fan, "A classification method for wood vibration signals of Chinese musical instruments based on GMM and SVM," *Traitement du Signal*, vol. 35, no. 2, pp. 137–151, 2018.

[4] M. Pei, X. Wu, Y. Guo, and H. Fujita, "Small bowel motility assessment based on fully convolutional networks and long short-term memory," *Knowledge-Based Systems*, vol. 121, no. 1, pp. 163–172, 2017.

[5] J. Kawahara, C. J. Brown, S. P. Miller et al., "BrainNetCNN: convolutional neural networks for brain networks; towards predicting neurodevelopment," *NeuroImage*, vol. 146, pp. 1038–1049, 2017.

[6] M. Ghafoorian, N. Karssemeijer, T. Heskes et al., "Location sensitive deep convolutional neural networks for segmentation of white matter hyperintensities," *Scientific Reports*, vol. 7, no. 1, 2017.

[7] M. T. Mccann, K. H. Jin, and M. Unser, "Convolutional neural networks for inverse problems in imaging: a review," *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 85–95, 2017.

[8] D. S. Tan, W. Y. Chen, and H. Kai-Lung, "DeepDemosaicking: Adaptive image demosaicking via multiple deep fully convolutional networks," *IEEE Transactions on Image Processing*, vol. 25, no. 5, 2018.

[9] Z. J. Wu and S. M. Naik, "DSP applications in engine control and onboard diagnostics: enabling greener automobiles," *IEEE Signal Processing Magazine*, vol. 34, no. 2, pp. 70–81, 2017.

[10] K. Lee, S. E. Kim, J. Doh, K. Kima, and W. K. Chung, "User-friendly image-activated microfluidic cell sorting technique using an optimized, fast deep learning algorithm," *Lab on a Chip*, vol. 21, no. 9, 2021.

[11] A. M. Stoica, T. V. Chelaru, F. Stoican, and B. D. Ciubotaru, "A Kalman filtering approach for systems subject to parametric modeling uncertainties," *IFAC-PapersOnLine*, vol. 52, no. 12, pp. 400–404, 2019.

[12] C. Gu, X.-Y. Qiao, H. Li, and Y. Jin, "Misfire fault diagnosis method for diesel engine based on MEMD and dispersion entropy," *Shock and Vibration*, vol. 2021, Article ID 9213697, 14 pages, 2021.

[13] S. Sneha, S. Potala, and A. R. Mohanty, "An improved method of detecting engine misfire by sound quality metrics of radiated sound," *Proceedings of the Institution of Mechanical Engineers-Part D: Journal of Automobile Engineering*, vol. 233, no. 12, pp. 3112–3124, 2019.

[14] X. Wang, Z. Li, H. Shan, Z. Tian, Y. Ren, and W. Zhou, "FastDerainNet: a deep learning algorithm for single image deraining," *IEEE Access*, vol. 8, p. 99, 2020.

[15] O. Janssens, V. Slavkovikj, B. Vervisch et al., "Convolutional neural network based fault detection for rotating machinery," *Journal of Sound and Vibration*, vol. 377, pp. 331–345, 2016.

[16] X. Ding and Q. He, "Energy-fluctuated multiscale feature learning with deep ConvNet for intelligent spindle bearing fault diagnosis," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 8, pp. 1926–1935, 2017.

[17] C. Qin, Y. Jin, J. Tao et al., "DTCNNMI: a deep twin convolutional neural networks with multi-domain inputs for strongly noisy diesel engine misfire detection," *Measurement*, vol. 180, Article ID 109548, 2021.

[18] K. Zhang, T. Jian-feng, C. Qin, W. Li, and C. Liu, "Diesel engine misfire diagnosis with deep convolutional neural network using dropout and batch normalization," *Journal of Xi'an Jiaotong University*, vol. 53, no. 8, pp. 159–166, 2019.

[19] J. Feng, Y. Yao, S. Lu, and Y. Liu, "Domain knowledge-based deep-broad learning framework for fault diagnosis," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 4, pp. 3454–3464, 2021.

[20] P. L. Wu, X. Y. Nie, and G. Xie, "Multi-sensor signal fusion for a compound fault diagnosis method with strong generalization and noise-tolerant performance," *Measurement Science and Technology*, vol. 32, no. 3, pp. 1–16, 2021.

[21] M. Long, J. Wang, Y. Cao, J. Sun, and P. S. Yu, "Deep learning of transferable representation for scalable domain adaptation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 8, pp. 2027–2040, 2016.

[22] C Huizhu, X. L. Liu, and J. A. Qiu, "Citespace based comprehensive analysis on debris flow risk of China during recent 30 years," *Journal of Engineering Geology*, vol. 26, no. 2, pp. 286–295, 2018.

[23] Z. G. Deng, H. D. Guan, H. John, M. A. Forster, Y. Wang, and C. T. Simmons, "A vegetation-focused soil-plant atmospheric continuum model to study hydrodynamic soil-plant water relations," *Water Resources Research*, vol. 53, no. 6, 2017.

[24] R. Duan, Y. Dong, P. Zhou, L. Wang, F. U. Yunmei, and S. Zhao, "Advances in application of hyperspectral remote sensing in hydrogeology," *Hydrogeology and Engineering Geology*, vol. 44, no. 4, pp. 23–29, 2017.

[25] Y. Zhang, S. Liu, C. Dong, X. Zhang, and Y. Yuan, "Multiple cycle-in-cycle generative adversarial networks for unsupervised image super-resolution," *IEEE Transactions on Image Processing*, vol. 29, p. 99, 2019.

[26] N. H. Son and M. S. Szczuka, "Neural networks design: rough set approach to continuous data," *Lecture Notes in Computer Science*, vol. 1263, pp. 359–366, 2019.

[27] J. Li, X. M. Wang, Y. H. Zhang, W. D. Wang, and S. J. Gai, "Research on the seismic phase picking method based on the deep convolutional neural network," *Chinese Journal of Geophysics*, vol. 63, no. 4, pp. 1591–1606, 2020.

[28] S. N. Pradhan, M. Anjum, and P. Jena, "Estimation of soil moisture content by remote sensing methods: a review," *Journal of Pharmacognosy and Phytochemistry*, vol. 7, pp. 1786–1792, 2018.

[29] P. H. Ma, J. B. Peng, Q. Y. Wang, X. Zhu, Q. Dong, and D. Zhai, "Formation mechanism, deposits and motion characteristics of the typical loess landslide in South Jingyang Platform," *Journal of Engineering Geology*, vol. 26, no. 4, pp. 930–938, 2018.

[30] S. J. Mi, Q. B. Wu, and B. Sheng, "Research on estimating soil Moisture by using thermal inertia," *Geomatics and Spatial Information Technology*, vol. 42, no. 10, pp. 11–14+18, 2019.

Hindawi
Scientific Programming
Volume 2022, Article ID 1011070, 10 pages
https://doi.org/10.1155/2022/1011070

Research Article

Research on Demand Forecasting of Engineering Positions Based on Fusion of Multisource and Heterogeneous Data

Ning Li,[1] Tianqi Wang ,[2] and Qianhui Zhang[3]

[1]*School of Economics and Management, Pingdingshan University, Pingdingshan 467000, Henan, China*
[2]*School of Logistics, Linyi University, Linyi 276000, Shandong, China*
[3]*Pingdingshan University, Pingdingshan 467000, Henan, China*

Correspondence should be addressed to Tianqi Wang; wangtianqi@lyu.edu.cn

Received 9 December 2021; Revised 2 January 2022; Accepted 16 January 2022; Published 2 February 2022

Academic Editor: Baiyuan Ding

Aiming at the project demand forecasting problem based on multisource data fusion, a multisource heterogeneous data fusion model is established, and the unified quantitative representation method of heterogeneous data based on triangular fuzzy numbers is studied, and the ordered weighted average operator is used to integrate the preferences of decision-makers. A multisource heterogeneous data fusion algorithm that supports multiuser decision-making is designed. Based on the analysis of the internal and external environment of human resources in a company's engineering positions, this paper qualitatively analyzes and selects the factors affecting the demand for talents in key positions in a company based on the characteristics of demand influencing factors and finds out the quantifiable and influential factors from the representative factors of talent demand for key positions in a company. Using historical data, statistical methods are used to process the eight related factors of a certain company, which confirms the factors that have a greater impact on the demand for talents in key positions of a company and influences the demand for talents in key positions in companies of the same type. The identification of factors provides a basic argument for a company. According to the results of statistical analysis and the characteristics of existing data, two variables of factory output and time are selected to be used in regression analysis forecasting model and gray system forecasting model of a certain company to predict the demand for key talents of a certain company. The company finally adopts combined forecasting. The method determines the predicted value of the talent demand for a certain company's key positions. According to the results of demand forecasting and the current status of human resource management in a company, this article proposes a company's key position talent management planning measures, in order to provide a reference for the management of a company's key company position talents and ensure a company's key company positions in the future talent demand reserve.

1. Introduction

In the process of enterprise informatization construction, due to the phased, technical, and other economic and human factors of the construction of various business systems and the implementation of data management systems, enterprises have accumulated a large number of business data in different storage methods during the development process. Data fusion is the multilevel, multifaceted, multilevel processing and combination of multiple sets of sensor data obtained from the same target to generate new meaningful information. The sensor here is in a broad sense, generally referring to the relevant databases of various data acquisition systems. Data fusion is a processing method of multisource information. Simply put, data fusion is a comprehensive algorithm of multiple data. The purpose of processing is to reason and identify the obtained information and make estimates and judgments accordingly. By fusing multisensor data, confidence can be increased, ambiguity can be reduced, and system reliability can be improved [1–7].

Data fusion originated in the 1970s, initially out of military needs, and rapidly expanded to areas such as automatic control, medicine, intelligent buildings, and commerce in research. The analysis objects are also extended from physical goals to information goals and even cognitive goals. The theoretical basis of data fusion is information

theory, detection and estimation theory, statistical signal theory, fuzzy mathematics, cognitive engineering, systems engineering, and so on. In 1986, Joint Directors of Laboratories established a function-oriented basic model and basic terminology, and in 1998, it was further improved. Although the JDL model is based on military proposals, it is also suitable for other application fields. The JDL model does not involve system structure. Bowman et al. extended this and proposed the concept of hierarchical data fusion tree, dividing the fusion problem into nodes [8–16]. Each node conceptually includes functions, such as data association, correlation, and evaluation. On this basis, Boss 6 colors further developed and proposed a set of modeling and simulation methods to realize the design of data fusion system, as can be seen in Figure 1, which mainly explains the whole system of data fusion system and its content.

Data fusion is essentially an integrated process of using computers to process, control, and make decisions on various information sources. The functions of the data fusion system mainly include detection, correlation, identification, and estimation. Data fusion can be divided into five levels: detection-level fusion, location-level fusion, attribute-(target recognition-) level fusion, situation assessment, and threat estimation. Related to this article is mainly the attribute -level fusion, also known as target-recognition-level fusion. It refers to the combination of target recognition data from multiple sensors to obtain a joint estimation of target identity. Attribute-level fusion uses multiple sensors to collect the data of the observed target, performs feature extraction and data combination, grouping the same target, and then using the content algorithm to synthesize the grouped data of the same target, and finally gets the joint attribute judgment of the target. That is, the type and category of the target are obtained. According to different fusion locations, attribute-level fusion is divided into three methods: decision-level fusion, feature-level fusion, and data-level fusion [17–20].

So far, researchers have proposed more than 30 data fusion models; the most cited is the JDL model of the US Department of Defense. Many mature applications of these models have appeared in the fields of target tracking, image fusion, and so on, but there are relatively few applications in data mining and natural language processing. Data fusion technology is an emerging interdisciplinary comprehensive theory and method. After decades of development, breakthrough progress has been made, but there are still many problems. For example, there is no unified definition, lack of systematic and complete basic theory, and so on. To sum up, in the face of the emerging scientific theory and method of data fusion, it is necessary to conduct in-depth systematic research on the existing data fusion technology and find its fit with the field of natural language processing from both theoretical and practical levels.

There are documents that have studied multisensor data fusion technology based on statistics and artificial intelligence methods and others that have studied the organization and management of multisource heterogeneous data in mobile geographic information systems and have established multisource heterogeneous data fusion models. The

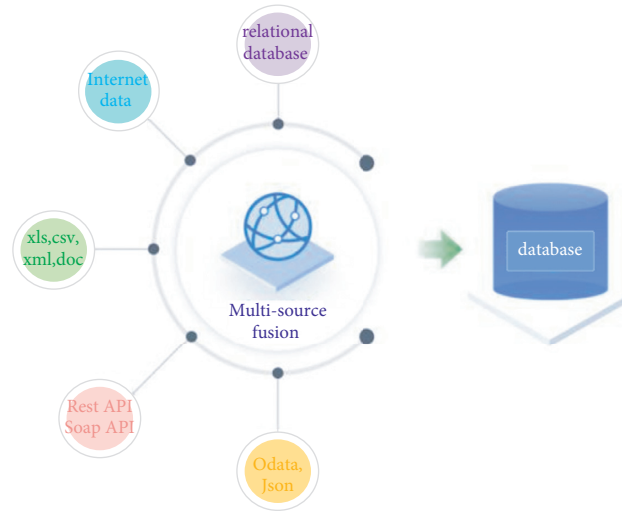


Figure 1: Data fusion system.

combination of sensor network and data fusion technology proposes a Kalman filter batch estimation fusion algorithm; some literature has studied the fusion method of massive multisource heterogeneous data in the Internet of Things environment and has been successfully applied in the process of target positioning and tracking. The literature has studied the intelligent maintenance decision-making architecture of the high-speed rail signal system based on heterogeneous data fusion, which has improved the accuracy and effectiveness of decision-making. Someone has studied the multisource heterogeneous data fusion technology in the construction of digital mines, ensuring that the construction of digital mines is in progress. The basic information platform is safe, stable, and efficient [21–25].

The type and structure of the fusion data is limited, and most work only incorporates an additional type of auxiliary information, which has domain limitations and sometimes relies on rules and expert knowledge and high labor costs. The more sufficient the auxiliary data is, the more comprehensive the hidden representation of users and items can be obtained. In the recommendation prediction, the rich feature relationship between the two can be integrated to obtain more accurate results. The structural dimensions of different data sources are also different, and the distribution of data is also very different, which also increases the difficulty of integrating more data in breadth. Including the adopted data management system is also very different; from simple file database to complex network database, they constitute the heterogeneous data source of the enterprise.

The fusion method is relatively preliminary, and the barriers between heterogeneous data have not been broken. In a related research, linear transformation is performed in the hidden semantic space of multisource data, and the method of adding and multiplying is used to integrate into the recommendation model, and it is impossible to fit the relationship between complex multisource heterogeneous data. To make matters worse, other auxiliary data may contain other information that has nothing to do with the recommendation. The mechanized fusion will introduce

unnecessary noise, which will reduce the accuracy of the recommendation model. In addition, multisource data only supplements the characteristics of the recommendation model itself, and the feature information between each multisource heterogeneous data source also lacks in-depth interaction and cannot achieve synergistic effects [26–29].

In addition to numerical data, there are other description forms such as language or symbols. Various descriptions lead to the ambiguity, difference, and heterogeneity of the structure and semantics of the data information. On the other hand, the decision-making process needs to comprehensively consider various heterogeneous data and information and make final decisions through the fusion of data and information. Therefore, starting from the characteristics of heterogeneous data, this paper studies a multisource heterogeneous data fusion method that supports multiuser decision-making.

Human resource forecasting refers to the assumption of the human resource situation in a certain period of time in the future based on the evaluation and prediction of the enterprise, as shown in Figure 2. It mainly includes the forecast of the quantity and type of human resource demand for the future development of the enterprise; the forecast of the future human resource status of the enterprise; the forecast of the future industry competition situation; the forecast of the supply and demand relationship of social human resources. Experience forecasting method, current situation planning method, model method, expert discussion method, quota method, and top-down method are commonly used methods for human resource forecasting. At present, the research on the special field of enterprise human resource forecasting needs to be in-depth. In practical applications, it is necessary to further improve the accuracy and feasibility of the forecast and at the same time increase the forecast of ability and quality based on the forecast of the number of personnel. In addition, there are few applied researches on human resource forecasting methods in specific enterprises.

2. Multisource Heterogeneous Data Fusion Model

Human resource forecasting can be divided into human resource demand forecasting and human resource supply forecasting, including the dual meanings of foreseeing and measuring the future. When studying the status of a large-scale system, the status of each part of the system is usually judged first and then integrated to comprehensively judge the overall status. Therefore, a fusion method is needed to fuse the data of each part. Subsequently, multisource data fusion technology emerged, which can associate and combine data from multiple sensors, and integrate them together for a unified evaluation. According to the characteristics of the fusion algorithm, it can be summarized as data-level fusion, feature-level fusion, and decision-level fusion. Data-level fusion directly integrates the original log information obtained by the detector, and then, the fused data is processed in the next step. In this way, many details of the original data are retained, the amount of information lost is relatively small,

Figure 2: Human resource.

and the granularity of fusion is relatively high. However, this fusion method is easily affected by the original data. When the original data is incomplete or the data stability is poor, it will directly affect the effect of the fusion, and the fused data must be homogeneous data. In addition, because many details in the original data are retained, the amount of calculation is relatively large and the processing cost is relatively high, which is not suitable for real-time fusion. This method has many applications in the field of image processing. Different from data-level fusion, feature-level fusion first performs data preprocessing and feature extraction on the data obtained by each detector and removes attributes that are weakly or irrelevant to the researched problem and then performs data extraction on the extracted feature data. Compared with the data-level fusion method, because the data fused by this fusion method is the data after feature extraction, the amount of data is small, the processing cost is low, the anti-interference ability is strong, the real-time performance is better, and heterogeneous data can be fused. Decision-level fusion first extracts the features of the original log information of each detector and analyzes and models it and then uses the single-source decision output from the model as a fusion factor to fuse, and the result of the fusion is the decision result of comprehensive multisource information. Compared with the other two fusion methods, this method has the smallest amount of data, so it has the best real-time performance and the lowest computational cost. And when the original data is unstable, the impact on fusion is minimal, and heterogeneous data can be fused. Enterprise human resources forecasting is a series of studies on the development trend, prospects, various possibilities, and consequences of enterprise human resources.

2.1. Multisource Heterogeneous Data Fusion Method. Data fusion is essentially the collaborative processing of data from multiple parties to achieve the purpose of reducing

redundancy, comprehensive complementation, and capturing collaborative information. Data fusion is divided into data-level fusion, feature-level fusion, and decision-level fusion according to operation level. This paper studies the fusion of multiple data sources at the decision-making level, and its methods mainly include weighted average method, D-S evidence theory, and voting.

2.1.1. Weighted Average Method. Calculate the support value of each data source for decision-making, wi is the weight of data source i, and t_{ij} is the support of data source i to the j-th decision. This method judges the pros and cons of decision-making schemes according to the degree of support, which is easy to operate, and considers the importance of the data source and other characteristics, but the determination of the weight contains subjective factors.

2.1.2. D-S Evidence Theory. The space formed by all possible results of the object to be recognized is defined as the recognition frame D, and its subset is marked as $2D$, and the definition is as follows:

$$m: 2^D \longrightarrow [0,1], \tag{1}$$

where

$$\begin{aligned} m(\Phi) &= 0, \\ \sum_{A \subseteq 2^D} m(A) &= 1. \end{aligned} \tag{2}$$

φ is the quasi-empty set, and then, m is the basic probability allocation function (BPAF) on 2^D, which actually assigns the trust degree of the subset of D based on the evidence.

In practice, different mi is often obtained for the same problem due to different evidences. After considering all the evidence, m can be obtained by the following formula:

$$m(A) = K^{-1} \sum_{\cap A_i = A} \prod m_i(A_i) (1 \le i \le n), \tag{3}$$

where

$$K = \sum_{\cap A_i \neq \phi} \prod m_i(A_i). \tag{4}$$

The D-S evidence theory is based on BPAF and can deal with the uncertainty caused by "not knowing." The disadvantage is that the elements in D must meet the mutually exclusive condition, and the calculation is complicated when there are too many BPAFs.

2.1.3. Voting Method. Consider each data source as a voter, and determine the pros and cons by comparing the number of votes obtained by each decision. The calculation method is

$$\mathrm{Sup}(a_i) = F\left[\mathrm{Sup}_j(a_i)\right]. \tag{5}$$

Among them, a_i is the i-th decision, and $\mathrm{Sup}(a_i)$ is the "number of votes"; $\mathrm{Sup}_j(a_i)$ is the support of the j-th data source for a_i. If it supports it, it is 1; otherwise, it is 0, and the function F can be defined as continuous add and sum. It is difficult to determine the BPAF for multisource heterogeneous data. The voting method cannot distinguish decisions with the same number of votes. Taking the preferences of decision makers into consideration, the OWA method is used to fuse the data in this article. The error is compared in Figure 3.

2.2. Multisource Heterogeneous Data Fusion Structure. The fusion structure of multiple data sources is shown in Figure 4. The data fusion process takes into account the characteristic factors expressing user needs and the reliability of information, uses context knowledge and domain knowledge, and uses voting to resolve data conflicts and other issues.

Aiming at the previously mentioned model, this paper designs a multisource heterogeneous data fusion structure model that supports multiuser decision-making. The data fusion engine in the model includes four modules: data warehouse, decision support calculation, OWA operator weight vector calculation, and data conversion and sorting. The specific descriptions are as follows. (1) The data warehouse implements data selection, feature extraction, and statistics operations: data integration, elimination of data heterogeneity and differences, and providing data sources for subsequent data processing. (2) The decision support calculation module obtains data of relevant dimensions from the data warehouse according to the decision attributes and calculates the impact of each data source on the decision: the support value s_{ij} (the support degree of the data source i for the j-th decision). (3) The OWA operator weight vector calculation module calculates the OWA weight wi according to the fuzzy semantic principle provided by the decision maker. The choice of fuzzy semantic parameters reflects the decision maker's: the preference attitude of the data source. (4) Data are converted and sorted according to the credibility or importance of the data source provided by the decision maker, combined with the OWA weight vector wi to convert sij and sort the converted results in order of size, and sort the result, which is calculated by summing the final decision value.

3. Multisource Heterogeneous Data Fusion Algorithm

3.1. Data Types and Their Characteristics. This technology has become a research hotspot in the fields of data processing, target recognition, situation assessment, and intelligent decision-making. Data can be described in terms of quantity and quality. The quantity is represented by numerical values, and the quality is described by linguistic variables. According to the different ways of data description, this paper divides the data into qualitative and quantitative types, focusing on the four types of descriptions of random variables, binary type, language level, and vocabulary terminology. The predicted values are compared in Figure 5. As can be found for these figures, the third one

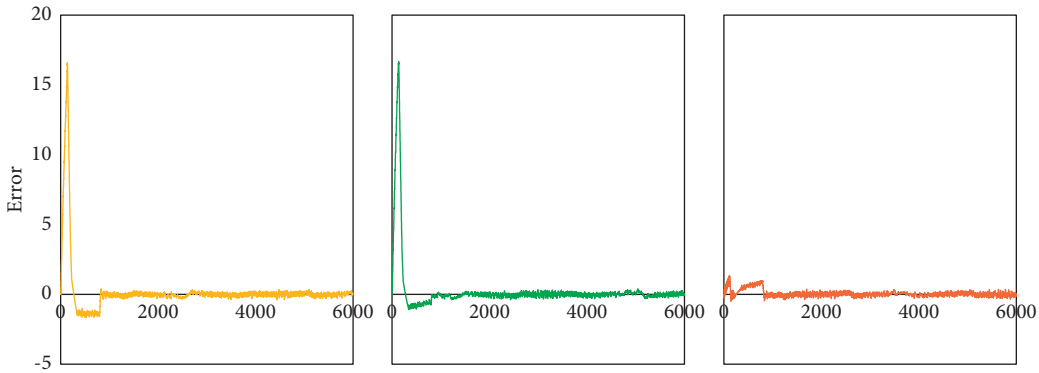


Figure 3: Error comparison.

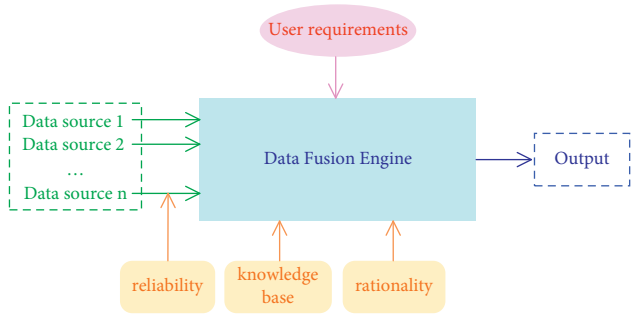


Figure 4: Multisource data source fusion structure.

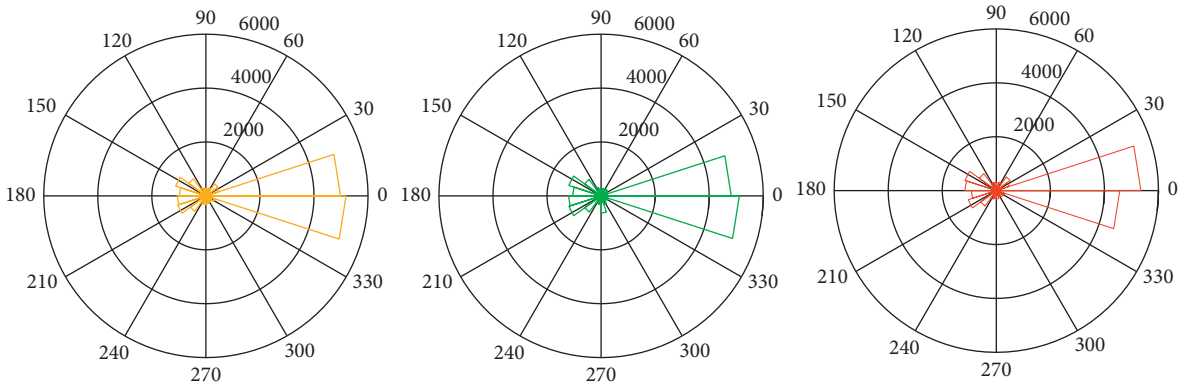


Figure 5: Value comparison.

exhibits the best performance of all, which is also consist with the previously mentioned analysis.

In the case of large samples, random variables follow a normal distribution. Binary data is used to describe the affirmation or negation of facts, and the value space is mostly {1, 0} or {True, False}. The data indicating the degree is generally expressed by Chinese adverbs of degree: very good, very poor, and so on. The degree level mostly uses 7 or 9 standards. The data based on the vocabulary term uses the vocabulary or term specified in the vocabulary space to give a qualitative description of things, and the number of vocabulary depends on the specific situation.

3.2. Support Calculation Based on Triangular Fuzzy Numbers. Taking into account the existence of ambiguity in the description of multisource data, triangular fuzzy numbers can be used to calculate the support value of the data for decision-making.

3.2.1. Conversion of Random Data. Suppose

$$x_0 = \mu - 3\sigma,$$
$$x' = \frac{x - x_0}{6\sigma}. \tag{6}$$

If the value of the random variable is larger, its support for decision-making is also greater. If the interval $[\mu - 3\sigma, \mu + 3\sigma]$ is divided into n equal parts, the conversion from random data to support can be defined as

$$s(x) = \begin{cases} (0,0,0), & x \le \mu - 3\sigma, \\ \left(\frac{i}{n}, x', \frac{i+1}{n}\right), & \frac{3\sigma i}{n} + x_0 < x \le \frac{6\sigma(i+1)}{n} + x_0, \\ (1,1,1), & x > \mu - 3\sigma. \end{cases} \tag{7}$$

If the value of the random variable is smaller, its support for the decision plan is greater, and then, the support is defined as

$$s'(x) = (\pi 1, 1, 1) - s(x). \tag{8}$$

3.2.2. Conversion of Binary Data. Binary data is described by 1 or 0. If the numbers of 1 and 0 in the data source are n and m, respectively, and the support is based on the value 1, and then, the support of the data source for decision-making is defined as

$$s(x) = \left(\frac{n}{n+m}, \frac{n}{n+m}, \frac{n}{n+m}\right). \tag{9}$$

3.2.3. Conversion of Degree Data. Generally speaking, 7 or 9 standards can be used to describe the quality of objects. This article adopts 7-level standards. The expression of degree adverbs can be divided into proportional type (the higher the efficiency, the better) and the inverse type (the higher the cost, the worse), and the degree of support for decision-making of each level can be quantified. The power in different situation is show in Figure 6, which shows the agreement between the prediction and the previously mentioned analysis in detail.

3.3. Data Fusion Algorithm. Suppose n decisions: A=($A1$, $A2$,...,An), m data sources: S=($S1$, $S2$,...,Sm), and the credibility (or importance) of each data source is pi. The data fusion algorithm is described as follows.

Step 1. Calculate the support of the data source for decision-making; extract data from the data warehouse, according to the different types of data and according to the previously mentioned method to convert it into the support for decision-making:

$$S_{ij} = \left(a_{ij}, b_{ij}, c_{ij}\right). \tag{10}$$

Among them, S_{ij} is the support degree of the i-th data source for the j-th decision target, and (a_{ij}, b_{ij}, c_{ij}) is the triangular fuzzy number representation of the support degree.

Step 2. Determine the weight vector of the OWA operator; select the appropriate fuzzy semantic quantization criterion according to the preference of the decision maker" and determine the corresponding parameter and value. The principle of fuzzy semantics is generally "majority," "at least half," or "as much as possible," and their parameter values are (0.3, 0.8), (0, 0.5), and (0.5, 1), which can be determined according to the parameters Fuzzy semantic quantization operator $f(x)$. According to $f(x)$, obtain the OWA weight vector w=($w1$, $w2$,..., wn); n is the number of data sources. Obtain the value of c.

Step 3. Convert s_{ij} according to the credibility (or importance) pi and support value s_{ij} of each data source; in order to use the OWA weight vector, each decision value needs to be converted according to p_i and s_{ij} and sorted in order of magnitude. The conversion method adopts the fuzzy judgment method. Assume

$$\begin{aligned} s_{ij\_\min} &= p_i s_{ij}, \\ s_{ij\_\max} &= p_i + s_{ij} - p_i s_{ij}, \\ s_{ij\_ave} &= \frac{n}{\sum_{i=1}^{n} p_i} p_i s_{ij}. \end{aligned} \tag{11}$$

Step 4. Fuse the data according to the OWA operator weight vector and the converted support, and calculate the final decision value of each decision:

$$s_j = \sum_{i=1}^{m} w_i b_{ij}, \quad j = 1, 2, \ldots, n. \tag{12}$$

Step 5. Make a decision based on the actual problem according to the decision value. The corresponding prediction is shown in Figure 7.

4. Forecast of Engineering Job Demand

Enterprise human resource demand forecasting is to predict human resources, which is a complex system, so, it must be completed based on a scientific forecasting model. This article has already sorted out the existing human resource demand forecasting models. In the qualitative and quantitative models of demand forecasting, each category contains many models with different forecasting focuses. To predict the needs of enterprise human resources, only those practical methods are scientific. After the previous analysis of the forecasting object and the internal and external human resources environment of the enterprise, the appropriate forecasting method can be determined according to the characteristics of the enterprise.

This article is a mid- and long-term forecast of the talent needs of a company's key positions. When choosing a forecasting method, it should be taken into account that due to the different degree of influence of internal and external factors, the results of the forecast of the talent demand for key positions will be different, so the main influencing

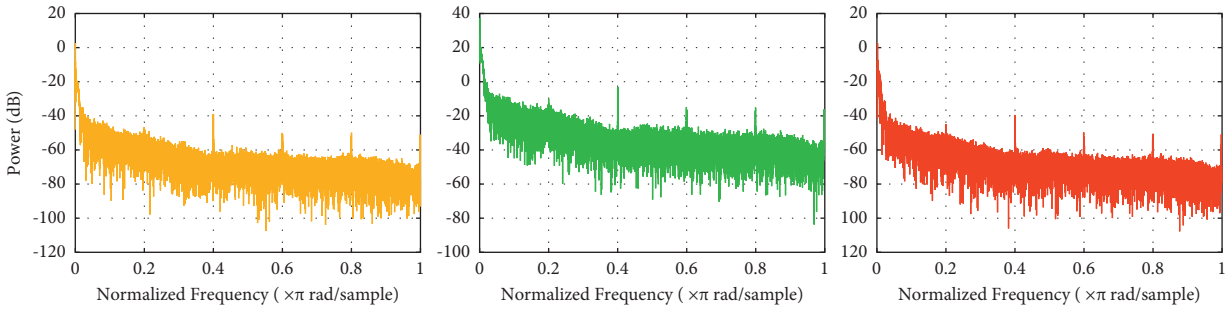


Figure 6: Power.

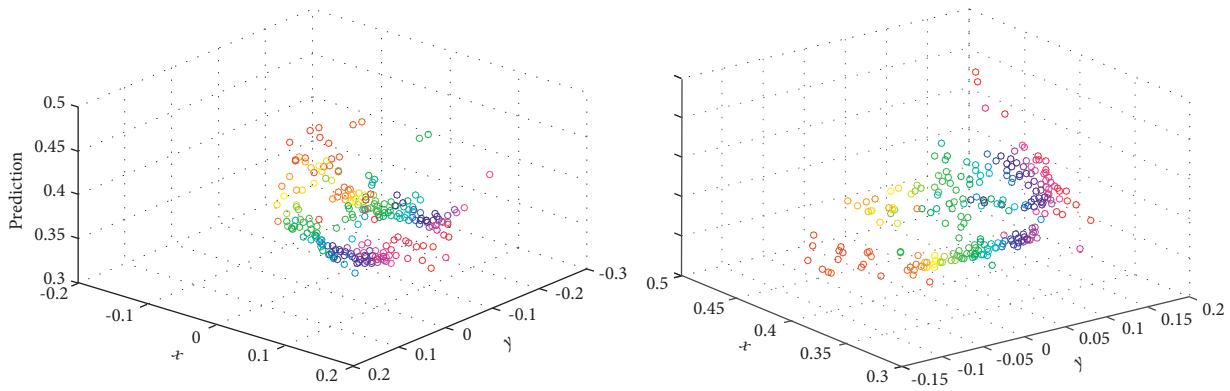


Figure 7: Prediction in different x and y.

factors need to be selected when forecasting. When predicting the total demand for talents, different variables are selected and a variety of forecasting schemes are used for forecasting, so that the information contained in various methods can be integrated to obtain more accurate forecasting values. Generally speaking, the selection of the method for forecasting talent demand for key positions in this article is mainly based on the following considerations.

(1) There are many factors that affect the demand for talents in key positions, but different factors sometimes have inherent correlations. Therefore, we should consider screening all factors to find out the main factors. In this way, we can consider using a few A variables to describe the nature of multiple variables.

(2) According to the data processing results, consider using one or more factors that have a greater impact on the demand for talents in key positions, and use mathematical models in statistical methods such as regression analysis and forecasting to predict the demand for talents, which will make the results more scientific, so as to obtain a better prediction effect.

(3) The development of enterprise human resources is a function with time as the basic variable. As time changes, the quantity and status of enterprise human resources are changing. Through the analysis of the internal and external environment of human resources of a company, it can be seen that the company is in a period of stable development, and the demand for talents in key positions has time continuity. Therefore, the time factor is an indispensable and important factor in the forecast of demand for talents in key position variable.

(4) Both theory and practical experience show that the combined forecasting method concentrates more relevant information and forecasting skills, so it can obtain better forecasting effects than single forecasting models, significantly improve the forecasting effect, and reduce the systematic error of the forecast. Therefore, this article will consider the use of combined forecasting methods to obtain the forecast value of the total demand for talents in order to reduce forecast errors and improve forecast accuracy.

Based on the previously mentioned considerations, this article will choose a quantitative forecasting method to predict a company's demand for key position talents and engineering professional and technical personnel (scarce talents) from 2006 to 2010. Comprehensive analysis and applicable methods are regression prediction model, gray system $GM(1, l)$, prediction model, and combination prediction. This article will first use the first two methods to forecast the demand for talents in key positions and finally use the combined forecasting method to comprehensively process the results of the two forecasts and obtain the forecast value of the demand for key positions in a company. The x y variation is shown in Figure 8.

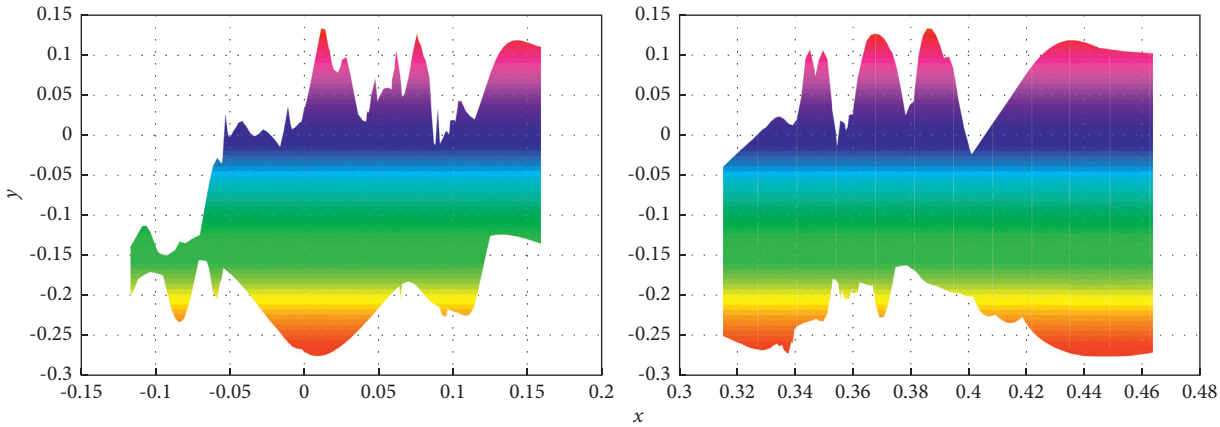


Figure 8: x and y variation.

Based on the analysis of the internal and external environment of a company's human resources, the preceding article qualitatively describes the factors that affect the talent needs of key positions. This qualitative analysis is only a preliminary identification of influencing factors and does not clarify the correlation and degree of influence on the number of talents in key positions. Therefore, it is necessary to carry out statistical analysis of various factors to find out the main factors affecting the demand for talents in key positions.

Based on the qualitative analysis of the internal and external influencing factors of talents in key positions, this paper selects some representative indicators that can be quantified among the factors to study the contribution of each factor to the demand for talents in key positions and engineering and technical personnel of a company degree and inner law.

Through the correlation test, we found that the A djus te *d* R Squ *s* re (① value) among the total talents in key positions is closest to 1 and has three important factors in qualitative analysis: the annual output of raw coal, the resource recovery rate, and the annual output of clean coal. Among them, the correlation coefficient between the annual output of raw coal and the demand for talents in key positions is the largest, with a value of 0.907. The resource recovery rate is negatively correlated with the demand for talents in key positions, with a value of -0.75. The correlation between the annual output of products is also not very significant, and the correlation coefficient value is 0.772 after the correlation analysis and processing of each factor index by SPSS software. There is an asterisk next to the correlation coefficient value of the annual output of raw coal, which indicates that when the specified significance level is 0.0 5. The associated probability of the statistical test is less than or equal to 0.0 5 (shown as 0.013 in the table); that is, the annual output of raw coal is talents in key positions, which are significantly correlated and positively correlated.

Therefore, the results of data processing show that, among the selected factors, the factor that has the greatest impact on the demand for talents in a company's key positions is the company's annual output value. This result is also the basis for the next personnel demand forecast. Based on this conclusion, we will use regression analysis methods to predict talents in key positions, which is the basis for selecting regression prediction methods in this article.

Analyzing the correlation test results between engineering professional and technical personnel and indicators of various influencing factors, it can be known that there are no indicators that have an important influence on engineering professional and technical personnel in terms of qualitative analysis among the selected indicators. That is to say, there is no index suitable for the regression prediction of engineering and technical talents among the selected indicators, so this article will use the gray prediction model to predict such talents.

Although the current coal industry has a good momentum of development, due to the country's relevant regulations on coal production, excessive exploitation of resources is not allowed (mine production is not allowed to exceed its approved production capacity), which is why the planned annual production value tends to stabilize. This planned value will be reserved relative to actual production, so the predicted value obtained by using the regression prediction model will be slightly smaller than the actual demand. At the same time, based on the actual situation of a certain company, since some people engaged in extractive work will meet the requirements for relocating extractive positions for 25 years in the next two years, considering the number of gaps in this part of personnel, there will be additional extractive positions in 2008 and 2009 number of talents. Based on the previously mentioned reasons, it can be seen that a company's actual demand for talents in key positions will increase based on the combined forecast value.

In addition, through the previous analysis, we know that another aspect of a company's demand for talents in key positions is the demand for competence and quality. Combining the development goals of the mine's future development plan for the personnel's educational structure, it can be known that the proportion of professional and technical personnel will increase in the personnel structure of key positions, and the proportion of high-level scientific and technological personnel will also increase to a certain extent.

The purpose of forecasting is to meet the demand for personnel in key positions and improve labor productivity. Only by strengthening the management of talents in key positions, fully mobilizing the enthusiasm of these personnel, and improving the overall competence and performance of talents in key positions can the goal of improving the overall performance of the enterprise be achieved. Therefore, it is necessary to use the theory of human resource management and combine the actual situation of a company to formulate corresponding planning measures for the management of talents in key positions of a company, in order to achieve the role of management and incentives for talents in key positions and promote the steady development of the company.

5. Conclusion

This paper qualitatively analyzes and selects the factors affecting the talent demand of key positions; a certain unit finds the quantifiable and influential factors according to the characteristics of demand influencing factors. Representative factors of talent demand for key positions in a certain unit. Using historical data, statistical methods are used to process the eight related factors of a certain unit, which confirms the factors that have a greater impact on the demand for talents in key positions of a certain unit, affecting the demand for talents in key positions of the same type of enterprise. The identification of factors provides a basic argument for a certain unit. Hence, the following conclusions can be obtained.

(1) Based on the results of statistical analysis and the characteristics of existing data, two variables of factory output and time are selected to be used in regression analysis forecasting model and gray system forecasting model for a certain unit to predict the demand for talents in a key unit; a certain unit finally adopts a combination forecast. The method determines the predicted value of talent demand for a certain unit's key positions.

(2) Besides, according to the results of demand forecasting and the current status of human resource management in a certain unit, this article proposes talent management planning measures for key positions in a certain unit, hoping to provide a reference for the management of key talents in a certain unit and ensure a certain unit's key positions in the future talent demand reserve.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors acknowledge the Henan Province Soft Science Research Project, Research on Evaluation Criteria of Scientific and Technological Talents in Henan Province Based on Industry Classification under the Background of Big Data (no.: 202400410290) and the Pingdingshan University Ph.D., Start Fund Project, Research on the Job Satisfaction and Turnover Rate of Employees in Private Enterprises in Pingdingshan (no.: PXY-BSQD-2019008).

References

[1] Y. Yu, C. Yang, Q. Deng, T. Nyima, S. Liang, and C. Zhou, "Memristive network-based genetic algorithm and its application to image edge detection [J]," *Journal of Systems Engineering and Electronics*, vol. 32, no. 5, pp. 1–9, 2021.

[2] Y. Ishida and S. Hashimoto, "Asymmetric characterization of diversity in symmetric stable marriage problems: an example of agent evacuation," *Procedia Computer Science*, vol. 60, no. 1, pp. 1472–1481, 2015.

[3] P. Zoha and R. Kaushik, "Image edge detection based on swarm intelligence using memristive networks," *IEEE Trans. on CAD of Integrated Circuits and Systems*, vol. 37, no. 9, pp. 1774–1787, 2018.

[4] J. Pais, "Random matching in the college admissions problem," *Economic Theory*, vol. 35, no. 1, pp. 99–116, 2018.

[5] J. J. Jung and G. S. Jo, "Brokerage between buyer and seller agents using constraint satisfaction problem models," *Decision Support Systems*, vol. 28, no. 4, pp. 291–384, 2020.

[6] Y. Liu and K. W. Li, "A two-sided matching decision method for supply and demand of technological knowledge," *Journal of Knowledge Management*, vol. 21, no. 3, JKM-05-2016-0183, 2017.

[7] J. Byun and S. Jang, "Effective destination advertising: matching effect between advertising language and destination type," *Tourism Management*, vol. 50, no. 10, pp. 31–40, 2015.

[8] A. N. Nagamani, S. N. Anuktha, N. Nanditha, and V. K. Agrawal, "A genetic algorithm-based heuristic method for test set generation in reversible circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 37, no. 2, pp. 324–336, 2018.

[9] C. Koch and S. P. Penczynski, "The winner's curse: c," *Journal of Economic Theory*, vol. 174, pp. 57–102, 2018.

[10] C. K. Karl, "Investigating the winner's curse based on decision making in an auction environment," *Simulation & Gaming*, vol. 47, no. 3, pp. 324–345, 2016.

[11] D. Ettinger and F. Michelucci, "Creating a winner's curse via jump bids," *Review of Economic Design*, vol. 20, no. 3, pp. 173–186, 2016.

[12] J. A. Brander and E. J. Egan, "The winner's curse in acquisitions of privately-held firms," *The Quarterly Review of Economics and Finance*, vol. 65, pp. 249–262, 2017.

[13] Z. Palmowski, "A note on var for the winner's curse," *Economics/Ekonomia[J]*.vol. 15, no. 3, pp. 124–134, 2017.

[14] B. R. Routledge and S. E. Zin, "Model uncertainty and liquidity," *Review of Economic Dynamics*, vol. 12, no. 4, pp. 543–566, 2009.

[15] D. Easley and M. O'Hara, "Ambiguity and nonparticipation: the role of regulation," vol. 22, no. 5, pp. 1817–1843, 2019.

[16] P. Klibanoff, M. Marinacci, and S. Mukerji, "A smooth model of decision making under ambiguity," *Econometrica*, vol. 73, no. 6, pp. 1849–1892, 2005.

[17] Y. Halevy, "Ellsberg revisited: an experimental study," *Econometrica*, vol. 75, no. 2, pp. 503–536, 2017.

[18] D. Ahn, S. Choi, D. Gale, and S. Kariv, "Estimating ambiguity aversion in a portfolio choice experiment," *Working paper*, vol. 5, no. 2, pp. 195–223, 2019.

[19] T. Hayashi and R. Wada, "Choice with imprecise information: an experimental approach," *Theory and Decision*, vol. 69, no. 3, pp. 355–373, 2010.

[20] K. Zima, E. Plebankiewicz, and D. Wieczorek, "A SWOT analysis of the use of BIM technology in the polish construction industry," *Buildings*, vol. 10, no. 1, 2020.

[21] P. Sun, B. Liu, and T. Sun, "Injury status and strategies of female 7-a-side rugby players in Anhui Province," *Sports Boutique*, vol. 38, no. 3, pp. 72–74, 2019.

[22] P. Guild, M. R. Lininger, and M. Warren, "The association between the single leg hop test and lower-extremity injuries in female athletes: a critically appraised topic," *Journal of Sport Rehabilitation*, vol. 30, no. 2, pp. 1–7, 2020.

[23] U. G. Inyang, E. E. Akpan, and O. C. Akinyokun, "A hybrid machine learning approach for flood risk assessment and classification," *International Journal of Computational Intelligence and Applications*, vol. 19, no. 2, Article ID 2050012, 2020.

[24] Q. Liu, S. Du, B. J. van Wyk, and Y. Sun, "Double-layer-clustering differential evolution multimodal optimization by speciation and self-adaptive strategies," *Information Sciences*, vol. 545, no. 1, pp. 465–486, 2021.

[25] H. R. Medeiros, F. D. B. D. Oliveira, H. F. Bassani, and A. F. R. Araujo, "Dynamic topology and relevance learning SOM-based algorithm for image clustering tasks," *Computer Vision and Image Understanding*, vol. 179, pp. 19–30, 2019.

[26] Y. Deng, D. Huang, S. Du, G. Li, C. Zhao, and L. Jun, "A double-layer attention based adversarial network for partial transfer learning in machinery fault diagnosis," *Computers in Industry*, vol. 127, Article ID 103399, 2021.

[27] J. J. Chan, K. K. Chen, S. Sarker et al., "Epidemiology of Achilles tendon injuries in collegiate level athletes in the United States," *International Orthopaedics*, vol. 44, no. 3, pp. 585–594, 2020.

[28] W. Li, G. G. Wang, and A. H. Gandomi, "A survey of learning-based intelligent optimization algorithms," *Archives of Computational Methods in Engineering*, pp. 1–19, 2021.

[29] G. G. Wang, A. H. Gandomi, A. H. Alavi, and D. Gong, "A comprehensive review of krill herd algorithm: variants, hybrids and applications," *Artificial Intelligence Review*, vol. 51, no. 1, pp. 119–148, 2019.

Hindawi
Scientific Programming
Volume 2022, Article ID 3102337, 10 pages
https://doi.org/10.1155/2022/3102337

Research Article

Multisensor Feature Fusion-Based Model for Business English Translation

Pingfei Zheng [1,2]

[1]*Universiti Utara Malaysia, Gubangbasu County, Kedah 06010, Malaysia*
[2]*Guangzhou Institute of Science and Technology, Guangzhou, Guangdong 510540, China*

Correspondence should be addressed to Pingfei Zheng; gjjy@gzist.edu.cn

Received 13 December 2021; Revised 6 January 2022; Accepted 8 January 2022; Published 31 January 2022

Academic Editor: Baiyuan Ding

With the wide use of computers, machine translation has been gradually applied in many fields from natural language processing, such as industry, education, and so on. Due to the increasing demand for multilanguage translation, it is an urgent problem to effectively improve the quality of text translation. Driven by the upsurge of artificial intelligence, neural network technology is increasingly integrated into the field of machine translation, which gradually expands the traditional machine translation method into neural machine translation method. With the continuous improvement of deep learning technology, machine translation has gradually integrated these methods and strategies and achieved good results in multiple tasks, but there are still some shortcomings. The most prominent problem is that, since word vector is the basis for the model to obtain semantic and grammatical information, the existing methods cannot obtain semantic and grammatical feature information, which greatly reduces the accuracy of English translation. Based on this, this paper proposed a method of splicing word vector with character- level and word-level encoding vector. The characterization of fusion of more word vector can effectively solve the word does not appear in the table, the word with some low frequency, can express meaning more complete information, performance directly affects the whole translation model, the results can be seen through the experiment, we put forward the characteristics of the fusion method and strategy, can effectively enhance the overall translation performance of the model.

1. Introduction

With the widespread use of computers, the speed of manual text translation is no longer enough to meet daily needs [1–3]. Therefore, how to use computers to realize the mutual conversion between multiple languages is an urgent problem to be solved. Machine translation technology has gradually attracted the attention of academic circles and developed rapidly. Driven by the upsurge of artificial intelligence, improving the quality of machine translation by using artificial intelligence technology is a hot research direction in the field of natural language processing. As the demand for multilingual text translation grows, many AI companies are developing translation software and providing online services. In the history of machine translation, British and French governments actively promote mutual translation between English and French, our country also supports the study of the machine translation between Chinese and English from the perspective of the theory of value, and machine translation is a blend of mathematical linguistics automation technology such as computer science discipline, studies of machine translation at the same time [4, 5].

Since the 1890s, data-driven statistical machine translation has replaced rule-driven method and gradually become the mainstream machine translation method. Compared with rule-driven methods, statistical machine translation methods do not require any rules, extract language knowledge from large-scale parallel corpora, construct efficient translation models by establishing statistical and probabilistic models for the whole translation process, and adjust model parameters during training. Statistical machine translation has become the core technology of online machine translation systems at home and abroad because of its advantages such as low labor cost, short development time,

and good robustness, which overcomes the bottleneck of rule-making based translation methods. In the process of statistical machine translation, various translation knowledge are needed by human experts set of language features to say, but different language leads to structure is different also, when translating, the conversion structure will become very important, only rely on artificial design features, will lead to the structure of language cannot get comprehensive coverage. At the same time, the traditional statistical machine translation technology still faces severe challenges due to the serious problems of data sparsity, strong dependence on corpus, and too much time cost [6].

So far, machine translation technology has been developed for decades, although various algorithm models have been introduced, the accuracy of machine translation is still low, and it cannot replace professional translators. Among them, the most prominent problem is the poor translation effect of long clauses with many words and complex sentence structure. In English, the structural components of long sentences are more complex. In addition to the main sentence structure, there are various modifiers and conjunctions. In addition, long sentences may contain more than one clause, and the relationships between clauses may be nested and parallel, so syntactic analysis is a necessary prerequisite for long sentence translation. Therefore, syntactic preprocessing of long and difficult sentences is an effective way to improve the translation quality of long sentences. At present, syntactic analysis translation methods can be roughly divided into two categories: language template-based translation methods and statistics-based translation methods. The translation method based on language template, which refers to the surface features of a sentence, is the earliest technical method of grammar translation. The advantage of language template-based translation method is that the translation is most accurate when the sentence template features are strong, while the disadvantage is that the translation is inaccurate or even impossible when the sentence template features are weak [7–9].

2. Related Works

In order to improve the translation methods based on language templates, scholars around the country have carried out in-depth studies and gradually evolved into a translation method based on statistics. This method can use machine learning method to conduct a lot of data mining and feature learning for weak features of sentences, such as the features of conjunctions and sentence patterns of long sentences and the features of punctuation usage, which can make up for the defects of incomplete matching of sentence features in template matching translation. However, the method based on statistics has its own limitations, if the sentence itself has a small number of punctuation marks or connectives, the accuracy of statistical translation methods will also decline [10–13].

Despite the promise of statistical machine translation, it is still difficult to translate statements in one language completely and correctly into statements in another language. Since 2013, deep learning composed of neural networks has become once powerful. The advantages of deep learning can well solve the disadvantages of statistical machine translation. Therefore, introducing deep learning methods into the field of machine translation has become a hot research direction. La et al. [14] put forward both short-term and long-term memory network, and its application in the frame of the end-to-end nerve machine translation, both short-term and long-term memory in circular neural network on the basis of introducing the door mechanism can control self-circulation, in order to achieve sustainable flow gradient can keep for a long period of time, avoids the disappear because of the gradient and the gradient semantic loss phenomenon caused by the explosion. Shuang et al. [15], for the first time, treated machine translation as an end-to-end learning task and effectively solved the problem of variable length of input and output by using recurrent neural network. Huang et al. [16] proposed a gated recurrent unit (GRU) to replace LSTM in handling machine translation tasks, GRU is actually an optimization of LSTM, which simplifies the internal structure, reduces the training parameters, and improves the training efficiency. Choi et al. [17] invented a new and simple network framework. Based on this, Ahmed et al. [18] proposed the attention mechanism, which effectively solved this problem and brought machine translation to a new lsevel. The attentional mechanism is essentially a small neural network trained simultaneously with the RNN-RNN network. Awad et al. [19] proposed a two-way simultaneous decoding method, in which the decoding direction of each word is dynamically determined by the model. Domestic scholars started relatively late in the field of machine translation, but they have also made a lot of achievements. The minimum risk training criterion was proposed by [20] to deal with the mismatch between training and testing in the codec framework of attention mechanism, and the performance of machine translation was significantly improved. Chen et al. [21] proposed an unsupervised adaptive method to solve the problem of domain migration, which fine-tuned the pretrained out-of-domain neural machine translation model by using pseudo-in-domain corpora. Specifically, the model first performed lexical induction to extract the dictionaries in the domain, and then adopted the pseudo-parallel intradomain corpus is constructed by reverse translation of the target sentences word by word.

From the above analysis, we know that the above methods have studied business English translation problem to some extent, some problem still exists [22–24]. For example, no scholar has applied multisensor feature fusion to this field till now, so the research here is still a blank, which has great theoretical research and practical application value for business English translation. In addition, almost all translation models are based on the encoder decoder framework. Although its structure achieves good results, all translation ends from left to right.

The contributions of this paper is: 1. multirepresentation fusion is proposed for the word vector of translation, encoding the input data with various granularity. 2. Meanwhile, a new neural machine translation model is proposed combining with the idea of refining neural networks and transformer network. Next, I will introduce this chapter according to the sequence of character-level coding refining neural networks.

This paper consists of five parts. The first and second parts give the research status and background. The third part is the translation model of fusion character encoding and negotiation network. The fourth part shows the experimental results and analysis. The experimental results of this paper are introduced and compared and analyzed with relevant comparison algorithms followed. Finally, the fifth part concludes the full paper.

3. Translation Model of Fusion Character Encoding and Negotiation Network

3.1. Multifeature Fusion Based on Transformer. Although the word-level vector in neural machine translation has achieved good results, there are still many inevitable defects, for example, it cannot solve the problem of accurate expression of rare words and words that do not appear in the training vocabulary, so it is generally only used in the translation of some languages with rich corpus, such as English, German, and French. Sometimes researchers solve this problem by increasing the vocabulary, but the algorithm complexity during training and decoding increases linearly with the vocabulary used, leading to a vicious circle [25]. Faced with these existing problems, some scholars proposed character-based neural machine translation model in 2017, as shown in Figure 1. One of the advantages of letter-level coding is that it is more suitable for multilanguage translation than word-level model, while word-level coding requires each language to use a separate vocabulary [26].

First, the input sequence is mapped to the corresponding character embedding vector and then the convolution operation is carried out using the convolution kernel with different window sizes, and then the output is connected. For example, in Figure 1, there are three convolution kernels whose window sizes are 3, 4, and 5, which are equivalent to learning character-based triples. Then the output of the convolution layer is input to the maximum pooling layer, which is equivalent to selecting the most important features to generate fragments and embed them. So far, from the initial input character embedding, we can get the fragment embedding which the system considers as language synthesis. Then, all fragments are embedded through the Highway Network layer (its function is similar to residual Network, and information flow is controlled by adding some gating systems) and two-way GRU, so that the encoder output can be obtained by the model at last. Finally, the decoder uses the attention mechanism and character-level encoding GRU Network decoding [27, 28].

3.2. Character Level Encoding Based on CNN. Since a single character contains little information and no rich semantic information, the input sequence is supplemented by convolution and GLU network, and finally the splicing word vector is input into transformer for training, as shown in Figure 2. As for GLU mentioned above, this section will introduce its specific structure and principle in detail. Its schematic diagram is shown in Figure 2, and the formula is as follows [29]:

$$h_l(X) = (X * W + b) \otimes \sigma(X * V + c), \tag{1}$$

where the input sequence is represented by X, L represents the number of layers, W and V represent different convolution kernels, respectively, and b and c represent the corresponding bias, respectively, and σ is the sigmoid activation function. Half of the convolution is normalized by sigmoid activation function, which controls what values of the other half of the convolution can be transferred to the next GLU layer. To put it simply, a gating mechanism is added to the original convolutional network, which is equivalent to an output gate in LSTM, enabling the model to learn and control the output of the convolutional layer and improving the interpretability of the model. The gating mechanism adjusts the flow direction of data information in the network, and some scholars have proved its reliability in the circular network.

The gradient calculation formula of GLU is shown in (2), this can generally be thought of as a kind of multiplicative hop connection that allows gradients to pass smoothly through each layer. This allows gradient information to flow smoothly in multiple time steps. Without these gating mechanisms, gradient disappearance may occur in the transition of the information. In addition, the network structure allows the model to perform nonlinear transformation. With GLU's multilayered stack structure, long-term dependencies can be captured without the need for gradient extinction.

$$\nabla[X \otimes \sigma(X)] = \nabla X \otimes \sigma(X) + X \otimes \sigma'(X) \nabla X. \tag{2}$$

3.3. Fusion of Multifeature Codes. In this chapter, the direct vector splicing strategy is adopted to connect word vector TW and character-level encoding vector (character-level encoding model output in the previous section) to obtain the final representation vector of word. In this chapter, this seemingly simple method is adopted, but the experimental results show that it is very efficient. This strategy is frequently used not only in the research of neural machine translation, but also in its other natural language processing tasks, through vector splicing:

$$x_t = [w_t; \mathrm{chr}_t], \tag{3}$$

x_t is the final representation vector, chr_t is character-level encoding vector. In this section, all corpus data collected will be fully utilized to encode the input at word level and character-level, respectively, and then be used as the feature vector of the final text by means of splicing. In the face of rare words and unknown words that do not appear in the word list of the training set, more information can be obtained by

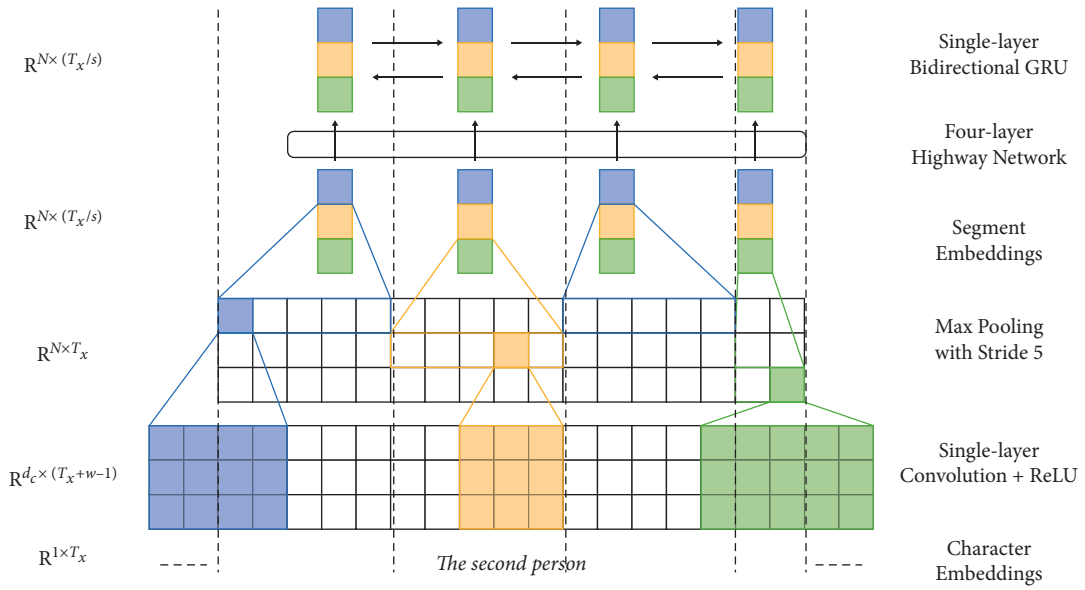


Figure 1: Word vector coding based on convolutional neural network.

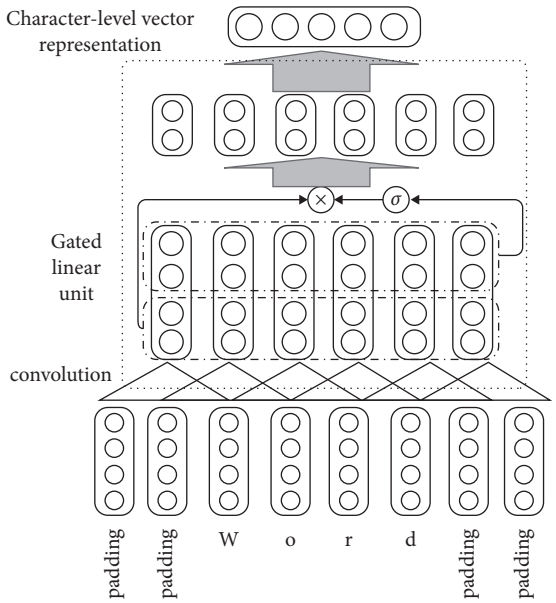


Figure 2: Schematic diagram of character-level coding model.

character-level coding vector, so as to alleviate the impact of this kind of problem. Retraining of word-level coding can obtain more semantic and sentential information, which is conducive to the integrity of the whole sentence information. When training the front part of neural network. The encoding of word vector is first trained. After the model converges, character-level encoding is added to ensure the high efficiency of the model. In the pretraining stage of the model, Transformer model structure is used, and the loss function of cross entropy type is adopted in this chapter for modal optimization. Just like the conventional neural machine translation model, attention mechanism is also added to the network. The formula for attention calculation is as follows:

$$ctx_e = \sum_{i=1}^{T_x} \alpha_i h_i,$$
$$\alpha_i \propto \exp\left(v_\alpha^T \tan h\left(W_{at,h}^c h_i + W_{att,\hat{s}}^c \hat{s}_{j-1}\right)\right) \forall i \in [T_x],$$
$$\sum_{i=1}^{T_x} \alpha_i = 1. \tag{4}$$

$\hat{s}_{j-1}$ is the corresponding hidden layer state quantity, T_x present the time point, α_i is the parameter, and W is the weight. Different from the conventional network based on encoding and decoding mechanism, there is more D2 part in the deliberative network. The specific model calculation details are as follows:

$$ctx_c = \sum_{j=1}^{T_j} \beta_j \left[\hat{s}_j; \hat{y}_j\right],$$
$$\beta_j \propto \exp\left(v_\beta^T \tan h\left(W_{dt,j,j}^{d} \left[\hat{s}_j; \hat{y}_j\right] + W_{dt,s}^d s_{t-1}\right)\right) \forall i \in \left[T_{\hat{y}}\right],$$
$$\sum_{j=1}^{T_j} \beta_j = 1. \tag{5}$$

From the above calculation process, it can be known that at moment T, the second decoding calculates the attention of each sequence generated by the first, which is equivalent to referring to the words after moment T in the second decoding process. Due to the RNN structure based on the neural network, there are two decoder parts and the target prediction space is very large, it is very difficult to calculate, gradient calculation is also very difficult, so the network uses Monte Carlo method to optimize the maximum likelihood loss.

However, in the syntactic feature model proposed in this paper, it is stored according to units. The syntactic feature

unit model will be established below, which directly stores the syntactic relations between words into corresponding syntactic units. The unit construction rules are as follows:

$$M_i = Mx_i, MPx_i, MCx_i, MBx_i. \quad (6)$$

M_i represents the storage syntactic unit of the ith word in a sentence. MPx_i, MCx_i, MBx_i denote the position of the ith parent node word, child node word, and adjacent node word in the sentence. The mathematical definition of conditional random fields is: suppose an undirected graph

$$G = T(V, G). \quad (7)$$

where V is the set of vertices and E is the set of edges. Let X be the condition that satisfies Y_v, then the variable Y_v satisfies the following equation:

$$p(Y_v|X, Y_u, u \neq v) = p(Y_v|X, Y_u, u: v), \quad (8)$$

where u, v represents two vertices contained in graph T, then (X, Y) is a random field of strips.

4. Experimental Results and Analysis

4.1. Introduction to Experimental Environment and Data Set. In order to compare the quality of multiattentional neural machine translation models, we conducted several experiments on the (a) WMT 14 English-German translation task. The training data set of the experiment is provided by WMT 14 English-German translation task. After preprocessing the corpus (deleting too long or too short sentences, removing blank lines, etc.), it contains a total of 4.5 million English-German sentence pairs, involving about 116M English words and 110M German words. Meanwhile, (b) newstest 2014 English-German Parallel Corpus database was used as the test set, which contained 2737 parallel statement pairs, and (c) Newstest2013 English-German parallel Corpus database was used as the verification set, which contained 3000 parallel statement pairs. Then byte pair encoding is used to process the parallel corpus training set test set and verification set, and a fixed dictionary with a size of 32000 is extracted from the training data set, which is mixed with the source language lists with the highest frequency words and target language words are shared between input text and output text, and the dictionary does not change during training [28]. All the models in this paper are coded by Python language, and all the experiments in this paper are carried out on a hardware device of NVIDIA 1080Ti GPU.

In the training process of the model, the model uses the stochastic gradient descent algorithm to optimize the model parameters, and the initial value of the learning rate is 1.0, and the maximum number of training steps is set to 340K. After 170K steps, the learning rate is halved every 17K steps. At the same time, the ownership weight parameters in the model are initialized to 0.1, and the deviation is initialized to 1.0. In order to prevent the generation of gradient explosion, gradient clipping algorithm and dropout algorithm (dropout probability set to 0.2) were used in the training so that the norm of the gradient would not exceed 5.0.

4.2. Experimental Results Analysis. In order to verify the performance of the proposed LSTM-based attention embedding translation model, the standard LSTM model and the LSTM model of attention embedding were used to train the experimental data set. The results are shown in Figure 3. As can be seen from the figure, compared with the standard LSTM model, the BLEU value of the attention-embedded LSTM model is higher, indicating that the attention-embedded LSTM model has better translation effect for long sentences and the model performance is effectively improved.

As shown in Table 1, the results are mainly shown and compared with some previously well-known models that are relevant to this study. The algorithm is briefly explained as follows:

RNNsearch: This model is the first to bring the attention mechanism to the field of neural translation, which has special significance for the development of the following attention mechanism and machine translation. Their encoder and decoder are based on the RNN of the gated mechanism. The encoder generates the required middle hidden layer vector, and the decoding process passes the note. The semantic mechanism solves the context representation of the output current moment, and finally combines the decoding output of the upper moment with the context representation to complete word prediction.

CovS2S: This model is a neural translation model proposed by Facebook based on convolutional neural network, which preserves the long-distance dependence of words in sentences through gating mechanism combined with multihop convolution and solves the problem of gradient descent disappearing in the training process. This model relies on the pattern of recurrent neural network for neural translation break, improve the model training parallel ability, and model training efficiency in the translation effect is also very good.

Deep-Att + posUnk: This model is a neural translation model proposed by Facebook based on convolutional neural network, which preserves the long-distance dependence of words in sentences through gating mechanism combined with multihop convolution and solves the problem of gradient descent disappearing in the training process. This model relies on the pattern of recurrent neural network for neural translation break and improve the model training parallel ability, and model training efficiency in the translation effect is also very good.

GNMT + RL: This network proposes corresponding solutions to the situation that the calculation cost is too high in the process of deep learning training and translation, and there are rare words in the translation sentences by adopting low-precision algorithms and cutting words into common wordpieces.

Transformer: This model is only based on the attention mechanism and the feedforward neural network

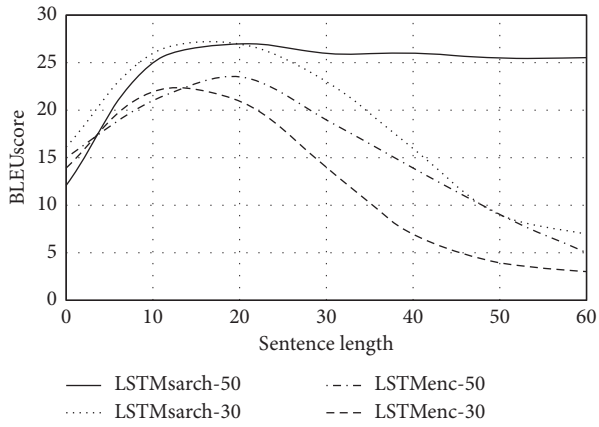


Figure 3: Translation performance of attention mechanism.

Table 1: BLEU value of different models.

Model	German-English translation	English-French translation
RNNsearch-50 + UNK	—	36.2
ConvS2S	25.2	40.8
Deep-Att + posUnk	—	39.5
GNMT + RL	24.9	39.8
Transformer	27.8	38.2
Transformer (big)	29.9	41.2
Ours	26.5	38.9

method, which can train the model in parallel to shorten the training time. The self-attention mechanism is adopted to shorten the distance between sentences. This model not only improves the translation effect, but also brings researchers a new model design idea.

In order to verify the translation effect of the proposed English machine translation model based on LSTM attention embedding, the research adopts the traditional LSTM English machine translation model RNN English machine translation model Gru-attention English machine translation model, respectively. As can be seen from the figure, the BLEU value of the proposed English machine translation model based on LSTM attentional force embedding on the development machine and test set. Both are higher than the BLEU values of the traditional LSTM and RNN Gru-attention machine translation models, indicating that the proposed translation model improves the performance and translation quality by embedding attention machine in the LSTM network compared with the comparative translation models. The validity of LSTM-based attention embedding model for English machine translation is verified (Figure 4).

Comparing the three backbone network RNNSearch, ConvS2S, and transformer models in the test set, the transformer model based on self-attention mechanism performs better than RNNSearch and ConvS2S model on translation tasks in both rich and low-resource scenarios, as shown in Figure 5. In Figure 5, from the overall translation performance, the transformer model, and GT; ConvS2S

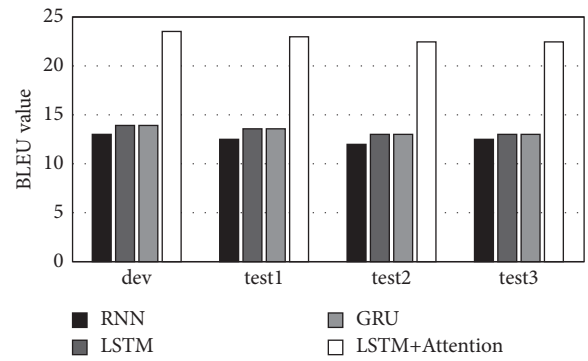


Figure 4: BLEU values of different translation models on experimental data sets.

model and gt; RNNSearch model is more obvious in the rich resource scenario, while the difference is not obvious in the low-resource scenario.

As can be seen from Figure 6, the degree of data dependence of NEURAL machine translation can be further verified. No matter in low-resource scenarios or resource-rich scenarios, BLEU value in newstest2016 test set also increases when the scale of training data increases.

However, the trend of BLEU value increase gradually slows down, which means that the performance improvement of pseudo-bilingual corpus constructed by artificial data enhancement is limited and cannot be increased indefinitely. The core of translation performance is still the size

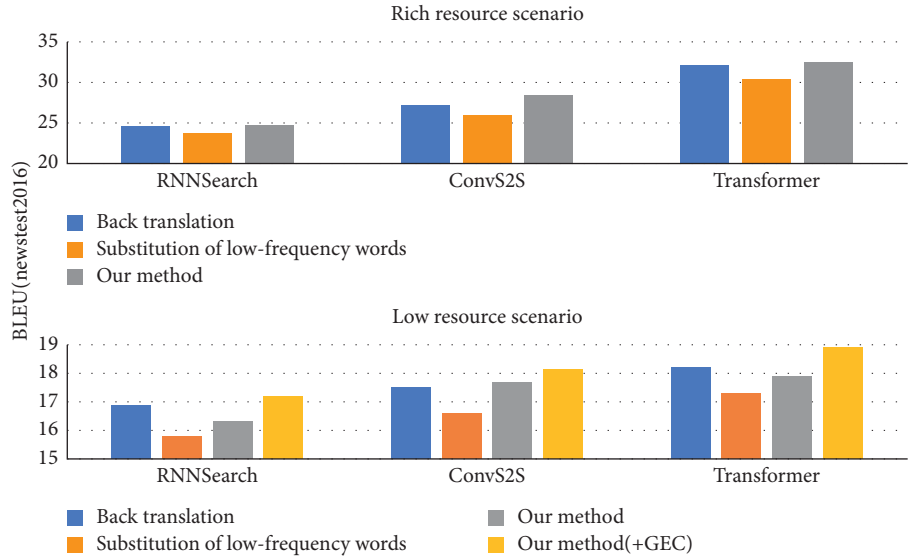


Figure 5: Different translation models apply different data enhancement methods to enhance translation performance in different scenarios.

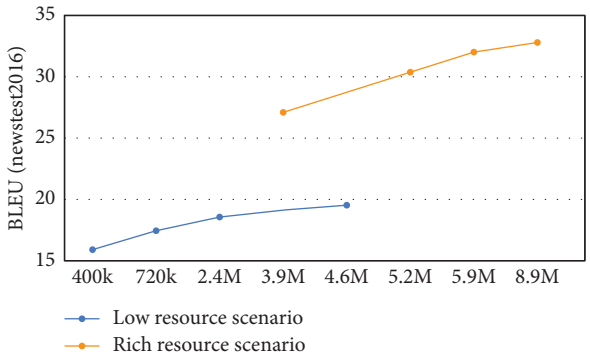


Figure 6: Variation relationship between corpus size and BLEU value.

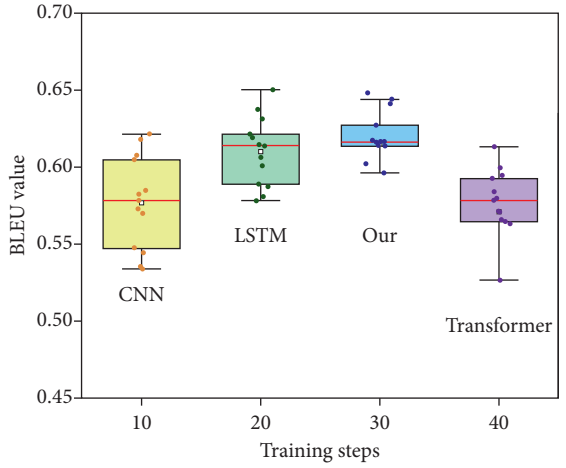


Figure 7: BLEU value line chart.

of the real bilingual corpus. Compared with the real parallel corpus scale of 400K in the low-resource scenario, the overall BLEU value of the translation task in the resource-rich scenario is more than 10 times higher than that in the low-resource scenario.

In order to further understand how the performance of the hyperbolic tangent neural machine translation model exceeds that of the comparative model, this paper uses a broken line graph to more intuitively reflect the training process.

As shown in Figures 7 and 8, the horizontal axis of the line graph is the training time step, the vertical axis is the BLEU score, the orange curve is the hyperbolic tangent neural machine translation model of this paper, the blue curve is the RNNSearch model, and the gray one is the Open NMT model. As can be seen from Figures 7 and 8, the curve of the model proposed in this paper is very steep at the beginning of training, indicating that BLEU score value of the model increases greatly during this period, and it is easy to see that the curve of BLEU value of the model proposed in this paper tends to be gentle after steep, indicating that it soon reaches convergence. However, the curve of the comparison model rises slowly and then reaches a gentle level. From the comparison of the line graph, it can be seen that the model proposed in this paper can greatly improve the training speed, achieve convergence quickly, and save the training time.

In this article, the attention thermal diagram after Sparsemax is adopted in the cross attention layer of decoder part is shown in Figures 9 and 10. As can be seen from the experimental results, after sparse normalization is used to calculate the probability, the corresponding attention score of some unrelated words in Figure 8 is zero (the color is white), which reduces the error size of model data distribution in the problem of induction bias. The accuracy of direct word alignment is improved, which not only improves the effect of the model but also improves the interpretability of the model. Some researchers also adopt the mechanism of local attention to deal with such bias. But in practice,

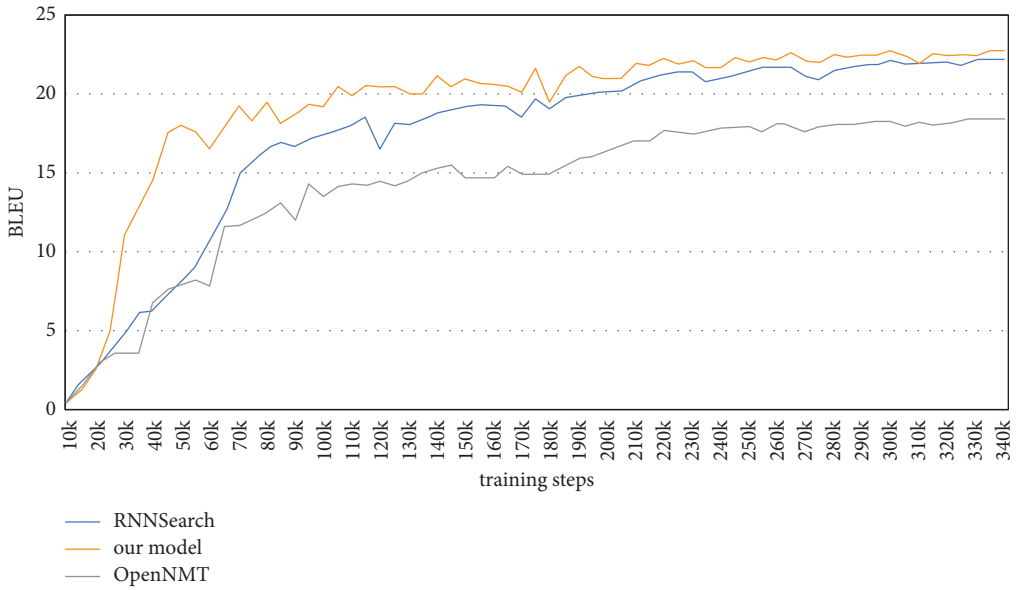


Figure 8: BLEU value line chart.

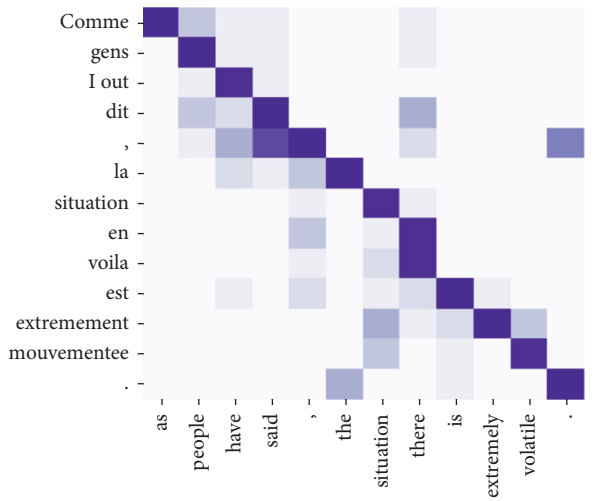


Figure 9: Matrix thermogram of attention score in French-English translation.

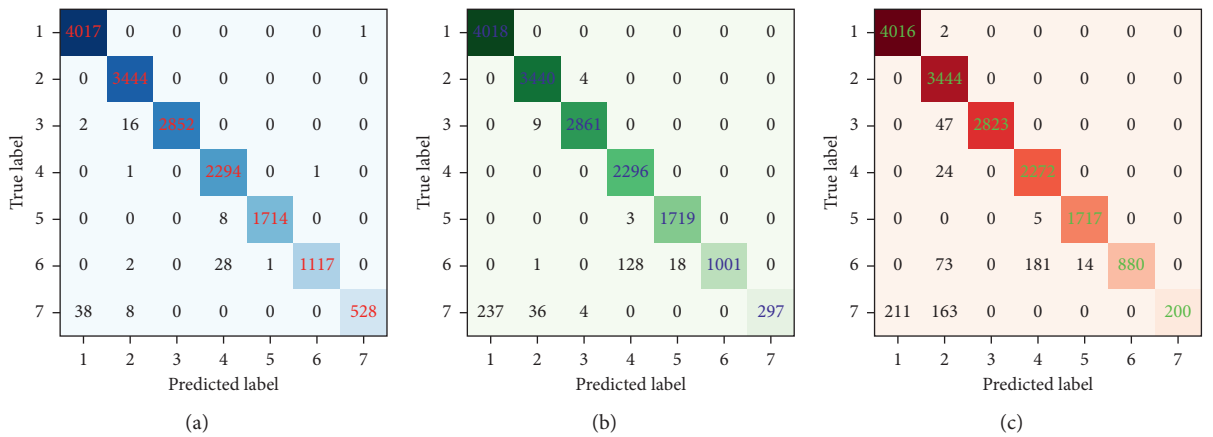


Figure 10: Confusion matrix of 7 words translation results: (a) Data set a; (b) Data set b; (c) Data set c.

discreteness and nondifferentiability require Monte-Carlo method to perform gradient approximation, which greatly increases the training complexity of the model. Sparsemax is differentiable, easy to calculate, and easy to use.

5. Conclusions

To study in depth and the rapid development of society now, people more and more demand for artificial intelligence brings the convenience of nerve has important role in machine translation, can help us faster to communicate in daily life, with the development of deep learning neural machine translation also got the attention of researchers and development. Especially in the application of encoder and decoder in natural language processing, the translation performance has been significantly improved. However, there are still many problems in deep learning, which is the focus and hotspot of current deep learning research, but also has many difficulties.

In this paper, transformer models, which are popular in various fields, and some existing problems of machine translation, are proposed to match the research methods and improvement strategies. Whether the text input sequence is represented well or not and whether the information contained is sufficient will directly affect the performance of the whole translation model. In most deep networks, word-level coding is carried out directly, although the effect is not wrong. However, if the embedded word vector is only at the word level, the quality of the corpus is required to be high while the learning effect is guaranteed. In the process of word vector training, the size of the model has a great influence on the number of iterations and translation effect. Therefore, on the basis of the current research work, this chapter uses the character-level vector embedding to ensure that the word vector can integrate more semantic information. Future research may study the specific translation performance of single decoding and integrates the global information of the first translation by integrating the idea of elaboration network, so that the model translation is more smooth and neat, and the generalization error of the whole model is smaller.

Data Availability

The data set can be accessed upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This study was supported by 2021 Educational Science Programming Projects (Special Program for Higher Education), Research on the Role Evolution and Core Competences Construction of Teachers of Application-oriented Undergraduate in Guangdong-Hong Kong-Macao Greater Bay Area (No. 2021GXJK316).

References

[1] R. Cattoni, M. A. Di Gangi, L. Bentivogli, M. Negri, and M. Turchi, "MuST-C: a multilingual corpus for end-to-end speech translation," *Computer Speech & Language*, vol. 66, Article ID 101155, 2021.

[2] C. Palladino, M. Foradi, and T. Yousef, "Translation alignment for historical language learning: a case study," *Digital Humanities Quarterly*, vol. 15, no. 3, pp. 11–24, 2021.

[3] W. Ling, "Research on the application of artificial intelligence in machine translation," *Journal of Frontiers in Engineering Technology*, vol. 1, no. 1, pp. 1–3, 2021.

[4] L. Deng, "The process-oriented assessment model of business English translation course in a flipped learning context," *Higher Education Studies*, vol. 10, no. 4, pp. 1–11, 2020.

[5] L. Bowker, "Machine translation literacy instruction for international business students and business English instructors," *Journal of Business & Finance Librarianship*, vol. 25, no. 1-2, pp. 25–43, 2020.

[6] L. Gao, "A study on the application of functional equivalence to business English E-C translation," *Theory and Practice in Language Studies*, vol. 8, no. 7, pp. 759–765, 2018.

[7] C. A. I. Ruizhen, "Status quo, problems and countermeasures of research on teaching of business English translation in China," *Journal of Ningbo University of Technology*, vol. 03, pp. 25–37, 2017.

[8] W. Hu, "On legal English translation from the perspective of legal linguistics," *Review of Educational Theory*, vol. 2, no. 3, pp. 6–10, 2019.

[9] J. Li, "Application of nord's functionalist translation theory to the translation of English business news discourses," *International Journal of Social Science and Education Research*, vol. 2, no. 5, pp. 46–56, 2019.

[10] L. Shi and Y. Liu, "A study on equivalence between English translation and cross-cultural communication based on translation coordination theory," *Revista Argentina de Clinica Psicologica*, vol. 30, no. 1, p. 283, 2021.

[11] Q. Xu and L. Deng, "Investigating the use of translation continuation tasks in commercial translation teaching: a study on translating user manuals," *Chinese Journal of Applied Linguistics*, vol. 44, no. 3, pp. 366–381, 2021.

[12] D. Viorela-Valentina, "Translation practice–A means for enhancing student employability," *Dialogos*, vol. 22, no. 38, p. 215, 2021.

[13] A. V. Novikova and L. A. Mylnikov, "Problems of machine translation of business texts from Russian into English," *Automatic Documentation and Mathematical Linguistics*, vol. 51, no. 3, pp. 159–169, 2017.

[14] S. Y. La and L. W. Wan, "Research on the LSTM Mongolian and Chinese machine translation based on morpheme encoding," *Neural Computing & Applications*, vol. 32, no. 1, pp. 41–49, 2020.

[15] K. Shuang, Z. Zhang, J. Loo, and S. Su, "Convolution-deconvolution word embedding: an end-to-end multi-prototype fusion embedding method for natural language processing," *Information Fusion*, vol. 53, pp. 112–122, 2020.

[16] Z. Huang, F. Yang, F. Xu, X. Song, and K.-L. Tsui, "Convolutional gated recurrent unit-recurrent neural network for state-of-charge estimation of lithium-ion batteries," *IEEE Access*, vol. 7, pp. 93139–93149, 2019.

[17] H. Choi, K. Cho, and Y. Bengio, "Fine-grained attention mechanism for neural machine translation," *Neurocomputing*, vol. 284, pp. 171–176, 2018.

[18] K. Ahmed, N. S. Keskar, and R. Socher, "Weighted transformer network for machine translation," 2017, https://arxiv.org/abs/1711.02132.
[19] M. Awad, S. M. Ibraheem, S. A. Napoleon, W. Saad, M. Shokair, and M. E. Nasr, "Secrecy enhancement of cooperative NOMA networks with two-way untrusted relaying," *IEEE Access*, vol. 8, pp. 216349–216364, 2020.
[20] E. Min, X. Guo, Q. Liu, G. Zhang, J. Cui, and J. Long, "A survey of clustering with deep learning: from the perspective of network architecture," *IEEE Access*, vol. 6, pp. 39501–39514, 2018.
[21] K. Chen and M. D. Sacchi, "Time-domain elastic Gauss-Newton full-waveform inversion: a matrix-free approach," *Geophysical Journal International*, vol. 223, no. 2, pp. 1007–1039, 2020.
[22] L. Deng, "The project-based flipped learning model in business English translation course: learning, teaching and assessment," *English Language Teaching*, vol. 11, no. 9, pp. 118–128, 2018.
[23] L. Wu and L. Wu, "Research on business English translation framework based on speech recognition and wireless communication," *Mobile Information Systems*, vol. 2, pp. 14–19, 2021.
[24] H. Wang, "Marine environment salinity measurement based on data classification system and features of business English translation," *Arabian Journal of Geosciences*, vol. 14, no. 15, pp. 1–14, 2021.
[25] Z. Cao, "Retracted article: evaluation of air pollution and business English vocabulary translation in coastal cities based on distributed storage," *Arabian Journal of Geosciences*, vol. 14, no. 17, pp. 1–7, 2021.
[26] Y. Jiang, "Critical thinking oriented teaching reform of business English translation," *Studies in Literature and Language*, vol. 19, no. 1, pp. 96–99, 2019.
[27] H. Feng, I. Crezee, and L. Grant, "Form and meaning in collocations: a corpus-driven study on translation universals in Chinese-to-English business translation," *Perspectives*, vol. 26, no. 5, pp. 677–690, 2018.
[28] J. Lyu, "The modularized construction on translation competence for business English majors in China," *English Language Teaching*, vol. 13, no. 7, pp. 124–129, 2020.
[29] Z. Li, "The construction of the turning classroom of business English translation teaching in higher vocational education under the Internet+ environment," *Frontiers in Educational Research*, vol. 2, no. 5, 2019.

Hindawi
Scientific Programming
Volume 2022, Article ID 7605125, 10 pages
https://doi.org/10.1155/2022/7605125

Research Article

Sentiment Analysis of English Text with Multilevel Features

Li Rao [1,2]

[1]*School of Foreign Studies, Nanjing University, Nanjing 210023, Jiangsu, China*
[2]*Jincheng College of Nanjing University of Aeronautics & Astronautics, Nanjing 211156, Jiangsu, China*

Correspondence should be addressed to Li Rao; raoli@nuaa.edu.cn

Received 11 December 2021; Accepted 8 January 2022; Published 28 January 2022

Academic Editor: Baiyuan Ding

In natural language processing, text sentiment analysis is one of the important branches. It refers to the use of text mining and other technologies to extract attitudes, opinions, and other information from texts containing emotional information for analysis. Traditional sentiment analysis methods can be roughly divided into two categories: one is dictionary-based methods, and the other is machine learning-based methods. The former relies on the quality of the sentiment dictionary, while the latter relies on a large amount of high-quality data, so both have certain limitations. In text sentiment analysis research, word-level and sentence-level sentiment information extraction is a basic research task and has important research value. Through research, it is found that domain knowledge and context are two important factors influencing the extraction of emotional information. To this end, this paper proposes a text sentiment analysis method that integrates multiple features and constructs three features, which are based on the sentiment value feature of the dictionary, the expression feature, and the improved semantic feature, which are combined to build a text sentiment classification model. Aiming at the colloquial, irregular, and diverse features of English social media texts, this paper proposes a multilevel feature representation method. The sentiment classification experiments on English text show that the multilevel features proposed in this paper can effectively improve the F1_macro and accuracy of multiple model classifications. Compared with the existing research, the model in this paper improves the effect the most obvious.

1. Introduction

English is the world's largest language in Europe, America, Oceania, Asia, and other dozens of countries, and the total number of people who use it as a mother tongue or a second language is about hundreds of millions. English short texts with subjective emotional colors, summarizing, analyzing, and reasoning about the emotional information contained in them, are conducive to business decision-making, political public opinion analysis, and social trend prediction in relevant countries, and are useful for preventing precision political marketing, building harmonious and stable international relations, and advancing the "One Belt One Road" strategy of crossborder and inter-regional economic and trade, and cooperation and win-win cooperation are of great value. However, most of the current research in this field focuses on the strong position of English. There are not many researches in the field of English sentiment analysis, as shown in Figure 1. Some studies try to use English-related tools to obtain the results of sentiment analysis of English-English translations. However, due to the loss of sentiment and even semantics in the translation stage, the analysis stage ignores the characteristics of English itself, so the result is not ideal [1–10].

There are two major difficulties in the sentiment analysis of English social media texts: (1) the characteristics of English itself make free word order, polysemy, complex morphology, and nonprojection relationships that often exist in the text, which increases the complexity of semantic analysis and sentiment extraction; (2) social media texts have the characteristics of colloquialism, multiple slang words, irregular language, and unobvious context information.

The method based on sentiment dictionary mainly calculates based on the prior information of sentiment dictionary to judge the emotion contained in the text, but the size of sentiment dictionary is limited, and because of ignoring semantics, it is often impossible to get accurate classification. The machine learning method is based on the

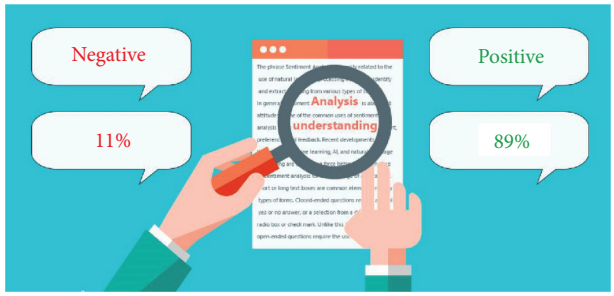


Figure 1: English text sentiment analysis.

idea of pattern classification to deal with this problem. Through the artificial design of features, the text is vectorized and input into various classifiers for classification [11–15].

Dictionary-based sentiment analysis is one of the classic methods. It treats the text as a collection of words and ignores the connections between words. This method uses grammar and common expressions to set syntactic rules, builds an emotional dictionary manually, finally weights emotional words through matching emotional dictionaries and words, and counts the scores to determine the emotional polarity of the text. At present, there are many emotional dictionaries, most of which are hand-labeled. Commonly used Chinese emotional dictionaries include HowNet emotional dictionaries. Thelwall et al. proposed the classic algorithm SentiStrength based on the sentiment dictionary, which adjusted the method of calculating the sentiment value of social networks and achieved good results. Some scholars think that sentiment analysis cannot be discussed without leaving the context, so they proposed an algorithm to explore the connection between modifiers and sentiment words, which is more effective in predicting the sentiment polarity of the text. Some scholars believe that ordinary sentiment dictionaries are the shortcomings of sentiment analysis. They supplemented the original dictionaries, seized most of the short content features in Weibo texts, and built a new dictionary on sentiment analysis with the help of new weighting rules algorithms. Although many researchers continue to improve and supplement emotional dictionaries and have made some progress, this method is always limited by the dictionary, it is impossible to include all emotional words, and it cannot adapt to the times. In addition, because it is not suitable for texts with implicit sentiment characteristics, the accuracy of this method has not been high when used in text sentiment analysis [16–19].

Sentiment analysis methods based on machine learning are usually not limited by dictionaries and are mainly used for model training. The general steps are as follows: prepare manually labeled text data for training an emotional classification model, optimize parameters, and finally predict the emotional tendency of unknown data through the model. Pang et al. used Naive Bayesian and support vector machine (SVM) algorithms for the first time in the field of sentiment analysis in 2002 and studied the related comments of movies as the object of sentiment classification. Significant progress has been made. When analyzing English reviews, Beineke et al. combined machine learning algorithms with manual text annotations and obtained excellent results. Then, Rao et al. proposed a model, the LDA topic generation model, which uses the bag-of-words method to recognize text or corpus. Mou et al. used three algorithms: Naive Bayes, K-nearest neighbor, and SVM in the model to judge the emotional tendency of English text. Dey et al. compared the two algorithms of Naive Bayes and K-nearest neighbors, and conducted experiments on hotel and movie review text sets [20–25]. The results showed that the effect of Naive Bayes is better. Mukras et al. proposed a method that can learn and mark part of speech from a text library. Mertiya et al. combined machine learning methods with dictionaries and performed sentiment analysis around comments on Facebook, providing a new idea of combining the two. Firstly, the features are extracted through the sentiment dictionary, and then, the text sentiment polarity is judged by Naive Bayes. In sentiment analysis, most of the methods based on machine learning are based on statistical theory, alleviating the problem of dictionaries that consume a lot of manpower and time and at the same time improving the accuracy of sentiment analysis, but there are still some shortcomings, such as the quality of data required for feature extraction—good and large numbers. In the current era of big data, its efficiency needs to be improved [26–29].

On the whole, machine learning methods perform better than rule methods. However, for complex English, traditional machine learning modeling methods cannot achieve satisfactory results. To this end, this paper proposes a text sentiment analysis method that integrates multiple features and constructs three features, which are based on the sentiment value feature of the dictionary, the expression feature, and the improved semantic feature, which are combined to build a text sentiment classification model.

2. Text Sentiment Analysis Combining Multiple Features

Due to the limited resources of English sentiment analysis and its unique complexity, it has become a challenging task to identify the sentiment of English comments. This paper proposes a multifeature fusion English text sentiment analysis method that combines machine learning and sentiment rules. The goal is to classify the sentiment of the existing review text, so as to find the user's evaluation information on products and topics. The emotional results are mapped to polarity. The flowchart of text sentiment analysis fused with multiple features is shown in Figure 2.

2.1. Building an Emotional Dictionary. The sentiment dictionary is a set of color markers with sentiment polarity. It is an indispensable part of the text sentiment analysis task. Generally, the more complete the sentiment dictionary, the more accurate the recognition results will be. In order to get better recognition results, the currently widely used emotional dictionaries (such as HowNet, NTUSD, and TSING) are integrated and expanded,

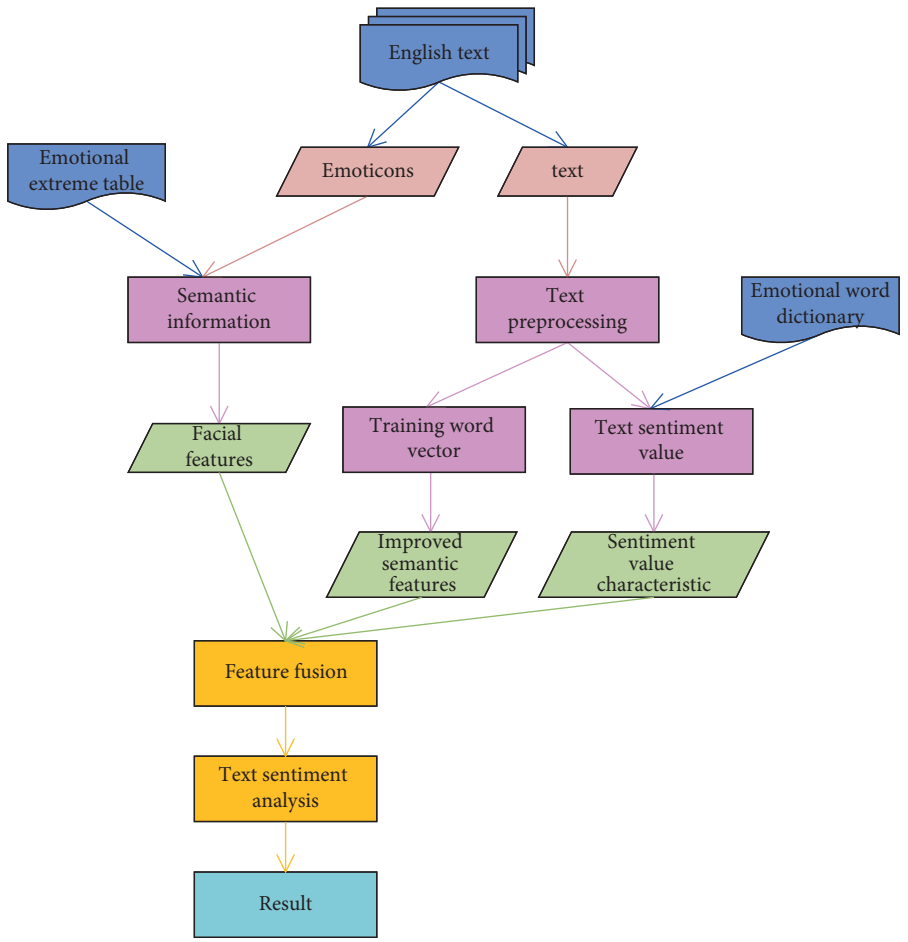


Figure 2: Flowchart of text sentiment analysis fused with multiple features.

and the basic emotional words, expression emotional words, degree adverbs, negative words, and transitional conjunctions have been established—comprehensive emotional dictionary. In addition, a dictionary of online emotional words has also been established. For the emergence of new words on the Internet, many documents have studied the method of expanding the emotional dictionary based on machine learning and have achieved certain results. However, due to the problems of word segmentation and candidate word selection, algorithms cannot be used to get a good processing effect for all kinds of network terms that are emerging in an endless stream. Therefore, based on the network term dictionary crawled by Zhihu, other network emotion words were sorted and expanded, and a network emotion dictionary with 726 emotion words was constructed. Dictionary-based sentiment value features refer to constructing specific rules based on sentiment dictionary and modifier dictionary, matching the sentiment words and modifiers contained in the text, and then performing weighting calculations to obtain sentiment value features as the expression form of text emotion. To calculate the sentiment value of the text, the formula is as follows:

$$\text{Score} = \sum_{j=1}^{m} \text{base} \times \prod_{i=1}^{n} \text{weight}_i. \quad (1)$$

Among them, m is the total number of emotional words contained in the text, n is the number of modifiers of a certain emotional word, base is the basic score, and weight is the degree adverb or the weight of the negative word.

2.2. Text Preprocessing. English text usually contains a strong personal style and personal emotional color, and the expression content is rich. In addition to being irregular, the grammar is basically biased toward daily life and colloquialism, but also contains a large number of irregular language, typos, links and expressions, symbols, etc., so it needs to be preprocessed before proceeding with the task of text sentiment analysis. In order to improve the efficiency of text sentiment analysis, the first step is to filter out URLs, tags, and irregular language and remove stop words. In the stage of text preprocessing, word segmentation is one of the very important components. Because the comment text has obvious colloquial characteristics and contains a large number of new words on the Internet, the effect of using general word segmentation tools is not very satisfactory. Therefore, the English word segmentation system ICTCLAS developed by the Chinese Academy of Sciences that can be added to user-defined dictionaries is used to segment the comment text to achieve better word segmentation effect.

2.3. Dictionary-Based Emotional Rule Classification Method. The most classic is to accumulate the sentiment words to obtain the sentimental value of the text. The formula is as follows:

$$S = \sum_{i=1}^{n} Sw_i. \tag{2}$$

Among them, Swi is the polarity of the i-th emotional word and n is the total number of emotional words.

According to formula (2), the polarity of all sentiment words is superimposed, and the sentiment tendency value of the text is judged according to the finally obtained value. However, it is not only the emotional words that determine the emotional polarity in the text. Others such as negative words, degree adverbs, and language structure will all have a certain impact on the emotional tendency. Aiming at the shortcomings of the classic methods, a dictionary-based emotional rule classification method is proposed. Since the review text is generally short, first, each clause in the text is used as a unit, the emotion calculation formula (3) obtained by the emotion rule method based on the emotion dictionary is calculated for each unit, and finally all the units are calculated. The scores are superimposed to get the emotional orientation of the entire review text.

$$S_{\text{unit}} = \sum_{i=1}^{n} \left(K * Pw_i * \prod_{j=1}^{m} \text{mod}_j \right). \tag{3}$$

Among them, n represents the total number of emotional words in the text; Pw_i represents the extreme value of the i-th emotional word; m represents the number of words that modify the i-th emotional word; modj represents the weight of its corresponding modifier; and k represents the strengthening and weakening coefficient, to avoid sentiment analysis bias caused by subject confusion. In various algorithms of text sentiment analysis tasks, often due to the lack of referential judgment, the sentiment polarity obtained is not the judgment of the subject, and the results are biased. The overall architecture of the MFCNN model is shown in Figure 3.

Extract the emoticons in the text and calculate the extreme value of text emotion as follows:

$$\text{Score} = \frac{\sum_{i=1}^{m} F(e_i, \text{pos})}{m} + \frac{\sum_{j=1}^{m} F(e_j, \text{neg})}{n}, \tag{4}$$

where m and n are the number of positive emoticons and negative emoticons in the text, e is emoticons, pos and neg are the extreme value tables of positive and negative emoticons, and the function of F is to take out the scores of corresponding emoticons in the extreme value. The x with different y is shown in Figure 4.

It is convenient to intuitively understand the relationship between the number of expressions in the text and the extreme tendency of the text, and the cumulative distribution function (CDF) is introduced. The definition formula is as follows:

$$F_X(x) = P(X \le x). \tag{5}$$

2.4. Improved Semantic Features. The text word vector is regarded as the semantic feature of the text, because it contains the semantic information of the word, it is regarded as the semantic feature of the text. The Word2vec model is used to convert the text into word vectors, which alleviates the problems of sparse matrix and excessive dimension, and retains the sequence information of the words in the text, but omits the different importance of different words to the text. The TF-IDF algorithm just solves this problem, so the TF-IDF and Word2vec are combined, and the text word vector trained by the model is called the semantic feature of text improvement. It combines the advantages of the two, not only retains the sequence information of the words in the text, but also gives different weights to different words in the text. Assuming a text di, the number of words after word segmentation is M, and the word vector dimension is N; the text is expressed as follows:

$$d_i = < w_1, w_2, \cdots, w_M >. \tag{6}$$

The word vector is generated through the Word2vec model. The text contains multiple words, and each word has its corresponding word vector. Join them to obtain the $M \times$ N-dimensional vector matrix of the text:

$$G(d_i) = \text{matrix}(M \times N). \tag{7}$$

Multiplying with the weight matrix is the vector matrix obtained by the improved Word2vec:

$$W\_G(d_i) = \text{matrix}(M \times N). \tag{8}$$

The expression formula is as follows:

$$\begin{aligned} G(d_i) &= \{W2v(w_1), W2v(w_2), \cdots, W2v(w_M)\}, \\ W\_G(d_i) &= \{\text{weight}(w_1)W2v(w_1), \cdots, \text{weight}(w_M)W2v(w_M)\}. \end{aligned} \tag{9}$$

Among them, each vector $W2v(wi)$ in the G(di) vector matrix is the word vector of the word wi in the text, which is obtained by training the Word2vec model; each vector in the $W - G(di)$ vector matrix weight $(wi)v(wi)$, where weight (wi) is the weight value of the word wi calculated by the TF-IDF algorithm; and multiplying weight (wi) and $W2v(wi)$ is the word vector of the improved Word2vec, which is the word vector of each word in the text. The composed text vector matrix $W - G(di)$ is used as the improved semantic feature of this article. The predicted value is compared in Figure 5.

2.5. Method Based on Machine Learning. The classification method based on machine learning regards sentiment analysis as a pattern classification problem and establishes a classification model to judge sentiment polarity. First, the machine learning method needs to label the text, use it as a training set, then extract features to train the classifier, and finally test the test corpus to get the classification result. Text

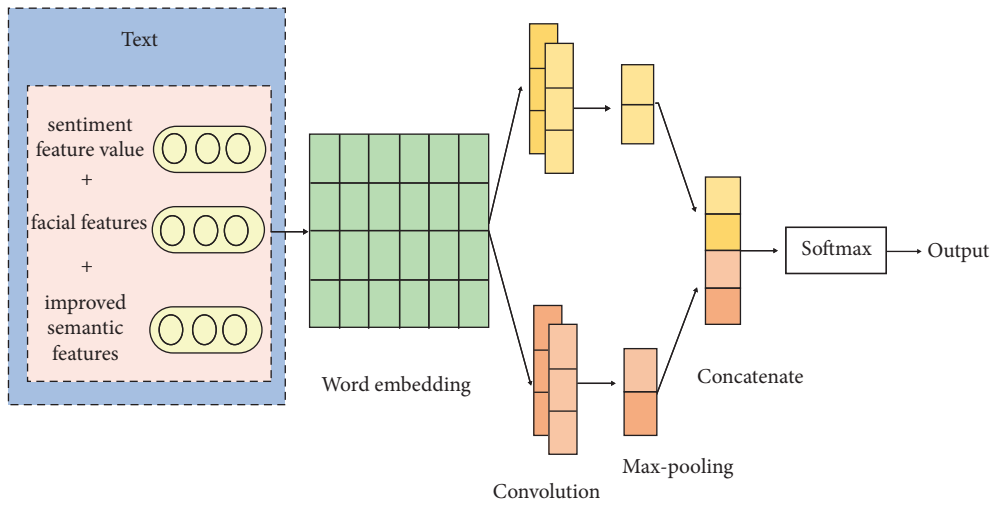


Figure 3: The overall architecture of the MFCNN model.

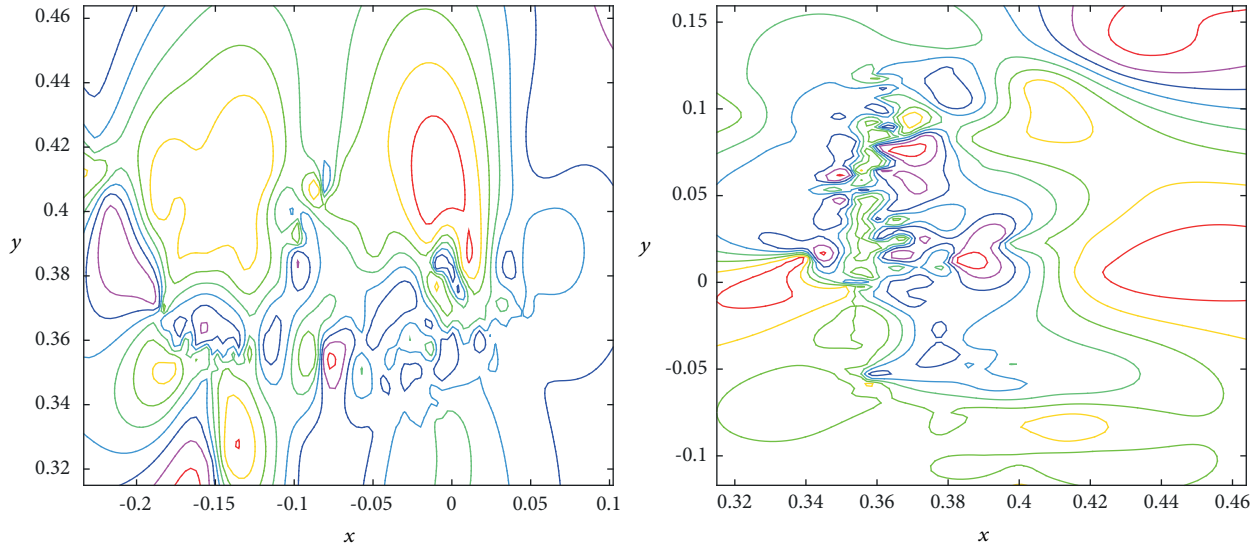


Figure 4: y vs. x.

feature selection is a key step of machine learning, which determines the accuracy of sentiment classification. Three types of features are selected in the article: unigram features, syntactic features, and dependent word collocation features. Among them, the syntactic feature is the feature of the research component and the order of arrangement. Considering that the phrase structure can reduce sentence ambiguity, the binary word (bigram) and its combined part-of-speech tag are added to the feature set as its feature; the dependency relationship feature is from the dependency parse tree. The dependence relationship identification obtained plays an important role in the labeling of emotional category information, and it can save the information directly related to emotional words and other hidden information. The amplitude variation is shown in Figure 6.

From the above analysis steps, three basic feature templates of the machine learning method can be obtained. In order to avoid the problem of the degradation of the classifier effect due to the large dimensionality of the original feature space, the feature selection method of information gain (IG) is used to reduce the dimensionality of the original feature space to select the corresponding features. The formula is as follows:

$$\begin{aligned} \mathrm{IG}(T) &= H(C) - H(C|T), \\ H(C) &= \sum_{i=1}^{n} P(C_i)\log_2(C_i). \end{aligned} \tag{10}$$

Among them, $P(C_i)$ represents the probability of category C_i.

2.6. Text Sentiment Analysis Method Fused with Multiple Features. The algorithm that combines machine learning methods and rule methods has attracted the attention of many researchers. For example, Qiu et al. used the dictionary

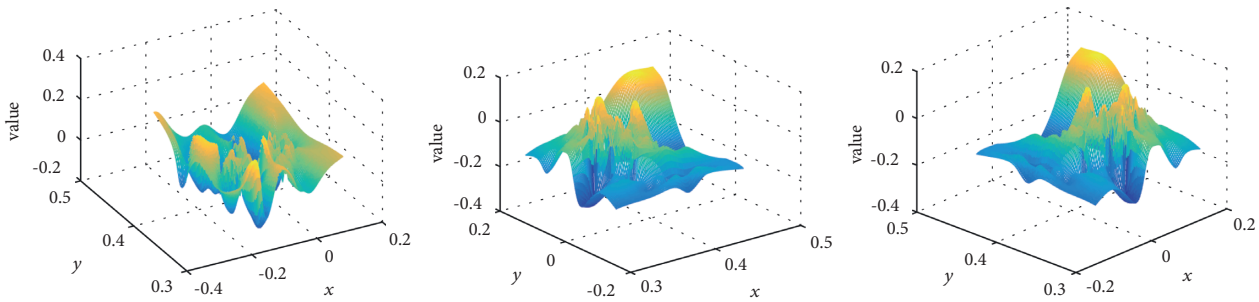


Figure 5: Predicted value comparison.

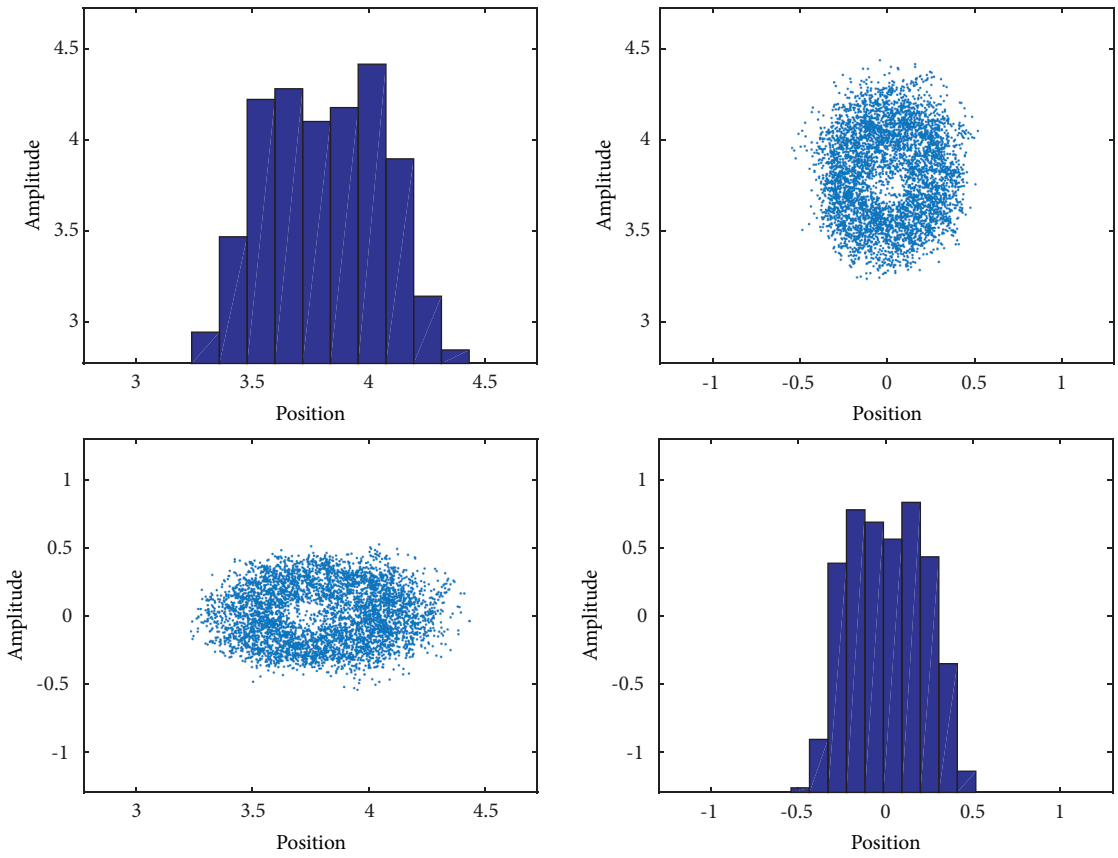


Figure 6: Amplitude in different positions.

classification results as the training corpus of the classification model to form a hierarchical iterative classification framework; Mohammad et al. accumulated emotional words and ended words—the polarity as a feature. Inspired by the predecessors, this paper proposes a multifeature fusion classification algorithm that combines machine learning and emotional rules. As a necessary step of the machine learning and emotional rule fusion method, after calculating the emotional score based on the improved emotional rule method, the effective information is extracted and expanded to integrate with machine learning features. Four characteristics are extracted in the article: emotional word score, the ratio of the number of positive/negative emotional words, the ratio of the number of enhancements to the number of weakenings, and the ratio of the number of praise/derogation emotional sentences, which are normalized and expanded to machine learning feature templates. Train the SVM classifier, and then use the test corpus to test. Through the above process, a multifeature fusion text classification method combining machine learning method and dictionary-based emotional rule method is realized, and multiple effective emotional information extracted from the rule algorithm is extended to the vector space, so that the machine learning algorithm can make full use of the characteristics of the rules and learn more emotional knowledge. The prediction is compared in Figure 7.

Emotional features play an important role in sentiment classification tasks. Its extraction is closely related to the

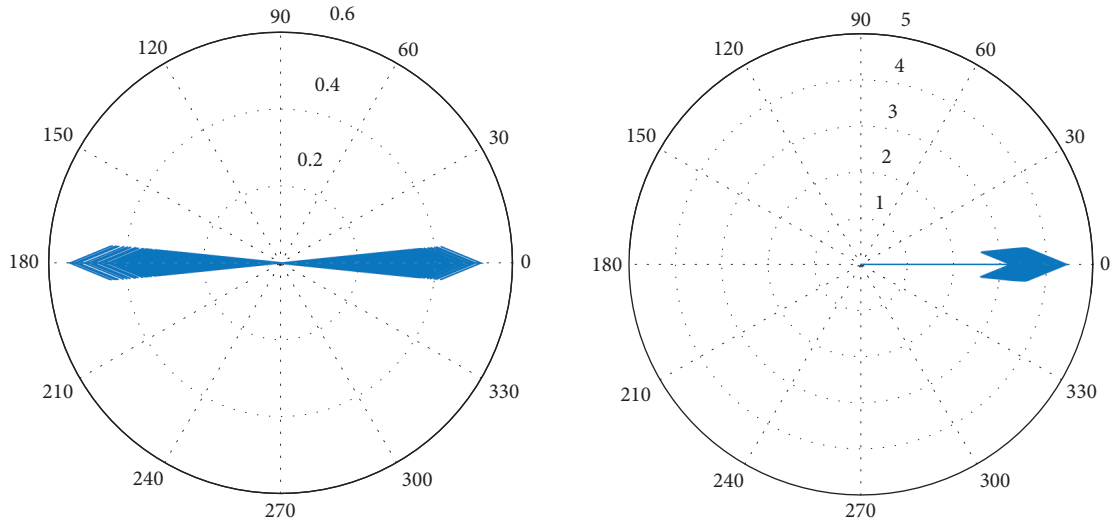


Figure 7: Prediction.

quality of the classification results. This paper constructs three kinds of emotional features: based on dictionary-based emotional value features, expression features, and improved semantic features. The feature-based text sentiment analysis method is to splice multiple features to form a multifeature vector matrix, which is used as the input of the convolutional neural network. This method can extract multidimensional sentiment information.

First, divide the text, store the text part in Dt, and store the expression part in De. Perform text preprocessing on Dt, and combine the sentiment dictionary and modifier dictionary to calculate the sentiment value characteristics of the text. Dt is trained through the improved Word2vec model to obtain the text word vector, which constitutes an improved semantic feature. De combines the emotional extreme value table of emoticons to calculate the emotional extreme value of the expression, plus the number of appearances of the expression and semantic information, which together constitute the expression feature. The three characteristics are merged to perform text sentiment analysis. TextCNN is one of the popular deep learning models. It is adjusted on the basis of CNN to make the TextCNN model more suitable for extracting text features and is often used in sentiment analysis. This chapter uses it as the core model and proposes an emotional classification model MFCNN based on multifeature fusion, which converts different features into corresponding vectors, uses splicing for feature fusion, constructs a multifeature vector matrix, and inputs it into the text convolutional neural network. Finally, get the classification result.

3. Experiment and Analysis

Research on sentiment analysis often requires a lot of text as support to facilitate training models. However, the amount of relevant text data on the Internet is not many and the quality is low, so this article crawled 10,000 texts through crawlers and labeled them as positive emotions or negatives according to their sentiment tendencies. Emotions constitute a two-category data set. Among them, there are 5497 positive sentiment texts and 4503 negative sentiment texts, and the sentiment tendencies are roughly balanced between positive and negative. The overall process of the crawler is roughly as follows: randomly take out an account from the account pool of the database to simulate the login of a Weibo user, obtain cookie information, so that the website can identify the user's identity, obtain the target URL resource, use the Python Requests Library to process HTTP requests, return a response object, then use the BeautifulSoup4 library to parse and process HTML, and finally use regular expressions to extract text data and save it to the database. The evaluated data are shown in Figure 8. As can be seen, the data and value vary in each condition, which shows the validation of the proposed method.

The text preprocessing is divided into three steps: data cleaning, Chinese word segmentation, and stop word removal, which are described in detail below. The first step is data cleaning. Data cleaning is to remove characters and data that have nothing to do with the text content, such as URL links, forwarding symbols //, topic symbols #, designated user symbols @, and other information in the text. This information has nothing to do with the content of the text, but it may interfere with the results of sentiment analysis and affect subsequent word segmentation. For these data, this article uses regular expressions to delete them all. The method of replacing Weibo emoticons is adopted. In order to facilitate the construction of emoticon features, the emoticons contained in the text are converted into "[emoticons]" format when crawling the Weibo text. For example, the emoticon that represents disappointment is transformed into "[disappointment]" in the text, and the emoticon is placed in "[ ]" to distinguish it from the text.

Text vectorization is to convert each word into a vector after preprocessing and dividing the word, and then, each

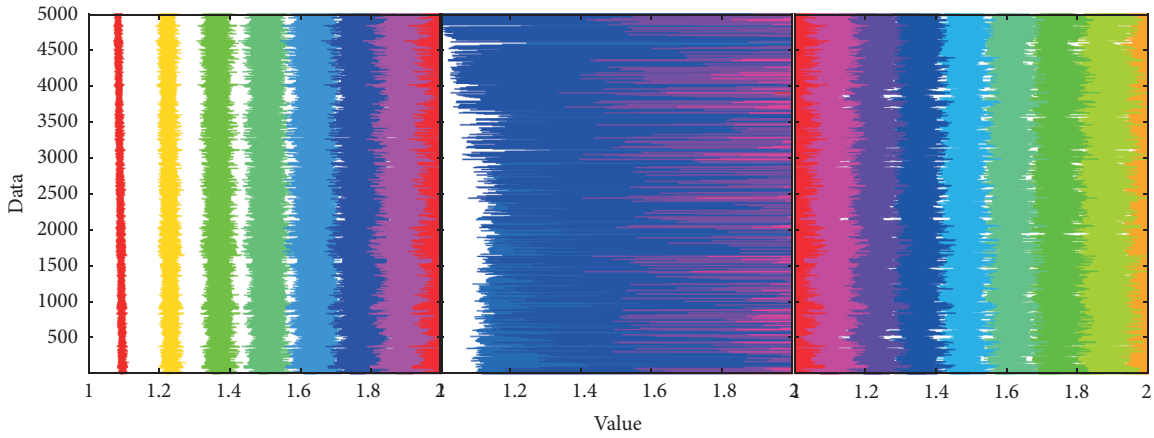


Figure 8: Evaluated data.

word vector is formed into a vector matrix according to the order of the words in the text. In this way, mapping the words to the vector space preserves its semantic information. Text vectorization is the cornerstone of text research. Whether or not the word vector can be expressed correctly will affect the judgment of text orientation. This paper uses the improved Word2vec model to vectorize the text and compares the final sentiment classification results. The specific configuration of the experiment is as follows: the processor is Intel(F)Core(TM)i5-6500CPU@3.2 GHz; the memory is 8 GB; the programming platform is Eclipse; the development language is Java; the database is SqlSever2008.

In order to verify the effectiveness of the text sentiment analysis method of fusion of multiple features, this section designs 7 models to conduct comparative experiments on the data set. The 7 models use different feature construction methods to form a vector matrix as the input of TextCNN and compare the final text—the result of sentiment classification. These seven models are as follows: (1) CNN model: Word2vec model trains word vectors, that is, the original semantic features, and then inputs them to TextCNN for text sentiment classification. (2) TCNN model: The word vector is trained by the Word2vec model weighted based on the TF-IDF algorithm, that is, the improved semantic features, and then input to TextCNN for text sentiment classification. (3) SCNN model: On the basis of the CNN model, fusion of emotional value features is based on dictionary, that is, the fusion of original semantic features and emotional value features. (4) ECNN model: On the basis of the CNN model, expression features are fused; that is, the original semantic features and expression features are fused. (5) TSCNN model: On the basis of the TCNN model, fusion of emotional value features is based on dictionary, that is, the fusion of improved semantic features and emotional value features. (6) TECNN model: On the basis of the TCNN model, the expression features are integrated; that is, the improved semantic feature and the expression feature are fused. (7) MFCNN model: Fusion of multiple features, fusion of dictionary-based emotional value features, expression features, and improved semantic features form a multifeature vector matrix, which can extract multidimensional emotional information as the input of TextCNN for emotional classification.

The following conclusions can be drawn: (1) Compared with the CNN model with only original semantic features, the accuracy of the TCNN model with improved semantic features is increased by 2.19%, indicating that the combination of TF-IDF algorithm improves the keyword in the text. The weight is beneficial to the improvement of sentiment classification performance and verifies the effectiveness of the improved semantic features. (2) The ECNN model, which integrates expression features on the basis of original semantic features, has an accuracy rate of 1.98% higher than that of the CNN model, which shows that emojis increase the effect of indicating emotions, and also proves the necessity of adding expression features. In summary, it can be seen that the improved semantic features and expression features do contain more emotional information that is conducive to classification for the self-built data set. (3) Compared with the CNN model, the accuracy of the SCNN model fused with emotional value features based on the original semantic features is increased by 0.53%, and the improvement effect is not very significant. It can be seen that the fusion of a single emotional value feature has limited improvement for the model. (4) The TSCNN model that combines improved semantic features and emotional value features has an accuracy rate of 2.42% higher than that of the CNN model. The TECNN model that combines improved semantic features and expression features has an accuracy rate of 3.17% higher than that of the CNN model. The accuracy of classification is improved. It has been further improved; the accuracy rate of the MFCNN model that finally fused the three features is 0.59% to 3.76% higher than other comparison models, indicating that the MFCNN model can learn more dimensional emotional information of the text from the vector matrix fused with multiple features. Prove the feasibility and effectiveness of this method.

Compared with the CNN model, the MFCNN model fused with multiple features has nearly 5% and 4% improvement in accuracy and F1 value, respectively, and an improvement in recall rate of 2%. Among them, the accuracy of the TCNN model and the MFCNN model has been significantly improved, and it can be seen that the improved semantic features

perform better. Regarding the recall rate, the ECNN model has achieved the best results, which is 4% higher than the CNN model. On the F1 value, comparing the results of the CNN model and the SCNN model, and the TECNN model and the MFCNN model, it can be seen that the individual emotion value feature has a small effect on the improvement of the F1 value, but the multifeatured model has achieved excellent results. This paper proposes a Word2vec model training word vector based on the TF-IDF algorithm. In order to prove its effectiveness, it is compared with the traditional Word2vec model in three sets of experiments.

4. Conclusion

In natural language processing, text sentiment analysis is one of the important branches. It refers to the use of text mining and other technologies to extract attitudes, opinions, and other information from texts containing emotional information for analysis. Traditional sentiment analysis methods can be roughly divided into two categories: one is dictionary-based methods, and the other is machine learning-based methods. The former relies on the quality of the sentiment dictionary, while the latter relies on a large amount of high-quality data, so both have certain limitations. In text sentiment analysis research, word-level and sentence-level sentiment information extraction is a basic research task and has important research value. Through research, it is found that domain knowledge and context are two important factors influencing the extraction of emotional information. To this end, this paper proposes a text sentiment analysis method that integrates multiple features and constructs three features, which are based on the sentiment value feature of the dictionary, the expression feature, and the improved semantic feature, which are combined to build a text sentiment classification model. Aiming at the colloquial, irregular, and diverse features of English social media texts, this paper proposes a multilevel feature representation method. The sentiment classification experiments on English text show that the multilevel features proposed in this paper can effectively improve the F1_macro and accuracy of multiple model classifications. Compared with the existing research, the model in this paper improves the effect the most obvious.

Data Availability

The data set can be accessed upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

[1] Y. Yu, C. Yang, Q. Deng, T. Nyima, S. Liang, and C. Zhou, "Memristive network-based genetic algorithm and its application to image edge detection," *Journal of Systems Engineering and Electronics*, vol. 32, no. 5, pp. 1–9, 2021.

[2] Y. Ishida and S. Hashimoto, "Asymmetric characterization of diversity in symmetric stable marriage problems: an example of agent evacuation," *Procedia Computer Science*, vol. 60, no. 1, pp. 1472–1481, 2015.

[3] P. Zoha and R. Kaushik, "Image edge detection based on swarm intelligence using memristive networks," *IEEE Trans. on CAD of Integrated Circuits and Systems*, vol. 37, no. 9, pp. 1774–1787, 2018.

[4] J. Pais, "Random matching in the college admissions problem," *Economic Theory*, vol. 35, no. 1, pp. 99–116, 2018.

[5] J. J. Jung and G. S. Jo, "Brokerage between buyer and seller agents using constraint satisfaction problem models," *Decision Support Systems*, vol. 28, no. 4, pp. 291–384, 2020.

[6] Y. Liu and K. W. Li, "A two-sided matching decision method for supply and demand of technological knowledge," *Journal of Knowledge Management*, vol. 21, no. 3, JKM-05-2016-0183, 2017.

[7] J. Byun and S. Jang, "Effective destination advertising: matching effect between advertising language and destination type," *Tourism Management*, vol. 50, no. 10, pp. 31–40, 2015.

[8] A. N. Nagamani, S. N. Anuktha, N. Nanditha, and V. K. Agrawal, "A genetic algorithm-based heuristic method for test set generation in reversible circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 37, no. 2, pp. 324–336, 2018.

[9] C. Koch and S. P. Penczynski, "The winner's curse: c," *Journal of Economic Theory*, vol. 174, pp. 57–102, 2018.

[10] C. K. Karl, "Investigating the winner's curse based on decision making in an auction environment," *Simulation & Gaming*, vol. 47, no. 3, pp. 324–345, 2016.

[11] D. Ettinger and F. Michelucci, "Creating a winner's curse via jump bids," *Review of Economic Design*, vol. 20, no. 3, pp. 173–186, 2016.

[12] J. A. Brander and E. J. Egan, "The winner's curse in acquisitions of privately-held firms," *The Quarterly Review of Economics and Finance*, vol. 65, pp. 249–262, 2017.

[13] Z. Palmowski, "A note on var for the winner's curse," *Economics/Ekonomia*, vol. 15, no. 3, pp. 124–134, 2017.

[14] B. R. Routledge and S. E. Zin, "Model uncertainty and liquidity," *Review of Economic Dynamics*, vol. 12, no. 4, pp. 543–566, 2009.

[15] D. Easley and M. O'Hara, "Ambiguity and nonparticipation: the role of regulation," *Review of Financial Studies*, vol. 22, no. 5, pp. 1817–1843, 2019.

[16] P. Klibano, M. Marinacci, and S. Mukerji, "A smooth model of decision making under ambiguity," *Econometrica*, vol. 73, no. 6, pp. 1849–1892, 2005.

[17] Y. Halevy, "Ellsberg revisited: an experimental study," *Econometrica*, vol. 75, no. 2, pp. 503–536, 2017.

[18] D. Ahn, S. Choi, D. Gale, and S. Kariv, "Estimating ambiguity aversion in a portfolio choice experiment," *Working paper*, vol. 5, no. 2, pp. 195–223, 2019.

[19] T. Hayashi and R. Wada, "Choice with imprecise information: an experimental approach," *Theory and Decision*, vol. 69, no. 3, pp. 355–373, 2010.

[20] K. Zima, E. Plebankiewicz, and D. Wieczorek, "A SWOT analysis of the use of BIM technology in the polish construction industry," *Buildings*, vol. 10, no. 1, p. 16, 2020.

[21] P. Sun, B. Liu, and T. Sun, "Injury status and strategies of female 7-a-side rugby players in Anhui Province," *Sports Boutique*, vol. 38, no. 03, pp. 72–74, 2019.

[22] P. Guild, M. R. Lininger, and M. Warren, "The association between the single leg hop test and lower-extremity injuries in female athletes: a critically appraised topic," *Journal of Sport Rehabilitation*, vol. 30, no. 2, pp. 1–7, 2020.

[23] U. G. Inyang, E. E. Akpan, and O. C. Akinyokun, "A hybrid machine learning approach for flood risk assessment and classification," *International Journal of Computational Intelligence and Applications*, vol. 19, no. 2, Article ID 2050012, 2020.

[24] Q. Liu, S. Du, B. Wyk, and Y. Sun, "Double-layer-clustering differential evolution multimodal optimization by speciation and self-adaptive strategies," *Information Sciences*, vol. 545, no. 1, pp. 465–486, 2021.

[25] H. R. Medeiros, F. D. Oliveira, H. F. Bassani, and A. Araujo, "Dynamic topology and relevance learning SOM-based algorithm for image clustering tasks," *Computer Vision and Image Understanding*, vol. 179, no. FEB, pp. 19–30, 2019.

[26] Y. Deng, D. Huang, S. Du, G. Li, and J. Lv, "A double-layer attention based adversarial network for partial transfer learning in machinery fault diagnosis," *Computers in Industry*, vol. 127, Article ID 103399, 2021.

[27] J. J. Chan, K. K. Chen, S. Sarker et al., "Epidemiology of Achilles tendon injuries in collegiate level athletes in the United States," *International Orthopaedics*, vol. 44, no. 3, pp. 585–594, 2020.

[28] W. Li, G. G. Wang, and A. H. Gandomi, "A survey of learning-based intelligent optimization algorithms," *Archives of Computational Methods in Engineering*, vol. 28, pp. 3781–3799, 2021.

[29] G. G. Wang, A. H. Gandomi, A. H. Alavi, and D. Gong, "A comprehensive review of krill herd algorithm: variants, hybrids and applications," *Artificial Intelligence Review*, vol. 51, no. 1, pp. 119–148, 2019.

Hindawi
Scientific Programming
Volume 2022, Article ID 7700467, 10 pages
https://doi.org/10.1155/2022/7700467

Research Article

Generative Adversarial Network-Based Short Sequence Machine Translation from Chinese to English

Wenting Ma,[1] Bing Yan,[1] and Lianyue Sun [2]

[1]*English Major, Art and Law Department, Hua Xin College of Geo University, Shijiazhuang, Hebei 050700, China*
[2]*Shijiazhuang Posts and Telecommunications Technical College, Shijiazhuang, Hebei 050020, China*

Correspondence should be addressed to Lianyue Sun; sunly@sjzpc.edu.cn

Received 19 November 2021; Accepted 18 December 2021; Published 28 January 2022

Academic Editor: Baiyuan Ding

With the acceleration of economic globalization, the economic contact, information exchange, and financial integration between countries become more and more frequent. In this context, the communication between different languages is also closer, so accurate translation between languages is of great significance. However, existing methods give little thought to short sequence machine translation from Chinese to English. This paper designs a generative adversarial network to solve the above problem. First, a conditional sequence generating adversarial net is constructed, which includes two adversarial submodels: a generator and a discriminator. The generator is designed to generate sentences that are difficult to distinguish from human-translated sentences, and the discriminator is designed to distinguish the sentences generated by the generator from human-translated sentences. In addition, static sentence-level BLEU values will be used as reinforcement targets for the generator. During training, both dynamic discriminators and static BLEU targets are used to evaluate the generated sentences, and the evaluation results are fed back to the generator to guide the generator's learning. Finally, experimental results on English-Chinese translation dataset show that the translation effect is improved by more than 8% compared with the traditional neural machine translation model based on recurrent neural network (RNN) after the introduction of generative adversative network.

1. Introduction

As one of the important tools for human communication, language is the fundamental ability for human beings to distinguish themselves from other creatures. There are more than 5,600 languages in the world. The diversity of languages promotes the development of cultural diversity, but it also sets up barriers for human communication in different regions [1–3]. With the acceleration of globalization, the contradiction between cross-language communication and language gap becomes more and more obvious. The existing human resources and translation capacity are insufficient to cope with the translation needs of large-scale texts and multiple language pairs in the future. Therefore, machine translation (machine translation) has the advantages of low cost, fast speed, and no language restriction. Translation technology is regarded as the most effective way to break the barrier of language translation. Machine translation (MT) is a process of automatic translation from one language (source Language) to another language (target language) using computers. In recent years, machine translation technology has developed vigorously and become one of the research hotspots of natural language processing (NLP) task [4, 5].

After more than 70 years of development, machine translation has developed from rule-based machine translation to instance-based machine translation and then to statistics-based machine translation (SMT) [6] and to the present neural network-based neural machine translation (NMT). With the rapid development of deep learning, the performance of neural machine translation has significantly surpassed. It has not only become the hotpot research method of language translation, but also the core technology of commercial online machine translation systems such as Google and Baidu. Neural machine translation models generally adopt encoder-decoder framework. The encoder

model translates the source sentences and encodes them into a set of hidden layer vector representations with fixed dimensions. The decoder generates the word sequence of the target end word for word according to the hidden layer vector output by the encoder. Encoders and decoders of neural machine translation models can be realized by different network structures, which are divided into self-attention network (SAN), convolutional neural networks (CNN), and recurrent neural networks (RNN) according to the characteristics of different network topologies. Compared with statistics-based and rule-based machine translation methods, the neural machine translation model does not require such steps as word extraction from translation rules, word alignment, and ordering but completely relies on the neural network to reflect the relationships between source language and target language automatically, which greatly simplifies the complexity of the model [7].

Although the results of neural machine translation have been enhanced largely under the condition of large-scale training corpus and powerful computing power, there are still many problems to be solved in neural machine translation, such as out of vocabulary (OOV), overtranslation, undertranslation [8], and exposure bias [9]. These problems become the development bottleneck of neural machine translation and restrict the further improvement of translation performance.

2. Related Works

Based on the coverage idea in statistical machine translation, coverage mechanism is introduced into neural machine translation model in [10], and coverage vector (CV) is used to store the historical translation in the decoding process and integrate it into the calculation process of attention weight, so as to guide the attention mechanism to give more resource to the untranslated words and reduce the weight of translated words. As for the phrase-based translation model, literature [11] marks the corresponding source language phrase as translated when adding the translation results of candidate phrases into the output sequence, so as to ensure that each source language phrase will not be repeated or omitted in the translation process. However, because there is no similar mechanism for explicitly storing translation history information in the neural machine translation model, and all source words are involved in the prediction of target words, overtranslation and missing-translation are inevitable. Literature [12] uses full coverage embedding vector to represent the translation degree of source language words in the decoding process and reduces the role of translated words in future decoding by reducing the encoding vector of translated words. In literature [12], it combines a recurrent attention mechanism and conditioned decoder to provide more ordering information in the translation process, thus reducing the phenomenon of repeated translation. In [13], the authors use two circulating neural networks to store the past information and future information in translation process, respectively, and use this information to guide the attention mechanism and decoding state. These methods can alleviate the phenomenon of overtranslation and missed translation to a certain extent but cannot completely eliminate the problem [14].

Since the neural machine translation model often selects sentences with high probability but short length as translation results, coverage can also be used as an evaluation index to screen translation results [15]. In [16], coverage penalty is introduced into the beam search algorithm to make the model consider the probability of sentence generation and fidelity to the original text at the same time when selecting the translation, so as to avoid the bias towards short sentences. On this basis, literatures [17–19] introduced the detection of coverage into each cluster search step and improved the calculation method of coverage score (CS), making it suitable for a variety of mapping relations between source language and target language. These methods select better translation results by improving search evaluation methods without changing the structure of the neural machine translation model [20, 21].

In order to give consideration to the fluency and fidelity of the target text, the context gate structure is introduced in [22, 23] to dynamically control the proportion of the influence of source language and target language context on the generation of words in the target language during decoding. Combined with the overlay mechanism, this method can improve the coverage of the translation results to the source language and the fluency of the sentence, but the target context does not fully utilize the generated translation information. Literature [24] has modeled the structural relations of all words, so that the neural machine translation model can make better use of the context features of source language and target language. In addition, literature [25] studies the features of omitted words in source languages, finds that words with higher translation entropy are more likely to be omitted, and proposes to use coarse-to-fine framework to improve the translation quality of sentences and words and reduce the number of missed translation of words with high entropy.

From the previously mentioned analysis, we know that the previously mentioned methods have alleviated the problem of overtranslation and overtranslation in NMT to some extent, but these problems are still unavoidable due to the soft alignment of attention mechanism and the imperfection of coverage mechanism in NMT model word-for-word prediction. On the other hand, no scholar has applied English-Chinese short sequence machine translation till now, so the research here is still a blank, which has a great theoretical research and practical application value [26–28].

The contribution of this paperis the GAN model, which is first used to solve the problem of short sequence machine translation from Chinese to English. In this paper, a survival adversarial network-based machine translation method is proposed. Specifically, BLEU value is used to generate D for network, and BLEU value is established as feedback to enhance the training of generator G. So, generator G can generate the translation result which is closer to the real sentence, thus improving the accuracy of translation. In future work, we may try to combine more neural machine translation models with generative adversarial networks or use multiadversarial network frameworks to experiment

with different parameters and construct different reinforcement feedback to improve translation effects.

This paper consists of five parts. The first and second parts give the research status and background. The third part is the short sequence machine translation by generative adversarial network. The fourth part shows the experimental results and analysis. The experimental results of this paper are introduced and compared and analyzed with relevant comparison algorithms followed. Finally, the sixth part summarizes the full paper. The study of this paper gives the user portrait system of operators and gives the realization scene of precision marketing. The paper can complete the analysis function according to the data provided by operators to form user portraits, which has the application value of precision marketing of operators.

3. Short Sequence Machine Translation by Generative Adversarial Network

3.1. GAN Model and Transformer Model. The process of short sequence machine translation by generative adversarial network is divided into two parts, and the overall architecture of the model is shown in Figure 1 [29]. The left half is made up of generator G and discriminator D, where G is the neural machine translation model, which generates target sentences. D discriminates between the sentences generated by G and the artificial translation sentences and generates the feedback results. The right part carries out strategy gradient training for G, and the final feedback is provided by D and Q, where Q is BLEU value.

Another type of encoding and decoding similar to the GAN model is the transformer model, which is shown in Figure 2 [30]. And the best translation effect of current neural machine translation is achieved. Similarly, a set of source language sequences X with input length m and target language sequences Y with corresponding length n are given. Transformer encoders are stacked by N identical network layers. Every network layer contains two subnetwork layers: the first layer is self-attention layer, and the second layer is a fully connected neural network. Residual network and layer normalization (LN) connections are used after each sublayer. Multihead attention network maps input values into various subspaces and then uses scaled dot product attention network to calculate context vectors of different spaces and splices these context vectors to the final output result.

Similar to the encoder, the transformer decoder also contains N isomorphic network layers, each containing three subnetwork layers: the first layer is self-attentional network because the decoder can only see the generated word information when decoding, so mask technology is used to shield the ungenerated word information. The second layer is multiattentional network, which models the hidden state of source language sentences and the hidden state of target language to generate the context vector of source language sentences. The third layer is a fully connected feedforward neural network, which also uses residual networks and hierarchically normalized connections after each sublayer. Since both encoders and decoders based on full attention networks do not consider location information, which is very important for language understanding and generation, the transformer model includes location coding in the input vector of the lowest level encoder and decoder. Position encoding can be fixed position encoding, relative position encoding, or learned position encoding.

3.2. Generator G. Generator G uses a recurrent neural network (RNN) based neural machine translation model RNN search, which consisted of two parts: the encoder and coder [31]. The encoder adopts bidirectional cyclic gate to control the unit, codes the input sequence formula $x = (x_1, \ldots, x_m)$, and calculates the hidden state of forward and backward propagation as follows:

$$\overrightarrow{h} = \left(\overrightarrow{h_1}, \ldots, \overrightarrow{h_m}\right), \tag{1}$$

$$\overleftarrow{h} = \left(\overleftarrow{h_1}, \ldots, \overleftarrow{h_m}\right), \tag{2}$$

where the final annotation vector h_p is calculated by $\overrightarrow{h_p}$ and $\overleftarrow{h_p}$ together.

The encoder uses recurrent neural network to predict target sequence $y = (y_1, \ldots, y_m)$. The current prediction of each word y_i is calculated based on the current hidden state s_i and the prediction y_{i-1} of the previous moment, as well as the context vector c_i, where, c_i is derived from the weighted sum of the annotated vectors h_p.

3.3. Discriminator D. The discriminator uses convolutional neural network structure to discriminate the results generated by the generator. Since the length of the sequence generated by the generator is variable, the sequence is filled with a fixed length of T. Source matrix $X_{1:T}$ and target matrix $Y_{1:T}$ are established, respectively, for source sequence and target sequence as follows:

$$X_{1:T} = x_1; x_2; \ldots; x_T, \tag{3}$$

$$Y_{1:T} = y_1; y_2; \ldots; y_T. \tag{4}$$

When x_t and $y_t \in R^k$ are both the k dimensional word vectors, the convolution kernel $W_1 \in R^{l \times k}$, and the convolution calculation formula is shown as follows:

$$c_{ji} = f\left(W_1^* X_{i:i+l-1} + b\right), \tag{5}$$

where W_1^* is the weight of the convolution kernel, $X_{i:i+l-1}$ is the vector matrix in the window from i to $i + l - 1$, b is the bias term, f is the activation function, and ReLu function is adopted in this paper.

After using different convolution kernels for convolution operation, we conducted the maximum pooling (Max) operation on feature vector c_{ji} of each convolution kernel, that is, proposing the maximum value of each feature vector, splicing the pooling result, and obtaining the feature vector of the source sequence, as follows:

$$c_x = [c_1, c_2, \ldots, c_{T-l+1}]. \tag{6}$$

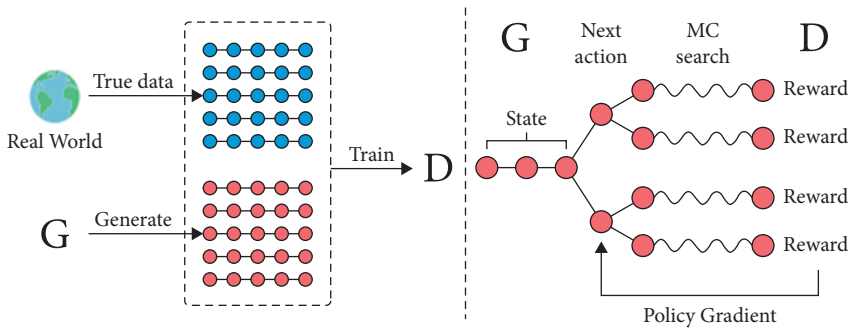


Figure 1: The process of short sequence machine translation by GAN.

Similarly, the feature vectors c_y of the target sequence are given by the matrix $Y_{1:T}$. The calculation of the probability of target sequence depends on c_x and c_y, which can be expressed as follows:

$$p = \varphi\left(V\left[c_x; c_y\right]\right), \tag{7}$$

where V is the parameter matrix, $[c_x; c_y]$ are converted into two-dimensional vectors, and φ is the softmax function.

3.4. Strengthening Target Feedback. BLEU value Q of bilingual translation quality assessment is used as the reinforcement target, and the guidance generator can generate higher BLEU value. Q is a static function and is not updated during training. We apply BLEU value as the specific target of generator and assign the sentence y_g generated by generator and the artificially translated sentence y_d. By calculating the n-gram accuracy of y_g, the feedback $Q(y_g, y_d)$ of target Q is obtained. The same as the output of discriminator D, $Q(y_g, y_d)$ also ranges from 0 to 1, which makes it easier for Q and D to merge.

3.5. Strategy Gradient Training. Since the goal of generator G is defined as maximizing the expected feedbacks from the initial state of the generated sequence. Formally, the objective function is as follows:

$$J(\theta) = \sum_{Y_{1:T}} G_\theta\left(Y_{1:T} \mid X\right) \cdot R_{D,Q}^{G_\theta}\left(Y_{1:T-1}, X, y_T, Y^*\right), \tag{8}$$

where θ is a parameter of generator G.

The discriminator provides N feedbacks for N sentences sampled, and the resulting feedbacks are the summation average of these feedbacks. For the target sentence with length T, the feedback calculation of y_T is shown as follows:

$$R_{D,Q}^{G_\theta}\left(Y_{1:t-1}, X, y_T, Y^*\right) = \begin{cases} \dfrac{1}{N}\displaystyle\sum_{n=1}^{N} \lambda\left(D\left(X, Y_{1:T_n}^n\right) - b\left(X, Y_{1:T_n}^n\right)\right) + (1-\lambda) Q\left(Y_{1:T_n}, Y^*\right), & t < T, \\ \lambda\left(D\left(X, Y_{1:t}\right) - b\left(X, Y_{1:t}\right)\right) + (1-\lambda) Q\left(Y_{1:t}, Y^*\right), & t = T. \end{cases} \tag{9}$$

The gradient calculation of the objective function $J(\theta)$ to the parameter θ of generator G is shown as follows:

$$\nabla J(\theta) = \frac{1}{T}\sum_{t=1}^{T}\sum_{y_t} R_{D,Q}^{G_\theta}\left(Y_{1:t-1}, X, y_T, Y^*\right), \tag{10}$$

$$\nabla_\theta\left(G_\theta\left(y_t \mid Y_{1:t-1}, X\right)\right) = \frac{1}{T}\sum_{t=1}^{T} E_{y_t \in G_\theta}\left[R_{D,Q}^{G_\theta}\left(Y_{1:t-1}, X, y_T, Y^*\right) \cdot \nabla_\theta \log\left(G_\theta\left(y_t \mid Y_{1:t-1}, X\right)\right)\right]. \tag{11}$$

4. Experimental Results and Analysis

4.1. Introduction to Dataset. In this part, experiments on Chinese-English translation tasks are carried out to verify the effectiveness of the proposed neural machine translation method. The training data set used in the Chinese-English translation task is 1.25 million Chinese-English parallel sentence pairs that arer extracted from Linguistic Data Consortium (LDC). The test sets are NIST02 NIST03 NIST04 NIST05 and NIST08 from the National Institute of Standards and Technology in 2002. NIST06 was used as the test set. Byte pair encoder (BPE) encoding is performed for

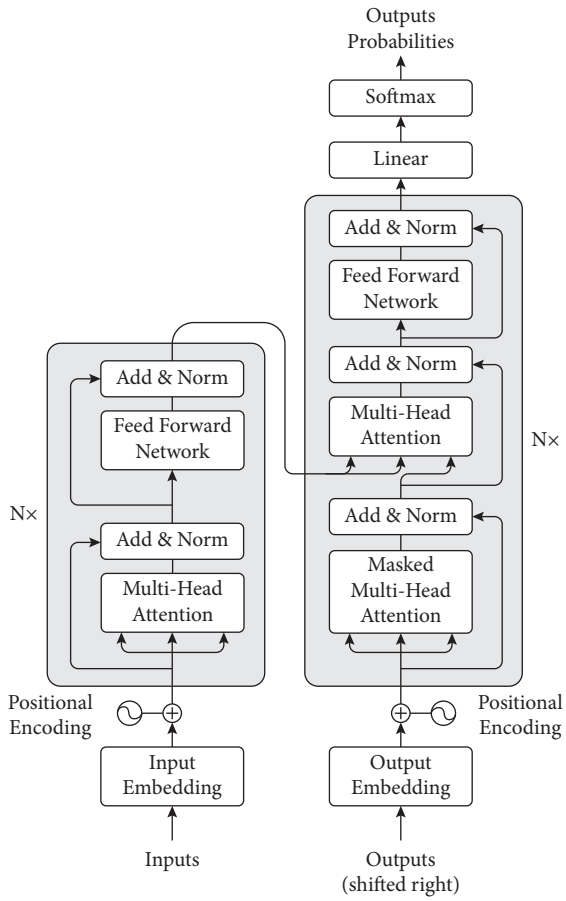


Figure 2: The structure of transformer model.

the Chinese and English materials, respectively. The BPE table is used in Chinese and English, respectively, and the size of both tables is 32000. The size of the Chinese is set to 40,000, and the size of the English is set to 30,000. The UNK takes the place of the low frequency of the words that are not in the word table.

4.2. Introduction to the Experimental Platform. In order to verify the effectiveness of the proposed method in machine translation, the experimental hardware environment is Intel (*R*) Core (TM) I7–8550U CPU @ 1.80 GHz 1.99 GHz, 16.0 GB memory, SSD 512 GB, Windows 10 Professional operating system model training and testing using Google's open source deep learning framework TensorFlow, experimental data analysis software environment, and test environment PyCharm 2019 Professional edition.

4.3. Experimental Results Analysis

4.3.1. Experimental Results of the Data Generalization Method. This section introduces the experimental results given by the data generalization based neural machine translation model. During the test, the 10 models with the highest BLEU value in the training stage are averaged to get the final experimental results of the model, as shown in Table 1 (the best result is shown in bold).

In Table 1, RNNSearch and Transformer are the experimental results of the baseline system on the original corpus, while RNNSearch_G and Transformer_G are the experimental results of processing the unknown words in the corpus using the data generalization strategy. The last row is the proposed method, and the experimental results (in bold) show that the performance of the machine translation model is significantly improved by the generalization of unknown words. The average BLEU values of the RNNSearch, RNNSearch_G, and Transformer_G models are 21 and 22.9, 22.8, 23.5, and 28.2, respectively. That is to say, the proposed methods is 4.7 percentage points higher than the suboptimal model, which has proved the superiority and effectiveness of the proposed method.

In order to examine the translation effect of unknown words in the NIST04test set, 200 sentences were randomly selected for manual evaluation. According to the experimental results, the proposed model has a better performance than Transformer_G models, so the experiment is only carried on the proposed model and Transformer_G model. The 200 sentences include 79 time expressions, 124 number expressions, 135 number expressions, and 46 website and special expressions, which are used to evaluate the translation effect of unknown words. The calculation formula is as follows:

$$\mathrm{ACC} = \frac{\text{number of untranslated words correctly}}{\text{total number of unlogged words}}. \tag{12}$$

The experimental performances are shown in Table 2 (the best result is shown in bold).

According to the experimental results in Table 2, the generative adversarial network can greatly improve the translation effect of unknown words in corpus. In terms of names, Chinese names account for 72% of the total number of names in the corpus, so the Chinese Pinyin translation method has the highest improvement, reaching 40.1%. Regarding the time, number, website, and special expressions, due to the rules-based method, the accuracy of all translations has been significantly improved, reaching 10.37%, 24.02%, and 34.44%, respectively.

4.3.2. Source Language Sentence Length Influences. The translation effect of long sentences is one of the important indexes to evaluate the performance of neural machine translation models. To study the translation performance of the multicoverage fusion model regarding the sentence length range of different source languages, the source languages in the test set were grouped according to the sentence length in the experiment. The BLEU values of different models in the range of source language length (0, 10], 10, 20], 20, 30], 30, 40, 50 and (50, +) were compared, as shown in Figure 3.

The following conclusions can be drawn from the experimental results shown in Figure 3. Overall, the performance of proposed model is better than that of RNNSearch and Transformer model. In the length interval of (0, 10), the BLEU value of proposed model is lower than that of the Transformer model, but higher than that of RNN search system. With the further increase of sentence length, BLEU values of RNNSearch system and proposed model decreased significantly.

Table 1: BLEU scores of different methods incorporating data generalization.

Method	NIST02	NIST03	NIST04	NIST05	NIST06	AVG
RNNSearch	26.2	21.7	19.0	18.4	19.7	21
RNNSearch_G	29.7	22.73	21.9	19.7	22.1	22.9
Transformer	23.4	24.2	23.5	21.6	20.6	22.8
Transformer_G	24.6	24.33	25.7	20.5	21.2	23.5
Proposed	**31.2**	**28.6**	**26.2**	**26.8**	**25.3**	**28.2**

Table 2: The translation results of different methods.

Method	Name (%)	Time expression (%)	Numeral expression (%)	URL and special expressions (%)
Transformer_G	31.82	68.29	52.24	52.37
Proposed	**71.92**	**78.66**	**76.26**	**86.81**

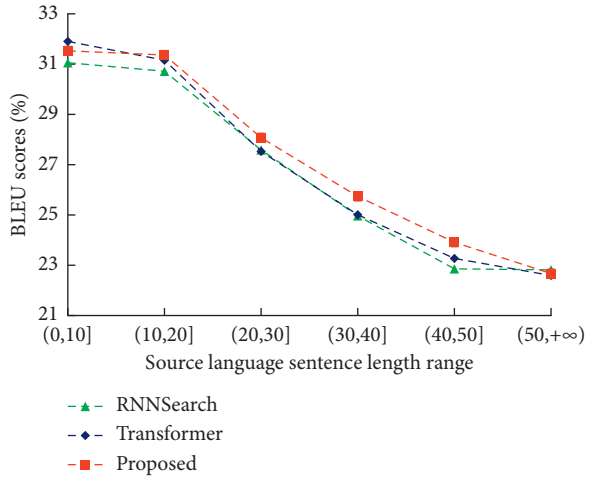


Figure 3: The comparison of BLEU scores with respect to the source length.

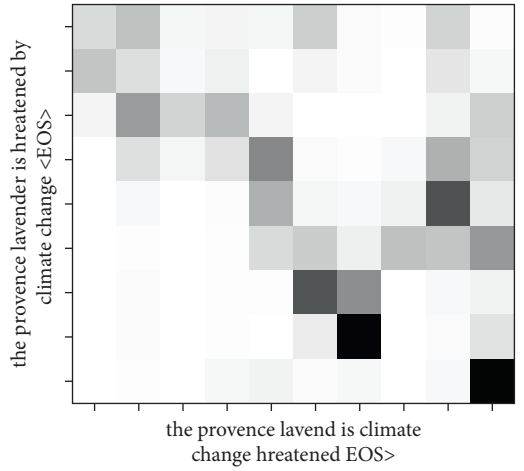


Figure 4: The proposed model alleviates the overtranslation phenomenon.

In the experiment, the overtranslation of different models was evaluated by counting the number of source language words in the test set. In Figure 4, the proposed model correctly translates Provence lavender as "Provence lavender." It not only eliminates duplication of translation but also correctly aligns attention between Provence and lavender, thus producing higher quality translation results.

As shown in Figure 5(a), in the baseline system, the translation of the word payment is omitted, leading to the loss of context meaning, so mobile is mistranslated as movement. The translation deviates completely from the

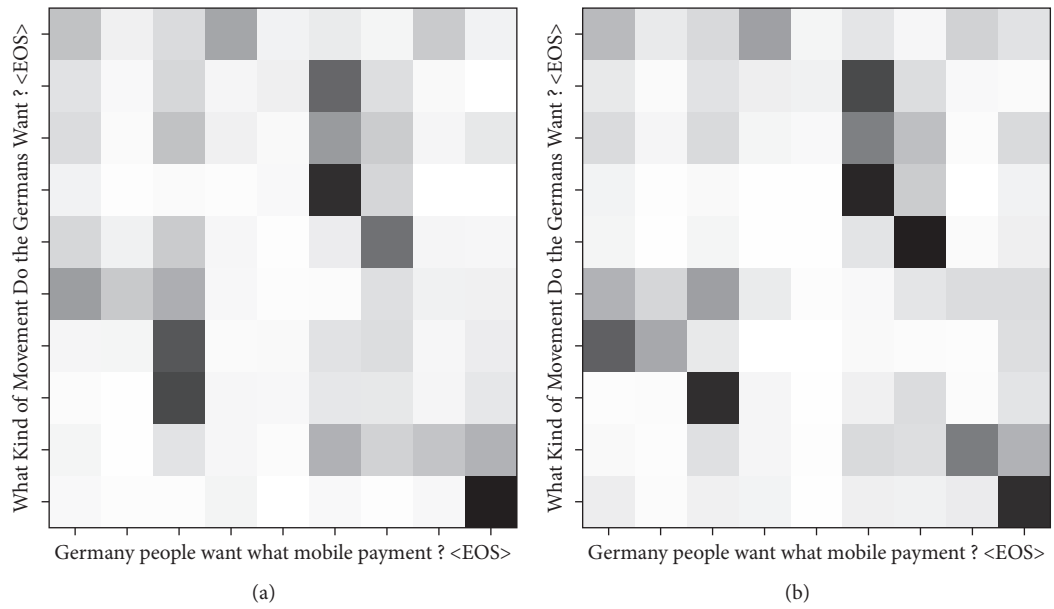


Figure 5: The proposed model alleviates the overtranslation phenomenon. (a) Baseline. (b) GAN.

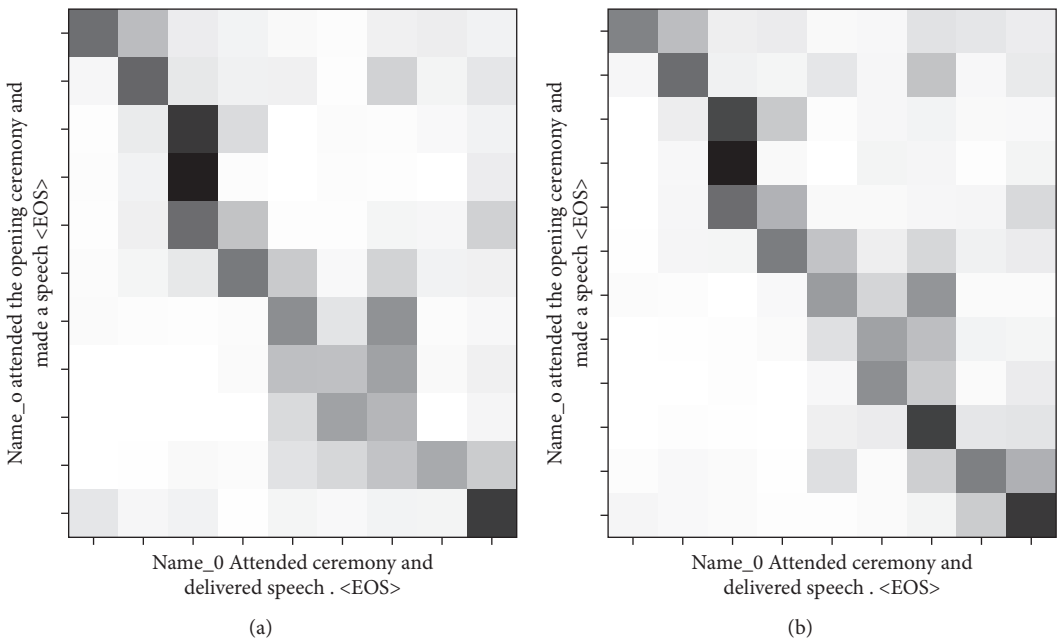


Figure 6: Alignment example.

meaning of the original sentence. In contrast, in Figure 5(b), the proposed model eliminates the phenomenon of translation leakage and correctly translates mobile payment into "mobile payment," thus ensuring the consistency of translation results with the original sentence meaning.

In addition to further alleviating the problems of overtranslation and overtranslation, the multicoverage fusion model can also improve the alignment quality of attentional mechanisms by studying the alignment relationship. As shown in Figure 6, in the baseline system, the second half of the text gives the keynote speech. There was an alignment error in translation. The keynote and speech were misaligned to speech and make, respectively, and the word keynote was omitted. The proposed model not only correctly establishes the alignment between keynote and speech but also eliminates the phenomenon of missing translation, which improves the quality of translation results.

4.3.3. Convergence Analysis of the Model. In order to better demonstrate the performance of the proposed method, its convergence is studied. Figures 7 and 8 show the variation

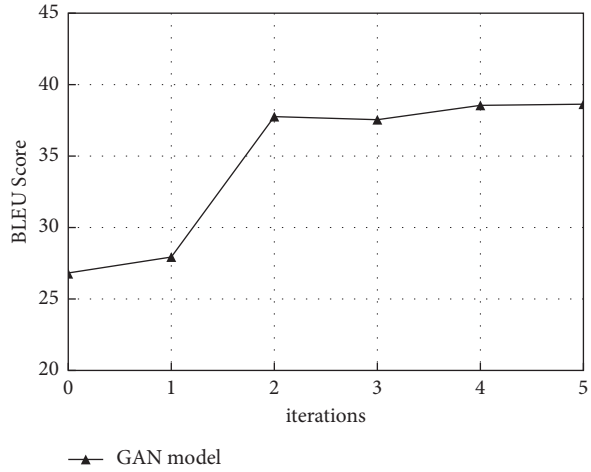


Figure 7: Performance curves of Chinese to English models.

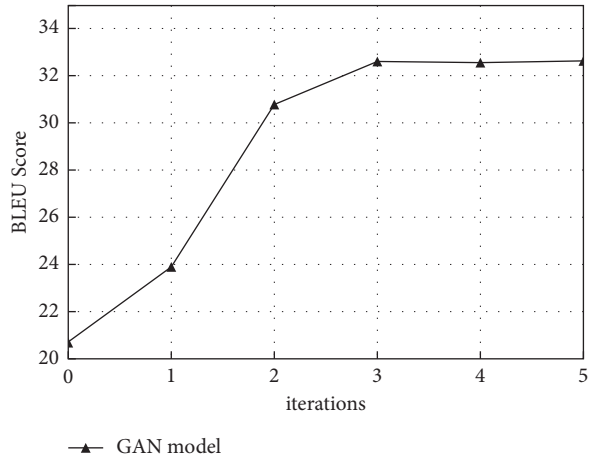


Figure 8: Performance curves of English and Chinese models.

curves of the model's performance with each iteration in the experimental process of different translation tasks. Figure 6 shows the model performance change curve of the proposed method in the Chinese-English translation task on the NIST2006 dataset.

Figure 8 shows the performance change curve of Transformer model with the number of iterations of EM algorithm in the translation task of English and Chinese on NIST2006 dataset. It can be seen from Figures 7 and 8 that, after five iterations, the model tends to converge, and it is difficult to improve the performance of the model, which also proves that the method in this paper tends to be stable and has a fast convergence speed after certain iterations. It shows that the proposed method has good performance.

5. Conclusions

Neural machine translation (NMT) model based on encoder and decoder structure is a mainstream approach in the field of machine translation. Although the performance of neural machine translation model has far exceeded that of traditional statistical machine translation methods, due to the lack of word list size and coverage mechanism, neural machine translation still has the problem of overtranslation and omitted translation of unknown words.

In this context, the communication between different languages is also closer, so accurate translation between languages is of great significance. However, existing methods give little thought to short sequence machine translation from Chinese to English. This paper designs a generative adversarial network to solve the previously mentioned problem. To avoid the loss of coverage information affecting the quality of attention alignment, both the coverage vector and the coverage score are obtained when calculating the attention score. According to the fusion mode of coverage vector and coverage fraction, the GAN model is constructed. Experimental results show that this model can further alleviate the problems of overtranslation and overtranslation, and GAN model has a more obvious improvement effect. In addition, the proposed GAN model has a fast convergence rate.

Although the previously mentioned methods can effectively improve the problem of unregistered words and

overtranslation, there are still some shortcomings. First, a large number of manual rules are needed to achieve the maximum alignment of unknown words, and the translation quality of people's names depends too much on dictionaries. Second, there are still many overtranslation and overtranslation phenomena in the multicoverage fusion model. In the following work, on the one hand, we will expand the scope of unknown words, such as the names of places, organizations, and proper nouns in specific fields, and explore other methods to translate unknown words, so as to reduce the dependence of thesaurus. On the other hand, phrase alignment in statistical machine translation is given to the neural machine translation model as an overlay reference to further improve overtranslation and overtranslation.

Data Availability

The dataset can be accessed upon request to the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] D. Shilton, M. Breski, D. Dor, and E. Jablonka, "Human social evolution: self-domestication or self-control?" *Frontiers in Psychology*, vol. 11, Article ID 134, 2020.

[2] N. M. Makhmudovna, N. G. Davronovna, E. D. Shonazarovna, and M. D. Shavkatovna, "language is a means of cognition and communication for teaching in education," *International Journal of Psychosocial Rehabilitation*, vol. 24, no. 1, pp. 247–256, 2020.

[3] V. M. Vysotki, E. V. Patalashko, V. E. Stepanenko, T. V. Makarova, and I. A. Balabanova, "Language as a communication resource and its place in the representation of world practices," *Linguistics and Culture Review*, vol. 5, no. S3, pp. 574–584, 2021.

[4] X. Qiu, T. Sun, Y. Xu, Y. Shao, N. Dai, and X. Huang, "Pre-trained Models for Natural Language Processing: A survey," *Science China Technological Sciences*, vol. 63, no. 10, pp. 1–26, 2020.

[5] C. Park, Y. Yang, K. Park, and H. Lim, "Decoding strategies for improving low-resource machine translation," *Electronics*, vol. 9, no. 10, Article ID 1562, 2020.

[6] Y. Xia, "Research on statistical machine translation model based on deep neural network," *Computing*, vol. 102, no. 3, pp. 643–661, 2020.

[7] Y. Belinkov, N. Durrani, F. Dalvi, H. Sajjad, and J. Glass, "On the linguistic representational power of neural machine translation models," *Computational Linguistics*, vol. 46, no. 1, pp. 1–52, 2020.

[8] K. Bringmann, M. KüNnemann, and A. Nusser, "Discrete fréchet distance under translation," *ACM Transactions on Algorithms*, vol. 17, no. 3, pp. 1–42, 2021.

[9] S. Knobloch-Westerwick, C. Mothes, and N. Polavin, "Confirmation bias, ingroup bias, and negativity bias in selective exposure to political information," *Communication Research*, vol. 47, no. 1, pp. 104–124, 2020.

[10] B. Zhang, D. Xiong, J. Xie, and J. Su, "Neural machine translation with GRU-gated attention model," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 11, pp. 4688–4698, 2020.

[11] Y. Liu, K. Wang, C. Zong, and K. Y. Su, "A unified framework and models for integrating translation memory into phrase-based statistical machine translation," *Computer Speech & Language*, vol. 54, pp. 176–206, 2019.

[12] H. Choi, K. Cho, and Y. Bengio, "Fine-grained attention mechanism for neural machine translation," *Neurocomputing*, vol. 284, pp. 171–176, 2018.

[13] M. Chen, Y. Li, and R. Li, "Research on neural machine translation model," *Journal of Physics: Conference Series*, vol. 1237, no. 5, Article ID 052020, 2019.

[14] S. Willis, "Mistranslation, missed translation: hélène cixous' vivre L'orange," *Rethinking Translation*, pp. 106–119, Routledge, England, UK, 1st edition, 2018.

[15] W. M. Bramer, D. Giustini, and B. M. R. Kramer, "Comparing the coverage, recall, and precision of searches for 120 systematic reviews in Embase, MEDLINE, and Google Scholar: a prospective study," *Systematic Reviews*, vol. 5, no. 1, pp. 1–7, 2016.

[16] Z. Ren, D. Dong, H. Li, and C. Chen, "Self-paced prioritized curriculum learning with coverage penalty in deep reinforcement learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 6, pp. 2216–2226, 2018.

[17] C. Leacock, M. Chodorow, M. Gamon, and J. Tetreault, "Automated grammatical error detection for language learners," *Synthesis lectures on human language technologies*, vol. 3, no. 1, pp. 1–134, 2010.

[18] S. Sun, C. Luo, and J. Chen, "A review of natural language processing techniques for opinion mining systems," *Information Fusion*, vol. 36, pp. 10–25, 2017.

[19] C. Leacock, M. Chodorow, M. Gamon, and J. Tetreault, "Automated grammatical error detection for language learners, second edition," *Synthesis lectures on human language technologies*, vol. 7, no. 1, pp. 1–170, 2014.

[20] F. Cairo, R. Rotundo, P. D. Miller, and G. P. Pini Prato, "Root coverage esthetic score: a system to evaluate the esthetic outcome of the treatment of gingival recession through evaluation of clinical cases," *Journal of Periodontology*, vol. 80, no. 4, pp. 705–710, 2009.

[21] F. Isaia, R. Gyurko, T. C. Roomian, and C. E. Hawley, "The root coverage esthetic score: intra-examiner reliability among dental students and dental faculty," *Journal of Periodontology*, vol. 89, no. 7, pp. 833–839, 2018.

[22] M. A. Marrs, B. D. Vogt, and G. B. Raupp, "Comparison of wet and dry etching of zinc indium oxide for thin film transistors with an inverted gate structure," *Journal of Vacuum Science and Technology: Vacuum, Surfaces, and Films*, vol. 30, no. 1, Article ID 011505, 2012.

[23] F. A. Gers and E. Schmidhuber, "LSTM recurrent networks learn simple context-free and context-sensitive languages," *IEEE Transactions on Neural Networks*, vol. 12, no. 6, pp. 1333–1340, 2001.

[24] Y. S. G. Kim, "Structural relations of language and cognitive skills, and topic knowledge to written composition: a test of the direct and indirect effects model of writing," *British Journal of Educational Psychology*, vol. 90, no. 4, pp. 910–932, 2020.

[25] P. Satasivam, B. Y. Poon, B. Ehdaie, A. J. Vickers, and J. A. Eastham, "Can confirmatory biopsy be omitted in patients with prostate cancer favorable diagnostic features on

active surveillance? The," *Journal of Urology*, vol. 195, no. 1, pp. 74–79, 2016.

[26] J. Su, J. Zeng, D. Xiong, Y. Liu, M. Wang, and J. Xie, "A hierarchy-to-sequence attentional neural machine translation model," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 3, pp. 623–632, 2018.

[27] M. Ghazvininejad, V. Karpukhin, and L. Zettlemoyer, "Aligned cross entropy for non-autoregressive machine translation," in *Proceedings of the International Conference on Machine Learning*, pp. 3515–3523, Vienna, Austria, April 2020.

[28] S. Wu, D. Zhang, Z. Zhang, N. Yang, M. Li, and M. Zhou, "Dependency-to-Dependency neural machine translation," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 11, pp. 2132–2141, 2018.

[29] H. Zhang, Y. Sun, L. Liu, X. Wang, L. Li, and W. Liu, "ClothingOut: a category-supervised GAN model for clothing segmentation and retrieval," *Neural Computing and Applications*, vol. 32, no. 9, pp. 4519–4530, 2020.

[30] N. Chiesa, B. A. Mork, and H. K. Høidalen, "Transformer model for inrush current calculations: simulations, measurements and sensitivity analysis," *IEEE Transactions on Power Delivery*, vol. 25, no. 4, pp. 2599–2608, 2010.

[31] J. C. W. Lin, Y. Shao, Y. Djenouri, and U. Yun, "ASRNN: a recurrent neural network with an attention model for sequence labeling," *Knowledge-Based Systems*, vol. 212, Article ID 106548, 2021.

Hindawi
Scientific Programming
Volume 2022, Article ID 1270700, 11 pages
https://doi.org/10.1155/2022/1270700

Research Article

Two-Way Neural Network Chinese-English Machine Translation Model Fused with Attention Mechanism

Jing Liang[1] **and Minghui Du**[2]

[1]*Yibin Vocational and Technical College, Yibin, Sichuan 644000, China*
[2]*Nanjing University of Chinese Medicine, School of Foreign Language Education, Nanjing, Jiangsu 210023, China*

Correspondence should be addressed to Minghui Du; 390052@njucm.edu.cn

Received 30 November 2021; Revised 24 December 2021; Accepted 27 December 2021; Published 25 January 2022

Academic Editor: Baiyuan Ding

This study uses an end-to-end encoder-decoder structure to build a machine translation model, allowing the machine to automatically learn features and transform the corpus data into distributed representations. The word vector uses a neural network to achieve direct mapping. Research on constructing neural machine translation models for different neural network structures. Based on the translation model of the LSTM network, the gate valve mechanism reduces the gradient attenuation and improves the ability to process long-distance sequences. Based on the GRU network structure, the simplified processing is performed on the basis of it, which reduces the training complexity and achieves good performance. Aiming at the problem that some source language sequences of any length in the encoder are encoded into fixed-dimensional background vectors, the attention mechanism is introduced to dynamically adjust the degree of influence of the source language context on the target language sequence, and the translation model's ability to deal with long-distance dependencies is improved. In order to better reflect the context information, this study further proposes a machine translation model based on two-way GRU and compares and analyzes multiple translation models to verify the effectiveness of the model's performance improvement.

1. Introduction

Machine translation usually combines natural language processing and artificial intelligence. It retains the original meaning. Realize the mutual translation between two natural languages and complete the equivalent conversion. As an important method to realize the equivalent transmission, translation is rather important in this process [1–6].

In the previous research, linguists manually wrote the conversion rules between the two languages to be translated. Although the research on rule-based methods has reached the syntactic stage, this method is almost entirely dependent on the quality of language rules, which requires linguists to be very high, and is easily restricted in the field of application and cannot summarize. all the rules that the language will use. The emergence of a translation model based on word alignment proposed by IBM in 1993 marked the birth of statistical machine translation. Compared with the previous model, statistical machine translation learns the conversion rules from the corpus, no longer needs to provide the language rules actively, and solves the bottleneck problem of knowledge acquisition. But in order to obtain perfect translation results, this method still has many problems. The translation results have poor utilization of global features and rely too much on data preprocessing steps. Each step is related to each other, and the accuracy of the processing steps affects the final translation result, greater impact. In fact, its related research has achieved good results. It is usually applied to provide new solutions for machine translation-related problems. The machine translation can be roughly divided into two types. One still uses the statistical machine translation system framework to improve key modules [7–11].

The research and exploration may be dated back to the birth of electronic computers in 1940s. Known as the pioneer of machine translation, Warren Weaver, as the head of the Natural Science Department of the Rockefeller Foundation, published "Machine Translation" in 1949, marking the

formal proposal of the idea of machine translation. He believes that the existence of multiple languages in the world is inherently consistent, and languages exist as a description tool of human society and objective things. For people who master a certain language, a foreign language is just a form of encoding of their mother tongue, and the foreign language can be translated into their mother tongue by cracking the password. Under the social background of significant achievements in the field of password cracking, people generally accepted Weaver's point of view and full of expectations for machine translation research. In the following ten years, the popularity of machine translation has continued to rise. The United States and the Soviet Union have caused a sudden upsurge in machine translation [12–15].

The development of machine translation has not been smooth sailing. For a period of time, the development of machine translation was put on hold. This was attributed to the "Language and Machine" report issued by the American Academy of Sciences Automated Language Processing Advisory Committee in 1966. In this report, the researchers of the committee denied the feasibility of machine translation with the shortcomings of slow speed, high accuracy, and high consumption and directly denied the possibility of the development of practical machine translation systems. At that time, the field of machine translation research was affected by the report, and the research investment was terminated one after another, making machine translation research into a period of underestimation. With the development of computer application technology and linguistics-related fields, coupled with the expansion of the demand for social information services at that time, machine translation has only begun to recover. The development of machine translation has mainly experienced the following aspects [16, 17].

The translation process is mainly composed of three parts: analysis, conversion, and generation. By parsing the source language sentence, the deep structure is obtained. The deep structure representation mentioned here can be divided into syntax-based, semantic-based, or intermediate language-based methods. Syntactic research focuses on analyzing the syntactic structure of the source language and generating the syntactic structure of the target language; semantic research uses semantic analysis on the source language to obtain the semantic representation and then transforms the semantic representation to generate the target language; the intermediate language method uses a mapping relationship with natural language. The universal intermediate language realizes through two mappings. Among them, machine translation based on syntactic rules has relatively mature syntactic theories such as generative transformation grammar and dependent grammar, which can obtain better translation results under certain conditions. The methods of semantic analysis and intermediate language are still in the preliminary research stage [18–22].

The rule-based machine translation method highly relies on the rules of language. Although it has a certain degree of versatility, the cost of obtaining the rules is relatively high. The quality of the language rules is too dependent on the experience and knowledge of the linguist, the maintenance of the rules, and the compatibility of the new and the old rules. It is a bottleneck problem that is difficult to break through. After the rule-based method, researchers built a large-scale bilingual corpus. At the same time, machine learning research methods appeared. The combination of the two has led to improved machine translation methods. This leads to the emergence of instance-based and statistical-based machines. Translation methods are based on bilingual corpus for research.

After years of research, statistical machine translation has achieved good results, but there are still many problems to be solved. Because the linear model is used, linearity is inseparable when processing high-dimensional complex data, making it difficult for training and search algorithms to approach the theoretical upper limit of the translation space; the conversion is mainly at the vocabulary, phrase and syntax level realization, and lack of proper semantic representation; use context-independent features to design funny dynamic programming search algorithms, which makes it difficult to accommodate nonlocal context information in the model; in addition, it also includes the difficulty of feature design, data sparseness, and error propagation caused by pipeline architecture problem. The application of deep learning theory can better solve these problems of statistical machine translation. Existing research methods mainly include two kinds: one uses deep learning technology to improve key modules and built the model to achieve direct source language to target language [23–27].

The well-known professor Bengio in the field of deep learning proposed a language model based on the neural network in 2003, which effectively alleviated the problem of data sparseness through distributed representation. Compared with the traditional language model that only considers the first $n-1$ words of the target language, Jacob et al. believed that not only the historical information of the target language but also the relevant parts of the source language also play a certain role. They proposed a neural network joint model to increase the BLEU value. About 6%, Peng et al. used a recursive autoencoder to obtain the distributed representation of the word string and then built a neural network classifier to alleviate the problem of feature design. Lu et al. transformed the existing feature set into a new feature set, which significantly improved the expression ability of the translation model. Nal et al. first proposed end-to-end neural machine translation in 2013 and proposed a "encoding-decoding" model framework to directly model the translation probability. For the input source language sentence, the encoder was used to map it into a continuous, dense vector, the continuous representation method solves the previous sparse problem, and then the vector is converted into a target language sentence through the decoder. They use convolutional neural networks to construct encoders, and decoders use recurrent neural networks to obtain historical information and process variable-length strings. In view of the gradient attenuation and gradient explosion problems that are prone to occur in the training process, Sutskever et al. introduced long and short-term memory LSTM into end-to-end neural machine translation and used the method of setting threshold switches to

improve the recurrent neural network, aiming at the problem of the encoder generating fixed-length vectors [28–33].

2. Neural Networks

The use of activation functions adds nonlinear modeling capabilities to neural networks. It can only perform linear mapping expressions. The entire network is only equivalent to a single-layer neural network, losing the actual meaning of the hidden layer. The activation function can make the deep neural network have the ability of nonlinear mapping learning. Under normal circumstances, the activation function should have the properties of differentiability, monotonicity, and limited output value range. Because the neural network training process is optimized based on the gradient, the activation function must be differentiable to ensure the calculation of the gradient. The monotonic activation function can ensure that the single-layer neural network is convex. Since the optimization is carried out on the basis of gradient calculation, the finite output of the activation function can make the feature representation in the neural network more significantly affected by the finite weight, thereby making the optimization method more stable. At present, the common activation functions are mostly piecewise linear and nonlinear functions with exponential shape, mainly including sigmoid, tanh, ReLU, ELU, and PReLU. Here, we will introduce the two activation functions used in this article.

The sigmoid function is the most widely used activation function, and its functional formula is

$$f(x) = \frac{1}{1 + e^{-x}}. \tag{1}$$

Tanh is a hyperbolic tangent function, and the function form is

$$f(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}}. \tag{2}$$

As another commonly used activation function, the output range of the tanh function is between [−1, 1], and the average output value is 0, which is the same as the sigmoid function. When the input is large or small, the output of the function is almost smooth. Although there will also be the problem of gradient disappearance, the tanh function converges faster and can reduce the number of iterations.

For the linear regression model, the hypothetical function is expressed as follows:

$$h_\theta(x_1, x_2, \ldots, x_n) = \theta_0 + \theta_1 x_1 + \cdots + \theta_n x_n. \tag{3}$$

where $\theta_i (i = 0, 1, \ldots, n)$ is the model parameter and $x_i (i = 0, 1, \ldots, n)$ is the n eigenvalue of each sample. Adding $x_0 = 1$, the above formula can be transformed into

$$h_\theta(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} \theta_i x_i. \tag{4}$$

The corresponding cost function is

$$J_\theta = \frac{1}{2m} \sum_{j=0}^{m} \left[h_\theta\left(x_0^j, x_1^j, \ldots, x_n^j\right) - y_j \right]^2. \tag{5}$$

The algorithm flow corresponding to the gradient descent method is shown in Figure 1:

In the system, the column search algorithm search tree is constructed. This method can greatly reduce the time and space resources occupied by the search process, but since each step is implemented by the greedy method, it cannot guarantee that the global optimal solution is obtained.

Then, it becomes

$$P(y_1, y_2, \ldots, y_{T'}) = \prod_{t'=1}^{T'} P(y_{t'} | y_1, ..., y_{t'-1}, c). \tag{6}$$

The dropout is shown as Figure 2.

Regarding the above systems, translation results cannot be directly reflected in the form of output text, and certain quantitative standards are needed for evaluation. In recent years, a variety of evaluation standards have appeared in the world, including BLEU, N1ST, and METEOR. The most commonly used is the BLEU (Bilingual Evaluation Understudy). This method obtains the evaluation value by calculating the similarity between the translation result of the computer translation system and the manual translation result. Compare and count the number of n-grams that cooccur in the system translation and the reference translation, and then divide the number of matched n-grams. The mathematical formula is

$$\text{BLEU} = \text{BP} \cdot \exp\left(\sum_{n=1}^{N} \omega_n \log\, p_n \right). \tag{7}$$

Among them, BP is the penalty factor, and N is the longest n-gram, ω_n means that the weight of the cooccurring n-gram word usually takes the constant value $1/n$, and the value of N is usually 4, and $\omega_n = 0.25$ is the most commonly used BLEU4 indicator.

3. Neural Network Translation Model Incorporating Attention

The rule-based machine translation method highly relies on the rules of language. Although it has a certain degree of versatility, the cost of obtaining the rules is relatively high. The quality of the language rules is too dependent on the experience and knowledge of the linguist, the maintenance of the rules, and the compatibility of the new and old rules. It is a bottleneck problem that is difficult to break through. After the rule-based method, researchers built a large-scale bilingual corpus. At the same time, machine learning research methods appeared. The combination of the two has led to improved machine translation methods. This leads to the emergence of instance-based and statistical-based machines. Translation methods are based on bilingual corpus for research.

After years of research, statistical machine translation has achieved good results, but there are still many problems

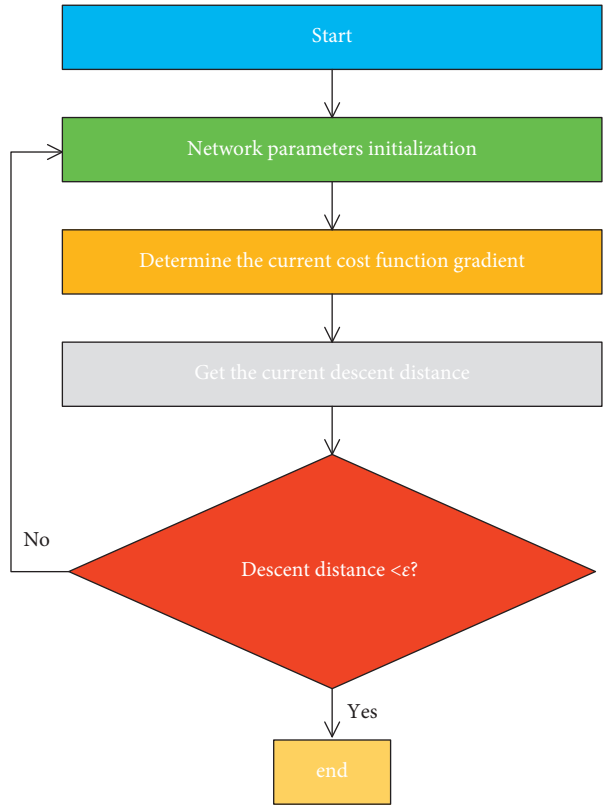


Figure 1: Gradient descent algorithm flow.

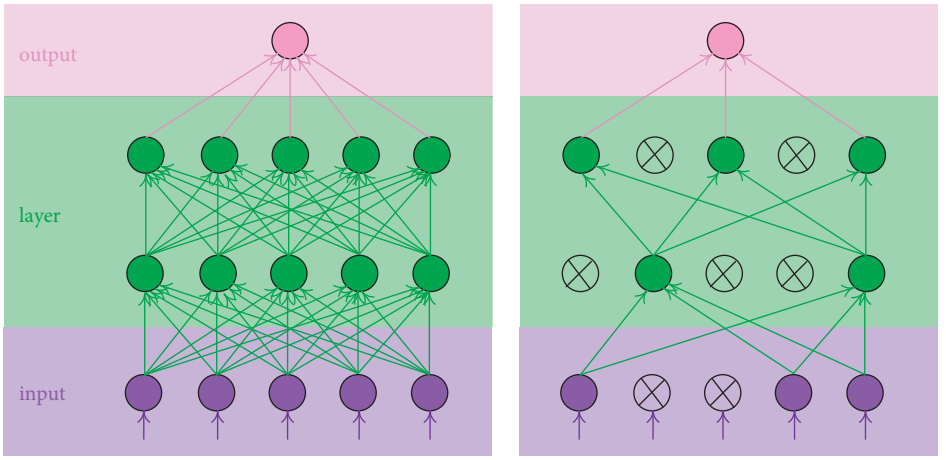


Figure 2: Dropout figure. (a) Traditional neural network. (b) Dropout neural network.

to be solved. Because the linear model is used, linearity is inseparable when processing high-dimensional complex data, making it difficult for training and search algorithms to approach the theoretical upper limit of the translation space; the conversion is mainly at the vocabulary, phrase and syntax level realization, and lack of proper semantic representation; use context-independent features to design funny dynamic programming search algorithms, which makes it difficult to accommodate nonlocal context information in the model; in addition, it also includes the difficulty of feature design, data sparseness, and error propagation caused by pipeline architecture problem. The application of deep learning theory can better solve these problems of statistical machine translation. Existing research methods mainly include two kinds: one uses deep learning technology to improve key modules and built the model to achieve direct source language to target language.

At present, many researchers have proposed many improved methods for NMT, among which the most effective is the codec system based on the attention

architecture. Figure 3(a) shows the structural principle according to the attention NMT model, which is devided into three parts: the coding and decoding layers and the structure of the middle cascade that introduces the attention mechanism.

The NMT system first converts all the divided symbols into a sequence, namely, word embedding. In this process, each character must be processed separately, and finally, the original sequence after the word embedding is generated. In the figure, the word embedding layer, NMT uses a bidirectional recurrent neural network (biRNN) to obtain a representation of the entire original sequence after training. Between the encoding layer and the decoding layer, an attention mechanism is added to fuse all time steps of the input sequence and focus on the current time step of the decoding layer. In the process of generating the target word, the controller will integrate three items: the last generated word, the current hidden layer state, and the context information calculated by the attention mechanism to determine the final target word.

The RNN coding layer is very important for NMT based on the attention model. However, it is difficult for traditional RNNs to achieve multilayer information integration, and machine translation increasingly requires this network structure. Therefore, this article proposes a multichannel attention mechanism encoder, the network of which is shown in Figure 3(b). This structure adds an external storage to assist the RNN to complete more complex integrated learning. In addition, the hidden layer state of the RNN and the word embedding sequence together generate gated annotations for the attention mechanism between the codec layers. From another perspective, integrating the word embedding sequence into the attention mechanism model can also be seen as establishing a short-circuit connection, which can alleviate the degradation problem. This short-circuit connection does not introduce any additional parameters while enhancing the network function and does not cause computationally complex upgrades.

Figure 4 shows the detailed rules for reading and writing the memory of the coding layer of the neural translation system. In each time step, the state node in the RNN queries the context information in the memory, and the memory is stored according to the attention-based mechanism. In this design, the previous state node is used to query and obtain context information as the input state of the gated recurrent unit (GRU), instead of directly feeding back the previous state to the GRU. This operation ensures that the controller can obtain more context information before generating the current state, which can potentially help the GRU make judgments. While designing read memory operation, write operation is also designed in the system. The purpose of this design, according to the research work of the Baidu team, is to hope that the RNN and NTM can learn different types of associations through different update strategies.

The function of the encoder is to convert an input sequence of variable length into a background vector C of fixed length.

The encoder part is mainly divided into three steps:

(1) First, the input source language sequence $X = \{x_1, x_2, \ldots, x_n\}$ is represented by one-hot encoding, and each word x_i in the source language sentence is represented as a column vector w_i. The value of the dimension represented by the index value is 1, and the rest are 0.

(2) Map the vector obtained in the previous step to the low-dimensional semantic space to form a word vector. Since the vector obtained by only one-hot encoding is too sparse and cannot describe the semantic similarity, it is necessary to use a word vector model of distributed representation to convert it and map it to the low-dimensional semantic space to generate a fixed-dimensional word vector.

(3) Using neural network coding to generate source language, as shown in the following equation:

$$h = \sigma(h_{i-1}, S_i). \tag{8}$$

Among them, h_0 is an all-zero vector, σ represents a nonlinear activation function, and the equation $h = \{h_1, h_2, \ldots, h_n\}$ obtained is the input state coding sequence of n source language words, which is the background vector C mentioned at the beginning of this section.

The task of the decoder is to obtain the maximum probability. The specific process is as follows:

(1) For a certain time i in the sequence, calculate the next hidden layer state Z_{i+} according to the background vector C generated by the source language sequence. Calculate by

$$Z_{i+1} = \sigma(c, u_i, z_i). \tag{9}$$

Among them, z_0 is an all-zero vector, and s represents a nonlinear activation function.

(2) The softmax function is used for normalization to obtain the probability distribution of the $Z_{i+1}{}^{\text{th}}$ word

(3) Calculate the cost function according to the obtained probability, and repeat the above sequence processed.

Statistical machine translation usually designs translation models based on linguistic knowledge. Neural machine translation focuses on the design of neural network mechanisms in the model. The encoder and decoder frameworks involve the use of neural networks to build models. In actual use, there are many types of neural networks, and choosing a suitable neural network has an important impact on the machine translation results. As the earliest network structure proposed in the field of neural machine translation, the recurrent neural network is a neural network that can store time series information. It stores historical information through hidden states. This structure makes it theoretically able to input sequences of arbitrary length processed. It evolved on the basis of the feedforward neural network. Each layer of the recurrent neural network not only outputs the next layer but also outputs a hidden layer state, which is used when processing the next sample for the current layer. The state of the hidden layer is a function representation

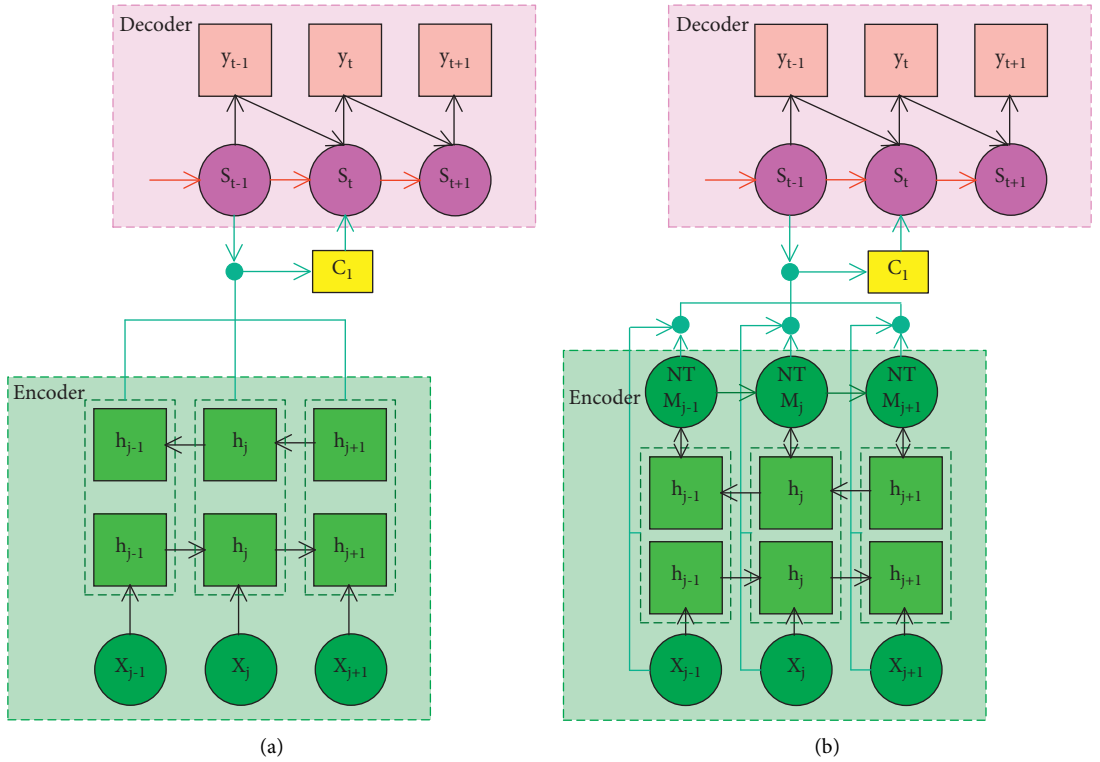


Figure 3: The structural principle of the NMT model based on attention. (a) Conventional NMT. (b) NMT with multichannel encoder.

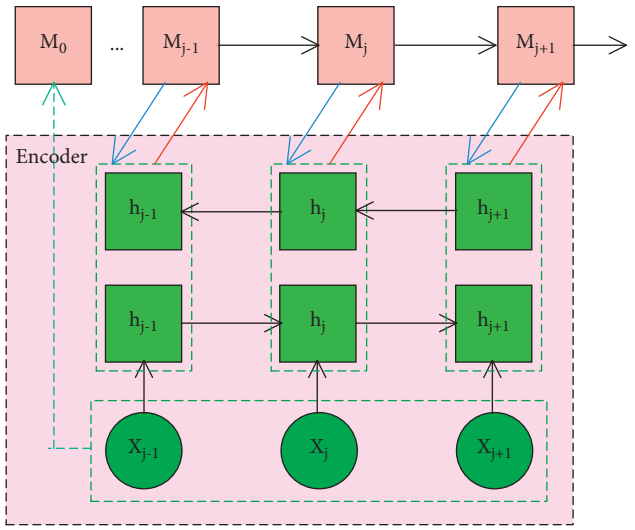


Figure 4: Coding layer of the neural translation system.

of the state of all the hidden layers at the previous moment. This structure enables the recurrent neural network to process data samples with front-to-back dependencies.

In a single hidden layer feedforward neural network, the vectorization method is used to calculate, where s represents the sigmoid function, that is, the activation function of the hidden layer; assuming that the length of the hidden layer is h, the number of samples is n, and the feature vector dimension is for batch data of x; the output of the hidden layer is represented as

$$H = \sigma\left(XW_{xh} + b_h\right),$$
$$\text{sigmod}\left(x\right) = \frac{1}{1 + e^{-x}}. \tag{10}$$

Figure 5 shows the basic processing flow of constructing an encoder-decoder machine translation model using a recurrent neural network. The encoder end input is a source language sequence with a start mark and an end representation, which is converted into a distributed representation of word vector data and then passed into the encoder end neural network, and the background vector containing the source language information is passed. As the input of the neural network on the decoder side, finally input the target language sequence through the neural network calculation. Add "<bos>" (beginning of sequence) in front of the input sequence to indicate the beginning of the sequence, and add the character "<eos>" (end of sequence) after the input and output sequences to indicate the termination of the sequence and character to determine the termination condition of the current sequence.

In the recurrent neural network, the hidden layer state at the current moment records the network information at the previous moment, and the output of the network is calculated based on this state. In actual use, the word represented by one-hot is first mapped into a word vector representation using a distributed representation method and then used as the input of the recurrent neural network at each time to achieve sequence output. In the training process, when the time sequence length of each time series training data sample is large or time t is small, the hidden layer variables of the

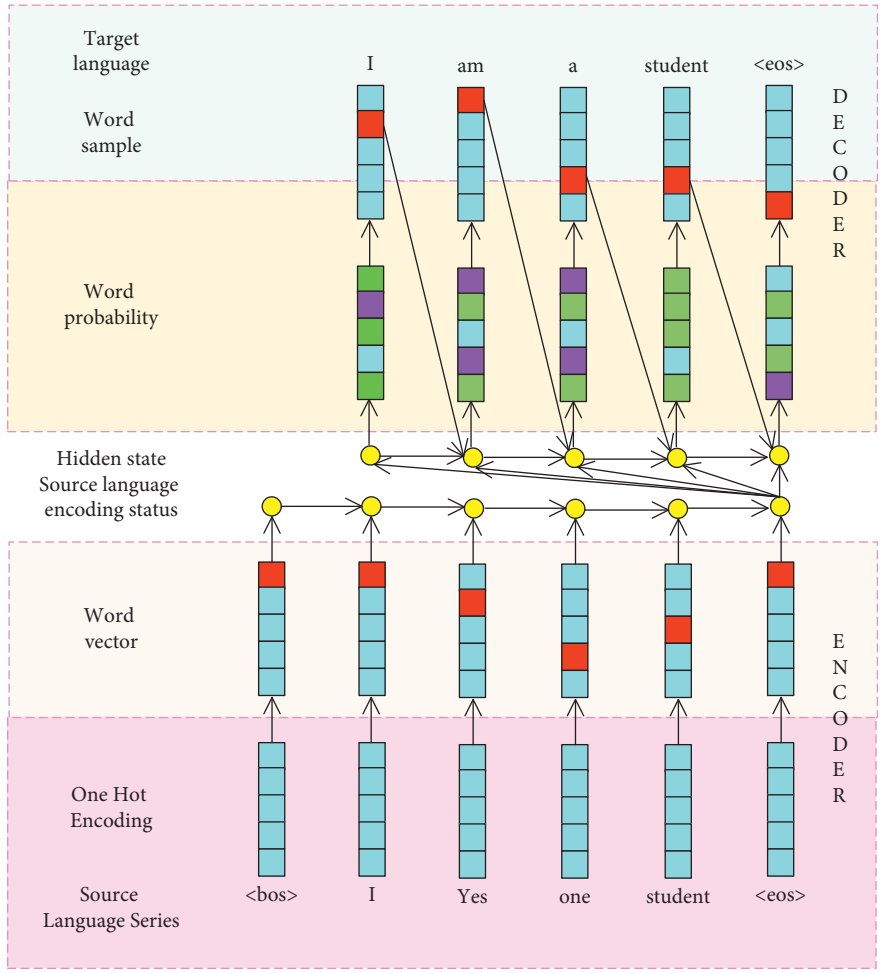


Figure 5: Encoder-decoder framework.

objective function at time t may have gradient decay or gradient explosion. Since the deep neural network training adopts the method of back propagation, chain derivation is required. The gradient calculation of each layer will involve continuous multiplication operation. In the face of multilayer networks, if the continuous multiplication factor is mostly less than 1, the final product may tend to 0, that is, gradient attenuation occurs, resulting in the subsequent network layer parameters not changing, and updated iterations cannot be performed; similarly, if the multiplication factor is mostly greater than 1, the final product may tend to infinity, that is, a gradient occurs and explode.

In the face of gradient explosion, gradient clipping is required. All gradients are spliced into one vector, and the clipping threshold is set to θ. According to (10), it can ensure that the magnitude of the gradient vector does not exceed the threshold:

$$g = \min\left(\frac{\theta}{|g|}, 1\right)g. \tag{11}$$

According to the translation model constructed in the previous content, it can be found that the output of the encoding stage is a fixed-dimensional vector, which leads to the use of the same-dimensional vector to encode the semantic syntax and other information, regardless of the length of the source language sequence. The input involved in the translation process is a sequence of variable length, and the constructed model may not fully fit the source language input sequence, and the processing effect is not ideal when dealing with long sentence sequences. At the same time, from the perspective of human translation, while translating, more attention will be paid to the part of the source language that is more closely related to the current translation. As the translation progresses, the part of concern will also change, but the use of fixed-dimensional vectors here means all parts of the source language sequence are given the same degree of attention, which is obviously not conducive to the improvement of machine translation performance. The decoder model strengthens the representation ability of the source language sequence. As shown in Figure 6, the neural machine translation model consists of three parts: the encoder, the decoder, and the attention mechanism, using a set of multiple vectors instead of a fixed vector to represent the source language information, the target sequence generation process dynamically selects the background vector, and the

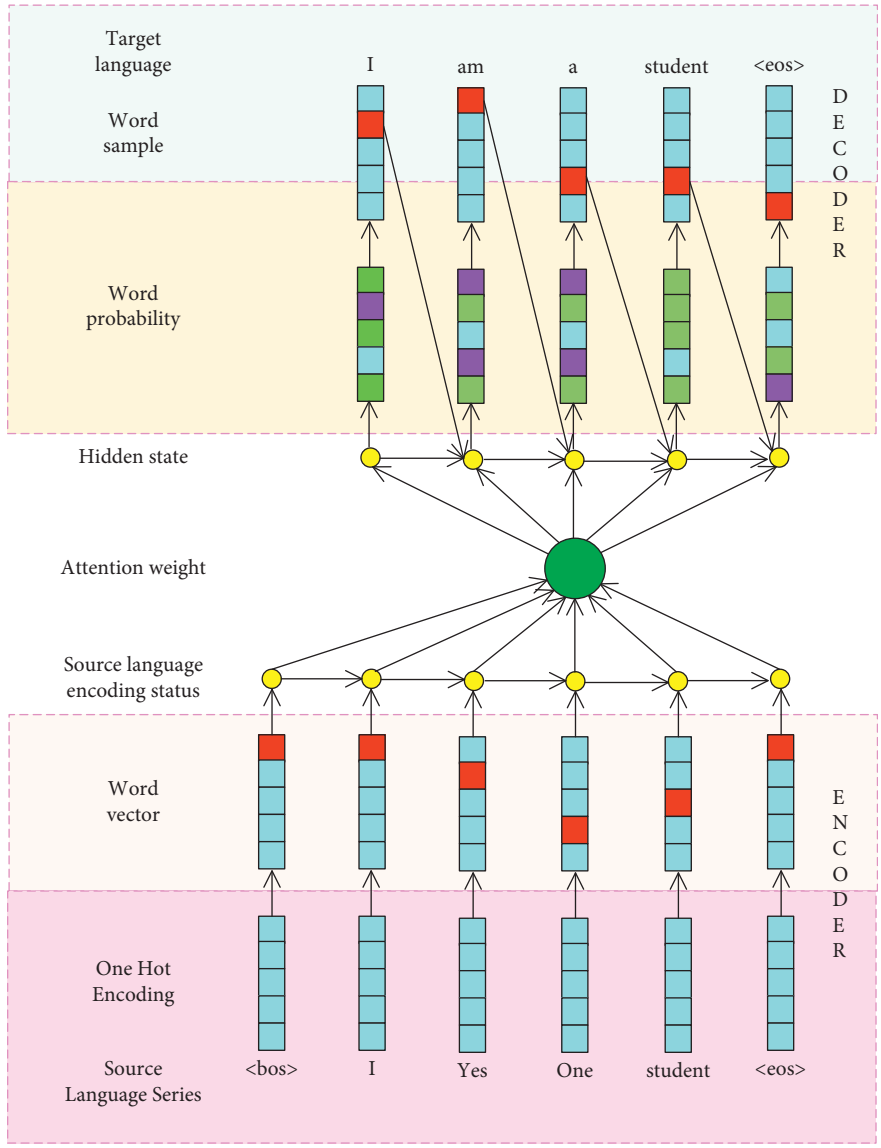


Figure 6: Translation model using the attention mechanism.

decoding process pays more attention to the part of the source language that is more relevant to the target sequence.

Similar to the decoder part proposed in this part, the calculation formula for the next hidden layer state of the target end is

$$Z_{i+1} = \sigma(c_i, u_i, z_i). \tag{12}$$

At this time, the background vector is no longer uniform and fixed, and C_i is the i^{th} word in the source language sentence, which makes it possible to find a unique background vector C_i corresponding to each word.

The calculation formula is

$$C_i = \sum_{j=1}^{T} \alpha_{ij} h_j. \tag{13}$$

This formula actually performs a weighted average calculation on the hidden layer state. The calculation formula is

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k=1}^{T} \exp(e_{ik})}, \qquad e_{ij} = a(z_i, h_j). \tag{14}$$

The attention mechanism allows the translation model to better solve the problem of long-distance dependence by adding different weights to the source language. According to the experimental results of Bahdanau et al., as shown in Figure 7, through the introduction of the attention mechanism, the translation model's translation effect for long sentences has been significantly improved compared to the model without the attention mechanism. The exact model is shown in Figure 8.

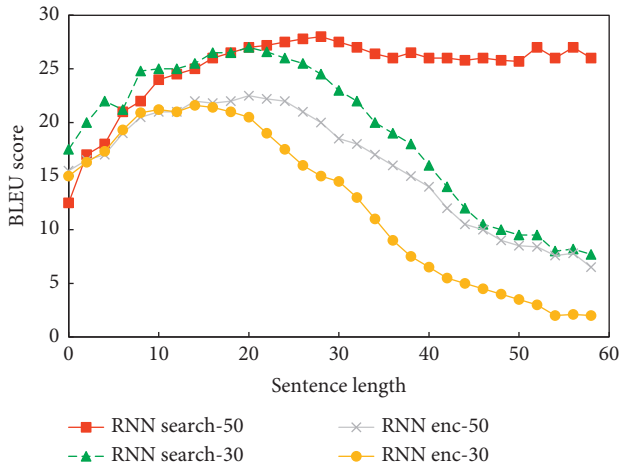


Figure 7: Attention mechanism performance.

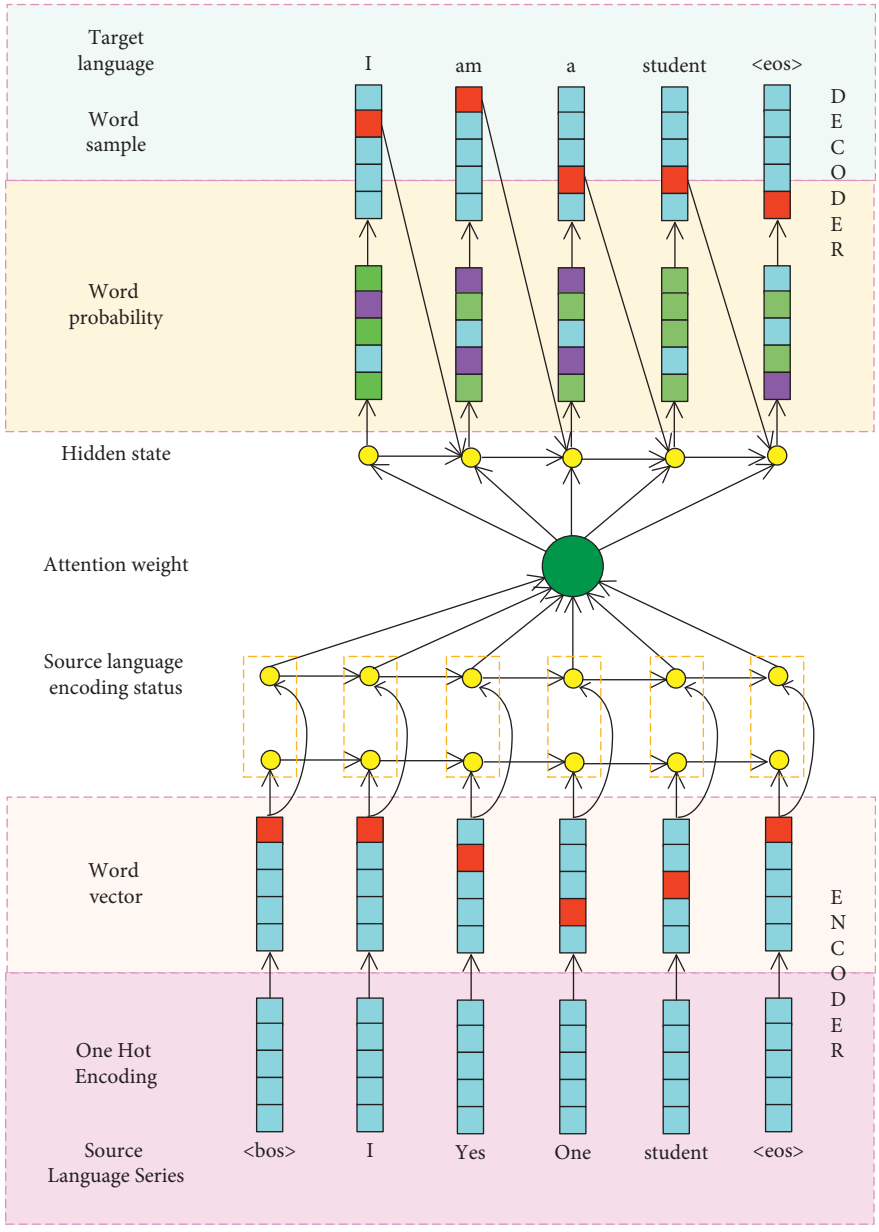


Figure 8: BiGRU model structure.

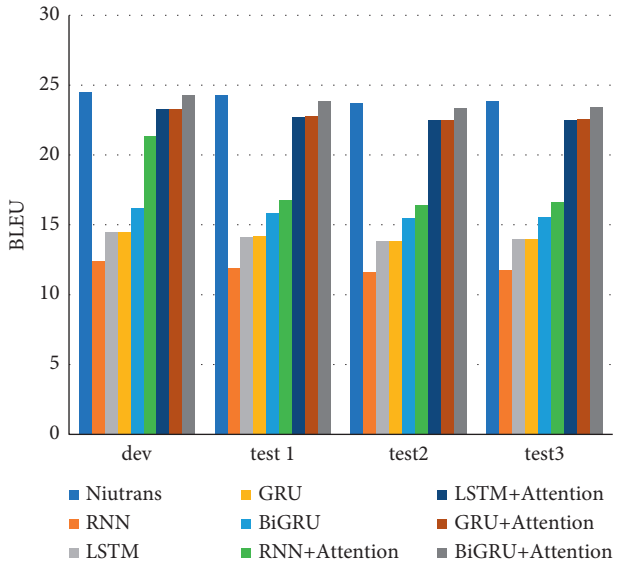


Figure 9: BLEU values of different translation models on the development set and test set.

4. Experimental Comparative Analysis

In view of the actual situation of the research in this article, the experimental data in this section use the IWSLT dataset with a smaller data scale for experiments. The International Workshop on Spoken Language Translation (IWSLT) is the most influential spoken language machine translation in the world. Evaluation of the game: a comparative analysis of the various neural network models mentioned in this chapter, and an experimental comparison of Chinese-English machine translation tasks. The experimental data use the IWSLT 2015 dataset, which includes 220000 Chinese-English parallel sentence pairs training data, 1 pair development set data, and 3 pairs of test set data. Build a neural machine translation system on the deep learning framework TensorFlow.

Different neural networks are used in the translation model used for training separately. The experimental results are shown in Figure 9. The first column in the figure shows the translation models of different types of neural networks, and the second to fourth columns show that several translation models are in the BLEU value performance on the development set and test set.

Comparing lines 2–5 in the figure, it can be found that under the premise that the attention mechanism is not used, the performance of the model based on the RNN structure is better than that of the model using the LSTM and GRU structure. The two-way GRU proposed in this study is in the original, improved on the basis. The 5–8th lines in the figure are all translation models that incorporate the attention mechanism on the original structure.

5. Conclusions

This article mainly focuses on Chinese-English machine translation tasks. The encoder and decoder parts are constructed by the neural network, and the word vector method of distributed representation is adopted, source language and target language sequence. Different translation models are constructed for comparison and analysis of three kinds of neural networks: recurrent neural network, long-term short-term memory, and gated recurrent unit. The attention mechanism is added to improve the translation model's processing of long-distance dependence. A neural machine translation model based on two-way GRU is proposed to improve the contextual representation ability of the source language, and experiments are carried out to verify the effectiveness of the improved method.

Neural machine translation uses neural networks to switch between the source language and the target language. The relatively closed translation process makes it difficult to explicitly use prior knowledge. Aiming at the problem that neural machine translation cannot make good use of linguistic knowledge, proposed a neural machine translation model that adds part-of-speech sequence information. On the basis of the two-way GRU translation model fused with the attention mechanism, Stanford Parser is used for syntactic analysis to obtain the part-of-speech sequence information, and then, it is integrated into the translation model in the form of two-way encoding. The encoder part uses vector splicing to form the background vector together. Verify the effectiveness of the analysis and improvement methods.

Data Availability

The dataset used to support the findings of this study can be accessed upon request from the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, Las Vegas, NV, USA, June 2016.

[2] M. Yun, J. Zhao, J. Zhao, X. Weng, and X. Yang, "Impact of in-vehicle navigation information on lane-change behavior in urban expressway diverge segments," *Accident Analysis & Prevention*, vol. 106, no. 1, pp. 53–66, 2017.

[3] S. Kumar Dwivedi, R. Amin, V. Satyanarayana, and C. Rashmi, "Blockchain-based secured event-information sharing protocol in internet of vehicles for smart cities," *Computers & Electrical Engineering*, vol. 86, no. 1, pp. 1–9, 2020.

[4] Z. Khan and S. Amin, "Bottleneck model with heterogeneous information," *Transportation Research Part B: Methodological*, vol. 112, no. 1, pp. 157–190, 2018.

[5] J. M. Cairney, K. Rajan, D. Haley et al., "Mining information from atom probe data," *Ultramicroscopy*, vol. 159, no. 1, pp. 324–337, 2020.

[6] J. Yu and P. Lu, "Learning traffic signal phase and timing information from low-sampling rate taxi GPS trajectories," *Knowledge-Based Systems*, vol. 110, no. 1, pp. 275–292, 2016.

[7] K. P. Wijayaratna, V. V Dixit, and L. Denant-Boemont, S. Waller, An experimental study of the Online Information Paradox: does en-route information improve road network performance?" *PLoS ONE*, vol. 12, no. 9, pp. 184–191, 2017.

[8] Z. Wang, H. Ren, Q. Shen, W. Sui, and X. Zhang, "Seismic performance evaluation of a steel tubular bridge pier in a five-span continuous girder bridge system," *Structures*, vol. 31, no. 1, pp. 909–920, 2021.

[9] S. Nakayama and J. Takayama, "Traffic network equilibrium model for uncertain demands," in *Proceedings of the 82nd Transportation Research Board Annual Meeting*, Washington, DC, USA, 2021.

[10] H. Shao, W. H. K. Lam, and M. L. Tam, "A reliability-based stochastic traffic assignment model for network with multiple user classes under uncertainty in demand," *Networks and Spatial Economics*, vol. 6, no. 3, pp. 173–204, 2019.

[11] A. Chen, J. Kim, S. Lee, and Y. Kim, "Stochastic multi-objective models for network design problem," *Expert Systems with Applications*, vol. 37, no. 2, pp. 1608–1619, 2020.

[12] H. Wang, W. H. K. Lam, X. Zhang, and H. Shao, "Sustainable transportation network design with stochastic demands and chance constraints," *International Journal of Sustainable Transportation*, vol. 9, no. 2, pp. 126–144, 2015.

[13] S.-M. Hosseininasab and S.-N. Shetab-Boushehri, "Integration of selecting and scheduling urban road construction projects as a time-dependent discrete network design problem," *European Journal of Operational Research*, vol. 246, no. 3, pp. 762–771, 2015.

[14] S.-M. Hosseininasab, S.-N. Shetab-Boushehri, S. R. Hejazi, and H. Karimi, "A multi-objective integrated model for selecting, scheduling, and budgeting road construction projects," *European Journal of Operational Research*, vol. 271, no. 1, pp. 262–277, 2018.

[15] M. Johnson, M. Schuster, and Q. V. Le, M. Krikun, Y. Wu, Z. Chen et al., Google's multilingual neural machine translation system: enabling zero-shot translation," *Transactions of the Association for Computational Linguistics*, vol. 5, no. 1, pp. 339–351, 2017.

[16] M. D. Moreno, "Translation quality gained through the implementation of the iso en 17100:2015 and the usage of the blockchain," *Babel*, vol. 1, no. 2, pp. 1–9, 2020.

[17] X. Wang, X. Yu, L. Guo, F. Liu, and L. Xu, "Student performance prediction with short-term sequential campus behaviors," *[J]. Information.*vol. 11, no. 4, p. 101, 2020.

[18] Q. Guo, Z. Zhu, Q. Lu, D. Zhang, and W. Wu, "A dynamic emotional session generation model based on Seq2Seq and a dictionary-based attention mechanism," *Applied Sciences*, vol. 10, no. 6, pp. 1–10, 2020.

[19] H. Ren, Xi Mao, W. Ma, J. Wang, and L. Wang, "An English-Chinese machine translation and evaluation method for geographical names," *ISPRS International Journal of Geo-Information*, vol. 9, no. 3, pp. 193–201, 2020.

[20] J. Arús-Pous, T. Blaschke, S. Ulander, J.-L. Reymond, H. Chen, and O. Engkvist, "Exploring the GDB-13 chemical space using deep generative models," *Journal of Cheminformatics*, vol. 11, no. 1, pp. 20–29, 2019.

[21] T. Moon, T. Ahn, and J. Son, "Long short-term memory for a model-free estimation of macronutrient ion concentrations of root-zone in closed-loop soilless cultures," *Plant Methods*, vol. 15, no. 1, pp. 1–12, 2019.

[22] N. Pourdamghani and K. Knight, "Neighbors helping the poor: improving low-resource machine translation using related languages," *Machine Translation*, vol. 33, no. 3, pp. 239–258, 2019.

[23] L. Bote-Curiel, S. Muñoz-Romero, A. Gerrero-Curieses, and J. L. Rojo-Álvarez, "Deep learning and big data in healthcare: a double review for critical beginners," *Applied Sciences*, vol. 9, no. 11, pp. 1–11, 2019.

[24] J. Zhang and T. Matsumoto, "Corpus augmentation for neural machine translation with Chinese-Japanese parallel corpora," *Applied Sciences*, vol. 9, no. 10, pp. 1–12, 2019.

[25] Y. Chen, Y. Ma, X. Mao, and Q. Li, "Multi-task learning for abstractive and extractive summarization," *Data Science and Engineering*, vol. 4, no. 1, pp. 14–23, 2019.

[26] Pu Zhou and Z. Jiang, "Self-organizing map neural network (SOM) downscaling method to simulate daily precipitation in the Yangtze and Huaihe River Basin," *Climatic and Environmental Research*, vol. 21, no. 5, pp. 512–524, 2016.

[27] X. Xiao, "Analysis on the employment psychological problems and adjustment of retired athletes in the process of career transformation," *Modern Vocational Education*, vol. 5, no. 12, pp. 216-217, 2018.

[28] S. Sahoo and M. K. Jha, "Pattern recognition in lithology classification: modeling using neural networks, self-organizing maps and genetic algorithms," *Hydrogeology Journal*, vol. 25, no. 2, pp. 311–330, 2016.

[29] Y. Zhou and B. Yang, "Sports video athlete detection using convolutional neural network," *Journal of Natural Science of Xiangtan University*, vol. 39, no. 1, pp. 95–98, 2017.

[30] J. Pang, "Research on the evaluation model of sports training adaptation based on self-organizing neural network," *Journal of Nanjing Institute of Physical Education*, vol. 16, no. 1, pp. 74–77, 2017.

[31] G. Querzola, C. Lovati, C. Mariani, and L. Pantoni, "A semi-quantitative sport-specific assessment of recurrent traumatic brain injury: the TraQ questionnaire and its application in American football," *Neurological Sciences*, vol. 40, no. 9, pp. 1909–1915, 2019.

[32] J. Wang, X. Luo, and H. Yan, "Correlation analysis between injuries and functional movement screening for athletes of the National Shooting Team," *Journal of Capital Institute of Physical Education*, vol. 5, no. 4, pp. 352–355, 2016.

[33] G. Ma, "Research on the design of juvenile football players' sports injury prediction model," *Automation Technology and Application*, vol. 277, no. 7, pp. 141–144, 2018.

Hindawi
Scientific Programming
Volume 2022, Article ID 8058390, 12 pages
https://doi.org/10.1155/2022/8058390

Research Article

Research on Promoting Visual Communication of Local Folk Culture by Using Digital Technology

Zhixiong Jia and Yingfa Yang

School of Marxism, Hebei University of Engineering, Handan, Hebei 056038, China

Correspondence should be addressed to Yingfa Yang; yangyingfa@hebeu.edu.cn

Received 21 November 2021; Accepted 27 December 2021; Published 20 January 2022

Academic Editor: Baiyuan Ding

Through the visual analysis of local customs and cultures by using digital technology, the total index of digital artificial intelligence, the environmental support of artificial intelligence (AI), and the competitiveness of AI industry have developed well, while the creativity of AI knowledge has developed steadily. According to the survey of digital technology audience education, it can be seen that the respondents with high school to undergraduate education have a higher audience rate of digital technology, while those with other academic qualifications have a lower audience rate, so digital technology can be popularized through education. Through the investigation of audience jobs, it is known that most of the audience of digital technology are engaged in enterprise staff, but they are less exposed to digital technology in political party organs. It can be seen that the loading time of visual images after clustering processing will be reduced by about 200 compared with that without clustering processing, which greatly improves the efficiency of visual graphics loading. As well as the analysis of the development of visualization, we can see that from the sixteenth century to the present, the visualization efficiency has increased from about 50% to 98%, which greatly improves the efficiency. According to the survey of local folk customs, from 2013 to 2019, local residents' awareness of local folk customs has been continuously improved, from 30% to about 48%, while the awareness of completely unclear folk customs has dropped from 10% to about 1%. Through the investigation of local folk culture, it can be seen that in the promotion of local folk culture, the governance system and measures are insufficient, and the governance effect cannot be effectively fed back.

1. Introduction

Through the use of digital technology to visually analyze the local folk culture, we know that there are some problems in the development of local folk customs, such as imperfect governance system and feedback system. We should improve the related problems and promote the protection, development, and inheritance of local folk culture.

Literature [1] realizes data transmission to memory by constructing a new meteor radar. In order to construct this kind of meteor radar, it is necessary to combine modern digital technology with computer method so that it can be realized, and the relevant position of meteor shower radiation source can also be inferred by this meteor radar. In order to apply the new digital technology in literature [2], economists put forward three explanatory methods to explain the paradox related to productivity. In this paper, generative pretrained transformer (GPT) technology and economic system transition method are used to understand these three effects. Literature [3] investigates 1234 educators and investigates the internal and external influencing factors of digital technology on educators through research path modeling. The results of this study show that digital technology has a great influence on educators and children's learning, and this digital technology can also improve teachers' teaching confidence and help teachers' classroom teaching. Literature [4] congress reformulated two schemes to protect copyright and creative works. To address the threat posed by digital technology, congress enacted the Digital Millennium Copyright Act (DMCA) to enable authors to protect their works. It may be unwise for the DMCA to prohibit the use and distribution of decryption

technology, but it also aims to reduce the impact of digital technology on copyright. Literature [5] provides an analytical framework to understand the development from e-marketing to mobile marketing. Through analysis, we realize that its growth is ultimately due to the growth of technology, including the performance and power of digital technology and the development of the times. These are the factors of marketing growth. With the development of the times, many children are also using network technology [6]. In order to study the experience of children and families in using new technologies such as mobile phones, tablets, computers, and other digital technologies, different countries are studied. Its research structure provides children and families with suggestions on the use of digital technology. Literature [7] studies single event reversals for testing heavy ions and protons. In this study, heavy ion and proton test structures showed that many cells were disordered and no latch-up phenomenon was observed. It is proved that charge sharing is a mechanism of multiunit flipping. Literature [8] manages the process of laboratory or network by combining digital technology management with printing technology, so as to effectively handle image input and workflow. Improve efficiency, thus reducing the requirement of physical batch processing. Therefore, the laboratory can effectively handle the relevant workflow. Literature [9] is a representative folk culture with vivid and diverse forms of expression. At the same time, it can best reflect the local folk customs and also has the development route of keeping pace with the times. Local folk culture is the concrete expression of cultivating local people's social value, which has very important educational significance and is a good embodiment of social value. Literature [10] shows that the local traditional folk culture is often formed by the collision, absorption, and acceptance of various customs and cultures through Professor Zhao Zongfu's related works. It exposes the unique and distinct regional customs and culture. Academy is the experience of famous customs and culture [11]. Through the study of academy, new ideas can be effectively disseminated through academy, and it is also helpful for the early introduction of new ideas. Academy is also a gathering place of intellectuals, which embodies the secular and down-to-earth development of people's daily life. It plays an important role in the introduction of new ideas. This paper takes the Snowtown Festival in Lhasa as the research object, and through analysis [12], it is known that this custom culture can protect and activate the local traditional culture. We should also attract the public's interest in local customs and cultures through customs and cultures and actively participate in traditional customs and cultural activities to promote the better development of customs and cultures. In order to study the signal transmission of neural connections in visual cortex, literature [13] can study the optical imaging of neural activity in rat cortex, which has the advantage of high spatial and temporal resolution. From this study, we can know that excitation is at least partially transmitted horizontally through vertical connections between neurons. Literature [14] focuses on the concept of learning vision in the process of knowledge transfer, which has become the most active research field at present. In order to explore the cross-category learning process, we can study the algorithm through cross-category tags, through which we can realize the knowledge dissemination between the source category and the target category and prove the effectiveness of this method. Literature [15] shows the intrinsic optical signals of frog retinal slices, which can work together by measuring different types of retinal neurons at the same time. Although the light sensing layer IOS is mainly limited to visible light, the inner retinal IOS has a characteristic near-to-far time history. Through the contents of the above literature, we find that the research points of traditional culture are more prominent, without using relevant research techniques and strategies for in-depth analysis. However, the application of digital technology is still analyzed in the relevant framework, and the visualization technology of traditional culture can further show the specific methods described in the article, and the visualization realizes the communication of folk culture and provides a suburban way.

2. Visual Analysis of Folk Customs with Digital Technology

2.1. Digital Development Status. Common digital technology types: application programming interface (API), artificial intelligence, big data, cloud computing, Internet of things, mobile technology, 5G technology, blockchain, etc. [16].

2.1.1. Digital Technology Artificial Intelligence. The artificial intelligence of digital technology operates individual-data behavior data by accumulating individual data; AI learns simulation and makes decisions for it. Through individual-behavior data, AI simulates group behavior and individual behavior. Finally, individual operation is carried out through individual behavior and group behavior [17]. Flow chart of artificial intelligence is shown as Figure 1.

The three modules are divided into individual data accumulation, AI learning simulation, and decision-making module. Accumulating individual data is to realize data collection module, and using the model needs enough data sets; the AI learning model simulates the collected data set to obtain the corresponding feature attributes; the decision-making module is to achieve the above simulation characteristics for decision-making results and give specific decision-making conclusions for the corresponding results.

2.1.2. Digital Technology Cloud Computing. Digital technology carries out basic features, development models, service models, and other computing models through cloud computing. The basic features include quick response and customized service. The development model includes private, public, community, and ecological aspects. Service model includes software, platform, infrastructure, and other services. Cloud computing is through digital technology to improve the computing efficiency of cloud computing-related models. Digital technology cloud computing diagram is shown as Figure 2.

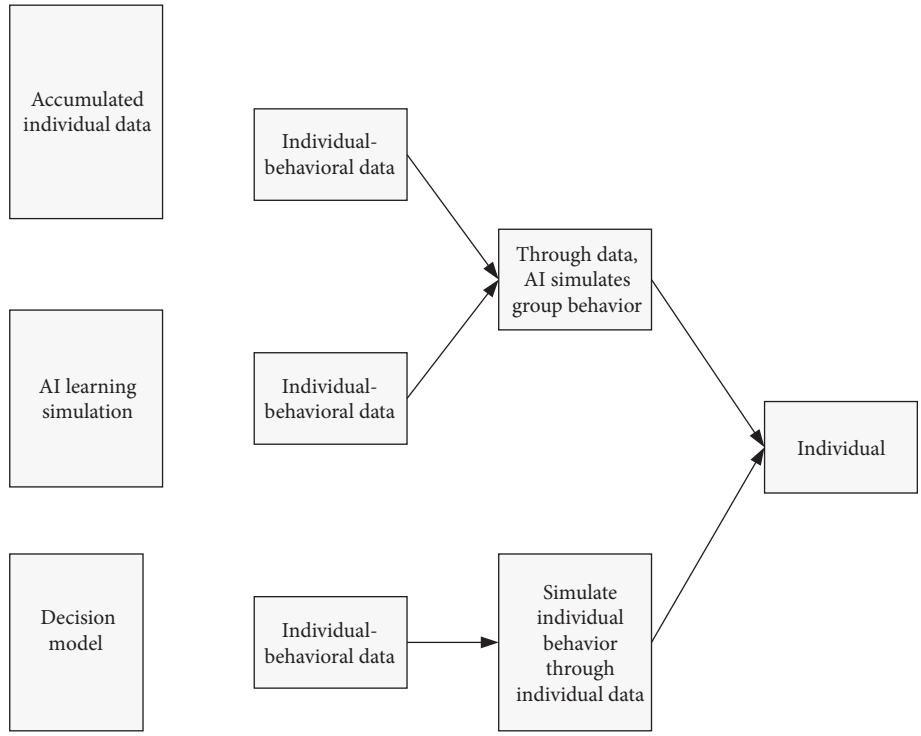


Figure 1: Flow chart of artificial intelligence.

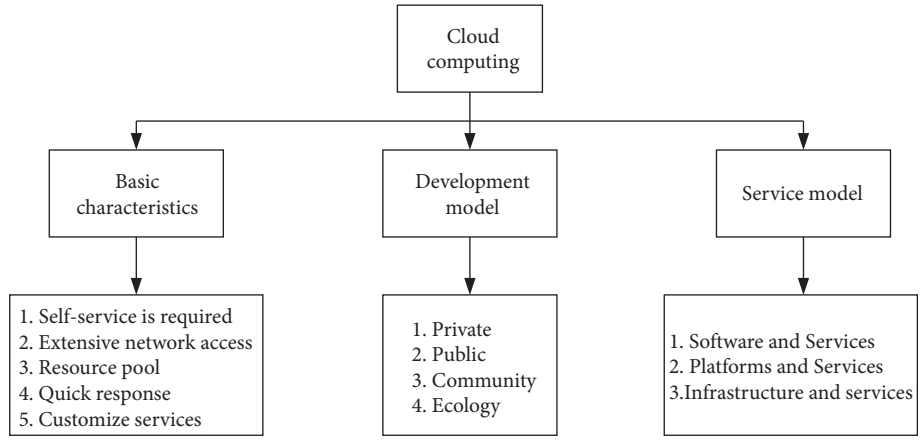


Figure 2: Digital technology cloud computing diagram.

2.2. Conceptual Model of Digital Technology Maturity. Digital maturity includes its digital readiness, intensity, and contribution. Among them, digital readiness includes strategic organization and infrastructure, through which digital technology is basically managed [18]. Digital intensity includes the management and integration of digital technology and improves the operation technology of digital technology through these two aspects. The digital contribution is reflected in the digital performance, through which the digital technology can be developed to a higher standard. Digital maturity is shown in Figure 3.

2.3. Promotion of Local Folk Culture. Local folk culture is a representative local folk culture formed by the gathering and mixing of people's customs of various nationalities in a region for several years. Through the folk culture of a place, we can learn about the living habits of local residents, local folk customs, and living habits [19]. Folk customs in various places nurture the local people and have an important influence on the spiritual values of the local people.

2.3.1. Protection of Folk Culture. Folk culture is an important symbol of local cultural customs, but in the continuous development of history, many folk cultures gradually declined and lost, which is the loss of local culture. We should strengthen the development of folk culture, combine folk culture with today, and strengthen people's understanding of folk culture. People should understand the history and culture of generations through the inherited folk

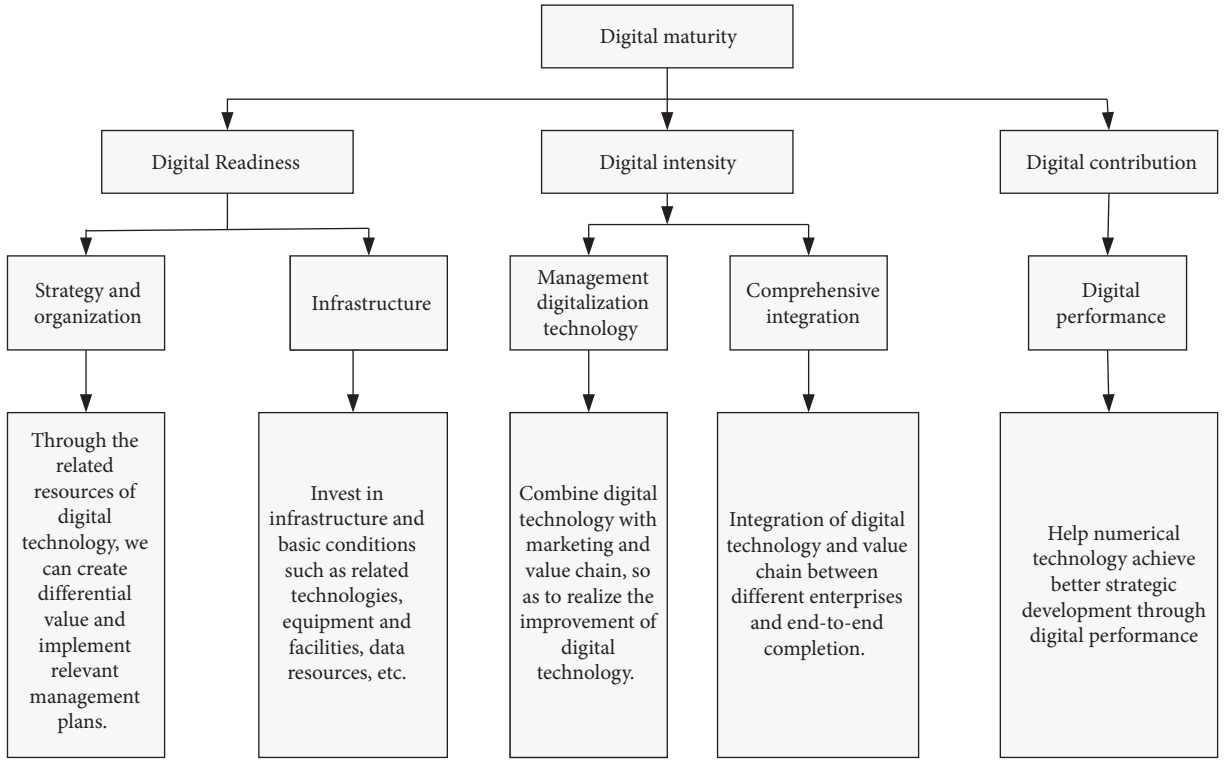


Figure 3: Digital maturity.

culture [20]. Therefore, we should expand the understanding of folk culture and then protect folk culture.

2.3.2. Promoting the Cognition of Folk Culture. Nowadays, many young people do not know enough about local folk culture and pay enough attention to it. The understanding of folk culture is limited to the cognition of living environment and parents' telling, and there are few opportunities to really contact folk culture, which leads to the existence of some people who do not understand or even know what local folk culture is. Therefore, more young people can understand and love folk culture through education and popularization.

2.3.3. Inheritance of Folk Culture. Many folk cultures have strict conditions for inheriting their culture, so special inheritors should be trained according to local folk customs. We should cultivate their cognition of local folk culture and their operation skills of folk culture [21]. Only by cultivating folk culture for a long time , can we have enough cognition of local folk culture, and can we keep inheriting national culture and not be declined.

3. Research on Digital Technology and Visual Communication

3.1. Digital Technology

3.1.1. Digital Level Measurement. v_{it} represents the original data, the maximum value of the data investigated by $v_{\max}$, and q and k are set by themselves in the distribution interval.

$$X_i = \frac{V_{it} - V_{\min 0}}{V_{\max} - V_{\min 0}} \times k + q. \tag{1}$$

M: $x_1, x_2, \ldots, x_m$. There are n evaluation objects in total.

$$\begin{aligned} \overline{x}_{ij} &= \frac{x_{ij} - \overline{x}_{ij}}{s_j} \ (i = 1, 2, \&, n, \ j = 1, 2, \&, m), \\ \overline{x}_j &= \frac{1}{n}\sum_{i=1}^{n} x_{ij}, \\ s_j &= \sqrt{\frac{1}{(n-1)}\sum_{i=1}^{n}\left(x_i - \overline{x}_j\right)^2}, \end{aligned} \tag{2}$$

where $\overline{x}_j$ is the sample mean and standard deviation s_j of the J-th index [22].

$$R = (r_\delta)_{m\times m},$$

$$r_\delta = \frac{\sum_{k=1}^{n} \overline{x}_{ki} \cdot \overline{x}_{kj}}{n-1} (i,\ j = 1, 2, \&, m),$$

$$r_\delta = 1,$$

$$r_\delta = r_{ji},$$

$$\begin{cases} y_1 = u_{11}\overline{x}_1 + u_{21}\overline{x}_2 + \ldots + u_{n1}\overline{x}_n \\ y_2 = u_{12}\overline{x}_1 + u_{22}\overline{x}_2 + \ldots + u_{n2}\overline{x}_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ y_m = u_{1m}\overline{x}_1 + u_{2m}\overline{x}_2 + \ldots + u_{mm}\overline{x}_n \end{cases} ' b_j y_j. \quad (3)$$

$$b_j = \frac{\lambda_j}{\sum_{k=1}^{m} \lambda_2} (j = 1,\ 2, \&, m),$$

$$a_p = \frac{\sum_{k=1}^{y} \lambda_i}{\sum_{k=1}^{m} \lambda_i},$$

$$Z = \sum_{j=1}^{p}$$

3.1.2. Digital Technology Artificial Intelligence. (1) Bayesian

$$\begin{aligned} \mathrm{LS} &= \frac{P(E|H)}{P(E|H)}, \\ \mathrm{LN} &= \frac{P(E|H)}{P(E|H)} = \frac{1 - P(E|H)}{1 - P(E|H)}, \\ P(A|B) &= \frac{P(A \cap B)}{P(B)}, \\ P(B|A) &= \frac{P(A \cap B)}{P(A)}, \\ P(A|B)P(B) &= P(A \cap B) = P(B|A)P(A), \\ P(A_i \mid B) &= \frac{P(A_i)P(B|A_i)}{\sum_{j=1}^{n} P(A_j)P(B|A_j)}. \end{aligned} \quad (4)$$

(2) Probability function

$$\begin{cases} D(x) = \dfrac{P(x)}{1 - P(x)} = \dfrac{P(x)}{P(x)}, \\ P(x) = \dfrac{D(x)}{1 + D(x)}, \\ D(H|E) = LS \times D(H), \\ D(H|E) = LN \times D(H). \end{cases} \quad (5)$$

(3) Uncertain reasoning

Evidence is determined to appear

$$\begin{aligned} P(E) &= 1, \\ P(H|E) &= \frac{\mathrm{LS} \times P(H)}{(\mathrm{LS} - 1) \times P(H) + 1}. \end{aligned} \quad (6)$$

The evidence is determined not to appear

$$\begin{aligned} P(E) &= 0, \\ P(H|E) &= \frac{\mathrm{LN} \times P(H)}{(\mathrm{LH} - 1) \times P(H) + 1}. \end{aligned} \quad (7)$$

3.2. Trajectory Data Visualization.

$$P_i = (x_i, y_i, t_i) i = 1,\ 2,\ 3, 4,$$

$$\begin{cases} \mathrm{Row} = \dfrac{2y_i}{r} \\ \qquad\qquad i = 1, 2, 3, 4, \cdots, \\ \mathrm{Column} = \dfrac{2x_i}{r} \end{cases}$$

$$\begin{cases} X = \dfrac{\sum_{i=1}^{n} x_i z_i}{\sum_{i=1}^{n} z_i} \\ Y = \dfrac{\sum_{i=1}^{n} x_i z_i}{\sum_{i=1}^{n} z_i}\ i = 1, 2, 3, 4, \ldots, n \\ Z = \sum_{i=1}^{n} z_i \end{cases} \quad (8)$$

3.3. Digital KANO Model.

$$\text{Better} = \frac{(A+O)}{(A+O+M+I)},$$
$$\text{Worse} = -1 * \frac{(O+M)}{(A+O+M+I)}. \tag{9}$$

Its model includes five functions, function 1 is in the second quadrant, functions 2 and 5 are in the first quadrant, and functions 3 and 4 are in the fourth and third quadrants, respectively. The first quadrant is also called one-dimensional demand, the second quadrant is also called charm demand, the third quadrant is necessary demand, and the fourth quadrant is indifference demand [23] in Figure 4.

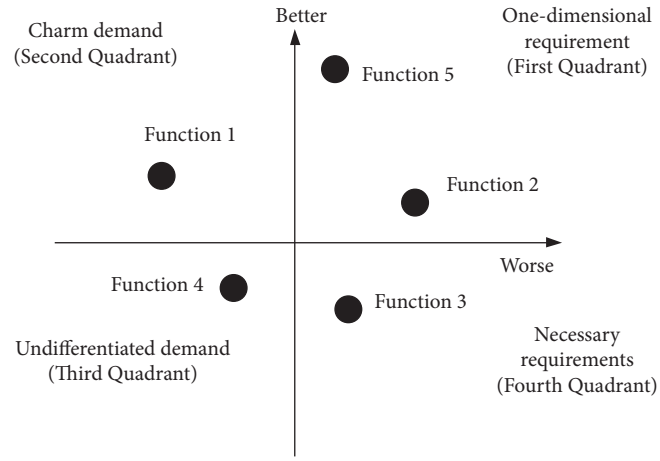


Figure 4: Quadrant diagram of KAON model.

4. Visual Analysis of Digital Local Folk Culture

4.1. Analysis of Digital Technology

4.1.1. Digital Artificial Intelligence. Through the investigation and analysis of digital artificial intelligence technology in each district, it can be seen that Region 2 ranks first with a growth rate of 379.74% and Region 1 ranks second with a growth rate of 377.51%, compared with Region 5 with a growth rate of 166.08% and Region 6 with a growth rate of 123.88%. Through data observation, it is known that the total index of digital artificial intelligence in Region 2 and Region 1 develops well, while that in Region 5 and Region 6 lags behind. However, according to the growth rate, it can be seen that the growth rate of the first to fifth total indexes in the regions exceeds 100%, indicating that the development prospect of digital artificial intelligence is more objective in several regions surveyed, and the total growth rate has increased. Comparison of total indexes is shown in Table 1.

Table 1: Comparison of total indexes.

Rank	Growth rate (%)	Economy
1	379.74	Area 2
2	377.51	Area 1
3	205.40	Area 3
4	168.34	Area 4
5	166.08	Area 5
6	123.88	Area 6

According to the AI environmental support, as shown in Table 2, in the growth rate ranking, Region 2 ranks first with 559.97%, and Region 1 ranks second with 537.07%. The three regions and six regions rank fifth and sixth, respectively, with growth rates of 71.15% and 59.40%, respectively. According to the overall chart perception, the growth rates of Regions 1 and 2 exceeded 500%, while those of Regions 4 and 5 remained above 100%, while those of Regions 3 and 6 did not exceed 100%. It can be seen from the analysis that the development of AI environmental support is unbalanced, and the development trends of the first and second regions are excellent, the fourth and fifth regions are good, and the third and sixth regions are qualified. It should be rectified according to the development methods and factors of the third and sixth regions corresponding to the first and second regions so that the growth rate of the third and sixth regions will rise greatly.

Table 2: AI environmental support.

Rank	Growth rate (%)	Economy
1	559.97	Area 2
2	537.07	Area 1
3	128.96	Area 4
4	117.15	Area 5
5	71.15	Area 3
6	59.40	Area 6

It can be seen from Table 3 that the fifth, first, second, and fourth regions rank first to fourth, respectively, and their AI industry competitiveness growth rates are 754.42%, 540.41%, 474.33%, and 404.22%, respectively. The AI industry competitiveness of these four regions all exceed 400%, and their development trend is excellent. However, the growth rate of the third and sixth regions is about 211%, and the development trend is good. Compared with the first four regions, the development of single phase has fault phenomenon. We should study and analyze the top four regions in response to development and find the relevant influencing factors of their growth rate to improve [24].

Table 3: AI industry competitiveness.

Rank	Growth rate (%)	Economy
1	754.42	Area 5
2	540.41	Area 1
3	474.33	Area 2
4	404.22	Area 4
5	210.92	Area 3
6	210.91	Area 6

Table 4 is a table of AI knowledge creativity growth rate. From this table, it can be seen that Region 1 ranks first with a growth rate of 165.44%, Region 2 ranks second with 71.53%, Region 5 ranks fifth with −15.70%, and Region 3 ranks sixth with −35.90%. In contrast, the growth rate of Area 1 is over 100%, and the development trend is good, but the growth

Table 4: AI knowledge creativity.

Rank	Growth rate (%)	Economy
1	165.44	Area 1
2	71.53	Area 2
3	27.22	Area 4
4	27.21	Area 6
5	−15.70	Area 5
6	−35.90	Area 3

rates of Area 5 and Area 3 are −15.70% and −35.90%, respectively, which shows that its development trend is retrogressing, so we should find out the reasons from the root to make it develop positively. Comparing AI knowledge creativity table with the other three tables, we can see that its development trend is lower than its environmental support, industrial competitiveness, and total index growth rate, and its development is relatively backward, so we should pay attention to the development of AI knowledge creativity.

According to the ranking (Figure 5) of the index growth rate, compared with AI environmental support and AI industry competitiveness, its total index and AI knowledge creativity develop smoothly in each district, without large-scale fault phenomenon, and develop smoothly. However, the AI environmental support has obvious fault phenomenon between regions 1 and 2 and regions 3, 4, 5, and 6, and the image is steep, which shows that the development trend of regions 1 and 2 is quite different from that of other regions. Several regions with low growth rate should refer to the development methods of regions 1 and 2 to improve their own development rate. However, the competitiveness image of AI industry is also steep, which is mainly aimed at between the three regions and the five regions. It should be adjusted for the development of this region by analyzing the regions with good development trends.

4.1.2. Age Distribution of Digital Effect. Age digitization reflects a better distribution effect map. It can be seen from Figure 6 that only about 4% of the people under 18 years old, 51–55 years old, and over 56 years old are digital audience. Its main audience is distributed in the age range of 19–50 years old, among which 19–25 years old and 36–40 years old are the majority. Through the analysis of this table, it is known that the number of people under 18 years old and over 51 years old is small. By analyzing the reasons, it can be seen that people under 18 years old and over 51 years old should share and understand their digital technology, so as to improve the audience level and perception of digital technology.

4.1.3. Educational Background of Digital Effect Audience. From Figure 7, the educational background of digital audience is mainly concentrated in high school, junior college, and undergraduate, while in primary school, junior high school, master's degree, and above, the effect of digitalization is relatively small. According to the corresponding data, digital technology should be popularized for those with primary school, junior high school, master's degree, or above, so as to comprehensively improve the audience rate of digital technology. We should make digital technology develop in all aspects.

4.1.4. Investigation on the Homework Situation of Digital Effect Audience. From this Figure 8, it can be seen that most of the audiences are enterprise staff, and there is less contact with digital technology in the work of political parties. It shows that there are many aspects that can be applied to digital technology in enterprise work, which is of great help to the work of enterprise staff and improves their work efficiency. At the same time, it is seldom used in the work of political parties, which shows that digital technology has not penetrated into all working levels. Digital technology should be popularized for jobs with less contact with digital technology to improve their work efficiency.

4.2. Visual Analysis

4.2.1. Comparison of Visual Loading Time. By comparing the loading time of pictures before and after visual clustering, it can be seen from Figure 9 that the maximum time before clustering is about 600, and the minimum is about 550. After clustering, the highest loading time is about 400, and the lowest is about 290. The highest time difference between clustered and unclustered pictures is about 200, and the lowest time difference is about 260, forming a cliff-like reduction. According to the data, after clustering processing, the image is loaded by visualization, which reduces the loading time in a large area, greatly improves the loading time of visual images, and improves the loading efficiency of images.

4.2.2. Visual Efficiency Analysis Chart. Image visualization has developed from the germination of Figure 10 in the sixteenth century to the physical measurement used in the seventeenth century, image symbols in the eighteenth century, data images in the nineteenth century, modernization enlightenment in the first half of the 20th century, and the second half of the 20th century. After a long period of development, its related efficiency has been greatly improved from about 50% of the budding efficiency in the sixteenth century to about 98% now. In this slow development process, visualization technology is constantly improving, its efficiency is constantly improving, and its accuracy rate is also constantly improving [25].

4.3. Analysis of Local Folk Culture

4.3.1. Understanding of Folk Culture Policy. According to the random inquiry of local residents' understanding of local folk culture policies, the clarity accounted for 30% in 2013, 35% in 2015, 42% in 2017, and 48% in 2019. From 2013 to 2019, the clarity of surveyed residents increased year by year, indicating that the inheritance and promotion of local folk culture are improving. It is completely unclear that from 10% in 2013 to 5% in 2015 to 1% in 2017 and 2019, it has also

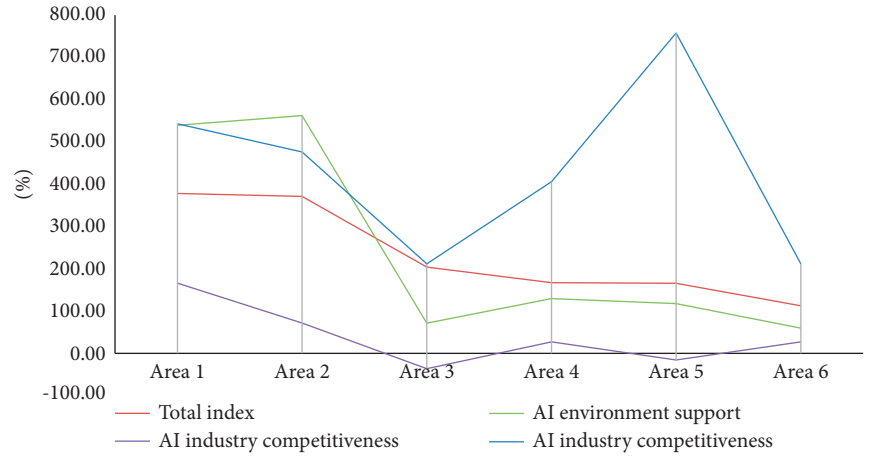


Figure 5: Ranking of artificial intelligence index growth rate.

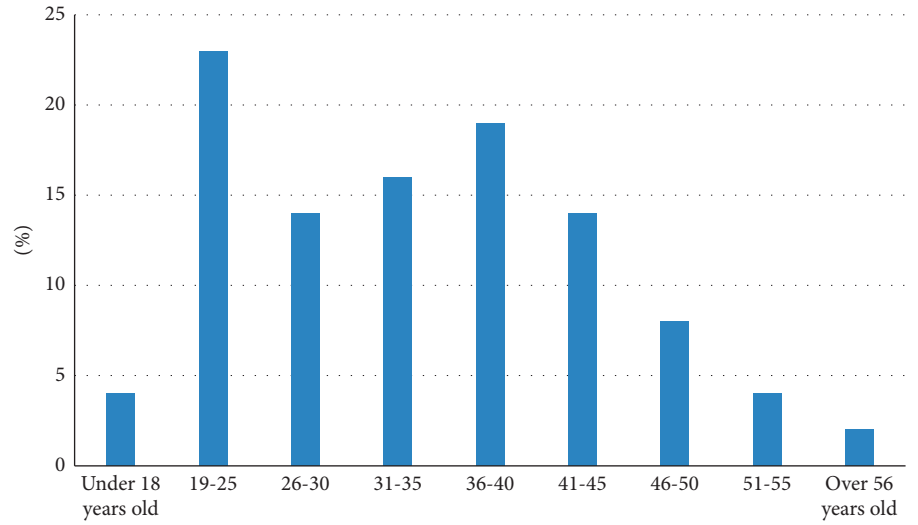


Figure 6: Age distribution of audience.

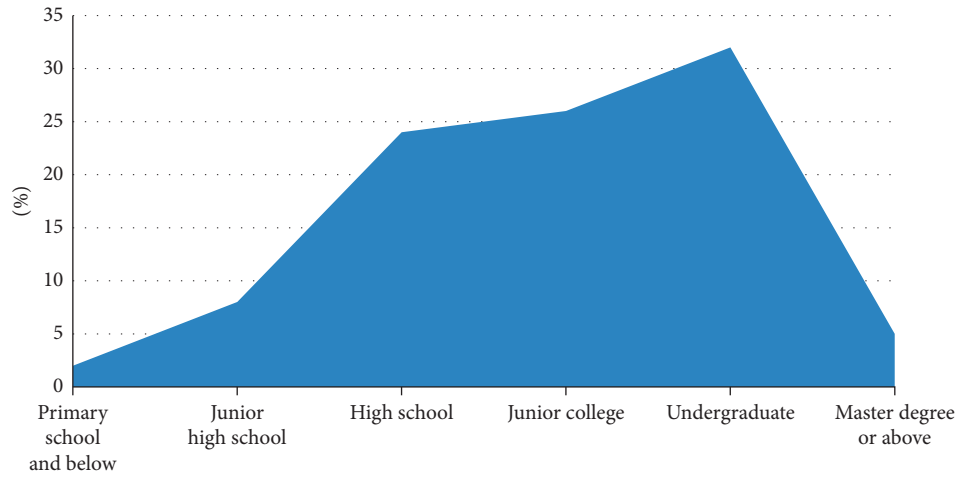


Figure 7: Educational background of respondents.

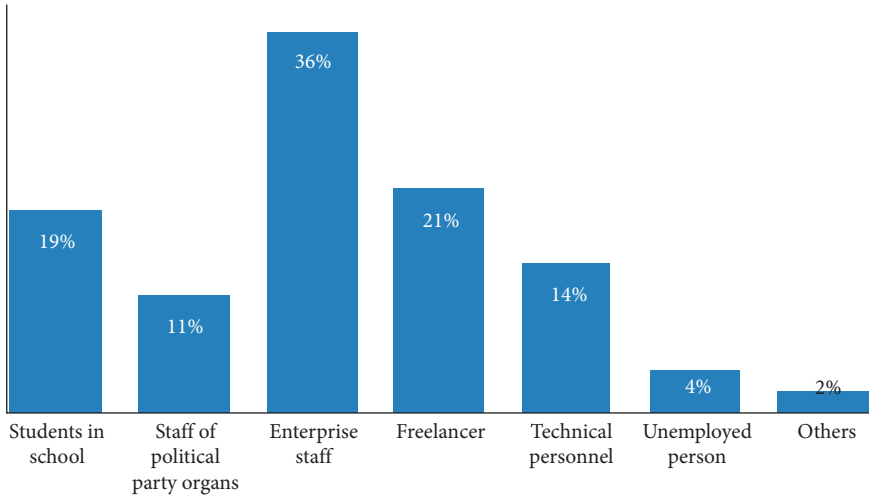


Figure 8: Survey of audience work.

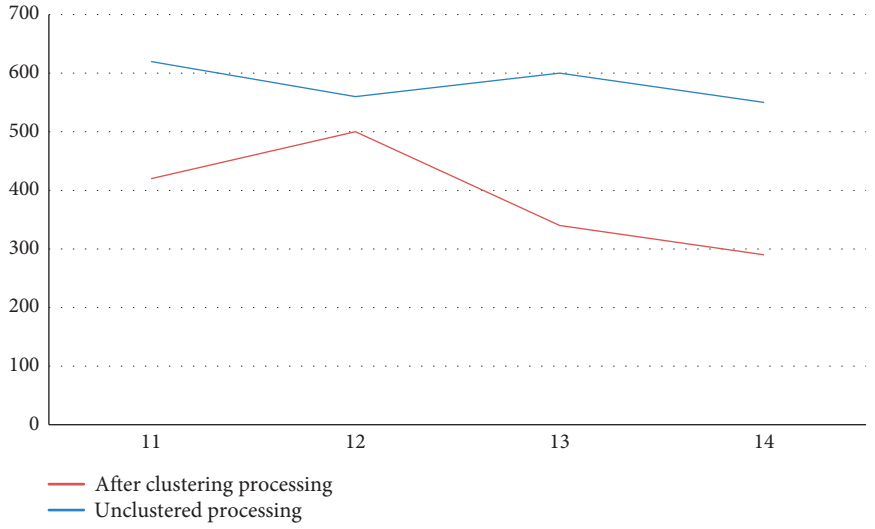


Figure 9: Comparison of image loading time.

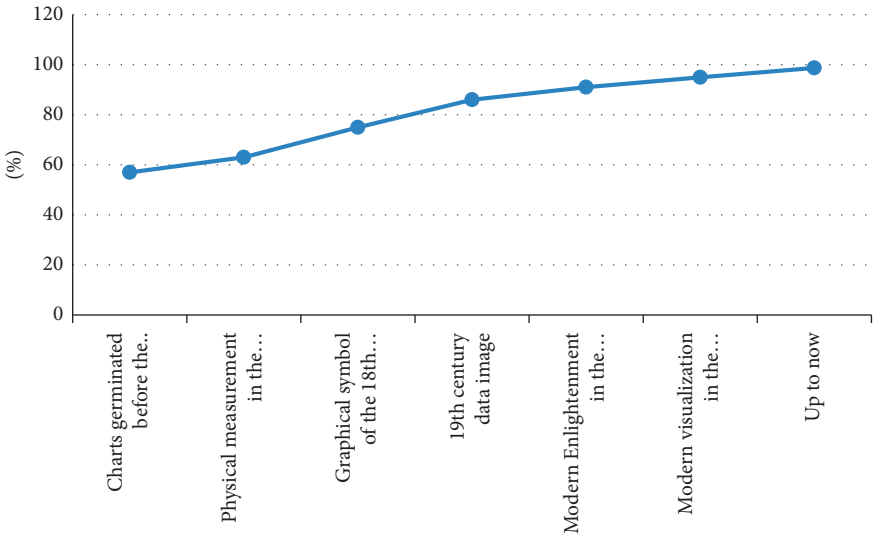


Figure 10: Visual efficiency analysis diagram.

Table 5: Understanding of folk culture policy.

Degree	2013 (%)	2015 (%)	2017 (%)	2019 (%)
Clear	30	35	42	48
Not quite clear	42	36	28	12
Be clear	18	24	29	39
Completely unclear	10	5	1	1

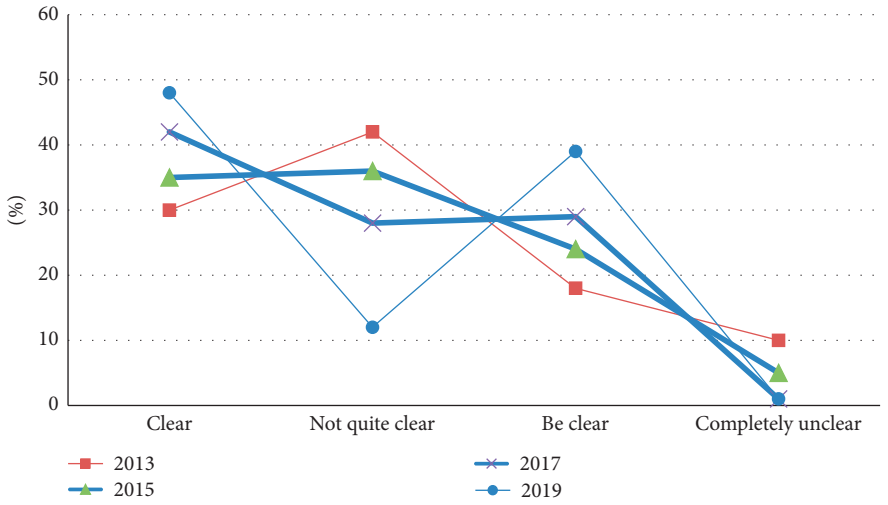


Figure 11: Investigation on the understanding of local folklore policy.

dropped sharply. It shows that the local people pay more and more attention to local folk culture, and their cognition of local folk culture is also improving. Understanding of folk culture policy is shown in Table 5.

It can be seen from Figure 11 that the proportion of clarity is increasing from 2013 to 2019, the unclear and completely unclear are decreasing year by year, and the relative clarity is also increasing. From the data, we can see that the development of local folk culture has become a good development trend.

4.3.2. Insufficient Promotion of Local Culture. Through the investigation, it can be seen that the development of local folk culture mainly faces the following problems: the governance system is uncoordinated, which is mainly due to the implementation of local folk culture, the lack of proper arrangement of venues for holding activities and citizens' participation, improper coordination, and low efficiency of activity arrangement and participation. The lack of governance measures focuses on the failure to register related folk activities, provide funding for related activities, and effectively publicize folk cultural activities. The unfairness of governance means is reflected in the participants of folk culture, whether it is carried out openly, whether the safety factors are considered carefully, and whether the funds for activities are unified. The ineffective feedback of governance effect is reflected in the failure to effectively feedback to relevant management departments after rectification through the above related factors, and there are ineffective

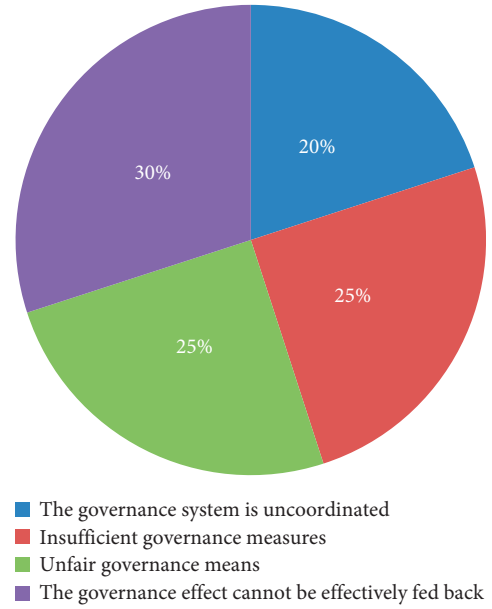


Figure 12: Insufficient promotion of local culture.

improvements, which is an unscientific and insufficient form of feedback. Insufficient promotion of local culture is shown in Figure 12.

The article shows the visual effect of folk culture communication by visualization. The data of this paper is carried out in the form of investigation, and the data are visually analyzed to analyze the current communication effect of folk

culture in the region and the communication effect among different groups of people. The analysis results can directly show that folk culture has different communication effects in different regions, different groups, educational background, and other factors, so as to provide an effective way of communication.

5. Conclusion

Through the visual analysis of digital technology and local folk culture, we know that there are still some problems in the process of promoting the development of local folk culture, such as management system, governance means, and governance effect, which cannot be fed back. We should make relevant development routes through deep understanding of local folk culture, so as to protect local folk culture, promote its development, promote cognition, and inherit folk culture.

Data Availability

The experimental data used to support the findings of this study are available from the corresponding author upon request.

Disclosure

The research result of the project of Hebei Water Ecological Civilization and Social Governance Research Center: "Research on the Value and Enlightenment of Rural Social Governance in the Construction of Beautiful Villages and Characteristic Towns—A Case Study of Handan," project no. 2018SZX1.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding this work.

References

[1] W. K. Hocking, B. Fuller, and B. Vandepeer, "Real-time determination of meteor-related parameters utilizing modern digital technology," *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 10, no. 2, pp. 9–17, 2001.

[2] P. A. David, "Understanding digital technology's evolution and the path of measured productivity growth: present and future in the mirror of the past," *Papers*, vol. 8, no. 02, pp. 20–28, 2000.

[3] C. K. Blackwell, A. R. Lauricella, and E. Wartella, "Factors influencing digital technology use in early childhood education," *Computers & Education*, vol. 77, no. aug, pp. 82–90, 2014.

[4] G. S. Lunney, "The death of copyright: digital technology, private copying, and the DMCA," *Virginia Law Review*, vol. 87, no. 5, pp. 17–32, 2001.

[5] G. S. Mort and J. Drennan, "Mobile digital technology: e," *The Journal of Database Marketing & Customer Strategy Management*, vol. 10, no. 1, pp. 9–23, 2002.

[6] D. Smahel, S. Chaudron, M. E. Beutel, and M. Šmahelová, "Young Children (0-8) and digital technology: a qualitative exploratory study across seven countries," *JRC*, vol. 5, no. 12, pp. 26–40, 2015.

[7] R. K. Lawrence and A. T. Kelly, "Single event effect induced multiple-cell upsets in a commercial 90 nm CMOS digital technology," *IEEE Transactions on Nuclear Science*, vol. 55, no. 6, pp. 3367–3374, 2009.

[8] W. C. Slater and T. J. Murray, "Photofinishing system and method incorporating digital technology," United States Patent Application Publication, Alexandria, VA, USA, 2003.

[9] Y. L. Ning, Q. Y. Pang, and X. M. Xie, "On the role of local folk culture in cultivating socialist core values in the new era: taking the junpo festival of hainan for example," *Journal of Zhengzhou Normal Education*, vol. 43, no. 21, pp. 19–32, 2018.

[10] J. Liang, "The poetic imagination and theoretical construction of local folk culture:review of studies of pluralistic folk culture circle in qinghai," *Northwestern Journal of Ethnology*, vol. 53, no. 12, pp. 89–102, 2014.

[11] L. I. Hong-Qi, "Academies in song and yuan dynasty and local culture: Academy, academic and folk religions in jizhou region," *Journal of Hunan University*, vol. 34, no. 7, pp. 97–121, 2006.

[12] K. Liu and D. Zhen, "Protection and display of festival activities and local folk culture——based on the research of shoton festival in Lhasa," *Journal of Leshan Normal University*, vol. 20, no. 5, pp. 17–25, 2018.

[13] M. Tanifuji, T. Sugiyama, and K. Murase, "Horizontal propagation of excitation in rat visual cortical slices revealed by optical imaging," *Science*, vol. 266, no. 5187, pp. 1057–1059, 1994.

[14] G. J. Qi, C. C. Aggarwal, R. Yong, and Q. Tian, "Towards cross-category knowledge propagation for learning visual concepts," in *Proceedings of the Computer Vision & Pattern Recognition*, vol. 27, no. 6, pp. 103–133, IEEE, Colorado Springs, CO, USA, June 2011.

[15] Y.-C. Li, C. Strang, F. R. Amthor et al., "Parallel optical monitoring of visual signal propagation from the photoreceptors to the inner retina layers," *Optics Letters*, vol. 35, no. 11, pp. 1810–1812, 2010.

[16] L. Molteni and A. Ordanini, "Consumption patterns, digital technology and music downloading," *Long Range Planning*, vol. 36, no. 4, pp. 389–406, 2003.

[17] I. Stevenson, "Tool, tutor, environment or resource: e," *Computers & Education*, vol. 51, no. 2, pp. 836–853, 2008.

[18] P. A. David, "A tragedy of the public knowledge 'common'? Global science, tellectual property and the digital technology boomerang," *International Institute of Infonomics*, vol. 18, pp. 1–40, 2000.

[19] L. R. Baker, "Links between local folklore and the conservation of sclater's monkey (*Cercopithecus sclateri*) in Nigeria," *Afr Primates*, vol. 8, pp. 17–24, 2013.

[20] T.-S. Yang and M.-C. Oh, "A study on impacts of selection attribute of j local folklore food on customers' behaviors -focusing on customer satisfaction, Re-visit, and word of mouth of jt," *Journal of the Korean Society of Food Science and Nutrition*, vol. 38, no. 5, pp. 636–643, 2009.

[21] E. J. Blown and T. Bryce, "The enduring effects of early-learned ideas and local folklore on children's astronomy knowledge," *Research in Science Education*, vol. 54, no. 12, pp. 118–123, 2018.

[22] A. E. Miller and D. M. West, "Where's the revolution? Digital technology and health care in the Internet age," *Journal of Health Politics, Policy & Law*, vol. 43, no. 2, pp. 19–27, 2009.

[23] C.-C. Chen and M.-C. Chuang, "Integrating the Kano model into a robust design approach to enhance customer

satisfaction with product design," *International Journal of Production Economics*, vol. 114, no. 2, pp. 667–681, 2008.

[24] G. Dennis, B. T. Sherman, D. A. Hosack et al., "DAVID: database for annotation, visualization, and integrated discovery," *Genome Biology*, vol. 40, no. 3, pp. 78–99, 2003.

[25] T. Helga, J. T. Robinson, and J. P. Mesirov, "Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration," *Briefings in Bioinformatics*, vol. 14, no. 2, pp. 178–192, 2013.

Hindawi
Scientific Programming
Volume 2022, Article ID 3138851, 10 pages
https://doi.org/10.1155/2022/3138851

Research Article

Music Recognition and Classification Algorithm considering Audio Emotion

Wang Na and Fang Yong

School of Music, Beihua University, Jilin City 132013, Jilin Province, China

Correspondence should be addressed to Fang Yong; jl_fy2855@beihua.edu.cn

Received 28 October 2021; Revised 8 December 2021; Accepted 21 December 2021; Published 20 January 2022

Academic Editor: Baiyuan Ding

At present, the existing music classification and recognition algorithms have the problem of low accuracy. Therefore, this paper proposes a music recognition and classification algorithm considering the characteristics of audio emotion. Firstly, the emotional features of music are extracted from the feedforward neural network and parameterized with the mean square deviation. Gradient descent learning algorithm is used to train audio emotion features. The neural network models of input layer, output layer, and hidden layer are established to realize the classification and recognition of music emotion. Experimental results show that the algorithm has good effect on music emotion classification. The data stream driven by the algorithm is higher than 55 MBbs, the anti-attack ability is 91%, the data integrity is 83%, the average accuracy is 85%, and it has good effectiveness and feasibility.

1. Introduction

In the field of music information retrieval, affective recognition and classification have been widely concerned. Emotional classification of music is a multidisciplinary field of research. It perceives emotions from songs and classifies them into specific emotional categories. Emotional recognition depends on acoustic features, linguistic features, and metadata-based filtering [1]. Genre of music can be classified to describe the type and characteristics of music, and music characteristics can be used to construct more digital form of music, so it is of great significance to music information retrieval. In the past classification methods of audio types, music emotion classification has always been performed manually, and music types have been identified by artificial emotion. However, this method will cause subjective evaluation errors due to the variety of instrumental music equipment and music rhythm structure and form related to statistical types. Therefore, it is necessary to describe the classification of audio signals with the help of automatic emotion classification algorithm [2]. The performance of these musical features can be evaluated by training statistical pattern recognition. A multi-type user interface browsing audio collection system based on automated classification has been developed to fully enhance human interaction. In the process of life and work, music emotion classification is mostly used in mobile music stores, radio stations, the Internet, and other terminals. Through the description of music emotion, it can effectively select the music type suitable for the current play [3, 4]. Although there may be some subjectivity and arbitrariness in the emotional classification of music, it is possible to classify music according to instrument type and rhythmic structure. However, due to the different subjective initiative and evaluation standards of individuals, there will be different classification standards for the same piece of music. Therefore, it is necessary to make use of the information processing ability of network big data to classify music types through specific templates and expand the music repertoire of the system database. In the previous analysis, it is found that automatic music type classifier can be used to improve the ability of music emotion recognition and classification by automatically indexing and retrieving audio data. The classification behavior is performed by selecting the frequency band that should be attenuated or re-attenuated according to the automatic music type classifier, through the type label assigned to the signal.

Music emotion classification can effectively make up for the lack of retrieval based on lyrics, singer names, and other ways and has become a research hotspot in recent years. Because multimedia information data cannot be directly matched with human emotion perception, it is necessary to analyze the relevance of music information data to human emotion from its own characteristics [5, 6]. At present, music emotion classification and recognition are mainly based on audio signals and lyrics, which are two different forms of data features [7]. The relationship between audio signals and music emotion is analyzed by machine learning method, and the characteristics of music rhythm and rhythm are discovered [8–10], so as to achieve the effective recognition of music emotion.

The essence of music is the latent emotion, and the recognition of music emotion is satisfied according to the mapping relationship between music characteristics and emotion, which is an effective promotion of music emotion recognition and can strengthen the effect of human-computer emotional interaction [4, 11]. At present, the classification and recognition of music emotion is mainly based on the analysis of data features in the form of music text and audio signal. Chaudhary et al. [2] proposed a spectral approach to transform musical features into visual representations using convolution neural networks, using CNN to extract music signal from the song as a specific feature, and then through the comparison of features of the song emotional classification. Quasim et al. [10], combining emotional maturity with existing Internet of Things systems, proposed an affection-based music recommendation and classification framework that allows for high precision classification of songs with memories and emotions. Chen et al. [12] aiming at the deficiency of single network classification model proposed a multi-feature combination network classifier based on CNN-LSTM to improve the effect of music emotion classification. By using convolution kernels to extract 2D musical information and using serialization to output single modal emotion classification, the heterogeneous features are effectively combined to improve the classification performance. Rajesh et al. [13] believed that music is one of the effective media to convey emotion, so the emotion recognition in music fragment is of great significance to the understanding of music. Therefore, a classification recognition method of music emotion based on deep learning technology is proposed. The Mel frequency cepstrum coefficient, chrominance energy normalization statistic, and chrominance short time Fourier transform are extracted from the musical instrument data set, and the neural network is trained to recognize emotion. Wang D et al. [1] using intelligent recognition and classification algorithm of human emotion studied different emotion classifications under different audio signals. In Haridas AV et al. [14] through the Taylor Series Deep Belief Network emotion recognition system, the noise in audio signal is removed, and then emotional information feature is extracted. Based on the above research, this paper presents a new algorithm of music emotion recognition and classification based on forward neural network, which is used to classify the lyrics and rhythm. The parameterized audio emotion feature is trained by mean square error and gradient descent algorithm, and the neural network is constructed by input layer, output layer, and hidden layer to recognize and classify music emotion feature. Experimental results show that the proposed method has strong classification accuracy and can effectively identify the multiple classification criteria and emotional complexity of musical emotion.

2. Music Emotion Feature Extraction

A piece of music mainly includes audio information and lyrics text information, and audio information includes time domain features, frequency domain features, and cepstrum features. Time-domain features usually refer to the time-domain parameters calculated in a fixed length of music, mainly including short-time energy, short-time average amplitude, and average amplitude difference function. Frequency-domain features refer to the characteristic parameters obtained by converting time-domain signals into frequency-domain signals through Fourier transform, mainly including spectral centroid, spectral rollover, and spectral flux. Cepstrum features include audio information and lyrics text information, and audio information includes time domain, frequency domain, and cepstrum features by using different loudness and tones of human ears. Time-domain features usually refer to the time-domain parameters calculated in a fixed length of music, mainly including short-time energy, short-time average amplitude, and average amplitude difference function. Frequency-domain features refer to the characteristic parameters obtained by converting time-domain signals into frequency-domain signals through Fourier transform, mainly including spectral centroid, spectral rollover, and spectral flux. Cepstrum features, which use the human ear on the sound of different loudness, tone, and timbre, have different perceptual effects. Therefore, it can be used to classify different musical emotions. The specific operation methods are as follows:

(1) The music is divided into specific lengths and transformed into frequency domain signals by Fourier transform.

(2) Calculating the logarithm of the signal in frequency domain and then inverse Fourier transform, the cepstrum characteristic of the music can be extracted.

In addition to the rhythmic characteristics of music, audio messages also imply certain emotions.

The characteristics of music can be divided into basic features, complex features, and overall features [15], as shown in Figure 1. The basic characteristics of music include the following: the complex characteristics of music include rhythm, melody, harmony, etc. The overall characteristics include musical form structure, musical style, emotional connotation, etc. [16]. Accordingly, the recognition of musical features can be divided into three levels: the first is to extract the basic characteristics of music, the second is to get the complex characteristics based on the analysis of basic characteristics, and the last is to identify the overall characteristics of music according to the basic and complex

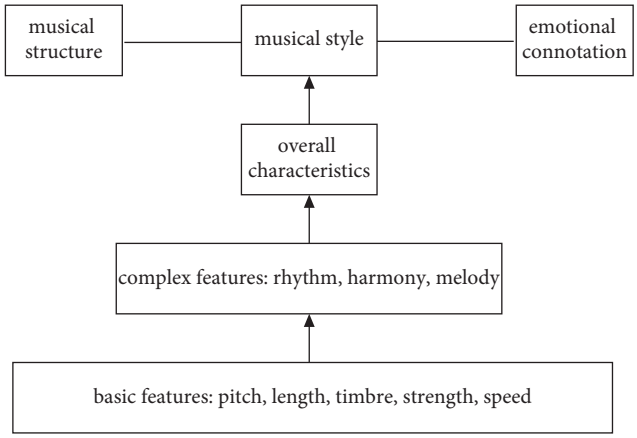


Figure 1: Composition of musical emotional characteristics.

characteristics. Firstly, the basic features of music are extracted and the complex features of music are calculated by using the feedforward neural network. Then, the emotion space is established according to the identified audio emotion features, and the emotion recognition model is constructed by using the emotion feature space to complete the classification of music emotion.

2.1. Overall Steps of Music Emotion Classification. Through constructing the music emotion classifier, complete the music emotion concrete classification. Using FNN [17] as the music intelligent classifier, the FNN can realize the multi-feature fusion of music emotion appreciation effectively by forming multi-modal data from different modal data.

The process of music emotion classification is as follows: a record is randomly selected from the database to find the similar information in the database, and the result is fed back to the system in the form of annotated record. Using these annotation records, the system trains the music emotion classifier until the music emotion segment search error reaches a controllable range. Then, the music emotion is labeled in the classifier.

The steps of music emotion classification are as follows:

Step 1. Extract feature vectors $d_1, d_2, \ldots d_n$ from n music fragments in the library, where $d_i = \{d_{i1}, d_{i2}, \ldots d_{im}\}$.

Because the physical meanings of the m components in d_i are different, we use Gaussian function [18] to normalize them:

$$d_i = \frac{1}{2} \times \frac{d_i - u_i}{e_i}, \quad 1 \le i \le m. \tag{1}$$

In the formula, u_i is the mean and e_i is the standard deviation, $m = 25$.

Step 2. The d_1 of one of the n music segments is randomly selected and its characteristic is vector d_{i1}. Calculate other music fragments in the library from d_1:

$$d = \sqrt{(d_i - e_i)}. \tag{2}$$

Return the distance calculation result to the system in turn.

Step 3. The system judges these results, marks the information with similar feelings to the system, and marks them in the database.

Step 4. In order to train music emotion classifier, the feature of music segment marked as similar emotion information in database is used as training sample.

Step 5. The trained music emotion classifier is used to classify the unlabeled music segments in the database and transmit the same type of emotion information to the user. The user judges the same type of emotion segment. If all of these segments are similar emotion segments, then after marking the segment in the database, the segment will jump to Step 6, and if the segment is different emotion segment, the segment will jump to Step 3.

Step 6. Update the tagged music fragment information to the new music library.

Step 7. Return to Step 1 and loop through the steps of Step 1 and Step 6 until all the songs in the database are categorized.

2.2. Audio Emotion Feature Parameterization. Musical features include pitch, duration and timbre, rhythm, melody and speed, as well as musical form, mode, and other semantic features. Music emotion analysis needs to consider music basic information, rhythm change, and structure form, etc. Music characteristics are divided into musical note characteristics and high-level characteristics [19]. Note features include pitch, duration, and intensity, while high-level features include speed, power, rhythm, melody, and interval information [20]. Pitch, duration, and timbre are the acoustic cues of musical emotion cognition. Therefore, they are regarded as the most basic components of musical emotion characteristics and are characterized by the mean square deviation [21]:

$$F_i = \sqrt{\frac{1}{x} \sum_{p=1}^{x} L_i}. \tag{3}$$

In the formula, L_i denotes the pitch of the i and x denotes the number of notes. Intensity is an important means of expressing emotion in music [22]. Different musical intensity can create different emotional experience for listeners. Therefore, the intensity can be used to express the pitch breadth of the music, and the average intensity and intensity change characteristics can be used to represent the music emotion. The intensity characteristics can be extracted as follows:

$$S_i = \left(f_i \times \frac{v}{G}\right) \times t. \tag{4}$$

In the formula, f_i represents the strength of the i note; G represents the total number of notes; v denotes pitch; and t stands for time. Through formula (4) to describe the

dynamics change in the form of music sections, we can effectively exclude the influence of light and light changes in the rhythm on the parameterization of emotional characteristics.

Melody is an organization and harmonious movement formed according to certain music rules, which can effectively reflect the organization form of music in time and space. The direction of the melody describes the variation in pitch. The direction of the melody can be expressed as follows:

$$R_i = \sum_{i=1}^{m} \frac{S_i}{H - h_i}. \tag{5}$$

In the formula, h_i stands for the i length and H stands for the sum of the tones. Rhythm is a regular intensity phenomenon that occurs alternately in music [23]. Different rhythms give different strains to the music. According to the tension change of rhythmic movement, the musical characteristics are extracted by using the density of pronunciation point [24] to reflect the basic state of music. The greater the density of the point of articulation, the stronger the intensity, and the more obvious the characteristics. The density of the rhythm is

$$Q_i = \frac{W_i - 1}{W_{\mathrm{Max}i}}. \tag{6}$$

In the formula, W_i means the energy value of the subbar of i and $W_{\mathrm{Max}i}$ means the maximum energy value of the bar.

2.3. Establishment of an Audio Emotion Feature Extraction Model. Emotion is the essential characteristic of music. The mathematical model of music emotion must be based on the research of its psychological model. From the perspective of information theory, the whole process of music psychological feeling is actually a process of information acquisition, transformation, transmission, processing, and storage [25], while from the perspective of cybernetics, every composer, performer, or listener can be regarded as a biological steady-state system, which has emotional sensation and self-control behavior. Forward neural network is a kind of neural network model imitating human brain, so it can be constructed by imitating human brain to recognize musical emotion.

Through the feature extraction model of audio emotion, any 10-second music fragment j can be identified emotionally. Formula (7) is used to calculate the sum of the squares of the difference between the eigenvalues of j and the average value of the music features of each class, and then the qualification vector value of j in each class is calculated by formula (8), and then the j is classified into the music features with the largest qualification vector value.

$$d_{j,l} = \sum_{b=1}^{25} \left(j_b - t\left(l, b\right)\right). \tag{7}$$

$d_{j,l}$ denotes the proximity of j to l, and j_b denotes the b eigenvalue of j.

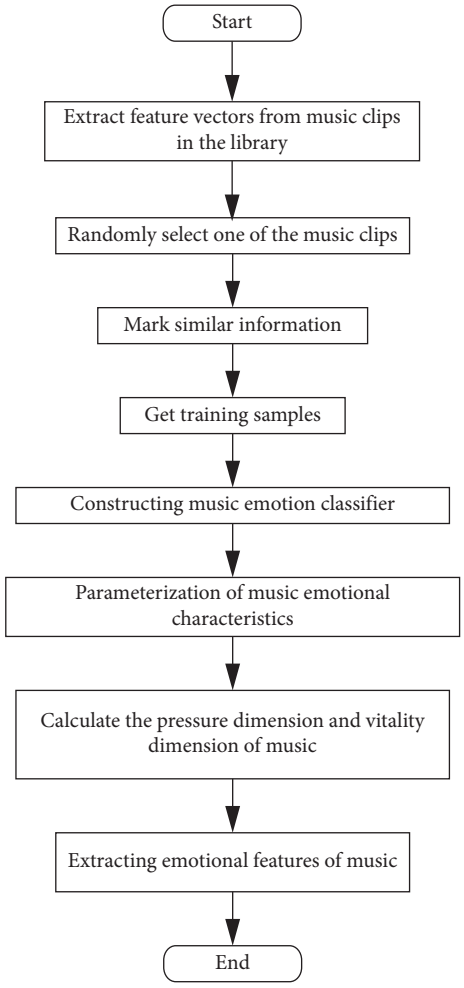


Figure 2: Flowchart of music emotion feature extraction.

$$g_{j,l} = \frac{d_{j,l} \times K}{\sum_{b=1}^{25} d_{j,b}}. \tag{8}$$

$g_{j,l}$ denotes the eligibility vector of j in l-like, and K denotes the ambiguity.

2.4. Emotional Feature Extraction of Whole Music. In order to sympathize with the listener and let the listener feel certain emotion, the composer must exaggerate certain emotion through different emotion to make certain emotion express more intensely, so there are different emotion changes in the same piece of music.

Firstly, the whole piece of music is divided into several segments, each segment is 10 seconds, and there is a 3-second repetition between two consecutive segments; then, the FNM method is used to calculate the four qualification vectors $g_{j,l}$ for each segment, formula (9) is used to calculate the value of the stress dimension, and formula (10) is used to calculate the value of the vitality dimension, so as to complete the emotional feature extraction of the whole piece of music.

$$j_w = g_{j1} + g_{j2} + g_{j3} + g_{j4}. \tag{9}$$

$$j_i == g_{j4} - g_{j2} - g_{j3} + g_{j1}. \tag{10}$$

Flowchart of music emotion feature extraction is shown in Figure 2.

3. Music Emotion Recognition Classification Algorithm Design

3.1. Forward Neural Network Architecture. Forward neural networks are multilayer neural networks that can be used to solve nonlinear problems [26, 27]. FNN is usually composed of input layer, output layer, and hidden layer. The structure of input layer is received by the hidden layer, and the result is transferred to the output layer as output of the whole network. When the actual output is inconsistent with the expected output, the calculated error is transmitted to the input layer as input feedback, and the weights of each layer in the network model are modified by the error gradient descent method. Aiming at the feature vector data of audio and lyrics, this paper improves the traditional feedforward neural network. Based on the special Chebyshev orthogonal polynomial clusters, a forward neural network model with special structure is constructed. In this model, a single hidden layer is used to reduce the complexity of the whole model. The stimulation function of each neuron in the hidden layer uses each function in Chebyshev's orthogonal polynomial cluster, and the stimulation function of neurons in other layers of the model uses linear stimulation function.

Assuming that the value of each neuron in the network is a_i, the layer and interlayer elements are connected with different weights, and the activation value b_i obtained by nonlinear activation function [28] is used to calculate the data of the upper layer, and the obtained results are transferred to the data of the lower layer to realize the forward calculation. Forward mental network model frame diagram is shown in Figure 3.

In the expression, I represents the input of the neural network, P represents the output of the neural network, and b_1 and b_2 are the weights of the input and output layers.

3.2. Neural Network Propagation Formula. The propagation of the neural network includes forward propagation and back-propagation [29]; forward propagation refers to the propagation of the model from the bottom up, calculated based on a given input value; back-propagation calculates losses based on forward propagation calculations and then calculates and trains each neuron parameter using a gradient descent algorithm [30].

The forward propagation formula of a neural network can be represented by a set of recurrent formulas, namely,

$$a_r^t = \sum_{i=1}^{25} q_i j_r^t + b_r^t. \tag{11}$$

In the formula, the input of the r neuron hiding layer at the time of t is marked as a_r^t, the output of the r neuron hiding layer at the time of t is marked as b_r^t FF, the input of the r neuron is marked as j_r^t, and q_i represents the parameter weights of the input layer and the hidden layer, thus

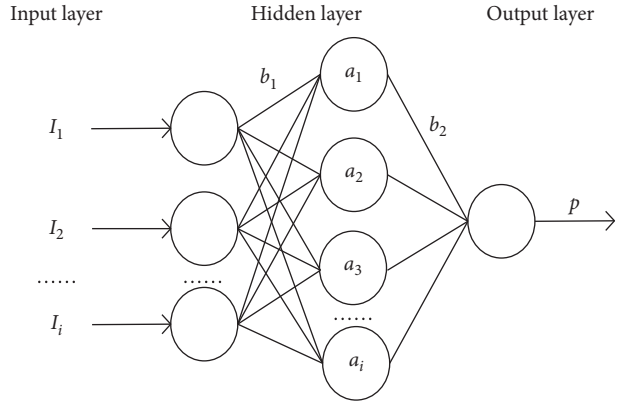


Figure 3: Structure diagram of forward neural network.

obtaining the specific forward propagation results of the neural network.

3.3. Softmax Activation Function. Neural networks are trained using Softmax as an activation function [31]. Softmax has the advantage of being computationally convenient and having a simple form of results for multiple classification problems. Because the output of each neuron of Softmax is positive and the sum is 1, the output of Softmax layer can be regarded as a probability distribution, and a more intuitive result can be obtained.

Assuming that the output of Softmax is $f(c)$, the input is marked c, and the k element in the array c is represented by an c_k, the formula for determining Softmax is

$$f(c) = \frac{e^c}{\sum_{k=1}^{25} e^{c_k}}. \tag{12}$$

Softmax's loss function is

$$\text{loss} = -\lg f(P_k). \tag{13}$$

In the expression, P_k is the correct input to the sample, but since there is no single data input at the time of training; assuming the input is $f(c) = e^c / \sum_{k=1}^{25} e^{c_k}$, there is $\text{loss} = \log \sum_{i=1}^{25} e^c - c_k$. In the case of applying the chain rule [32], the main steps of solving the partial derivative of loss from the weight matrix are as follows:

(1) Calculating the activation value of each node in the network.

(2) The gradient is propagated through the direction propagation algorithm to obtain the gradient values of each parameter.

(3) Adopt gradient descent algorithm to update model parameters.

(4) Iterate the process until it converges.

3.4. Overall Affective Recognition Framework. Music emotion recognition based on forward neural networks includes the following flow, as shown in Figure 4.

(1) Characteristics of learning samples: through controlling the forward neural model, the thresholds of

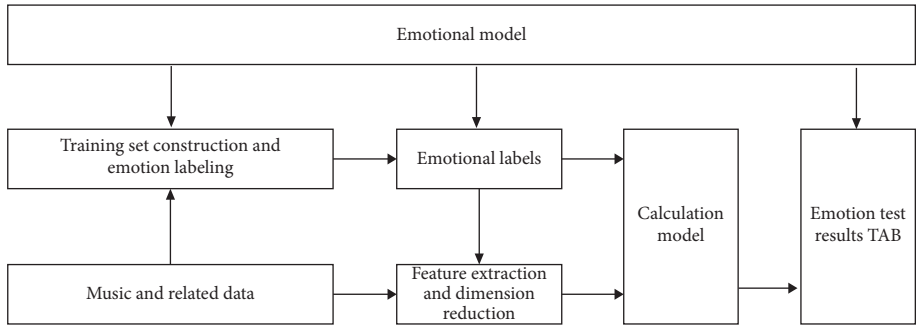


Figure 4: Music emotion recognition system framework.

neurons and the connection weights between neurons in the forward neural network model are iterated under the standard of sample data. Under the constraint of objective function minimization, the FNN can learn the features of sample data by a large margin.

(2) Music emotion is classified in combination with relevant music data and information, and a type data set is constructed to label emotion. The emotion classification process proposed in this paper is divided into two parts: model generation and emotion classification. The music emotion classifier is obtained by training the classifier according to the features of training samples, and the emotion classifier classifies the input music segments according to the emotion classifier.

(2) The preprocessing of audio information mainly includes signal filtering, etc. The preprocessing of lyrics text is the operation of the vocabulary of word emotion recognition and the addition of similar words. The feature extraction and dimensionality reduction are carried out for the data set, and the neural network model is established.

(3) Input training sample set, training output value, calculation error, and reverse update model parameters until the error is less than the expected error.

(4) Enter the test samples, conduct the test, and obtain the label of the test results.

(5) Judge whether the requirements for precision and iteration times are met. If they are satisfied, the model can be directly used to test the discrimination of musical emotion types; if not, it can increase the number of iterations and adjust the parameters of the network model.

4. Experimental Section

This experiment is a simulation experiment, the experimental software platform is MATLAB 2020b; the proposed method will be input into the simulation software to achieve operation, set the parameters of the experimental environment, and get the test results of the proposed method. The experimental indices include training by sampling, training time analysis of FNN model, accuracy analysis of emotion classification by FNN model, transmission rate analysis driven by FNN model, anti-attack analysis of FNN model, data integrity analysis of FNN model, and accuracy analysis of different music emotion recognition.

4.1. Experimental Environment. Collect 300 multitrack MIDI files on YouTube through Web Crawler [33]. Perform with a variety of instruments. Play music in a variety of styles with accurate emotional descriptions. Randomly selected music files were divided into four categories according to the emotional labels at the time of download: happy, anxious, calm, and depressed. Data were divided into five categories according to the specific needs of the experiment: provocative, happy and painful, happy, humorous, and fanatical, as shown in Tables 1 and 2.

4.2. Simulation Results

4.2.1. Sample Training. Then, 1,000 volunteers of different ages, genders, and professions were randomly selected to subjectively categorize the music in the sample library. If more than 50% of people categorized the same piece of music into a certain category, the music was used as a sample. If 40% of people chose one type of emotion and 30% chose another, the piece was considered an invalid sample. The average of all the features in each category is calculated as follows:

$$R(o,e) = \frac{1}{N_o}\sum_{e=1}^{25} H_{o,e,n}. \tag{14}$$

Among them, $R(o,e)$ represents the average value of e eigenvalue in $o = 1,2,3,4,5$ in o-like, N_o represents the number of fragments in o-like, and $H_{o,e,n}$ represents the e eigenvalue in o-like n fragment.

4.2.2. Training Time Analysis of Forward Neural Network Model. The training time of FNN model is not only related to the supervised algorithm, but also related to the number of layers of the model and the number of neurons in each layer. In this paper, 8 hidden layers and 60 neurons are used as experimental objects to compare and analyze the training

Table 1: Data emotions divided into four categories.

	Category 1(exuberance)	Category 2(anxious)	Category 3(contentment)	Category 4(depression)
Quantity	76	69	83	72
Total		300		

Table 2: Data emotions divided into five categories.

	Category 1(aggressive)	Category 2(bittersweet)	Category 3(happy)	Category 4(humorous)	Category 5(passionate)
Quantity	67	62	54	58	59
Total			300		

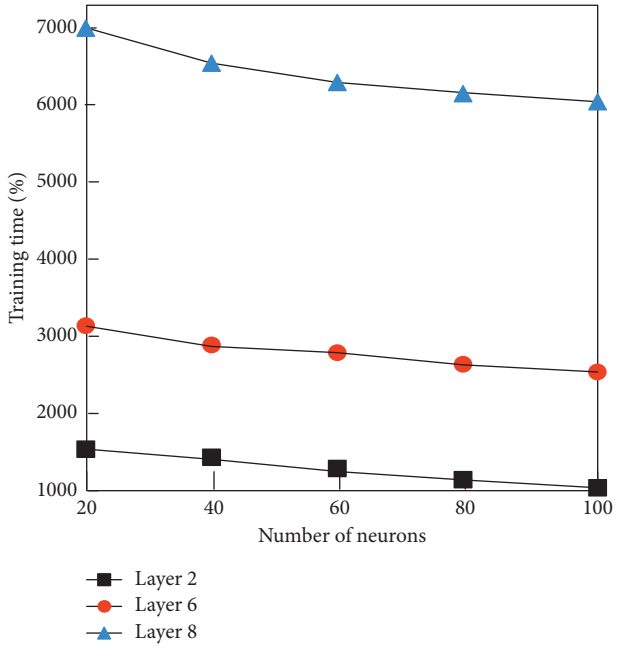


Figure 5: Training time of forward neural network model.

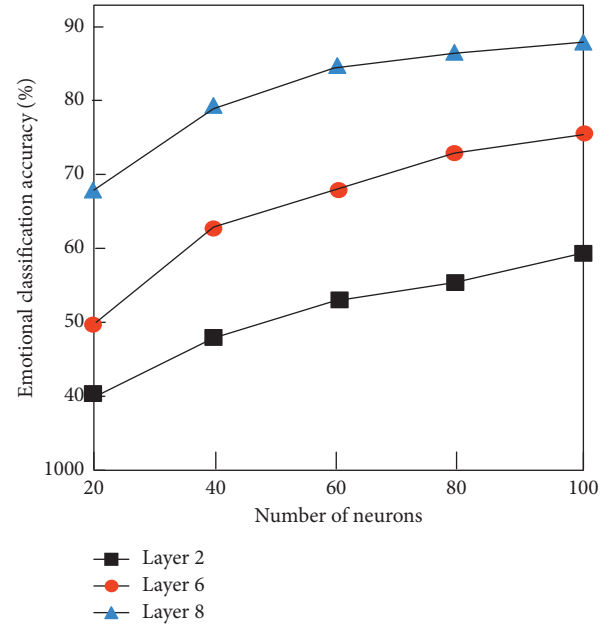


Figure 6: Emotional classification accuracy of forward neural network model.

time of FNN model. The simulation results are shown in Figure 5.

Figure 5 shows that the training time is inversely proportional to the number of layers and the number of neurons. The average training time is 6400s when the number of layers increases to 8. But with the increase of the number of neurons, the training time will be reduced gradually, the number of neural network layers is 8, the number of neurons is 100, and the training time is 6000 s. Therefore, in order to achieve a better training effect, we can reduce the training time and improve the efficiency by increasing the number of neurons in the forward neural network.

4.2.3. Accuracy Analysis of Emotion Classification Using Forward Neural Network Model. In the calculation of FNN model, the complexity of emotion classification is different due to the difference of layers and neurons, and the accuracy is different. The simulation result is shown in Figure 6.

Figure 6 shows that the increase in the number of layers and neurons will improve the network model's learning ability, that is, improve the accuracy of classification. When the number of neural networks is 8 and the number of neurons is 100, the accuracy of emotion classification can reach 89%. This is because the neural parameters are trained and calculated by gradient descent algorithm, and the probability distribution is obtained by activation function, which improves the accuracy of emotion classification effectively.

4.2.4. Forward Neural Network Model Driven Transmission Rate Analysis. Data transmission rate is the number of data transmitted by data path in unit time. By analyzing data transmission rate, special bandwidth limitation can be prevented during high load, and the model can run safely and effectively. Therefore, the forward neural network model driven transmission rate is measured using 32 inputs and outputs as the standard. The result is shown in Figure 7.

Figure 7 shows that the data flow of each input and output of the application object is higher than 55MBS in the process of driving debugging, which shows that the

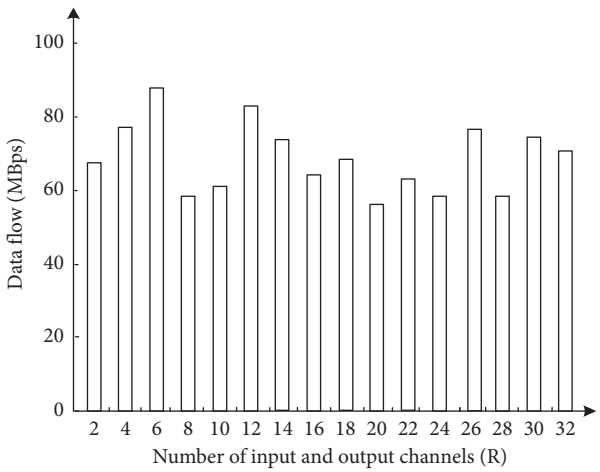


Figure 7: Forward neural network model driven transmission rate results.

proposed method has good performance in practical application and obvious advantages in application.

4.2.5. Attack Resistance Analysis of Forward Neural Network Model. In order to test the resistance of the forward neural network model, the attacker is set to perform brute force cracking and wormhole attack on the forward neural network model; the result is shown in Figure 8.

Figure 8 shows that the FNN model has the highest anti-attack ability, 91%, at the time of 20s, and the anti-attack ability is in the high anti-attack ability in different music emotion classification period. The results show that the FNN model has better security performance in the data transmission.

4.2.6. Data Integrity Analysis of Forward Neural Network Model. The larger the value of F is, the more complete the music segment feature data is. The formula is as follows:

$$F = \frac{d}{L} \times 100\%. \tag{15}$$

In the formula, d is the feature data extracted from the music segment in formula (1) and L is the overall feature data of the music. Using data integrity as a test metric, the results are shown in Figure 9.

Figure 9 shows that the F value is above 83 when using this method to extract the music fragment feature data. According to the test results, Gaussian function is used to normalize the feature vectors of music fragments, so the integrity of feature extraction data is improved.

4.2.7. Analysis of the Accuracy of Different Music Emotion Recognition. In order to verify the superiority of the proposed algorithm, the following simulation experiments focus on the accuracy of 300 music emotion segments; the specific simulation results are shown in Figure 10.

Figure 10 shows that the accuracy of music emotion recognition is influenced by the number of categories and

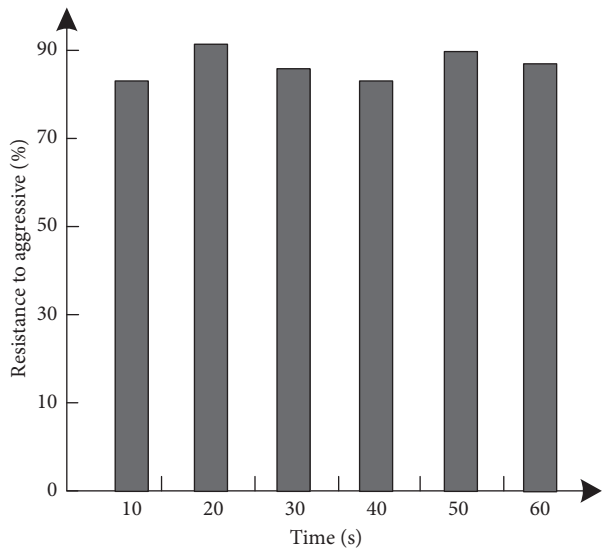


Figure 8: Results of forward-neural network model attack resistance.

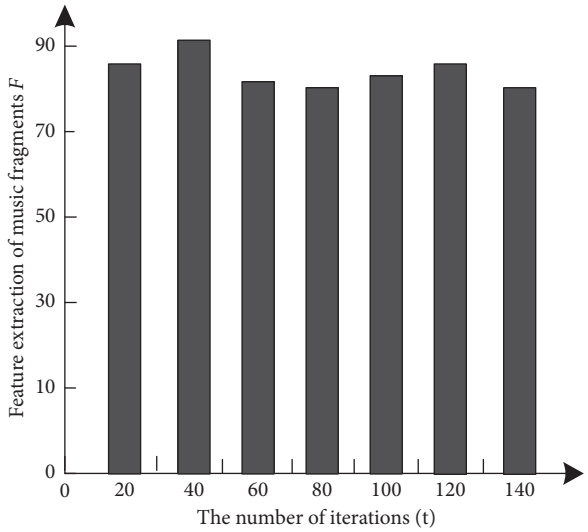


Figure 9: Data integrity test results of forward neural network model.

the complexity of music emotion. The accuracy of the tests is above 85%.

The above experimental results show that the training time of the feedforward neural network applied in this study is inversely proportional to the number of layers and neurons. When the number of layers increased to 8, the average training time was 6400 s. But as the number of neurons increases, the training time will gradually decrease. The number of neural network layers is 8, the number of neurons is 100, and the training time is 6000 seconds. Therefore, in order to achieve better training effect, we can reduce training time and improve training efficiency by increasing the number of neurons in the forward neural network. When the number of neural networks is 8 and the number of neurons is 100, the accuracy of emotion classification can reach 89%. The data stream of each input and

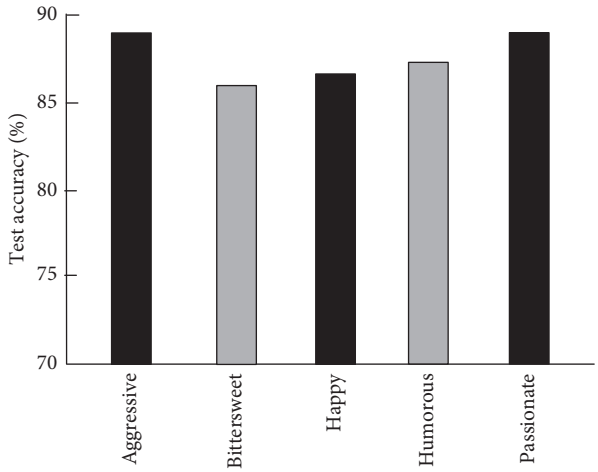


Figure 10: Musical emotion fragment test accuracy.

output of the application object is higher than 55MBS, which shows that the method has good performance and obvious application advantages in practical application. The FNN model has the highest anti-attack ability, up to 91%, and has the higher anti-attack ability in the different music emotion classification stage. The accuracy of music emotion recognition is affected by the number of music emotion categories and the complexity of music emotion, but the accuracy of this method is above 85%. The experimental data show that the proposed method has ideal application effect and provides a reliable basis for further research in related fields.

5. Conclusion

Music emotion cognition is an important part of multimedia content understanding and intelligent human-computer interaction, which has been widely used in human-computer interaction, entertainment robot, game industry, and music education. With the development of computer technology, the application of music emotion in stage performance art has become a new trend of the development of multimedia technology. Emotion recognition plays an unshakable role in music retrieval and data service, so it is of great practical significance to study emotion recognition. In this paper, we construct a forward neural network model to extract the musical emotion features and identify and classify them according to the characteristics of intensity, melody, and rhythm density. Through the experimental data, we can conclude that the forward neural network model has better use value, has higher accuracy for music emotion recognition, and has important significance for music emotion recognition classification.

Data Availability

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding this work.

Acknowledgments

This work was supported by Jilin Province Department of Education 13th Five-Year Plan, "A research on the Application of Music Therapy in Rehabilitation Education of Special Children" (JJKH20180407SK) and Jilin Province Department of Education 13th Five-Year Plan, "The Exploration and Research on Vocal Music Teaching Mode in Normal Universities in the New Era" (JJKH20170090SK).

References

[1] D. Wang and X. Guo, "Research on intelligent recognition and classification algorithm of music emotion in complex system of music performance," *Complexity*, vol. 2021, no. 6, Article ID 4251827, 10 pages, 2021.

[2] D. Chaudhary, N. P. Singh, and S. Singh, "Development of music emotion classification system using convolution neural network," *International Journal of Speech Technology*, vol. 24, no. 2, pp. 1–10, 2020.

[3] D. Zheng, "Music emotion recognition classification algorithm based on forward neural network," *Journal of Information Technology*, vol. 43, no. 12, pp. 57–61, 2019.

[4] T. Li, "Visual classification of music style transfer based on PSO-BP rating prediction model," *Complexity*, vol. 2021, no. 6, Article ID 9959082, 9 pages, 2021.

[5] S. Werner and G. N. Petrenko, "Speech emotion recognition: humans vs machines," *Discourse*, vol. 5, no. 5, pp. 136–152, 2019.

[6] A. Fh, B. Ma, and E. Smcd, "Continuous emotion recognition during music listening using EEG signals: a fuzzy parallel cascades model," *Applied Soft Computing*, vol. 101, Article ID 107028, 2020.

[7] S. Panwar, P. Rad, K.-K. R. Choo, and M. Roopaei, "Are you emotional or depressed? Learning about your emotional state from your music using machine learning," *The Journal of Supercomputing*, vol. 75, no. 6, pp. 2986–3009, 2019.

[8] T. Saha, D. Gupta, S. Saha, and P. Bhattacharyya, "Emotion aided dialogue act classification for task-independent conversations in a multi-modal framework," *Cognitive Computation*, vol. 13, no. 3, pp. 1–13, 2021.

[9] X. Liu, "An improved particle swarm optimization-powered adaptive classification and migration visualization for music style," *Complexity*, vol. 24, no. 7, p. 0872, 2020.

[10] M. T. Quasim, E. H. Alkhammash, M. A. Khan, and M. Hadjouni, "Emotion-based music recommendation and classification using machine learning with IoT Framework," *Soft Computing*, vol. 25, no. 7, Article ID 12249, 2021.

[11] S. Demircan and H. K. Örnek, "Comparison of the effects of mel coefficients and spectrogram images via deep learning in emotion classification," *Traitement du Signal*, vol. 37, no. 1, pp. 51–57, 2020.

[12] C. Chen and Q. Li, "A multimodal music emotion classification method based on multifeature combined network classifier," *Mathematical Problems in Engineering*, vol. 2020, no. 8, 11 pages, Article ID 4606027, 2020.

[13] S. Rajesh and N. J. Nalini, "Musical instrument emotion recognition using deep recurrent neural network," *Procedia Computer Science*, vol. 167, pp. 16–25, 2020.

[14] A. V. Haridas, R. Marimuthu, V. G. Sivakumar, and B. Chakraborty, "Emotion recognition of speech signal using Taylor series and deep belief network based classification," *Evolutionary Intelligence*, no. 4, pp. 11–14, 2020.

[15] S. Hizlisoy, S. Yildirim, and Z. Tufekci, "Music emotion recognition using convolutional long short term memory deep neural networks," *Engineering Science and Technology an International Journal*, vol. 24, no. 3, pp. 760–767, 2020.
[16] D. Lee, L. Robinson, and D. Bawden, "Modeling the relationship between scientific and bibliographic classification for music," *Journal of the Association for Information Science and Technology*, vol. 70, no. 3, pp. 230–241, 2019.
[17] S. Mohammed, C. Ramos, W. C. Fang, and T. H. Kim, "Emerging higher-level artificial neural network-based intelligent systems," *Neural Computing and Applications*, vol. 33, no. 3, pp. 4595–4597, 2021.
[18] G. Wang, C. Lopez-Molina, and B. Debaets, "Multiscale edge detection using first-order derivative of anisotropic Gaussian kernels," *Journal of Mathematical Imaging and Vision*, vol. 61, no. 5, pp. 1096–1111, 2019.
[19] M. Kaliakatsos-Papakostas and E. Cambouropoulos, "Conceptual blending of high-level features and data-driven salience computation in melodic generation," *Cognitive Systems Research*, vol. 58, pp. 55–70, 2019.
[20] A. Viloria, O. B. Pineda Lezama, and D. Cabrera, "Segmentation process and spectral characteristics in the determination of musical genres," *Procedia Computer Science*, vol. 175, pp. 96–101, 2020.
[21] D. d. F. P. Melo, I. D. S. Fadigas, and H. B. d. B. Pereira, "Graph-based feature extraction: a new proposal to study the classification of music signals outside the time-frequency domain," *PLoS One*, vol. 15, no. 11, Article ID e0240915, 2020.
[22] M. Jabreel and A. Moreno, "A deep learning-based approach for multi-label emotion classification in tweets," *Applied Sciences*, vol. 9, no. 6, p. 1123, 2019.
[23] S. Ma, "Music rhythm detection algorithm based on multipath search and cluster Analysis," *Complexity*, vol. 2021, no. 4, pp. 1–9, Article ID 5627626, 2021.
[24] K. Ohara, K. Saito, M. Kimura, and H. Motoda, "Critical node identification based on articulation point detection for uncertain network," *International Journal of Networks and Communications*, vol. 9, no. 2, pp. 201–216, 2019.
[25] M. A. Khan and V. S. Patel, "Knowledge and approximations: a formal study under the perspective of information systems and rough set theory," *Information Sciences*, vol. 524, no. 5, pp. 97–115, 2020.
[26] A. Benavoli, A. Facchini, D. Piga, and M. Zaffalon, "Sum-of-squares for bounded rationality," *International Journal of Approximate Reasoning*, vol. 105, pp. 130–152, 2019.
[27] A. González-Espinoza, G. Martínez-Mekler, and L. Lacasa, "Arrow of time across five centuries of classical music," *Physical Review Research*, vol. 2, no. 3, Article ID 33166, 2020.
[28] S. Bharadwaj, S. Hegde, N. Dutt, and A. P. Rajan, "Nonlinear signal processing method detects emotional changes induced by Indian classical music," *International Journal of Engineering and Advanced Technology*, vol. 9, no. 1, pp. 6200–6206, 2019.
[29] Y. Zheng, "The use of deep learning algorithm and digital media art in all-media intelligent electronic music system," *PLoS One*, vol. 15, Article ID 0240492, 2020.
[30] V. Bondarev, "Training a digital model of a deep spiking neural network using backpropagation," *E3S Web of Conferences*, vol. 224, no. 38, Article ID 01026, 2020.
[31] S. Totaro, A. Hussain, and S. Scardapane, "A non-parametric softmax for improving neural attention in time-series forecasting," *Neurocomputing*, vol. 381, pp. 177–185, 2020.
[32] R. Cushman and J. Niatycki, "On subcartesian spaces Leibniz' rule implies the chain rule[J]," *Canadian mathematical bulletin = Bulletin canadien de mathématiques*, vol. 63, no. 2, pp. 1–9, 2019.
[33] D.-H. Han and Y.-K. Lee, "Design of action-based Web crawler structural configuration for multi-website management," *KIISE Transactions on Computing Practices*, vol. 27, no. 2, pp. 98–103, 2021.

Hindawi
Scientific Programming
Volume 2022, Article ID 1054294, 8 pages
https://doi.org/10.1155/2022/1054294

Research Article

Learning of Martial Arts Action Decomposition Method Based on Image Recognition

Yuanbing Zhou,[1] Yufeng Du,[1] Xiaochun Lu,[1] and Haiou Chen [2]

[1]*Xi'an Academy of Fine Arts Sports Department, Xi'an, Shaanxi 710065, China*
[2]*Department of Physical Education, Guangdong Medical University, Zhanjiang, Guangdong 524023, China*

Correspondence should be addressed to Haiou Chen; jinling@gdmu.edu.cn

Received 1 December 2021; Revised 23 December 2021; Accepted 27 December 2021; Published 19 January 2022

Academic Editor: Baiyuan Ding

In order to improve the effective guidance of martial arts movements, a decomposition method of martial arts movements based on image recognition is proposed. First of all, visual image target sampling of martial arts action is performed, and most of the noise background is eliminated through morphological gradient operation. Then, the contour edge of the human body is obtained, the contour edge of each frame of the video is extracted, and the accumulation is realized in the same image. Using accumulation, the edge image calculates the grid-based HOG to obtain the image action feature vector. Secondly, using the improved dynamic time warping theory combined with the characteristics of the angle change of each joint under the action time sequence, the joint change sequence among various martial arts movements can be identified in order to realize the decomposition process of martial arts actions based on image recognition. The experimental results show that the use of image recognition can effectively decompose martial arts movements.

1. Introduction

At present, with the vigorous development of sports, the competition rules of martial arts competitions are constantly improving in the process of internationalization. The martial arts routine competitions are developing in the direction of rapid movement combinations and difficult movements [1, 2]. The real-time extraction and optimization method of human martial arts action images can extract the contour features of human martial arts action images, and based on this, the decomposition learning of martial arts actions is the fundamental way to solve the above problems, which has attracted the attention of many experts and scholars [3–5]. Due to the far-reaching developmental significance of the decomposition method of human martial arts movements, it has also become the focus of research by industry insiders and has received extensive attention. At the same time, many good results have appeared [6–9]. It is the ultimate goal of computer vision research to enable computers to have the same visual perception functions as humans and to recognize external things, perceive scenes, and analyze the activities of surrounding things like humans. Using computers to analyze and understand human movements is challenging. This subject involves multiple disciplines, including cognitive science, pattern recognition, and machine learning, and has certain academic research value. The application of this technology will enable the computer to have the ability to observe the external world and then make a decision response through the automatic analysis and understanding of image information, and the computer will have a better ability to adapt to the environment [10–14]. Three-dimensional images of martial arts movements have developed into a very active research topic in the field of image research. Each three-dimensional image of martial arts movements has its own fixed characteristics. They are located at a certain point in the three-dimensional image of martial arts movements, but there are many uncertainties. The interference of factors makes the three-dimensional image of martial arts movement have insignificant areas, which increases the difficulty of studying the three-

dimensional image of martial arts movements, and the adaptive enhancement of the insignificant areas is to solve the insignificant areas of the three-dimensional image of martial arts movements. The most effective method to study the trouble has become the focus of many scholars' research.

Literature [15] pointed out that the analysis and understanding of human actions in videos can be summarized as follows: by extracting movement and performance characteristics from the video, reasonable judgments are made on the type of action and the direction in which the action occurs, and the semantic information corresponding to the action is analyzed in detail. Finally, the person's behavioral intentions are determined. Action representation, action segmentation, action recognition, and action positioning are the most important studies in the process of human action analysis and understanding. Action representation refers to obtaining the feature vector describing the video by extracting information such as the motion and structure of the input video. In [16], the definition of action segmentation is to divide the continuous video stream into several groups of subvideo segments containing only one action instance. Human action recognition is realized by establishing the association between the video content and the action category. Literature [17] proposed a real-time extraction method of human martial arts difficult action images based on extended edge features. This method first establishes the network structure of the difficult action image of human martial arts based on the principle of lateral suppression competition, prompts the human action pixel in the grid to be connected with its surrounding pixels, extracts the edge features of the difficult action image of human martial arts, and calculates the suppression competition coefficient of the pixel. Based on this, the real-time extraction of images of difficult martial arts movements of the human body is completed. This method is relatively simple, but it has the problem of large limitations. Literature [18] focuses on a real-time extraction method of human martial arts difficult action images based on the improved Canny method. This method firstly integrates the color features of the human martial arts difficult action image to initially locate the outline of the human martial arts difficult action image and performs edge detection on the human martial arts difficult action image. The Canny method is used to calculate the gradient amplitude. Difficult martial arts action images are extracted in real time. This method has high extraction efficiency, but the current method is used for the real-time extraction of actions, which cannot extract accurate contour information of difficult martial arts action images, and there is a problem of large errors in real-time action extraction. Literature [19] proposed a real-time extraction method of human martial arts difficult action images based on HSV color space and mathematical morphology. This method first uses color information to filter the areas that may contain the difficult action images of martial arts, then uses mathematical morphology technology to generate connected areas, judges and generates the correct difficult action image areas of martial arts, and finally uses radon transform to perform tilt correction in order to complete the real-time extraction of images of difficult martial arts movements on the human body. This method has high extraction accuracy, but the current method performs real-time extraction of actions, which cannot extract accurate image contour information of martial arts actions, and there is a problem of large errors in the real-time extraction of actions.

Therefore, this technique is used in this article to propose a method of decomposing and recognizing human martial arts movements based on image recognition [20]. Firstly, the morphological gradient operation is used to remove most of the noise background, the edge boundary of the human body contour is obtained, the shape edge of each frame of the image is obtained in the video, it is accumulated in an image at the same time, and the grid-based image is calculated by accumulating the edge image. HOG, the image action feature vector is obtained; secondly, combined with the change characteristics of each joint angle under the action time sequence, the improved dynamic time warping theory is used to identify the similarity of the joint change sequence among various martial arts actions, and the classifier is designed, and the time-varying feature data of human actions in the image is input into it, and the decomposition process of martial arts action based on image recognition is completed, and finally, the simulation test analysis is carried out, and the validity conclusion is reached [21–25].

2. Sampling of Martial Arts Action Visual Image Targets

2.1. Principles of Martial Arts Action Image Target Collection. In the process of real-time acquisition of human martial arts action image targets, first determine the spatial position of the martial arts action image, track the spatial position of each body part of the human martial arts action image, build a calculation model for the position of the human body part, and extract the human martial arts action image contour feature, matching different features, selecting a set of matching sequences of the most neighboring action images, as the real-time extraction of the difficult action images of human martial arts. The specific steps are as follows.

Suppose p represents that the camera will reflect the mark, $[x_w, y_w, z_w, 1]^T$ represents the world coordinates of the reflective mark p, and formula (1) is used to convert the world coordinates of the reflective mark p to the three-dimensional coordinates of the camera:

$$P_c = \left[R_C^W, T_C^W\right]\left[x_w, y_w, z_w, 1\right]^T \times P. \tag{1}$$

In the above formula, $[R_C^W, T_C^W]$ represents the external parameters of the camera during the calibration process.

Assuming that (u_0, v_0) represents the reference point of the camera image, equation (2) is used to track the spatial position of each body part in the human martial arts action image:

$$\eta(\zeta) = \frac{(u_0, v_0) \times f(P)}{N(f)} \times L_{(n_i)} \times \lambda(\xi). \tag{2}$$

In the above formula, $f(P)$ represents its gray value information, $N(f)$ represents the state vector of the martial arts

action image, $L(n_i)$ represents the target area of the human martial arts action image, and $\lambda(\xi)$ represents the edge of the image.

Assuming that $\sigma[j]$ represents the position of the space where the human body martial arts difficult action image part is located, $\sigma[j]$ represents its direction, and Λt represents the body's movement speed; then, formula (3) is used to construct the calculation model of the human body part position:

$$p(k+1) = \frac{p(k) + v(k)\Lambda t}{\sigma[j] \cdot \sigma[i]} \times \Lambda t. \tag{3}$$

In the above formula, $p(k)$ represents the acceleration of the body part and $v(k)$ represents the coefficient of the camera's radial distortion.

Suppose that $\Sigma(q)$ represents the position offset of the martial arts action image, and $\eta(\lambda)$ represents the gravity acceleration of the action. Equation (4) is used to extract the contour features of the martial arts action image, the different features are matched, and the equation is given as follows:

$$R(a,b) = \frac{f_{ij} \cdot g(a,b)}{f_{ij} \times g_2(a,b)} \times \frac{\Sigma(q) \times \eta(\lambda)}{\omega(\mu) \times E(j)}. \tag{4}$$

In the above formula, f_{ij} represents the correlation coefficient of the contour edge matching of the human martial arts action image, $g(a,b)$ represents the reference image of the human martial arts action image, $\omega(\mu)$ represents the real-time image, and $E(j)$ represents the reference image and real-time image of the action image. Assuming that $d(p)$ represents the similarity measurement threshold between the two images, equation (5) is used to match different features, and a set of action image matching sequences that are closest to each other is selected:

$$\overline{\omega}(W) = \frac{d(p) \times \eta(w)}{R(a,b) \times V(E)} \times p(k+1) \times \Lambda t. \tag{5}$$

In the above formula, $V(E)$ represents the length of the matching sequence of human martial arts action images and $\eta(w)$ represents the optimal matching sequence of images.

The above formula can explain the principle of real-time extraction of martial arts action images, and this principle is used to complete the real-time extraction of martial arts difficult action image targets.

3. Image-Based Martial Arts Action Decomposition and Recognition Process

3.1. Action Feature Extraction Based on Cumulative Edge Image. The combined morphological operation can eliminate part of the background on the video image, the morphological features can be preserved intact, and the silhouette image of the human body can be obtained, which is very similar to the background subtraction technology [26–29]. The combined morphology operation formula can be expressed as follows:

$$G(x,y) = F(x,y) \cdot B(x,y) - F(x,y). \tag{6}$$

In the above formula, $G(x,y)$ represents the image processed through combined morphological operations, $F(x,y)$ represents the frame of image in the original video, $B(x,y)$ represents the structural element, and $\cdot$ represents the closing operation. Through the closing operation of formula (6), the area darker than the background and smaller than the size of the resulting element in the original image can be removed. the appropriate structural element is selected, and the remaining background image is obtained through the closing operation, and it is subtracted from the original image. Then, target extraction is completed. One frame of image in the video image of human martial arts action cannot fully express an action. Generally, it is necessary to extract the features of multiple frames to fully demonstrate a human action. Due to the difference in the action rate, even the same action, the number of frames of each video image may be different. In order to deal with the changes of these two rates, taking into account the characteristics of rate changes, the paper accumulates the gray features of each frame of the edge image in the same time window into the same image and uses the established cumulative edge image to extract its features, which are used to represent human martial arts movements. The formula for accumulating the edge image at the point (x, y) at time t can be described as follows:

$$\begin{aligned} I(x,y) &= G(x,y)E(x,y), \\ H(x,y,t) &= \max(H(x,y,t-1), I(x,y)). \end{aligned} \tag{7}$$

The cumulative edge image is to multiply the binary image $E(x, y)$ and the morphological gradient image $G(x, y)$ at each pixel to obtain the edge image $I(x, y)$ with gray information, and all edge images are accumulated into one image, instead of accumulating every frame of a binary image into one image. 0 and 1 are the only two values of the pixel gray value of the binary image $E(x,y)$. If the pixel value of the binary image $E(x,y)$ corresponding to the edge image $I(x,y)$ is 1, the gray value range at this point has more information than the binary image. The edge image is accumulated for the target image, and the image information center already contains the edge information of more frames of images, so there is no need to extract the edge features, and the direction gradient histogram can be solved directly at each point of the accumulated edge image. Calculating the directional gradient histogram based on the grid is to solve the directional gradient of all points on the cumulative edge image. The cumulative edge image is divided into $i \times i$ spatial grids, the histogram vector is calculated on each grid, one of the scale feature vectors is extracted, and it is used as an action feature, the local shape obtained by the target is counted, and then, the cumulative edge is obtained.

3.2. Human Martial Arts Action Recognition Based on Dynamic Time Warping. Action expression has time continuity, that is, action can be a collection of a series of static actions in a certain period of time. The movement process of

the human body can reflect the change trend of the action through the change of the joint angle curve, and the angle change curve of the joint with the time change can be called the "joint angle time series." Human motion characteristics are described by the time series of joint angles. If the duration of a martial art action is set to T, the motion characteristics can be defined as follows:

$$\text{action feature} = \{A_1, A_2, \ldots, A_M\}^T, \tag{8}$$

where the time series of a certain joint angle is represented by the row vector A_i; the row vector when the number of motion features is M is represented by A_M; the number of motion features is represented by M, and the range is $1 \le M \le 16$. The row vector A_i can be understood as a time-varying one-dimensional signal and then evolved into a classification problem of simple action recognition classified as time-varying feature data. It can be seen from the prior data that when the tester freely demonstrates martial arts movements, the same movements have different waveforms and amplitudes, and the possibility of similarity to A_i cannot be ruled out. Therefore, action recognition is realized by comparing the similarity of time series; that is, the determination of martial arts action decomposition is realized by comparing the distance between vectors of different lengths.

The comparison of the similarity between the curves is the focus of the time-series change trend. Because there are uncertain factors in the video feedback system and the tester, etc., which will cause the deviation and fluctuation of the data, the following formula is used to smooth the sequence:

$$x_i = \frac{x_1 + x_2 + \cdots + x_n + x_{n+1}}{n}. \tag{9}$$

In the above formula, the joint angle value of the i-th time point in the sequence is represented by x_i; x_n+x_{n+1} are the joint angle values of n and $n+1$, respectively; and n is an integer greater than 0.

The dynamic time warping theory based on the idea of dynamic programming aims to find the optimal matching path and the shortest distance between two test samples of different lengths and the reference template. The reference time series is set to $R = \{r_1, r_2, \ldots, r_i, \ldots, r_{L1}\}$, and the test sample is set to $T = \{t_1, t_2, \ldots, t_j, \ldots, t_{L2}\}$. The joint angle values of time i and j are represented by r_i and t_j, respectively; L_1 and L_2 represent the vector length. If the vectors R and T are nonlinearly matched, the cumulative distance moment $D(i,j)$ can be described as follows:

$$D(i,j) = \min\left\{\begin{array}{c} D(i, j-1) \\ D(i-1, j) \\ D(i-1, j-1) \end{array}\right\} + d\left(r_i, t_j\right), \tag{10}$$

where $d(r_i,t_j)$ represents the distance function between r_i and t_j and $D(i,j-1)$, $D(i-1,j)$, and $D(i-1,j-1)$ are the cumulative distance matrix elements.

To make the points r_i and t_j on the time series have different joint angle Y-axis values, a three-dimensional vector needs to be constructed based on the points r_i and t_j to redefine $d(r_i,t_j)$ instead of the original Euclidean distance, that is, $r_i = [r_i, r_i', r_i'']$ and $t_j = [t_j, t_i', t_j'']$. The following formula describes the first derivative of the reference sequence r_i' and the second derivative of the reference sequence r_i'' in turn:

$$\begin{aligned} r_i' &= \frac{\left(r_i - r_{i-1}\right) + \left(\left(r_{i+1} - r_{i-1}/2\right)\right)}{2}, \\ r_i'' &= r_{i+1} + r_{i-1} - 2r_i. \end{aligned} \tag{11}$$

In the above formula, r_{i-1} represents the joint angle value at the $i-1$th time point and r_{i+1} represents the joint angle value at the $i+1$th time point. Since the construction of the above vector is conducive to the accuracy of the mapping, $d(r_i,t_j)$ can be defined as follows:

$$d\left(r_i, t_j\right) = w_1\left(r_i - t_j\right)^2 + w_2\left(r_i' - t_j'\right)^2 + w_3\left(r_i'' - t_j''\right)^2, \tag{12}$$

where t_j' represents the first derivative value of the joint angle of the test sample sequence, t_j'' represents the second derivative value of the joint angle of the test sample sequence, and w_1, w_2, w_3, respectively, represent the weight of the shortest distance adjustment of the joint angle value and the adjustment of the joint angle, the shortest distance weight of the first-order derivative value, and the shortest distance weight of the second-order derivative value of the adjusted joint angle.

According to formula (8), there are motion template feature matrix $\text{AR} = \{R_1, \ldots, R_k, \ldots, R_M\}^T$ and the sample to be tested $\text{AT} = \{T_1, \ldots, T_k, \ldots, T_M\}^T$, and if D_k is R_k and T_k, the distance between AR and AT can be described as follows:

$$D(\text{AR}, \text{AT}) = \{D_1, \ldots, D_k, \ldots, D_M\}, \tag{13}$$

where D_k represents the improved distance between samples, k represents the number of motion features with improved distance, and D_M is the improved distance between AR and AT when the number of motion features is M.

The expected distance value is calculated as follows:

$$\text{ED} = \sum_{k=1}^{M} w_k D_k, \tag{14}$$

where W_k is the weight value of the desired distance. Given a martial arts action image test sample, the martial arts action corresponding to the template with the minimum expected distance ED is the recognition result.

$$\text{Decision} = \arg_{i \in C} \min\{\text{ED}\}. \tag{15}$$

In the above formula, C represents a known template in the reference library.

In summary, the feature vector of the martial arts action features in the video image is extracted by accumulating edge images, and then, the martial arts action time sequence is calculated using the dynamic time warping theory. After the martial arts action to be recognized is matched with the reference time sequence sample, the process of decomposition and recognition of martial arts movements is completed.

4. Simulation Experiment Results and Analysis

In order to accurately decompose martial arts movements, this paper uses a dynamic time warping method based on accumulated edge images to identify martial arts movements, and the feasibility of this method is verified by simulation experiments [30].

Experiment 1. Two martial arts action images are given as experimental objects in the article. In order to effectively extract the target contour from the image, the morphological operation and active contour model method in the article are used to extract the target contour of the image. The specific image processing effect is shown in Figures 1~6.

From the first set of images (Figures 1 to 3), it can be seen that Figure 1 is the original image, and Figure 2 is the effect diagram after contour extraction of the image using the morphological operation in the text. By observing the morphological operation in Figure 2, the first after the image is transformed into a binary image, and the outline of the martial arts action is extracted. The edge image of the action can be clearly identified from Figure 2; after comparing Figure 2 with Figure 3, it can be seen that Figure 3 uses the active contour model. The method realizes the extraction of the outline of the martial arts action, but this method is not accurate in the extraction of the outline edge, and the outline of the action cannot be clearly recognized.

The images of the second set of experiments (Figures 4~6) can also prove that when using the morphological operation in this paper to process the image, it can effectively extract the contour edges of martial arts movements, indicating that the morphological operation in the text is an effective method to extract the contour edges of the image.

The experiment gives 5 groups of images and uses the morphological operation and the active contour model method in this paper to realize the contour extraction processing of the 5 groups of images and compares the time consumption and sharpness of the image contour extraction. The specific data are shown in Table 1.

By observing Table 1, we can see that using the morphological operation and the active wheel model method in this paper to achieve contour edge feature extraction on 5 groups of images with a given number of images, the method in this paper is used to process the 5 groups of images, and the average image contour extraction time is 1.2 s; while using the active contour model method to process the image, the average image contour extraction time is serious, which is more than 10 s. Comparing the image output definition, it can be seen that the image definition after the morphological operation in this paper is obviously much higher than that after the active contour model method, which shows that the performance of the morphological operation in this paper is superior.

Experiment 2. The experiment gives a set of different martial arts action sample sets, each set of sample sets including 4 actions. By using this article's dynamic time warping method and sports history image recognition method to perform action recognition on the martial arts action samples, the two methods are compared to recognize actions. The specific data are shown in Table 2.

Figure 1: Decomposed image of martial arts movements.

Figure 2: Morphological contour extraction in the text.

Figure 3: Active contour model contour extraction.

Through the use of this article's dynamic time warping method and sports history image method to recognize the actions of martial arts action concentration, as can be seen from Table 2, the success rate is less than 50%, which shows that the dynamic time warping method in this paper can effectively identify martial arts decomposition actions.

Figure 4: Decomposed original image of martial arts action.

Figure 5: Morphological contour extraction in the text.

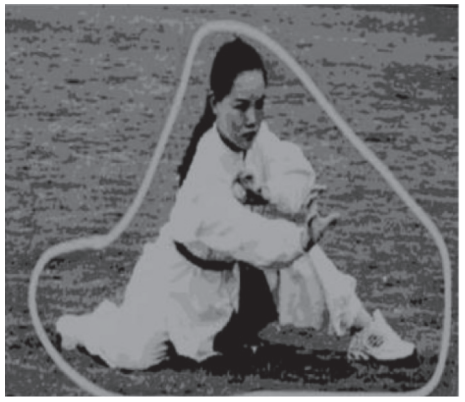


Figure 6: Active contour model contour extraction.

Table 1: Comparison table of time consumption and sharpness of image contour extraction.

Number of images/groups	Morphological operation		Active contour model	
	Contour extraction (s)	Clarity (%)	Contour extraction (s)	Clarity (%)
1	1.21	98.06	10.24	51.07
2	1.08	99.12	11.31	50.34
3	1.34	98.53	10.42	49.71
4	1.03	96.97	12.36	45.33
5	1.31	97.12	15.67	46.78

Table 2: Successful recognition data of the first set of martial arts action sample sets.

Number of actions in sample set/piece	Recognition success rate of dynamic time warping method (%)	Motion history image method recognition success rate (%)
1	98.31	49.87
2	96.78	49.08
3	95.49	48.31
4	98.69	49.06

5. Conclusion

Image processing and computer vision feature recognition methods are combined to decompose martial arts actions, and a method of decomposing and recognizing human martial arts actions is proposed based on image recognition. Firstly, through the morphological gradient operation, most of the noise background can be eliminated, and then, the contour edge of the human body is obtained. The contour edge of each frame of the video is extracted and accumulated in the same image. The accumulated edge image is used to calculate the grid-based HOG, and the image action feature vector is obtained; secondly, using the improved dynamic time warping theory combined with the characteristics of the angle change of each joint under the action time sequence, the similarity of the joint change sequence among various martial arts movements can be identified, and then, the classifier is designed and directed. It inputs the time-varying feature data of human actions in images so as to realize the decomposition process of martial arts actions based on image recognition. The experimental results show that the use of image recognition can effectively decompose martial arts movements.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] C. Wang, H. Chen, R. Zhang, D. Zhu, Q. Wang, and S. Mei, "Research on DTW action recognition algorithm with joint weights," *Journal of Graphics*, vol. 37, no. 4, pp. 537–544, 2016.

[2] F. Huang, J. Cao, and X. Ji, "Two-person interactive action recognition algorithm based on multi-channel information fusion," *Computer Technology and Development*, vol. 26, no. 3, pp. 58–62, 2016.

[3] C. Chao, "Multi-threshold optimization of moving image contour feature extraction method," *Journal of Shenyang University of Technology*, vol. 41, no. 3, pp. 315–319, 2019.

[4] X. Lu, M. Liu, Q. Long, and Z. Chen, "Moving target feature detection algorithm based on gray histogram," *Computer and Modernization*, no. 6, pp. 71–75, 2019.

[5] Y. Cao, S. Bai, and M. Cao, "A compressed sensing sampling algorithm for image adaptive blocking," *Chinese Journal of Image Graphics*, vol. 21, no. 4, pp. 416–424, 2016.

[6] J. Gao and X. Chen, "Research on human motion capture technology based on AR dynamic images," *Modern Electronic Technology*, vol. 41, no. 8, pp. 144-145+150, 2018.

[7] Y. Qian and Y. Shen, "Hybrid application of gesture feature and depth feature in image action recognition," *Acta Automatica Sinica*, vol. 45, no. 3, pp. 626–636, 2019.

[8] S. Chen, W. Wei, and B. He, "Human action recognition method based on improved deep convolutional neural network," *Application Research of Computers*, vol. 36, no. 3, pp. 945–949, 2019.

[9] X. Zhang and S. Hou, "Image weighted fusion algorithm based on guided filtering and fractal dimension," *Packaging Engineering*, vol. 39, no. 9, pp. 220–227, 2018.

[10] H. Zhang and L. Chang, "Infrared and visible image fusion based on directional guided wave enhancement," *Laser and Infrared*, vol. 50, no. 4, pp. 507–512, 2020.

[11] Y. Zhou and C. Li, "On the development of martial arts in scientific fitness," *Fighting: Sports Forum*, vol. 6, no. 10, pp. 88–90, 2014.

[12] P. Zhang and D. Yu, "The fitness function and mechanism of altitude fitness and aerobic power," *Journal of Shanghai Sport University*, vol. 40, no. 2, pp. 51–55, 2016.

[13] P. Zhang, "Application of new physical training methods in wushu sanda team of mianyang normal university," *Journal of Suzhou Education College*, vol. 19, no. 1, pp. 160-161, 2016.

[14] Y. Xue and W. Ruan, "Color image retrieval method based on artificial fish swarm algorithm," *Journal of Xi'an Polytechnic University*, vol. 5, no. 30, pp. 651–656, 2016.

[15] J. Wan and H. Li, "Production line parts sorting system based on image recognition," *Modern Electronic Technology*, vol. 39, no. 12, pp. 62–65, 2016.

[16] P. Yu, J. Zhao, and J. Zhang, "Research on convolutional neural network image recognition based on nonlinear correction function," *Science Technology and Engineering*, vol. 15, no. 34, pp. 221–225, 2015.

[17] Q. Tao, S. Wenlei, K. Jinsheng, and Y. Liang, "A seat comfort evaluation method based on optical motion capture technology," *Journal of Hebei University of Science and Technology*, vol. 36, no. 5, pp. 459–466, 2015.

[18] G. Zhu and L. Cao, "Human action recognition based on the skeleton information of the Kinect sensor," *Computer Simulation*, vol. 31, no. 12, pp. 329–333, 2014.

[19] S. Song, Q. Xu, and B. Liao, "Research on the impact of wind power connection on current protection of distribution network," *Computer Simulation*, vol. 32, no. 1, pp. 137–142, 2015.

[20] L. Rong and Y. Xu, "Research on image feature extraction algorithm based on visual information," *Electronic Design Engineering*, vol. 24, no. 9, pp. 188–190, 2016.

[21] L. Xu and D. Liu, "Visual image recognition and simulation of bus dangerous actions," *Computer Simulation*, vol. 32, no. 6, pp. 150–153, 2015.

[22] M. Wang and H. Sun, "Human action recognition based on layered motion pose covariance," *Computer Application Research*, vol. 32, no. 12, pp. 3794–3797, 2015.

[23] B. Chen and J. Sun, "Human action recognition based on local constraint linear coding," *Journal of Beijing University of Aeronautics and Astronautics*, vol. 41, no. 6, pp. 1122–1127, 2015.

[24] T. Lin, X. Liu, X. Zhang, M. Yan, and L. Zhexing, "DTI image segmentation algorithm based on new morphological gradient parameters," *Television Technology*, vol. 39, no. 6, pp. 5–7+35, 2015.

[25] J. Song, H. Zhang, Z. Gao, Y. Zhang, Y. Xue, and G. Xu, "Human action description algorithm based on deep and dense spatiotemporal interest points," *Pattern Recognition and Artificial Intelligence*, vol. 28, no. 10, pp. 939–945, 2015.

[26] Y. Ren, X. Dong, and Y. Su, "Image saliency region extraction algorithm based on adaptive manifold similarity," *Journal of Shandong University*, vol. 47, no. 3, pp. 56–62, 2017.

[27] H. Xie, T. Tang, D. Xiang, and Y. Su, "Scale-adaptive SAR image saliency detection method," *Computer Engineering and Applications*, vol. 51, no. 20, pp. 145–152, 2015.

[28] R. Wang, J. Chen, L. Jiao, and Y. Su, "Block adaptive compression perception algorithm based on visual saliency," *Journal of Huazhong University of Science and Technology (Nature Science Edition)*, vol. 43, no. 1, pp. 127–132, 2015.

[29] H. Wang, "Video image saliency detection method fused with motion features," *Journal of Science and Economics Guide*, vol. 28, no. 11, pp. 22–25, 2016.

[30] M. Wang, F. Huang, Y. Song, Z. Liu, and N. Ma, "Adaptive enhancement algorithm for hybrid image of artificial fish swarm and particle swarm," *Computer Measurement & Control*, vol. 20, no. 10, pp. 2805–2807, 2012.

Hindawi
Scientific Programming
Volume 2023, Article ID 9782074, 1 page
https://doi.org/10.1155/2023/9782074

Retraction

Retracted: Remote English Teaching Resource Sharing Based on Internet O2O Model

Scientific Programming

Received 1 August 2023; Accepted 1 August 2023; Published 2 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Z. Hou, "Remote English Teaching Resource Sharing Based on Internet O2O Model," *Scientific Programming*, vol. 2022, Article ID 1217807, 10 pages, 2022.

Hindawi
Scientific Programming
Volume 2022, Article ID 1217807, 10 pages
https://doi.org/10.1155/2022/1217807

Research Article

Remote English Teaching Resource Sharing Based on Internet O2O Model

ZuoXun Hou

School of International Studies, University of Science and Technology, Anshan 114051, Liaoning, China

Correspondence should be addressed to ZuoXun Hou; zuoxun@ustl.edu.cn

Received 8 November 2021; Revised 8 December 2021; Accepted 21 December 2021; Published 13 January 2022

Academic Editor: Baiyuan Ding

Aiming at the problems of low success rate, delay, and high communication cost in distance English teaching resource sharing, this paper puts forward a method of distance English teaching resource sharing based on Internet O2O mode. Based on the model of distance English teaching resource sharing, this paper designs four processes: query, reply, resource substitution, and resource sharing optimization. Experimental results show that the proposed method can achieve high success rate of resource sharing, low latency, communication cost, and high transmission efficiency. Therefore, it is an effective method.

1. Introduction

English is a practical course. In Modern English teaching practice, we should not only keep and carry forward the advantages of traditional humanistic teaching, but also make full use of the media and network resources under the condition of modern educational technology, so as to ensure that English learners have active and full practice opportunities. With the continuous improvement of educational concepts and the rapid development of information technology, the traditional teaching model has been unable to meet the teaching needs; distance teaching model came into being [1–3]. Distance teaching is a new teaching mode which is based on the supplement or reformation of network technology to the traditional classroom teaching and learning [4]. English online learning platform has become an indispensable tool for autonomous learning and daily English teaching. Especially when the Internet is connected with the campus network, the role of the network has changed from the teaching assistant resource to the teaching platform, which makes the online learning become a brand-new learning method and realizes the interaction of English teaching under the network environment [5]. In the process of distance online learning, resource sharing can enhance the user's learning experience and promote the popularity of high-quality teaching resources. The contradiction between the scarcity of high-quality teaching resources and the popularization of higher education restricts the improvement of the quality of higher education. Distance learning resources are not only inadequate in total quantity, but also unbalanced in structure. The main factors affecting the level and efficiency of distance education resource sharing are resource construction, design of resource sharing mechanism, and technical support of resource sharing. Resource sharing is a complex system engineering [6, 7]. At present, scholars have carried out the research on teaching resource sharing. Chen et al. [8], based on 5G and FPGA system, studied the sharing of martial arts teaching resources and realized the educational goal of the reform of martial arts teaching system. Yuan [9] puts forward an improved algorithm of neural network path sorting based on path sorting method and realized the sharing of network teaching resources through link prediction of online learning knowledge base. Yao et al. [10] found that, in the world of major public health emergencies, distance education and resource sharing platform building is becoming more and more urgent and important. Based on the teaching practice, the paper analyzes the main problems and development bottlenecks and shares the experience of distance teaching through practical application. In a paper by He et al. [11], the construction of the teaching resource database of vocational education and its popularization and application in other

relevant colleges and universities play an important role in promoting the sharing of high-quality teaching resources, speeding up the innovation of teaching methods and improving the service capability of vocational education industry. Although the above research has made some progress, the success rate of distance English teaching resource sharing is low, the delay and communication cost are high, and the distance English teaching resource sharing has limitations and cannot be widely popularized in practical application. Therefore, a distance English teaching resource sharing method based on Internet O2O mode is proposed.

O2O mode is the thinking mode that applies to commercial domain at first; namely, the entity part below the line and the Internet part on the line are integrated with each other. In recent years, O2O model has been widely used in education and teaching, "Internet plus education" is an effective combination of information technology and school teaching. It can effectively integrate network information resources and modern teaching, and it is an innovation to the traditional teaching mode.

2. Resource Sharing Model of Distance English Teaching

The "O2O model" is the most valuable and favored business model in the Internet era. With the development of Internet technology, the "O2O model" of online education has developed from 1.0 era to today's 3.0 era. In the 3.0 era of O2O model, online and offline collaboration, and gradually forming a complete teaching environment, teachers and students becoming more and more involved in information interaction, transmission and feedback, teacher-student, student-student exchange, inquiry, and cooperative learning will become a new normal.

In the distance English teaching resource sharing system, there are a total of N users (teachers, students), represented by a set of N users who search for and use Internet resources in their shared spaces according to their needs [12]. Each user is installed with a resource sharing system, which can provide each user with a unique ID, location, current time, and other information. In addition, each user's Internet client has a certain amount of computing power, storage capacity, and communication capacity. Teachers make teaching courseware, teaching media video, test paper, literature, and other teaching resources in their personal identity. Through the management module of system resource database, the data of various database manufacturers can upload batch resources. Each user's local storage stores some resources, which can be pictures, audio, video, etc. [13–15]. Each resource can be shared through the Internet "O2O mode" to users who need this resource. In the Internet, when a user needs a resource, if there is no such resource in the user's local memory, it generates a query message and sends the query message in the Internet. This user can be called the source user. For a query message, the source user sets a query message lifetime for it. After any other user receives this query message, if he finds that the local storage has the resource, he will reply the resource to the source user through the Internet to realize the resource sharing. Resource sharing in the Internet, include how to query and how to reply. The O2O model stripped the browse-download-exchange link in the traditional resource sharing process chain and replaced one connection mode (online to offline) with another one (offline to offline). Then it is necessary to analyze the contract between resources and users. Establish the reply process under the new connection mode. Only in this way can the resource sharing in O2O mode ensure the minimum communication cost. In this mode, the Internet provides all the English skills resources, such as the fine course of English in listening, speaking, grammar part in accordance with the easy steps for video shooting or acquisition by video, and by importing the platform, users can select the corresponding course content in the course selection system for extended learning according to their own needs. The course selection system can link other teaching resources for users to use. In summary, the above is also the basis for the construction of this paper. Therefore, the model of distance English teaching resources sharing in Internet O2O mode is shown in Figure 1.

3. Distance English Teaching Resource Sharing Based on Internet O2O Model

3.1. Inquiry Strategies for Distance English Teaching Resources. According to the principles of interactivity, innovation ability training, and scientificity, the query strategy of distance English teaching resources is designed. The source user sends the query message according to this strategy. The principle of interactivity is to have good interactivity and give corresponding feedback to students' learning activities in time. The performance of English teaching knowledge should be operable, not the electronic relocation of English teaching resources. The cultivation principle of innovative ability is the ability of knowledge innovation and information acquisition, which is the core of contemporary quality education. Educational software should adopt a variety of teaching strategies, so as to fully reflect the role of students' cognitive subject and enable students to think actively in the learning process rather than passively accept knowledge, so as to play its due role in cultivating students' innovative ability and enhancing information cultural literacy. The principle of scientificity is that the knowledge to be expressed in distance English teaching resources should be scientific, the wording should be accurate, and the writing should be smooth, in line with the internal logical system of knowledge and students' cognitive structure. The goal of designing query strategy is to improve the success rate of resource sharing and reduce the latency and communication cost. The schematic design of the data query is shown in Figure 2.

When the source user needs a resource, if the resource is not in local storage, it generates a query message and sets a query message ID for the query message. The query message ID consists of the source user ID and the query ID, where the query ID can be generated in order of the source user query. The query message ID set in this way is globally unique. The content of the query message includes the query message ID, the query request to get the resource,

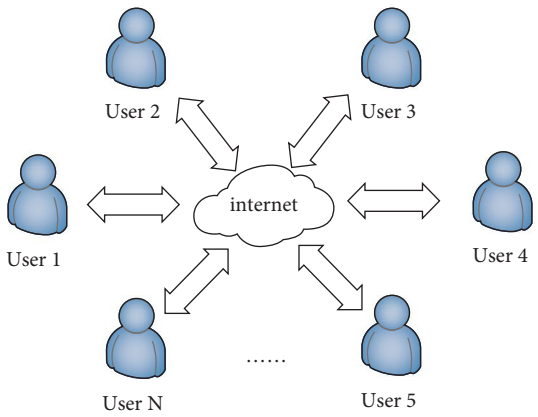


Figure 1: Distance English teaching resource sharing model.

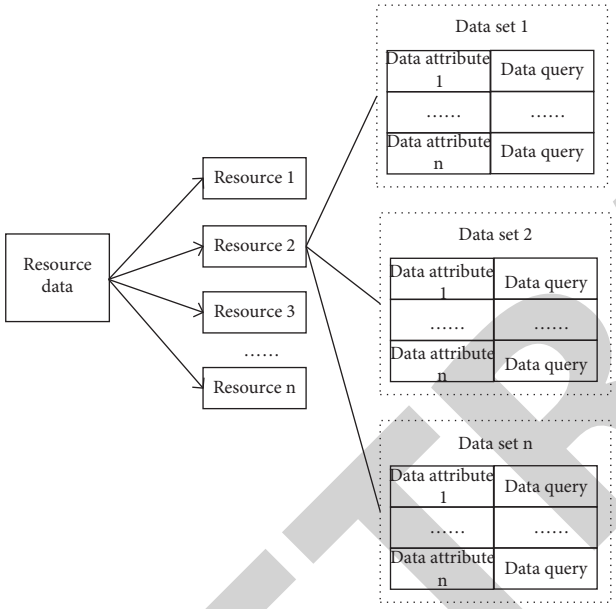


Figure 2: Schematic diagram of data query composition.

and the query message lifetime. In addition, in order to facilitate replying to the source user by the user who owns the resource, the content of the query message includes current time and location of the source user and source user's requirements resource information. Before the existing resource data in the Internet is shared, M is defined as a separate resource pool for resource request user i and c is a collection of users in a given Internet. M is divided by using the following formula:

$$M = \frac{i \times q}{c \times G_i(c)}, \tag{1}$$

where q represents the number of tasks expected by the job, and $G_i(c)$ represents the number of resource-sharing jobs.

In each resource pool of the Internet, resources are allocated using the following methods:

$$w = \sum_{i=1}^{N} (M_i \times J), \tag{2}$$

where J stands for the balance of tasks assigned to each job.

Assuming that l represents the task assignment priority constraint, g represents the query request, and Δw_i represents the communication energy consumption, the query results are as follows:

$$C_X = \sum_{i=1}^{N} [g - l(\Delta w_i)] \times w. \tag{3}$$

3.2. Methods for Responding to Distance English Teaching Resources. If each user forwards a resource along the network channel, there will be a very high communication cost. To reduce communication costs, distance-based forwarding strategy is adopted [16]. The advantages of designing the response method of distance English teaching resources are as follows.

3.2.1. Maximizing Resource Utilization. Distance English teaching is the modern expression of distance education. Based on network technology and combined with excellent educational resources of colleges and universities, it spreads the best teachers and teaching achievements of colleges and universities in all directions.

3.2.2. Flexible Learning Methods. Distance English learning is not limited by time and space. The response method of distance English teaching resources is generally a combination of online counseling and face-to-face teaching. Online counseling can teach and learn online through students and teachers. At the same time, it can communicate, answer questions, and complete homework and exams through the platform.

3.2.3. Individualization of Teaching Form. It is more conducive to the development of students to realize a complete system tracking record of each student's data, learning process, and stage, and put forward different personalized learning suggestions or plans according to the data of different students.

3.2.4. Teaching Management Automation. The response method of distance English teaching resources has the functions of automatic management and remote interactive processing. The whole process is automatically processed by a special management system, which makes up for the shortcomings of large amount of manual operation and low efficiency. Suppose the user of the forwarding resource is i, and other users of the forwarding resource i compute their waiting time t after i' forwards the resource.

$$t = \sum_{i=1}^{N} (d_1 - d_{ii'}) \times t_g. \tag{4}$$

In the formula, d_1 is the communication distance, $d_{ii'}$ is the distance between i and i', and t_g stands for time granularity and can be set to several milliseconds (e.g., 1 ms).

The farther away from the user's i, the shorter the wait for forwarding. Therefore, the user farthest from the user i will first forward the resource. Other demanders can reduce communication costs by not forwarding resources when they hear that users have forwarded them [17]. If the Internet is not connected, that is, the user i cannot get any other user information and continues to forward the resource, and the user will adopt the store-port-forward strategy.

Therefore, the reply strategy of distance English teaching resources is designed to avoid the conflict between multi-users and make full use of remote online resources.

3.3. Establishment of an Internet Resource Data Replacement Storage Model. As a resource returns from the user who owns the resource to the source user along the routing path, some users who continue searching for the resource in the routing path receive the resource. In the Internet's existing storage mode, when a user receives a resource, even if the user does not need the resource, it stores the resource in the local storage for other users to query. As resources accumulate over time, you may run out of local storage space, so you need to delete some resources. This section designs a replacement strategy for a resource that deletes the resource when the local store runs out of space.

In the resource data substitution storage model, the effective scheduling of resource data is firstly carried out [18–20]. The specific methods are as follows:

(1) A shared resource data scheduling model shall be established based on the different status types of the acquired resources:

$$S = \frac{v}{E}(t+1), \tag{5}$$

where v represents resource job tasks, and E represents dependencies between tasks.

(2) Assuming that the computational capacity of the scheduling nodes in the model is proportional to the storage capacity, the nodes are divided into K levels according to the computational capacity level, which is expressed by the following formula:

$$l_p = S\sum_{i=1}^{N}\left(\frac{L}{\Delta w_i + E\rho}\right), \tag{6}$$

where ρ represents a constant factor greater than 1 and L represents the computational power level of the node.

Assuming that p_i represents the computing power of any node in the model, p_i is calculated using the following formula:

$$p_i = \frac{T_i \times N}{B'} \times l_p, \tag{7}$$

where B' represents the size of the data processed on the node i, N represents the number of nodes, and T_i represents the time taken by the node to process the data B'.

(3) Scheduling tasks of shared resources is by using the method of upward ranking value, defining the average value of $\overline{x}_i$ representation as the overhead of scheduling shared resource data in the Internet [21, 22], and calculating $\overline{w}_i$ by using the following formula:

$$\overline{w}_i = \frac{\sum_{i=1}^{N} w_{i,j}}{P_i \times \overline{x}_i}, \tag{8}$$

where $w_{i,j}$ represents the estimated execution time of the task n_i on the processor p_j.

According to formula (8), the upward sorting value of each task in the scheduling of shared resource data in the Internet on different types of processors can be obtained. Use formula (9) to state the following:

$$F(n_i) = \frac{1}{|k|} \times \overline{w}_i, \tag{9}$$

where k stands for heterogeneous cluster characteristics.

(4) Assume that $Q_{g,r}$ represents the X task set of the completed task g on the node with computational power of r level and T_x represents the completion time of the task $X_m(Q_{g,r})$ in the X set. Calculate the average completion time of the task using the following formula:

$$T_p = \frac{X_m\left(Q_{g,r}\right) \cdot T_x}{Q_{g,r}}. \tag{10}$$

(5) The minimum completion time for assigning task n_i to a cluster is T_c, and T_c is calculated using the following formula:

$$T_c = \min \times \frac{T_p \cdot n_i}{X}. \tag{11}$$

Thus, the optimal scheduling of shared resource data in the Internet is accomplished effectively, and a resource value is calculated for each resource on the basis of sufficient scheduling. When a user's local memory space is less than a preset threshold (such as 100 MB), the user's local memory space is insufficient. When a user runs out of local storage space, the user deletes the least valuable resource in turn, based on the resource value of each resource, until the user's local storage space is greater than or equal to the threshold. One user calculates the resource value of a resource in local storage according to the following formula:

$$R = \frac{Q_V}{Q_s}, \tag{12}$$

where Q_V is the resource prevalence of this resource and Q_s is the storage space occupied by this resource. The initial value of resource popularity for any resource is 1. When a user replies to a resource that the source

user needs, the user adds 1 to his or her resource popularity with respect to the resource. From formula (12), it can be concluded that the resource value of a resource is proportional to its resource prevalence and inversely proportional to the storage space it occupies [23]. Therefore, the greater the resource popularity of a resource, or the smaller the storage space it occupies, the greater the resource value of that resource.

3.4. Remote Mapping of English Teaching Resources. In the process of optimizing the sharing of database resources, the method of principal component analysis is used to filter the components of data resources, retain the components of data resources with large variance and more information, build the interclass matrix and intraclass matrix of the attributes of data resources, and transform the original nonlinear data resources space into linear space with higher dimensions through a given linear mapping, and solve the various data distribution points in high-dimensional space through the calculation of Euclidean distance [24, 25]. The specific steps are detailed below.

Assuming that R_h represents the high-dimensional space of the database, Q_r represents the resource vector corresponding to the high-dimensional space of the database and satisfies the conditions of $Q_r = \{r_1, r_2, \ldots, r_n\}^T$, and Q_l represents its low-dimensional space vector and satisfies the conditions of $Q_l = \{l_1, l_2, \ldots, l_n\}^T$, then the resource components in distance English teaching are filtered by means of formula (13) combined with principal component analysis [26–28], and the data resource components with large variance and much information are retained:

$$R_i = \frac{G}{\gamma} \times \frac{R_h \times \delta}{Q_r^T} \times \frac{Q_l}{E_i \times L_i}, \tag{13}$$

where E_i represents the amount of component information contained in the original data resource, L_i represents the amount of principal component contained in the data resource, γ represents the probability of having the same principal component data, G represents the variable of the principal component data, and δ represents the contribution rate of variance of the sample data.

It is assumed that Z represents the covariance matrix of the normalized variables of each data resource, $c(j, f)$ represents the j component of the f attribute of c, N_j represents the number of samples of the j attribute, F_k represents the number of types of data attributes, η represents the distribution probability of different data resource attributes in all data of a heterogeneous data source, and R_q represents the set of neighboring points of data resources. The interclass matrix P_e and the intraclass matrix P_d of data resource attributes are constructed by using the following formulas:

$$P_e = Z \times \frac{N_j * F_k}{P_b \pm c(j, f)} \cdot \eta, \tag{14}$$

$$P_d = \frac{R_q \times P_e}{\eta}, \tag{15}$$

where P_b stands for quantified interclass optimization criteria.

Assuming that Q_c represents the nonlinear data resources of the original space and Q_R represents the linear space with higher dimensions, then formula (16) is used to transform the original nonlinear data resources space into the linear space with higher dimensions through a given linear mapping:

$$Q_k = \frac{Q_c}{c'} \times Q_R \cdot U_c, \tag{16}$$

where c' represents the number of times the attributes of data resources are superimposed and U_c represents the weight vector of different attributes of data resources.

Suppose that $Q_{zj'}$ represents the j' neighborhood data resources of z, and the distribution of each data resource point in the high-dimensional space is calculated by Euclidean distance:

$$Q_X = \frac{\left\{Q_{zj'\,ij} \times X_{\|Y\|}\right\}^{\vartheta}}{Q_k} \times Y, \tag{17}$$

where $X_{\|Y\|}$ represents the distribution of regular terms, Y represents the local weight matrix of data points, and ϑ represents the approximation of linear combination of data points.

To sum up, in the process of optimizing the sharing and access of English teaching resources, the method of principal component analysis is applied to filter the data resource components, to preserve the data resource components with large variance and more information, and to construct the interclass and intraclass matrices of the data resource attributes [29–31], to transform the original nonlinear spatial data resources into linear space with higher dimensions through a given linear mapping, and to calculate the distribution points of each data point in the high-dimensional space by Euclidean distance, which lays a foundation for optimizing the sharing and access of distance English teaching resources.

3.5. Optimization of Distance English Teaching Resources Sharing. In the process of optimizing distance English teaching resource sharing, the multidimensional state space of data resource structure is established based on the distribution point Q_X of data resource obtained in Section 3.4 and the theory of reconstruction of chaotic phase space, and the optimal time delay of the reconstruction of phase space of data resource structure is calculated. The probability distribution curve of the characteristics of data resource structure is obtained, the characteristics of chaotic correlation dimension of data resource structure are extracted, and the cluster center value of different data resource structure features is given [32–34]. The specific steps are detailed below.

Assuming that $\{t_1, t_2, \ldots, t_n\}$ represents the time series of the distribution of each data resource in a resource, F_s represents the result of phase space reconstruction of each data resource structure, and ε represents the embedded

dimension, the multidimensional state space of data resource structure is established based on the distribution point Q_X of each data resource in high-dimensional space obtained in Section 3.4 and integrated into the theory of chaotic phase space reconstruction [35, 36]:

$$I_Y = \frac{Q_X \times \{t_1, t_2, \ldots, t_n\}}{F \times \varepsilon \otimes O_p} \times_s t_c, \tag{18}$$

where O_p stands for delay time mutual information and t_c stands for optimal delay time.

Assuming that W_{O_p} represents the weight of mutual information of delay time and $d_{m\mu}$ represents the μ vector formed in m dimension phase space, formula (19) is used to calculate the optimal time delay t_c for reconstructing phase space of data resource structure:

$$t_c = \frac{W_{O_p} \times \kappa}{I_Y \pm d_{m\mu}}, \tag{19}$$

where κ represents the minimum embedding dimension of phase space reconstruction.

Assuming that $d(f)_{m\mu}$ represents the false nearest neighbor of $d_{m\gamma}$ and $C_{s\,dj}$ represents the proportion curve of the false nearest neighbor, the probability distribution curve of the structural characteristics of data resources is obtained by using the following formula:

$$C_r = \frac{d(f)_{m\mu} \times t_c}{C_d} \times \kappa. \tag{20}$$

Assuming that N_p represents the characteristics of chaotic correlation dimension of data resource structure, R_k represents the logarithm of associated phase points, and G_u represents the number of vector points in phase space reconstruction, formula (21) is used to calculate the characteristics of chaotic correlation dimension N_p of data resource structure:

$$N_p = R_k + \frac{G_u}{C_r}. \tag{21}$$

Assuming that the N_G represents the number of cluster centers selected from the sample, the optimal result of sharing characteristics of different data resources is given by using the following formula:

$$R_n = \frac{N_G \times N_p}{C_r}. \tag{22}$$

Based on the calculation results of formula (22), data resource sharing and optimization can be completed. The specific process is shown in Figure 3.

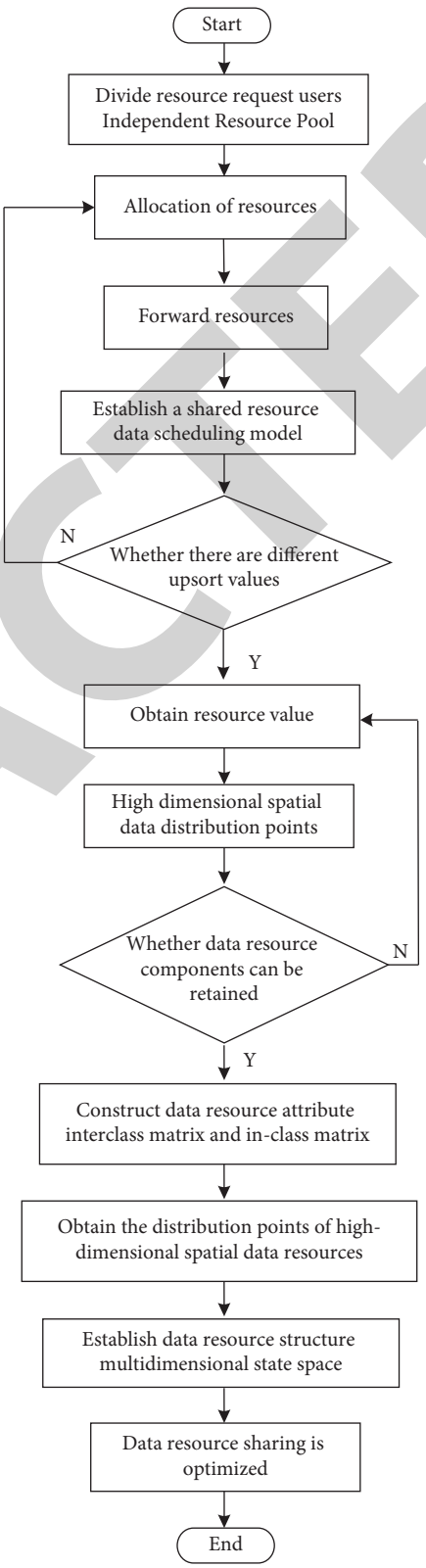


Figure 3: Detailed flowchart.

4. Simulation Experiment Section

In order to verify the effect and feasibility of distance English teaching resource sharing method based on Internet O2O mode, a simulation experiment is set, and the parameters required for the experiment are shown in Table 1.

The indicators of resource sharing include success rate, communication cost, and unit throughput. As an effective resource sharing method, we should achieve high success rate and low average delay as much as possible. In addition, lower communication cost and higher efficiency should be realized. Therefore, this method is compared with the

Table 1: List of simulation parameters.

Parameter name	Parameter value
Network layer protocol	WAVE short-message protocol
Range of communication	250 m
Number of query messages	500–1000
Size of each English teaching resource	10 MB
Number of nodes	20–100
Memory size	256 GB
Data transfer rate	100 Mbs

methods in literature [8, 9] to verify the performance of the designed method.

The implementation process of distance English teaching resource sharing method based on Internet O2O mode is shown in Figure 4.

According to the above process, the success rate, communication cost, and unit throughput of distance English teaching resource sharing method are set, and the results are as follows.

The success rate is defined as the number of resources successfully obtained by the source user, divided by the total number of query messages of the source user. The results of query message quantity and success rate of different methods are shown in Figure 5.

As can be seen from Figure 5, the success rate of the proposed method is higher than that of the two methods in the literature. In the proposed method, more than 90% of the resources are successfully obtained, which shows that the proposed method has a high query success rate and is more suitable for practical application.

The communication cost is defined as the total number of messages sent; efficiency is defined as the number of resources successfully obtained by the source user divided by the communication cost. The comparison results of communication costs of different methods are shown in Figure 6.

As can be seen from Figure 6, compared with the two methods in the literature, the method proposed in this paper has lower communication cost. It can be seen from the figure that the communication cost of the method in reference [9] is low, but combined with the average delay results in 3.2, it can be found that it cannot be used as a good resource sharing method.

Unit throughput refers to the average successful packets sent in the network per unit time after the experiment. Figure 7 shows the change of network unit throughput with the change of data volume in different data quantities.

As can be seen from Figure 7, with the increase of the amount of data, the average throughput in the network shows a decline in varying degrees, but the method proposed in this paper is always maintained at the top, indicating that the proposed method has more transmission processes with higher average transmission rate, so as to improve the total number of successful data packets shared by distance English teaching resources in the network.

To sum up, the designed remote English teaching resource sharing method based on Internet O2O mode has a high query success rate and low communication cost. With the increase of the amount of data, the average throughput in the network shows a downward trend, which improves the total number of successful data packets of remote English teaching resource sharing and transmission in the network, and has a good effect.

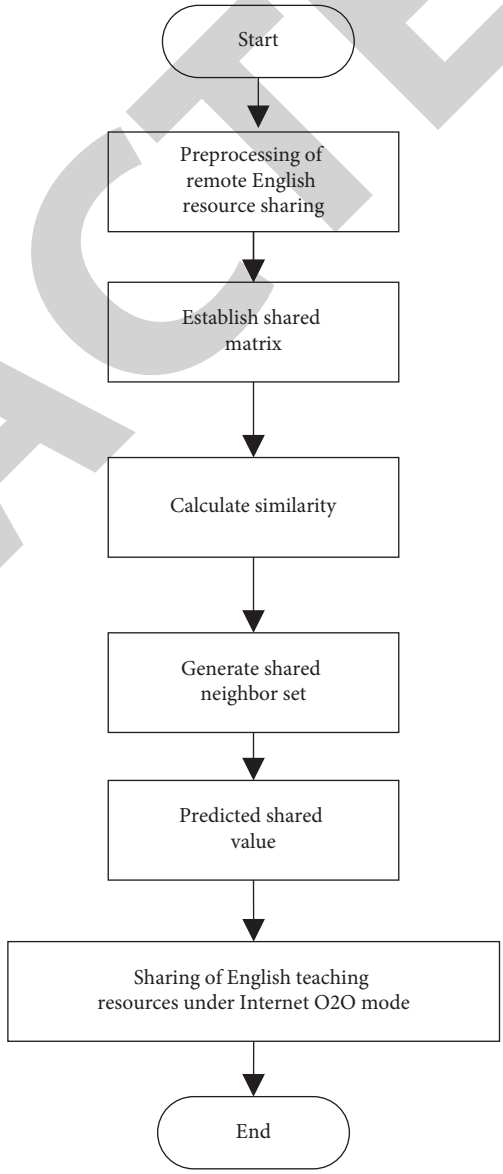


Figure 4: Implementation flowchart of distance English teaching resource sharing method.

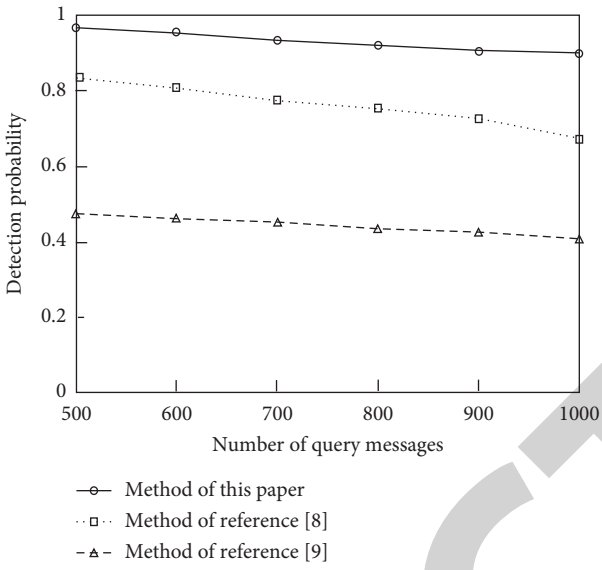


Figure 5: Comparison of success rates of different methods.

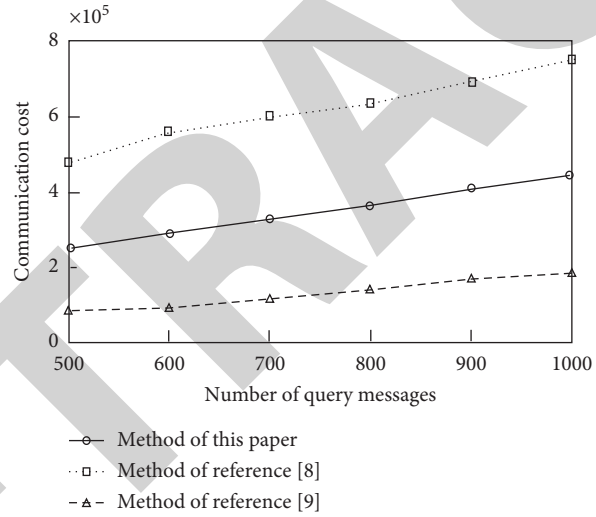


Figure 6: Comparison results of communication costs of different methods.

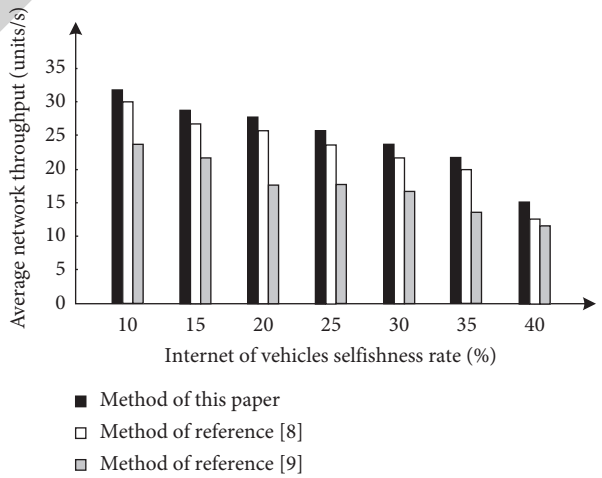


Figure 7: Average network throughput.

5. Conclusion

In order to solve the problems of low success rate, high delay, and high communication cost in Distance English teaching resource sharing, a distance English teaching resource sharing method based on Internet O2O mode is proposed. The experimental results are as follows:

(1) The resource sharing method of distance English teaching based on Internet O2O mode can achieve high success rate of resource sharing and low communication cost

(2) It can ensure that the sharing process has high sending efficiency and provides a certain reference for the sharing of distance English teaching resources

The sharing methods of distance English teaching resources are still developing, and various types of markets are gradually opened. The methods need to be optimized according to the development of Internet O2O mode, so as to truly and accurately provide the basis for the sharing of distance English teaching resources. In the following work, we study the security and omnipotence of sharing in order to further optimize the sharing performance of distance English teaching resources.

Data Availability

The raw data supporting the conclusions of this article will be made available by the author, without undue reservation.

Conflicts of Interest

The author declares no conflicts of interest regarding this work.

References

[1] P. Kawinkoonlasate, "Online language learning for Thai EFL learners: an analysis of effective alternative learning methods in response to the covid-19 outbreak," *English Language Teaching*, vol. 13, no. 12, p. 15, 2020.

[2] A. J. Evans, N. Depeiza, S. G. Allen, K. Fraser, and R. Chetty, "Use of whole slide imaging (WSI) for distance teaching," *Journal of Clinical Pathology*, vol. 74, no. 7, p. 206763, 2020.

[3] V. Domenici, "A course of history of chemistry and chemical education completely delivered in distance education mode during epidemic COVID-19," *Journal of Chemical Education*, vol. 97, no. 9, pp. 2905–2908, 2020.

[4] P.-E. Danjou, "Distance teaching of organic chemistry tutorials during the COVID-19 pandemic: focus on the use of videos and social media," *Journal of Chemical Education*, vol. 97, no. 9, pp. 3168–3171, 2020.

[5] Y. Gao, "Computer-aided instruction in college English teaching under the network environment," *Computer-Aided Design and Applications*, vol. 18, no. 4, pp. 141–151, 2021.

[6] J. Niu, Y. Sun, X. Jia, and Y. Ji, "Key-size-driven wavelength resource sharing scheme for QKD and the time-varying data services," *Journal of Lightwave Technology*, vol. 39, no. 9, pp. 2661–2672, 2021.

[7] B. Zhao, R. Hou, J. Dong et al., "Venice," *ACM Transactions on Computer Systems*, vol. 36, no. 1, pp. 1–26, 2019.

[8] S. Chen and L. Liang, "WITHDRAWN: online resource sharing of martial arts teaching based on 5G network and FPGA system," *Microprocessors and Microsystems*, no. 8, p. 103447, 2020.

[9] Q. Yuan, "Network education recommendation and teaching resource sharing based on improved neural network," *Journal of Intelligent and Fuzzy Systems*, vol. 39, no. 4, pp. 5511–5520, 2020.

[10] S. Yao, D. Li, A. Yohannes, and H. Song, "Exploration for network distance teaching and resource sharing system for higher education in epidemic situation of COVID-19," *Procedia Computer Science*, vol. 183, pp. 807–813, 2021.

[11] B. B. Saravana-Bawan, C. Fulton, B. Riley et al., "Evaluating best methods for crisis resource management education," *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*, vol. 14, no. 6, pp. 366–371, 2019.

[12] J. Tao, "Evaluation of service-oriented English practice teaching under resource sharing -- A case study of English in hotel industry," *Advances in Higher Education*, vol. 3, no. 2, p. 134, 2019.

[13] S. Tsianikas, N. Yousefi, J. Zhou, M. D. Rodgers, and D. Coit, "A storage expansion planning framework using reinforcement learning and simulation-based optimization," *Applied Energy*, vol. 290, p. 116778, 2021.

[14] K. Zheng, "Simulation of emergency resource optimization scheduling for differential distributed storage systems," . *Computer Simulation*, vol. 036, no. 7, pp. 415–418, 2019.

[15] Q. Huang, C. Huang, J. Huang, and H. Fujita, "Adaptive resource prefetching with spatial-temporal and topic information for educational cloud storage systems," *Knowledge-Based Systems*, vol. 181, p. 104791, 2019.

[16] O. Rui, M. Luís, and S. Sargento, "On the performance of social-based and location-aware forwarding strategies in urban vehicular networks," *Ad Hoc Networks*, vol. 93, p. 101925, 2019.

[17] D. G. Zhang, P. Z. Zhao, Y. Y. Cui, L. Chen, and H. Wu, "A new method of mobile ad hoc network routing based on greed forwarding improvement strategy," *IEEE Access*, vol. 7, no. 10, pp. 1–10, 2019.

[18] H. W. Wang, J. R. Lin, and J. P. Zhang, "Work package-based information modeling for resource-constrained scheduling of construction projects," *Automation in Construction*, vol. 109, no. 1, pp. 102958.1–102958.20, 2020.

[19] A. Bz and B. Dc, "Resource scheduling of green communication network for large sports events based on edge computing," *Computer Communications*, vol. 159, no. 6, pp. 299–309, 2020.

[20] J. Zhang, F. Xiong, and Z. Duan, "Research on resource scheduling of cloud computing based on improved genetic algorithm," *Journal of Electronic Research and Application*, vol. 4, no. 2, pp. 4–9, 2020.

[21] A. Khelifa, T. Hamrouni, R. Mokadem, and F. B. Charrada, "Combining task scheduling and data replication for SLA compliance and enhancement of provider profit in clouds," *Applied Intelligence*, vol. 51, no. 10, pp. 7494–7516, 2021.

[22] C. K. Swain, B. Gupta, and A. Sahu, "Constraint aware profit maximization scheduling of tasks in heterogeneous datacenters," *Computing*, vol. 102, no. 6, pp. 2229–2255, 2020.

[23] L. Shi, X. Wang, and R. T. B. Ma, "On multi-resource procurement in internet access markets: optimal strategies and market equilibrium," *Performance Evaluation*, vol. 143, no. 11, p. 102139, 2020.

[24] S. Pang, Q. Feng, Z. Lu et al., "Hippocampus segmentation based on iterative local linear mapping with representative

Hindawi
Scientific Programming
Volume 2022, Article ID 6048088, 9 pages
https://doi.org/10.1155/2022/6048088

Research Article

Identification and Modeling of College Students' Psychological Stress Indicators for Deep Learning

Yuan Tian

School of Education, Xinyang University, Xinyang, Henan 4640, China

Correspondence should be addressed to Yuan Tian; 160208230@stu.cuz.edu.cn

Received 8 November 2021; Revised 29 November 2021; Accepted 9 December 2021; Published 11 January 2022

Academic Editor: Baiyuan Ding

Aiming at the problems of low accuracy of recognition results, long recognition time, and easy interference in traditional methods, a deep learning-oriented recognition modeling method of college students' psychological stress indicators is proposed. First, the ECG signal is collected by the ECG signal acquisition system, and the wavelet transform method is used to denoise the collected ECG signal. Then, the sequential backward selection algorithm is used to select the features of psychological stress indicators to reduce the feature dimension. Finally, based on the convolutional neural network in deep learning technology, a mental pressure indicator recognition model is established and the model parameters are optimized to realize the recognition of college students' mental pressure indicators. Experimental results show that the method in this paper has high recognition accuracy, has high recognition efficiency, is not susceptible to interference, and has certain feasibility and effectiveness.

1. Introduction

People in modern society are exposed to various pressures, and often need to bear pressures from work, life, economy, interpersonal relations, and so on [1, 2]. Continuous stress will bring psychological and physiological obstacles and defects. Tracking and recording individual's daily psychological stress state in real time, timely warning in case of abnormalities, and giving appropriate intervention and guidance can well help individuals maintain a good psychological state [3, 4]. For a long time, people have used interviews, psychological questionnaires, and other methods to monitor their psychological stress status. This method often requires the participation of psychologists. Real-time monitoring is not possible, and it is more difficult to extend to a large number of people under pressure. At present, there are no mature indicators and methods recognized in the industry that can accurately assess the level of psychological stress. Therefore, it is very important to design a method that can accurately identify indicators of psychological stress [5].

In recent years, the use of noninvasive sensors to collect physiological indicators and identify psychological stress has become a major research hotspot. These studies aim to collect relevant physiological indicators and identify psychological stress indicators by wearing sensors that do not affect their daily activities. With the help of wearable devices, users can collect relevant physiological indicators anytime and anywhere and upload data to the service platform via the mobile Internet. The service platform uses the data received to evaluate the psychological stress level of users and, under the guidance of experts, provides users with effective intervention and guidance through the mobile Internet. This is a typical mode of psychological stress tracking services for mobile health [6, 7]. In addition, a pressure identification algorithm based on improved particle swarm optimization BP neural network is proposed in reference [8]. Based on the basic particle swarm optimization model, the shrinkage factor is introduced. Under the action of the shrinkage factor, the boundary limit of speed disappears, and appropriate parameters are selected to ensure the boundedness and convergence of the PSO algorithm, so as to realize the optimization of BP neural network. The stress was induced by mental arithmetic task, the ECG signals under high and low pressure were collected, the eigenvalues of heart rate variability related to psychological stress were extracted, and

the characteristic data were compared and analyzed. The classification model of psychological stress degree is established, and the BP neural network is optimized by the improved PSO model to identify psychological stress. The results show that the recognition rate of psychological stress can reach 94.83%, and the recognition effect is good. Reference [9] proposes a psychological index recognition modeling method based on social media data, summarizes its feasibility in psychological measurement, introduces feature extraction methods, common machine learning algorithms, and application scenarios, and summarizes and respects the advantages and disadvantages of psychological index recognition modeling. This measurement method is based on social media data and has unique advantages compared to the self-report method, such as high timeliness, retrospective measurement, and good ecological validity. However, the method of identifying and modeling mental indicators based on social media also has limitations in terms of learning costs and hardware costs. In the future, researchers need to further explore the association mechanism between social media information and user psychological variables and combine the psychological indicator recognition model with traditional psychological research methods for more exploration and application.

Although the above methods can realize the identification of psychological stress indicators to a certain extent, there are problems of low recognition accuracy, susceptibility to interference, and long recognition time. It can be seen that it is necessary to conduct research in this area. In this context, this article takes college students as the research object and proposes a deep learning-oriented method for identifying and modeling the psychological stress indicators of college students, aiming to improve the effect of identifying psychological stress indicators. Before establishing the model, the ECG signal after the change of college students' psychological pressure is collected by the heart electrical signal acquisition system and the collected ECG signal is preprocessed by the wavelet transform algorithm. So far, the ECG signal preprocessing is realized. After preprocessing, the sequential backward selection algorithm is used to select the characteristics of psychological stress indicators. Finally, the convolution neural network in deep learning technology is used to establish the identification model of college students' psychological stress indicators.

2. Preprocessing of Psychological Stress Indicators for College Students

2.1. ECG Signal Acquisition and Processing

2.1.1. ECG Signal Acquisition. The acquisition and processing of ECG signals are the prerequisite for the identification of psychological stress indicators. When collecting data, the design of the pressure induction scheme and the different pressure states of the collection require scientific selection and judgment. In order to accurately analyze the psychological stress indicators of college students and extract features, it is necessary to process the interference noise in the collected original ECG signals [10, 11].

The ECG lead system of the ECG signal acquisition system selects the I lead system, and the acquisition block diagram is shown in Figure 1. Among them, the lead electrodes RA and RL are respectively connected to the right upper limb and right lower limb, and LA is connected to the left upper limb; the left and right upper limb ECG signals pass through the predifferential amplifier circuit and the main amplifier circuit, band-pass filter circuit, and power frequency trap circuit; the right lower limb is connected to the right leg drive circuit, also through the predifferential amplifier circuit, and finally output from the output of the power frequency trap circuit, and the output signal is connected to the AIO port in NI My RIO to achieve ECG data acquisition [12, 13].

For the study of psychological stress identification methods, the correctness and validity of data are the key to subsequent processing and analysis. Therefore, the collection of ECG signals is the basis for identifying psychological stress. This article builds the circuit according to the block diagram of the ECG acquisition module shown in Figure 1. Then, on this basis, the NI My RIO embedded development platform is used as the acquisition card, and the FPGA technology in NI My RIO is used to control the acquisition card to collect the ECG signals. Connect the output of the acquisition circuit to the analog input of NI My RIO. Finally, use the LabVIEW FPGA module to design the acquisition program and store the collected original ECG data through the FIFO (First Input First Output) memory in the FPGA. The ECG data stored in FIFO memory are finally implemented by the RT terminal programs read and display.

2.1.2. ECG Signal Preprocessing. Due to the characteristics of the ECG signal itself and the influence of the equipment, the original ECG signal collected above often contains a lot of noise. These noises mainly include the noise generated by the equipment itself, respiratory wave noise, and myoelectric interference. These noises will cause changes in the ECG signal waveform to varying degrees, including amplitude changes, baseline drift, and so on. Therefore, these noises must be filtered out before further analysis of the ECG signal [14].

As a signal analysis tool developed in recent years, wavelet transform has its characteristics including orthogonality, direction selectivity, small amount of analysis data, and resolution variability. These characteristics of wavelet transform make it often used for noise filtering and waveform detection of ECG signals [15, 16].

The basic wavelet definition is as follows: suppose $F(t) \in S^2(t)$ represents a square integrable real number space, its Fourier transform is $F(k)$, and $F(k)$ satisfies the allowable condition:

$$A_F = \sqrt{\frac{|F(k)|^2}{|k|}}. \tag{1}$$

Here, $F(k)$ represents a basic wavelet. The wavelet generated by the basic wavelet is

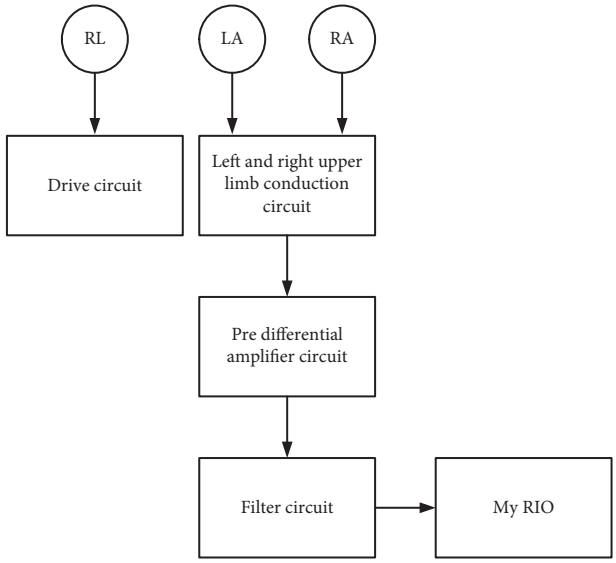


Figure 1: Working principle diagram of the ECG signal acquisition system.

$$F_{ij}(n) = F\left(\frac{n-j}{i}\right) \times A_F. \tag{2}$$

For a continuous signal $F(t) \in S^2(t)$ (a space with limited energy), the continuous wavelet transform (CWT) of $F(t)$ is

$$S(a,b) = \frac{R(a,b)\sum_{k=1}^{n} a_k b_k}{\sqrt{\left(\sum_{k=1}^{n} a_k^2\right)\left(\sum_{k=1}^{n} b_k^2\right)} \times F(t)}. \tag{3}$$

Here, a_k represents the scale factor; b_k represents the delay factor; and $R(a,b)$ represents the continuous wavelet basis function.

In actual ECG signal analysis, discrete wavelet transform (DWT) is generally used to analyze the signal. Discrete wavelet transform generally uses binary transform to calculate the time and scale wavelet coefficients of discrete intervals; that is to say, the choice of scale and time is usually carried out in the exponential way of 2 [17, 18]. The discrete wavelet transform of any function $F(t)$ is

$$W(a,b) = \int_t F_{ij}(a,b) \times F_{ij}(t)\mathrm{d}t. \tag{4}$$

Here, $F_{ij}(t)$ represents the shrinkage threshold function, and its calculation formula is as follows:

$$F_{ij}(t) = V_i^S(j) + (1-\alpha)\mathrm{Re}_{ij}. \tag{5}$$

The two key factors for processing the ECG signal by wavelet transform are the choice of wavelet base and the determination of the number of decomposition layers. Commonly used wavelet bases include Haar wavelet, Daubechies (dbN) wavelet, Mexican Hat (mexh) wavelet, Morlet wavelet, and Symlets wavelet. This article uses the wavedec function of the Matlab analysis software to select the db5 wavelet to decompose the ECG signal in three layers, and on this basis, the wdencmp function is used to achieve the denoising of the original signal. The denoising effect of the original ECG signal is shown in Figure 2. Figure 2(a) is the original ECG signal of a certain subject collected, and Figure 2(b) is the ECG signal after wavelet denoising.

2.2. Feature Selection of Psychological Stress Indicators. According to the ECG signal denoising results obtained in Section 2.1, further select the characteristics of psychological stress indicators. When classifying numerous ECG signal features through the classifier, not every feature is useful, and this part of the feature is considered to be original features [19]. In the field of deep learning, the dimension of the feature space sent into the classifier should not be too high; otherwise, it will cause the so-called "dimension disaster," make the classifier model complex, and reduce the generalization ability [20–22]. Therefore, it is necessary to select the features with less correlation between features from the original feature set as the feature set. The essence of feature selection is to search the optimal or suboptimal subset and select the features with the highest classification contribution to the data set as the feature subset. In this paper, sequential backward selection (SBS) is selected as the feature selection algorithm.

Backward selection is a bottom-up feature selection algorithm. Its core idea is to take the complete set of features to be selected as the initial feature set, delete a feature with the least contribution to classification each time, and gradually eliminate and judge it. If the dimension of the feature decreases to 1, it will stop [23, 24]. The algorithm flow of the backward selection algorithm mainly includes the following processes:

Step 1. Take all the features as the initial $Q_0 = P$.

Step 2. Remove a feature from the feature set, so that the removed feature set Q_g is optimal, namely,

$$\overline{P} = \mathrm{argmax}\left[D\left(Q_g - x\right)\right]. \tag{6}$$

Here, $\overline{P}$ represents removing the feature from the feature set; the minus sign is not a subtraction operation for feature values in the usual sense.

Step 3. Update the feature set so that

$$Q_{g+1} = \sum_{i,j=1}^{M} E\left[W_{ij}(t)\right]. \tag{7}$$

Here, E represents the frequency domain index of the psychological stress index feature and $W_{ij}(t)$ represents the single time domain of the psychological stress index feature.

Step 4. Return to the second step until the feature dimension drops to 1.

According to the above steps, redundant features are eliminated, the accurate selection of psychological stress index characteristics is realized, and the data basis for the identification of the psychological stress index of college students is provided.

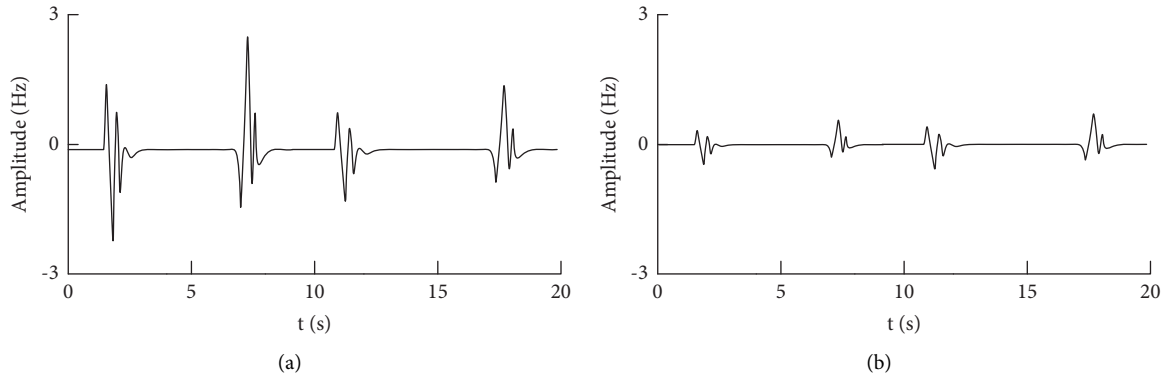


Figure 2: ECG signal comparison before and after denoising. (a) Original ECG signal. (b) ECG signal after denoising.

3. Recognition and Modeling Method of Psychological Pressure Indicators for College Students Oriented to Deep Learning

The identification of psychological stress indicators can be used as the basis for psychological doctors to assist in diagnosis, help psychological doctors treat psychological patients more effectively, and enable people to understand their own stress, reduce pressure in time, and relieve daily stress according to the understanding. In addition, it can also help the school to understand the physical and psychological state of college students for effective management. It can also enable the government to correctly count the general psychological pressure of college students, so as to take some measures to improve social medical care. Therefore, based on ECG signal acquisition and processing and psychological stress index feature selection, this paper uses deep learning technology to accurately identify college students' psychological stress index.

3.1. CNN Model Construction. Deep learning technology subverts the development rules of traditional artificial intelligence systems, enabling computers to simulate the operation mode of the brain and learn and recognize abstract patterns through multilayer convolution neural networks, so as to solve some general pattern recognition problems [25–27], which means that any task involving a large amount of data may benefit from deep learning. Based on the idea of deep learning, this paper puts forward a technical idea of target recognition of college students' psychological stress indicators [28, 29].

Convolutional neural network CNN is a deep learning model that can effectively extract targets. Therefore, CNN is used to identify the psychological stress indicators of college students and establish a CNN model [30]. The CNN model is generally composed of an input layer, a convolutional layer, a subsampling layer, a fully connected layer, and an output layer. The feature extraction work of convolutional neural network is mainly realized by convolutional layer and subsampling layer [31, 32]. In the convolutional layer, a feature plane is composed of the arrangement of neurons, and the weight matrix shared by all neurons in a feature plane is also called a convolution kernel. The feature surface of the upper layer can be output feature surface through convolution and activation function. In traditional neural networks, sigmoid is often used as the activation function, but sigmoid is easy to cause saturation during the gradient descent process, thereby terminating the gradient transfer [33]. The ReLU function has the advantages of fast convergence speed and simple gradient solution. Therefore, the ReLU function is commonly used as the activation function in CNN at present, and the calculation formula is as follows:

$$U(x) = \sum_{i=1}^{M} D_i(w) w_k. \tag{8}$$

Here, $D_i(w)$ represents the radial basis function and w_k represents the neuron.

The convolutional layer gradually extracts different features through the process of convolution. The higher the number of layers, the more advanced and abstract the extracted features.

The subsampling layer is also called the pooling layer, in which the feature surfaces correspond to the feature surfaces in the convolutional layer, and the number is the same. The pooling layer further extracts the output features of the convolutional layer, which can effectively reduce the feature dimension. Commonly used pooling methods include random pooling, average pooling, and maximum pooling. The fully connected layer can summarize the local features extracted by the convolutional layer and the sampling layer and finally enter the classifier for classification [34]. CNN uses the loss function to measure the robustness of the model. It can estimate the difference between the predicted value and the true value in the model. The smaller the value, the better the robustness of the model. Most models use the mean square error as the loss function, and its calculation formula is as follows:

$$E_j^2 = \left| \frac{\max\left(b_{1j}, \ldots, b_{nj}\right) + \min\left(b_{1j}, \ldots, b_{nj}\right)}{2} - \partial_j \right|. \tag{9}$$

Here, b_{ij} represents the supervision information of the mean square error loss and ∂_j represents the cross entropy.

The deep features extracted by CNN contain high-level semantic information and maintain certain invariance to

zoom, translation, and so on, which significantly reduces the gap between high-level semantics and low-level features [35]. The multilayer convolutional neural network structure used in this article includes an input layer, a three-layer convolutional layer, a three-layer sampling layer, a fully connected layer, and an output layer. The size of the convolution kernel is 5×5, and the information of the fully connected layer is extracted as the global feature, that is, the psychological pressure index feature of college students.

3.2. Model Parameter Optimization. For the designed CNN model, it is very important to optimize each node into the best state. This requires a large amount of training in combination with the training samples. Each training is a correction iteration of the whole network parameters, so that the constructed network can conduct a recognition learning on the concerned goals, until the goal cognition learning under various states of the training samples forms a representative description of the concerned goals. The better the network model is optimized, the deeper the learning of the target is, and the more transparent the description is, the higher the accuracy of the model is for detection [36]. Therefore, model parameter optimization is very important.

Suppose there is a fixed training set:

$$C = \left\{\left(c^1, v^1\right), \left(c^2, v^2\right), \ldots, \left(c^m, v^m\right)\right\}. \tag{10}$$

Here, m represents the number of training samples. The batch gradient descent algorithm can be used to train the convolutional neural network. Specifically, for a single training sample (c, v), the cost function can be defined as

$$\mu(W, b; c, v) = \sqrt{\sum_{i,j=1}^{m} \left(Y_{ij} - E_{ij}(Q_i)\right)^2}. \tag{11}$$

Here, Y_{ij} represents the location of the optimized region of the convolutional neural network; Q_i represents the network parameters; and E_{ij} represents the classification probability.

The cost function c is the cost function of a single sample after the conversion between W and b, and then the cost function for m training samples can be defined as

$$\mu(W, b) = \frac{\sum_{i=1}^{m} Y'_{ij} \omega_j}{\sum_{i=1}^{m} \sigma(Q_m)}. \tag{12}$$

Here, Y'_{ij} represents the output value distribution of each neuron in the convolutional layer in response to different inputs and $\sigma(Q_m)$ represents the cost function of the feedforward neural network.

The first term $\mu(W, b)$ of the cost function is the average of the sum of squared differences. The second item is an adjustment item (also called the weight decay item). The purpose is to reduce the magnitude of the weight and prevent overfitting. Assuming that H_l represents the number of neurons in layer l and H_{l+1} represents the number of neurons in layer $l + 1$, the sum of the square weights between each neuron in layer l and all neurons in layer $l + 1$ is G, and its calculation formula is

$$B_{ij}^{l} = \frac{\sum_{i=1}^{l} \sum_{j=1}^{l+1} \left(H_{ij}^{l}\right)^2}{\left(R_r / G_r\right)^2}. \tag{13}$$

Here, R_r represents the weight connection value between neurons and G_r represents the mean value of the corrected distance between neurons.

In each iteration of the gradient descent method, the parameters W and b are updated according to the following formula:

$$\begin{aligned} W_{ij}^{l} &= I_{ij}^{2} \times \mu(W, b), \\ b_{ij}^{l} &= \psi_{ij}^{2} \times \mu(W, b). \end{aligned} \tag{14}$$

Among them b_{ij}^{2} represents the output information of neurons; ψ_{ij}^{2} represents the amount of information transferred between networks.

Finally, the iterative steps of the gradient descent method are repeated to reduce the value of the cost function $\mu(W, b)$, and then the optimal parameters of the CNN model are solved.

3.3. Realization of the Recognition of Psychological Pressure Indicators for College Students. Based on the CNN model constructed in Section 3.1 and the obtained model parameter optimization results, the psychological pressure indicators of college students are identified. The following is a specific analysis of the process of identifying the psychological pressure indicators of college students.

First, they need to obtain the self-report score of college students' psychological characteristics. At present, most of the self-report scores used in the research are self-assessment scale scores. The number of users using social media data for psychological modeling is generally large. For example, the number of users in the research on Facebook application MyPersonality is generally more than 1000, up to 390,000 users. Therefore, in this paper, the psychological characteristics of college students are processed in the form of quantitative coding, that is, the step of feature extraction. Commonly used coding methods include classification, frequency, frequency, building sparse matrix, and so on. Based on the selection results of psychological stress index features obtained in Section 2.2, this paper further obtains the psychological stress index features of college students by establishing a sparse matrix. The sparse matrix is

$$\psi_{(W,b)} = \frac{\omega_s \lambda_m}{\sqrt{\left(R_r / s\right)^2 + \left(\omega_s\right)^2}}. \tag{15}$$

Here, ω_s represents the sparse coefficient of the sample to be identified and λ_m represents an over-complete dictionary.

Second, select appropriate deep learning methods to establish a mapping relationship between the user's self-report score and the corresponding psychological stress index characteristics, and use cross-check to verify the calculation effect of the model. Cross-checking is the most commonly used model performance evaluation method in the deep learning modeling process. The specific operation is to divide the data set into a training set and a test set, and use

the training set to model and evaluate the model performance with the test set. The data set will be divided multiple times until each data set has been done on the training set and also on the test set. Cross-validation can make full use of the original data and avoid the impact of unbalanced random partitioning on model performance, and at the same time, it can also try to avoid overfitting of the model.

Finally, a feature recognition model for college students' psychological stress indicators is obtained. When data related to college students' psychological stress are input, feature extraction, model calculations, and output of the user's psychological feature values can be automatically performed according to the characteristics of the model, and the results of psychological stress indicator recognition can be obtained.

In summary, the general process of identifying and modeling college students' psychological stress indicators includes several main parts such as ECG signal acquisition and processing, feature selection and extraction, modeling and parameter optimization, and result output. The specific process of identifying and modeling college students' psychological stress indicators is shown in Figure 3.

4. Experimental Verification Research

In order to verify the validity and application value of the proposed deep learning-oriented college students' psychological stress index identification modeling method, experimental verification is carried out. In the experiment, the pressure recognition algorithm based on the improved particle swarm optimization BP neural network (method in reference [8]) and the psychological indicator recognition modeling method based on social media data (method in reference [9]) are used as comparison methods. The application effects of different methods are compared and experimental conclusions are drawn.

4.1. Experiment Preparation. Participants in this experiment are undergraduates and graduate students (between 19 and 25 years old) of a certain university. The participants are required to have no mental illnesses and be able to follow the guidance of the experimenters and cooperate to complete the experiment. To ensure the universality of the experiment, testers include engineering students and liberal arts students. There are 60 subjects in this experiment, including 36 males and 24 females. The distribution of test subjects is shown in Table 1.

In order to more effectively evaluate the recognition effects of different methods, in the experiment, the tester's calm state is regarded as a low pressure state, and this part of the mental arithmetic task process is regarded as a high pressure state. During the mental arithmetic task test, 20 mental arithmetic questions will be tested to induce the tester's psychological stress state. At the same time, in order to ensure the effectiveness of the tester's stress state, the time will be set for the questions in the mental arithmetic task during the experiment, and the subjects are required to complete the mental arithmetic questions within a certain time range, which can also enable the tester to enter the test state faster. In this experiment, data collection is combined with questionnaire survey and waveform diagram to intercept effective data.

4.2. Analysis of Experimental Results

4.2.1. Anti-Interference Effect. This article uses the ECG acquisition circuit and NI the RIO embedded development platform as the ECG acquisition equipment, among which the ECG signal leads use the right upper limb, right lower limb, and left upper limb. The ECG signals of 60 students were collected, and a total of 300 groups of ECG data were collected. After removing unstable signals, there were a total of 180 groups of valid data, including 90 groups of high-voltage data and 90 groups of low-voltage data. The 3000 sampling points during the calm phase of the tester were used as low pressure data, and the 3000 sampling points during the pressure induction process were used as high pressure data. Figure 4 shows the ECG waveforms of a subject under low pressure and high pressure.

As can be seen from Figure 4, the collected ECG signals under low-voltage and high-voltage conditions contain interference, but the degree of interference is low, which can more clearly obtain the ECG waveform state of students, indicating that this method can effectively identify the psychological stress state of college students. This is because the wavelet transform method is used to process the ECG signal in Section 2.1.2 of this paper. The ECG waveform without noise can be obtained from the processed waveform, so as to extract the pressure related features.

4.2.2. Recognition Accuracy Rate. During the recognition accuracy test, keep the laboratory environment quiet and the mobile phone silent. After the preparation is completed, start the mental arithmetic experiment. Mental arithmetic experiment is a quantitative analysis of psychological stress, which is divided into two states: no and there. The subjects first have a 60 s resting state, adjust to a calm psychological state, and then start the mental arithmetic task. The task content is the addition of six random three digits. The time of each task is 5 seconds, and the mental arithmetic is repeated 10 times, with a total time of 50 seconds. After the mental arithmetic experiment, the subjects took a short rest to adjust their psychological stress state. *M* P150 physiological parameter recorder was used to collect the ECG signals of subjects. The total number of samples collected in the recognition accuracy test is 60, and the proportion of the number of samples in the training set and the test set is set to 35 : 25. The methods in this paper, reference [8] and reference [9] are used to identify the psychological stress indicators of college students, respectively, and the recognition accuracy results are shown in Figure 5.

By analyzing the experimental result data in Figure 5, it can be seen that there are some differences in the accuracy of identifying s students' psychological stress indicators by using the methods of this paper, method in reference [8], and method inreference [9]. Among them, the highest accuracy of this method in identifying college students' psychological

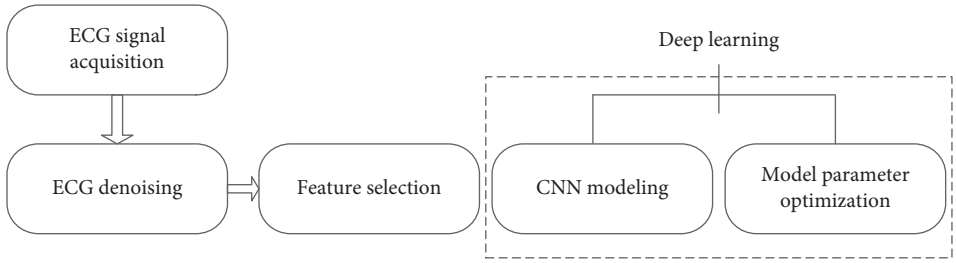


Figure 3: The process of identifying and modeling college students' psychological stress indicators.

Table 1: Subject information.

Education	Subject	Gender	
		Male	Female
Undergraduate	Engineering	12	7
Master	Liberal arts	8	5
Undergraduate	Engineering	9	6
Undergraduate	Liberal arts	7	8

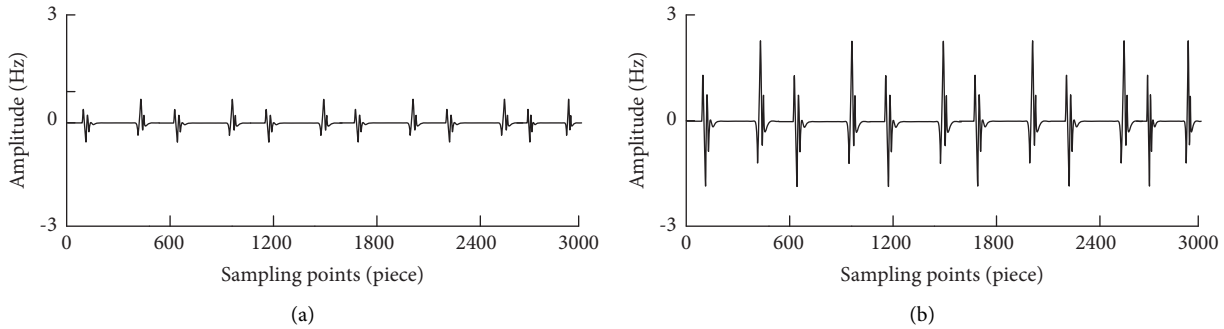


Figure 4: ECG waveforms of subjects in different states: (a) low pressure state; (b) high pressure state.

stress indicators is about 98%. Although the accuracy of method in reference [8] and method in reference [9] is within a reasonable range, it is always lower than this method, which verifies the effectiveness of this method. The reason for the high recognition accuracy of this method is that this method preprocesses with the psychological stress index of college students before constructing the model and uses the sequential backward selection algorithm to select the characteristics of psychological stress index, optimize the input data of the model, and further improve the recognition accuracy.

4.2.3. Identification Time. The optimal training parameters are as follows: the learning rate is 0.001, the number of iterations is 40, the inactivation rate is 0.5, and the vector dimension is 300. Based on the above parameter settings, the identification time of college students' psychological stress indicators identified by the methods in this paper, reference [8] and reference [9] is tested. The results are shown in Figure 6.

By analyzing the experimental data in Figure 6, it can be seen that with the change in the number of experiments, the time used to identify the psychological stress indicators of college students by this method is always less than 1.5 s, which is significantly lower than the methods in reference [8] and reference [9]. In contrast, the time-consuming of this method is relatively short, because the ECG signal is preprocessed before index recognition, which reduces the negative impact of interference signal and then reduces the recognition time of this method.

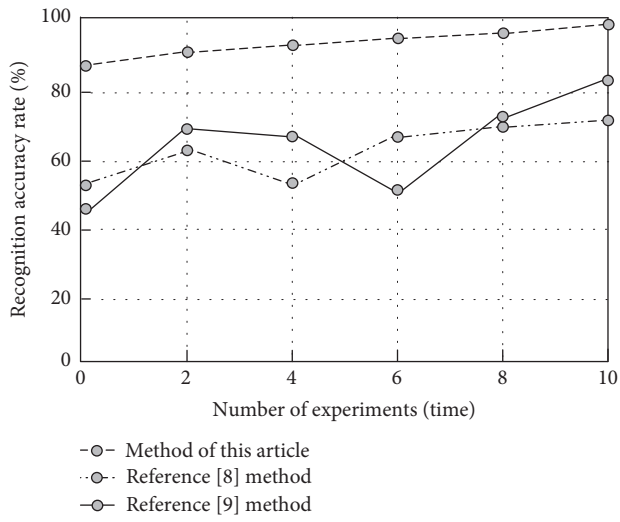


Figure 5: Comparison results of recognition accuracy of different methods.

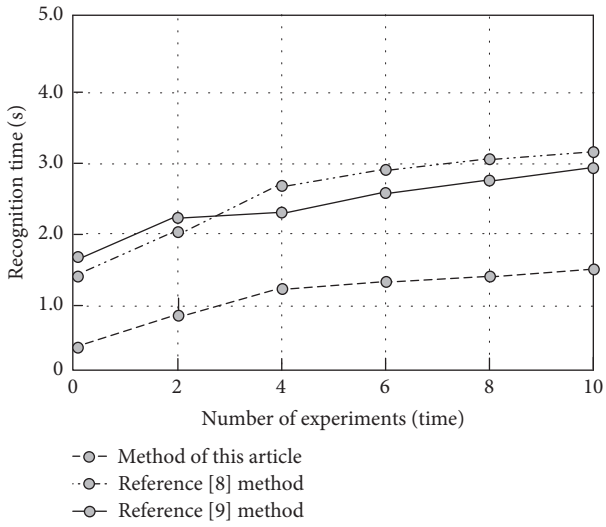


Figure 6: Comparison results of recognition time of different methods.

To sum up, the recognition result of this method has high accuracy, has high recognition efficiency, and is not easy to disturb, which shows that the recognition result of this method for college students' psychological stress index is stable and reliable.

5. Conclusion

Psychological index recognition and modeling, combined with the basic principles of psychological measurement and machine learning technology in computer field, will open a new door for psychological research. In order to improve the recognition effect and provide help for the diagnosis and treatment of college students' psychological problems, this paper proposes a deep learning-oriented identification and modeling method of college students' psychological stress indicators. Experiments show that this method has high recognition accuracy, has high recognition efficiency, and is not easy to be disturbed.

Although the experimental results show that the proposed method is feasible and its calculation results have been tested for reliability and validity, as a new method, it still has shortcomings. For example, massive user data can reach TB level. When processing and analyzing, the computing performance and storage performance of the computer will face higher requirements, which is also the focus of the next research.

Data Availability

The raw data supporting the conclusions of this article will be made available by the author, without undue reservation.

Conflicts of Interest

The author declares that there are no conflicts of interest regarding this work."

Acknowledgments

This work was funded by Henan Province Philosophy and Social Science Program Youth Project: "Developmental dyslexia group text and face cognitive neural mechanism research" (2018CJY035).

References

[1] N. Liu, H. Liu, and H. Liu, "Mental health diagnosis of college students based on facial recognition and neural network," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 3, pp. 1–12, 2020.

[2] J. Liu, G. Shi, J. Zhou, and Q. Yao, "Prediction of college students' psychological crisis based on data mining," *Mobile Information Systems*, vol. 2021, no. 23, 7 pages, Article ID 9979770, 2021.

[3] E. P. Ji and J. H. Kim, "Nursing students' experiences of psychological safety in simulation education: a qualitative study," *Nurse Education in Practice*, vol. 55, no. 1, Article ID 103163, 2021.

[4] S.-N. Ahn and J. Lee, "Effects of sensory awareness, imagery and observation on electroencephalography in adult with psychological stress," *Journal of Physical Therapy Science*, vol. 31, no. 1, pp. 17–19, 2019.

[5] E.-J. Kim, J. J. Kim, and S.-H. Han, "Understanding student acceptance of online learning systems in higher education: application of social psychology theories with consideration of user innovativeness," *Sustainability*, vol. 13, no. 2, p. 896, 2021.

[6] F. Wang, Y. S. Wang, J. T. Wang, H. Y. Xiong, J. F. Zhao, and D. Q. Zhang, "Mental stress assessment approach based on smartphone sensing data," *Journal of Computer Research and Development*, vol. 56, no. 3, pp. 611–622, 2019.

[7] H. F. Liu, X. Y. Wang, and Z. H. Gao, "The moderating effect of emotional regulation strategies in the relationship between emotional experiences and physical health of college students," *Chinese Journal of Behavioral Medicine and Brain Science*, vol. 28, no. 2, pp. 166–171, 2019.

[8] Y. Shang and N. Yang, "An improved particle swarm optimization back propogation neural network algorithm for psychological stress identification," *Science Technology and Engineering, Science Technology and Engineering*, vol. 20, no. 4, p. 6, 2020.

[9] Y. Su, M. M. Liu, N. Zhao, X. Q. Liu, and T. S. Zhu, "Identifying psychological indexes based on social media data: a machine learning method," *Advances in Psychological Science*, vol. 29, no. 4, p. 15, 2021.

[10] S. Thakran, "A hybrid GPFA-EEMD_Fuzzy threshold method for ECG signal de-noising," *Journal of Intelligent and Fuzzy Systems*, vol. 39, no. 1, pp. 1–10, 2020.

[11] K. Kido, T. Tamura, N. Ono et al., "A novel CNN-based framework for classification of signal quality and sleep position from a capacitive ECG measurement," *Sensors*, vol. 19, no. 7, p. 1731, 2019.

[12] M. Pelc, Y. Khoma, and V. Khoma, "ECG signal as robust and reliable biometric marker: datasets and algorithms comparison," *Sensors*, vol. 19, no. 10, p. 2350, 2019.

[13] D. Zhang, S. Wang, F. Li et al., "An efficient ECG denoising method based on empirical mode decomposition, sample entropy, and improved threshold function," *Wireless Communications and Mobile Computing*, vol. 2020, no. 2, pp. 1–11, 2020.

[14] S. Taran and V. Bajaj, "Motor imagery tasks-based EEG signals classification using tunable-Q wavelet transform," *Neural Computing & Applications*, vol. 31, no. 11, pp. 6925–6932, 2019.

[15] J.-J. Liaw, C.-P. Lu, Y.-F. Huang, Y.-H. Liao, and S.-C. Huang, "Improving census transform by high-pass with haar wavelet transform and edge detection," *Sensors*, vol. 20, no. 9, p. 2537, 2020.

[16] X. Dong, G. Li, Y. Jia, and K. Xu, "Multiscale feature extraction from the perspective of graph for hob fault diagnosis using spectral graph wavelet transform combined with improved random forest," *Measurement*, vol. 176, no. 2, Article ID 109178, 2021.

[17] C. Zhang, H. Wang, J. Zeng, L. Ma, and L. Guan, "Short-term dynamic radar quantitative precipitation estimation based on wavelet transform and support vector machine," *Journal of Meteorological Research*, vol. 34, no. 2, pp. 413–426, 2020.

[18] A. Rezapour, A. Ortega, and M. Sahimi, "Upscaling of geological models of oil reservoirs with unstructured grids using lifting-based graph wavelet transforms," *Transport in Porous Media*, vol. 127, no. 3, pp. 661–684, 2019.

[19] A. A. Ewees, M. Aziz, and A. E. Hassanien, "Chaotic multiverse optimizer-based feature selection," *Neural Computing & Applications*, vol. 31, no. 1, pp. 1–16, 2019.

[20] M. Waheed, H. Afzal, and K. Mehmood, "NT-FDS-A noise tolerant fall detection system using deep learning on wearable devices," *Sensors*, vol. 21, no. 6, p. 2006, 2021.

[21] J.-M. Kim, J. Bae, S. Son, K. Son, and S.-G. Yum, "Development of model to predict natural disaster-induced financial losses for construction projects using deep learning techniques," *Sustainability*, vol. 13, no. 9, p. 5304, 2021.

[22] X. Guan, Y. Fan, Q. Qin, K. Deng, and G. Yang, "Construction of science and technology achievement transfer and transformation platform based on deep learning and data mining technology," *Journal of Intelligent and Fuzzy Systems*, vol. 39, no. 1, pp. 1–12, 2020.

[23] R. Mohammad and M. K. Alsmadi, "Intrusion detection using Highest Wins feature selection algorithm," *Neural Computing and Applications*, vol. 33, no. 2, pp. 1–12, 2021.

[24] B. Chen, H. Chen, and M. Li, "Improvement and optimization of feature selection algorithm in swarm intelligence algorithm based on complexity," *Complexity*, vol. 2021, no. 7, 10 pages, Article ID 9985185, 2021.

[25] Y.-T. Cheng, A. Patel, C. Wen, D. Bullock, and A. Habib, "Intensity thresholding and deep learning based lane marking extraction and lane width estimation from mobile light detection and ranging (LiDAR) point clouds," *Remote Sensing*, vol. 12, no. 9, p. 1379, 2020.

[26] J. Zhang, T. Xie, C. Yang et al., "Segmenting purple rapeseed leaves in the field from UAV RGB imagery using deep learning as an auxiliary means for nitrogen stress detection," *Remote Sensing*, vol. 12, no. 9, p. 1403, 2020.

[27] A. Nasif, Z. A. Othman, and N. S. Sani, "The deep learning solutions on lossless compression methods for alleviating data load on IoT nodes in smart cities," *Sensors*, vol. 21, no. 12, p. 4223, 2021.

[28] R. Theagarajan, B. Bhanu, T. Erpek, Y. K. Hue, and Y. E. Sagduyu, "Integrating deep learning-based data driven and model-based approaches for inverse synthetic aperture radar target recognition," *Optical Engineering*, vol. 59, no. 5, p. 1, 2020.

[29] Y. Long and C. Han, "Transaction processing and value evaluation of carbon emission rights based on wavelet transform image and deep learning," *Journal of Intelligent and Fuzzy Systems*, vol. 39, no. 4, pp. 5821–5832, 2020.

[30] M. Kim, S. H. Choi, K.-B. Park, and J. Y. Lee, "A hybrid approach to industrial augmented reality using deep learning-based facility segmentation and depth prediction," *Sensors*, vol. 21, no. 1, p. 307, 2021.

[31] A. Kashefi, D. Rempe, and L. J. Guibas, "A point-cloud deep learning framework for prediction of fluid flow fields on irregular geometries," *Physics of Fluids*, vol. 33, no. 2, Article ID 027104, 2021.

[32] Y. Qing, N. Geng, and Y. H. Zhu, "Patient examination demand forecasting based on lasso and tabu search[J]," *Computer Simulation*, vol. 38, no. 1, pp. 387–393, 2021.

[33] R. Banupriya and D. Kannan, "A convolutional neural network based feature extractor with discriminant feature score for effective medical image classification," *NeuroQuantology*, vol. 18, no. 7, pp. 1–8, 2020.

[34] Z. Huang, J. Zhu, J. Lei, X. Li, and F. Tian, "Tool wear monitoring with vibration signals based on short-time fourier transform and deep convolutional neural network in milling," *Mathematical Problems in Engineering*, vol. 2021, no. 6, 14 pages, Article ID 9976939, 2021.

[35] J. Wang, W. Zhang, and J. Zhou, "Fault detection with data imbalance conditions based on the improved bilayer convolutional neural network," *Industrial & Engineering Chemistry Research*, vol. 59, no. 13, pp. 5891–5904, 2020.

[36] S. Arvidsson, M. Gullstrand, B. Sirmacek, and M. Riveiro, "Sensor fusion and convolutional neural networks for indoor occupancy prediction using multiple low-cost low-resolution heat sensor data," *Sensors*, vol. 21, no. 4, p. 1036, 2021.

Hindawi
Scientific Programming
Volume 2022, Article ID 3101614, 11 pages
https://doi.org/10.1155/2022/3101614

Research Article

Application of Data Mining in Effect Evaluation of Lean Management

Song Ding,[1] Jun Li,[2] and Jiye Li[3]

[1]*School of Economics and Management, Chengdu Technology University, Chengdu, Sichuan 611730, China*
[2]*Office of Teaching Construction and Quality Control in Chengdu Technological University, Chengdu, Sichuan 611730, China*
[3]*School of Computer Engineering in Chengdu Technology University, Chengdu, Sichuan 611730, China*

Correspondence should be addressed to Jiye Li; ljyybf123@cdtu.edu.cn

Received 9 November 2021; Revised 22 November 2021; Accepted 21 December 2021; Published 10 January 2022

Academic Editor: Baiyuan Ding

Quantitative evaluation is an important part of enterprise diagnosis, which promotes the scientific and modern management of enterprises. At present, the existing enterprise management evaluation methods cannot complete the mining of enterprise index data, which leads to large error and low significance coefficient in enterprise management evaluation. Therefore, the application of data mining in enterprise lean management effect evaluation is put forward. The process and main functions of data mining are analyzed; data mining algorithm is used to establish the evaluation index system of lean management effect and calculate the index weight. Using the association rules method in data mining, according to the parameters of enterprise lean management level evaluation index and weight value, through the fuzzy set transformation idea, the fuzzy boundary of each index and factor is described by the membership degree, the fuzzy judgment matrix is constructed, and the final evaluation result is obtained by multilayer compound calculation. Experimental results show that this study has a high significance coefficient, and the proposed evaluation method of enterprise lean management effect has ideal accuracy and short time consumption. In practical application, the cumulative contribution rate is higher and has higher stability.

1. Introduction

With the continuous improvement of the national economy, a large number of new types of enterprises have emerged. There are still some problems in the management of enterprises, such as improper management mode and imperfect related systems. Therefore, it is necessary to adopt lean management mode, innovate enterprise management mode and technology continuously to ensure the increasing economic benefits of enterprises, and highlight the pursuit of high quality and high efficiency. Lean management requires lean thinking in all aspects of business. The core is to create a large amount of value with the minimum investment of resources and to provide users with new products and timely services [1]. Lean management not only strengthens the informationization control to the manufacture and the production link of the company's products, but also pays more attention to the overall operation and the optimization of the operation link, involving all the management details of the company and the production details of the products, and all the project flows of these enterprises are related to the cost. The focus of lean management is to optimize the product processing and manufacturing, daily operation and management, material management and storage, technological innovation and development, and other aspects during the stable operation stage of the company [2].

Lean management is a new type of company management and operation, concepts and methods, and scientific and technological tools, such as a unified scientific system. In this way, value stream is the focus of management, and striving for perfection is an important content of management pursuit. We shall build up a "people-oriented" corporate culture, adopt the purpose of systematic thinking, and fully absorb and adopt a variety of advanced methods and scientific and technological methods under the guidance of the broad sense of lean philosophy in the strategic technical

level and each subsystem of the company. Through business process reengineering, innovate the value stream, fully adapt to users and meet user requirements, maximize the company's value, and achieve win-win with relevant partners. Create a learning organization, implement lean strategy, establish lean operation system, develop economic and financial software, understand lean technology, strengthen lean cost control, create a kind of lean management style and harmonious company environment, and continuously optimize.

Reference [3] explores the choice of cooperation strategies among participants in new R&D institutions by using evolutionary game model under the premise of bounded rationality hypothesis. By analyzing the multiagent income payment matrix and solving the replication dynamic equation and Jacobian matrix, the local equilibrium point of the model is obtained so as to obtain the final strategy choice of all participants over time, verify the strategy evolution trend of the game system through numerical simulation, and sort out the influencing factors of the strategy choice of new R&D institutions under cooperative innovation. On this basis, this paper puts forward strategic suggestions to realize the sustainable development of new R&D institutions from the perspectives of cooperation mode, enterprise demand, and policy environment. This method completes the selection of cooperation strategy among participants, but it is unable to complete the index data mining of R&D institutions. In [4], the credit risk evaluation model of small enterprises based on ELECTRE III is studied. Firstly, based on the evaluation principle of ELECTRE III, the credit risk evaluation score of new loan customers is calculated according to the net credibility of new loan customers better than all historical loan customers. It not only solves the problem of credit risk evaluation of new loan customers, but also ensures that the evaluation model has the ability to learn from historical data. Secondly, referring to the characteristics that Theil's index can not only reflect the overall income difference, but also decompose the overall difference into intragroup difference and intergroup difference; the evaluation indexes of small enterprise credit risk are weighted. It embodies the weighting idea of "the more you can distinguish the default status of customers, the greater the weight of indicators." Finally, based on the weighted intragroup difference between default and nondefault samples, the preference threshold of ELECTRE III is determined. Determine the indifference threshold of ELECTRE III based on the degree of difference within the nondefault group. Based on the difference degree within the default sample group, the rejection threshold of ELECTRE III is determined. This method not only reflects the influence of the data difference of different evaluation indexes on the evaluation results, but also avoids the deficiency of artificial and subjective determination of the existing threshold. However, it is easy to lead to large error in the evaluation of enterprise management effect. A nonrelational distributed big data mining algorithm is proposed in [5]. Firstly, the distributed maximum frequent item algorithm model is introduced. Under the physical decentralized logic, the distributed method is applied to the nonrelational data for data analysis, the search conditions of candidate frequent items are established to reduce the number of data detected, and the partition projection method is adopted to calculate the partition number corresponding to each frequent item and plan it into different partitions so as to optimize the redundant frequent items. Then, according to the data characteristics of big data in the database, the correlation degree of each attribute is established. According to the boosting clustering method, the local model obtained by weak clustering is upgraded to the global model in each iteration, and then it is divided into blocks. Using the partition quality to readjust the iterative sampling rate, the final clustering results are obtained and the data mining is completed. It has a good effect, but the significance coefficient of this method is low.

However, the above reference methods cannot complete the mining of enterprise index data, resulting in large error and low significance coefficient in the evaluation of enterprise management effect. Therefore, this paper puts forward the application research of data mining in enterprise lean management effect evaluation.

2. Enterprise Lean Management Effect Evaluation Based on Data Mining

2.1. Data Mining Process and Main Functions. Data mining is a kind of technology that is mainly used by different subjects, and it is of great importance in every field. Data mining refers to the process of extracting useful knowledge from large, incomplete, noisy, and vague data [6, 7]. It comes from knowledge extraction in database, namely, KDD. People use KDD to depict the whole process of data mining and use data mining to describe the basic process of mining algorithm to achieve data mining [8]. The process of discovering valuable knowledge in the database, known as the KDD process, is shown in Figure 1.

According to Figure 1, KDD is composed of data collection and processing, data information mining and result analysis, etc. Data mining is mainly based on association analysis, cluster analysis, and prediction to find useful knowledge in large-scale data. At the same time, through model evaluation, valuable models are used as knowledge to assist related personnel to make scientific and rational decisions.

The main function of data mining is to make decisions based on valuable knowledge by predicting the future development trend. The main function of data mining technology is not realized by a single way, but by a group of methods. The main task of data mining is to find valuable knowledge or information from relevant databases. The main functions of data mining are as follows [9, 10]:

(1) *Correlation Analysis of Data*

In a database, the analysis of the correlation between data is the key link to discovering important knowledge. If there is some special law between two or more variable values, it can be called an association.

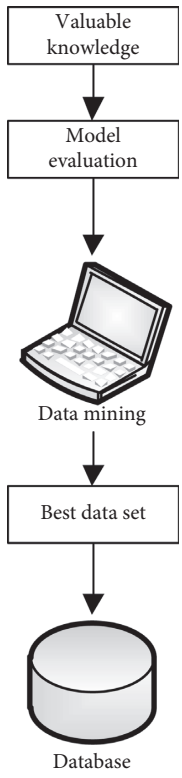


Figure 1: Overall flow of KDD.

(2) *Cluster Analysis between Data*

In a database, the data information records can be divided into different meaningful data subsets; the process is data clustering. Among them, the smaller the distance between the data in the same data class, the more similar they are, and dissimilar they are. Data clustering can enhance people's cognition of objective facts, and it is the precondition of concept representation and deviation analysis. At present, there are several data clustering algorithms in common use: mean clustering, density clustering, fuzzy clustering, and so on [11].

(3) *Forecast Analysis*

In some instances, the public may have to predict some vacant values. If the data for the vacant value is a numeric type, it is generally called a forecast. The process of forecasting is as follows: searching for the laws between the data and information according to the historical data, constructing a corresponding model, and forecasting the following data based on the model [12].

(4) *Data Evolution Analysis*

In a database, the evolution analysis of data information indicates that there are laws and trends in the target of a certain behavior changing with time, and at the same time, the trend is modeled.

2.2. Evaluation Index System of Enterprise Lean Management Effect. The enterprise management effect evaluation index system can measure the intensity of the impact of scientific and technological production factors on industrial transformation and development. According to the characteristics of industrial upgrading and scientific and technological production, and on the basis of the basic principles of index system construction, the enterprise management effect evaluation index is selected from factor flow and factor composition [13]. The content of Constructing Indicators mainly includes input dimension, status dimension, and evaluation dimension. Among them, human input, financial input, and material input constitute the input dimension. Personnel activities, fund use, implementation form, R&D output, and achievement transformation constitute the state dimension. The upgrading of industrial structure, the rationalization of industrial structure, the dual orientation of industrial development, and the transformation of industrial development constitute the evaluation dimensions.

The basis for establishing the enterprise operation effect evaluation index system shall follow the following basic principles:

(1) Scientific principle: this is the basic principle that should be followed in the setting of enterprise operation effect evaluation index system. According to this principle, the setting of indicators and indicator system should be consistent with the idea of circular economy and the concept of enterprise operation effect evaluation. Therefore, the concept of indicators should be accurate, the connotation and extension should be clear, and the calculation method should be scientific and feasible.

(2) The principle of combining qualitative measurement with quantitative measurement: this principle is a higher-level requirement for the setting of index system. Because only quantitative indicators can be used to evaluate the enterprise's operation effect from one level, the evaluation results are not comprehensive and one-sided, and the enterprise's operation effect cannot be truly evaluated. If qualitative indicators are set on the basis of quantitative indicators, from a higher level, by investigating various measurement factors that have a direct impact on the operation effect of enterprises but are difficult to be uniformly quantified, the combination of quantitative analysis and qualitative analysis can avoid the defects caused by relying solely on quantitative methods.

(3) Operability principle: there must be a clear calculation method and expression method when setting indicators so that all indicators can be easily calculated for easy operation.

(4) Comparability principle: when setting indicators, they should maintain relative stability in meaning, scope, and methods within a certain period of time, so as to facilitate the comparability of evaluation results, so as to determine the level and position of enterprises in the same industry, the same scale and in the national economy, so as to find out the gap and tap the potential.

The constructed industrial effect evaluation index system is shown in Figure 2.

It can be seen from Figure 2 that the enterprise operation effect evaluation itself is a complex system engineering, which cannot be simply determined by a certain index. After introducing the strategic idea of lean management into the process of enterprise management, the goal of enterprise business activities has shifted from simply pursuing economic benefits to enterprises pursuing both their own economic benefits and environmental benefits. Therefore, the external performance of enterprise management effect evaluation oriented to data mining should be multidimensional, and its goal should be diversified. Therefore, the industrial effect oriented evaluation index system is a multilevel, multifactor, and multiobjective comprehensive evaluation.

Suppose $X = [x(1), x(2) \ldots x(n)]$ describes the vector of national or regional industrial structure, and $X_0 = [x_0(1), x_0(2) \ldots x_0(n)]$ describes the vector of optimized objective structure of enterprise management at a certain level of economic development by means of the following primitive point zeroing:

$$\begin{aligned} x^0(k) &= x(k) - x(1), \\ x_0^0(k) &= x_0^0(k) - x_0^0(1). \end{aligned} \tag{1}$$

The following can be obtained through formula (1):

$$\begin{aligned} X^0 &= \left[x^0(1), x^0(2), \ldots, x^0(n)\right], \\ X_0^0 &= \left[x_0^0(1), x_0^0(2), \ldots, x_0^0(n)\right]. \end{aligned} \tag{2}$$

Let ς represent the degree of industrial order, which can be calculated by the following formula:

$$\varsigma = \frac{1 + |s_0| + |s_1|}{1 + |s_0| + |s_1| + |s_0 - s_1|}. \tag{3}$$

Here, ς describes the proximity between the proportion of the target industry structure and the current proportion of the industrial structure. The closer the two proportions are, the smaller the value of $|s_0 - s_1|$ is, and the closer the degree of industrial order ς is to 1; s_0 represents the level of nullification of the target industry:

$$|s_0| = |x_0^0(2)| + \frac{1}{2}|x_0^0(3)|. \tag{4}$$

s_1 represents the level of nullification corresponding to the actual industrial structure of the region, with the following expression:

$$|s_1| = |x^0(2)| + \frac{1}{2}|x^0(3)|. \tag{5}$$

The relationship between s_0 and s_1 can be described by the following formula:

$$|s_0 - s_1| = |x^0(2) - x_0^0(2)| + \frac{1}{2}|x^0(3) - x_0^0(3)|. \tag{6}$$

The specific operation steps of data mining process are as follows:

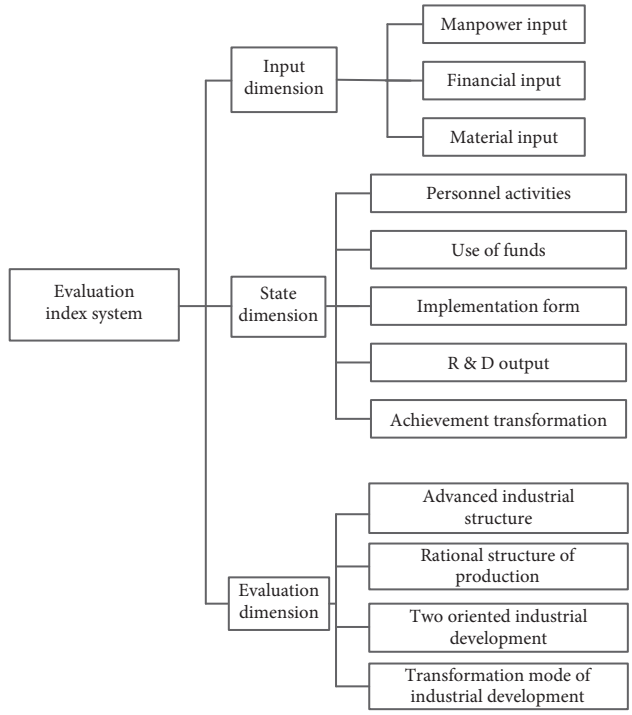


Figure 2: Industrial effect evaluation index system.

Step 1. Set the initial weight vector as random acquisition.

Step 2. Update the weight vector through formula (6).

Step 3. Normalization.

Step 4. If the target industrial structure proportion is close to the current industrial structure proportion, the data mining stops convergence. If it is not achieved, jump back to Step 2 and continue to execute until an independent component is estimated.

If you want to realize multicomponent extraction, you can repeatedly execute the data mining algorithm for separation. At the same time, if you need to judge that the components extracted each time are not similar components, you need to complete one component extraction at a time, remove the component from the observation signal, and repeat the elimination operation; that is, you can extract all the required independent components. After execution, it can realize the denoising of the original spatial attribute data.

2.2.1. Single Indicator Evaluation. There are different positive and negative trends and different units of measurement between different index values. The index value of enterprise management effect evaluation is nondimensionalized and trended by single index evaluation [14, 15].

The following formula shall be used to make the evaluation indicators of negative enterprise management effect co-trend:

$$y_{ij} = \max_{1 \le i \le m} \{x_{ij}\} - x_{ij}. \quad (7)$$

In the formula, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$; y_{ij} is the value after the j industrial effect evaluation index of the i evaluated object in the same trend processing; x_{ij} describes the statistical value corresponding to the j industrial effect evaluation index of the i evaluated object.

Dimensionless average treatment shall be carried out on the evaluation indexes of management effect of a single enterprise through the following formula [16]:

$$Y_{ij} = \frac{y_{ij}}{\overline{y}_j}. \quad (8)$$

Here, Y_{ij} describes the value obtained without quantization processing and $\overline{y}_j$ describes the average value of the j industrial effect evaluation index.

2.2.2. Index Weight. The research method of enterprise management effect evaluation calculates the weight of the enterprise management effect evaluation index through coefficient of variation method [17].

The coefficient of variation method assigns the value of each industrial effect evaluation index through the variation degree of the observed value of each industrial effect evaluation index on the evaluated object [18]. Therefore, the difference of each evaluation object can be reflected by the index of large variation degree of observation value.

Set x_{ij} to represent the evaluation value of item j industrial effect evaluation index of item i; V_j to represent the coefficient of variation corresponding to the industrial effect evaluation index of Item j, and the calculation formula is as follows:

$$V_j = \frac{S_j}{M_j} \times 100\%. \quad (9)$$

Here, M_j represents the average value of the indices for evaluating the industrial effect of item j of the evaluated object i; S_j represents the standard deviation, which is calculated by the following formulas:

$$M_j = \frac{1}{m} \times \sum_{i=1}^{m} x_{ij},$$
$$S_j = \sqrt{\sum_{i=1}^{m} (x_{ij} - M_j) \times \frac{1}{m-1}}. \quad (10)$$

The coefficient of variation V_j of the industrial effect evaluation index of normalized treatment shall be obtained, and the weight w_j of the industrial effect evaluation index shall be obtained:

$$w_j = \frac{V_j}{\sum_{i=1}^{n} V_j}. \quad (11)$$

Only long-term implementation and continuous improvement of enterprise lean management can achieve the ideal final result. Therefore, the implementation of lean management is a long-term process, which needs to go through different implementation extremes. It is also a systematic project, which needs to take into account different aspects such as personnel, technology, environment, and culture. It has established a lean management evaluation index system that takes into account results and processes.

2.3. Quality Evaluation of Enterprise Lean Management Based on Data Mining. From the above analysis, we can see that association rules are one of the most significant methods in data mining. Association rules are widely used in enterprise lean management quality management. There are a lot of matters involved in enterprise lean management quality management, such as courses and evaluation index, etc. Association rules can extract valuable knowledge hidden in data, assist educational administrators to make effective decisions, and improve the management level [19–21].

According to the determination of such parameters as the evaluation index and weight of lean management level of enterprises, the fuzzy boundary of each index and factor is described by the membership degree through the idea of fuzzy set transformation, and the fuzzy judgment matrix is constructed, and the final evaluation result is obtained by multilayer compound calculation [22].

There are n evaluation grades and m Grade I evaluation indexes. Each Grade I index contains several Grade II indexes, which are represented by Grade U, Grade I index scope V , and Grade II index scope V_i, and are as follows:

$$U = \{u_1, \ldots, u_n\},$$
$$V = \{v_1, \ldots, v_n\}, \quad (12)$$
$$V_i = \{v_1, \ldots, v_k\}.$$

The fuzzy matrix can be expressed as follows [23]:

$$U = \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_m \end{bmatrix} \begin{bmatrix} U_{11} & u_{12} & \cdots & u_{1n} \\ U_{21} & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ U_{m1} & u_{m2} & \cdots & u_{mn} \end{bmatrix} = \left[u_{ij}\right]_{mn}. \quad (13)$$

Here, u_{ij} represents the degree of membership of the evaluation index i to Grade j [24–26].

Based on the above calculation and analysis, the fuzzy relation matrix is obtained, and the weight value of the first-grade evaluation index can be defined as follows:

$$\mathfrak{R} = (\mathfrak{R}_1, \mathfrak{R}_2, \ldots, \mathfrak{R}_m). \quad (14)$$

The Grade I evaluation index is divided into very good (A), good (B), better (C), poor (d), and very poor (E). The evaluation index grade standard is shown in Table 1.

Combined with the above calculation and Table 1, the enterprise lean management quality evaluation model can be expressed as follows [27, 28]:

$$W = \mathfrak{R} \times C \times U. \quad (15)$$

Table 1: Evaluation index grade standard.

Evaluation grade	Critical value (%)	Weight (%)
A	100	31.7
B	80	29.2
C	60	26.2
D	40	18.5
E	20	14.3

The fuzzy comprehensive evaluation and data mining methods are used to construct the enterprise lean management quality evaluation model, which fully considers the fault-tolerance of the evaluation process and other aspects, and can effectively improve the modeling freedom and reduce the modeling complexity [29–31]. From the above evaluation results of the lean management effect of the enterprise, it can be seen that the implementation of data mining technology in the enterprise has improved the overall quality of the enterprise, improved the enterprise management system, and achieved obvious economic and environmental benefits. Lean management is the mainstream economic model for the development of the new era. If enterprises want to survive, develop, and make profits in the new century, they must infiltrate the idea of lean management into every corner of the enterprise. In order to evaluate the effect of lean management oriented to data mining, it is necessary to establish an enterprise operation evaluation index system oriented to lean management. In practical application, enterprises can increase or decrease the evaluation indicators according to the specific situation so as to better evaluate the effect of lean management.

3. Experimental Results and Analysis

In order to further study the impact of internal control, customer relationship maintenance, and financial management on the effect of lean management innovation, this paper constructs a method of lean management effect evaluation.

3.1. Selection of Variables and Establishment of the Model. The explanatory variable in the regression model [32] is enterprise lean management effect MAP. In order to simplify the calculation of the model, this paper substitutes variable EPS per share for enterprise lean management effect [33]. The explanatory variables are internal control effect $ICE34$, customer relationship maintenance effect CRM [35], and financial management level FML [36], and the relevant control variables are selected and designed as shown in Table 2.

After the control variables of the regression model are determined, the unknown parameters of the model are estimated by using the sample data, and the multiple regression linear regression model is designed as follows:

$$\begin{aligned} \mathrm{EPS} = {} & e + \gamma_0 + \gamma_1 \mathrm{ICE} + \gamma_2 \mathrm{CRM} + \gamma_3 \mathrm{FML} + \gamma_4 \mathrm{ESIZE} \\ & + \gamma_5 \mathrm{BSIZE} + \gamma_6 \mathrm{DR} + \gamma_7 \mathrm{EC}. \end{aligned} \tag{16}$$

Here, γ_0 is the model constant, $\gamma_1-\gamma_3$ are the explanatory variable coefficient, $\gamma_4-\gamma_7$ are the control variable coefficient, and e is the residual term of the model. Whether the model really reveals the relationship between the explained variables and the explanatory variables requires the establishment of multiple linear regression equations based on the actual observation data of dependent variables and multiple independent variables, and the linear relationship between the dependent variable and multiple independent variables is only a hypothesis. Although this hypothesis is often not groundless, after establishing the multiple regression linear regression model, it is also necessary to test the significance of the hypothesis of the linear relationship between dependent variables and multiple independent variables, that is, to test the significance of multiple linear regression relationship or to test the significance of multiple linear regression model.

3.2. Selection of Sample Data and Descriptive Statistical Analysis of Variables. From the sample data of lean management behaviors of enterprises in a city from 2017 to 2020, 75 enterprises and 86 enterprises are selected, respectively, and the time node for the selected enterprises to conduct lean management behaviors shall be June 2018. ST enterprises with abnormal operating performance shall be excluded, samples of abnormal financial data or financial reports shall be excluded, enterprises of less than three years of listed companies shall be excluded, financial and insurance listed enterprises with obvious industry specificity shall be excluded, and the last 123 research samples shall be excluded. The selected sample enterprises have assets of more than 2 billion people, among which there are 36 enterprises with assets of 2–5 billion and 87 enterprises with assets of more than 5 billion. From the scale of the sample, we can see that the asset strength of the enterprise is strong, and it has the precondition of lean management, and it is helpful to control the financial risk.

Based on the SPSS 24.1 statistical analysis software, the descriptive statistical analysis of each variable is performed as shown in Table 3.

From 2017 to 2020, the change trend of the EPS-means of the variables can be seen, from −0.0357 to 0.5499, which shows that the EPS index after M&A shows an increasing trend, and the level of business risks and financial risks of enterprises is also decreasing. Then, analyze the correlations among the explanatory variables, control variables, and explained variables B of the enterprise before and after the merger, and the statistical analysis results are as shown in Tables 3–6. The symbols *, **, and *** in Tables 4–7 indicate statistical significance in the range of 10%, 5%, and 1%.

From the correlation data of each explanatory variable and EPS of the explanatory variable, it can be seen that there is a negative correlation between ICE, CRM, and FML before M&A in 2014 and 2015 and the proportion of the explanatory variable, while in 2016 and 2017 after the enterprise implements lean management, the correlation coefficient between the explanatory variables and the coefficient of the explanatory variable is high, which

Table 2: Selection of control variables of the regression model.

Control variable	Code	Relevant description
Enterprise scale	*ESIZE*	Natural logarithm of asset scale in the year before lean management
Lean management scale	*BSIZE*	Lean management transaction volume/total assets at the end of the year
Debt ratio	*DR*	Total liabilities/total assets of the enterprise
Equity concentration	*EC*	Shareholding ratio of the top 5 shareholders of the enterprise

Table 3: Descriptive statistical analysis results of variables.

Variable		Number of samples	min	max	$\overline{\mu}$	σ
EPS	2017	200	−0.2534	0.4552	−0.0357	0.0035
	2018	200	−0.5524	0.5254	0.0450	0.0240
	2019	200	0.5324	0.3524	0.5979	0.2558
	2020	200	0.2357	5.4255	0.5499	0.5329
ICE		200	500	3.3037	2.2548	0.3358
CRM		200	500	0.3425	0.5254	0.0005
FML		200	500	0.3984	0.5949	0.5205
ESIZE		200	500	0.8524	0.3255	0.0047
BSIZE		200	500	5.2545	0.8748	0.5534
DR		200	500	52.2545	4.8455	5.3325
EC		200	500	3.3254	4.2544	0.0354

Table 4: Correlation statistics between model variables in 2017.

Variable	*EPS*	*ICE*	*CRM*	*FML*	*ESIZE*	*BSIZE*	*DR*	*EC*
EPS	1	—	—	—	—	—	—	—
ICE	−0.231*	1	—	—	—	—	—	—
CRM	−0.235**	0.025	1	—	—	—	—	—
FML	0.057	−0.359	−0.811	1	—	—	—	—
ESIZE	0.015**	0.256*	0.369	0.125**	1	—	—	—
BSIZE	−0.0415	0.287**	0.448***	0.484*	0.015*	1	—	—
DR	0.027	0.590***	0.458*	0.744	−0.999	0.362	1	—
EC	−0.371	0.581*	0.012*	0.559***	0.072	0.005	0.168	1

Table 5: Correlation statistics between model variables.

Variable	*EPS*	*ICE*	*CRM*	*FML*	*ESIZE*	*BSIZE*	*DR*	*EC*
EPS	1	—	—	—	—	—	—	—
ICE	−0.066*	1	—	—	—	—	—	—
CRM	−0.128**	0.036*	1	—	—	—	—	—
FML	−0.005***	−0.093*	0.285	1	—	—	—	—
ESIZE	0.149*	0.245*	−0.283**	0.356**	1	—	—	—
BSIZE	0.018*	0.858**	0.448***	0.484*	0.015*	1	—	—
DR	0.158*	0.699**	0.414*	0.778**	−0.115**	0.758**	1	—
EC	−0.121	0.238*	0.021**	0.550**	0.654*	0.152**	0.256*	1

Table 6: Correlation statistics between model variables in 2019.

Variable	*EPS*	*ICE*	*CRM*	*FML*	*ESIZE*	*BSIZE*	*DR*	*EC*
EPS	1	—	—	—	—	—	—	—
ICE	0.282**	1	—	—	—	—	—	—
CRM	0.369**	0.079*	1	—	—	—	—	—
FML	0.269*	0.331*	0.689***	1	—	—	—	—
ESIZE	0.545**	0.667*	0.969***	1.365**	1	—	—	—
BSIZE	0.118*	0.864*	0.997**	3.497	19.626*	1	—	—
DR	0.988*	2.554*	1.215**	0.826**	−0.001**	0.758**	1	—
EC	1.128	−0.784*	0.007**	0.598**	0.025**	0.152*	0.148**	1

Table 7: Correlation statistics between model variables in 2020.

Variable	*EPS*	*ICE*	*CRM*	*FML*	*ESIZE*	*BSIZE*	*DR*	*EC*
EPS	1	—	—	—	—	—	—	—
ICE	2.655*	1	—	—	—	—	—	—
CRM	0.885***	0.900*	1	—	—	—	—	—
FML	2.268***	2.398*	0.789**	1	—	—	—	—
ESIZE	2.669*	2.362*	1.699***	1.380***	1	—	—	—
BSIZE	9.287**	2.202*	6.562*	5.569*	19.596**	1	—	—
DR	3.336*	5.565	7.254**	1.112	0.888*	7.784*	1	—
EC	5.542**	21.362	2.112*	1.251*	0.283*	2.834**	4.189*	1

indicates that it is scientific and reasonable to evaluate the performance level of M&A in an all-round way from a multidimensional perspective. From the analysis of the data changes between the 4 control variables and the explanatory variable *EPS*, it also shows an overall trend of improvement.

3.3. Regression Analysis. The multivariate linear regression data analysis between variables before and after M&A of sample enterprises from 2017 to 2020 is shown in Table 8.

To sum up, there is a high correlation between the changes of multidimensional evaluation index factors selected in this paper and enterprise lean management performance. It is more accurate and reliable to comprehensively evaluate the M&A effect of enterprises based on multiple dimensions.

3.4. Comparison of Significance Coefficients of Different Methods. The enterprise lean management effect evaluation research method (method 1), enterprise management evaluation method based on evolutionary game model (method 2), and enterprise management evaluation method based on the third-party recycling model (method 3) proposed in this study are used to test respectively. The significance coefficients of indicators selected by different methods are compared. The test results are as shown in Figure 3.

According to the analysis of Figure 3, in multiple iterations, the corresponding significance coefficient of the enterprise management effect evaluation index selected by method 1 is more than 0.8, the corresponding significance coefficient of the industrial performance evaluation index selected by method 2 is low to 0.4 in the fourth iteration, and the corresponding significance coefficient of the industrial performance evaluation index selected by method 3 is low to 0.5 in the third iteration. Comparing the test results of method 1, method 2, and method 3, it can be seen that the significance coefficient obtained by method 1 is high because this method selects the enterprise management effect evaluation index from the element flow and element composition on the basis of the basic principles of index system construction according to the characteristics of industrial upgrading and scientific and technological production, which improves the rationality of the enterprise management effect evaluation index.

Table 8: Multiple linear regression analysis between variables.

Variable	Explained variable *EPS*			
	2017	2018	2019	2020
ICE	−0.005**	0.023*	0.125	0.635***
CRM	−0.276	−0.005**	0.236	0.551*
FML	0.005*	0.202	0.287**	0.744
ESIZE	−0.178	0.066*	0.145	0.222*
BSIZE	−0.883**	−0.331	0.156***	1.451
DR	−1.289	1.223	2.253	3.692
EC	0.980*	0.775	0.985	1.112
T	200	200	200	200

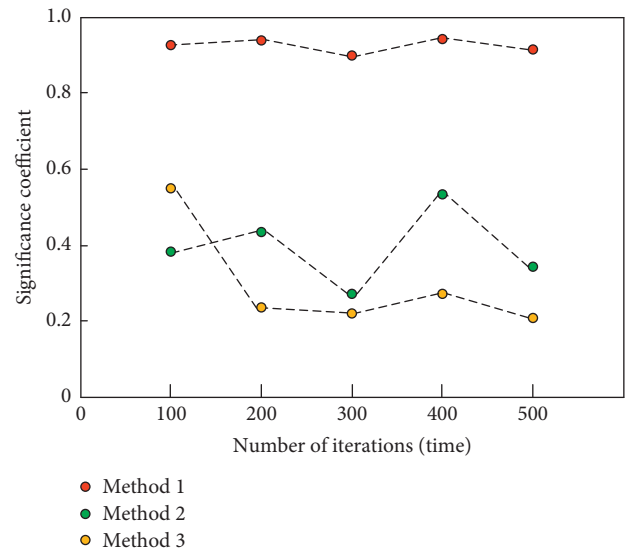


Figure 3: Significance coefficients of different methods.

3.5. Comparison of Evaluation Accuracy of Different Methods. Taking the evaluation accuracy as the index, methods 1, 2, and 3 are tested. The test results are as shown in Figure 4.

Analysis of the data in Figure 4 shows that the evaluation accuracy of method 1 fluctuates between 80% and 100%, that of method 2 between 40% and 60%, and that of method 3 between 60% and 80%. Comparing the test results of method 1, method 2, and method 3, the evaluation accuracy of method 1 is the highest because the method uses the coefficient of variation method to calculate the weight of the evaluation index of enterprise management effect, improves

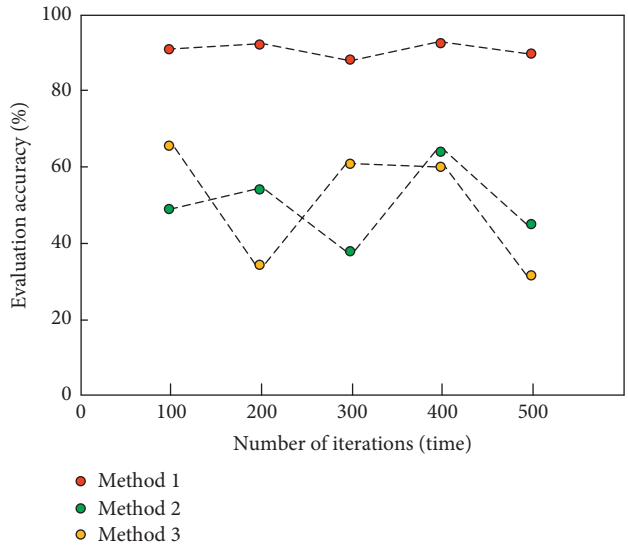


Figure 4: Evaluation accuracy of different methods.

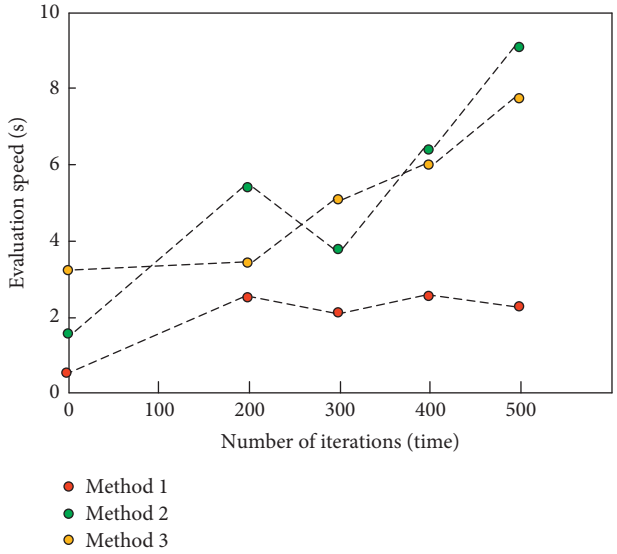


Figure 5: Comparison of evaluation speed of different methods.

the accuracy of weight calculation, and thus improves the evaluation accuracy of method 1.

3.6. Comparative Test of Evaluation Speed and Satisfaction with Different Methods. Take the evaluation speed and evaluation satisfaction as comparison indicators to compare the application performance of different methods. The results are shown in Figure 5.

As can be seen from Figure 5, under the eight iterations, the evaluation model constructed by the proposed method is faster, and the time consumption is more stable with the increase of the number of experiments, while the time consumption of other methods is higher and the evaluation speed is slower, which shows that the proposed method has greater advantages in evaluation efficiency. Figure 6 shows that in 500 iterations, the proposed method resulted in

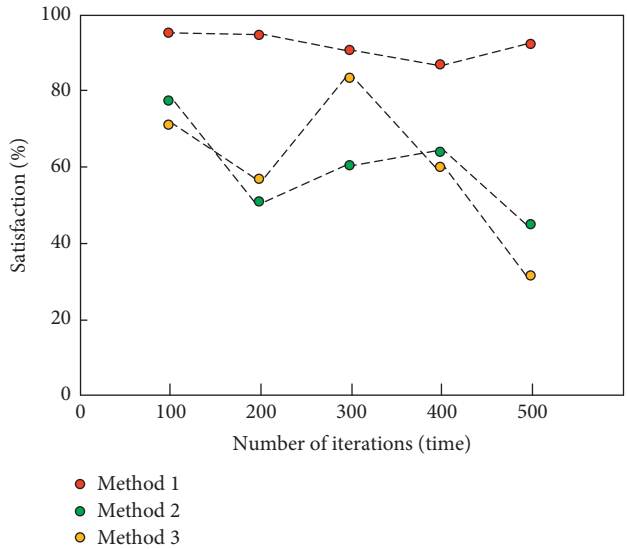


Figure 6: Analysis of satisfaction results of evaluation methods.

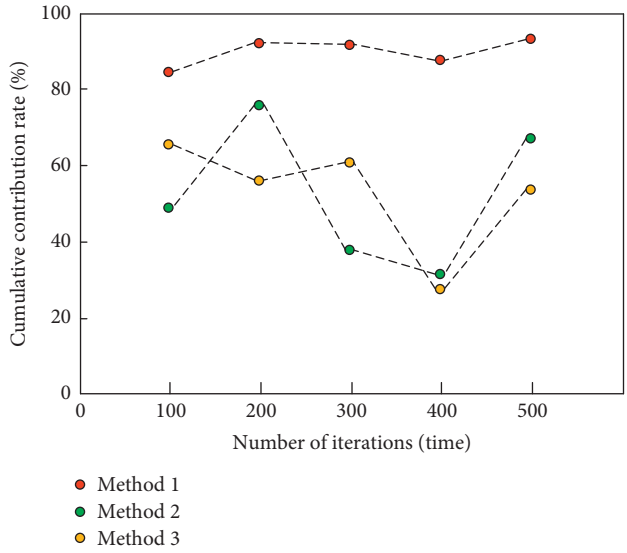


Figure 7: Comparison of cumulative contribution rates of different evaluation methods.

higher user satisfaction and a smaller degree of satisfaction fluctuation.

3.7. Comparative Test of Cumulative Contribution Rate and Root-Mean-Square Error of Different Methods. The study adopted the lean management effect evaluation method (Method 1), the evolutionary game model-based enterprise management evaluation method (Method 2), and the third-party recycling model-based enterprise management evaluation method (Method 3). The cumulative contribution rate and root-mean-square error of different methods are compared, and the results are used to measure the effectiveness of different methods. The results are shown in Figures 7 and 8.

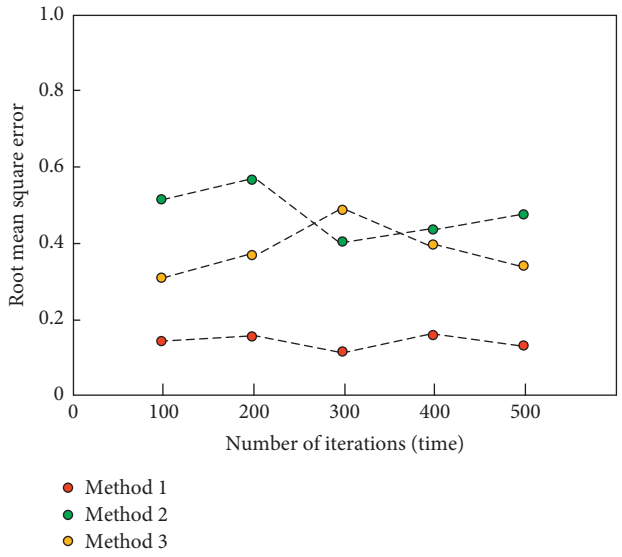


Figure 8: Stability comparison of different evaluation methods.

From the experimental results of Figures 7 and 8, it can be seen that compared with the evaluation method based on the evolutionary game model (Method 2) and the evaluation method based on the third-party recycling model (Method 3), the method proposed in this study (Method 1) has a higher cumulative contribution rate and application stability. The cumulative contribution rates of Methods 2 and 3 are low, which indicates that less effective information is obtained in the application of this method, and that indirectly results in the error of evaluation results.

To sum up, there is a high correlation between the changes of multidimensional evaluation index factors selected in this paper and enterprise M&A performance. It is more accurate and reliable to comprehensively evaluate the effect of enterprise M&A based on multiple dimensions. The significance coefficient obtained by this method is higher, the evaluation accuracy is higher, and the evaluation speed of the evaluation model is faster, and with the increase of the number of experiments, the time consumption is more stable, with higher cumulative contribution rate and application stability.

4. Conclusion

A more accurate evaluation and analysis of the evaluation methods commonly used in enterprise management is conducive to the enterprise in the actual evaluation of a reasonable choice of evaluation methods. This paper puts forward the application research of data mining in enterprise lean management effect evaluation. Through the data mining algorithm to establish the enterprise lean management effect evaluation index system, the index weight is calculated. Association rules and fuzzy set transformation are used to describe the fuzzy boundary of each index and factor, and fuzzy judgment matrix is constructed to evaluate the effect of lean management. Experimental results show that the significance coefficient and precision of the results are high, and the cumulative contribution rate is high, which shows that the proposed method has good practicability.

Data Availability

The raw data supporting the conclusions of this article will be made available by the corresponding author without undue reservation.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this work.

References

[1] C. X. Zhang, "Management innovation of tourism complex development and operation enterprises," *Journal of Landscape Research*, vol. 11, no. 03, pp. 110–112, 2019.

[2] V. K. Potemkin, "People-oriented management of enterprises and organizations," *Economics and Management*, vol. 4, no. 2, pp. 165–176, 2020.

[3] Y. Li, Y. X. Liu, G. Y. Zhang, H. Zhu, and S. O. Management, "Selection of cooperative strategies for participating subjects of new R&D institutions: evolutionary game theory," *Science and Technology Management Research*, vol. 39, no. 10, pp. 51–55, 2019.

[4] Y. Q. Cheng and Z. D. Xu, "Credit risk evaluation of small enterprises based on revised ELECTRE III by Theil index," *Chinese Journal of Management Science*, vol. 35, no. 10, pp. 51–55, 2019.

[5] j Z. Zou, N. Zhao, C. R. Lu, and L. Wu, "Complex relational environment non relational distributed large data mining simulation," *Computer Simulation*, vol. 36, no. 10, pp. 334–338, 2019.

[6] S. Abarajithan and S. V. Mohan, "Cockroach swarm optimization algorithm for high utility association rule mining," *International Journal of Swarm Intelligence Research*, vol. 12, pp. 459–469, 2021.

[7] I. Sugiyarto, B. Sudarsono, and U. Faddillah, "Performance comparison of data mining algorithm to predict approval of credit card," *SinkrOn*, vol. 4, no. 1, p. 149, 2019.

[8] K. K. Pandey, D. Shukla, and R. Milan, "Data mining algorithm and new HRDSD theory for big data," *International journal of computer sciences and engineering*, vol. 7, no. 3, pp. 76–81, 2019.

[9] J. Nalić, G. Martinović, and D. Žagar, "New hybrid data mining model for credit scoring based on feature selection algorithm and ensemble classifiers," *Advanced Engineering Informatics*, vol. 45, Article ID 101130, 2020.

[10] S. Wahyuni and M. Marbun, "Implementation of data mining in predicting the study period of student using the naïve bayes algorithm," *IOP Conference Series: Materials Science and Engineering*, vol. 769, no. 1, Article ID 012039, 2020.

[11] Y. Abboud, A. Brun, and A. Boyer, "C3Ro: an efficient mining algorithm of extended-closed contiguous robust sequential patterns in noisy data," *Expert Systems with Applications*, vol. 131, no. 10, pp. 172–189, 2019.

[12] S. Samsir, S. Suparno, and M. Giatman, "Predicting the loan risk towards new customer applying data mining using nearest neighbor algorithm," *IOP Conference Series: Materials*

Science and Engineering, vol. 830, no. 3, Article ID 032004, 2020.

[13] P. Liu, Q. Wang, and W. Liu, "Enterprise human resource management platform based on FPGA and data mining," *Microprocessors and Microsystems*, vol. 80, Article ID 103330, 2020.

[14] H. Ma, "Enterprise human resource management based on big data mining technology of internet of things," *Journal of Intelligent and Fuzzy Systems*, vol. 6, no. 1, pp. 1–7, 2021.

[15] Y. Meng, "Establishment and application of Enterprise management maturity model based on multimedia data information systems," *Multimedia Tools and Applications*, vol. 78, no. 4, pp. 4503–4525, 2019.

[16] B. Arivazhagan and R. Sankara Subramanian, "Hybrid model for pattern discovery in data communication to enhance customer relationship management using data mining t," *Journal of Physics: Conference Series*, vol. 1979, no. 1, Article ID 012048, 2021.

[17] K. Munonye and P. Martinek, "Microservices data mining for analytics feedback and optimization," *International Journal of Enterprise Information Systems*, vol. 17, no. 1, pp. 22–43, 2021.

[18] A. Barbara, A. Pimonov, and L. Sluder, "A review of methods for processing unstructured data in the assessment of mining personnel," *E3S Web of Conferences*, vol. 174, no. 1, Article ID 04040, 2020.

[19] S. Selamat, S. Prakoonwit, and W. Khan, "A review of data mining in knowledge management: applications/findings for transportation of small and medium enterprises," *SN Applied Sciences*, vol. 2, no. 5, pp. 1–15, 2020.

[20] A. Brunello, P. Gallo, E. Marzano, A. Montanari, and N. Vitacolonna, "An event-based data warehouse to support decisions in multi-channel, multi-service contact c," *Journal of Cases on Information Technology*, vol. 21, no. 1, pp. 33–51, 2019.

[21] O. Moscoso-Zea, J. Castro, J. Paredes-Gualtor, and S. Lujan-Mora, "A hybrid infrastructure of enterprise architecture and business intelligence & analytics for knowledge management in education," *IEEE Access*, vol. 7, pp. 38778–38788, 2019.

[22] H. M. Alshira, M. Al-Omari, and B. Igried, "Usability evaluation of learning management systems (LMS) based on user experience," *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, vol. 12, no. 11, pp. 6431–6441, 2021.

[23] S. Suhartina, L. M. Saleh, S. Sirajuddin, S. Baja, and A. Mallongi, "Evaluation of the application of health and safety management system (SMK3) in the mining company of PTX. Based on government regulation number 50 of 2012," *Open Access Macedonian Journal of Medical Sciences*, vol. 8, no. T2, pp. 183–187, 2020.

[24] A. Uzun, A. Onur, and S. Alabay, "Students' views on database management systems course designed according to problem-based learning," *International Journal of Evaluation and Research in Education*, vol. 9, no. 1, p. 177, 2020.

[25] T. Fatyani and Z. Yahia, "TAKA news search engine: a proposed DEMO-based performance evaluation system," in *Proceedings of the International Conference on Industrial Engineering and Operations Management*, Bangkok, Thailand, March 2019.

[26] R. Ramadiani, A. Kurniawan, Z. Arifin et al., "Evaluation of student academic performance using e-learning with the association rules method and the importance of performance analysis," *Journal of Physics: Conference Series*, vol. 1524, no. 1, Article ID 012107, 2020.

[27] S. Shukla, B. K. Mohanty, and A. Kumar, "A fuzzy approach to prioritise DEA ranked association rules[J]," *International Journal of Business Intelligence and Data Mining*, vol. 13, no. 1-2, pp. 155–176, 2019.

[28] A. Nie, "Design of English interactive teaching system based on association rules algorithm," *Security and Communication Networks*, vol. 2021, no. 1, pp. 1–10, 2021.

[29] A. Kumar, S. Shukla, and B. K. Mohanty, "A fuzzy approach to prioritize DEA ranked association rules," *International Journal of Business Intelligence and Data Mining*, vol. 1, no. 1, pp. 1–8, 2019.

[30] B. Keith Norambuena and C. Meneses Villegas, "An extension to association rules using a similarity-based approach in semantic vector spaces," *Intelligent Data Analysis*, vol. 23, no. 3, pp. 587–607, 2019.

[31] B. S. Neysiani, N. Soltani, N. Soltani, R. Mofidi, and M. H. Nadimi-Shahraki, "Improve performance of association rule-based collaborative filtering recommendation systems using genetic algorithm," *International Journal of Information Technology and Computer Science*, vol. 11, no. 2, pp. 48–55, 2019.

[32] T. Beridze, A. Cherep, Z. Baranik, V. Korenyev, and I. Vasylchuk, "Analysis of the regression model of the enterprise's financial activity by research on residual error," *Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu*, vol. 1, no. 2, pp. 193–197, 2021.

[33] M. Zeeshan, G. Mohapatra, and A. K. Giri, "The effects of non-farm enterprises on farm households' income and consumption expenditure in rural India," *Economía Agraria y Recursos Naturales*, vol. 19, no. 1, pp. 195–222, 2019.

[34] A. U. Widyaningdyah and L. Ezra, "Enterprise resource planning (ERP) sfor internal control effectiveness," *Jurnal Reviu Akuntansi dan Keuangan*, vol. 10, no. 2, p. 234, 2020.

[35] M. Kim, K. Sudhir, K. Uetake, and C. Rodrigo, "When salespeople manage customer relationships: multidimensional incentives and private information," *Journal of Marketing Research*, vol. 56, no. 5, pp. 753–766, 2019.

[36] R. F. Supardianto, R. Ferdiana, and S. Sulistyo, "The role of information technology usage on s financial management and taxation," *Procedia Computer Science*, vol. 161, pp. 1308–1315, 2019.

Hindawi
Scientific Programming
Volume 2022, Article ID 8063652, 9 pages
https://doi.org/10.1155/2022/8063652

Research Article

Research on Personalized Minority Tourist Route Recommendation Algorithm Based on Deep Learning

Guanglu Liu

University of Science and Technology Liaoning, Anshan 114051, Liaoning, China

Correspondence should be addressed to Guanglu Liu; 319973600034@ustl.edu.cn

Received 24 November 2021; Revised 6 December 2021; Accepted 7 December 2021; Published 7 January 2022

Academic Editor: Baiyuan Ding

With the improvement of living standards, more and more people are pursuing personalized routes. This paper uses personalized mining of interest points of ethnic minority tourism demand groups, extracts customer data features in social networks, and constructs data features of interesting topic factors, geographic location factors, and user access frequency factors, using LDA topic models and matrix decomposition models to perform feature vectorization processing on user sign-in records and build deep learning recommendation model (DLM). Using this model to compare with the traditional recommendation model and the recommendation model of a single data feature module, the experimental results show the following: (1) The fitting error of DLM recommendation results is significantly reduced, and its recommendation accuracy rate is 50% higher than that of traditional recommendation algorithms. The experimental results show that the DLM constructed in this paper has good learning and training performance, and the recommendation effect is good. (2) In this method, the performance of the DLM is significantly higher than other POI recommendation methods in terms of the accuracy or recall rate of the recommendation algorithm. Among them, the accuracy rates of the top five, top ten, and top twenty recommended POIs are increased by 9.9%, 7.4%, and 7%, respectively, and the recall rate is increased by 4.2%, 7.5%, and 14.4%, respectively.

1. Introduction

As the self-service travel group continues to grow, more and more users organize the travel information they want to obtain from the internet [1]. Personalized ethnic minority tourism route recommendation refers to the generation of ethnic minority tourism routes that meet their travel conditions for each user based on the user's personalized factors. However, with the rapid development of technologies such as cloud computing and big data, the scale of data has shown an explosive growth trend. When faced with massive amounts of network data, users cannot quickly select information. The typical solutions to the problem of information overload are search engines and personalized ethnic minority travel route recommendation systems.

The personalized minority travel route recommendation algorithm is the core of the recommendation system. The traditional recommendation algorithm mainly includes collaborative filtering, content-based recommendation algorithm, and hybrid recommendation algorithm. Collaborative filtering is the most widely used recommendation algorithm. Its principle is to use the interactive information between users and items to make recommendations. The advantage of collaborative filtering recommendation is that there is no need to perform complex feature modeling of users or items, and only the number of users' historical feedback is required, so it is simple and efficient. The disadvantage is that there are serious data sparseness and cold start problems. The principle of a content-based recommendation algorithm is to use the items selected by the user to find items with similar attributes for recommendation [2]. The content-based recommendation algorithm does not have the problem of sparse scoring data, and at the same time avoids the cold start problem of new users. The disadvantage is that complex feature engineering is required for feature extraction. Multimedia data such as images and audio often face the problem of feature extraction difficulties, and there is still a cold start problem for new users [3].

A hybrid recommendation algorithm refers to the fusion of multiple recommendation algorithms to achieve a better recommendation effect. Common combination strategies include weighted combination, result mixing, feature combination, cascade, and so on. This combination algorithm not only enhances the feature crossover ability of the model but also effectively avoids the problems of combination explosion and high computational complexity [4]. Deep learning originated from the study of artificial neural networks, and its concept was proposed by Hinton et al. in 2006 [5]. In recent years, in the fields of computer vision, natural language processing, and speech recognition, deep learning technology has made breakthroughs, and it has also brought a new technological revolution to the research of recommender systems. Integrating deep learning into the recommendation system can effectively make up for the shortcomings of the traditional personalized ethnic minority travel route recommendation model and improve the quality of recommendation results. Wang et al. used the influence of image features shared by tourists' historical tours on POI recommendation and proposed a POI recommendation system based on image content enhancement [6]. Zheng et al. proposed a recommendation model of deep CNN. The model extracts different features by defining two parallel convolutional neural networks. One convolution kernel is used to extract the user behavior pattern features, and the other convolution kernel is used to extract the features from items reviewed by users in history. Then the features extracted by the two convolution kernels are fused for the recommendation. This method has achieved good results in the field of item recommendation [7]. Shen et al. constructed an e-learning recommendation system by using convolutional neural networks to extract features from text information [8]. Tang et al. constructed a sequential recommendation model through a convolutional neural network [9]. Lei et al. proposed a deep learning model for image recommendation based on convolutional neural networks [10].

This chapter proposes a deep neural network recommendation framework that integrates DNN network [11] with LDA topic model and matrix factorization algorithm and uses word embedding technology for user preference features, geographic factor features, and probabilistic topic features in social networks [12] being integrated into the minority tourist routes recommendation task, through the neural network to learn the high-level interaction between features and then make personalized recommendations to users.

2. Related Work

2.1. LDA Topic Model. Latent dirichlet allocation is a topic model algorithm based on a probability model proposed by Blei et al. in 2003. The LDA is an unsupervised machine learning technology that can be used to identify the potentially hidden topic information in a large-scale document set or corpus. The method assumes that each word is extracted from a potentially hidden topic behind it.

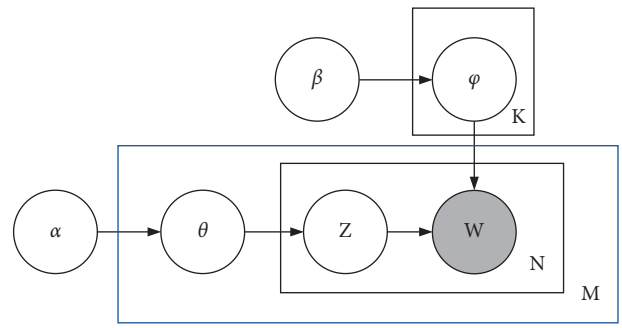


Figure 1: LDA topic model structure.

The shaded circle in Figure 1 represents the observable variable, the unshaded circle represents the latent variable, the arrow represents the conditional dependence between the two variables, the box represents repeated sampling and the number of repetitions is in the lower right corner of the box. This corresponds to the production process of LDA.

2.2. Matrix Factorization Algorithm. The matrix factorization algorithm is the core algorithm that won the 2006 Netflix recommendation competition. It has a pivotal position in the history of the entire recommendation system and has contributed to the large-scale development and industrial application of the recommendation system. Assume that the set of (u, v) pairs (u represents the user and$v$ represents the subject matter) combination of which all users have a score is A.

$$A = \{(u\ ,\ v)|r_{uv} \neq \varnothing\}. \tag{1}$$

The vectors that embed the user u and the subject v into the k-dimensional implicit feature space through matrix decomposition are

$$\begin{aligned} u &\longleftarrow p_u = (u_1, u_2, \ldots, u_k), \\ v &\longleftarrow q_v = (v_1, v_2, \ldots, v_k). \end{aligned} \tag{2}$$

Then the prediction score of user u on the subject v is

$$\widehat{r_{uv}} = p_u * q_v. \tag{3}$$

The error between the true value and the predicted value is

$$\Delta r = r_{uv} - \widehat{r_{uv}}. \tag{4}$$

If the prediction is more accurate, then the smaller $||\Delta r||$, for all users rated (u, v), if we can ensure that the sum of these errors is as small as possible, then there is a reason to believe that our prediction is accurate. Matrix factorization can be transformed into a machine learning problem, that is, an optimization problem of finding the minimum value.

$$\min_{p^*,\ q^*} \sum_{(u,v)\in A} \left(r_{uv} - p_u * q_v^{\ t}\right)^2 + \lambda\left(\left\|p_u\right\|^2 + \left\|q_v\right\|^2\right). \tag{5}$$

2.3. DNN Model. The fully connected neural network is the simplest neural network, it has the most network parameters

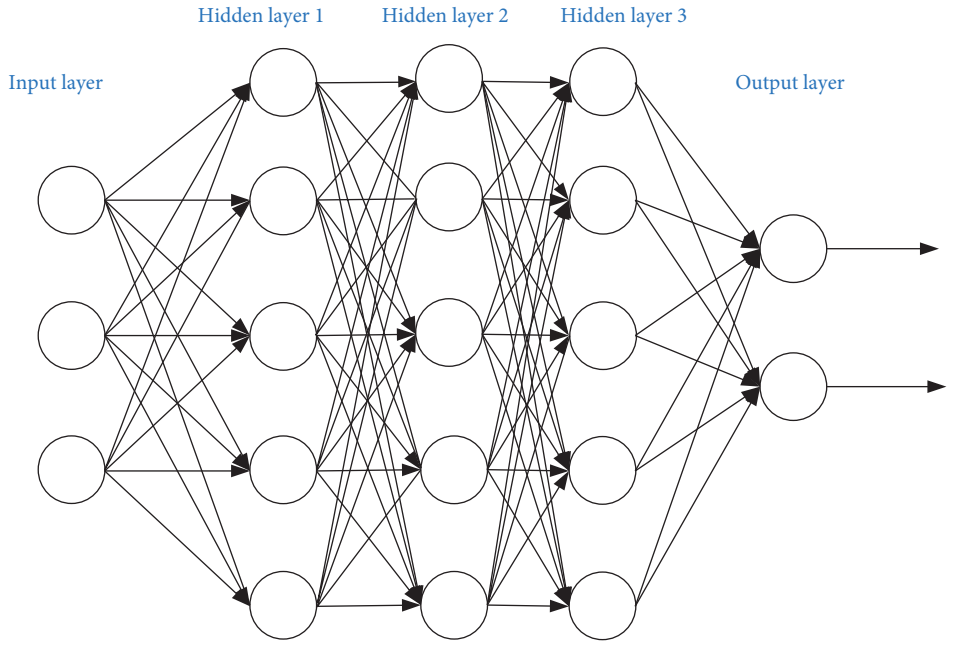


Figure 2: DNN model structure.

and the largest amount of calculation. The neural network layers inside the DNN can be divided into three categories, input layer, hidden layer, and output layer, as shown in the example below. Generally speaking, the first layer is the input layer, the last layer is the output layer, and the number of layers in the middle are all hidden Floors are shown in Figure 2.

The layers are fully connected, that is, any neuron in the i^{th} layer must be connected to any neuron in the $i+1^{\text{th}}$ layer. Although the DNN looks complicated, it is still the same as the perceptron from a small local model, that is, a linear relationship $z = \sum \omega_i x_i + b$ plus an activation function $\sigma(z)$. The commonly used activation function types are (1) sigmoid activation function; (2) tanh activation function; (3) softmax activation function; (4) Relu function.

2.3.1. Sigmoid Activation Function.

$$f(x) = \frac{1}{1 + e^{-x}}. \tag{6}$$

The three main defects of the sigmoid function: *a*. When the neural network uses the sigmoid activation function for backpropagation, the neuron whose output is close to 0 or 1 has a gradient close to 0. If a large neural network contains sigmoid neurons, many of them are in a saturated state, then the network cannot perform backpropagation, and the network cannot be learned and optimized. *b*. The sigmoid output is not centered at zero, which will easily cause the result of the subsequent activation function output to shift from the center POI. The effective range of the sigmoid function is between (−4, 4). This will cause the gradient to disappear. *c*. Compared with other nonlinear activation functions, the sigmoid function is computationally expensive.

2.3.2. Tanh Activation Function.

$$f(x) = \frac{2}{1 + e^{-2x}} - 1. \tag{7}$$

Tanh function will also have the problem of vanishing gradient and high computational cost.

2.3.3. Softmax Activation Function.

$$f(x_i) = \frac{e^{x_i}}{\sum_k e^{x_k}}. \tag{8}$$

2.3.4. Relu Function.

$$f(x) = \begin{cases} x, x > 0 \\ 0, x \le 0 \end{cases}. \tag{9}$$

The defect of the Relu function is not centered at zero, similar to the sigmoid activation function. The output of the Relu function is not centered at zero, which will cause the center to shift; during the forward pass, if $x < 0$, the neuron remains It is inactive and "kills" the gradient in the backward pass. In this way, the weights cannot be updated and the network cannot learn.

2.4. Data Vectorization. Social network data contains rich user access information and POI information. How to find the effective features from the user's historical check-in data and integrate the effective features into the user preference model is the key to improving the effect of personalized ethnic minority travel route recommendations. This section analyzes the features from three aspects: theme features, geographic factor features, and user access features, and

determines the effective features that affect the users' travel preferences. Extract better user preferences from the analyzed effective features to improve the accuracy of personalized ethnic minority travel route recommendations.

2.4.1. Theme Feature Analysis. In the check-in records of users on social networks, different users often have different theme preferences, that is, everyone has their own preferred theme. For check-in locations under the same theme, different users have their own personalized visit hobbies. For example, user u loves shopping and often visits major shopping malls, while user v is a food lover and is not keen on shopping. Then user u's favorite shopping mall is higher than user v. That is, foodies may be more interested in gourmet POI, while shopaholics are more concerned with shopping mall POI . According to the historical check-in and visit records of the user in the social network, the theme characteristics of the user's visit are analyzed, and according to the analysis result, the theme preference characteristic can fully display the user's personalized preference in the social network. Therefore, the topic features can be integrated into the personalized recommendation model, and the user's preference for different topics of the POI can be used as one of the criteria for measuring the user's preference for the POI. According to the set of POI names in the user's historical visit records, the topic vector characteristics of each user are obtained, which characterizes the user's preference for different topics of POI, and enables deep learning to better learn the topic characteristics. The probabilistic topic model digs out the personal preference for POI based on the user's historical visit frequency to different topics. Given the theme set $T[t_1, t_2, \ldots, t_n]$, according to the set H of all keywords appearing in the check-in record, the LDA model is used to obtain the keyword distribution βi corresponding to each theme t:

$$\beta_{t_i} = T \times H_{t_i}. \tag{10}$$

H_{t_i} represents a matrix of keywords belonging to the topic t_i. For the user u_i, every social network record has its probability topic distribution $P[p_{t_1}, p_{t_2}, \ldots, p_{t_2}]$ in the topic set T. This distribution describes the probability of each record to the theme of the set T and represents the degree of relevance between the points of interest (POI) and the theme.

For the check-in record w in the check-in set W of a certain user in the data set, the probability $T_w\ [t]$ that the check-in record w belongs to the topic t is obtained according to formula (11). The topic vector feature T_w in the check-in record w is used to reflect the user's personalized access preferences for different topics.

$$T_w[t_i] = \frac{N_\omega \in \beta_{t_i} + \alpha}{|W| + |T| * \alpha}, \tag{11}$$

where $N_\omega \in \beta_{t_i}$ represents the number of keywords belonging to H_{t_i} in the check-in record W, α represents the prior probability of symmetric Dirichlet, generally $\alpha = 0.1$, $|W|$ and |T|, respectively, represent the number of keywords ω in W in the check-in record, and the number of potential topics t_i.

2.4.2. Geographical Factor Analysis. In the check-in records of social networks, first, according to all the historical check-in records of each user, the user's historical TOP-K check-in record POI is calculated. The user's access frequency to the POI can reflect the user's personal preference from the side. TOP-K POI of interest and all historical records are visualized and analyzed using ArcGIS tools. All points of interest in the user's TOP-10 are distributed in the center of the user's activity, that is, the user's activities are basically carried out around these centers. It can be concluded that the user's historical activities are all in his frequent area, and the user is usually used to visiting the POI in the frequent area, that is, the user always likes to visit the POI that is close to his frequent area. Therefore, geographic factor features can be integrated into the personalized recommendation model, and the user's preference for POI in different geographic locations can be used as one of the criteria for measuring the user's preference for this POI.

The geographic distance affects the choice of human visits. The closer the distance to the center of the frequent activity area, the greater the possibility of the user's choice [13]. According to formula (11), the longitude and latitude information in the user sign-in data are respectively normalized, and the data is normalized to one appropriate range, and input into the model as the geographic factor feature of each POI.

$$\text{geo} = \frac{(\text{loc} - \text{loc}_{\min})}{(\text{loc}_{\max} - \text{loc}_{\min})}. \tag{12}$$

where loc is the longitude or latitude of the POI, $\text{loc}_{\max}$ represents the largest value in the longitude or latitude of the location information, $\text{loc}_{\min}$ is the smallest value in the longitude or latitude, as the data format after the neural network is normalized. POI geographic factor characteristics. In addition, normalization is used to process the location information in the social network data, and the normalized latitude and longitude eliminate the influence of singular data on model training and increase the learning ability of the model.

2.4.3. Analysis of User Access Characteristics. In social networks, users will not show their favorite POI like movie ratings, and the user's check-in record does not show their favorite POI. Even if the user has visited a POI, the user may not necessarily like it. For the POI, the performance of user preference is often implicit feedback [14]. In social networks, the more frequently a user visits a certain POI, the more likely the user is to prefer the POI. Based on the preference of user visit frequency, user visit characteristics in social networks can be determined, and the relationship between users can still be analyzed based on the idea of collaborative filtering, that is, users with the same visit pattern may have similar interests and preferences. Therefore, similar interests

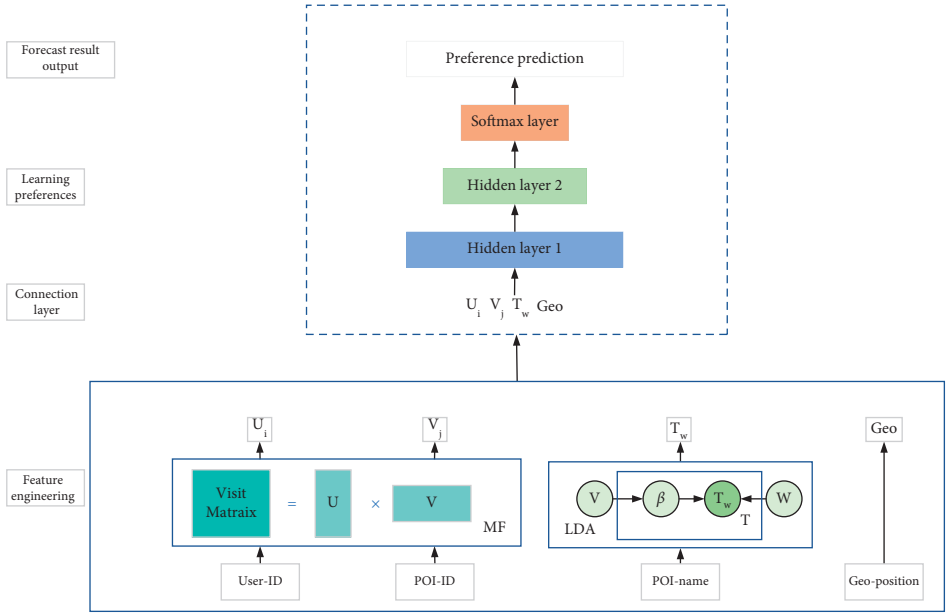


Figure 3: DLM model structure.

among similar users can be used as one of the criteria for measuring users' preference for this POI.

The singular value matrix decomposition technology is used to vectorize the user's sign-in data. After the user's sign-in record matrix is matrix-decomposed, the high-dimensional sparse data is transformed into a low-dimensional user potential vector. For users with similar sign-in records, the user potential vector obtained by matrix decomposition is closer in the vector space.

According to the user records in the social network, the user-POI sign-in matrix $N \times M$ is formed, where N represents the number of all users in the city, M represents the number of all POI of interest in the city, and the value in the check-in matrix Rj represents the user u The total frequency of visits at POI is v. According to formula (13), the historical visit matrix of user-POI is matrix-decomposed.

$$R_{N\times M} = U_{N\times M} \sum_{N\times M} V^{T}_{M\times N}. \tag{13}$$

3. Personalized Minority Travel Route Recommendation Algorithm

3.1. DLM Recommendation Algorithm Framework. The DNN network can adaptively learn high-level features and their interactions from the input of a specific task. Therefore, this chapter proposes to use the DNN deep learning model for personalized POI recommendations. The overall framework of the model is shown in Figure 3. DLM personalized POI recommendation model is divided into feature extraction module and network learning module. Among them, the feature extraction module uses word embedding technology to achieve the extraction and construction of LSBN features; the network learning module includes a network connection layer and a network layer, the connection layer in the network learning module realizes the feature fusion of the extracted feature vectors, and the network learning module The network layer implements DLM model training and user preference score prediction for POI.

The DNN network can adaptively learn high-level features and their interactive features from the input of a specific task. Therefore, this chapter proposes to use the DNN deep learning model for personalized ethnic minority travel route recommendations. The overall framework of the model is shown in Figure 3. The DLM personalized ethnic minority travel route recommendation model is divided into a network training module and a feature vectorization module. Among them, the feature vectorization module uses word embedding technology to realize the extraction and vectorization of points of interest features in social networks; the network training module includes a network connection layer and a network layer, and the network training module first uses the connection layer to feature the extracted feature vector integration, and then through the network layer to achieve the DLM model training and the user preference score prediction of ethnic POI.

The feature vectorization module uses word embedding technology to extract topic feature vectors, geographic factor feature vectors, and user-visit feature vectors in the user's social network. The network connection layer in the network training module fully connects the vectors extracted by the feature vectorization module and sends them to the DNN training network. The connection layer ensures the scalability of the entire model. If other relevant context information needs to be added, it can be automatically fully connected through the fully connected layer, and the characteristics of the input layer can be sent to the network for training. The network layer in the network training module contains two functions: training and prediction. Among them, the DNN network is used in the training stage to extract and learn implicit features; then the feature vectors

extracted in the hidden layer are input to the softmax layer for classification task learning. The DLM model transforms personalized ethnic minority travel route recommendation into a two-category task, in which the user's check-in record is defined as a positive sample and personalized ethnic minority travel routes that the user has not visited as a negative sample. The output of the softmax layer is a two-dimensional probability vector $P = [Q1, Q2]$, where $Q1$ represents the user's preference probability for personalized ethnic minority travel routes, and $Q2$ represents the user's non-preference probability for personalized ethnic minority travel routes, a cross-entropy loss function is constructed based on the network output results and positive and negative samples, and the gradient descent method is used to optimize the function; in the prediction stage, input user and personalized ethnic minority travel route information, and the network outputs a probability vector P, ranking top according to the probability of $Q1$ k, recommend to users.

3.2. DLM Model Learning Optimization. The connection layer of the DLM learning module fully connects the existing features and sends them to the neural network for learning. For any user-POI pair $< U, V >$, its fully connected vector representation is shown in formula.

$$X0 = \text{Merge}\left(\langle U_i, V_j \rangle, T_w, \text{Geo}\right). \tag{14}$$

Among them, the feature contains the user-POI latent vector $< U_i, V_j >$, the user preference feature, the topic feature T_w, and the geographic factor feature Geo. Merge connects all the feature vectors into a one-dimensional vector and sends it to the model. According to formula (15) calculated in the hidden layer of the model

$$X1 = \text{Dropout}\left(\text{Relu}\left(W^{\tau} X0 + b\right)\right), \tag{15}$$

where τ represents the number of hidden layers in the model, and the activation function used is Relu. Avoiding the problem of gradient disappearance in multi-layer network training, reducing the impact of this problem on model training, and keeping the model in a stable state during iteration [15]. In addition, in each hidden layer training, the dropout technology is added to prevent over-fitting problems. The addition of this technology can effectively increase the generalization ability of the model so that the model still has strong adaptability when dealing with deep network training. In the output layer of the model, the predicted access probability of the user u to the POI v is obtained, as shown in formula .

$$Y = \text{Soft max}\left(W_{\text{out}} * X1 + b_{\text{out}}\right). \tag{16}$$

Among them, W_{out} is the weight value of the output layer, and b_{out} is the bias value of the output layer. The output of the softmax output layer of the model is two probability values, which respectively represent the possibility of the user accessing the POI and the possibility of the user not accessing it. Here, cross-entropy is used as the loss function of model adjustment, as shown in formula.

$$E = -\sum \left(y \text{In} y + (1 - y) \text{In}(1 - y)\right). \tag{17}$$

The number of times of the summation of the loss function is related to the dimension of the input data. The optimization of the model is achieved through the minimization formula (17), and the sorted top-k is output as the recommendation result. Among them, k is the number of recommended POI results. Finally, the recommended results are sorted by the prediction results after model training according to the probability that the user may visit. The greater the probability, the more likely the user is to visit. When the value of k is determined, the corresponding top-k prediction results are selected to recommend to the user according to the sorting result of the probability.

DLM personalized minority travel route recommendation algorithm. The model training code is shown in Table 1.

4. Simulation Experiment

This chapter designs and implements two parts of experiments to verify the effectiveness of the methods in this chapter, including comparison experiments of the characteristics of the DLM algorithm and comparison experiments with existing methods on real data sets and verification and evaluation of experimental results with algorithm recommendations. In the self-feature comparison experiment of the DLM algorithm, the different feature combinations in the algorithm of this chapter are compared in the comparison experiment with the existing method, the comparison between the DLM algorithm and the existing POI recommendation algorithm is realized in the chapter. Finally, the DLM algorithm and other recommendation algorithms are compared and analyzed.

4.1. Description of the Experimental Data Set. From the real data set of Foursquare, users whose check-in data location is Beijing were selected as the experimental data. In addition, the data was preprocessed and denoised. In the division of the experimental data set, in this chapter, the historical check-in record data of the user u will be randomly selected according to 8 : 2. Among them, 80% of the training set is used for model training, and 20% of the test set is used for model evaluation. Before data application, a good job of data denoising and filtering is performed, and the inconsistent data are deleted through the test of the distance between the check-in point and the positioning point of the social network.

4.2. Description of the Comparison Algorithm. Traditional recommendation algorithms UCF, PMF, LCARS, etc. are selected. To compare the promotion model of this article, the optimization effect of this algorithm is studied, and the control variable method is used to construct a recommendation algorithm lacking a certain data feature recognition module, and the importance of data features is evaluated by comparing the recommendation effect.

Table 1: DLM personalized minority travel route recommendation algorithm.

Algorithm 1: DLM training
Premise
1. Construct user POI access matrix according to the frequency of user access to POI.
2. User topic vector matrix representation
3. MF algorithm iteration number N1
4. DLM recommended model training iteration number N2
5. The normalized geographic factor feature vector matrix
6. Input: user ID, POI ID, POI name, geographic factor information
7. Cycle
8. Matrix factorization
9. for $n \leftarrow 1$ to N1 do
10. $R_{N\times M} = U_{N\times M}\sum_{N\times M}V^T_{M\times N}$ $_k$
11. Loss_MF $= \sum_{(n,m)\in R}(r_{n,m} - \sum U_{n,k}V_{m,k})^2$
12. Use gradient to optimize Loss(Loss_MF)
13. return U, V
14. Train the DLM mode
15. for $n \leftarrow 1$ to N2 do
16. Predict the probability of user preference
17. $E = -\sum (y\mathrm{In}y + (1-y)\mathrm{In}(1-y))$
18. Use gradient to optimize Loss(E)
19. End
20. End

UCF: This is a user-based collaborative filtering method that improves the efficiency of personalized ethnic minority travel route recommendations by taking into account the influence of interest among similar users.

PMF: Explain the feasibility of matrix decomposition from the perspective of the probability generation process, and then recommend it to users.

LCARS: This method integrates topic features into the recommendation system and realizes travel route recommendations to users by considering the comprehensive interests of personal interests and user preferences [16].

Rank-GeoFM: Matrix decomposition method based on ranking, through this method, the number of user access matrices is increased, the sparseness of data is alleviated, and the matrix decomposition method is used to realize travel route recommendation to users [17].

SGFM: Based on the social relationship and geographic influence between users, this method designs a travel route recommendation method based on social geographic factors [18].

DLM: This method is the recommended method proposed in this chapter. The recommended method of user interest feature, topic factor feature, and geographic factor feature is added to the feature fusion.

DLM_MF: This method is the recommended method proposed in this chapter, and only the pass matrix is included in the feature fusion Decomposed user preference characteristics.

DLM_MF + Geo: This method is the recommended method proposed in this chapter, and only the passing moment is included in the feature fusion user preference characteristics and geographic factor characteristics obtained by matrix decomposition.

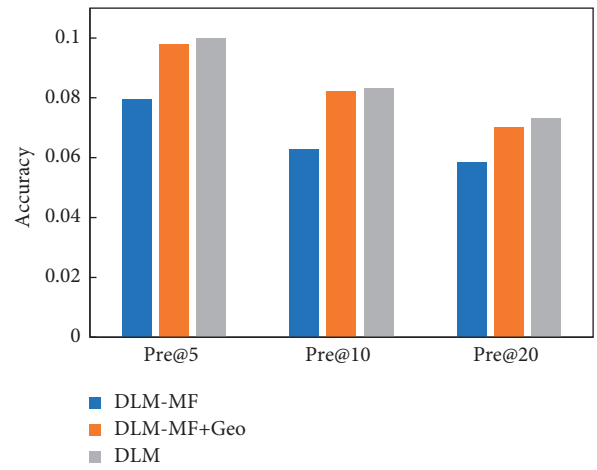


Figure 4: Comparison of @N results before DLM model feature factors.

4.3. Evaluation Criteria. In this chapter, two broad indicators are used to evaluate the performance of different recommendation algorithms, namely, accuracy and recall (represented by pre@N and Rec@N, respectively) [19], as shown below:

$$\begin{aligned} \mathrm{pre@}N &= \frac{1}{\sigma}\sum_{u\in U}\frac{\mathrm{Top}-N\cap K}{N}, \\ \mathrm{Rec@}N &= \frac{1}{\sigma}\sum_{u\in U}\frac{\mathrm{Top}-N\cap K}{K}. \end{aligned} \tag{18}$$

Among them, δ represents the number of users, N represents the number of recommended points, Top- N represents the list of the top N points of interest recommended by the recommendation model to the target user, and K represents the actual check-in list in the user test set, that is, the user's actual history in the visit record, the set of POI that the user has actually visited.

4.4. Experimental Results and Analysis. By comparing the recommendation results of the three models of DLM, DLM_MF, and DLM_MF + Geo, the impact of each feature in the social network on the recommendation results was studied, and the comparison was made in the case of Pre@5, Pre@10, and Pre@20. The experimental results are shown in Figure 4.

Experiments show that the three DLM methods that incorporate the three features are better than the DLM_MF + Geo model. The recommendation effect of the DLM_MF + Geo model is significantly better than that of DLM_MF, which shows that in the recommendation system, the selection of geographic features is more important than the selection of features that are of interest to users. The best DLM effect shows that adding topic features to the recommendation system is more conducive to accurately positioning user needs.

By comparing the DLM method with other methods on the accuracy and recall rates on the Foursquare data set, The recommended results are shown in Figures 5 and 6.

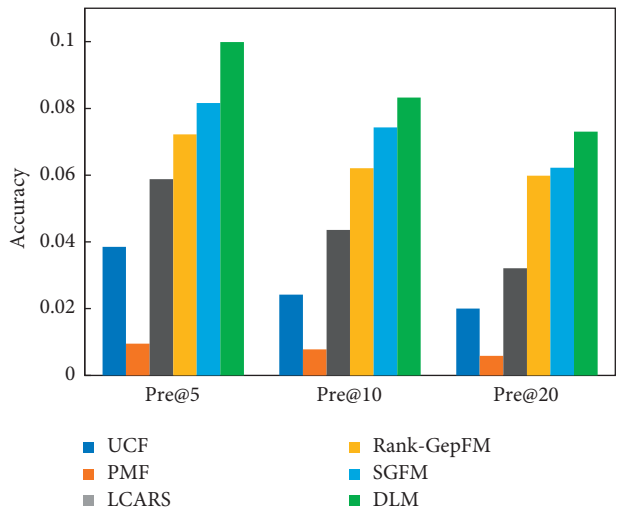


Figure 5: Comparison of recommended accuracy rates of various algorithms.

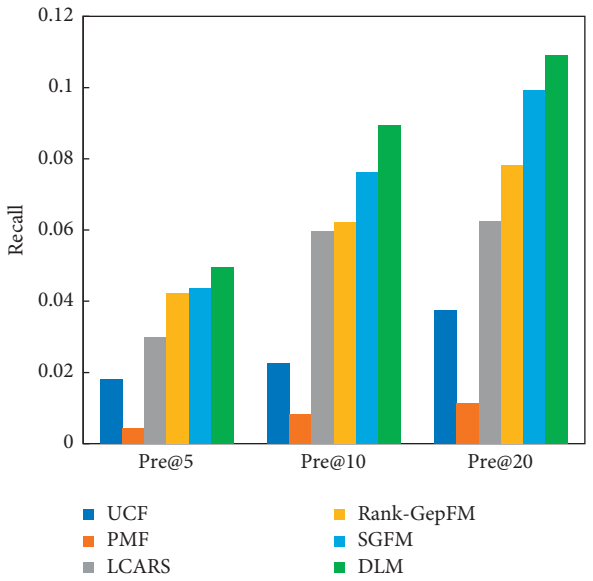


Figure 6: Comparison of recommended recall rates of various algorithms.

The results show that on the Foursquare data, the DLM algorithm is significantly better than other personalized ethnic minority tourism route recommendation algorithms, which shows that the recommendation system based on deep learning is effective for the recommendation of ethnic minority tourism routes; among them, the user-based collaborative filtering algorithm and the matrix factorization algorithm had the worst recommendation effect, and neither algorithm makes use of the geographic influence or other characteristics in social networks to make effective personalized recommendations. Compared with the UCF and PMF algorithms, the LCARS algorithm adds theme factor features, and its recommendation effect is effectively improved. In the case of Pre@5, its recommendation accuracy rate is 50% higher than that of the UCF. Rank-GeoFM and SGFM also have a significant improvement due to the addition of geographic factors. In the case of Pre@5, its recommendation accuracy rate is improved by 110% compared with the UCF. In the task of personalized travel route recommendation, how to better integrate the characteristics of social networks is the key to improving the recommendation performance. The recommendation effect of the DLM model based on deep learning algorithms is due to other traditional algorithms, one is the integration of three data characteristics, and the other is the advantages of deep learning algorithms, which makes the recommendation effect more accurate.

5. Conclusion

This paper uses feature extraction of user check-in data in social networks to construct data features of interest topic factors, geographic location factors, and user access frequency factors. The LDA topic model and matrix decomposition model are used to perform feature vectorization processing on user check-in records. The word embedding technology and the DNN network model are jointly constructed as a DLM model and this model is used to compare with the traditional recommendation model. The experimental results show the following:

(1) In the personalized recommendation system of ethnic minority tourism routes, the characteristics of geographic factors recorded by users are more important to the recommendation effect than the topics and frequency of user visits.

(2) In this method, the performance of the DLM model is significantly higher than other POI recommendation methods in terms of the accuracy or recall rate of the recommendation algorithm. Among them, the accuracy rates of the top five, top ten, and top twenty recommended POIs are increased by 9.9%, 7.4%, and 7%, respectively, and the recall rate is increased by 4.2%, 7.5%, and 14.4%, respectively.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This study was supported by the Liaoning University of Science and Technology Fund (Research on the New Development of Rural Tourism Industry under the "Two-Wheel Drive" Integration of Agriculture and Tourism) (Item no. 2019RW08), Liaoning Provincial Department of Education Project (Research on the Composition and Influencing Factors of Retired People's Tourism Happiness) (Item no. LJKR0127), and Liaoning Provincial Department of Education Project (Research on the Formation

Mechanism and Influence of Tourist Awe Emotional Experience Based on Embodied Cognition) (Item no. 2020LNJC10).

References

[1] O. Iatrellis, A. Kameas, and P. Fitsilis, "A novel integrated approach to the execution of personalized and self-evolving learning pathways," *Education and Information Technologies*, vol. 24, no. 1, pp. 781–803, 2019.

[2] P. Resinick and H. R. Varian, "Recommender systems," *Communications of the ACM*, vol. 40, no. 3, pp. 56–58, 1997.

[3] B. Alejandro, C. Iván, D. R. Fernando, C. Pablo, and C. Enrique, "An empirical com-parison of social, collaborative filtering, and hybrid recom-menders," *ACM Transactions on Intelligent Systems and Technology*, vol. 4, no. 1, pp. 1–29, 2013.

[4] H. T. Cheng, L. Koc, J. Harmsen et al., "Wide & Deep Learning for Recommender Systems," in *Proceedings of the 1st Workshop on Deep Learning for Recom-mender Systems*, pp. 7–10, Boston, MA, USA, September 2016.

[5] B. Gavade Anil and S. Rajpurohit Vijay, "A hybrid optimization-based deep belief neural network for the classification of vegetation area in multispectral satellite image," *International Journal of Knowledge-Based and Intelligent Engineering Systems*, vol. 24, no. 4, 2021.

[6] B. Mohamed, E. F. Brahim, and B. Mohammed, "Toward a personalized learning path through a services-oriented approach," *International journal of Emerging Technologies in Learning*, vol. 14, no. 15, pp. 52–66, 2019.

[7] V. Vanitha and P. Krishnan, "A modified ant colony algorithm for personalized learning path construction," *Journal of Intelligent and Fuzzy Systems*, vol. 37, no. 5, pp. 6785–6800, 2019.

[8] H. Bal´Azs, K. Alexandros, B. Linas, and T. Domonkos, "Session-based recommendations with recurrent neural networks," in *Proceedings of the International Conference on Learning Representations*, pp. 10–15, San Juan, Puerto Rico, May 2016.

[9] D. Robin and B. Hugues, "Collaborative filtering with recurrent neural networks," *Computer Science*, pp. 1–8, 2016.

[10] S. E. Mohammad and R. D. Edward, "Effect of separate sampling on classification accuracy," *Bioinformatics*, vol. 30, no. 2, pp. 242–250, 2014.

[11] J. Basilico and Y. Raimond, "Deja Vu:The Importance of Time and Causality in Recommender Systems," in *Proceedings of the Eleventh ACM Conference*, Como, Italy, August 2017.

[12] G. Pedro, F. D Campos, and I. Cantador, "Time-aware recommender systems: a comprehensive survey and analysis of existing evaluation protocols," *User Modeling and User-Adapted Interaction*, vol. 24, no. 1-2, pp. 67–119, 2014.

[13] H. Li, Y. Ge, and R. Hong, "Point-of-Interest Recommendations: Learning Potential Check-Ins from Friends," in *Proceedings of the 22nd ACM SIGKDD International Conference*, August 2016.

[14] R. M. Bell and Y. Koren, "Lessons from the Netflix prize challenge," *ACM SIGKDD Explorations Newsletter*, vol. 9, no. 2, pp. 75–79, 2007.

[15] L. Castillo, E. Armengol, E. Onaindia et al., "samap: an user-oriented adaptive system for planning tourist visits," *Expert Systems with Applications*, vol. 34, no. 2, pp. 1318–1332, 2008.

[16] S. Chintapalli, D. Dagit, B. Evans et al., K. Nusbaum, K. Patil, Peng, and Poulosky, Benchmarking Streaming Computation Engines: Storm, Flink and Spark Streaming," in *Proceedings of the IEEE International Parallel and Distributed Processing Symposium Workshops*, pp. 1789–1792, Chicago, IL, USA, May 2016.

[17] A. Perrot, R. Bourqui, N. Hanusse, and D. Auber, "HeatPipe: High Throughput, Low Latency Big Data Heatmap with Spark Streaming," in *Proceedings of the International Conference Information Visualisation*, pp. 66–71, July 2017.

[18] I. R. Brilhante, J. A. Macedo, F. M. Nardini, R. Perego, and C. Renso, "On planning sightseeing tours with TripBuilder," *Information Processing & Management*, vol. 51, no. 2, pp. 1–15, 2015.

[19] K. H. Lim, J. Chan, C. Leckie, and S. Karunasekera, "Personalized trip recommendation for tourists based on user interests, points of interest visit durations and visit recency," *Knowledge and Information Systems*, vol. 54, no. 2, pp. 375–406, 2018.

Hindawi
Scientific Programming
Volume 2022, Article ID 5641782, 11 pages
https://doi.org/10.1155/2022/5641782

Research Article

Construction of Moral Education Evaluation Model Based on Quality Cultivation of College Students

Xiaolin Yuan

School of Psychology, Hainan Normal University, Haikou, Hainan 571158, China

Correspondence should be addressed to Xiaolin Yuan; 040029@hainnu.edu.cn

Received 9 November 2021; Revised 23 November 2021; Accepted 3 December 2021; Published 7 January 2022

Academic Editor: Baiyuan Ding

Contemporary young college students are greatly impacted in the aspects of moral cognition and moral choice, which results in the weak moral will of some college students, vague moral concepts, and weak ideals and beliefs, which seriously affect the formation and development of college students' moral quality. Therefore, the moral education evaluation model based on college students' quality cultivation is constructed. Firstly, the present situation and defects of college students' quality training are analyzed. Based on this, association rules in data mining method are constructed and introduced to extract valuable knowledge hidden in the data to assist education managers to make effective decisions and improve management level. Finally, the evaluation index is selected and the weighted principal component TOP-SIS model is constructed to realize the evaluation of moral education based on college students' quality cultivation. The experimental results show that the evaluation results of the model are consistent with the actual situation, high degree of fit and freedom, and good practical performance.

1. Introduction

With the comprehensive deepening of economic globalization, the surging of new ideas, and the intersection and collision of multiculturalism, it has a profound impact on people's original moral concepts [1, 2]. Therefore, whether colleges and universities can grasp the pulse of the times and solve the moral confusion and value conflict of college students not only plays a vital role in comprehensively promoting the ideological and political work of colleges and universities, but also concerns whether college students can practice the socialist core values with practical actions [3].

Reference [4] uses Amos to test the doctoral education quality evaluation model. The results show that the education quality evaluation model can be divided into four parts: input quality, process quality, output quality, and development quality. There is a significant positive effect between these four parts. Reference [5] takes multiple regression analysis of time series as the main means to explore the linear relationship between the scientific and technological innovation service function of local universities and the characteristics of innovation and entrepreneurship education resources and gives relevant conclusions and policy analysis. In [6], aiming at the phenomenon of "soft evaluation" in college curriculum academic evaluation, taking the evaluation triangle as the theoretical basis, and based on the comprehensive analysis of relevant research results at home and abroad, this paper constructs an "evidence" based academic evaluation model of college mixed learning curriculum (e-abc model). Based on the above research, this paper constructs an evaluation model of moral education based on college students' quality training, introduces association rules, extracts valuable knowledge, and assists education managers to make effective decisions and improve management level, building a weighted principal component TOP-SIS model to realize the evaluation of moral education. Compared with previous studies, the advantages of the model are that the evaluation results of the model are basically consistent with the actual situation, high degree of fit and freedom, and good practical performance.

2. Current Situation of College Students' Moral Education

(1) Moral education is divorced from the reality of society. Morality originates from social life and is developed and perfected in social practice. Practice is the precondition for the existence and development of morality. Therefore, the moral character is the human forms and the development in the interactive contact practice; the personal practice cannot be less [7]. However, at present, the moral education of college students in some schools is often superior to the students' real life and divorced from the students' moral knowledge and behavior. In view of this serious situation, some colleges and universities still have weak response to the social reality in the process of moral education, and the theory of moral education lacks keeping pace with the times and self-innovation [8]. This kind of moral education is out of touch with the real society and lacks appeal and influence. It is impossible to make a convincing interpretation of the puzzles brought by the complicated real world to the moral cognition of college students from the theoretical level, which leads to the conflict between the moral behaviors of college students and their moral ideas, or even a set of ideas to switch between the superficial "moral spiritual world" and the inner "moral real world" and become a moral "double-faced person."

(2) The blending degree of moral education and knowledge education is insufficient. Herbart, a German philosopher and educator known as "the father of educational science," argues that teaching without moral education is a means to no end, and moral education (or character education) without teaching is an end without means [9, 10]. Obviously, in Herbart's teaching idea, the moral education and the knowledge education need to unify. That is to say, moral education cannot fight alone but can play the role of "leading" and "running through" knowledge education only when it permeates into the teaching contents and process of various subjects. As the carrier of spreading morality, moral education should be embodied in the rich content of students' real life, which coincides with the function of knowledge education. Therefore, moral education and knowledge education complement each other in purpose and are unified in whole [11]. However, due to such problems as social value function orientation, teacher education concept, teaching evaluation index, or teacher assessment, in actual teaching practice, some teachers only pay attention to impart professional knowledge and train thinking ability to students but neglect to cultivate and inspire moral spirit and value rationality. Students only learn instrumental knowledge and lack the promotion of moral character [12]. Moral education is basically out of line with knowledge education. What is more regrettable is that some moral educators understand moral education narrowly as the education of moral knowledge, inculcate moral knowledge into education instead of moral character, and lack the consideration of the true feelings of college students, so that moral education evolves into the study of moral concepts, moral rules, and moral feelings, moral education becomes pure knowledge teaching, and education becomes the materialization of human beings.

(3) Moral education is separated from virtue cultivation. Colleges and universities cultivate moral education. It is the duty and responsibility of moral education in colleges and universities to cultivate the moral character of college students. This kind of moral education should be made up of the interaction between the external normalization and the internal cultivation of virtue, and the two are integrated into each other, so as to make the moral subject conscious, voluntary, and self-motivated in the historical circumstances and the practice of the times [13]. However, some scholars always emphasize the enlightenment and restraint of external norms, equate moral education in colleges and universities with "educators turn the moral norms and requirements respected by the society into a kind of education of individual moral character of the educated," regard it as the ultimate goal of moral education to train college students to obey the basic moral rules and constraints, and neglect the cultivation of inner moral character, which leads to the confusion and perplexity of cultivating college students into "obligatory moral" who blindly obey the basic social rules. It is precisely because of separating the cultivation process of external norms and internal virtues, ignoring the difference and connection between external norms and internal virtues, paying attention only to the obedience of external norms, and ignoring the moral value pursuit of the main body of university students that the two-way integration of the two aspects is deficient, there is no humanistic concern, only empty requirements of moral norms, and it is difficult to penetrate and touch the moral mind of university students, and it also deviates from the "original heart" of cultivating and promoting the moral character of university students [14].

(4) Moral education is divorced from moral subjects. There is no doubt that people's moral character is not born but is gradually formed in the process of learning and practicing. Because of the differences in growing environment and life experience, people's moral cognition and behavior show greater differences. College moral education must also pay attention to this difference, study the students' moral development level, teach students according to their needs, and use various methods synthetically. While moral indoctrination is essential in this process, it may lead to a gradual distance between moral

education and the educated [15]. In moral education in colleges and universities, some moral educators, without fully understanding the students' living situation, actual needs, thoughts, and their different needs for morality, regard students as pure objects, ignore the subjective role of students and the differences of students, adopt monologue indoctrination education, and impose moral truth on students, and teachers seem to become the mouthpiece of morality, thus making it difficult for moral education to touch students' souls and obtain value identification, aggravating students' dissatisfaction and disgust for moral education and contradicting moral education from the heart. The biggest weakness of this educational model is that it neglects the subjectivity of moral education, mental development, and ideological reality of college students, which easily leads to the disconnection of moral knowledge and practice.

3. Moral Education Evaluation Model Data Mining

3.1. Data Mining Process and Main Functions. Data mining refers to the process of searching hidden information from a large number of data through algorithms. It is mainly a technology used by various disciplines. It plays a very important role in various fields. Data mining is a hot issue in the field of artificial intelligence and database. The so-called data mining refers to a nontrivial process of revealing implicit, previously unknown and potentially valuable information from a large amount of data in the database. Data mining is a decision support process of moral education evaluation. It highly automatically analyzes the data of moral education evaluation, makes inductive reasoning, excavates potential models, and helps decision makers adjust education and teaching strategies, reduce risks, and make correct decisions The process of extracting effective knowledge hidden in noisy and fuzzy data information [16, 17] comes from the knowledge extraction in the database, namely, KDD (Knowledge Discovery Database) [18], and it is also the theoretical basis of the proposed data mining. People use KDD to describe the whole process of data mining and use data mining to describe the basic process of data mining using mining algorithm. The process of discovering valuable knowledge in the database, known as the KDD process, is shown in Figure 1.

According to Figure 1, KDD is composed of data collection and processing, data information mining, and result analysis, etc. Among them, data mining is mainly based on association analysis and cluster analysis and prediction of various statistical analysis tools to find useful knowledge in large-scale data and through model evaluation will be valuable model as knowledge to assist relevant personnel to make a scientific and rational decision.

The main function of data mining mainly refers to making valuable knowledge-based decisions by predicting future development trends. The main function of data mining technology is not realized by a single way, but by a group of methods. The main task of data mining is to find valuable knowledge or information from relevant databases. Its main functions are as follows.

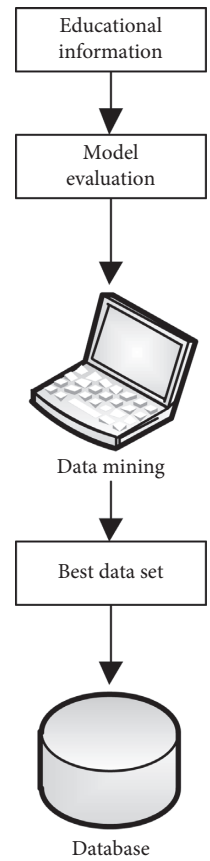


Figure 1: Overall process of KDD.

(1) Correlation analysis of data

In a database, the analysis of the correlation between data is the key link to discover the important knowledge. If there is some special law between the values of two or more variables, then it can be called correlation [19].

(2) Cluster analysis between data

In the database, the data information records can be divided into different meaningful data subsets; the process is data clustering. Among them, the smaller the distance between the data in the same data class, the more similar they are, and the less similar they are by contraries. Data clustering can enhance people's cognition of objective facts, and it is the precondition of concept representation and deviation analysis. At present, there are several data clustering algorithms in common use: mean clustering, density clustering, and fuzzy clustering, etc. [20].

(3) Forecast analysis

In some instances, the public may have to predict some vacant values. If the data for the vacant value is a numeric type, it is generally called a forecast. Prediction process is according to historical data to

find the laws between the data information, build a corresponding model, and based on the model predict the next data.

(4) Data evolution analysis

In the database, the analysis of the evolution of data information indicates that there are laws and trends in the target of a certain behavior changing with time, and the model of such trends is established [21].

3.2. Data Mining and Quality Evaluation of Moral Education. From the above analysis, we can see that association rules are one of the most significant methods in data mining. The application of association rules in the management of moral education quality is more and more extensive [22]. There are many objects involved in the management of college moral education quality, such as curriculum and evaluation index, etc. Association rules can extract valuable knowledge hidden in data, assist educational administrators to make effective decisions, and improve the management level.

4. Constructing the Quality Evaluation Model of Moral Education

4.1. Selection of Evaluation Indicators. The evaluation index of moral education level is the key to construct a high-precision evaluation model, and the index system is the concrete regulation of moral education level.

In order to construct an objective and accurate evaluation model of moral education quality, a hierarchical and multiobjective evaluation index system is constructed by AHP [23]. The AHP is selected to obtain the evaluation index to ensure the completeness of the evaluation system.

AHP operation is divided into four steps: design and build a hierarchical structure model, build a comparison matrix, calculate the weight value, and get the combination weight vector and implement consistency test.

Based on the comprehensive analysis of each evaluation index, the preparation of the curriculum, the effect of moral education, and the content are taken as the first index. See Table 1 for details.

According to the above AHP running steps, we can see that it is necessary to construct the criterion judgment matrix. Based on experience, use 1 to 9 as the corresponding judgment matrix scale [24]. The judgment matrix scale is as follows: when the scale is 1, it means i and j are equally important; when the scale is 3, it means that i is slightly more important than j; when the scale is 5, it means that i is significantly more important than B; when the scale is 7, it means that i is very important than j; when the scale is 9, it means that A is extremely important than j; when the scale is 2, 4, 6, 8, i and j are in the middle of two adjacent judgments [25–27].

4.2. Evaluating the Properties of Matrix Composition Operator. The properties of the synthesis operator are analyzed through the following nonlinear fuzzy matrix [28]:

$$B = A \cdot R = a_n \cdot \begin{matrix} r_{11} & r_{12} & \ldots & r_{1m} \\ r_{21} & r_{22} & \ldots & r_{2m} \\ r_{n1} & r_{n2} & \ldots & r_{nm} \end{matrix}. \quad (1)$$

In the above matrix, A represents the index salience influence coefficient vector, and $A = (\lambda_1, \lambda_2, \ldots, \lambda_n)$, where $\lambda_i \geq 1$. With higher impact of indicator R, λ_i will increase accordingly.

Assuming $\lambda = \max\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$, the nonlinear matrix composition operator is expressed in the form of

$$f(\lambda) = a_1 x_{1^1}^{\lambda} + a_2 x_{2^2}^{\lambda} + \ldots + a_n x_{n^n}^{\lambda}. \quad (2)$$

Because in the nonlinear evaluation, there is generally membership grade $r_{ij} \in [0, 1]$, so in the process of nonlinear matrix synthesis, it is necessary to change the membership grade of the evaluation target matrix so that the membership grade is higher than 1, then the matrix synthesis operator has the following properties:

$$f(\lambda)_i = f(\lambda) r_{ij} + Mc^{\lambda}. \quad (3)$$

In the formula, $M = n(a_1 a_2 \ldots a_n)^{1/n}$. When $na_1 = na_2 = \ldots = na_n$, the equation holds; then $M = 1$, $f(\lambda)_i \geq c^{\lambda}$.

And because $0 < 1/n^{\lambda} < 1$, and the function $f(\lambda)_i = a^x$, at $0 < a < 1$, it is monotonically decreasing, indicating that the higher the λ value, the higher the M value, so the minimum value of $f(\lambda)_i$ is greater [29].

In the nonlinear synthesis matrix, when the influence of all the indexes is different, some indexes have greater influence.

Assuming $x_i < x_i' (i = 1, 2, \ldots, n)$, there is the following expression:

$$f(x_i) < f(x_i'). \quad (4)$$

The above expression is obtained by monotonically decreasing function $f(x_i)$ with respect to X.

$$\lim_{x_i \longrightarrow x_i} f(x_i) = f(\lambda)_i. \quad (5)$$

Formula (5) is successively obtained on $[1, +\infty]^n$ according to function $f(x_i)$ and X.

Assuming that all the evaluation indexes of one evaluation target are larger than other targets, the former evaluation result should be larger than the latter. An increase in the value of a single target indicator leads to an increase in the final indicator value, which remains steady and does not lead to sudden jumps [30].

4.3. Moral Education Evaluation Model of TOP-SIS

4.3.1. Weighted Principal Component TOP-SIS Model. The weighted principal component TOP-SIS value moral education model is selected to evaluate the moral education of a city. The main processes of data screening and statistics are as follows: first, analyze the educational subjects, including schools, society, and families. At the same time,

Table 1: Evaluation index system.

First-level evaluation index	Secondary evaluation index
Course preparation	Master the details of moral education The lecture process is consistent with the schedule
Effect of moral education	Clear learning objectives Cultivate students' self-study ability Cultivate students' ability to analyze problems Cultivate students' application ability
Content of moral education	The lecture is well organized Expand knowledge on the basis of new knowledge The content of moral education is perfect and focused Reasonably arrange students' homework The content of moral education has depth
Professional conduct and ethics	Responsible and patient to answer questions The examination method is scientific and fair Discipline is lax Pay attention to communication with students Pay attention to the problems reflected by students Pay attention to being a teacher
Moral education method	The way of moral education is scientific and appropriate The course explanation is vivid Clean and tidy facilities related to moral education

three teams of big data technology, analysis, and education should be established to collect and summarize data information through the daily learning and life of college students, comprehensively grasp and understand the moral performance of college students, guide college students to establish a correct moral outlook, and abide by the correct TOP-SIS value and moral code of conduct. Second, from the analysis of the form of education, it mainly covers the content of TOP-SIS value and moral education, the practice of TOP-SIS value and moral education, and the networking of TOP-SIS value and moral education.

Based on the above research, the weighted principal component TOP-SIS model is constructed, and the process is as follows:

(1) Select SPSS software to carry out principal component analysis on the secondary indicators moral education factors in the evaluation index system, and according to the principle that the eigenvalue and cumulative contribution rate are greater than 1 and 80%, respectively, the principal component with the quantity of j is represented by Y_{i_j}, $i = 1, 2, 3, 4$, $j = 1, 2, \cdots, n$, and the extracted principal component is set as the secondary indicator of the evaluation model to establish the principal component expression [31, 32]. Input the original data of standardization processing into the principal component expression; get the final principal component score matrix expressed by $Y_{p\times i_j}$, $p = 1, 2, \cdots, 16$.

(2) The normalized principal component decision matrix $Z = \left\{z_{p\times i_j}\right\}$ shall be established by means of vector normalization according to the corresponding principal component scores of each first-grade index. The weight matrix of principal component contribution rate and cumulative contribution rate expressed by $W = (w_1, w_2, \cdots, w_j)^T$ is established, and the weighted canonical matrix expressed by $X = \left\{x_{p\times i_j}\right\}$ is obtained.

(3) Obtain the maximum reasonable value of each level of indicators and the relative proximity between the maximum reasonable value and different schemes through the TOP-SIS method, and use the maximum reasonable value and the relative proximity to evaluate the comprehensive evaluation indicators of moral education.

(4) Obtain the moral education indices of 10 cities of a certain province and the final ranking of each city through the calculation formula of relative moral education [33].

4.3.2. TOP-SIS Value Function Model. The TOP-SIS model of value moral education is often used in the decision analysis of fixed scheme determination. The TOP-SIS model of value moral education is a method to minimize and maximize the distance between positive ideal scheme and negative ideal scheme and the optimal scheme. The TOP-SIS model of value moral education is used to rank many objects to be evaluated which have measurement attributes.

(1) Select the vector normalization method to obtain the standardized decision matrix $Z = \left\{z_{p\times i_j}\right\}$.

(2) The weight and weight coefficient of vector matrix are represented by $W = (w_1, w_2, \cdots, w_j)^T$ and $w_j = \eta_j / \sum_{j=1}^{n} \eta_j$, and the weighted gauge matrix is established and represented by $X = \left\{x_{p\times i_j}\right\}$.

(3) X^+ is the positive ideal solution; X^- is the negative ideal solution and can be obtained; $X_{i_j}^- = \min(x_{i_j})$.

(4) Obtain the distance from different schemes to the positive ideal solution and the distance to the negative ideal solution, and evaluate the pros and cons of each scheme by using the relative proximity of the positive ideal solution to different schemes. The range of relative proximity is [0, 1]; when the value is close to 0 and 1, the scheme is close to the worst level and the optimal level, respectively.

Using the relative proximity of 10 cities in a province to obtain the comprehensive relative proximity and using the comprehensive relative proximity to evaluate the moral education of each city in a province, the higher the comprehensive relative proximity is, the higher the moral education of the city is, otherwise the contrary [34].

4.3.3. Index Correlation Judgment. Before nonlinear evaluation, it is necessary to clarify which indicators in the evaluation objectives have linear correlation, because the indicators with linear relationship not only have no additional information, but also easily bring more redundant data [35, 36].

Combined with the grey correlation analysis method, the pairwise comparison sequence is regarded as two indicators with linear relationship among the objectives to be evaluated. The detailed process is as follows:

(1) Experts study the known assessment indicators and mark the indicators with linear relationship as the target of grey relational grade analysis.

(2) Assume that x_i and x_j are two indicators with linear relationship, and b, x, w represent ideal scheme, evaluated scheme, and negative ideal scheme, respectively. Using the index as the comparison unit, all the indexes are dimensionless. If the distance between the corresponding points is small, the consistency of the change of the series is strong, otherwise weak. So, according to $X_{ij}(k) = |x_i(k) - x_j(k)|$, the distance between indices x_i and x_j is in b, x, w.

(3) Define operation expression:

$$\varepsilon_{ij}(k) = \frac{\left(X_{\min} + \rho X_{\max}\right)}{\left(X_{ij}(k) + \rho X_{\max}\right)}. \tag{6}$$

In the above formulas, $\varepsilon_{ij}(k)$ represents the grey correlation coefficients of indices x_i and x_j in any scheme, $X_{\min}$ and $X_{\max}$ represent the maximum and minimum values of absolute difference, respectively, and ρ is a resolution coefficient, which can control the influence of $X_{\max}$ on data transformation and enhance the difference between the correlation coefficients.

(4) The association coefficient shall be averaged, and the definition of association degree shall be

$$g_{ij} = \frac{1}{3}\sum_{k=1}^{3} \varepsilon_{ij}(k). \tag{7}$$

(5) If the correlation degree is higher than a certain threshold, it indicates that indicators x_i and x_j have great correlation, and one of them can be eliminated.

Accurate value evaluation type: suppose that x represents the evaluation value, b represents the best evaluation value obtained by each ideal value index, and e represents the worst evaluation value obtained by all negative ideal value indexes. Therefore, $(x - w)/(b - w)$ is the optimization degree p_b; $(b - x)/(b - w)$ describes the degree of deterioration p_w. Then the calculation formula of comprehensive score is

$$P_f = 5 \times p_b + p_w. \tag{8}$$

Interval value evaluation type: suppose $[x_1, x_u]$ represents interval evaluation value, and $[b_1, b_u]$ and $[w_1, w_u]$ represent ideal and negative ideal scheme evaluation values, respectively. And x_1, b_1, w_1 represent the lower limit of interval value, x_u, b_u, w_u are the upper limit, and $b_u > w_u$. At this time, the following three situations will occur: the value range of the scheme to be evaluated is all in the ideal value range or negative ideal value range, and there may be intersection with the two ranges. When the third condition occurs, $x_u \in [b_1, b_u]$, $x_1 \in [w_1, w_u]$, or any of the following conditions are met:

$$\begin{aligned} p_b &= \frac{(x_u - b_1)}{(x_u - x_1)}, \\ p_w &= \frac{(w_u - x_1)}{(x_u - x_1)}. \end{aligned} \tag{9}$$

When $w_u < x_i < b_1$, the calculation formula of intermediate rate is

$$p_m = \frac{(b_1 - w_u)}{(x_u - x_1)}. \tag{10}$$

Then the comprehensive score is expressed as

$$M = 5 \times p_b + p_w + 3 \times p_m. \tag{11}$$

4.3.4. Relative Development of Moral Education Calculation. There are great differences in the maximum reasonable values of population, resources, and other indicators in 10 cities in a province. The use of unified evaluation criteria makes the evaluation results more reasonable. Select the unified evaluation index to process the data to obtain the relative score formula of the evaluation index for the development of moral education, as follows:

$$R_{pi} = \frac{F_i^+ - f_{pi}}{F_i^+} = 1 - \frac{f_{pi}}{F_i^+}, \tag{12}$$

where R_{pi} and F_i^+, respectively, represent the standardized score and the maximum reasonable value of the indicators within the first-level indicators of each city in a province; f_{pi} represents the corresponding value of each city in a province in the weighted norm matrix.

The weighted summation formula is as follows:

$$\omega_p = \sum_{i=1}^{n} \eta R_{pi}, \tag{13}$$

where ω_p and η, respectively, represent the comprehensive score of moral education of each city in a province and the primary index weight in the comprehensive evaluation index system of moral education.

4.3.5. Index Contribution Rate Operation. For complex evaluation objects, not all the indicators have the significance of participating in the evaluation. According to the information theory, the function of each index in the evaluation system depends on the amount of decision information it has. The more the information is, the higher the effect is.

The Del entropy method combines the information of all indexes to judge the importance of indexes, that is, entropy weight. Assume that the initial indicator attribute matrix is represented by $D' = (z_{ij})_{m\times n}$, and z_{ij} is the attribute value of Plan i under the j indicator, so the contribution of Plan i to the j indicator attribute p_{ij} is expressed as

$$p_{ij} = \frac{z_{ij}}{\sum z_{ij}, j \in (1, m);\ i = 1, 2, 3}. \tag{14}$$

This contribution rate includes a kind of information that can describe the sum of contribution rates of the three methods to the j index through entropy E_j.

$$E_j = -k \sum_{i=1}^{m} p_{ij} \ln\ p_{ij},\ i = 1, 2, 3. \tag{15}$$

In a formula, k represents a constant, usually a value of $1/\ln 3$, and $E \in [0, 1]$ is guaranteed.

If the index belongs to interval type, the contribution rate and entropy of the interval upper bound and lower bound of the index in all schemes are obtained, and the value of interval entropy is taken as the total contribution rate.

Assuming that the contribution degree of two indicators is basically the same, that is, E_j is close to 1, it indicates that the indicator does not play any role in the decision-making process, and the weight attribute of the indicator is 0; otherwise, if the contribution degree of an indicator is small, it indicates that the information content of the indicator is larger, and the role is larger. When the contribution of the index is higher than 0.85, the index is less important in the decision-making process. The contribution of other indicators can be obtained according to the following formula:

$$\chi = \frac{\left(1 - E_j\right)}{m - \sum E_j}. \tag{16}$$

In the formula, m represents the total number of indicators corresponding to level n indicators.

The final score of the final development of moral education in each city is obtained by using the first-level indicators of each city in a province relative to the moral education index. The level of moral education in a province can be reflected by the final score of moral education. The higher the score obtained, the higher the development of moral education in the education industry of the city.

5. Experimental Design and Test

Bring the original index data of a city into SPSS software for principal component analysis. Due to the different units of measurement of the original indicators, it is impossible to compare directly. Therefore, it is necessary to standardize the original indicator data of 10 cities to eliminate the impact of dimensions on the evaluation results.

5.1. KMO and Bartlett Test. In order to ensure the validity of the empirical analysis results of factor analysis and extraction of comprehensive factors, it is necessary to conduct correlation test on the standardized index data, using the commonly used KMO and Bartlett sphericity correlation test methods, and the test results are shown in Table 2.

Table 2 shows that the final KMO value is 0.643, and the KMO value satisfies the prerequisite requirements for the quality of PCA data, and the PCA can be carried out for the index data; the Bartlett sphericity test result for the index data is $P = 0$, and the index data is verified again to meet the PCA requirements.

5.2. Validation of Accuracy of Principal Component Analysis Results. The obtained component score coefficient matrix results are shown in Table 3.

By analyzing the component score coefficient matrix in Table 3, we have the following.

The formula of component score F_1 is as follows:

$$F_1 = X1 + X2 + X3 + X4 + X5 + X6 + X7 + X8 - X9 + X10. \tag{17}$$

The formula of component score F_2 is as follows:

$$F_2 = -X1 - 0.275 \times X2 - X3 - X4 + X5 + X6 + X7 + X8 + X9 + X10. \tag{18}$$

The formula of component score F_3 is as follows:

$$F_3 = X1 + X2 - X3 - X4 - X5 - X6 + X7 + X8 + X9 + X10. \tag{19}$$

Table 2: KMO and Bartlett's inspection.

KMO value	Bartlett sphericity test		
	Approximate chi square	df	*P* value
0.643	230.307	66	0

Table 3: Component score coefficient matrix.

Name	Component			
	Component 1	Component 2	Component 3	Component 4
X1	0.242	−0.228	0.059	−0.257
X2	0.220	−0.275	0.327	0.006
X3	0.229	−0.076	−0.26	−0.299
X4	0.229	−0.232	−0.228	−0.257
X5	0.226	0.222	−0.247	0.296
X6	0.095	0.264	−0.306	0.375
X7	0.244	0.022	0.002	−0.209
X8	0.047	0.442	0.007	−0.436
X9	−0.224	0.232	0.222	−0.027
X20	0.093	0.308	0.258	−0.002

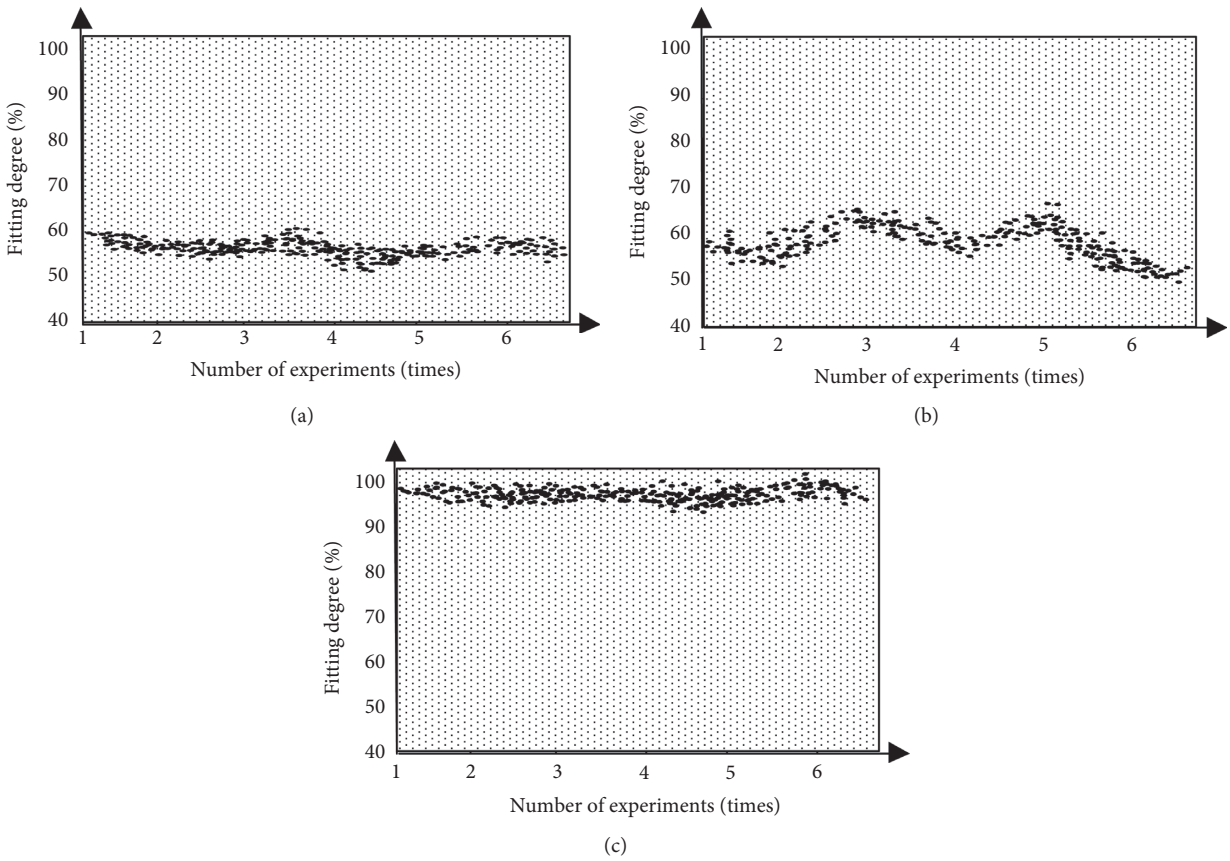


Figure 2: Comparison of the fitting degree between the evaluation results of different models. (a) Education quality evaluation model based on Amos. (b) Educational quality evaluation model based on multiple regression of time series. (c) Designed model.

The formula of component score F_4 is as follows:

$$F_4 = -X1 + 0.006 \times X2 - X3 - X4 + X5 + X6 - X7 - X8 - X9 - X10. \quad (20)$$

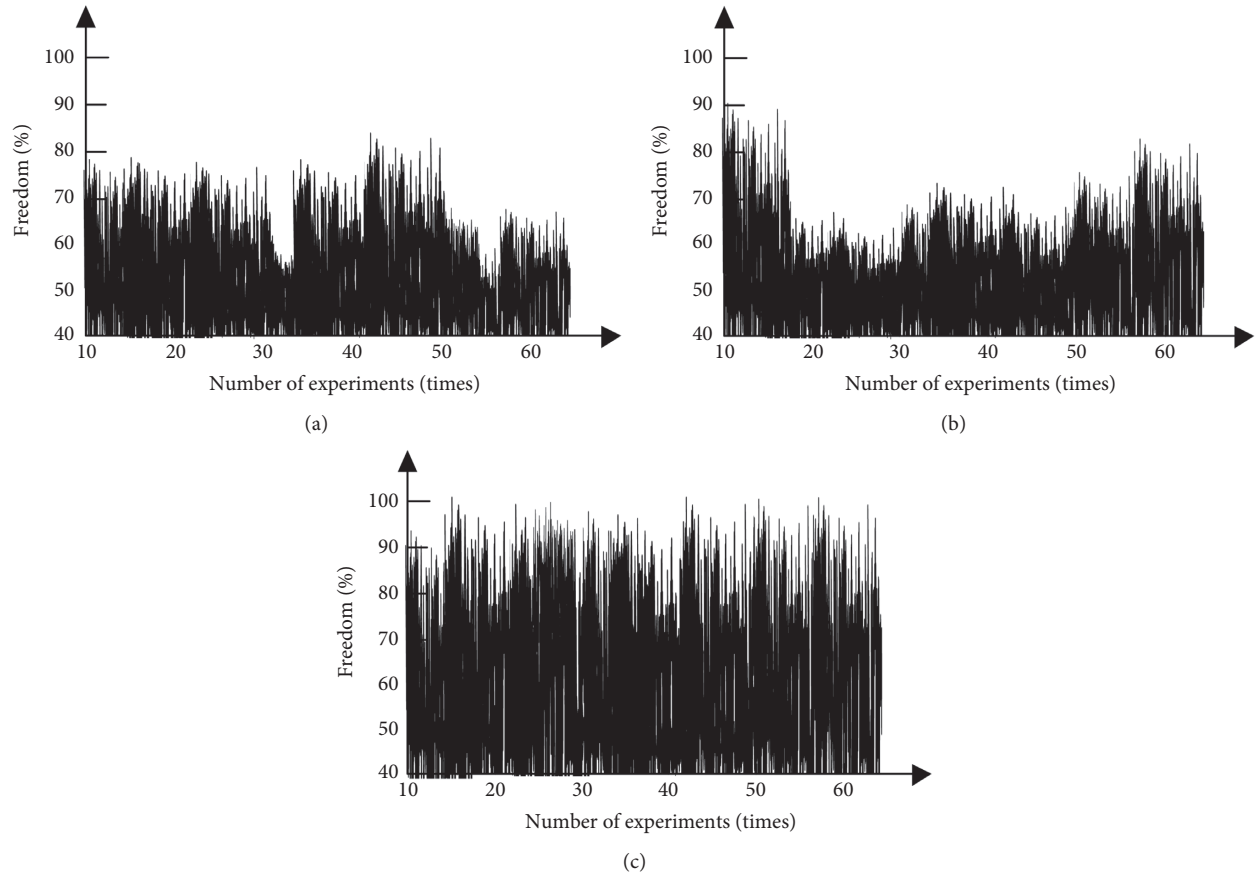


Figure 3: Comparison of degrees of freedom of different models. (a) Education quality evaluation model based on Amos. (b) Educational quality evaluation model based on multiple regression of time series. (c) Designed model.

The comprehensive factor score is obtained by using the linear combination of the original indicators. The comprehensive factor score formula is as follows:

$$F = \sum F_i \varepsilon_i. \tag{21}$$

In formula (21), F is the comprehensive factor, F_i is the main factor score, and ε_i is the score weight. The evaluation equation of moral education obtained by formula (22) is as follows:

$$F = F_1 + F_2 + F_3 + F_4. \tag{22}$$

The single and comprehensive scores of principal components are positive, indicating that the urban moral education is stronger than the average development of a province, and a negative value indicates that the urban moral education is weaker than the average development of 10 cities in a province. The higher the comprehensive score and individual score, the better the city's moral education and the stronger its competitiveness.

5.3. Comparison of Fitting Degree of Different Models. Using the Amos based education quality evaluation model proposed in [4] and the time series multiple regression based education quality evaluation model proposed in [5] as the control experimental group, and integrating the experimental environment and data, the fitting experimental results of different models are shown in Figure 2:

The higher the fitting degree between the evaluation result and the actual situation is, the higher the evaluation precision is. In Figure 2, the accuracy of the model is lower than that of the paper. Based on the data mining technology, the model extracts the valuable knowledge hidden in the data of moral education quality evaluation by using the association rules and assists the education managers to make effective decisions, which provides a reliable support for improving the fitting degree between the evaluation results and the actual situation. In order to set up an objective and accurate evaluation model of moral education quality, a hierarchical and multiobjective evaluation index system is established by using AHP. Selecting AHP to obtain the evaluation index not only ensures the completeness of the evaluation system, but also improves the accuracy of evaluation results to a certain extent [37].

5.4. Comparison of Degrees of Freedom between Different Models. The higher the degree of freedom is, the lower the evaluation complexity is. As can be seen from Figure 3, the moral education level evaluation model constructed in this paper has a higher degree of freedom under different experimental times, but the evaluation complexity is lower than that of literature achievements. The proposed model

uses the fuzzy comprehensive evaluation method to construct the moral education quality evaluation model, which not only fully considers the fault tolerance of the evaluation process, but also effectively improves the degree of freedom of modeling, that is, reduces the complexity of evaluation.

To sum up, the moral education evaluation model based on college students' quality training has a high degree of fit and freedom, which can effectively reduce the complexity of education evaluation and has good applicability.

6. Conclusion

In order to strengthen the moral will and moral concept of college students, an evaluation model of moral education is established based on the analysis of college students' quality cultivation, using data mining method to extract valuable knowledge hidden in the data, to assist education managers to make effective decisions, and improve management standards. According to the evaluation index, the weighted principal component TOP-SIS model is constructed to complete the evaluation of moral education. The following experimental results are obtained:

(1) The moral education evaluation model based on college students' quality training has a high degree of fit between the evaluation results and the actual situation, and the evaluation accuracy is high.

(2) The model fully considers the fault tolerance of the evaluation process and effectively improves the degree of freedom of modeling, and the evaluation complexity is low, which verifies the practicability of the model.

Data Availability

The raw data supporting the conclusions of this article will be made available by the author, without undue reservation.

Conflicts of Interest

The author declares that there are no conflicts of interest regarding this work.

Acknowledgments

This work was supported by Hainan Provincial Natural Science Foundation of China "Research on the cognitive and emotional processing mechanism of moral judgment" (Project no. 719MS056).

References

[1] X. R. Wu, B. Q. Liu, and T. T. Yuan, "A new generation of smart class: concept, platform and system Architecture," *China Educational Technology*, vol. 3, pp. 81–88, 2019.

[2] S. K. Mydhili, S. Periyanayagi, and S. Baskar, "Machine learning based multi scale parallel K-means++ clustering for cloud assisted internet of things," *Peer-to-Peer Networking and Applications*, vol. 13, no. 2, pp. 2023–2035, 2020.

[3] T. Yilmaz, R. Ozcan, I. S. Altingovde, and Ö. Ulusoy, "Improving educational web search for question-like queries through subject classification," *Information Processing & Management*, vol. 56, no. 1, pp. 228–246, 2019.

[4] Y. Z. Luo, X. M. Chen, and J. L. Zhang, "Evaluation of doctoral education quality:model construction and empirical test," *Research in Higher Education of Engineering*, vol. 2, pp. 171–177, 2019.

[5] J. L. Tu and Q. X. Tao, "Relevance analysis and evaluation model of innovation and entrepreneurship education resources and service function of science and technology innovation in local universities," *Science & Technology Progress and Policy*, vol. 37, no. 8, pp. 159–166, 2020.

[6] H. J. Wang and N. Zhang, "Construction of "Evidence-based" Academic assessment model for blended learning courses in higher education," *E-education Research*, vol. 41, no. 5, p. 7, 2020.

[7] S. J. Andriole, "Social media analytics, wearable technology, and the internet-of-things," *IT Professional*, vol. 21, no. 5, pp. 11–15, 2019.

[8] M. D. Abdulrahaman, N. Faruk, A. A. Oloyede et al., "Multimedia tools in the teaching and learning processes: a systematic review," *Heliyon*, vol. 6, no. 11, Article ID e05312, 2020.

[9] Z. Chen and N. Polytechnic, "Research on the new characteristics of ideological and political education in colleges under the background of network self-media," *Journal of Ningbo Polytechnic*, vol. 55, no. 13, pp. 560–572, 2019.

[10] Q. H. Zheng, B. Dong, B. Y. Qian et al., "The state of the art and future tendency of smart education," *Journal of Computer Research and Development*, vol. 56, no. 1, pp. 213–228, 2019.

[11] D., A. Satarupa, "Smart university utilising the concept of the internet of things(IoT)," *Neurochemistry International*, vol. 129, no. 10, p. 2022, 2019.

[12] H. Tanveer, T. Balz, N. S. Sumari, R.-u. Shan, and H. Tanweer, "Pattern analysis of substandard and inadequate distribution of educational resources in urban-rural areas of Abbottabad, Pakistan," *Geojournal*, vol. 85, no. 5, pp. 1397–1409, 2020.

[13] T. L. Friesz, "Socio-economic determinants of student mobility and inequality of access to higher education in Italy," *Networks and Spatial Economics*, vol. 19, no. 1, pp. 125–148, 2019.

[14] V. Mkrttchian, L. Gamidullaeva, A. Finogeev, and S. Chernyshenko, "Big data and internet of things (IoT) technologies' influence on higher education: current state and future prospects," *International Journal of Web-Based Learning and Teaching Technologies*, vol. 16, 2021.

[15] C. Wiley, "Standardised module evaluation surveys in UK higher education: establishing students' perspectives," *Studies In Educational Evaluation*, vol. 61, pp. 55–65, 2019.

[16] K. Khosravi, J. R. Cooper, P. Daggupati, B. Thai Pham, and D. Tien Bui, "Bedload transport rate prediction: application of novel hybrid data mining techniques," *Journal of Hydrology*, vol. 585, no. 6, Article ID 124774, 2020.

[17] N. W. Borsato, S. L. Martell, and J. D. Simpson, "Identifying stellar streams in Gaia DR2 with data mining techniques," *Monthly Notices of the Royal Astronomical Society*, vol. 492, no. 1, pp. 1370–1384, 2020.

[18] K. Snehal, S. Gaikwad, S. Iyer, and D. Lunga, "Proceedings of KDD 2020 workshop on data-driven humanitarian mapping: harnessing human-machine intelligence for high-stake public policy and resilience planning," *Papers*, vol. 9, Article ID 210900435, 2021.

[19] S. Pyne and T. Laskey, "Broadening the cybercartographic research and education network," *Cybercartography in a*

Reconciliation Community - Engaging Intersecting Perspectives, vol. 8, pp. 217–236, 2019.

[20] M. Dimitris, K. Georgoulias, and G. Chryssolouris, "The Teaching Factory Network: a new collaborative paradigm for manufacturing education," *Procedia Manufacturing*, vol. 97, no. 7, pp. 120–132, 2019.

[21] W. L. Shi and Y. H. Zhang, "Intelligent education platform design based on big data analysis technology," *Modern Electronics Technique*, vol. 43, no. 9, pp. 158–161, 2020.

[22] J. Lai and M. Bower, "Evaluation of technology use in education: findings from a critical analysis of systematic literature reviews," *Journal of Computer Assisted Learning*, vol. 36, no. 3, pp. 1459–1471, 2020.

[23] G. Bodur, S. Gumus, and N. G. Gursoy, "Perceptions of Turkish health professional students toward the effects of the internet of things (IOT) technology in the future," *Nurse Education Today*, vol. 79, pp. 98–104, 2019.

[24] D. A. Tamburri and G. Casale, "Cognitive distance and research output in computing education: a case-study," *IEEE Transactions on Education*, vol. 62, no. 2, pp. 99–107, 2019.

[25] S. Pundir, M. Wazid, D. P. Singh, A. K. Das, J. J. P. C. Rodrigues, and Y. Park, "Intrusion detection protocols in wireless sensor networks integrated to internet of things deployment: survey and future challenges," *IEEE Access*, vol. 8, pp. 3343–3363, 2020.

[26] A. Gee and H. Abbass, "Transparent machine education of neural networks for swarm shepherding using curriculum design," *International Joint Conference on Neural Networks*, vol. 67, no. 13, pp. 57–69, 2019.

[27] G. R. Browne, H. Bender, J. Bradley, and A. Pang, "Evaluation of a tertiary sustainability experiential learning program," *International Journal of Sustainability in Higher Education*, vol. 35, no. 5, pp. 2–13, 2020.

[28] K. Moriarty, "Collective impacts on a global education emergency: the power of network response," *Prospects*, vol. 49, no. 1, pp. 81–85, 2020.

[29] J. O'Hara, M. Brown, G. Mcnamara, and P. Shevlin, "The potential, limitations and evaluation of education networks in a monocentric system," *Revista de Investigación Educativa*, vol. 44, no. 1, pp. 34–57, 2020.

[30] L. I. Ye, "Research on the existing problems and solutions of network ideological and political education in colleges and universities," *Journal of Hlongjiang College of Education*, vol. 56, no. 1, pp. 781–792, 2019.

[31] Ö. K. Kalkan, A. Altun, and B. Atar, "Role of teacher-related factors and educational resources in science literacy: an international perspective," *Studies In Educational Evaluation*, vol. 67, Article ID 100935, 2020.

[32] A. Ghaleb, "Rasch validation of the Arabic version of the teacher efficacy for inclusive practices (TEIP) scale - ScienceDirect," *Studies In Educational Evaluation*, vol. 62, pp. 104–110, 2019.

[33] S. Wild and L. Schulze Heuling, "Re-evaluation of the D21-digital-index assessment instrument for measuring higher-level digital competences," *Studies In Educational Evaluation*, vol. 68, no. 1, Article ID 100981, 2021.

[34] A. Liqiong, Y. B. Jing, B. Xiaoyan, H. Luoya, S. Shaomei, and Z. Yan, "Based on Delphi method and Analytic Hierarchy Process to construct the Evaluation Index system of nursing simulation teaching quality," *Nurse Education Today*, vol. 79, pp. 67–73, 2019.

[35] I. Klima Ronen, "Action research as a methodology for professional development in leading an educational process," *Studies In Educational Evaluation*, vol. 64, Article ID 100826, 2020.

[36] C. Spoden, J. Fleischer, and A. Frey, "Person misfit, test anxiety, and test-taking motivation in a large-scale mathematics proficiency test for self-evaluation," *Studies In Educational Evaluation*, vol. 67, Article ID 100910, 2020.

[37] S. A. Wind, E. Jones, C. Bergin, and K. Jensen, "Exploring patterns of principal judgments in teacher evaluation related to reported gender and years of experience," *Studies In Educational Evaluation*, vol. 61, pp. 150–158, 2019.

Hindawi
Scientific Programming
Volume 2022, Article ID 7940755, 12 pages
https://doi.org/10.1155/2022/7940755

Research Article

Research on the Allocation Method of Regional Science and Technology Resources from the Perspective of Rationality

Huozhong Zhang[1] and Yongjun Zhou[2]

[1]*Minnan Science and Technology University, Business College, Quanzhou, Fuzhou 360000, China*
[2]*Fujian Normal University, Organization Department, Fuzhou, Fujian 350000, China*

Correspondence should be addressed to Yongjun Zhou; zyj@mku.edu.cn

Received 8 November 2021; Revised 29 November 2021; Accepted 14 December 2021; Published 7 January 2022

Academic Editor: Baiyuan Ding

At present, the allocation efficiency of regional scientific and technological resources is low, and there are few research studies on social equity and economic efficiency. Therefore, this paper puts forward the allocation method of regional scientific and technological resources based on rationality. With the support of rationality perspective, the evaluation model of reform path of regional scientific and technological resource allocation is constructed to analyze the economic benefits and equity benefits of regional scientific and technological resource allocation. According to the principle of optimum allocation of regional science and technology resources, three-dimensional structure is constructed to maximize national investment and benefit and determine the optimal Pareto of resource allocation to measure the efficiency of resource allocation. The evaluation index system of the reform path of resource allocation is constructed by selecting the evaluation index of the reform path of resource allocation. The benchmark platform of big data was selected to generate data sets to be processed, and the spark on yam platform was used to submit jobs and generate spark job running data sets. The operation performance prediction model was established to optimize the configuration parameters of regional science and technology resources. The analysis results show that the designed method has high configuration capability and good effectiveness.

1. Introduction

Science and technology are the primary productive forces, and scientific and technological resources are the material basis of scientific and technological creative labor. As the primary resource, the allocation of scientific and technological resources has attracted more and more attention [1, 2]. Therefore, it is necessary to comprehensively evaluate the allocation efficiency of this specific productivity factor and its regional differences [3]. Countries all over the world have established their own STR sharing platform, which can match various STR services and needs in a lower cost and more convenient way and create the interaction between supply and demand [2]. These sharing platforms make it possible to realize the sharing economy in the field of scientific and technological innovation. The efficiency of scientific and technological input-output is one of the important standards to measure a region's scientific and technological strength. Therefore, using reasonable methods to analyze and study the efficiency of regional scientific and technological resource allocation and putting forward countermeasures and suggestions according to the research results is of great significance to improve regional scientific and technological strength and promote economic development.

In this regard, some scholars have used different methods to quantitatively study the efficiency of scientific and technological resource allocation. Reference [4] puts forward the three-dimensional operation mechanism and system dynamics simulation of regional scientific and technological resource allocation system. A new framework for the operation mechanism of STR distribution system is constructed, and its three submechanisms are designed. The system dynamics model of the operating mechanism is designed and experimentally studied. Simulation results show that the proposed mechanism can effectively improve the sharing rate and value. Finally, some policy suggestions are put forward to help the government improve the management

mode and optimize the allocation of resources. Reference [5] puts it forward in addition to multilayer online resource allocation and offline planning and configuration in man. The network supporting software defined network and network function virtualization needs to make decisions on two time scales: short-term online resource allocation and medium- and long-term offline planning. In this paper, we first discuss the scale of man supporting software defined network and network function virtualization, with particular attention to the role of delay in capacity planning. A delay-aware multilayer service chain allocation algorithm is proposed to explore a series of maximum delay requirements and their impact on man scale resources. As more data center facilities need to be expanded to approach the edge of the network, reducing the economies of scale of the IT infrastructure, the design cost of low latency requirements will increase. This paper reviews the recent joint calculation of multisite VNF layout and multilayer resource allocation when deploying network services in metropolitan area networks. A set of subroutines included in the multilayer service chain allocation are experimentally verified in the network optimization as a service architecture, which can help the metro network test the open source MANO instance, virtual infrastructure manager, and WAN controller in the mosaic platform.

Most scholars only use a single evaluation method to evaluate the allocation efficiency of regional scientific and technological resources, but there are often differences in the results obtained by using different methods to evaluate the same object. At the same time, each method has its own advantages and disadvantages, and it is impossible to explain which method is good or bad. It is undoubtedly one-sided to use only one method for evaluation. If the first mock exam is the first mock exam, the first mock exam will be able to make use of more information. It can make full use of more information and make full use of the advantages of a single model to make up for the deficiency of single model and make the comprehensive evaluation results more scientific and reasonable. Supported by the rationality perspective, the research method constructs the path evaluation model of regional science and technology resource allocation reform and analyzes the economic and fair benefits of regional science and technology resource allocation. Build a three-dimensional architecture, determine the optimal resource allocation Pareto, and measure the efficiency of resource allocation. Select the evaluation index of resource allocation reform path and build the evaluation index system of resource allocation reform path. Establish spark operation performance prediction model and optimize regional science and technology resource allocation parameters. The results show that the design method has good configuration effect.

2. Evaluation Index System and the Model of Regional Scientific and Technological Resource Allocation Method Reform Path

2.1. Constructing the Evaluation Index System of the Reform Path of Regional Science and Technology Resource Allocation. Establishing a reasonable and feasible path evaluation system for regional science and technology resource allocation reform can select an appropriate evaluation index for regional science and technology resource allocation data to reduce the evaluation error of regional science and technology resource allocation reform path [3, 6]. First, clarify the criteria and basis for the establishment of the system. When designing the evaluation index, according to the principle of system establishment, we can improve the scientificity of the evaluation system of the path of regional scientific and technological resource allocation reform [7]. The construction principle of the path evaluation system of regional science and technology resource allocation reform should not only meet the theoretical design requirements, but also consider the practical experience of regional science and technology resource allocation. Based on the rationality perspective, it is summarized that the following four principles should be followed in the construction of the path evaluation index system of regional science and technology resource allocation reform. They are the principle of combining comparability and operability, the principle of combining systematicness and comprehensiveness, the principle of combining dynamics and foresight, and the principle of combining objectivity and purpose [8, 9].

The rationality perspective mainly analyzes the evaluation indicators of regional science and technology resources and selects the appropriate indicators for the evaluation of the path of regional science and technology resource allocation reform [10, 11]. On the basis of the rational perspective, the assessment can reflect the regional science and technology resources allocation reform path of key contents of fairness, mainly from the human, material, and financial resources and other resources' evaluation index of regional science and technology resources allocation reform path choice; the object of the evaluation index system refers to the planned fixed scientific and technological resource object for which market analysis, technical analysis, and economic analysis are carried out. It is mainly based on the review and reevaluation of the feasibility study conclusion of the planned regional scientific and technological resource project. It is the evaluator's final judgment on the feasibility study from a long-term and objective perspective, forming a path evaluation system for the reform of regional scientific and technological resource allocation from the perspective of rationality, as shown in Table 1.

In Table 1, the data represent the evaluation coefficients. The evaluation coefficient is the data to measure the fluctuation of the reform path of regional science and technology resource allocation from the perspective of human resources, material resources, and financial resources. According to the design requirements of the path evaluation model of regional science and technology resource allocation reform, when constructing the path evaluation index system of regional science and technology resource allocation reform, the appropriate construction principles are adopted to complete the construction of the path evaluation index system of regional science and technology resource allocation reform [12, 13]. By evaluating the path indicators of scientific and technological resource allocation reform in different regions, the analysis of the path evaluation model of regional scientific and technological resource allocation reform based on the perspective of rationality is realized.

Table 1: Evaluation index system of regional scientific and technological resource allocation reform path.

Primary index	Secondary index	Tertiary indicators	Evaluation coefficient
Select evaluation indicators from the perspective of human resources	Regional science and technology resource allocation management personnel	Proportion of total managers	0.1360
		Education of management personnel	0.1922
		Administrative level of managers	0.1751
	Regional science and technology resource allocation instructor	Number of instructors	0.1140
		Instructor structure	0.1694
		Instructor level	0.1504
	Auxiliary personnel for allocation of regional scientific and technological resources	Proportion of total auxiliary personnel	0.1922
Select evaluation indicators from the perspective of material resources	Area of regional science and technology resource allocation site	Area of various scientific and technological resources such as government and enterprises	0.1543
		Area of scientific research institutes, intermediary organizations, and other scientific and technological resources	0.1359
	Allocation of regional scientific and technological resources venue facilities	Quantity of various scientific and technological resources such as government and enterprises	0.1869
		Number of special funds for regional science and technology resource allocation allocated by the government per capita	0.1300
			0.1542
Select evaluation indicators from the perspective of financial resources	Government financial allocation	Number of special funds allocated by the government to allocate regional science and technology resources	0.1349
	Social sponsorship and fund raising	Amount of social sponsorship and fund raising for allocation of regional scientific and technological resources per capita	0.1849
	Capital investment	Number of funds per capita absorbed for allocation of regional scientific and technological resources	0.1092

2.2. Constructing the Path Evaluation Model of Regional Science and Technology Resource Allocation Reform. With the in-depth development of regional scientific and technological resource allocation, how to improve the evaluation quality of scientific and technological resource reform path has attracted more and more attention. Select appropriate evaluation indicators to reduce the evaluation error of the path of regional scientific and technological resource allocation reform by evaluating the path indicators of scientific and technological resource allocation reform in different regions [14]. Firstly, collect the information of the regional science and technology resource allocation server, calculate the weight of the server state, and compare it with the internal weight. When the difference value is less than the set threshold, the weight inside the server does not need to be updated; otherwise, the weight inside the server needs to be updated. The new weight is used to allocate regional scientific and technological resources to achieve the best allocation effect. The specific allocation methods are as follows. Based on the analysis of the application principle of rationality, the rationality perspective principle is adopted to decompose the reform path of regional science and technology resource allocation layer by layer, so that the meaning of elements at each level is complementary and cross, the upper and lower elements are parent-child relations, and the lower elements are used as evaluation indicators [15, 16]. The hierarchical structure of the evaluation of the reform path of regional science and technology resource allocation is shown in Figure 1.

According to the hierarchical structure of the evaluation of the reform path of regional science and technology resource allocation, the combination weight of the bottom elements in the evaluation index system is calculated, and the weight of the lower elements of adjacent levels to the upper elements is calculated by the accurate method [17], and then the combination weight of the bottom elements to the reform path of regional science and technology resource allocation is calculated. The calculation formula is

$$W = (w_1, w_2, w_3, \ldots, w_n)^T. \quad (1)$$

Based on the combination weight of the underlying elements to the reform path of regional science and technology resource allocation, the evaluation coefficient of the reform path of regional science and technology resource allocation is calculated as follows:

$$n_J^{(A)} = \sum_{i=1}^{k} b_i^{(A)}. \quad (2)$$

Here, A represents the evaluation index, J represents the evaluated person, and $b_i^{(A)}$ represents the evaluation coefficient of the path of regional scientific and technological resource allocation reform. For the evaluation index A of the reform path of regional science and technology resource allocation, the total evaluation coefficient $C_{\mathrm{JK}}^{(A)}$ of the evaluated person belonging to each category of rationality perspective is

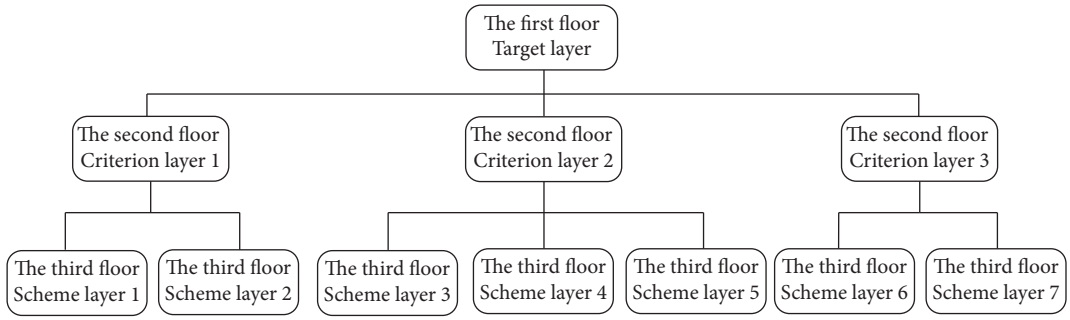


Figure 1: Hierarchical structure of path evaluation of regional science and technology resource allocation reform.

$$C_{\mathrm{JK}}^{(A)} = \sum_{I=1}^{i} f_K\left(d_{\mathrm{JI}}\right). \tag{3}$$

Here, $C_{\mathrm{JK}}^{(A)}$ represents the total evaluation coefficient, d_{JI} represents the weight of each evaluation index, and f_K represents the weight of a single index.

According to the total evaluation coefficient, the evaluation path weight vector and weight matrix of regional scientific and technological resources allocation are calculated [18]. $n_J^{(A)}$ and $C_{\mathrm{JK}}^{(A)}$ can directly calculate the evaluation weight $r_{\mathrm{JK}}^{(A)}$ and weight vector $r_J^{(A)}$ of the j-th evaluator in the evaluation index A belonging to the k-th rationality perspective:

$$r_{\mathrm{JK}}^{(A)} = \frac{C_{\mathrm{JK}}^{(A)}}{n_J^{(A)}}. \tag{4}$$

Here, $K = 1, 2, 3, \ldots, k$; then the weight line vector $r_{\mathrm{jk}}^{(A)}$ of rationality perspective evaluation can be obtained:

$$r_{\mathrm{jK}}^{(A)} = \left[r_{j1}^{(A)}, r_{j2}^{(A)}, r_{j1}^{(A)}, \ldots, r_{\mathrm{jk}}^{(A)}\right]. \tag{5}$$

Considering $J = 1, 2, 3, \ldots, j$, the weight column vector $r_{\mathrm{JK}}^{(A)}$ of rationality perspective evaluation can be obtained:

$$r_{\mathrm{JK}}^{(A)} = \left[r_{1k}^{(A)}, r_{2k}^{(A)}, r_{3k}^{(A)}, \ldots, r_{\mathrm{jk}}^{(A)}\right]^T. \tag{6}$$

Thus, the Atlas evaluation weight matrix $R^{(A)} = [r_{\mathrm{JK}}^{(A)}]$ of the evaluators for the evaluation index A of the path of regional scientific and technological resource allocation reform can be obtained:

$$R^{(A)} = \begin{bmatrix} r_{11}^{(A)} & r_{12}^{(A)} & r_{13}^{(A)} & \cdots & r_{1k}^{(A)} \\ r_{21}^{(A)} & r_{22}^{(A)} & r_{23}^{(A)} & \cdots & r_{2k}^{(A)} \\ \cdots & & & & \\ r_{j1}^{(A)} & r_{j2}^{(A)} & r_{j3}^{(A)} & \cdots & r_{\mathrm{jk}}^{(A)} \end{bmatrix}. \tag{7}$$

Use the Atlas evaluation weight matrix to complete the evaluation of the evaluation index of the path of regional scientific and technological resource allocation reform, which is obtained by $R^{(A)}$:

$$r_J^{*(A)} = \max_K \left\{r_{\mathrm{JK}}^{(A)}\right\}. \tag{8}$$

Thus, the evaluation weight vector of the path of regional scientific and technological resource allocation reform can be obtained:

$$r^{*(A)} = \left\{r_1^{*(A)}, r_2^{*(A)}, r_3^{*(A)}, \ldots, r_j^{*(A)}\right\}. \tag{9}$$

Here, $r^{*(A)}$ represents the evaluation vector. By synthesizing all evaluation factors, determining the Atlas category of the evaluator, and automatically arranging $r^{*(A)}$ into a matrix, the comprehensive evaluation vector of the evaluator can be obtained:

$$r_J = \sum_{K=1}^{k} B_K \times R_{\mathrm{JK}}. \tag{10}$$

Here, B_K represents the weight coefficient of different maps, and the specific value can be determined before evaluation; R_{JK} refers to the total evaluation right of the evaluated person rated as different map categories.

Based on the above calculation, the periodicity of the regional science and technology resource allocation server node is based on the collection of regional science and technology resources. According to the calculated weight of the resource allocation load, it is transmitted to the scheduler. The specific derivation process is shown in Figure 2.

According to the above process, in order to meet the evaluation model of regional science and technology resources allocation reform path design requirements, on the basis of the rational perspective, we summed up the evaluation index system of regional science and technology resource allocation reform path construction principles, and from the aspects of human, material, and financial resources, the evaluation index of regional science and technology resources allocation reform path to choose completed the construction of the evaluation index system of the reform path of regional science and technology resource allocation. According to the hierarchical structure of the evaluation of regional science and technology resources allocation reform path, the computing system of regional science and technology resource allocation reform path evaluation coefficient, using the weight vector and weight matrix, evaluation of the different areas of science and technology resource allocation reform path index, completed the evaluation model of regional science and technology resources allocation reform path construction, realizing the evaluation of

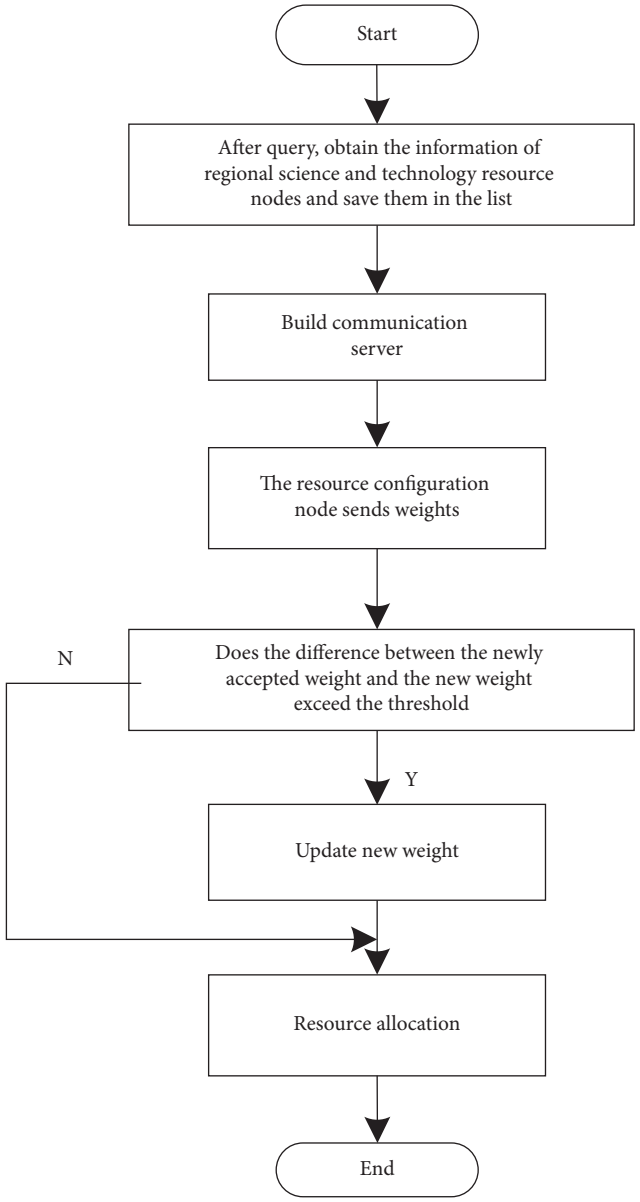


Figure 2: Construction flowchart of the path evaluation model for regional science and technology resource allocation reform.

regional science and technology resources allocation reform path analysis.

3. Spark Operation Performance Prediction Model for Regional Science and Technology Resource Allocation under the Market Economy System

Under the mechanism of market economy, individuals and social organizations are the main trading media, which plays a key role in the allocation of regional scientific and technological resources [19–21]. On the basis of perfecting the basic competitive links, the mechanism studies the specific system and plays a regulatory role in the allocation of regional scientific and technological resources.

For the research on the allocation efficiency of regional scientific and technological resources, build a market economic system framework, as shown in Figure 3.

Give full play to the positive role of economic analysis in optimizing the allocation of regional scientific and technological resources, study its characteristics and relevant mechanisms, ensure the fair distribution of regional scientific and technological resources, take into account economic interests, and achieve rational allocation.

3.1. Analysis of Economic Benefits. The allocation of regional scientific and technological resources requires a certain cost, which is prone to economic problems [22]. At the economic level, the allocation of regional scientific and technological resources should follow the principle of maximizing benefits at the lowest cost. Combined with local reality, whether the allocation of scientific and technological resources reaches the best state, comprehensively evaluate the allocation efficiency of regional scientific and technological resources and increase resources on the premise of ensuring that it does not harm the economic interests of others [23, 24].

3.2. Analysis of Equity Benefit. One of the important issues in analyzing the allocation of regional science and technology resources of local governments from an economic point of view is fairness. Under the market economy system, under the conditions of equality and harmony, the distribution of regional scientific and technological resources also needs to be fair [25, 26]. In view of the large gap between urban and rural areas, local governments should actively formulate relevant policies, give full play to the role of government mechanisms, standardize the market environment, and stop the vicious price rise in the allocation of regional scientific and technological resources.

3.3. Research on the Allocation Efficiency of Regional Science and Technology Resources of Local Government. With the support of market economy mechanism, analyze the problems of economic benefits and fair benefits, and determine the optimal Pareto of regional scientific and technological resource allocation on the premise of ensuring that it does not harm the economic interests of others, so as to realize the rational allocation of regional scientific and technological resources by local governments.

3.3.1. Building a Three-Dimensional Architecture. Under the market economy system, the allocation efficiency of regional science and technology resources of local governments is studied in a three-dimensional structure. The local government that allocates regional scientific and technological resources is the subject of responsibility and management [27]. According to the principle of optimal allocation of resources, local resources are allocated in order to maximize national investment and benefits. It breaks the current practice of setting according to administrative divisions, greatly reduces the number, improves the construction

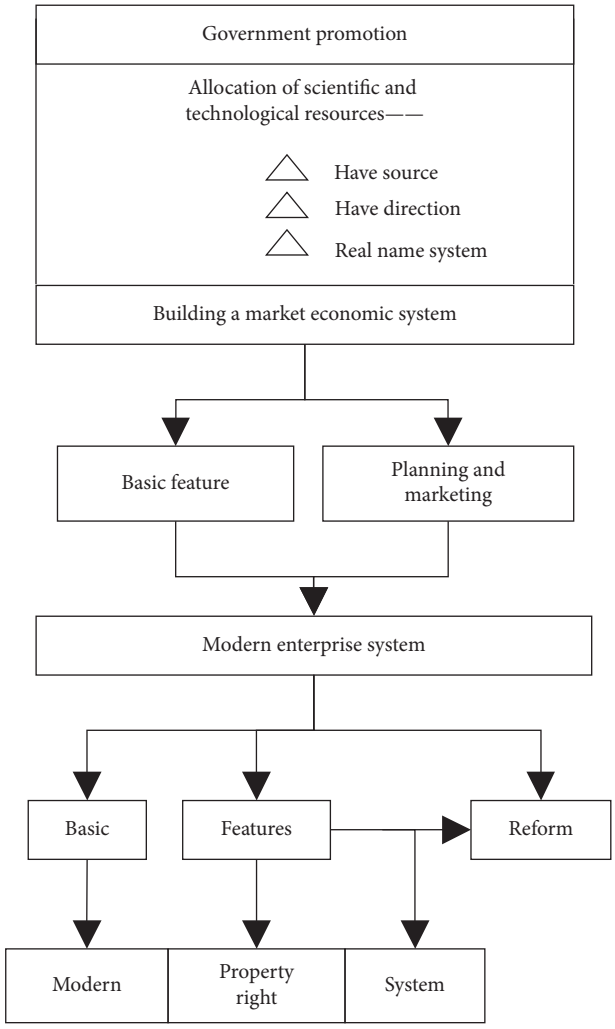


Figure 3: Institutional framework of market economy.

standards, and leaves a certain space for the development of local governments [28, 29]. Developing and applying advanced means representing the industry level and stabilizing market prices are the main functions of local governments in allocating regional scientific and technological resources [30]. The local government will also return this part of the tax to the insurance fund. The insurance institution can be a commercial operation. While ensuring its own development, on the one hand, it accepts social individuals to pay insurance premiums, mainly the quota fund of the local government, which is used to allocate local regional science and technology resources.

3.3.2. Pareto Optimality. On the premise of investing in the market economic system, the allocation of regional science and technology resources of local governments reaches Pareto optimal. Because of the particularity of regional science and technology resources, we can choose the projects and plans that can achieve specific goals and have the lowest cost and change the cost-benefit analysis method to the cost-effectiveness analysis method to measure the efficiency of resource allocation.

It is assumed that the two allocation modes of regional scientific and technological resources are X and Y, and the given quantities are X1 and Y1. National science and technology parks and local science and technology parks are A and B, respectively. Figure 4 shows the resource distribution between the two science parks.

The horizontal length of the block diagram represents the number X1 of the first type of regional science and technology resource allocation X, and the vertical height represents the number Y1 of regional science and technology resource allocation Y. OA is the origin of national science park A, and local science park B is the origin of OB. The consumption of X in the figure from left to right represents the regional science and technology resource allocation X of national science and technology park A, while the vertical figure represents the science and technology resource allocation Y of public medical institution A. From the OB level to the left, it represents the consumption of private science and technology parks in regional science and technology resource allocation X and XB, and from the vertical direction it represents the consumption of private science and technology park B in allocation Y and YB. For any point, such as point A, it is equivalent to the cost of national science park A and local science park B [31]. Therefore, the following formula holds:

$$\begin{aligned} XA + XB &= 1, \\ YA + YB &= 1. \end{aligned} \tag{11}$$

Determine a group of quantities at any point in the block diagram, which represents the consumption of each scientific and technological resource allocated by national and local science and technology parks and conforms to the above formula.

When the transaction reaches a certain state, any regional science and technology resource allocation transaction will reduce the satisfaction of at least one person, so this state is the best state of the transaction [32, 33]. In terms of economic benefits, the allocation efficiency of scientific and technological resources in this region is the highest.

In view of a series of problems in the development of the research method of allocation efficiency supported by theory, this paper studies the efficiency of the government's allocation of regional scientific and technological resources under the local market economy system from the perspectives of social equity and economic efficiency and briefly analyzes the government mechanism and the role of market regulation.

3.4. Spark Job Performance Prediction Model

3.4.1. Spark Job Performance. Spark operation performance formula is as follows:

$$\text{perf} = F(p, d, r, c). \tag{12}$$

In formula (12), perf and p represent the job operation performance and the job itself, respectively, and the spark job operation time is used to reflect the job operation

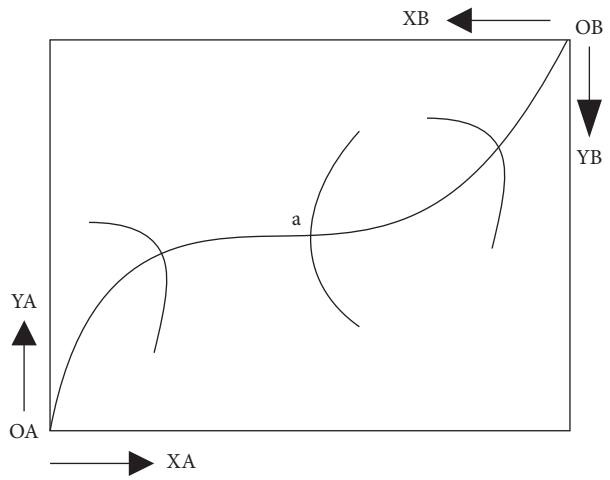


Figure 4: Distribution of regional science and technology resources between the two science and technology parks.

performance; d and r, respectively, represent the data scale to be processed by the job and the available hardware resources of the cluster; c indicates the configuration parameters of regional science and technology resources when submitting the operation. This paper studies the impact of regional science and technology resource allocation parameter c and data scale d of job p on job performance when hardware resource r is sufficient.

3.4.2. Job Acquisition and Preprocessing. Select the big data benchmark platform to generate the data set to be processed and implement sampling in the regional science and technology resource allocation parameter space; the spark on yam platform is used to submit jobs. The spark on yam platform is based on the management mechanism of yarn for the whole big data benchmark platform, analyzes the relationship between the resource manager yarn and the computing engine spark, and studies how to monitor the spark task in the big data platform by monitoring yarn, so as to ensure the availability of the whole spark on yam platform. This generates the spark job run dataset.

The structure diagram of spark platform acquisition operation and preprocessing operation is shown in Figure 5.

There are great differences in the value range of science and technology resource allocation parameters in each region of spark platform. Before establishing the operation performance prediction model, the values of science and technology resource allocation parameters in each region need to be normalized. The formula of normalized values is as follows:

$$C_j = \frac{C_{j,\text{ori}} - C_{j,\min}}{C_{j,\max} - C_{j,\min}}. \tag{13}$$

In formula (13), $C_{j,\text{ori}}$ and $C_{j,\max}$, respectively, represent the original value of regional science and technology resource allocation parameters and the maximum value of all values; $C_{j,\min}$ refers to the minimum value of all science and technology resource allocation parameters in the region. All

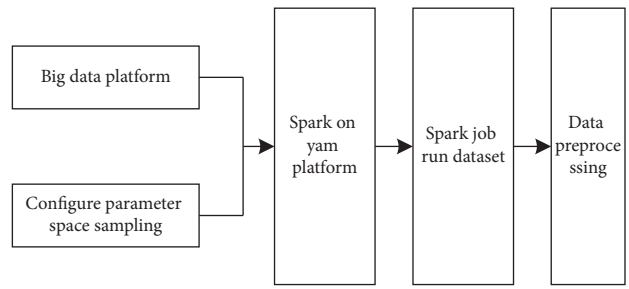


Figure 5: Spark job acquisition and preprocessing.

values are converted to the same data range through the normalization of regional science and technology resource allocation parameters [34]. The data in the same order of magnitude range can effectively improve the comparability of difference characteristic data.

3.4.3. Establishment of Job Performance Prediction Model. The support vector machine algorithm is selected to establish the spark operation performance prediction model. The vector formed by the spark operation data scale and regional science and technology resource allocation parameters is taken as the independent variable, the spark operation time vector is taken as the dependent variable, and the radial basis function $K(x_i, x)$ is selected as the kernel function trained by the support vector regression algorithm. The formula is as follows:

$$K(x_i, x) = \exp\left(-\frac{\|x_i - x\|^2}{2\sigma^2}\right). \tag{14}$$

By mapping data from low dimensional space to high dimensional space through formula (14), nonlinear data can be effectively processed.

It is necessary to normalize and preprocess the original data set, divide the preprocessed original data set, divide the data in the original data set into training set and test set in the proportion of 8 : 2, use the empirical cross validation method to not repeatedly sample the training set to avoid overfitting of the model, divide the training set into subsets of the same scale, and select one of them as the validation set in order. The remaining sets are used as training sets, and the minimum cross validation error support vector machine model is used as the final prediction model.

$A = \{a_1, a_2, \ldots, a_n\}$ is used to represent the data set. The number of observed values in the data set is n. The predicted values of the corresponding model of each observed value are expressed by b_1, $b_2 \cdots$ and b_n, respectively. The average observed value formula is as follows:

$$\overline{a} = \frac{1}{n}\sum_{i=1}^{n} b_i. \tag{15}$$

The evaluation indexes applied to spark operation performance prediction model are as follows:

(1) Goodness of fit

The fitting degree of spark operation performance prediction model is better when the goodness of fit

value is closer to 1. The goodness of fit formula is as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (b_i - a_i)^2}{\sum_{i=1}^{n} (b_i - \overline{a})^2}. \quad (16)$$

(2) Root mean square error

The higher the prediction accuracy of the job performance prediction model, the smaller the root mean square error. The root mean square error result is the square sum of the mean value of the prediction error, and then the square is processed, and its formula is as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (a_i - b_i)^2}. \quad (17)$$

(3) Average absolute percentage error

The average absolute percentage error is the average value of absolute error, and its formula is as follows:

$$\mathrm{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \frac{|a_i - b_i|}{a_i}. \quad (18)$$

3.5. Parameter Optimization Algorithm of Regional Science and Technology Resource Allocation. Taking the minimization of operation cost as the optimization goal of resource regional science and technology resource allocation for activity performance prediction, taking full account of the two dimensions of cluster resource utilization and activity running time, the optimization goal is to minimize cluster consumption cost.

Define job execution cost and optimize resource utilization and job execution time. The optimization objective formula is as follows:

$$O = \min\left(\mathrm{ET} * \left(\frac{\text{used Mem}}{\text{total Mem}} + \frac{\text{used CPU}}{\text{total CPU}} \right) \right). \quad (19)$$

In formula (19), ET and used Mem, respectively, represent the execution time of spark job and the cluster memory occupied by the job; total Mem and used CPU, respectively, represent the total memory of the cluster and the CPU occupied by jobs, and total CPU represents the total CPU of the cluster.

Based on the relationship between the allocation parameters of scientific and technological resources in the operation area and the operation execution time, the optimization objective formula is as follows:

$$O = \min\left(F(p, d, r, c) * \left(\frac{\text{used Mem}}{\text{total Mem}} + \frac{\text{used CPU}}{\text{total CPU}} \right) \right). \quad (20)$$

The input data scale and resource utilization characteristics of spark operation are known. The optimization problem of science and technology resource allocation in resource area can be transformed into the combination of parameters of science and technology resource allocation in all regions, which belongs to the solution space problem of the least cost optimization search.

Aiming at the problem of market failure, this paper completes the research on the allocation method of regional science and technology resources from the perspective of rationality.

4. Experimental Analysis

In order to better verify the effect and feasibility of the regional science and technology resource allocation method designed from the perspective of rationality and explore whether it can be applied to the actual resource allocation in the future, design experiments. In order to make full use of the data advantages of the regional science and technology resource allocation system, cooperate with the regional science and technology resource allocation reform path evaluation work and synchronize the regional science and technology resource allocation reform path evaluation information to the rationality perspective; the appraisers shall be notified regularly of the evaluation information with simple evaluation data structure, so as to facilitate the appraisers to successfully complete the evaluation. Due to the complex business logic of the evaluation of the path of regional science and technology resource allocation reform, a large amount of background data needs to be called during the experiment. In terms of evaluation data processing, a reasonable perspective is adopted to alleviate the pressure of a large amount of data. Firstly, a large number of data in the regional science and technology resource allocation database are extracted from the data set, then the relevant data are processed through the model for evaluation and analysis, and finally the analysis results are displayed. The specific functions between each level constitute the integrity of the evaluation of the path of regional scientific and technological resource allocation reform and ensure the smooth progress of the evaluation work.

In terms of data collection, the evaluation data required by the path evaluation of regional science and technology resource allocation reform can be obtained through the regional science and technology resource allocation system. The regional science and technology resource allocation system processes the evaluation data provided by the evaluation report declared by the regional science and technology resource allocation personnel through the data warehouse technology, provides the basic data source for the evaluation of the path of regional science and technology resource allocation reform, and is also the basic guarantee for the smooth development of the path evaluation of regional science and technology resource allocation reform.

Suppose Spark + Hadoop is set as the spark cluster environment, which contains three nodes, one master node and two slave nodes. The storage system is HDFS, the resource manager is yarn, and the resource manager yarn version is 2.6.5. Use Cent OS linux release 7.6.1810 as the node resource to configure the operating system. According to the resources occupied during the operation of jobs, Word Count, Sort, K-means, and Naive Bayes jobs are selected as

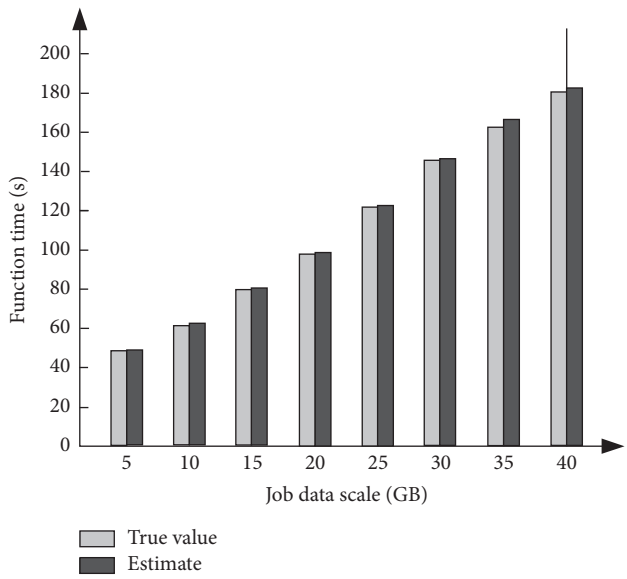


Figure 6: Comparison between predicted execution time and actual execution time.

experimental objects to test the optimization of resource allocation when the above jobs are completed by this method, and the studies in [4] and [5] are selected as comparison methods. The grid resource scheduler in GridSim simulation platform is extended by EFRB, and 100 tasks are constructed to form a resource scheduling set. The execution time of each resource scheduling task is 120 s, and the total task is divided into 10 queues. Science and technology resources are divided into several regions, with an average of two LANs in each region. The storage capacity of the primary site is 300 GB, and the number of nodes varies from 200 to 2000, with an increase of 200. To ensure network connectivity, δ is set to $\delta = 0.65$.

After clarifying the source of experimental data, the following experimental steps are formulated:

Step 1: organize and preprocess the experimental data.

Step 2: randomly generate the evaluation data format by using the evaluation model of the reform path of regional science and technology resource allocation based on the rationality perspective.

Step 3: process the generated evaluation data.

Step 4: calculate the path evaluation coefficient and comprehensive evaluation vector of regional scientific and technological resource allocation reform.

Step 5: take the complexity of regional science and technology resource allocation data as the independent variable, and analyze and count the evaluation error value.

The comparison results between the predicted sort job execution time and the actual execution time are shown in Figure 6.

As can be seen from the experimental results in Figure 6, there is a small error between the predicted job running time by this method and the actual job running time. The overall trend shows that the actual job running time is very consistent with the job running time predicted by this method, which verifies that the job performance prediction model adopted by this method has high prediction effectiveness.

Table 2 shows the prediction performance of regional science and technology resource allocation under different training set sizes by using the three methods.

According to the experimental results in Table 2, using this method to predict the operation time of regional science and technology resource allocation has higher prediction accuracy. The three indicators of the operation time predicted by this method are the best, which effectively verifies that this method has excellent prediction performance. With the increase of the scale of training data set, the root mean square error and average percentage error of the job prediction model used in this method are reduced, which verifies that the job prediction model used in this method has high fitting effect. This method has high stability and can be applied to the actual job running scenario.

Three methods are used to run different jobs 100 times, respectively, compare the quality of the optimal solution obtained by optimizing the configuration parameters by different methods, and select the average value of convergence results as the evaluation index of the effectiveness of configuration parameter optimization.

The smaller the average value of the convergence results, the better the convergence effect of the representation method. The average value of convergence results of optimized configuration parameters by different methods is counted, and the statistical results are shown in Figure 7.

In Figure 7, experimental results show that the proposed method has better convergence effect and stability. This method has high stability for different jobs, and the standard deviation of the convergence result is very close to 0, which validates the high quality and strong adaptability of this algorithm.

Table 2: Comparison of prediction performance of regional science and technology resource allocation.

Training set size (%)	Paper method			Reference [4] method			Reference [5] method		
	Goodness of fit	Root mean square error	Average percentage error	Goodness of fit	Root mean square error	Average percentage error	Goodness of fit	Root mean square error	Average percentage error
10	0.81	31.52	3.52	0.69	65.25	9.52	0.68	67.52	10.52
20	0.82	31.25	3.42	0.71	64.25	9.43	0.69	67.15	9.52
30	0.83	30.95	3.28	0.72	63.52	9.05	0.71	66.85	9.05
40	0.83	30.85	3.12	0.73	62.43	8.94	0.73	66.24	8.76
50	0.84	30.52	3.05	0.74	61.52	8.61	0.74	65.42	8.52
60	0.86	30.25	2.95	0.75	60.85	7.52	0.75	65.02	8.13
70	0.87	29.52	2.85	0.77	60.25	7.13	0.76	64.23	7.85
80	0.88	29.34	2.76	0.78	59.45	7.25	0.77	64.02	7.46
90	0.89	29.14	2.67	0.79	59.05	7.05	0.78	63.52	7.25

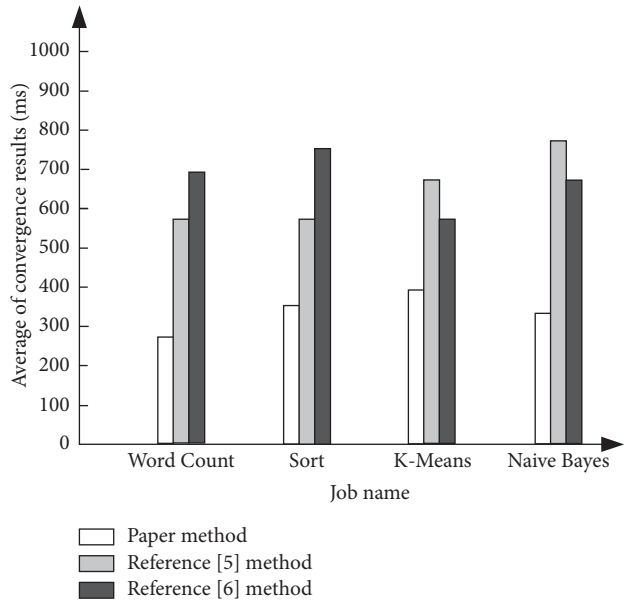


Figure 7: Comparison of the average convergence results of different methods.

In conclusion, there is a small error between the predicted operation time and the actual operation time of the regional science and technology resource allocation method based on the rationality perspective, which has high prediction effectiveness. Using this method to predict the operation time has higher prediction accuracy and fitting effect. At the same time, the convergence effect and stability are good, which can be applied to the actual operation scenario.

5. Conclusion

Today, with the integration of science and technology, economy, and social development, it has become a common strategic choice for all countries in the world to continuously increase investment in science and technology resources. At present, government departments at all levels are constantly increasing the investment in related funds of regional scientific and technological resources, gradually accelerating the growth rate of scientific and technological resources, and paying attention to their impact on economic and social development and supporting ability. While making efforts to increase input, we must attach great importance to the utilization efficiency of scientific and technological resources. If we just blindly increase scientific and technological input without paying attention to the rational allocation of scientific and technological resources, it is bound to cause a huge waste of scientific and technological resources. Experiments show that this method is effective to optimize the allocation of regional science and technology resources.

Data Availability

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding this work.

Acknowledgments

This work was supported by Educational and Scientific Research Projects for Young and Middle-Aged Teachers of the Education Department of Fujian Province, Research on the Promotion Strategies of Scientific and Technological Innovation Policies of Private Enterprises in Quanzhou (JAS21449) and Policy Decision and Consultation Projects of Quanzhou Association for Science and Technology, Research on the Innovation of Allocation Mechanism of Scientific and Technological Resources in Quanzhou City in Response to the 14th Five-Year Plan (Quanzhou Association for Science and Technology [2020] no. 45).

References

[1] G. Sun, H. Chen, and Y. Mao, "The efficiency analysis model of technological innovation resource allocation of coastal Enterprises based on genetic algorithm," *Journal of Coastal Research*, vol. 103, no. 1, p. 691, 2020.

[2] J. Wang and L. Yang, "Does factor endowment allocation improve technological innovation performance? An empirical

study on the Yangtze River Delta region," *The Science of the Total Environment*, vol. 716, no. 5, Article ID 137107, 2020.

[3] W. Zhen, Z. X. Gang, and Z. Ying, "Biased technological progress and total factor productivity growth: from the perspective of China's renewable energy industry," *Renewable and Sustainable Energy Reviews*, vol. 146, no. 1, pp. 1–10, 2021.

[4] J. Lu, G. Jia, and W. Liu, "A three-dimensional operating mechanism for regional science and technology resource allocation system and its system dynamics simulation," *IEEE Transactions on Engineering Management*, no. 99, pp. 1–14, 2020.

[5] M. Garrich, J. L. R. Gazquez, F. J. M. Muro et al., "IT and multi-layer online resource allocation and offline planning in metropolitan networks," *Journal of Lightwave Technology*, vol. 38, no. 12, pp. 3190–3199, 2020.

[6] I. Park, B. Yoon, S. Kim, and H. Seol, "Technological opportunities discovery for safety through topic modeling and opinion mining in the fourth industrial revolution: the case of artificial intelligence," *IEEE Transactions on Engineering Management*, vol. 68, no. 5, pp. 1–16, 2019.

[7] A. I. Rudskoi, V. V. Mishin, and I. A. Shishov, "Scientific and technological basis of production of thin beryllium foils with improved operational properties," *Doklady Chemistry*, vol. 497, no. 2, pp. 55–58, 2021.

[8] Y. Fan, X. Zhao, Z. Han et al., "Scientific and technological progress and future perspectives of the solar assisted heat pump (SAHP) system," *Energy*, vol. 229, no. 8, Article ID 120719, 2021.

[9] T. C. Pimentel, L. R. Brandão, M. P. D. Oliveira, W. K. A. D. Costa, and M. Magnani, "Health benefits and technological effects of Lacticaseibacillus casei -01: an overview of the scientific literature," *Trends in Food Science & Technology*, vol. 114, no. 8, pp. 722–737, 2021.

[10] D. Yu and Z. Zou, "Empirical research on the interaction between marine scientific and technological innovation and marine economic development," *Journal of Coastal Research*, vol. 108, no. 1, pp. 1–8, 2020.

[11] P. Galeazzi and A. Galeazzi, "The ecological rationality of decision criteria," *Synthese*, vol. 198, no. 6, Article ID 11241, 2020.

[12] S. Budiyanti, H. M. Siahaan, and K. Nugroho, "Social communication relation of Madurese people in Max Weber rationality perspective," *Jurnal Studi Komunikasi (Indonesian Journal of Communications Studies)*, vol. 4, no. 2, p. 389, 2020.

[13] T. W. Hung, "Nonhuman rationality: a predictive coding perspective," *Cognitive Processing*, vol. 279, no. 1, pp. 353–362, 2021.

[14] J. Zhang and N. University, "Analyzing conflict talk from the rationality perspective," *Modern Foreign Languages*, vol. 42, no. 01, pp. 25–36, 2019.

[15] B. Chen, Y. Zou, and L. I. Zhi, "The development history and prospect forecast of characteristic towns from the perspective of planning rationality," *Journal of Landscape Research*, vol. 11, no. 3, pp. 41–44, 2019.

[16] Q. Shao, H. C. Man, and J. Huang, "Multimedia crowdsourcing with bounded rationality: a cognitive hierarchy perspective," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 7, pp. 1478–1488, 2019.

[17] X. Jia and F. Gao, "Exploration of decision-making mechanism of national science and technology tasks under governance framework," *Science and Technology Management Research*, vol. 40, no. 08, pp. 35–39, 2020.

[18] D. Vagiona and G. Pozoukidou, "Preface to the topical collection on the 16th international conference on environmental science and technology," *Euro-Mediterranean Journal for Environmental Integration*, vol. 6, no. 1, pp. 1-2, 2021.

[19] M. Z. Allam, "Underlining the role of data science and technology in supporting supply chains, political stability and health networks during pandemics - ScienceDirect," *Surveying the Covid-19 Pandemic and its Implications*, no. 7, pp. 129–139, 2020.

[20] C. Zheng, C. Lixian, L. Chunhong, and M. Guangyuan, "Thoughts on science and technology service in agricultural research institutes: taking institute of plant protection of hebei academy of agricultural and forestry sciences as an example," *Asian Agricultural Research*, vol. 12, no. 10, pp. 23–26, 2020.

[21] R. Hirst, "Writing, in English, for publication in science and technology, and policy: the example of nuclear security," *Journal of Technical Writing and Communication*, vol. 50, no. 1, Article ID 004728161986515, 2020.

[22] K. Katagiri, "Report on 27th AIRAPT international conference on high pressure science and technology," *The Review of High Pressure Science and Technology*, vol. 30, no. 1, pp. 49-50, 2020.

[23] Y. Wu and Y. Wu, "Research on the influencing factors of collaborative innovation of science and technology in the pearl river delta urban agglomeration," *Journal of Physics: Conference Series*, vol. 1616, no. 8, Article ID 012036, 2020.

[24] C. J. Obi, Q. Xiong, and M. Y. Appiah, "Indian journal of science and technology Using genetic algorithm to solve multiple traveling salesman problemconsidering Carbon emissions," *Indian Journal of Science and Technology*, vol. 13, no. 36, pp. 1–9, 2020.

[25] T. Thawarom and W. Singhasiri, "Lexical richness of one-minute speaking task by science and technology university students," *Journal of Asia TEFL*, vol. 17, no. 1, pp. 70–86, 2020.

[26] M. M. Zheleznov, O. I. Karasev, S. S. Trostyansky, and E. A. Shitov, "Critical technologies in the system of science and technology priorities of the railway industry: global experiences," *World of Transport and Transportation*, vol. 17, no. 5, pp. 16–37, 2020.

[27] A. Mamhoori, "Science and technology parks of developing countries as a new way for return on talents (RoT) case study: pardis technology park (PTP)," *SSRN Electronic Journal*, no. 1, pp. 15–20, 2020.

[28] J. C. Orji, C. S. Gana, V. S. Ezema et al., "Assessment of the efficacy of creativity-based instructional model on scientific attitude in basic science and technology among pupils[J]," *Global Journal of Health Science*, vol. 12, no. 5, pp. 1–10, 2020.

[29] J. T. Pruonosa, J. M. Raya, and R. D. Fernández, "The economic and social value of science and technology parks. The case of tecnocampus," *Frontiers in Psychology*, vol. 11, no. 1, pp. 1–10, 2020.

[30] M. Asad, R. Muhammad, N. Rasheed, S. D. M. Chethiyar, and A. Ali, "Unveiling antecedents of organizational politics: an exploratory study on science and technology universities of Pakistan," *International Journal of Advanced Science and Technology*, vol. 26, no. 6, pp. 2057–2066, 2020.

[31] K. E. Supriya and R. K. Rao, "IoT based real time water level monitoring using Texas instruments' CC3200," *Indian Journal of Science and Technology*, vol. 13, no. 17, pp. 1720–1729, 2020.

[32] L. O. Raji, M. Babashani, G. J. Akorede et al., "Changes in semen, hormonal profile and testicular morphology of west african dwarf goat bucks treated with danazol," *Turkish Journal of Agriculture - Food Science and Technology*, vol. 8, no. 12, pp. 2570–2573, 2020.

[33] G. Pak, Y. Noh, M. I. Lee et al., "Korea institute of ocean science and technology earth system model and its simulation characteristics," *Ocean Science Journal*, vol. 56, no. 1, pp. 87–92, 2021.

[34] L. Y. Qiao and Z. Y. Qiu, "Algorithm for electromagnetic spectrum resource scheduling based on cognitive radio technology," *Computer Simulation*, vol. 37, no. 10, pp. 407–411, 2020.

Hindawi
Scientific Programming
Volume 2022, Article ID 9326563, 12 pages
https://doi.org/10.1155/2022/9326563

Research Article

Performance Appraisal Management System of University Administrators Based on Hybrid Cloud

XiuQing Wu

Minnan Science and Technology University, Quanzhou, Fujian Province 362332, China

Correspondence should be addressed to XiuQing Wu; xiuqing@mku.edu.cn

Received 9 November 2021; Revised 29 November 2021; Accepted 9 December 2021; Published 7 January 2022

Academic Editor: Baiyuan Ding

Aiming at the problems of poor functionality, high occupancy, and low real-time performance of the currently designed performance appraisal management system for university administrators, a performance appraisal management system for university administrators based on hybrid cloud is designed. According to the characteristics of hybrid cloud technology, the overall functional requirements and feasibility of the system are analyzed, and the assessment scheme of administrative personnel according to the provisions of relevant documents is analyzed. The performance appraisal index system is designed using an analytic hierarchy process. Using the hybrid cloud architecture and B/S mode based on each component as a service model and J2EE development framework, the basic information management subsystem of administrative personnel, performance appraisal information management subsystem, information analysis and data mining subsystem, and platform system management subsystem are developed. Using XML technology and database technology, the system is integrated with the performance appraisal management system of administrators in colleges and universities. Through the design of data flow diagram and E-R diagram, the design of performance appraisal management system for university administrators based on hybrid cloud is realized. The experimental results show that the proposed method has good functionality, can effectively reduce the system occupancy, and can improve the real-time performance of the system.

1. Introduction

The annual performance appraisal of university administrators is one of the basic works of examining administrators and an important content of cadre management, and the appraisal result is one of the important bases for the organization/personnel department to hire and employ cadres. However, with the deepening of the internal personnel system reform in colleges and universities in China, it is urgent to establish and perfect the performance appraisal system, implement the performance appraisal, and establish the incentive mechanism in the development and management of human resources in colleges and universities, and it is also the foundation for a university to maintain/consolidate and develop. As colleges and universities cultivate high-level and high-quality talents, they are facing unprecedented development opportunities and challenges. People pay more and more attention to the theory and practice of human resource management, and performance management has gradually developed into a widely recognized human resource management behavior [1–3]. The performance appraisal of university administrators is difficult to effectively evaluate their work performance. The traditional performance appraisal method cannot effectively meet the needs of human resource management [4–6]. With the deepening of the strategy of rejuvenating the country through science and education, higher education has developed rapidly. As an important part of human resources in colleges and universities, managers shoulder the important task of serving the teachers and students of the whole university, and their management level is becoming more and more important [7]. Colleges and universities are different from enterprises. How to apply the good practice of performance management in enterprises to colleges and universities continuously strengthen the performance appraisal of university administrators, improve the overall

quality of managers, and reasonably measure the performance, and work efficiency of managers is of great significance to the performance appraisal of colleges and universities.

At present, the research on performance appraisal management system has also made great progress. Nobari et al. [8] designed the staff performance evaluation system of the Iranian National Library and Archives. Using the soft operation research method, combined with the soft system methodology and important performance analysis method, the challenges of employee performance evaluation are determined. The opinions and experience of system stakeholders are used, and the important performance analysis methods are used to determine and prioritize improvement actions. The system designed by this method has certain effectiveness. Demartini and Otley [9] designed a beyond performance management system. Using loose coupling theory, this study studies the relationship types in performance management system. A performance management system coupling index is conceptually developed and verified with a sample of 140 managers operating in different industries. This method designs the system to provide the best results for effectiveness and innovation. However, the above methods still have the problems of poor system functionality, high occupancy rate, and low real-time performance.

To optimize the functionality of the performance management system, the CPU occupancy rate is reduced and the real-time performance of the system is improved, and this study designed a new hybrid cloud-based performance appraisal management system for university managers. The basic information management subsystem, performance evaluation information management subsystem, information analysis subsystem, data mining subsystem, and platform system management subsystem are designed through the hybrid cloud architecture and B/S mode of component as a service model and J2EE development framework. By analyzing the overall functional requirements of the system, the evaluation scheme, and feasibility of management personnel, the performance evaluation index system is designed using an analytic hierarchy process. Using hybrid cloud architecture, B/S mode, J2EE development framework, XML technology, and database technology, through the design of data flow chart and E-R diagram, the performance appraisal management system of university administrators is developed. To verify the application effectiveness of the research system, a simulation experiment is designed to test the system. The experimental results show that the system has an ideal application function and low occupancy rate and can effectively improve the real-time performance of the system.

2. Related Technologies

2.1. Hybrid Cloud Technology. At present, the concept of hybrid cloud widely accepted in the industry comes from the definition of cloud computing released by NIST, which gives cloud computing five characteristics, 3 business models, and 4 deployment models. In particular, five major features are resource pooling, on-demand self-service, broadband network access, rapid flexibility, and measurable services; three types of business models are IaaS (infrastructure as service), PaaS (platform as service), and SaaS (software as service); and four types of deployment models are public cloud, community cloud, private cloud, and hybrid cloud. Hybrid cloud is a combination of several other cloud deployment patterns that support data sharing, automated deployment, flexible migration, and on-demand scaling of customer applications across the cloud. The main business drivers for hybrid clouds are cost savings, on-demand capacity expansion, disaster preparedness, and accelerated deployment of new services.

The hybrid cloud can be configured as either a public cloud environment such as Amazon, Alibaba, or Tencent, or a private cloud environment based on the OpenStack platform, or a hybrid cloud environment with multiple public and private clouds. Service organizations may select cloud computing and cloud storage services of different types and technologies based on the specific circumstances such as resources, costs, and risks, and may achieve unified management of the entire cloud environment, isomorphism of heterogeneous cloud platform resources, and cross-platform collaborative work of businesses. The hybrid cloud interface of this layer can implement the uniform encapsulation and protocol conversion of the underlying cloud services, adapt to the development interfaces of different cloud service providers, and provide a standardized cloud environment for the upper platform. By constructing a unified base cloud service pool, the differences between different technologies and protocols can be eliminated effectively, and the heterogeneous computing and storage resources can be integrated.

Hybrid cloud is a combination of at least one private cloud and at least one infrastructure based on public cloud, which generates an environment, provides transparent user access to hybrid cloud, and has dynamic scalability to manage unbalanced requirements [10–12]. Its basic model is the software as a service (SaaS) model. On this basis, the following models emerged: every component as a service (XaaS), infrastructure as a service (IaaS), platform as a service (PaaS), etc. "Each component as a service" is an ecosystem delivered in the form of components. The component is a ready-made, end-to-end independent business function module. The meaning of XaaS is to make resources consumable, which has become the core concept of hybrid cloud computing. The cloud provides complete components, as shown in Figure 1.

2.2. XML Technology. XML is a language developed by the World Wide Web Consortium (W3C) to describe the structure of the organization and arrangement of data in data documents. It defines a general syntax for tagging data using simple, easy-to-understand tags and provides a standard format for computer documents. The data contained in an XML document are a text string surrounded by text labels that describe the data. Data and tags have a special unit called an element. XML is a meta-markup language for text documents, so you are free to define tags in XML to adequately express the content of the document.

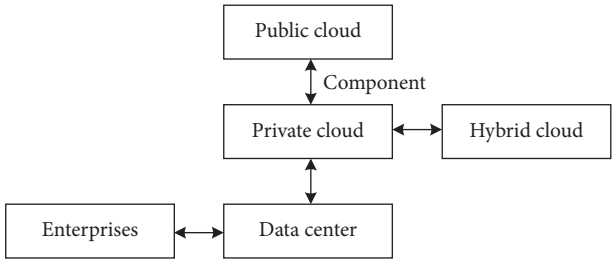


Figure 1: Cloud provides complete components.

The main advantage of Extensible Markup Language (XML) is its universality. It has an unshakable position in the IT field. It is difficult to imagine an important application that does not use XML to save its configuration file or data file [13–15]. XML has two main applications: the first is to use XML to express the underlying data, such as configuration files; the second is to use XML to add metadata to documents. One of the goals of XML is to completely separate data from its display. The same data can be displayed in different ways. When data are transmitted through the Internet, the bandwidth will not be wasted on the style information of the transmitted data. The main advantage of XML is not the display mode with embedded data like HTML. It separates the data from the display mode, which is very useful to realize the data independence of interface program.

3. Demand Analysis of Performance Appraisal Management System for University Administrators

Demand analysis is the key stage of developing the performance appraisal management system for university administrators. It is a process of continuous understanding and gradual refinement. It provides the basis for the physical scheme design in the next stage, which is mainly reflected in the two aspects of "physical demand" and "expressing demand."

3.1. Functional Requirement Analysis. System analysis focuses on the analysis from the perspective of the whole business process. Its main task is to investigate the current system and assessment system, classify and sort out the documents obtained in the investigation, analyze the overall management status and information processing process within the organization, and provide the required data for system development. From the overall performance analysis, the administrative personnel assessment management system has the following requirements:

(1) Accuracy rate: this is an important indicator to measure the system performance. The evaluation indicators of various administrative links of administrative personnel should be accurate; otherwise, it will affect the objectivity and impartiality of the assessment and lose the significance of the system development [16–18].

(2) Scalability: it refers to the expansion of system design functions. The functions realized by the design and development of the system are not limited to the evaluation between administrators and leaders, but also include the evaluation of service objects, and finally realize the "three-level" evaluation system.

(3) Ease of use: the installation, configuration, and maintenance of the system should be simple and convenient, with good user interface, convenient for all kinds of users, and convenient for system transplantation.

(4) Fault tolerance and robustness: on the one hand, the system itself has security performance requirements. On the other hand, when users operate, they can handle corresponding errors and give prompts or warnings, which can not affect the normal operation of the system.

3.2. Assessment Scheme Analysis. According to the provisions of relevant documents, based on the analysis of the original evaluation scheme of administrative personnel in colleges and universities, a set of administrative personnel evaluation schemes in line with the current work reality and development needs of colleges and universities is designed to provide a basis for the functional design of administrative personnel evaluation management system.

3.2.1. Assessment Principle and Organization

(1) The annual assessment shall adhere to the principles of objectivity, impartiality, democracy, and openness, and pay attention to the performance of morality and talent and work performance.

(2) The college sets up an assessment leading group to lead the assessment work of the whole school. The leader of the institute in charge of personnel work shall serve as the leader, and the deputy leader shall serve as the leader of the institute in charge of teaching work. The members are composed of 10 heads of the party and government office, organization department, personnel office, academic affairs office, discipline inspection commission (supervision office), trade union, logistics office, student work office, and staff representatives. The assessment is organized, guided, and supervised, the assessment results are reviewed, and the major problems encountered in the assessment are solved. The assessment leading group of the college has an assessment office, which is located in the personnel department, and the daily affairs are specifically undertaken by the personnel department.

(3) Each department establishes a service object assessment team composed of party and government leaders, faculty representatives, and student representatives. The main person in charge of each department acts as the leader of the assessment team. The members of the assessment team are required to report to the assessment office of the college for

review and filing [19–21]. The representatives of the service object assessment team and the examinee have the relationship of being served and serving, which should be representative and extensive. Before the assessment, the members of the service object assessment team should widely solicit the opinions of the faculty and staff of their department or the students of their class and evaluate the examinee objectively, fairly, and truthfully.

3.2.2. Assessment Scope and Objects

(1) The formal managerial personnel in colleges and universities and the managerial personnel appointed according to labor contracts shall participate in the annual performance assessment.

(2) Those who leave their posts (retire) due to illness (except for work-related injuries), private affair leave, going abroad to visit relatives, and nonpublic suspension of study for more than half a year or (retire) before the end of June of the current year shall not participate in the annual assessment. The early retirees of colleges and universities shall not participate in the annual performance assessment.

3.2.3. Assessment Content. The annual assessment of administrative personnel is mainly carried out according to the requirements of the state and the college on the post responsibilities of various personnel. The content mainly includes five aspects: morality, ability, diligence, performance, and integrity, focusing on the actual performance of the work. The specific contents are as follows:

(1) Morality: it mainly assesses the political and ideological performance, social ethics, and professional ethics of administrative personnel.

(2) Ability: it mainly evaluates the professional level, the application and play of management ability, the improvement of the professional technical level, and the renewal of knowledge in line with the responsibilities of administrative posts.

(3) Diligence: it mainly assesses the working attitude, diligence and professionalism of administrative personnel, and their compliance with labor discipline.

(4) Achievements: it mainly assesses the performance of post responsibilities by administrative personnel, the quantity, quality, and efficiency of work tasks, the level of achievements, and social and economic benefits.

(5) Integrity: it mainly assesses the administrative personnel's compliance with discipline and law, self-discipline, integrity, and self-discipline.

3.2.4. Assessment Criteria. The annual assessment of administrative personnel should determine the corresponding assessment focus according to the basic requirements of management education and service education, the basic performance of political ideology and professional ethics, and in combination with the characteristics of each department.

(1) The assessment of party and government managers focuses on honesty and self-discipline, policy and theoretical level, work ability, work attitude, work style, cooperation spirit, work performance, etc.

(2) The assessment of administrative professional technicians focuses on the attitude of serving teaching and scientific research, professional and technical level, performance of post responsibilities, etc.

(3) For personnel who concurrently hold multiple positions, their performance of main post responsibilities shall be assessed, and their other work shall be taken into account.

The assessment results are divided into three grades: excellent, qualified, and unqualified. Among them, the number of people with excellent grades is less than 15% of the total number of people in the department. The basic criteria for assessing excellence, qualification, and disqualification are as follows.

3.2.5. Excellent. Morality: it supports the line, principles, and policies of the party and the state, has a strong sense of responsibility and exemplary compliance with national laws, regulations, and various rules and regulations, and is loyal to the educational cause of the party and the people. As a leading cadre above the deputy section chief, the personnel of the department have no fighting or violation of social security management, and all the staff of the department under management have responded well.

Ability: strong working ability and greatly improved business ability.

Diligence: hardworking, without being criticized in a notice throughout the year, no absenteeism, personal leave within 7 days, and sick leave within 15 days.

Achievements: well perform the post responsibilities and complete the work tasks specified by the employed post with high quality and full load, innovation, outstanding performance, due diligence, and remarkable achievements in scientific research, management, and logistics services.

Integrity: model compliance with national laws, regulations, and various rules and regulations, honest performance of official duties, and no violation of discipline such as favoritism, corruption, and bribery.

3.2.6. Qualified. Morality: support the line, principles, and policies of the party and the state, and consciously abide by various rules and regulations. Work actively, seriously, and responsibly, be able to complete work tasks without liability accidents.

Ability: strong working ability, improved professional ability, and basically competent for the administrative position.

Diligence: work actively, seriously, and responsibly, abide by labor discipline, and have no late arrival, early leave, or unexplained absenteeism.

Achievements: perform post responsibilities well and basically complete the work tasks specified by the employed post.

Integrity: be able to abide by national laws, regulations, and various rules and regulations, no violation of discipline and law, and no criticism and punishment throughout the year.

3.2.7. Unqualified. Morality: the political and professional quality is low, the organizational discipline is poor, and the rules and regulations cannot be basically observed.

Ability: difficult to adapt to work requirements and poor performance of post responsibilities.

Diligence: absenteeism or failure to return without justified reasons for more than 10 consecutive days, or absenteeism for more than 10 days in a year, or leaving without the approval of the college.

Achievements: poor performance of post responsibilities, weak sense of responsibility, and unable to complete work tasks.

Integrity: those who violate discipline and law, cause adverse consequences, commit serious mistakes, and are punished by the administrative record of major demerit or serious warning within the party.

3.2.8. Design Method of Assessment Index Weight System. There are generally many methods to design the performance appraisal index system [22–24], such as the Delphi method, ranking method, direct judgment method, analytic hierarchy process, weight factor judgment table method, fuzzy coordinated decision-making method, and frequency statistics method. This study uses the analytic hierarchy process (AHP) to design the weight of each index.

(1) Judgment matrix: AHP organizes and hierarchizes the problems by analyzing the factors and related relationships contained in the complex system and constructs an analytic hierarchy process structure model. Generally, the factors contained in the problems are divided into the highest level, the middle level, and the lowest level, that is, the target level, the criterion level, and the index level [25–27]. Assuming that B_k in the structural model is used as a criterion, the element value of the matrix is judged by comparing the element $C_{ki}, \ldots, C_{km}$ of the next level. That is to say, a_{ij} represents the ratio of the influence of factors C_{ki} and C_{kj}, and then, the judgment matrix is as follows:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}. \tag{1}$$

The above element a_{ij} of the judgment matrix represents the relative importance of the elements C_{ki} to C_{kj} from the perspective of the evaluation criterion B_k, and its value can be determined by the "1–9" scale method.

(2) Index weight vector: the maximum eigenvalue and eigenvector are calculated by the square root method, and the calculated eigenvector is the decision index weight vector [28–30]. The specific calculation steps are as follows:

Step 1: calculate the product $\mathbf{M}_i$ of each row element of the judgment matrix:

$$\mathbf{M}_i = \prod_{j=1}^{m} a_{ij}, \quad i = 1, 2, \ldots, m. \tag{2}$$

Step 2: calculate the m root $\overline{W}_i$ of $\mathbf{M}_i$:

$$\overline{W}_i = m\sqrt{\mathbf{M}_i}. \tag{3}$$

Step 3: normalize the vector $\overline{\mathbf{W}} = [\overline{W}_1, \overline{W}_2, \ldots, \overline{W}_m]^T$:

$$W_i = \frac{\overline{W}_i}{\sum_{j=1}^{m} \overline{W}_j}, \quad i = 1, 2, \ldots, m. \tag{4}$$

Then, $\mathbf{W} = [W_1, W_2, \ldots, W_m]^T$ is the required feature vector. The feature vector is the weight vector of the element $C_{k1}, C_{k2}, \ldots, C_{km}$ under the B_k criterion [31].

Step 4: calculate the largest characteristic root $\lambda_{\max}$ of the judgment matrix:

$$\lambda_{\max} = \frac{1}{m} \sum_{i=1}^{m} \frac{(\mathbf{A}W_i)}{W_i}. \tag{5}$$

In formula (5), (AW_i) represents the i element of the vector $\mathbf{AW}$. AW is expressed as follows:

$$\begin{aligned} \mathbf{AW} &= \begin{bmatrix} (AW)_1 \\ (AW)_2 \\ \cdots \\ (AW)_m \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} W_1 \\ W_2 \\ \cdots \\ W_m \end{bmatrix}. \end{aligned} \tag{6}$$

(3) Consistency test: due to the complexity of objective things or one-sided understanding of things, it is necessary to test the consistency and randomness of the judgment matrix whether the eigenvector (weight) obtained through the constructed judgment matrix is reasonable. The test formula is as follows:

$$C_R = \frac{C_I}{R_I}. \tag{7}$$

In formula (7), C_R is the random consistency ratio of the judgment matrix and R_I is the average random consistency index of the judgment matrix. R_I is obtained from a large number of experiments. For the low-order judgment matrix, the value of R_I is shown in Table 1.

C_I is expressed as the consistency index of the judgment matrix [32–34], which is as follows:

$$C_I = \frac{1}{m-1}(\lambda_{\max} - m). \tag{8}$$

In formula (8), m is expressed as the order of the matrix. When $C_I < 0.1$, the judgment matrix is considered to have satisfactory consistency, indicating that the weight distribution is reasonable. Otherwise, the judgment matrix needs to be adjusted until a satisfactory consistency is achieved.

3.3. Feasibility Analysis

(1) Economic feasibility: it mainly compares the funds invested in the development of the new system with the economic benefits brought by the system after it is put into use to confirm whether the new system will bring certain economic benefits to the college. At present, computers are equipped in all departments, and the modern technology center has relevant equipment, and a group of professional teachers and technical teams. To develop this system, you only need to buy relevant software and pay *R&D* personnel. The development of the system will promote the reform and improvement of the administrative staff assessment system of the college and promote the long-term development of the college, and the benefits are far greater than the required expenses. Therefore, the system is economically feasible.

(2) Operational feasibility: the college administrative personnel assessment management system is mainly to assist the personnel department in the assessment of administrative personnel. The operation is simple. Users can quickly be familiar with the functions of each module of the system. Moreover, the system can integrate the scores of all assessors and automatically generate the assessment results, saving time, manpower, and material resources. In addition, the system has good compatibility, which can be applied under general application system conditions, and colleges and universities also have these conditions. Therefore, the system is feasible in operation.

(3) Legal feasibility: the design of this system is carried out within the scope permitted by law and does not involve any aspects that conflict with the law, such as contract, liability, or copyright infringement. Therefore, the system is legally feasible.

(4) Technical feasibility: it is to measure whether the required technology is available according to the objectives of the new system, mainly including four aspects: whether the current relevant technology can support the developed new system and the number and level of new system developers, namely human resources, and hardware and software resources.

To sum up, the design and development of this system meet the technical and hardware conditions. Therefore, it is technically feasible.

4. Design of Performance Appraisal Management System for Administrative Personnel in Colleges and Universities

4.1. System Architecture Design. To realize the requirements obtained through the analysis of real objects, this study adopts the hybrid cloud architecture [35] and B/S mode based on each component as a service model and uses the J2EE development framework with the advantages of strong portability, security, and stability [36]. A basic information management subsystem for administrative personnel, a performance appraisal information management subsystem, an information analysis and data mining subsystem, and a platform system management subsystem are developed. XML technology and database technology are used to integrate the system with the performance appraisal management system of administrators in colleges and universities. The overall architecture of the system is shown in Figure 2.

The system is equipped with four types of authorized users: ordinary users, administrative personnel data administrators, performance appraisal data administrators, and system super administrators. All subsystems are deployed on the public cloud and the private cloud, respectively. VPN is established between the public cloud and the private cloud through password technology to realize interoperability. Among them, the clients of the basic information management subsystem of administrative personnel, the performance appraisal information management subsystem, and the information analysis and data mining subsystem are deployed on the public cloud, ordinary users and data administrators realize data entry and query through these clients, and the platform system management subsystem, all functional components, and all databases are deployed on the private cloud and realized the actual management function of data information.

4.2. Software Structure Design. Due to the hybrid cloud architecture of each component as a service model, the performance appraisal management system of university administrators based on hybrid cloud is divided into

Table 1: Average random consistency index value of analytic hierarchy process.

Matrix order	1	2	3	4	5	6	7	8	9	10	11
R_I	0	0	0.58	0.96	1.12	1.24	1.32	1.41	1.45	1.49	1.51

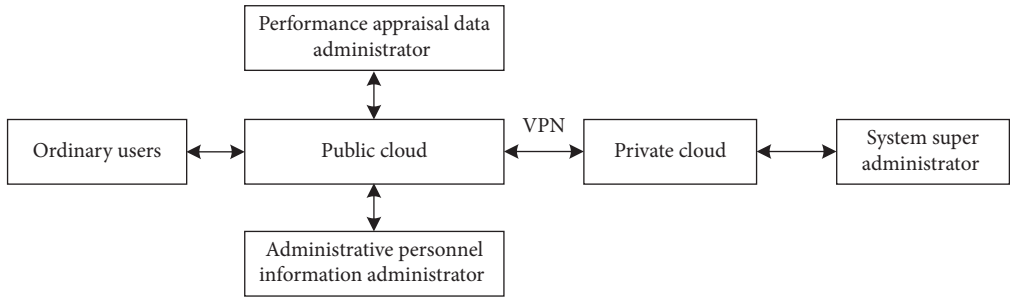


Figure 2: Overall system architecture.

three layers at the software level, which are data cluster, functional component group, and human-computer interaction software group from bottom to top. The system software framework structure is shown in Figure 3.

The data cluster is implemented in the system support database server, which is specifically composed of administrative personnel basic information database, performance appraisal information database, platform system dictionary database, platform system operation management database, and other personnel information management system interface database, and is deployed on the private cloud. The functional component group is realized by many functional components, specifically including administrative personnel basic information management functional components, performance appraisal information management functional components, information analysis and data mining functional components, platform system management functional components, and XML data interface functional components. Among them, XML data interface and information analysis and data mining functional components are deployed in the public cloud, and other functional components are deployed in the private cloud. The human-computer interaction software group is a group of interface programs realized by HTML + JSP, and the corresponding functions are realized by calling functional components. The interface programs related to platform system management are deployed in the private cloud and others in the public cloud.

4.3. System Function Design. The main functions of the whole assessment management system are divided into five modules: personnel management, assessment questionnaire management, assessment management, report output, and statistical query. The functional structure of the system is shown in Figure 4.

(1) Personnel management: evaluation object maintenance, evaluation appraiser maintenance, and administrator maintenance.

(2) Appraisal questionnaire management: questionnaire maintenance, answer maintenance, questionnaire publishing, and judgment matrix setting.

(3) Appraisal management: appraisal indicator maintenance, indicator weight calculation, and personnel appraisal.

(4) Report output: statistics report output.

(5) Statistics query: result statistics, questionnaire statistics, and statistics result query.

4.4. Data Flow Diagram Design. The data flow diagram is the main tool to describe the logical function of the system. Through several specific symbols, it can comprehensively reflect the overall situation of information use, processing, transmission, and storage in the system. The data flow diagram is an effective tool for system structural analysis. It abstractly describes the system data processing, but it cannot express the details of each processing. Therefore, it is necessary to further supplement the data flow and processing in the data flow diagram, which is the data dictionary and transformation logic description. The appraiser data flow chart includes the administrative personnel, leaders in charge, and service objects of the same department. The appraiser needs to enter the user name and password to log in to the system, manage his own basic information, and fill in comments, opinions, and query comments. The data flow chart of administrative personnel needs to enter the user name and password to log in to the system, manage their basic information, fill in self-evaluation and management summary, modify personal information, and query the evaluation content.

4.5. E-R Drawing Design. E-R method is the abbreviation of "entity relation method." It is an effective method to describe the conceptual structure model of the real world. It uses a rectangle to represent the entity type, and the entity name is written in the rectangular box. Ellipses are used to represent the properties of entities, and undirected edges

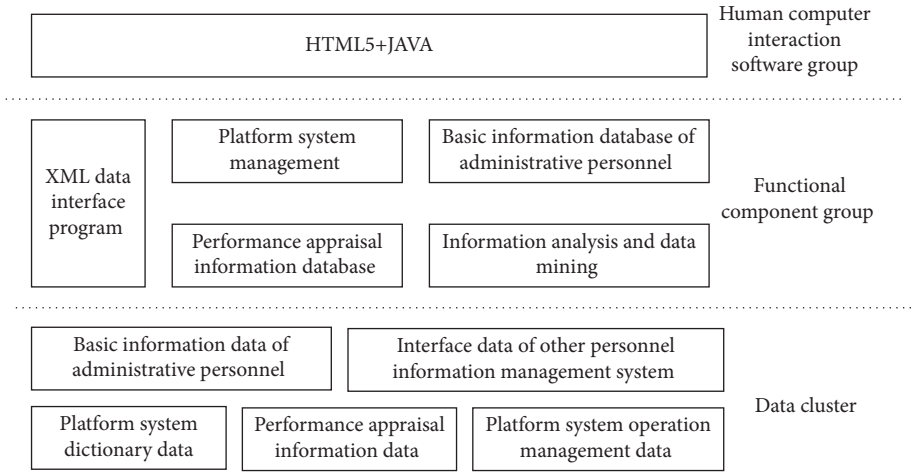


Figure 3: System software framework structure.

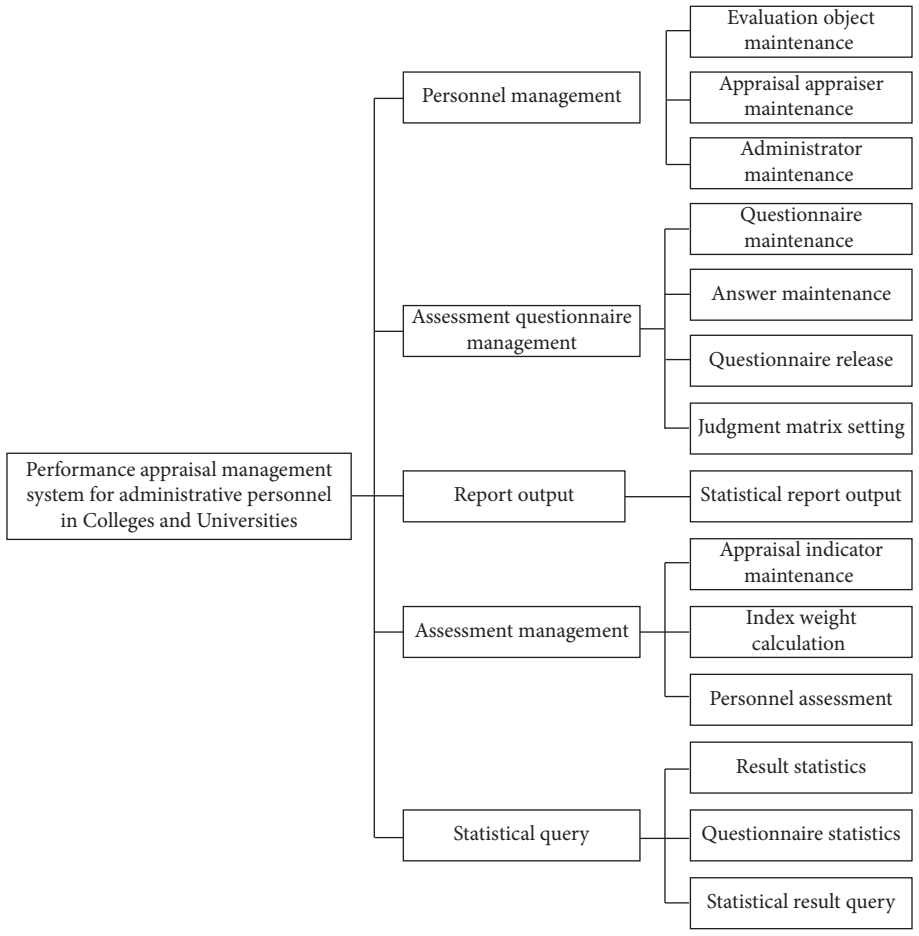


Figure 4: System function structure diagram.

are used to connect them with the corresponding entity types. Diamond is used to indicate the connection between entity types, write the contact name in the diamond box, connect with relevant entity types with undirected edges, and mark the type of connection (1 : 1, 1: n, or M: n) next to undirected edges. The overall E-R diagram of the system is shown in Figure 5.

5. System Test Results and Analysis

5.1. Setting System Test Environment. To test the effectiveness of the designed performance appraisal management system for university administrators based on hybrid cloud, a trial operation was carried out in a university, and several staff were arranged to participate in the test. Problems found

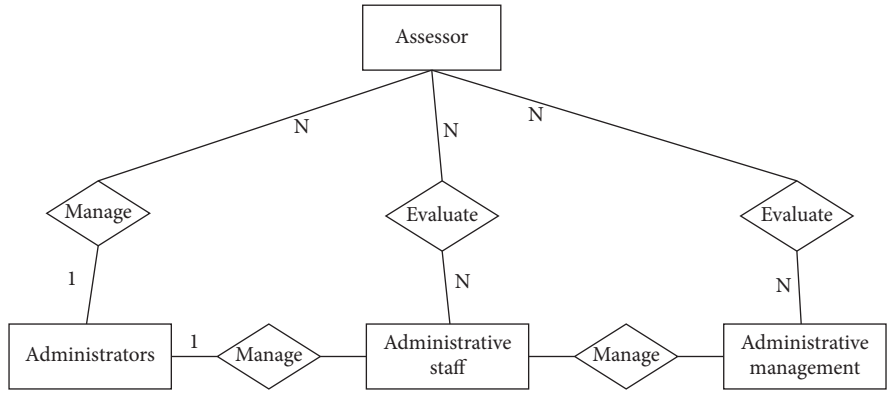


Figure 5: Overall E-R diagram of the system.

through test collection shall be updated and corrected in time during trial operation. When testing, the staff try to operate according to the real business scenario, write corresponding test cases for each module, find problems, and collect feedback, using CentOS 6.5 operating system, Apache + tomcat, MATLAB application software, and MySQL 5.6 database software. The test method of this project adopts black box testing. From the perspective of actual users, according to the software specifications (from demand analysis), software testers test the software functions by inputting various relevant data and observing the output results, regardless of the internal operation principle of the software and how to realize the program, so as to find its defects.

5.2. Test Methods. The test indexes are system functionality, occupancy rate, and response time. The higher the qualified rate of the system function is, the lower the CPU occupancy rate is. The better the real-time performance is, the better the system application performance is. The methods of reference [8], the methods of reference [9], and the proposed methods are compared to verify the performance of the proposed method. For the simulation experiments designed in this study, the simulation software is used to test the application effects of different methods. The simulation software is MATLAB 2020b. The above 3 methods are input into the simulation software to simulate the application and output the test results.

5.3. System Functional Test Results. To verify the functionality of the system, the system function test status is taken as the evaluation index, and the test status is qualified, which indicates that the system functionality designed by the method is good. By comparing the methods of reference [8], the methods of reference [9], and the proposed methods, the system functional test results of different methods are shown in Table 2.

According to the data in Table 2, the function test status of assessment management in the system function test status of the methods of reference [8] is unqualified, and its error type is function error. In the system function test status of the method of reference [9], the function test status of the examination questionnaire management is unqualified, and its error type is syntax error. The system function test status of the proposed method is qualified, which indicates that the system function designed by the proposed method is good.

5.4. System Occupancy Test Results. To further verify the occupancy rate of the system designed by the proposed method, the methods of reference [8], the methods of reference [9], and the proposed method are compared, respectively, and the comparison results of the system occupancy rates of different methods are obtained, as shown in Figure 6.

Analyzing Figure 6 shows that as the number of administrative staff increases, the system occupancy rate of different methods increases. When the number of administrative personnel is 250, the system occupancy rate of the methods of reference [8] is 15.5%, the system occupancy rate of the methods of reference [9] is 11%, and the system occupancy rate of the proposed method is only 4.5%. It can be seen that, compared with the methods of reference [8] and the methods of reference [9], the system occupancy rate of the proposed method is lower.

5.5. System Real-Time Test Results. On this basis, the real-time performance of the proposed method is further verified, and the system response time is taken as the evaluation index. Among them, the shorter the system response time, the higher the real-time performance of the system. The methods of reference [8], the methods of reference [9], and the proposed methods are compared, respectively, and the system real-time comparison results of different methods are obtained, as shown in Figure 7.

By analyzing Figure 7, it can be seen that with the increase in the number of administrative personnel, the system response time of different methods increases. When the number of administrative personnel is 250, the system response time of the methods of reference [8] is 18 s and the system response time of the methods of reference [9] is 21.5 s, while the system response time of the proposed method is only 9.6 s. It can be seen that the system response time of the proposed method is shorter than that of the methods of reference [8] and the methods of reference [9], indicating that the system has high real-time performance.

Table 2: System functional test results of different methods.

Different modules	The proposed method		The methods of reference [8]		The methods of reference [9]	
	Test status	Error type	Test status	Error type	Test status	Error type
Personnel management	Qualified	Nothing	Qualified	Nothing	Qualified	Nothing
Assessment questionnaire management	Qualified	Nothing	Qualified	Nothing	Unqualified	Syntax error
Assessment management	Qualified	Nothing	Unqualified	Function error	Qualified	Nothing
Report output	Qualified	Nothing	Qualified	Nothing	Qualified	Nothing
Statistical query	Qualified	Nothing	Qualified	Nothing	Qualified	Nothing

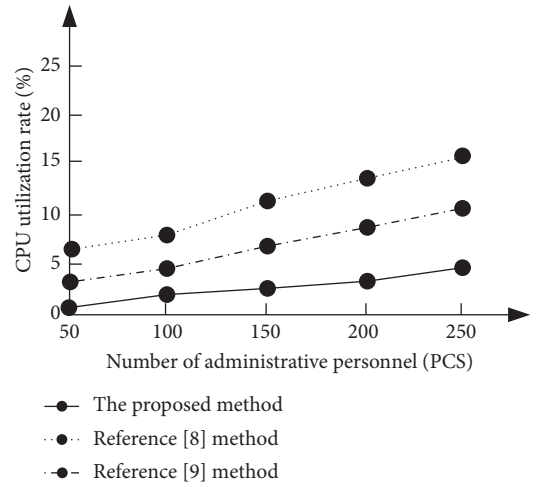


Figure 6: Comparison results of system occupancy rate of different methods.

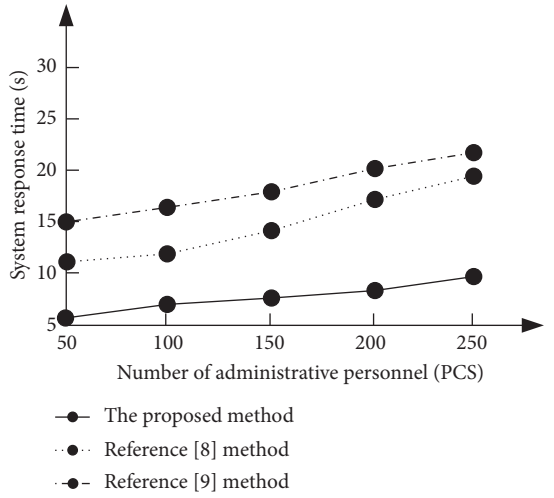


Figure 7: Comparison results of system real-time performance of different methods.

6. Conclusion

Performance appraisal is an effective means of modern management in colleges and universities. If it can be used reasonably and effectively, the overall efficiency and performance of colleges and universities will be improved without increasing the input of human capital. The performance appraisal management system based on hybrid cloud is designed in this study, which fully displays the advantages of hybrid cloud technology. According to the characteristics of hybrid cloud technology, the overall functional requirements and feasibility of the system are analyzed to optimize the system application and reduce the CPU usage. Using hybrid cloud architecture and B/S model, the basic information management subsystem and performance evaluation information management subsystem are developed, and the hardware design is completed. In the part of software design, data flow graph and E-R graph are introduced to optimize the real-time performance of the system.

However, in the performance appraisal management system of university administrators, due to the limitations of technical level and the complexity of the work content, specific posts, positions, influencing factors, and service objects of university administrators, the system has not fully considered the relevant factors, and many problems need to be further studied and discussed in specific practice. Therefore, in the next research, we should also consider the division of assessment results and the selection of weight proportion, the division of administrative personnel types, and the differences in assessment indicators and weights according to the changes in the actual situation.

Data Availability

The raw data supporting the conclusions of this article will be made available by the author, without undue reservation.

Conflicts of Interest

The author declares that there are no conflicts of interest regarding this work.

Acknowledgments

This work was supported by the Research and Implementation of Personnel Management System Design Based on B/S Frame (MK2021002).

References

[1] N. Yadav, "Application of system dynamics methodology in performance management system: a case study of Indian automotive firm," *International Journal of Business Performance Management*, vol. 21, no. 4, pp. 385–399, 2020.

[2] S. Joomunbaccus, "Investigating the intention to use an electronic performance management system at XYZ holdings limited: towards a revised technology acceptance model," *Archives of Business Research*, vol. 9, no. 1, pp. 75–96, 2021.

[3] F. A. Yovan, "Decision support system for flood risk assessment and public sector performance management of emergency scenarios," *International Journal of Public Sector Performance Management*, vol. 1, no. 1, Article ID 10026215, 2020.

[4] T. B. Febrian and A. Simangunsong, "Decision support system employee performance appraisal method using topsis," *Journal of Computer Networks, Architecture, and High-Performance Computing*, vol. 2, Article ID 47709, 2020.

[5] J. P. Pereira, E. Natalya, and I. Slesarenko, "The analysis of competency model for a performance appraisal system in the management of food service industry," *Advances in Intelligent Systems and Computing*, vol. 1137, pp. 162–171, 2020.

[6] W. He, L. Chen, and W. Liu, "Does new performance appraisal system (EVA) affect earnings management?" *Nankai Business Review International*, vol. 11, no. 2, pp. 191–216, 2020.

[7] J. Liu, C. Wang, and Y. Wu, "Construction and optimization of higher education management system based on Internet video online technology," *Scientific Programming*, vol. 2021, no. 1, Article ID 5520662, 11 pages, 2021.

[8] B. Z. Nobari, A. Gholipour, E. Ebrahimi, and A. Shoja, "Employee performance appraisal system development in the national library and archives of Iran (NLAI): soft operational research approach," *Performance Measurement and Metrics*, vol. 22, 2021, ahead-of-print(ahead-of-print).

[9] M. C. Demartini and D. Otley, "Beyond the system vs. package dualism in performance management systems design: a loose coupling approach," *Accounting, Organizations and Society*, vol. 86, Article ID 101072, 2020.

[10] E. Barbierato, L. Campanile, M. Gribaudo, M. Iacono, M. Mastroianni, and S. Nacchia, "Performance evaluation for the design of a hybrid cloud based distance synchronous and asynchronous learning architecture," *Simulation Modelling Practice and Theory*, vol. 109, Article ID 102303, 2021.

[11] M. Bouache, "Hybrid cloud: workload placement performance," *Hybrid Cloud Migration*, vol. 20, no. 9, pp. 135–139, 2020.

[12] T. Rajagopal, M. Venkatesan, and A. Rajivkannan, "An improved efficient dynamic load balancing scheme under heterogeneous networks in hybrid cloud environment," *Wireless Personal Communications*, vol. 111, no. 4, pp. 1837–1851, 2020.

[13] A. Yousfi, M. Hafid, and A. Zellou, "xMatcher: matching extensible markup language schemas using semantic-based techniques," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 8, 2020.

[14] W. C. Hsu and I. E. Liao, "UCIS-X: an updatable compact indexing scheme for efficient extensible markup language document updating and query evaluation," *IEEE Access*, vol. 8, Article ID 176375, 2020.

[15] G. Polancic and B. Orban, "A BPMN-based language for modeling corporate communications," *Computer Standards & Interfaces*, vol. 65, pp. 45–60, 2019.

[16] V. Karimli, "Administrative management theory at high educational institutions and its following stages," *International Journal of Management*, vol. 11, Article ID 34218, 2021.

[17] Y. Tang, D. Xie, W. Zhang, X. Guan, and Y. Han, "Research on the informationization construction of college admission management," *Canadian Social Science*, vol. 15, no. 2, pp. 273–278, 2019.

[18] M. S. Jani, J. Shahril, M. I. Muszali, R. Adnan, and S. A. Mai, "Administrative management quality standard and teachers' practices compliance on school sports safety," *International Journal of Academic Research in Progressive Education and Development*, vol. 10, no. 2, pp. 648–664, 2021.

[19] R. K. Upadhyay, K. R. Ansari, and P. Bijalwan, "Performance appraisal and team effectiveness: a moderated mediation model of employee retention and employee satisfaction," *Vision: The Journal of Business Perspective*, vol. 24, no. 4, pp. 395–405, 2020.

[20] E. Y. Huang, D. Paccagnan, W. Mei, and F. Bullo, "Assign and appraise: achieving optimal performance in collaborative teams," *Social and Information Networks*, vol. 22, Article ID 09817, 2020 https://arxiv.org/abs/2008.09817.

[21] S. Ling, "Research and practice of performance appraisal of specialist nursing staff," *Advanced Journal of Nursing*, vol. 2, no. 2, pp. 313–318, 2021.

[22] S. Susanto and N. K. Darmasetiawan, "Designing a performance appraisal system for management consultant office "D"," *Dinasti International Journal of Digital Business Management*, vol. 1, no. 3, pp. 361–374, 2020.

[23] L. J. Xiao and Z. X. Xun, "Performance appraisal of technological innovation in military-civilian enterprises—empirical evidence from jiangsu military-civilian integration enterprises," *International Journal of Business and Management*, vol. 15, no. 11, p. 79, 2020.

[24] L. Jin, J. Liu, and D. Kong, "Evaluation of the incorporation of gross ecosystem product into performance appraisals for ecological compensation," *Acta Ecologica Sinica*, vol. 39, no. 1, pp. 24–36, 2019.
[25] B. Zhang, C. C. Li, Y. Dong, and W. Pedrycz, "A comparative study between analytic hierarchy process and its fuzzy variants: a perspective based on two linguistic models," *IEEE Transactions on Fuzzy Systems*, vol. 29, no. 99, Article ID 3018110, 2020.
[26] Z. J. Wang, "Eigenproblem driven triangular fuzzy analytic hierarchy process," *Information Sciences*, vol. 578, no. 1, pp. 795–816, 2021.
[27] A. Zjw, Y. B. Xuan, and A. Xtj, "And-like-uninorm-based transitivity and analytic hierarchy process with interval-valued fuzzy preference relations," *Information Sciences*, vol. 539, pp. 375–396, 2020.
[28] Z. A. Pan, M. A. Teng, J. A. Gang, and Z. W. Yang, "An efficient interference steering matrix estimation scheme for weight vector correction," *Digital Signal Processing*, vol. 114, Article ID 103039, 2021.
[29] X. Ma, Y. Yu, X. Li, Y. Qi, and Z. Zhu, "A survey of weight vector adjustment methods for decomposition-based multiobjective evolutionary algorithms," *IEEE Transactions on Evolutionary Computation*, vol. 24, no. 4, pp. 634–649, 2020.
[30] Y. Xue and L. Zhang, "Laplacian pair-weight vector projection for semi-supervised learning," *Information Sciences*, vol. 573, no. 2, pp. 1–19, 2021.
[31] E. Li and R. Chen, "Multi-objective decomposition optimization algorithm based on adaptive weight vector and matching strategy," *Applied Intelligence*, vol. 50, no. 6, pp. 4206–4222, 2020.
[32] T. Wahyuningrum, A. Azhari, A. Azhari, and S. Suprapto, "Modified LFPP to improve the accuracy of matrix pairwise comparison consistency index in the usability evaluation," *International Journal of Intelligent Engineering and Systems*, vol. 13, no. 3, pp. 397–406, 2020.
[33] J. Aguarón, M. T. Escobar, and H. Jiménez, "Reducing inconsistency measured by the geometric consistency index in the analytic hierarchy process," *European Journal of Operational Research*, vol. 288, pp. 576–583, 2021.
[34] X. Chen, L. Peng, Z. Wu, and W. Pedrycz, "Controlling the worst consistency index for hesitant fuzzy linguistic preference relations in consensus optimization models," *Computers & Industrial Engineering*, vol. 143, Article ID 106423, 2020.
[35] H. Pang and W. Zhao, "Big data efficient hybrid cloud storage model and algorithm simulation under 5 G network," *Computer Simulation*, vol. 37, no. 7, pp. 350–353+379, 2020.
[36] Y. Zhao and R. Wang, "Design and implementation of scitech project management system based on J2EE spring MVC framework," *Shanxi Science and Technology*, vol. 143, Article ID 106423, 2019.

Hindawi
Scientific Programming
Volume 2022, Article ID 5379646, 7 pages
https://doi.org/10.1155/2022/5379646

Research Article

Artificial Intelligence Education System Based on Differential Evolution Algorithm to Optimize SVM

Weilin Long[1] **and Yi Gao**[2]

[1]*Tianjin Modern Vocational Technology College, Tianjin 300350, China*
[2]*College of Information Technical Science, Nankai University, Tianjin 300071, China*

Correspondence should be addressed to Yi Gao; gaoyi@nankai.edu.cn

Received 15 November 2021; Accepted 13 December 2021; Published 7 January 2022

Academic Editor: Baiyuan Ding

The artificial intelligence education system promotes the rooting of artificial intelligence in the education field and accelerates its entry into the era of intelligent education. This article focuses on the development of the artificial intelligence education system and proposes an artificial intelligence education system based on differential evolution algorithm optimization support vector machine. First, the processing of educational demand information data is automated, then a differential evolution algorithm is built to optimize the support vector machine model, and the model is used to implement various educational tasks to achieve automated education. The test results show that the model classification accuracy, classification recall rate, classification accuracy rate, and $F1$-score value are 4 items. Performances have been improved to improve the efficiency of education work and provide a reference for exploring the application and practice of artificial intelligence in education.

1. Introduction

As the core driving force of technological progress, artificial intelligence has brought profound changes to social development. The education field is also actively exploring how to deeply integrate with artificial intelligence, promote educational innovation, and build a new education ecology in the intelligent era. However, the specific empowerment of artificial intelligence in education is not achieved overnight [1–4]. The education field is currently facing problems such as lack of mature application models, lack of artificial intelligence experts, and lack of technical platform support. Artificial intelligence has not been able to effectively support the construction of educational tools and education systems [5–10]. How to promote the application of artificial intelligence and realize the convenient development of artificial intelligence systems will be an important link in the development of educational artificial intelligence [11–16].

Automation is a sign of the progress of human civilization and social modernization. With the progress of human society, automation technology continues to develop under the impetus of social demand. At present, the construction of artificial intelligence systems through automated methods is an important direction that the field of artificial intelligence has begun to pay attention to [17, 18]. As a new research method, artificial intelligence methods such as automated machine learning and automated deep learning use the design idea of "training artificial intelligence with artificial intelligence" to expand the scope of artificial intelligence research and application and achieve the goal of open and inclusive artificial intelligence. The user only needs to provide the input data and task type, and the professional tasks such as algorithm and model construction in the system modeling are automatically completed by the machine, thereby effectively reducing the threshold of artificial intelligence applications and system development and promoting the standardization and module of artificial intelligence applications and automation [19–25].

The current research on artificial intelligence in education generally presents the status quo that there are many theoretical discussions and few practical applications [26, 27]. Automated artificial intelligence methods provide ideas for

artificial intelligence research and applications in the education field. Based on the analysis of existing research and problems in the design and application of educational artificial intelligence systems, this paper discusses design methods based on automation ideas and proposes an artificial intelligence education system based on differential evolution (DE) algorithm optimization support vector machine (SVM). The educational demand information is processed automatically, and then a feature model is generated, and the model is used to implement various educational tasks so as to realize automatic education, improve the efficiency of educational work, and provide a reference for exploring the application and practice of educational artificial intelligence.

2. Function Analysis of Artificial Intelligence Education System

2.1. Function Frame Structure. The development function of the educational artificial intelligence application system includes five modules: educational needs, educational data, educational features, intelligent models, and educational applications. It mainly organizes relevant data and information according to the needs of educational work, digs out important information from them, uses the extracted educational features information, builds an intelligent education model, and invests it in education. Figure 1 shows the system function frame structure.

Educational Needs. Educators and related technical personnel are responsible for related work, setting up related issues according to the current requirements for educational work, and listing educational needs information.

Educational Data. Educators and related technical personnel are responsible for related work, collecting educational data related to this information, and using data processing technology to complete data cleaning and conversion operations based on educational demand information.

Educational Characteristics. Educational experts and related technical personnel are responsible for related work, extracting educational characteristics from the collated data information, and constructing an educational characteristics system at the same time.

Intelligent Model. Relevant technical personnel are responsible for related work, constructing intelligent models, and training them, according to the model evaluation results, optimizing the current model design plan to obtain the best model.

Educational Application. Educators and related technical personnel are responsible for related work and putting the optimized model into use, so as to play the role of the educational intelligence system.

2.2. Problems in Current System Development. At present, China has invested a lot of research funds for the development of educational artificial intelligence systems and encouraged major scientific research institutions to participate in the research of such projects. Although some research results have been obtained, there are still many problems:

(1) *Technical Aspects*. The system has poor portability, relying on manual debugging and selecting the corresponding operation mode according to the operation requirements. These control technologies are difficult to meet the needs of the model.

(2) *System Development Cost Issue*. System development requires a lot of money. In addition to the purchase and leasing of system hardware and software-related equipment, it also includes operation and maintenance, human resources, and other costs. Judging from the current development situation, the system development and operating costs have exceeded the expected range.

(3) *System Development Complexity*. There are many people involved in system development, involving experts and staff in multiple fields. In order to ensure the effective connection of various links, almost every link requires the participation of all staff, which increases the complexity of the system development.

(4) *System Operation Limitation and Efficiency Issues*. Most of the systems currently developed only support the current parameter setting requirements. If the data or educational features are changed, the system operating procedures need to be readjusted, and some hardware devices should be replaced according to the needs, resulting in the system application range restricted. In the test, it was found that the operating efficiency of the system was low and the mature system development plan was lacking.

3. Design of Artificial Intelligence Education System Based on DE-SVM

3.1. DE. DE is a random search heuristic technology based on group differences. It is a new and efficient technology. The DE algorithm is developed on the basis of genetic algorithm. Similar to the genetic algorithm, it has operations such as mutation, crossover, and selection.

The detailed process of the DE algorithm is as follows.

3.1.1. DE Algorithm Initializes the Population. To determine the population size m, each individual is an n-dimensional vector, and the initial population can be expressed as the following formula:

$$X_t(0) = X_{t1}(0) + X_{t2}(0) + \cdots + X_{tj}(0). \quad (1)$$

Among them, X_t (0) is the t-th individual of the 0th generation and $X_{tj}(0)$ is the j-th gene of the t-th individual of the 0th generation, $t = 1, 2, ..., m$.

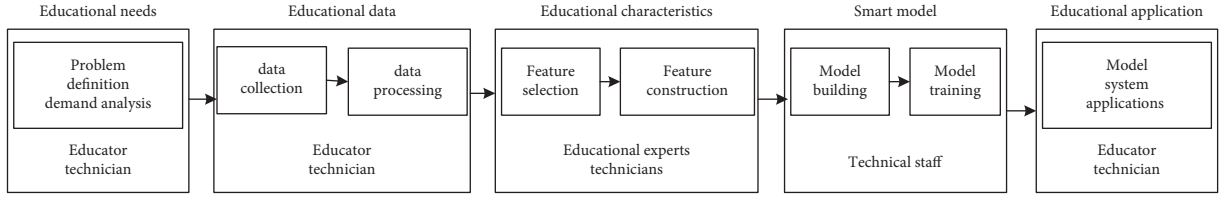


Figure 1: The functional framework structure of the artificial intelligence education system.

3.1.2. DE Algorithm Mutation. In the kth iteration, for an individual $X_t\ (k) = (X_{t1}\ (k + X_{t2}(k) + \ldots + X_{tn}(k))$, an intermediate vector Lt(k) is generated as follows and one is randomly selected from the population. For three individuals $X_{y1}(k)$, $X_{y2}(k)$, and $X_{y3}(k)$ and $y1 \neq y2 \neq y3$, there are the following formulas:

$$F_t(k) = X_{y1}(k) - E * \left(X_{y2}(k) - X_{y3}(k)\right), \tag{2}$$

where $\Delta = X_{y2}(k)\text{-}X_{y3}(k)$ is the difference vector.

E is the scaling factor, used to control the influence of the difference vector, generally between [0, 1].

It can be seen intuitively to construct the variation vector *Ft(k)* in the two-dimensional parameter space as shown in Figure 2.

3.1.3. DE Algorithm Crossover Operation. In the kth iteration, each individual crosses the first intermediate vector generated by it. Specifically, for each identical component, the first intermediate vector is selected with a certain probability and the second intermediate vector is generated. Generate the second intermediate vector $S_t(k)$, and calculate each component of $S_t(k)$ according to the following formula:

$$S_{tj}(k) = \begin{cases} F_{tj}(k), & \text{rand}\ (0,1) \leq P, \\ X_{tj}(k), & \text{else}, \end{cases} \tag{3}$$

where P is the crossover probability, $F_{tj}(k)$ is a gene in the number one intermediate vector, and $X_{tj}(k)$ is a gene in the original individual.

3.1.4. DE Algorithm Selection Operation. The selection of the differential evolution algorithm adopts the greedy mode. In this process, according to the value of the fitness function, in the kth iteration, the second intermediate vector $S_t(k)$ and the original vector X_t of each individual are selected for the next generation (k) the one with higher fitness, so that the population evolves to the optimal solution; the selection method is as follows:

$$X_t(k+1) = \begin{cases} S_t(k), & f\left(S_t(k) > X_t(k)\right), \\ X_t(k), & \text{else}. \end{cases} \tag{4}$$

After completing this step, continue to loop through the differential evolution operation until the number of iterations or fitness function requirements are met, and then the algorithm stops.

3.2. SVM. SVM is a supervised learning model that can analyze data between classification and regression analysis. Its basic idea is to define a linear classifier with the largest interval in the function space. The SVM classifier also includes a kernel technique that allows nonlinear classification. The learning strategy of the SVM classifier is the optimal classification hyperplane, where this hyperplane must meet the classification requirements to maximize the blank space on both sides of the hyperplane while ensuring the classification accuracy. The main idea of SVM is as follows: given a set of data set $T = \{(x\_1, y\_1), (x\_2, y\_2), \ldots, (x\_N, y\_N)\}$, where $x\_i \in Rn$, $y\_i \in \{-1, +1\}$, $i = 1,2, \ldots, N$, which satisfy

$$y_i\left(w \cdot x_i + b\right) \geq 1, \tag{5}$$

making

$$\min_{w,b} \frac{w^2}{2}. \tag{6}$$

According to Lagrangian duality, the optimal solution can be obtained by solving the dual problem of the original problem, which is transformed into

$$\begin{cases} \max\limits_{\alpha} & \left(\dfrac{1}{2}\displaystyle\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j\left(x_i \cdot x_j\right) + \sum_{j=1}^{N}\alpha_i\right) \\ \text{s.t.} & \displaystyle\sum_{j=1}^{N}\alpha_i y_i = 0, \quad \alpha_i \geq 0,\ i = 1, 2, \ldots, N. \end{cases} \tag{7}$$

After adding a minus sign to the target formula, the problem of solving the maximum value is converted into a minimum value problem, and after conversion, it becomes

$$\begin{cases} \min\limits_{\alpha} & \left(\dfrac{1}{2}\displaystyle\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j\left(x_i \cdot x_j\right) + \sum_{j=1}^{N}\alpha_i\right) \\ \text{s.t.} & \displaystyle\sum_{j=1}^{N}\alpha_i y_i = 0, \quad \alpha_i \geq 0,\ i = 1, 2, \ldots, N. \end{cases} \tag{8}$$

After calculating the solution α, we further solve w and b according to α to obtain the maximum separation hyperplane and the classification decision function.

3.3. SVM Optimization Process Based on DE Algorithm. The function of the DE algorithm is to optimize the parameters of the SVM, that is, to find the optimal parameter combination (C, gamma), so that the SVM model has the best classification performance. The detailed steps of SVM

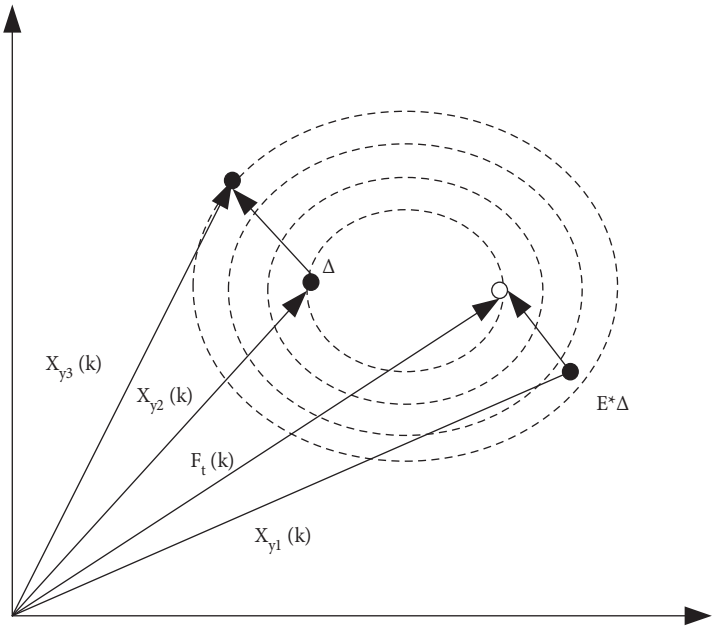


Figure 2: Schematic diagram of mutation vector construction.

parameter optimization based on the DE algorithm are as follows:

Step 1. Define the optimization objective function and initialize the size of the population N, the scaling factor k, the evolution algebra g, the crossover probability CR, the end condition, and the range of the parameters C, and gamma to randomly generate a combination (C, gamma).

Step 2. Use the current parameter combination (C, gamma) as the SVM parameter, use the SVM model to train and test the sample data, and obtain the classification result.

Step 3. Compare the above classification results with the actual classification results, calculate the value of the objective function, and judge whether the value reaches the predetermined accuracy or the maximum evolutionary algebra. If any of them is satisfied, go to Step 8; otherwise, $g = g + 1$; enter the next generation of evolution.

Step 4. Randomly select 4 different individuals $X_i(g)$ from the population of the current g generation and perform the mutation operation on the individuals according to formula (2) to generate the variant individual $V_i(g + 1)$ of the $g + 1$ generation.

Step 5. Perform crossover operation on individual $V_i(g + 1)$ according to formula (3) to generate the experimental individual $U_i(g + 1)$ of the $g + 1$ generation.

Step 6. Perform a greedy selection on the experimental individual $U_i(g + 1)$ of the $g + 1$ generation according to formula (4) and generate the individual $X_i(g + 1)$ of the $g + 1$ generation.

Step 7. Calculate and generate a new parameter combination (C, gamma) in the population of the $g + 1$ generation and then go to Step 2.

Step 8. Obtain the optimal parameter combination (C, gamma) of the SVM and then use the DE-SVM model to train and test the sample data to apply it to the passenger flow transfer model. The detailed process of using the DE algorithm to optimize SVM parameters is shown in Figure 3.

3.4. Model Building of Artificial Intelligence Education System. This system is based on self-contained thinking and learning theory. It automatically processes education demand information in the cloud and then generates a feature model, which is used to implement various educational tasks, thereby realizing automated education and improving the efficiency of educational work. Among them, model training is the key to improve educational work. Based on data features, it automatically combines educational information to create interpretable features. When the characteristics of education change, the education model is automatically adjusted to improve the efficiency of the system and expand the scope of system application conditions. The system model design scheme is shown in Figure 4.

The system model is mainly composed of 5 parts, namely, demand analysis and problem definition, education data, automation platform, application interface, and education system. Among them, the demand analysis and problem definition module is the extraction of intelligent education information and the summary of the problem as the input port of system operation data. The education data module is to extract important information about the development of education from the first section, as an information extraction and transmission tool, and send this part of the information to the automation platform. It completes various operations in the cloud. These operations all adopt automated processing technologies to clean up

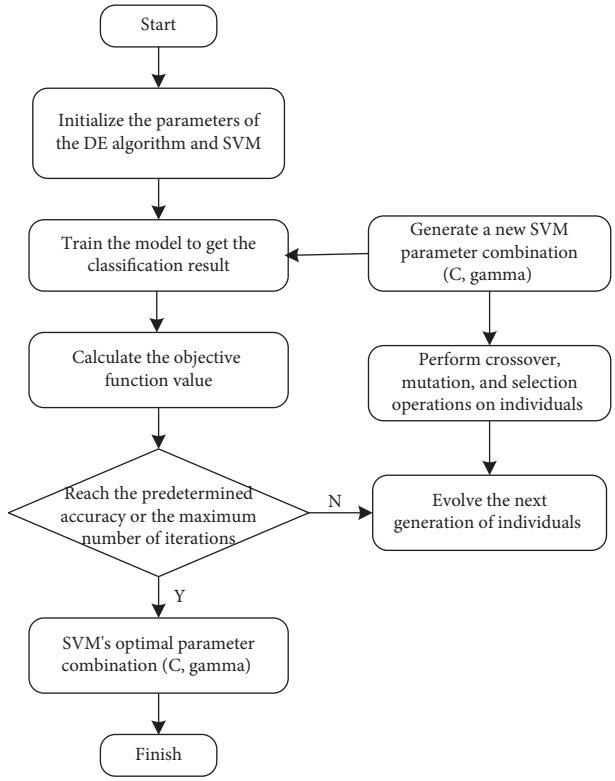


Figure 3: Flow chart of the DE-SVM model.

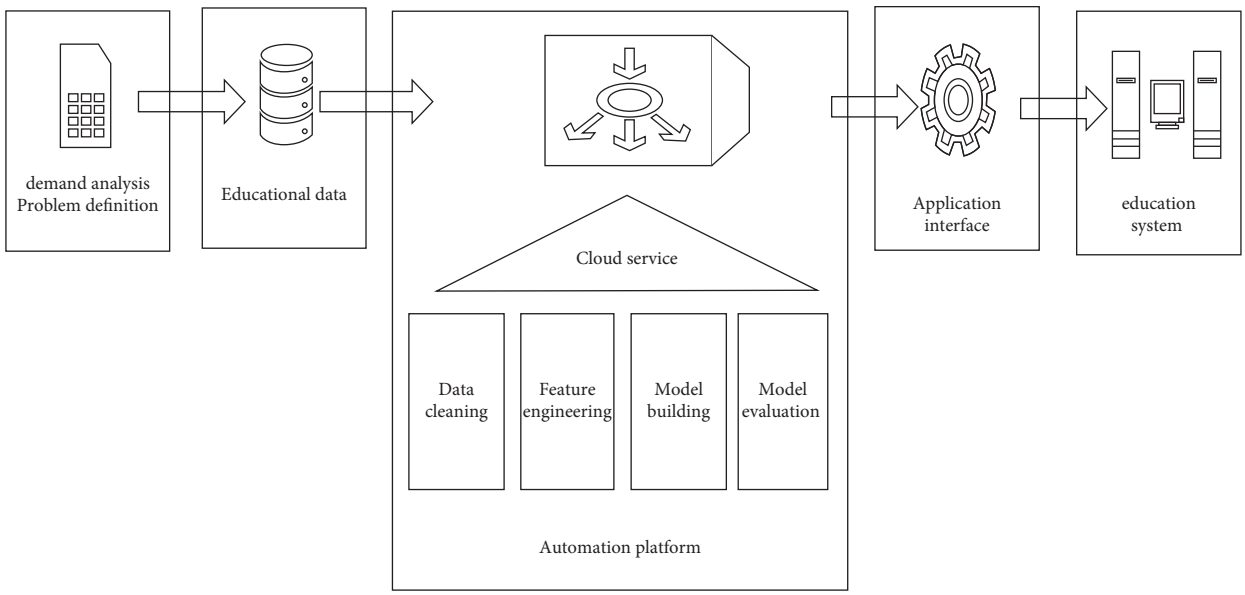


Figure 4: Artificial intelligence education system model.

data, create feature projects, build automated education models, and evaluate the role and performance of the models. The communication connection between the automation platform and the education system is established through the application interface. Through this interface, the automation platform operation status information is sent to the education system. According to the changes in the educational characteristics, the education system adjusts the operating parameters of the automation platform and completes the automatic update in the cloud.

4. Case Application Analysis

This research applies the system model to the MOOC and collects a large amount of data about the MOOC to provide a data source for system testing and analysis. Among them,

Table 1: Forecast results.

Number	Accuracy (%)	Recall rate (%)	Accuracy rate (%)	$F1$-score (%)
1	91.42	92.32	93.52	91.55
2	91.56	92.45	93.64	91.58
3	91.67	92.61	93.68	91.67

data information mainly includes learning process, learning experience, and learning evaluation. This system mainly provides convenient conditions for MOOC study and discussion. With the help of teaching assistants and teachers, different types of text data were marked to form 4 types of information, namely, personal feelings, learning evaluation, content goals, and curriculum questions.

$$F - \text{score} = \left(1 + \beta^2\right) \frac{\text{accuracy} \times \text{recall rate}}{\beta^2 \times \text{accuracy} + \text{recall rate}}. \quad (9)$$

The test indicators for model job performance include four items: information classification accuracy, classification recall rate, classification accuracy rate, and classification F value ($F1$-score). There are about 6,000 courses forum data, of which 80% are used as the training set, 10% of the system operation verification set, and the remaining 10% are used as evaluation data support. These data are brought into the system created in this article. The indicators of the currently developed system are around 70%, and there is no more than 80%. If the test result of this system is more than 80%, it is considered that this system is better than the system developed in the past. Performance has improved. The predicted results of various indicators are shown in Table 1.

Judging from the above three sets of test results, the operating performance of this model is better than that of the traditional system. The values of all indicators are more than 91%, and the classification accuracy rate is more than 93%.

5. Result

This article focuses on the development of the education artificial intelligence system. Aiming at the problems of current system development, from the perspectives of operating efficiency, cost, application range, and so on, this paper proposes an artificial intelligence education system based on the differential evolution algorithm to optimize SVM. First, the processing of education demand information is automated, then a feature model is generated, and the model is used to implement various educational tasks to achieve automated education. The test results show that the model classification accuracy, classification recall rate, classification accuracy rate, and $F1$-score value are 4 items. The performance has been improved to improve the efficiency of education work and provide a reference for exploring the application and practice of artificial intelligence in education. Automated methods provide directions to promote the application of artificial intelligence in education. Combining educational scenarios and educational needs, the practice of smart education applications should be the focus of the next step of research.

Data Availability

The dataset used to support the finding of this study can be accessed upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

[1] D. Liu, "Application scenario analysis and system design of intelligent assistants in education," *China Audio-Visual Education*, vol. 392, no. 9, pp. 27–35, 2019.

[2] Z. Zhang, L. Zhang, and Q. Luo, "Actual analysis of artificial intelligence education application: methods and limits of teaching automation," *China Distance Education*, vol. 530, no. 3, pp. 5–17, 2019.

[3] Q. Sun and X. Li, "Research on the design of self-education system of resource intelligence course," *Research in Audio-Visual Education*, vol. 6, pp. 98–104, 2019.

[4] P. Wang, "Application analysis and design of artificial intelligence in educational videos," *Audio-Visual Education Research*, vol. 323, no. 3, pp. 95–102, 2020.

[5] J. Zhao, X. Xue, and C. Xu, "Review of the application of a new generation of artificial intelligence technology in power system scheduling operation," *Automation of Electric Power Systems*, vol. 694, no. 24, pp. 6–15, 2020.

[6] J. Wang and X. Qu, *In-depth Understanding of AutoML and AutoDL: Building an Automated Machine Learning and Deep Learning Platform*, Machinery Industry Press, Beijing, China, 2019.

[7] L. WuQ. Liu et al., "Research on the semantic analysis model and application of MOOC comments from the perspective of big data," *Audio-visual Education Research*, vol. 38, no. 11, pp. 43–48, 2017.

[8] M. Shen, G. Cen, W. Zhou, R. Zhu, and Y. Liang, "Early education system design based on CNN intelligent AI assistant," *Journal of Zhejiang University of Science and Technology*, vol. 32, no. 6, pp. 590–594, 2020.

[9] S. Jiang, *Application of Deep Learning in Student Education Recommendation System*, Changchun University of Technology, Changchun, China, 2020.

[10] X. Yu, "The application of automated methods in the design of educational artificial intelligence systems," *Electronic Technology and Software Engineering*, vol. 17, pp. 186-187, 2021.

[11] Y. Liu, "Research on the application of artificial intelligence industry based on big data in education innovation," in *Proceedings of the 18th Shenyang Science Conference*, p. 8, Shenyang, China, October 2021.

[12] G. Li, "A performance prediction decision-making system using machine learning technology under the background of smart education," *Journal of Ningde Normal University (Natural Science Edition)*, vol. 33, no. 1, pp. 36–41, 2021.

[13] W. Ma and J. Ma, "The application and prospect of artificial intelligence in the field of physical education," *Journal of Sports Adult Education*, vol. 36, no. 6, pp. 42–45+99, 2020.

[14] Q. Hu, *Design and Performance Optimization of AI Algorithm Education System for Educational Robots*, Donghua University, Shanghai, China, 2020.

[15] W. Zhong and Z. Li, "Research on network education system based on machine learning," *Journal of Communications*, vol. 39, no. S1, pp. 135–140, 2018.

[16] M. Yu, F. Xiang, and Z. Zhu, "Educational application and innovative exploration of machine learning in the perspective of artificial intelligence," *Big Data Era*, vol. 1, pp. 64–73, 2018.
[17] Z. Wang, "Analysis of the application of artificial intelligence in the field of education," *China New Telecommunications*, vol. 20, no. 22, p. 179, 2018.
[18] X. Yue, X. Cheng, and J. Li, "Research on "artificial intelligence + education" system based on big data," *Computer Products and Circulation*, vol. 5, p. 133, 2018.
[19] L. Shi, "Educational application of open source artificial intelligence system tensorflow," *Modern Educational Technology*, vol. 28, no. 1, pp. 93–99, 2018.
[20] E. Chen, L. Qi, S. Wang et al., "Key technologies and applications of adaptive learning for intelligent education," *Journal of Intelligent Systems*, vol. 1-13, 2021.
[21] W. Lan and Z. Li, "Thoughts on the teaching reform of the "database" course integrating the cultivation of computational thinking," *Computer Age*, vol. 11, pp. 110–116, 2020.
[22] S. Zhang, "Talking about the application of artificial intelligence based on expert system in the field of education," *Digital World*, vol. 3, p. 167, 2020.
[23] S. Zhang, "Exploration of the application of "AI+education" in the United States," *Shanghai Information Technology*, vol. 1, pp. 54–57, 2020.
[24] J. Ran, "The future has come—looking at artificial intelligence and education," *Communication Power Technology*, vol. 36, no. 11, pp. 169–171, 2019.
[25] W. Li, "Research on artificial intelligence and future education development under the background of new economy," *Marketing Industry*, vol. 47, pp. 79-80, 2019.
[26] A. Li, Z. Guo, S. Xie, D. Zhao, and S. Zhang, "Research on the classification of steel plate surface defects based on SVM with ant colony and particle swarm optimization," *China Test*, vol. 46, no. 1, pp. 110–116, 2020.
[27] C. Yun, S. Song, Y. Pan, and Y. Li, "Credit risk assessment model based on SVM hybrid integration," *Computer Engineering and Applications*, vol. 52, no. 4, pp. 115–120, 2016.

Hindawi
Scientific Programming
Volume 2022, Article ID 1371778, 10 pages
https://doi.org/10.1155/2022/1371778

Research Article

Unbalanced Big Data-Compatible Cloud Storage Method Based on Redundancy Elimination Technology

Tingting Yu

College of Internet of Things Technology, Wuxi Institute of Technology, Wuxi, Jiangsu 214121, China

Correspondence should be addressed to Tingting Yu; yutt@wxit.edu.cn

Received 18 November 2021; Revised 9 December 2021; Accepted 23 December 2021; Published 6 January 2022

Academic Editor: Baiyuan Ding

In order to meet the requirements of users in terms of speed, capacity, storage efficiency, and security, with the goal of improving data redundancy and reducing data storage space, an unbalanced big data compatible cloud storage method based on redundancy elimination technology is proposed. A new big data acquisition platform is designed based on Hadoop and NoSQL technologies. Through this platform, efficient unbalanced data acquisition is realized. The collected data are classified and processed by classifier. The classified unbalanced big data are compressed by Huffman algorithm, and the data security is improved by data encryption. Based on the data processing results, the big data redundancy processing is carried out by using the data deduplication algorithm. The cloud platform is designed to store redundant data in the cloud. The results show that the method in this paper has high data deduplication rate and data deduplication speed rate and low data storage space and effectively reduces the burden of data storage.

1. Introduction

In the big data environment, data security and privacy protection are facing great impact and challenges. In recent years, with the explosive growth of digital information, data occupies more and more storage space [1]. It is found that the redundancy in the big data saved by the application system is as high as 60%, and the redundancy increases with the passage of time. The traditional data storage technology reduces the redundancy of data through coding mapping according to the internal relationship of data, so as to increase the data density and finally reduce the space occupied by data [2, 3]. Moreover, the traditional data storage technology can only eliminate the redundant data inside the file, but it cannot do anything about the data redundancy between different files. It can be seen that the traditional data storage technology and management method have been difficult to meet the requirements of big data in terms of speed, capacity, storage efficiency, and security. Therefore, it is necessary to study an effective data storage method [4, 5].

Reference [6] presents a redundant optical fiber data storage optimization method based on traditional genetic algorithms and data compression algorithms. Combined with the Doppler transformation, the optimal basis function is globally optimized, and the redundant properties of the optical fiber data are analyzed and filtered. Furthermore, the fiber redundant data is initially compressed using the traditional genetic algorithm. The load reduction processing of optical fiber data storage based on K-L characteristics completes the optimization of optical fiber redundant data compression and realizes the optimal storage of optical fiber redundant data. The method can effectively compress the fiber redundant data and handle it more efficiently. However, the method only compresses redundant data and does not really delete redundant data, so redundant data still occupies a large storage space. Reference [7] proposes a design and optimization method of spatial vector data storage model based on HBase. Based on the analysis of relational spatial database storage model, the conversion rules from relational database storage mode to HBase storage mode are applied to the field of spatial vector data management, and a conversion method from spatial vector data relational storage mode to HBase storage mode is proposed, and a space vector data HBase storage model is designed. The model is optimized and improved by using the characteristics of HBase, such as entity nesting,

antinormalization, and no pattern. Using this method for data storage without auxiliary indexing, the data query is highly efficient. But the method only changes the storage method, and deduplication is poor. Reference [8] proposed a digital library data storage method based on compressed sensing, introduced the theoretical basis and mathematical model of compressed sensing, and tried to apply compressed sensing technology to the digital resource management of the library for the first time. This method applies the compression sensing method of the orthogonal matching tracking algorithm to the acquisition of scanning text resources and image electronic resources, with relatively high compression. However, the method does not completely delete the deduplication data, and the duplicate data occupies a large storage space.

Although the above methods improve the efficiency of data storage to a certain extent, there are some problems such as poor redundancy elimination effect and high data storage space. Aiming at the above problems, this paper explores a nonequilibrium big data compatibility cloud storage method based on redundancy elimination technology. Taking unbalanced big data as the main research object, it is hoped that the analysis of cloud storage technology will lay the foundation for further promotion of computer cloud computing data storage technology in the later period. The main research contents and innovations of the thesis include the following:

(1) Based on Hadoop and NoSQL, a data collection platform was designed, and multiple concurrent data collection function modules were opened on multiple machines at the same time, which improved the data collection efficiency of the entire platform

(2) Use the Huffman algorithm to compress the data, which can significantly reduce the storage space of the data and improve the data query speed in the storage mode

(3) Redundancy elimination algorithm in nonequilibrium big data encryption technology is for big data elimination, detecting duplicate data objects in the data flow according to redundancy, transmitting and storing only unique copies of data objects, and replacing other duplicates with the unique data object copy, so as to eliminate the same files or data blocks in the big data set, effectively reduce the storage space of big data, and reduce the amount of data transmitted by the network

2. Unbalanced Big Data Compatibility Cloud Storage Method

Unbalanced data usually means that the number of negative samples in the two types of problems is much larger than the number of positive samples. In reality, examples include credit card transaction fraud identification, telecom equipment failure prediction, enterprise bankruptcy prediction, and radar image monitoring of offshore oil pollution. However, most traditional classification methods are based on the assumption of data balanced distribution in data storage design, When these methods are applied to unbalanced data, they will lead to poor data storage performance. Therefore, unbalanced data storage has become a research hotspot in the field of data processing.

2.1. Design of Unbalanced Big Data Collection Platform. Big data, as the name implies, has a very high demand for the efficiency of data collection, storage, and retrieval. Normally, the data collection efficiency can reach MB/s or GB/s, and the data storage can reach the order of TB or even PB. Because of consistency constraints, traditional relational databases have been difficult to meet such high-intensity requirements [9, 10]. In response to this, a large number of nonrelational databases have emerged. Their common feature is that they perform read and write operations based on Key-Value, but they lack support for complex operations such as multicolumn retrieval and multitable joint statistical analysis. On the other hand, due to the limitation of its caching mechanism, it does not support rapid collection and retrieval of large amounts of data, resulting in low overall efficiency. In order to solve these problems, a new type of big data collection platform is designed based on Hadoop and NoSQL technologies, through which efficient unbalanced data collection is realized, and the data foundation is provided for the subsequent storage and processing of data [11, 12]. Figure 1 is a schematic diagram of the overall architecture of the unbalanced big data collection platform.

According to Figure 1, the overall architecture of unbalanced big data acquisition platform can be started from three levels: physical layer, logical processing layer, and network layer. The construction of functional modules should be considered from the aspects of unbalanced big data acquisition, audit, management, sharing, and security control. Among them, unbalanced big data acquisition module is the most front-end, and its task is to actively collect external information of the platform. There are two collection methods, one is automatic collection, and the other is manual collection. Both methods incorporate the unbalanced big data resources obtained on LAN, intranet, and Internet into the database of the collection platform [13, 14].

The big data collection cluster is equivalent to the internal entrance of the whole system. Taking the process as the execution unit, multiple concurrent data collection function modules are opened on multiple machines at the same time to improve the data collection efficiency of the whole platform; the concurrent data collection function module uses the 5CSEMA5F31C6FPGA core board of Altera cycloneV series as the processor and uses FPGA parallel collection technology to collect several bits of data code collected by the automatic data collection module and the manual data module simultaneously through multiple machine parallel channels. A character was divided into 8 bits to transmit an 8 bit signal once in parallel. Parallel acquisition module mainly includes FPGA acquisition module, digital demodulation module, and high-speed data transmission module. The FPGA terminal will determine whether the data can be collected in parallel at high speed,

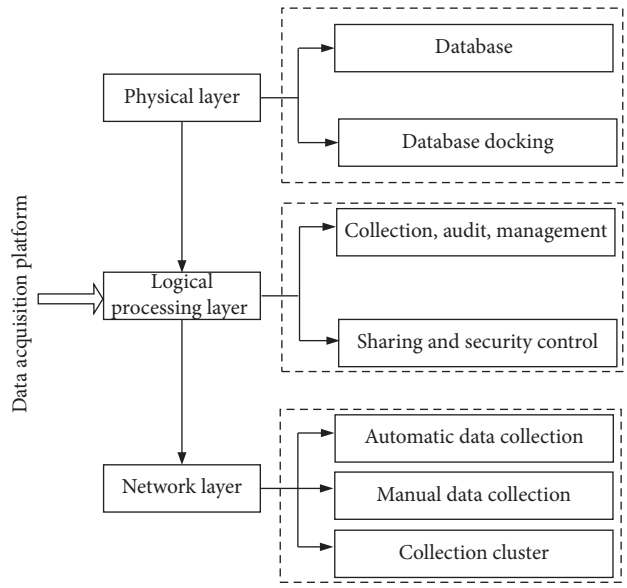


Figure 1: Overall architecture of unbalanced big data collection platform.

FIFO with clock frequency of 25M, after reading 4 bit check bit data, and Hamming coding, and the remaining data packets are collected in parallel, so as to ensure that the collected bit bit rate is 25 Mbps; digital demodulation module codes data according to Hamming code principle divides data by frequency and completes parallel collection and demodulation of multimachine data; high-speed data transmission module is mainly in P2P, based on backplane bus and coaxial transmission line, and vector signal analyzer sends IQ data to FPGA in parallel, trimmed by FPGA, and sends to database in MDA; the parallel acquisition process reduces signal attenuation, improves anti-interference capability, and realizes high-speed parallel data acquisition. After the completion of the data collection, the data analysis stage gradually receives the big data obtained in order, in this step, to complete the unbalanced data classification, operate according to the type, and finally complete the data storage.

2.2. Unbalanced Big Data Classification. Based on the data collected by the unbalanced big data collection platform, in order to further improve data storage efficiency and storage effectiveness, this article uses a classifier to classify the collected data, so as to improve the pertinence of data storage in the subsequent process.

The integration of the classifier in this paper is mainly divided into three steps: the first step is to use the MapReduce framework to process the initial unbalanced data set, thereby obtaining k balanced training subset; the second step is to use these k trains of the training subset to train the classifiers to obtain N classifiers; the third step is to use the majority voting strategy to integrate these N classifiers to complete the classification of unknown sample data [15–17]. The specific algorithm is described as follows:

(1) Enter the initial unbalanced data set D.

(2) Construct m balanced data sets $D = \{d_1, d_2, \ldots, d_m\}$ according to the unbalanced rate in data set D.

(3) Apply the ELM algorithm to train the training set $D = \{d_1, d_2, \ldots, d_m\}$ to obtain n subclassifiers $S = \{s_1, s_2, \ldots, s_n\}$.

(4) When classifying unknown sample data, each subclassifier obtains a classification result and votes on different subclassifiers. The final classification result of the sample is the one with the highest vote.

The overall algorithm process shown in Figures 2 and 3 is a schematic diagram of partial data classification results.

As can be seen from Figure 2, in the unbalanced big data classification process, in the construction of the initial unbalanced data set, each training set contains some examples in the negative class and all the positive examples, ensuring that all available data information in the training set will not be wasted. Classification results according to the data obtained by Figure 3 in the m balanced data set are trained, inspired by the principle of the MapReduce framework, the training of the m data set is distributed to each node, and the ELM algorithm is applied for parallel calculation, which greatly improves the classification efficiency [18]. Extreme Learning Machine (ELM) is an algorithm for single-implied layer feed-forward neural networks. The biggest feature is that both input weights and bias of implicit nodes are randomly generated within a given range, proven to be learning efficient and generalization. The main purpose of training is to solve the weights of the output layer. ELM has the advantages of high learning efficiency and strong generalization ability and is widely used in problems like classification, regression, clustering, and feature learning. For single hidden layer neural networks, ELM can randomly initialize the input weights and bias and get the corresponding output weights. Since only the required solution output weights, the ELM is essentially a linear parameter mode, and its learning process is easy to converge at global minima. For a given group N group training data, the SLFN containing L implied layers and M output layers using ELM learns in the following steps:

(1) Randomly assign the input weight vector and imply layer nodes of the ELM to complete the initialization

(2) Calculate the implied layer output parallel data matrix H

(3) Classify the matrix and calculate the output weight matrix in parallel

That is, the parallel calculation of the ELM neural network is completed.

2.3. Unbalanced Big Data Compression. Although the optimization of unbalanced big data can improve the efficiency of data processing to a certain extent, due to the huge capacity of big data, only classification cannot meet the needs of data processing. Therefore, Huffman algorithm is used to compress the classified unbalanced big data.

Huffman algorithm is a data compression method. When using the Huffman algorithm for data compression,

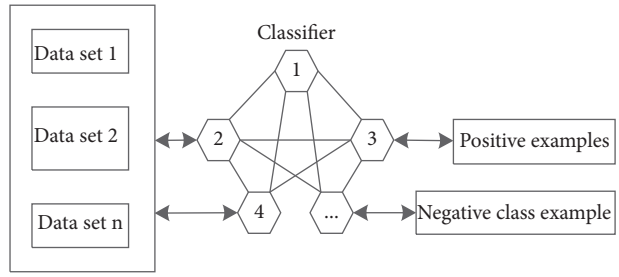


Figure 2: Unbalanced big data classification process.

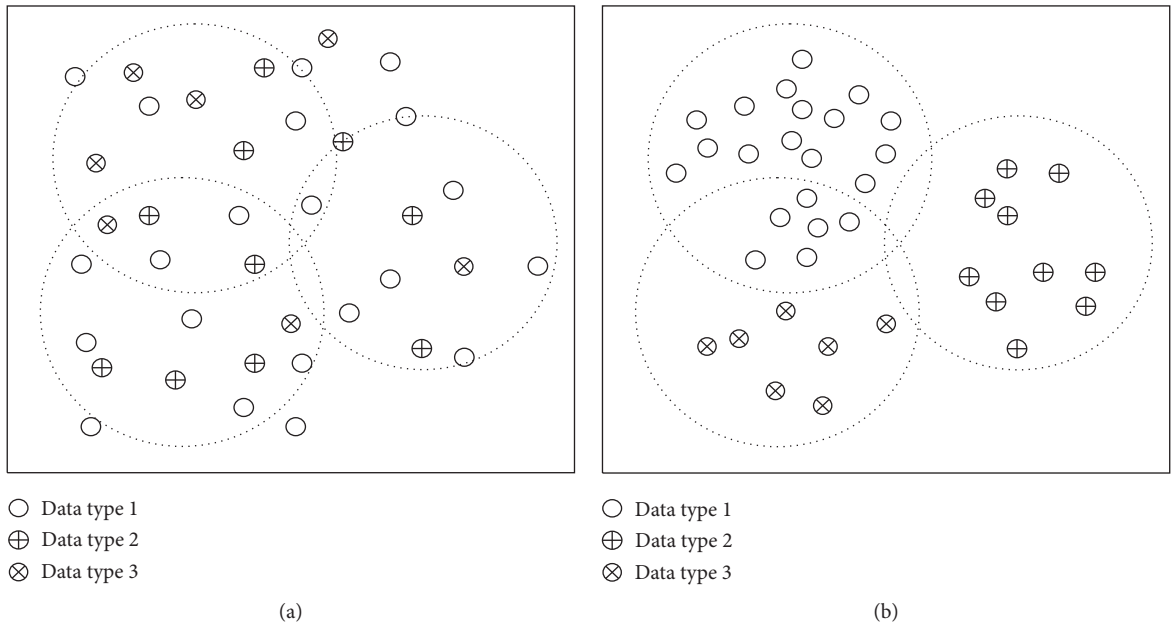


Figure 3: Schematic diagram of data classification results. (a) Before classification. (b) After classification.

the average code length of the data will not change, so this advantage is the coding efficiency uniqueness, which can significantly reduce the data storage space and improve the speed of the data query in the storage mode. In addition, the Huffman algorithm constructs the codeword with the shortest average length of different characters on the basis of the character appearance probability, which is more accurate. After the above classification processing, the data is transformed by wavelet decomposition method, and then Huffman coding. In the process of data compression using the combination of these two methods, the scale of wavelet decomposition should be smaller to reduce the amount of calculation of wavelet transform; binary coding of the transformed data can further improve the compression ratio [19, 20].

Huffman coding algorithm adopts optimized static coding technology, and the binary tree generated by the algorithm has the minimum weighted sum. The algorithm first arranges all the data in descending probability order, establishes a list, and then constructs a tree from bottom to top. Figure 4 is the flowchart of Huffman algorithm data compression.

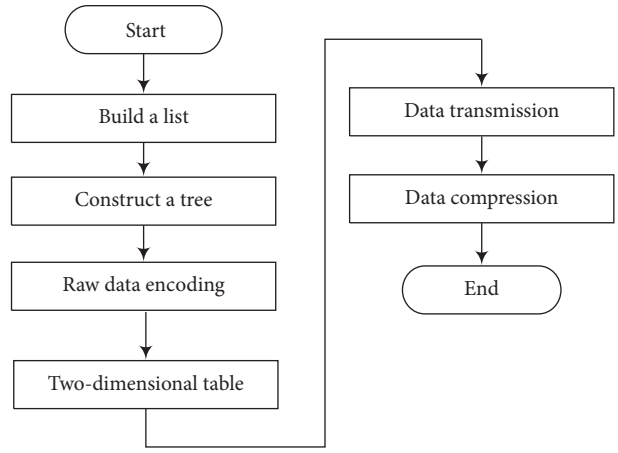


Figure 4: Flowchart of unbalanced big data compression.

According to the unbalanced big data compression process as shown in Figure 4, Huffman algorithm of each leaf is placed in a tree, and then the tree determines the coding of the original data. All the obtained codes form a two-

dimensional table. The size of the table is very small relative to the original data. The two-dimensional table and the encoded data are stored together or sent to the remote end through the communication network. The decoder does not need to traverse the tree but decompress it by looking up the table [21, 22].

2.4. Unbalanced Big Data Encryption. As users pay more and more attention to personal privacy, only realizing data processing can no longer meet the actual needs of users. Therefore, in order to meet the needs of user data security, it is necessary to encrypt the compressed data [23, 24]. The general chaotic encryption algorithm implements data encryption by superimposing a deterministic chaotic sequence on the original text. The decryption party uses the chaotic sequence reproduction method to achieve decryption. This approach is actually a process of superimposing white noise on the baseband signal. After smoothing and denoising, methods such as least squares processing and Kalman filtering can get the approximate original text, which reduces the confidentiality [25, 26]. At the same time, the chaotic sequence reappears in synchronization with the original text, which increases instability factors. The encryption algorithm proposed in this paper successfully solves these problems. The algorithm transforms the original data into a chaotic random sequence in the interval (0, 1) through a series of nonlinear transformations, which means that the baseband signal becomes white noise, so that the smoothing algorithm is useless. In addition, due to the good independence of the chaotic sequence itself, it increases the difficulty of being illegally deciphered [27, 28].

Suppose that the basic function of logistic mapping is

$$q(x) = w_t\delta_t + w_c\delta_c(D_h - S_h). \tag{1}$$

Among them, w_t and w_c represent the odd/even correlation characteristics of logistic map chaotic spread spectrum sequence, respectively; δ_t and δ_c respectively represent the influence of the finite length effect of the sequence on the odd/even correlation characteristics; D_h and S_h represent the balance rate of unbalanced data set and the balance rate of balanced data set, respectively.

The inverse mapping of formula (1) is

$$h = \frac{1}{2} \pm \sqrt{1 - \frac{q}{x}}. \tag{2}$$

When $0 < x \leq 1$, the Logistic map shows the desired chaotic characteristics, and this interval is the chaotic zone. Because the logistic mapping is simple in form and easy to analyze, the encryption algorithm described in this article uses it as the basic mapping. The bifurcation parameter c and the iteration parameter v of the mapping will be used as the encryption algorithm key, and their expressions are

$$\begin{aligned} c &= \sum_{i}^{N} (y_i - \hat{y}_i)^2 \\ v &= \sum_{i=1}^{N} t_i p_i \times T. \end{aligned} \tag{3}$$

Among them, y_i and $\hat{y}_i$ respectively represent the initial value and critical value of the bifurcation parameter; t_i represents the periodic response; p_i represents the chaotic response; T represents the parameter interval corresponding to the chaotic response.

After unbalanced big data is encrypted into ciphertext, it must be restored to the original text through the corresponding decryption algorithm. But logistic inverse mapping is a one-to-many mapping. To restore it, the one-to-one mapping problem of the mapping must be solved. This paper adopts the method of fixed word length discretization and adding disturbance term, which can successfully solve this problem. Finally, the corresponding encryption vector and decryption vector are generated from the key, and the original text encryption and ciphertext decryption can be performed.

2.5. Big Data Elimination Algorithm. Data redundancy technology refers to a large amount of data with the same content in the process of storing data, and the process of deleting redundant files and data blocks through repeated data detection, so that only unique data is stored in the system [29]. Compared with the traditional data compression technology, the redundancy elimination algorithm in the nonequilibrium big data encryption technology can not only eliminate the data redundancy in the files, but also eliminate the data redundancy between the files in the datasets, effectively reduce the storage space of big data, and provide new ideas for the storage and processing of big data [30].

2.5.1. The Basic Principle of the Deduplication Algorithm. Data deduplication technology detects duplicate data objects in the data stream based on the redundancy of the data itself, only transmits and stores the only data object copy, and uses a pointer to the unique data object copy to replace other duplicate copies. Data deduplication technology is committed to saving storage space and network bandwidth resources. In the process of unbalanced big data storage, the introduction of data deduplication technology can optimize the storage space of unbalanced big data and eliminate the same files or data blocks in the big data set, so as to reduce the workload of encryption processing [31, 32]. On the other hand, data deduplication effectively compresses the data, reduces the amount of data transmitted by the network, and reduces the bandwidth consumption. In order to improve the deduplication rate and deduplication efficiency of massive data deduplication system, this paper carries out big data deduplication.

Generally, the ratio of the number of bytes before deduplication B_{in} to the number of bytes after processing B_{out} is used to measure the data elimination ratio (DER) G_u, as shown in formula (5):

$$G_u = \frac{B_{\text{in}}}{B_{\text{out}}}. \tag{4}$$

G_u is usually determined by two factors: the type of partitioning strategy used and the average data block size. Although the data reduction rate shown in formula (5) takes into account the repeated data between data blocks after blocking and the data compression within a single data block, the metadata overhead is not considered. However, the metadata cost in the deduplication system cannot be ignored. Therefore, a correction formula for data reduction rate is proposed:

$$G_u' = \frac{B_{\text{in}}\left(U_i \times r\right)^2 Y}{B_{\text{out}}\left(U_j \times r\right)^2 Y'}. \tag{5}$$

Among them, U_i and U_j respectively represent the same data and similar data; Y and Y' respectively represent the space utilization rate before deletion and the space utilization rate after deletion; r represents the overhead size of metadata, and the calculation method is as follows:

$$r = \frac{\mu_k\left[G_u\left(x_k\right) - G_u'\right]}{\mu_k^{\text{avg}}}. \tag{6}$$

Among them, μ_k represents the size of the metadata; μ_k^{avg} represents the average value of the size of the metadata.

According to formula (6), there are two types of duplicate data: the same data and similar data. For these two types of data, duplicate data deletion technology is used to detect them, respectively. The specific processing methods are as follows:

(1) Same data detection technology: it divides the data set according to certain rules and replaces the same part of the data set with a pointer. According to different granularity, the same data detection technology is divided into complete file detection technology and data block detection technology. According to different blocking methods, data block detection technology is divided into fixed length block detection technology, content-based variable length block detection technology, and sliding window block detection technology [33].

(2) Similar data detection technology: according to the inherent similarity characteristics of data, the data set is detected. Delta coding technology is used to compress similar data to reduce data storage space and bandwidth occupied during transmission.

2.5.2. Implementation of Unbalanced Big Data Compatibility Cloud Storage. Section 2.5.1 discusses data deduplication technology, based on which a cloud platform is built to realize unbalanced big data compatibility cloud storage. The cloud platform has the ability of data reception, collection, and storage. It mainly obtains unbalanced big data resources through the Internet. The platform uniformly manages all kinds of unbalanced big data resources, establishes meta database and data catalog, and provides data browsing, query, download, and other services based on large screen for internal users and provides data sharing services for users based on the data distribution service submodule [34, 35]. Figure 5 shows the overall architecture of the cloud platform.

According to the overall architecture of the cloud platform in Figure 5, the software design of each platform is elaborated:

(1) Platform software support: cloud platform software support is the operating environment of various application software, and its components include operating systems, software operating platforms, application middleware, virtualized storage systems, and security systems. Among them, the data dynamic allocation strategy of the RSDO model supports the operation of the platform, which is mainly responsible for allocating unbalanced big data to different storage spaces according to the current storage situation of the data center to achieve the purpose of data balanced storage. The cloud scheduling service monitors the data of the entire cloud platform at the software support layer and provides task scheduling services for the storage of data in the cloud platform.

(2) Data distribution service submodule: the data distribution service submodule is the main portal for users to access the data storage center, and it is also the main platform for the center to provide external cloud services such as data sharing and application promotion. Users use the data resource cataloging service provided by the platform to search, query the metadata information of various resources, and understand the basic form of the data concerned. In addition, the system forms a data storage set through the classification and integration of data attribute characteristics, which is convenient for users to locate and understand the data of interest more quickly [36].

(3) Compatibility: the design of the unbalanced big data storage platform needs to organize the underlying database information in order to improve the compatibility of the data in the platform and avoid problems such as information corruption during the data storage process. Users can access the database uniformly through the data management portal and use the big data access unified interface to extract and store the required data. Finally, the data can be cached at the data relationship mapping layer and stored permanently in the bottom layer of the database. In addition, data storage platforms usually use nonrelational data characteristics to classify cached data to achieve data expansion and conversion, so as to optimize the data communication mode of the client. In the case of a relatively large

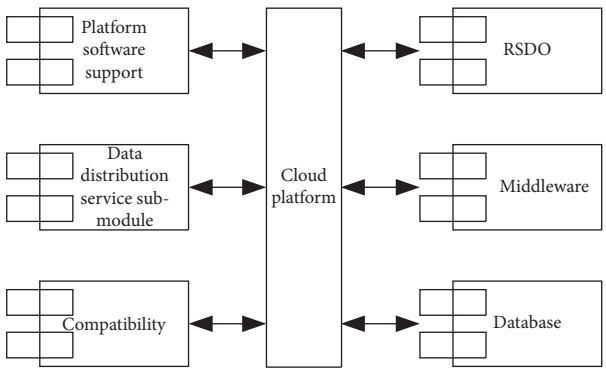


Figure 5: Schematic diagram of the overall architecture of the cloud platform.

amount of data, the cluster method needs to be used in the database cache system to optimize and store temporary data, so as to achieve the design goal of converting compatible large amounts of data.

Based on the above-built cloud platform, perform unbalanced big data-compatible cloud storage, and the specific steps are as follows:

(1) Assuming that the product of the CPU processing frequency, number of processors, and number of CPU cores of the a-th server in the cloud platform are represented by F_{all}, the expression is

$$F_{\text{all}} = F^{a}_{\text{CPU}} \cdot E_n \cdot L_{\text{CPU}}. \tag{7}$$

Among them, F^{a}_{CPU} represents the CPU processing frequency of the a-th server; E_n represents the number of processors; L_{CPU} represents the number of CPU cores.

(2) Assuming that the memory capacity of the cloud platform is represented by C_α, the disk capacity is represented by C_β, the disk read and write rate is represented by V_ε, and the network bandwidth throughput is represented by ∂_ω; then, the ratio between each parameter and the maximum value of the corresponding parameter of all servers is

$$\begin{cases} F^{a}_{\text{CPUC}} = \dfrac{C_\alpha}{C_{\alpha\max}}, \\ E_n = \dfrac{C_\beta}{C_{\beta\max}}, \\ L_{\text{CPU}} = \dfrac{V_\varepsilon}{V_{\varepsilon\max}}, \\ F_{\text{all}} = \dfrac{\partial_\omega}{\partial_{\omega\max}}. \end{cases} \tag{8}$$

The above ratio can reflect the performance of a certain parameter of the a-th server in the cluster server. The following uses the weighted average method to integrate these data together to calculate the overall performance value ℓ_a of the a-th server:

$$\ell_a = a_1 \times F^{a}_{\text{CPU}} + a_2 \times E_n + a_3 \times L_{\text{CPU}} + a_4 \times F_{\text{all}}. \tag{9}$$

From this, the weight value of the a-th server relative to the entire server cluster can be obtained:

$$\theta = \sum_{a \in A}^{n} \ell_a \times O_a. \tag{10}$$

According to the real-time performance of the cluster server statistics, define CPU_a as the actual CPU utilization of the a-th server, RAM_a represents the actual utilization of the memory, Disk_a represents the actual utilization of the disk capacity, Rate_a represents the average disk read and write rate, and Throughput_a represents the actual throughput of the network bandwidth, and it can calculate the actual load weight threshold of the server in the cloud platform:

$$\theta' = \frac{\theta}{\sum_{a \in A}^{n} \ell'_a} \times O'_a. \tag{11}$$

This article collects, classifies, compresses, and encrypts unbalanced big data and builds a cloud platform for data objects. In addition, the technology of eliminating redundancy is applied to unbalanced big data storage, so that running a cloud platform for data storage is self-contained. Features such as adaptability, fast response, and high efficiency were studied.

3. Simulation Experiment

In order to verify the effectiveness of the unbalanced big data compatibility cloud storage method based on redundancy elimination technology proposed in this paper, simulation experiments are carried out. In the experiment, reference [6] method and reference [7] method are used as comparison methods. The data redundancy elimination effect and data storage space are taken as the experimental indexes. The data redundancy elimination effect specifically refers to the duplicate data deletion rate and deletion rate.

3.1. Experimental Data Set. In this paper, six data sets are selected in the public machine learning database UCI, and one of the data sets in each data set is regarded as a positive class, and the rest are classified as negative classes, thus forming an unbalanced data set with different degrees of imbalance. In order to ensure the accuracy of the experimental results, all data sets are preprocessed with 0 mean and standard deviation 1. The total number of data is 74589, and the total size is 1130070.49 KB. The average size of the data block in the experiment is 8 KB, and the sliding window size is It is 48 Byte. Simulate the MapReduce computing environment in a stand-alone environment to verify the effectiveness of the method in this paper.

The specific experimental data set description is shown in Table 1.

Table 1: Experimental data set.

Data set	Number of samples/pieces	Category ratio (%) (positive/negative)
Letter	10237	5.24
Image	9892	10.29
Iris	11369	20.31
Class	10543	4.17
Wine	8446	9.30
Pima	24102	12.09

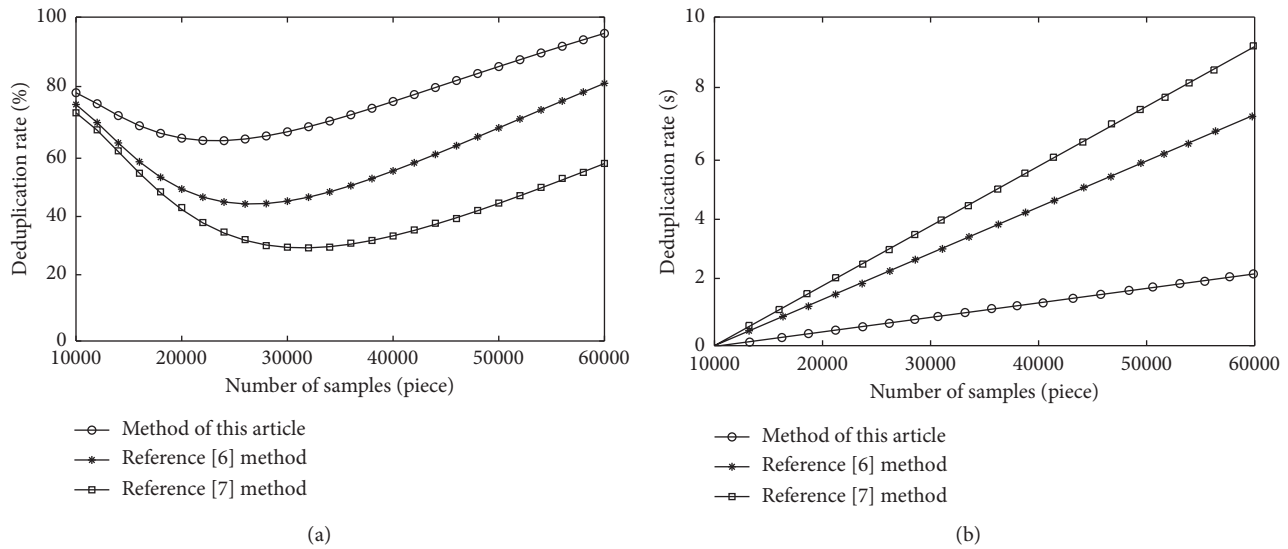


Figure 6: Comparison of data redundancy elimination effect. (a) Deduplication rate. (b) Deduplication speed rate.

3.2. Analysis of Experimental Results

3.2.1. Data Elimination Effect. The performance of this method, reference [6] method, and reference [7] method in data deduplication rate and deduplication speed rate is compared.

The performance test methods of the three methods are as follows: given a certain amount of data, stored in different directories and files, read these data, then store them in different methods, count the space occupied after storage, and calculate the deduplication rate. At the same time, count the total time from the beginning of reading data to the completion of the last storage, and calculate the deduplication speed rate. The data source is from Table 1, and the results are shown in Figure 6.

It can be seen from Figure 6(a) that, after using the method in this paper to eliminate redundancy, the deduplication rate is higher than that in reference [6] method and reference [7] method, and the highest value of the deduplication rate of the method in this paper is 96%. It can be seen from Figure 6(b) that the deduplication time is much lower than reference [6] method and reference [7] method, and the highest value of the duplicate data deletion time of the method in this paper is only 2.2s. The comparison results show that the method in this paper can not only ensure a high deduplication rate of duplicate data, but also improve the deduplication speed rate of duplicate data.

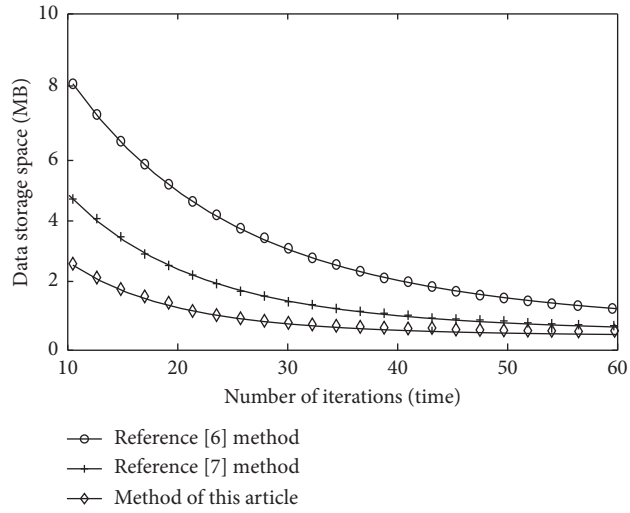


Figure 7: Data storage space.

3.2.2. Data Storage Space. In order to further verify the effectiveness of the method in this paper, the three methods are compared with the data storage space as an experimental indicator, and the results are shown in Figure 7.

It can be seen from the curve trend in Figure 7 that when processing the same amount of unbalanced big data, the data

storage space occupied by the method in this paper is significantly lower than that of reference [6] method and reference [7] method; although with iteration with the increase in the number of times, the gap between different methods is gradually narrowing, but the method in this paper can still maintain its advantages, indicating that it has a significant advantage in unbalanced big data storage space.

4. Conclusion

In order to improve the elimination capacity of the redundant data and reduce the storage space of the data occupied, a nonequilibrium big data-compatible cloud storage method based on the redundant elimination technology is proposed. Big data acquisition platform was designed based on Hadoop and NoSQL technologies, and data was collected in parallel using FPGA parallel acquisition technology. Based on this platform, the collected data is processed using the parallel utilization of the classifier classification. The Huffman algorithm is used to compress the redundancy elimination algorithm in the nonequilibrium large data encryption technique. Design a cloud platform to store big data. Experimental results show that the proposed method works better than traditional data storage methods, has high deduplication rate, and solves the problem of large data storage.

Data Availability

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding this work.

References

[1] A. Singh, S. Garg, K. Kaur, S. Batra, N. Kumar, and K. K. R. Choo, "Fuzzy-folded bloom filter-as-a-service for big data storage in the cloud," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 4, pp. 2338–2348, 2019.

[2] D. Gupta and R. Rani, "A study of big data evolution and research challenges," *Journal of Information Science*, vol. 45, no. 3, pp. 322–340, 2019.

[3] H. Chergui and C. Verikoukis, "Big data for 5G intelligent network slicing management," *IEEE Network*, vol. 34, no. 4, pp. 56–61, 2020.

[4] K. Kaur, S. Garg, G. Kaddoum, E. Bou-Harb, and K. K. R. Choo, "A big data-enabled consolidated framework for energy efficient software defined data centers in IoT setups," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 4, pp. 2687–2697, 2020.

[5] S. Kalaivani, C. Tharini, K. Saranya, and K. Priyanka, "Design and implementation of hybrid compression algorithm for personal health care big data applications," *Wireless Personal Communications*, vol. 113, no. 1, pp. 599–615, 2020.

[6] Z. P. Huang, L. Wang, S. X. Zhang, T. Z. Yu, and Q. R. Zhang, "Redundant optical fiber data storage optimization based on traditional genetic algorithm and data compression algorithm," *Laser Journal*, vol. 40, no. 3, pp. 135–139, 2019.

[7] P. Xie, C. C. Yang, S. Xiong, L. S. He, and X. D. Zhou, "Design and optimization of spatial vector data storage model based on HBase," *Acta Geodaetica et Cartographica Sinica*, vol. 49, no. 10, pp. 1365–1373, 2020.

[8] T. Chen, D. S. Wang, Z. J. Wang, and W. Liu, "An efficient data storage method in digital library based on compressed sensing," *Library Journal*, vol. 38, no. 09, pp. 4–11, 2019.

[9] D. de Santana Nunes, J. L. V. de Brito, and G. N. Doz, "A low-cost data acquisition system for dynamic structural identification," *IEEE Instrumentation and Measurement Magazine*, vol. 22, no. 5, pp. 64–72, 2019.

[10] F. Santoni, A. De Angelis, A. Moschitta, and P. Carbone, "A multi-node magnetic positioning system with a distributed data acquisition architecture," *Sensors*, vol. 20, no. 21, p. 6210, 2020.

[11] A. Elzanaty, A. Giorgetti, and M. Chiani, "Limits on sparse data acquisition: RIC analysis of finite Gaussian matrices," *IEEE Transactions on Information Theory*, vol. 65, no. 3, pp. 1578–1588, 2019.

[12] E. Grimberg, A. Botzer, and O. Musicant, "Smartphones vs. in-vehicle data acquisition systems as tools for naturalistic driving studies: a comparative review," *Safety Science*, vol. 131, no. 6, Article ID 104917, 2020.

[13] A. Mehto, S. Tapaswi, and K. K. Pattanaik, "Virtual grid-based rendezvous point and sojourn location selection for energy and delay efficient data acquisition in wireless sensor networks with mobile sink," *Wireless Networks*, vol. 26, no. 5, pp. 3763–3779, 2020.

[14] M. Rapin, F. Braun, A. Adler et al., "Wearable sensors for frequency-multiplexed EIT and multilead ECG data acquisition," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 3, pp. 810–820, 2019.

[15] M. Irfan, Z. Jiangbin, M. Iqbal, Z. Masood, M. H. Arif, and S. R. u. Hassan, "Brain inspired lifelong learning model based on neural based learning classifier system for underwater data classification," *Expert Systems with Applications*, vol. 186, no. 1, Article ID 115798, 2021.

[16] D. Griffiths and J. Boehm, "A review on deep learning techniques for 3D sensed data classification," *Remote Sensing*, vol. 11, no. 12, p. 1499, 2019.

[17] N. Gomathi and N. P. Karlekar, "Ontology and hybrid optimization based SVNN for privacy preserved medical data classification in cloud," *The International Journal on Artificial Intelligence Tools*, vol. 28, no. 3, Article ID 1950009, 2019.

[18] F. Bensaid and A. M. Alimi, "Online feature selection system for big data classification based on multi-objective automated negotiation," *Pattern Recognition*, vol. 110, no. 1, Article ID 107629, 2020.

[19] R. Corda and C. Perra, "Hologram domain data compression: performance of standard codecs and image quality assessment at different distances and perspectives," *IEEE Transactions on Broadcasting*, vol. 66, no. 2, pp. 292–309, 2020.

[20] J. Uthayakumar, T. Vengattaraman, and P. Dhavachelvan, "A new lossless neighborhood indexing sequence (NIS) algorithm for data compression in wireless sensor networks," *Ad Hoc Networks*, vol. 83, no. 2, pp. 149–157, 2019.

[21] M. Tegmark and T. Wu, "Pareto-optimal data compression for binary classification tasks," *Entropy*, vol. 22, no. 1, p. 7, 2019.

[22] D. Greenfield, V. Wittorff, and M. Hultner, "The importance of data compression in the field of genomics," *IEEE Pulse*, vol. 10, no. 2, pp. 20–23, 2019.

[23] R. Geetha, T. Padmavathy, T. Thilagam, and A. Lallithasree, "Tamilian cryptography: an efficient hybrid symmetric key

encryption algorithm," *Wireless Personal Communications*, vol. 112, no. 1, pp. 21–36, 2020.
[24] M. S. Khoirom, D. S. Laiphrakpam, and T. Tuithung, "Audio encryption using ameliorated ElGamal public key encryption over finite field," *Wireless Personal Communications*, vol. 117, no. 2, pp. 809–823, 2021.
[25] X. Wang, N. Guan, and J. Yang, "Image encryption algorithm with random scrambling based on one-dimensional logistic self-embedding chaotic map," *Chaos, Solitons & Fractals*, vol. 150, no. 3, Article ID 111117, 2021.
[26] J. Mou, F. Yang, R. Chu, and Y. Cao, "Image compression and encryption algorithm based on hyper-chaotic map," *Mobile Networks and Applications*, vol. 5, no. 3, pp. 1–13, 2019.
[27] I. Yasser, M. A. Mohamed, A. S. Samra, and F. Khalifa, "A chaotic-based encryption/decryption framework for secure multimedia communications," *Entropy*, vol. 22, no. 11, p. 1253, 2020.
[28] X. Wang, S. Lin, and Y. Li, "A chaotic image encryption scheme based on cat map and MMT permutation," *Modern Physics Letters B*, vol. 33, no. 27, Article ID 1950326, 2019.
[29] S. Kumar and V. K. Chaurasiya, "A strategy for elimination of data redundancy in Internet of things (IoT) based wireless sensor network (WSN)," *IEEE Systems Journal*, vol. 13, no. 2, pp. 1650–1657, 2019.
[30] Z. A. Pan, J. F. Wang, and M. F. Wang, "Research on cloud storage method of multi process spatial data based on SDN technology," *Computer Simulation*, vol. 38, no. 5, pp. 375–379, 2021.
[31] X. Xu, L. Liu, X. Zhang, W. Guan, and R. Hu, "Rethinking data collection for person re-identification: active redundancy reduction," *Pattern Recognition*, vol. 113, no. 4, Article ID 107827, 2021.
[32] S. Chandak, K. Tatwawadi, I. Ochoa, M. Hernaez, and T. Weissman, "SPRING: a next-generation compressor for FASTQ data," *Bioinformatics*, vol. 35, no. 15, pp. 2674–2676, 2019.
[33] M. Besseghier, A. B. Djebbar, A. Zouggaret, and I. Dayoub, "Joint channel estimation and data detection for OFDM based cooperative system," *Telecommunication Systems*, vol. 73, no. 4, pp. 545–556, 2020.
[34] Z. Qian, X. Wang, X. Liu, X. Xie, and T. Song, "An approach to dynamically assigning cloud resource considering user demand and benefit of cloud platform," *Computing*, vol. 102, no. 8, pp. 1817–1842, 2020.
[35] J. Yun, K.-W. Park, D. Koo, and Y. Shin, "Lightweight and seamless memory randomization for mission-critical services in a cloud platform," *Energies*, vol. 13, no. 6, p. 1332, 2020.
[36] L. S. Subhash and R. Udayakumar, "Sunflower whale optimization algorithm for resource allocation strategy in cloud computing platform," *Wireless Personal Communications*, vol. 116, no. 4, pp. 3061–3080, 2021.

Hindawi
Scientific Programming
Volume 2022, Article ID 6461538, 11 pages
https://doi.org/10.1155/2022/6461538

Research Article

Animation Design Based on 3D Visual Communication Technology

Feng Shan and Youya Wang

Wanjiang College, Anhui Normal University, Wuhu 241008, China

Correspondence should be addressed to Feng Shan; 1301220955@pku.edu.cn

Received 10 November 2021; Revised 10 December 2021; Accepted 21 December 2021; Published 5 January 2022

Academic Editor: Baiyuan Ding

The depth synthesis of image texture is neglected in the current image visual communication technology, which leads to the poor visual effect. Therefore, the design method of film and TV animation based on 3D visual communication technology is proposed. Collect film and television animation videos through 3D visual communication content production, server processing, and client processing. Through stitching, projection mapping, and animation video image frame texture synthesis, 3D vision conveys animation video image projection. In order to ensure the continuous variation of scaling factors between adjacent triangles of animation and video images, the scaling factor field is constructed. Deep learning is used to extract the deep features and to reconstruct the multiframe animated and animated video images based on visual communication. Based on this, the frame feature of video image under gray projection is identified and extracted, and the animation design based on 3D visual communication technology is completed. Experimental results show that the proposed method can enhance the visual transmission of animation video images significantly and can achieve high-precision reconstruction of video images in a short time.

1. Introduction

3D visual communication technology is immersive, interactive, and conceivable for users to create high-quality visual experience. Video based on 3D visual communication technology needs to go through different processes such as content making, coding compression, network transmission, and terminal display. In recent years, visual communication technology has been widely used in various fields because of its function of fast and real-time measurement of information [1]. With the development of multimedia technology, visual communication technology is widely used in the field of video and animation image management, such as network media operation, animation release, etc. Film and television animation and video images are developing constantly in the field of film and television animation and video images. The core of film and television animation and video is visual communication, and it is difficult to obtain target film and television animation and video images in complex background in real time and accurately. It is a common problem in multiframe film and television animation and video visual communication [2].

Reference [3] proposes an innovative design of visual communication system for animated figures and images in virtual reality environment. In the hardware design, the adaptation parameters of the main board of the renderer are designed to optimize the experience of visual communication. In the software design, by introducing Sobel edge operator, the gray function is established to solve the gradient amplitude, and the threshold is selected to compare it to complete the recognition and thinning of the edge data of animated character graphics and images. Design the motion capture module, establish the behavior control model, generate and manage the motion capture files, and complete the overall design of the system. Reference [4] proposes a new optimization method—plane visual communication

effect optimization method based on wavelet change. The image is decomposed by wavelet, and the wavelet is reconstructed. In the process of reconstruction, the modulus diagram and phase angle diagram are calculated, and the edge images of each scale are extracted. The corresponding edge points of the semireconstructed image are enhanced through the edge image. On the above basis, the graphic beautification vector of slip model is used for operation. In order to simplify the operation, the above operation method is transformed into simple mathematical operation, and the visual communication effect is optimized through the reflected light graphic mode. Reference [5] proposes a new automatic graphic language arrangement algorithm in visual communication design. The display size of buffer image in visual communication design is calculated by fixed value method, and the minimum number of layout and the maximum surface utilization are taken as the layout objective function. Ant colony algorithm is used to solve it, get the optimal solution, and obtain the best display position of graphics in visual communication design. The description of graphic language is realized by rule-based syntax description and ASM (abstract state machine) semantic description. The parallel process and selection process of graphic language in visual communication design are described by using parallel and selection marks, and the parallel process and selection process are used to realize the automatic arrangement of graphic language. The proposed algorithm is programmed and run in C language to achieve the purpose of automatic arrangement of graphic language.

However, when the above image visual communication technologies ignore the depth synthesis of image texture, the image visual effect is not ideal. Therefore, a film and television animation design method based on 3D visual communication technology is proposed. Build the video processing process under 3D visual communication technology, and mainly process film and television animated video through 3D video content production, and server and client links; based on the visual distance between 2D image and scene, calculate the zoom factor of 3-dimensional scene transformation; build a deep learning-based deep feature extraction model PCANet, filter images and extract features of film and television animation video images, use PCAN deep network to create a low-resolution feature dictionary, use high-resolution video image reconstruction method, realize LR animated video images and HR animation video images, and introduce nonlocal similarity empirical constraints to optimize HR images; analyze gray and detail features based on wavelet transform, extract the video image features, separate the prospect and background of animated video images, transform K-L (Karhunen–Loeve) as the core, dynamically identify the similarity between evaluation features, and complete the animation design based on three-dimensional visual communication technology. The experimental results show that the proposed method can significantly improve the visual transmission of animated video images and improve the speed of high-precision video image reconstruction.

2. Film and Television Animation Design Based on 3D Visual Communication Technology

3D visual communication technology mainly uses the computer to simulate the environment, to make people feel immersive. Three-dimensional visual communication technology, on the basis of 2D images, reads the main information, compresses and decomposes, and transforms it into three-dimensional images, increasing the measurement and display function of arbitrary sections. In the human-computer interaction module, we can zoom in the three-dimensional image to facilitate users to browse and change the image.

2.1. Video Processing Process under 3D Visual Communication Technology. Figure 1 shows the flowchart of video processing under 3D visual communication technology.

The 3D visual communication technology can be divided into three parts: 3D visual communication content production, server and client, and 3D visual communication content production which is the core of the whole video processing process [6]. In the production of 3D visual communication content, a set of cameras or a set of camera devices and audio sensors containing multiple cameras and sensors are used to capture the sound and visual scene in the real physical world. Cameras and sensors, for example, produce a set of digital video and audio signals. In general, the camera captures content within 360° of the periphery of the unit [7, 8].

Aiming at the video processing under the 3D visual communication technology, the 3D visual communication video image at the same time is combined with mosaic, projection mapping, and frame texture to generate a package, as shown in Figure 2.

2.2. The Scaling Factor of the 3D Scene Transformation. After determining the triangular meshes on a two-dimensional plane through screen projection, a continuous scaling field shall be assigned to each triangular facet; that is, a scaling factor shall be assigned to each triangular facet, and the scaling factor between adjacent triangular facets shall be continuously variable. The specific process is as follows [9, 10]:

S_i is used to represent the distance from the triangle to the viewpoint in 3D scene. S_i is defined as visual distance, and all S_i can be integrated into a threshold of discrete function S_i.

Determine the apparent distance S_i and S_{i+1} between any adjacent triangles i and $i + 1$, and the difference ∂_i between the two can be expressed as follows:

$$\partial_i = \left|S_i - S_{i+1}\right|. \tag{1}$$

Obtain the upper limit value max ∂ of ∂_i through formula (1).

When the value of min Scale is 0.4, the sample size changes gradually from 0.4–1 to 0.4. Based on the value of min Scale, pyramid shrinkage is applied to a given texture

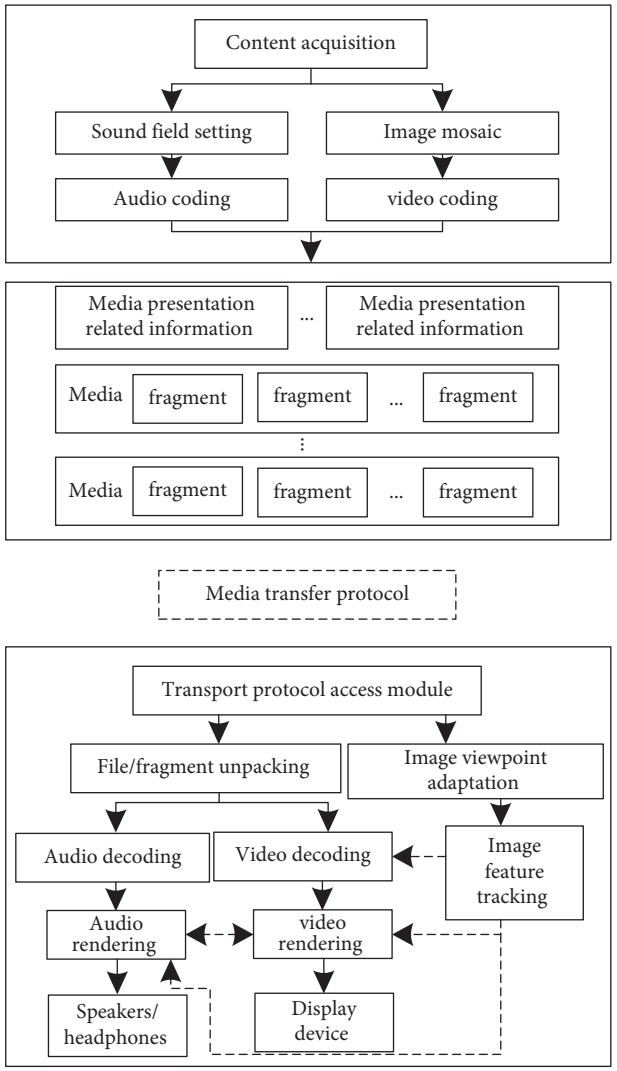


Figure 1: Video processing flowchart under 3D visual communication technology.

sample drawing to obtain a multilevel progressive shrinkage sample drawing, which is used as a scaling factor in the process of texture synthesis to obtain the basis for the corresponding sample drawing [11, 12].

In order to make the neighboring triangle mesh correspond to the neighboring layer, based on max ∂, set up: the value of S function can only be different one layer within the sample layer if the difference is max ∂ that is, according to max ∂ to divide the value range of S function, the value of the function is divided into different layers according to max ∂, and the layers corresponding to the neighboring triangle mesh are also adjacent, thus ensuring the continuity of texture in the process of composition.

The value of S_i for each triangle can obtain a value of Scale (min Scale $\leq$ Scale ≤ 1) according to the above setting, which is the value of Scale corresponding to the scaling of the meta-texture sample image [13]. Based on this, each Scalevalue corresponds to a Scale-fold shrunk layer, and a Scale-fold shrunk image can be mapped to a Scale-fold shrunk film and TV animation image, where each triangle has a scaling factor.

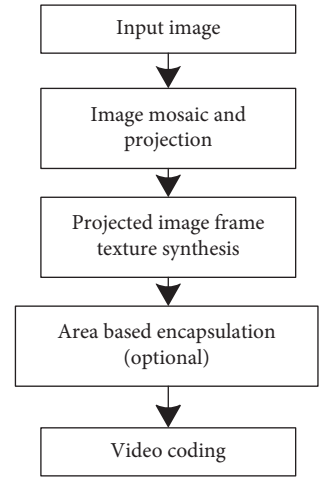


Figure 2: 3D visual communication animation video image mosaic, projection, and area-based packaging.

2.3. Deep Feature Extraction Model Based on Deep Learning. PCANet is a relatively simple method of deep learning based on the theory of deep learning and convolutional neural networks [14, 15]. The PCANet structure is composed of two PCA (principal component analysis) filter layers, a hash layer and a histogram calculation layer, which can extract the deep-seated features of high-resolution animation and video images. But unlike the usual networks, PCANet's filters are more computationally efficient. PCANet is not obtained through training, but through obtaining the mapping results of parts of high-resolution film and television animation video images, and then adopts the PCA (principal component analysis) method to extract the principal components of the high-resolution film and television animation video images, each of which is an independent filter [16]. When the neural network of the filter is used to extract the features, the optimal weights can be obtained without iterative operation; thus, the calculation time is reduced [17]. In order to further discuss the advantages of PCANet depth network in extracting the features of high-resolution animated cartoon video images, a mapping eigenvalue of N, 8×8 and 1 moving distance is set up for the input of animated video images [18, 19]. The main parts of the mapping features L_1 were computed using the PCANet algorithm, and a two-layer filter filters L_1. The steps are as follows:

2.3.1. First Layer: PCA Filter Layer 1. For the N training picture L_i^N with size of $m1m2$, PCANet selects a $k1 \times k2$ window, and each training image is compared by row and column to obtain the local features of the image. Get $m1 \times m2$ picture blocks of $k1 \times k2$ size. Dislodge the mean value of the picture block, the output behind dislodging the mean value of L_i:

$$\overline{A}_i = \left[\overline{a}_{i,1}, \overline{a}_{i,2}, \cdots, \overline{a}_{i,m_1,m_2}\right]. \quad (2)$$

In formula, each column of the matrix $\overline{A}_i$ represents a block of image, and each column contains $k1 \times k2$ elements. There are $m1 \times m2$ columns in total. After performing the above operations separately for all N training images, a new

data matrix A was obtained: $A = [\overline{A}_1, \overline{A}_2, \cdots, \overline{A}_N] \in R^{k1k2\times L}$, and A has $N \times m1 \times m2$ columns. Then, the A matrix was PCA and the ahead L_i eigenvectors of the eigenvector matrix W were taken as the convolutional kernel of the filter group.

$$W_{L_i} = ma_{k1k2}\left(q_l\left(AA^T\right)\right) \in R^{k1k2}, l = 1, 2, \cdots, L_i. \quad (3)$$

In formula, $q_l(AA^T)$ represents the l-th eigenvector of the matrix AA^T. ma_{k1k2} is a function, representing rearranging each of the L_i eigenvectors into the matrix. Each eigenvector contains $k1 \times k2$ elements, amount to get L_i convolutional nuclei of length $k1$ and width $k2$. Finally, for each input image L, the L_i convolutional nuclei were filtered separately. Thus, the first layer of the PCA filter group is completed.

2.3.2. Second Layer: PCA Filter Layer 2. The second layer of the PCA filter group is the same principle as the first layer. This yields the convolutional kernel of the second layer:

$$W_{L_i}^2 = ma_{k1k2}\left(q_{l_2}\left(BB^T\right)\right) \in R^{k1k2}, l_2 = 1, 2, \cdots, L_i^2. \quad (4)$$

For each input image L_i^2 in the second layer, the PCA filter group is all complete through the filter $W_{L_i}^2$ output features.

Finally, the L_i and L_i^2 feature mappings are obtained in the same way as CNN. PCANet is characterized by these L_1 mapping features, so it can provide a reliable data for the film and television animation video image processing [20]. PCANet uses learning to get multilayer network filters. After the input video image is filtered through two layers of filters, it can have very high-dimensional data output and can be regarded as the input of video image characteristics of the film and television animation and be used for the reconstruction of high-resolution video image of the film and television animation [21].

Therefore, when PCANet extracts the features of film and television animation video images, it is actually a direct processing of the pixels of the film and television animation video images, the operation stage has joined the process of blocking, and the number of film and television animation video image block output from the deep network has increased [22]. PCANet uses the deep network learning process to extract the features of the animation video image, and the video image features extracted by PCANet and the features obtained by artificial rules are more abundant in detailed information and more prominent in texture structure, which provides rich prior knowledge for subsequent reconstruction and fills in the details of the animation video image of low-resolution film and television so as to facilitate the reconstruction of the animation video image of super-resolution film and television.

2.4. Image Reconstruction Method of Multiframe Animation Video Based on Visual Communication. Combining the advantages of PCANet depth network and sparse representation-based reconstruction method, this paper proposes a visual communication-based multiframe animation video image reconstruction method [23, 24]. For the multiframe animation video image reconstruction method, it is assumed that the high-resolution image and the low-resolution image are sparse expressions for their respective dictionaries, and then, the image sample features are obtained by PCANet depth network. D_h represents the dictionary feature of a high-resolution image, and D_1 represents the dictionary feature of a low-resolution image. In the stage of high-resolution image reconstruction, the low-resolution image is processed in the same way, the deep-seated feature is extracted by PCANet, the sparse representation of D_1 above the low-resolution dictionary feature of the high-resolution image is obtained, the coefficients expressed by the sparse D_1 feature of the low-resolution image are directly applied on D_h, and the corresponding high-resolution feature image is obtained [25]. Low-resolution video animation images are used to achieve high-resolution reconstruction. Through the PCANet depth network, we can get better film and TV animation video image features than the nondepth network. The deep feature dictionary can also improve the description ability of the feature dictionary and improve the quality of film and TV animation video image reconstruction significantly [26].

K blurred film and TV animation video image need to be sampled in the training data set, which is adjusted to the same size to get the corresponding low-resolution film and TV animation video image, which is combined into a model sample pair: $T = \{X_h, X_1\}$, $X_h = \{x_h^i\}_{i=1}^k$ represents high-resolution features, and $X_1 = \{x_h^i\}_{i=1}^k$ represents low-resolution features. Compute the matrix into blocks for all samples in the data set, and select a sliding window of size $k_1 \times k_2$, (normally, the pixel square window of the film and television animation video image used is 3, 5, or 7). After feature extraction for all the film and television animation video images is carried out through the aforesaid sliding window, the new data matrix X of column $N \times m \times n$ can be obtained, and each column of the matrix represents a film and television animation video image block with a total of $k_1 \times k_2$ elements.

To obtain the training sample of i high-resolution animation video image, the formula is

$$\overline{X}_{hi} = \left[\overline{x}_{hi,1}, \overline{x}_{hi,2}, \cdots, \overline{x}_{hi,mn}\right]. \quad (5)$$

Features were extracted with the above data matrix X, and the extracted features were regarded as feature samples in SCSR model and substituted into dictionary features of PCANet.

Suppose that T_{hi}^L represents the results of the high-resolution animated video image after sparse coding, the $H(\cdot)$ function is regarded as a unit step function, the results of sparse coding are quantized, the histogram coding is completed, and the deep detail features of the high-resolution animated video image are extracted, that is, [27, 28]

$$F_{hi} = \left[Bhist\left(T_{hi}^1\right), \cdots, Bhist\left(T_{hi}^{L_1}\right)\right]^T \in R^{\left(2^{L1}\right)L_1 B}. \quad (6)$$

Similarly, it is inferred that through the same processing process as the high-resolution film and television animation video image, the results of extracting the deep-seated detail features of the low-resolution film and television animation video image are given, namely,

$$F_{li} = \left[\text{Bhist}\left(T_{hi}^{1}\right), \cdots, \text{Bhist}\left(T_{hi}^{L_1}\right)\right]^T \in R^{(2^{L_1})L_1 B}. \quad (7)$$

Here, F_{hi} and F_{li} on behalf of the film and television animation video image feature extraction results; Bhist on behalf of the histogram coding; and B on behalf of the film and television animation video image segmentation block number of samples.

In this paper, the dictionary is trained in the SCSR framework by combining the sparse coding method. The goal is to obtain a set of dictionary pairs D_h and D_1 that can represent complex feature samples. Make K have the same sparse representation on D_h and D_1 for the deep-level detail features F_{hi} and F_{li} generated by film, television, animation, and video images, and F_{hi} and F_{li} have the same description coefficient, that is,

$$\begin{aligned} \{D_h, a\} &= \|F_{hi} - D_h \cdot a\| + \sum_{i=1}^{K} \lambda_i \|\alpha_i\|, \\ \{D_1, a\} &= \|F_{li} - D_1 \cdot a\| + \sum_{i=1}^{K} \lambda_i \|\alpha_i\|. \end{aligned} \quad (8)$$

In the above formula, $a = \{a_i\}_{i=1}^{k}$ represents sparse matrix and λ_i represents equilibrium coefficient [29]. In order to make the high-resolution film and television animation video image features and low-resolution film and television animation video image features have the same sparse description method as their respective dictionaries, the combined training method is adopted through formula (5) and (6), that is,

$$\{D_h, D_1, \alpha\} = \frac{1}{N}\|F_{hi} - D_h \cdot a\| + \frac{1}{M}\|F_{li} - D_1 \cdot a\| + \sum_{i=1}^{K} \lambda_i \|\alpha_i\|. \quad (9)$$

In the above formula, N and M, respectively, represent the dimension rearrangement of the column vectors by the element values of high- and low-resolution feature blocks of film, television, and animation video images, and $1/N$ and $1/M$ are the cost between D_h and D_1 of balanced formula (9). In order to facilitate subsequent calculation, formula (10) is reconstituted:

$$\{D_h, a\} = \sum_{i=1}^{K} \lambda_i' \|\alpha_i\|, \quad (10)$$

$$D_C = \begin{bmatrix} 1/\sqrt{N}\, D_h \\ 1/\sqrt{M}\, D_1 \end{bmatrix}, X_C = \begin{bmatrix} 1/\sqrt{N}\, F_h \\ 1/\sqrt{M}\, F_1 \end{bmatrix}, \lambda_I' = \left(\frac{1}{N} + \frac{1}{M}\right)\lambda_i. \quad (11)$$

Equation (11) is solved by iteration. Given dictionary D_C, calculate the sparse representation coefficient a_i of each data sample F_C on D_C, obtain matrix $a = \{a_i\}_{i=1}^{k}$, and finally update dictionary D_C through a.

After obtaining the dictionary pair D_C, the reconstruction of LR film and television animation video image Y and HR film and television animation video image X can be obtained by using the high-resolution film and television animation video image reconstruction method based on sparse regular model.

In the stage of reconstructing the video image of film and television animation, some noise exists in the video image of film and television animation due to the interference of external environment noise, and block effect and blurred artifact will appear in the reconstructed video image of film and television animation [30]. Considering that the conventional back-projection model cannot guarantee the quality of the reconstructed image, the details of the image can be better preserved by matching the similar blocks between the image and the image without the prior constraints of partial similarity. Relative to the image block effect and blurred artifacts after the reconstruction, the nonlocal similarity empirical constraint is introduced to optimize the HR image based on the global optimization of back-projection. The global and nonglobal constraint models in this paper are as follows [31, 32]:

$$X^* = \text{argmin}_X \|SHX - Y\|_2^2 + c\|X - X_0\|_2^2 + \gamma\|(I - W)X\|_2^2. \quad (12)$$

In the above formula, S represents sampling operation, H represents fuzzy filtering, $\|X - X_0\|_2^2$ represents global constraint term, $\|(I - W)X\|_2^2$ represents nonlocal self-similarity constraint term, I represents identity matrix, c and γ represent normalization parameters, W represents nonlocal weight matrix, and the element formula in W is as follows:

$$W_{ij} = \frac{1}{z(i)} \exp\left(-\frac{\left\|N(x_i) - N\left(x_j\right)\right\|_2^2}{h}\right). \quad (13)$$

In the above formula, $N(x_i)$ represents the i-th film and television animation video image block in the film and television animation video image, x_j represents the searched film and television animation video image block similar to $N(x_i)$, h represents the attenuation factor, and $z(i)$ represents the normalized value.

$$z(i) = \sum_{j} \exp\left(-\frac{\left\|N(x_i) - N\left(x_j\right)\right\|_2^2}{h}\right). \quad (14)$$

2.5. Dynamic Recognition of Frame Features of Film, Television, and Animation Video Images

2.5.1. Frame Feature Extraction of Video Image under Gray Projection. The first step of frame feature extraction is to analyze the features of gray level and detail in the video image based on wavelet transform and then to extract the feature of the video image by gray projection [33]. Finally, the problem of frame feature extraction of film and television animation video image evolves into the problem of foreground and background classification of film and television animation video image, and the separation coefficient is determined by the ratio function of the variance of feature distribution of film and television animation video image in the foreground and background region. The process of

extracting frame feature of film and television animation video image is as follows:

The input of preprocessed grayscale animation video images shall be represented by $f_k(x, y)$, the grayscale image shall be decomposed by Mallat algorithm with wavelet, and a grayscale pyramid $L_k(s)$ shall be established. According to the high-frequency component, the decomposed grayscale image shall be built into detail pyramid $O_k(s, d)$, the size shall be represented by $s \in [0, N]$, and the three detail directions of the video image shall be represented by $d \in [\mathrm{HL}, \mathrm{LH}, \mathrm{HH}]$. The establishment of the original video grayscale image shall be completed by the pyramid $L_k(s)$ and detail pyramid $O_k(s, d)$ calculated in the above process, and the multiple features of two detail features can be obtained simultaneously [34].

In a video image conveyed visually, the local contrast of a certain image at different positions within a certain range is determined by the difference between scales, so as to improve the contrast effect of the animation and video images. The formula is as follows:

$$\mathrm{DOG}(x, y, \sigma_c, \sigma_s) = \frac{1}{2\pi\sigma_c^2}\exp\left[-\frac{r^2}{2\sigma_c^2}\right] - \frac{1}{2\pi\sigma_s^2}\exp\left[-\frac{r^2}{2\sigma_s^2}\right], \tag{15}$$

where r represents the feature contrast measure, σ_c represents the range of feature area, and σ_s represents the range of film, television, and animation video image suppression area. $F(c)$ represents the coarse-scale film and television animation video image, which can suppress the features other than the target area [35]. $F(s)$ represents the fine-scale film and television animation video image, which can describe the detail features of the target area, and then the scale difference between $F(c)$ and $F(s)$ is shown in the following equation:

$$F(c, s) = F(c)\Theta F(s). \tag{16}$$

$F(s)$ adjusts the difference value to the film and television animation video image area with the same scale as $F(c)$, which is represented by $L_k(c)$ and $L_k(s)$, and subtracts each point to obtain the corresponding absolute values $O_k(c, d)$ and $O_k(s, d)$, so as to obtain the gray-level features and detail features of the film and television animation video image. The formula is as follows:

$$\begin{cases} \widehat{L}_k(c, s) = L_k(c)\Theta L_k(s), \\ \widehat{O}_k(c, s, d) = O_k(c, d)\Theta O_k(s, d). \end{cases} \tag{17}$$

$G_k(i, j)$ represents the gray value of the (i, j)-th pixel in the film and television animation video image, $\widehat{d}x$ and $\widehat{d}y$ represent the evaluation estimator between the film and television animation video images, and the image feature display diagram is calculated according to the translation estimator, as shown in the following equation:

$$\widehat{M}_k(i, j) = \left|G_k(i, j) - G_{k-m}(i - \widehat{d}y, j - \widehat{d}x)\right|. \tag{18}$$

In the formula, $G_{k-m}(\cdot)$ represents the translation difference, the normalized function (No) is used to unify the scale of the image, and then a group of weighting factors α, β, and γ are determined to complete the weighted fusion of the video image through the above method. The formula is shown in the following equation:

$$S_k = \alpha \cdot N\left(\widehat{L}_k\right) \cdot \beta \cdot N\left(\widehat{O}_k\right) \cdot \gamma \cdot N\left(\widehat{M}_k\right), \tag{19}$$

where $N(\widehat{L}_k)$ represents grayscale, $N(\widehat{O}_k)$ represents detail, and $N(\widehat{M}_k)$ represents the normalization function of motion feature image and represents the feature value of film and television animation video image frame extracted at last.

2.5.2. Dynamic Identification. Dynamic recognition methods mainly include global and local recognition. Because the dimension of video image is high, it is difficult to identify the whole video image by global method, and there are noise interference and information occlusion in the image after dynamic recognition. Therefore, in order to improve the efficiency of dynamic recognition, local recognition method is adopted [36]. This method takes *K-L* (Karhunen–Loeve) transformation as the main core and uses the similarity between evaluation features to realize dynamic recognition.

K-L transform is based on the statistical characteristics of the transformation, through the incoming vector orthogonal transform and the outgoing vector to decouple the relevance of the data, which is the best way in the sense of variance.

Suppose, f_1 represents the $N \times M$-dimensional matrix, Γ represents the sample statistic $\Gamma = \{\Gamma_1, \Gamma_2, \Gamma_3 \ldots \Gamma_L\}$, $NM \times 1$ represents the dimension vector, and the average value μ of Γ is calculated and obtained from the set of L sample values, as shown in the following equation:

$$\mu = E(\Gamma). \tag{20}$$

Its variance matrix $C(\Gamma)$ is shown in the following equation:

$$C(\Gamma) = E\left\{(\Gamma - \mu)(\Gamma - \mu)^T\right\}. \tag{21}$$

λ_i represents the eigenvalue of the variance matrix, and Φ_i represents the eigenvector of the variance matrix. $L - 1$ eigenvectors and their corresponding eigenvalues are obtained, as shown in (19):

$$C_\Gamma \Phi_i = \lambda_i \Phi_i, 0 \le i \le L - 1. \tag{22}$$

The order of feature vector Φ_i is based on the order of the size of feature values. The first m eigenvectors are determined and then normalized to form a change matrix $\mathfrak{R}$. Then, the transformation matrix is used to transform vector Γ_i into vector $\mathfrak{I}$ in row m through (20):

$$\mathfrak{I} = \Gamma_i \times \mathfrak{R}. \tag{23}$$

Thus, achieve the effect of dimensionality reduction, and effectively complete the film and television animation video image enhancement design.

3. Experimental Results and Analysis

The simulation environment is Pentium m1.60 Ghz CPU and 760 m RAM. The dual-core Intel second-generation core processor OptiPlex 3010 is used as the main frequency, the system is win10 flagship 32 bits, and the simulation software is MATLAB 2020 B. In order to prove the effect of the method in this paper, it is compared with current advanced visual communication methods of animation images (the plane visual communication design method based on graphic beautification technology and the visual communication design method of animated character image in virtual reality environment).

3.1. Comparative Analysis of Visual Communication Image Reconstruction Effects under Different Methods. To make the experimental results more intuitive, two animated video images of 512×512 pixels shall be selected as reference animated video images, as shown in Figures 3 and 4(a), and the simulation experiment shall be carried out with the relevant simulation experimental tools. If we use Gaussian fuzzy model, set the Gaussian filter of 3×3 region invariant, set its sampling factor to 4, add Gaussian noise to all low-resolution images, and satisfy the signal-to-noise ratio of 30 dB. Figures 3(a) and 4(a) are the reference film and television animation video images, Figures 3(b) and 4(b) are the reconstructions of film and television animation video images by using the visual communication design method of virtual reality environment animation character images, Figures 3(c) and 4(c) are the reconstructions of film and television animation video images by using the graphic beautification technology-based graphic visual communication design method, and Figures 3(d) and 4(d) are the reconstructions of this method.

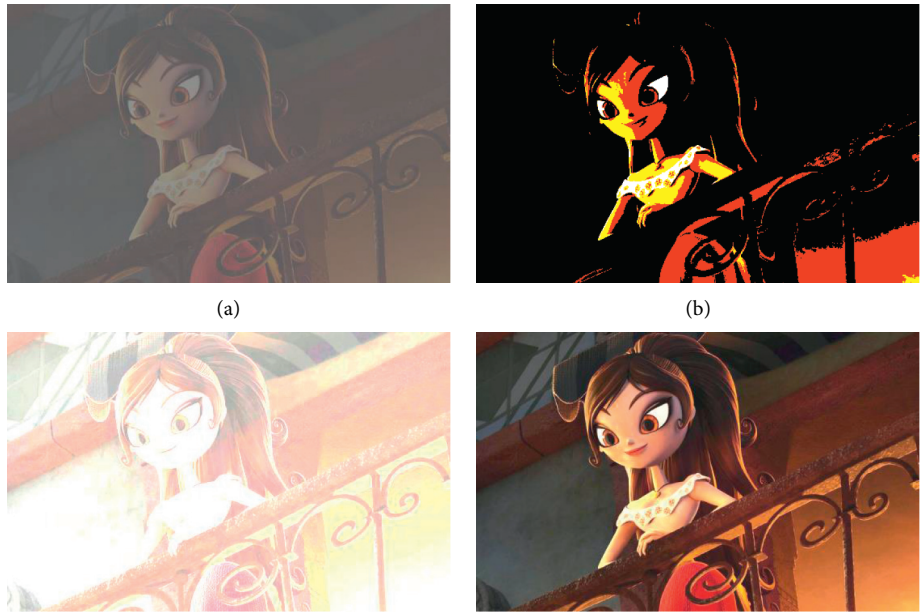
According to Figures 3 and 4, it can be seen that the visual communication design method of VR environment animated character image cannot avoid blurring edges and the reconstruction effect is not good. The method of plane visual communication design based on graphic beautification technology has a large error when matching some pixels, which leads to the lack of rich and clear details of the final reconstructed animation video image, and the reconstructed effect is inferior to the visual communication design method of virtual reality animation image. Compared with the other two methods, the proposed method has obvious advantages in the reconstruction of high-resolution animation video images.

3.2. Comparative Analysis of Recognition Rate. Compared with the visual communication design method based on VR environment animation and graphic landscaping technology, the effectiveness of visual communication video image dynamic recognition is tested. First, according to the actual situation, 50 kinds of license plate Chinese characters are segmented into two kinds of film and TV animation video images, which are divided into 10 sets of fuzzy similar Chinese characters according to the special structure features of Chinese characters, and then recognized. There are 270 fuzzy film and TV animation video images, and the recognition rates of the two methods are shown in Table 1.

As can be seen from Table 1, the dynamic recognition rate of the proposed method is 91%, which is higher than other methods. The dynamic recognition rate of the graphic beautification technology-based graphic visual communication design method is 62.2%, and the dynamic recognition rate of the virtual reality environment animation figure visual communication design method is 67.5% for the Chinese character images with the same resolution and blurring degree of film and television animation video images. This is mainly due to the fact that the graphic visual communication design method based on graphic beautification technology and the virtual reality environment animation figure visual communication design method do not take into account the influence of illumination color on film and television animation video images, so that the feature dimension is reduced. At the same time, in order to prevent overlapping areas, 2×1 block is used as a block at the time of establishing a block, which also reduces the dimension, but when constructing a histogram, the gradient of each cell is divided into 9 parts. According to the experimental analysis of Dalal, if the area is reduced at this time, the feature dimension can be reduced, but the description ability is also reduced, and the computational complexity is increased. The advantage of the proposed method in recognition rate is that it arranges the eigenvalues in an orderly way and determines the eigenvectors of the eigenvalues as Chinese characters. This method not only reduces the feature dimension, but also retains the key information of the animation video image. In the process of feature extraction, it highlights the main details of the animation video image, deepens the description of the feature, and effectively improves the recognition rate.

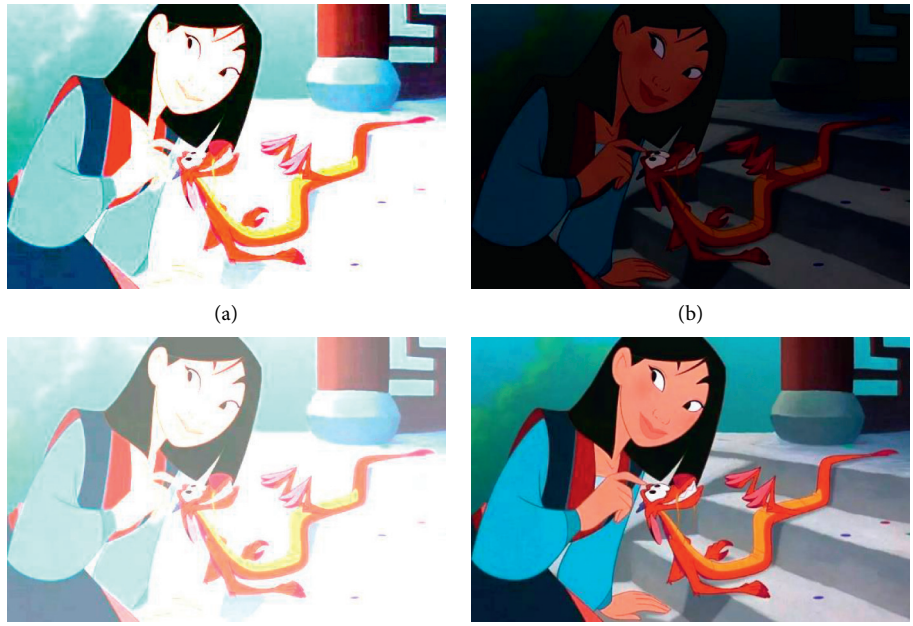
3.3. Space-Time Performance Test. Taking the visual communication design method of animated character images in virtual reality environment proposed in reference [3] and the plane visual communication design method based on graphic beautification technology proposed in reference [4] as the comparison method, and taking the grassland texture sample synthesis as an example, the temporal and spatial performance of this method and the comparison method are compared. The results are shown in Tables 2 and 3.

Analysis of Table 2 shows that the length of time of the three methods used in 3D visual communication is different. The length of time of the suggested method is below 3.8 s, the length of time of the visual communication design method of animated character image in virtual reality environment is up to 14.93 s, and the length of time of the graphic visual communication design method based on graphic beautification technology is up to 20.06 s. The length of time of the suggested method is smaller than that in VR environment. Compared with the two comparison methods, this method has obvious advantages in time. This is because this 3D visual communication animation video images through stitching, projection mapping, and animation video image frame texture synthesis projection, resulting in packaging. There is



(a) (b) (c) (d)

Figure 3: Simulation results. (a) Video image sample of film and television animation. (b) Visual communication design method of animated character image in virtual reality environment. (c) Graphic visual communication design method based on graphic beautification technology. (d) Paper method.



(a) (b) (c) (d)

Figure 4: Simulation results. (a) Video image sample of film and television animation. (b) Visual communication design method of animated character image in virtual reality environment. (c) Graphic visual communication design method based on graphic beautification technology. (d) Paper method.

no need to preprocess the texture sample to improve the efficiency of texture synthesis.

Analysis of Table 3 shows that the size of the animation video image after the three methods is quite different. The size of the synthesized image that is determined by this method is below 3.8 MB, the image size under visual communication design method of animated character image in virtual reality environment is up to 14.93 MB, and the image size under image synthesis based on graphic beautification technology is up to 20.06 MB. The size of the

Table 1: Comparison and analysis of recognition rate.

Method	Film animation video image resolution	Characteristic dimension	Recognition rate (%)	Identification time (s)
Visual communication design method of animated character image in virtual reality environment	32×16	70	62.2	1.8
Graphic visual communication design method based on graphic beautification technology	32×16	62	67.5	1.7
Proposed method	32×16	95	91.0	0.9

Table 2: Time performance simulation results.

Number of triangular meshes	Synthesis time of this method (s)	Visual communication design method of animated character image in virtual reality environment (s)	Graphic visual communication design method based on graphic beautification technology (s)
512	2.1	16.9	12.8
1024	2.7	22.4	16.7
1536	2.0	23.2	15.0
2048	2.4	34.6	76.2
2560	2.6	51.1	26.6
3072	3.5	87.3	907
3584	3.8	92.2	88.3
4096	3.2	167.2	135.4

Table 3: Space performance simulation results.

Number of triangular meshes	The size of the synthesized image is determined by this method (MB)	Image size under visual communication design method of animated character image in virtual reality environment (MB)	Image size after image synthesis based on graphic beautification technology (MB)
512	0.37	1.36	1.78
1024	1.22	2.58	2.81
1536	1.19	4.27	4.89
2048	1.92	5.17	5.74
2560	0.79	7.33	7.63
3072	1.27	9.17	10.08
3584	1.17	11.18	14.27
4096	2.04	14.93	20.06

synthesized image that is determined by the proposed method is smaller than other methods. This is because the present method constructs the scale factor fields for 3-dimensional transformations. Based on this, the identified and extracted video image frame feature scale is smaller and occupies less space.

3.4. Real Example Experiment of Video Image Reconstruction. In order to reflect the performance of this method in practice, the video image reconstruction performance is tested. The 500 images from the Lion King animated video were selected as subjects and iterated 200 times to test the peak signal-to-noise ratio of the HD reconstructed image, and the results are shown in Figure 5:

As shown in Figure 5, the image peak signal-to-noise ratio of the present method is 31.2dB–40.9 dB, which is higher than the original image, proving that this method meets the practical needs of high-resolution image reconstruction.

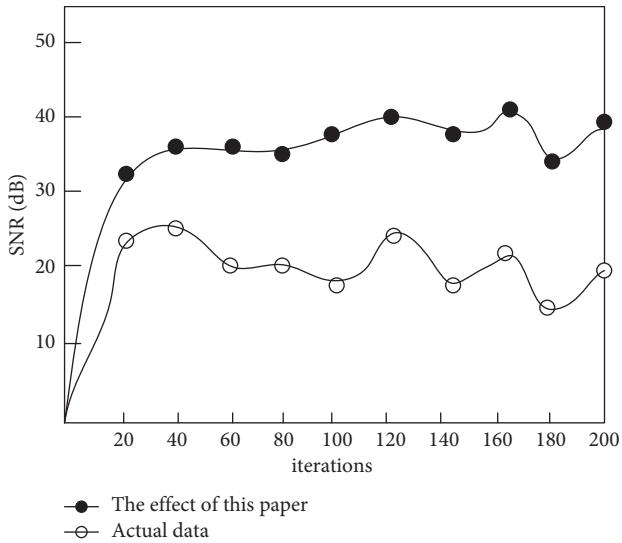


Figure 5: Peak SNR of the HD reconstructed images.

4. Conclusion

A new animation design method based on 3D visual communication technology is designed to solve the problem of poor visual effect in the existing image visual communication technology. Build a video processing process under 3D visual communication technology; based on the visual distance from a 2D images to a 3D scene, compute the scaling factor for the 3D scene transformation; building a deep learning-based PCANet model, images were used to filter for processing by using the model and to extract the characteristics of the film and television animation video images; deep features of the image samples were acquired using the PCAN deep network; create a low-resolution feature dictionary for high-resolution images; using a high-resolution video image reconstruction method based on a sparse rule model, realize the reconstruction of animated video images; nonlocal similarity empirical constraints are also introduced to optimize the HR images; analyze gray-scale and detail features based on wavelet transform; extract the features of the video images; with the K-L (Karhunen–Loeve) transformation as the core, the similarity between the picture features is used to dynamically identify the film and television animation video images, and to complete the high-precision film and television animation design. The experimental results show that the proposed method can improve the visual communication effect of animation and video images, with small time cost, small space capacity, and good application performance.

Data Availability

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding this work.

Acknowledgments

This work was supported by Provincial Quality Project of Higher Education Institutions in Anhui Province in 2019 (Boutique Offline Open Course Projects), "Introduction to Animation" (2019kfkc309), Feng Shan (host).

References

[1] J. Park and B.-U. Lee, "Color image enhancement with high saturation using piecewise linear gamut mapping," *Journal of Visual Communication and Image Representation*, vol. 67, p. 102759, 2020.

[2] M. Yu, "Film and television culture dissemination based on ZYNQ embedded digital image processing," *Microprocessors and Microsystems*, vol. 82, no. 9, p. 103921, 2021.

[3] X. Q. Zhu, "Innovative design of graphic image visual communication system for animated character in virtual reality environment," *Modern Electronics Technique*, vol. 43, no. 10, p. 3, 2020.

[4] M. M. Liu, "Optimization of plane visual communication effect simulation based on graphic beautification technology," *Computer Simulation*, vol. 36, no. 09, pp. 426–429, 2019.

[5] F. Sun, "Automatic programming algorithm of graphic language in visual communication design," *Science Technology and Engineering*, vol. 19, no. 02, pp. 149–154, 2019.

[6] L. U. Mei, Z. D. Wang, and P. Z. Zhang, "A multi-object tracking algorithm based on the combination of neighborhood frame matching and kalman filtering," *Science Technology and Engineering*, vol. 19, no. 15, pp. 204–210, 2019.

[7] C. Rodrigues, Z. Peixoto, and F. Ferreira, "Ultrasound image denoising using wavelet thresholding methods in association with the bilateral filter," *IEEE Latin America Transactions*, vol. 17, no. 11, pp. 1800–1807, 2019.

[8] W. Sun, D. Yan, J. Huang, and C. Sun, "Small-scale moving target detection in aerial image by deep inverse reinforcement learning," *Soft Computing*, vol. 24, no. 11, pp. 1–12, 2019.

[9] I. J. Kadhim, P. Premaratne, and P. J. Vial, "Improved image steganography based on super-pixel and coefficient-plane-selection," *Signal Processing*, vol. 171, pp. 107481–107489, 2020.

[10] M. Meem, A. Majumder, and R. Menon, "Multi-plane, multi-band image projection via broadband diffractive optics," *Applied Optics*, vol. 59, no. 1, p. 38, 2020.

[11] H. Zhang and J. Han, "Mathematical models for information classification and recognition of multi-target optical remote sensing images," *Open Physics*, vol. 18, no. 1, pp. 951–960, 2020.

[12] A. Yuan, X. Li, and X. Lu, "3G structure for image caption generation," *Neurocomputing*, vol. 330, pp. 17–28, 2019.

[13] A. Mik and J. Novák, "Three-component zoom lens with fixed position of optical center," *Applied Optics*, vol. 58, no. 14, p. 3844, 2019.

[14] G. Kertész, S. Szénási, and Z. Vámossy, "Comparative analysis of image projection-based descriptors in Siamese neural networks," *Advances in Engineering Software*, vol. 154, no. 6, p. 102963, 2021.

[15] Y. Feng, H. Wang, Z. G. Qu et al., "Trajectory simulation of hypersonic gliding vehicle based on reference ellipsoid," *Computer Simulation*, vol. 37, no. 10, pp. 18–23, 2020.

[16] K. G. Dhal, A. Das, S. Ray, and J. Gálvez, "Randomly Attracted Rough Firefly Algorithm for histogram based fuzzy image clustering," *Knowledge-Based Systems*, vol. 216, no. 1, p. 106814, 2021.

[17] W. Zhou, Y. Zhou, W. Qiu, T. Luo, and Z. Zhai, "Perceived quality measurement of stereoscopic 3D images based on sparse representation and binocular combination," *Digital Signal Processing*, vol. 93, pp. 128–137, 2019.

[18] J. Yang, W. Zhao, Y. Han et al., "The aircraft tracking based on fully conventional network and kalman filter," *IET Image Processing*, vol. 13, no. 8, pp. 1259–1265, 2019.

[19] N. R. Zhou, X. X. Liu, Y. L. Chen, and Ni-S. Du, "Quantum K-Nearest-Neighbor image classification algorithm based on K-L transform," *International Journal of Theoretical Physics*, vol. 60, pp. 1209–1224, 2021.

[20] S. Gu, L. Wang, W. Hao, Y. Du, J. Wang, and W. Zhang, "Online video object segmentation via low-rank sparse representation," *IET Computer Vision*, vol. 13, no. 5, pp. 469–479, 2019.

[21] A. Rahimi, P. Moallem, K. Shahtalebi, and M. Momeni, "Using Kalman filter in the frequency domain for multi-frame scalable super resolution," *Signal Processing*, vol. 155, pp. 108–129, 2019.

[22] Y. Pu, X. Zhao, G. Chi et al., "Design and implementation of a parallel geographically weighted k -nearest neighbor classifier," *Computers & Geosciences*, vol. 127, pp. 111–122, 2019.

[23] J. Zhao, "Art visual image transmission method based on cartesian genetic programming," *Scientific Programming*, vol. 30, no. 10, pp. 1–10, 2021.
[24] B. Pandey, S. Thakur, H. Joshi, A. Pradhanga, Y. Akiyama, and J. Peethambaran, "Towards video based collective MotionAnalysis through shape tracking andMatching," *Electronics Letters*, vol. 56, no. 17, pp. 2157–2168, 2020.
[25] A. Mustafa, H. Kim, and A. Hilton, "MSFD: multi-scale segmentation-based feature detection for wide-baseline scene reconstruction," *IEEE Transactions on Image Processing*, vol. 28, no. 3, pp. 1118–1132, 2019.
[26] M. K. Sadabad, A. J. Akbarfam, and B. Shiri, "A numerical study of eigenvalues and eigenfunctions of fractional sturm-liouville problems via laplace transform," *Indian Journal of Pure and Applied Mathematics*, vol. 51, no. 3, pp. 857–868, 2020.
[27] H. Ullah, M. A. Khan, and J. Peari, "New bounds for soft margin estimator via concavity of Gaussian weighting function," *Advances in Difference Equations*, vol. 2020, no. 1, pp. 1–10, 2020.
[28] S. Kumar and R. Mathusoothana, "Robust multi-view videos face recognition based on particle filter with immune genetic algorithm," *IET Image Processing*, vol. 13, no. 4, pp. 600–606, 2019.
[29] M. Lee, T. Budavari, I. Sullivan, and A. Connolly, "Sub-band image reconstruction using differential chromatic refraction," *The Astronomical Journal*, vol. 157, no. 5, p. 182, 2019.
[30] Y. Guo, "Moving target localization in sports image sequence based on optimized particle filter hybrid tracking algorithm," *Complexity*, vol. 2021, no. 7, pp. 1–11, 2021.
[31] R. Ma, H. Hu, S. Xing, and Z. Li, "Efficient and fast real-world noisy image denoising by combining pyramid neural network and two-pathway unscented kalman filter," *IEEE Transactions on Image Processing*, vol. 29, no. 3, pp. 3927–3940, 2020.
[32] N. Ghosh, R. Paul, S. Maity, K. Maity, and S. Saha, "Fault Matters: sensor data fusion for detection of faults using Dempster-Shafer theory of evidence in IoT-based applications," *Expert Systems with Applications*, vol. 162, no. 4, p. 113887, 2020.
[33] Z. Ji, Y. Zhang, Y. Pang, X. Li, and J. Pan, "Multi-video summarization with query-dependent weighted archetypal analysis," *Neurocomputing*, vol. 332, pp. 406–416, 2019.
[34] X. Hua, X. Q. Wang, Z. Y. Ma et al., "Video multi-target detection technology based on recursive neural network," *Application Research of Computers*, vol. 37, no. 02, pp. 301–306, 2020.
[35] A. Nicolson and K. K. Paliwal, "Spectral distortion level resulting in a just-noticeable difference between an a priori signal-to-noise ratio estimate and its instantaneous case," *Journal of the Acoustical Society of America*, vol. 148, no. 4, pp. 1879–1889, 2020.
[36] G. Lei, R. Yao, Y. Zhao, and Y. Zheng, "Detection and modeling of unstructured roads in forest areas based on visual-2D lidar data fusion," *Forests*, vol. 12, no. 7, p. 820, 2021.

Hindawi
Scientific Programming
Volume 2021, Article ID 5089236, 14 pages
https://doi.org/10.1155/2021/5089236

Research Article

Research on the Construction of Intelligent Community Emergency Service Platform Based on Convolutional Neural Network

Yu Chen and Zhong Tang

School of Humanities and Social Science, Guangxi Medical University, Guangxi Nanning 530021, China

Correspondence should be addressed to Zhong Tang; tangzhong@stu.gxmu.edu.cn

Received 10 November 2021; Accepted 3 December 2021; Published 28 December 2021

Academic Editor: Baiyuan Ding

Aiming at the shortcomings of the existing community emergency service platform, such as single function, poor scalability, and strong subjectivity, an intelligent community emergency service platform based on convolutional neural network was constructed. Firstly, the requirements analysis of the emergency service platform was carried out, and the functional demand of the emergency service platform was analyzed from the aspects of community environment, safety, infrastructure, health management, emergency response, and so on. Secondly, through logistics network, big data, cloud computing, artificial intelligence, and all kinds of applications, the intelligent community emergency service platform was designed. Finally, a semantic matching emergency question answering system based on convolutional neural network was developed to provide key technical support for the emergency preparation stage of intelligent community. The results show that the intelligent community emergency service platform plays an important role in preventing community emergency events and taking active and effective measures to ensure the health and safety of community residents.

1. Introduction

In China's 2021 Intelligent Community Construction and Operation Guide, intelligent community is defined as the basic unit of intelligent city, which is guided by the intelligence, green, and humanism of the community, integrates people, land, and other elements in the community scene, and promotes the communication and mutual assistance of community residents with the public interests of community residents. An innovative model of community management and service integrates various resources such as public management, public service, and commercial service, provides community management and service applications for the government, property management, residents, and enterprises, and improves the scientific, intelligent, and refined level of community management and service. Community is the basic unit of national governance system, community emergency service platform is the key to break through the "last mile" of intelligent community, and community emergency management ability is also an important issue of grass-roots governance [1]. When emergency medical treatment, extreme weather coping, and extreme emotional and psychological comfort prevent waterlogging control, mosquito control, and epidemic prevention, such as sudden event, the community often becomes the subject of prevention and treatment of sudden events, which is to prevent the forefront of sudden events and crisis. Improving the emergency management ability and responsibility of the community, building intelligence community emergency management system, we will enhance people's awareness of disaster prevention and mitigation [2]. Improving people's capacity for emergency response and disaster reduction plays an important role in safeguarding the safety of people's lives and property.

Community emergency tube is to combat the threat of unexpected accidents that have community organization personnel activated within their respective areas, office, family, and society, and the measures and management

mechanism, through extensive, positive, and orderly participation in emergency management work, smoothly and orderly solves emergencies, to ensure social stability and people's happiness and property safety [3]. Community emergency management, as the cornerstone of crisis management system, has a profound impact on the grass-roots management of society, public services, and the maintenance of residents' public interests. Throughout the COVID-19 outbreak, communities, as grass-roots units, have become the forefront of epidemic control and development, which has played a huge role in the practice of fighting the epidemic. However, at the same time, there is no participation in emergency management, no material guarantee, and the lack of correct concept guidance. Emergency management team quality is uneven; ability is not fine; community residents' awareness of emergency is not strong, and the ability to help each other is weak. How to make up the short board of community emergency management, use advanced management ideas and modern information technology means, develop a simple, practical, and efficient community emergency service system, and enhance the community to deal with the emergency handling ability of unexpected events, for improving the level and effect of community management personnel emergency service [4]? It is of scientific and academic value and social and economic significance to reduce loss and harm as much as possible.

Therefore, by analyzing the deficiencies of the existing community emergency service platform and the needs of the intelligent community emergency service platform, this paper builds an intelligent community emergency service platform based on convolutional neural network (CNN), improves the social network of community emergency management, builds a community information resource platform, and establishes an internal information exchange system, to form the technical support and convenient service for cross-department networked collaborative problem solving, continuously improve the community's scientific, networked and all-round service capacity and emergency response effect, realize the technical level of emergency response measures from the traditional passive defense to active early warning, and realize the improvement of community service level through information means.

2. Platform Requirement Analysis

Requirement analysis is an important link in software engineering. Requirement analysis of each application subsystem of intelligent community emergency service platform lays a foundation for subsequent system design and software implementation [5].

2.1. Functional Requirements. The functional requirements description focuses on what the developer needs to achieve. The functions of the intelligent community emergency service platform mainly include some functions.

(a) Intelligent emergency management: using sensors, multimedia, and broadband network information tools such as access to a large number of community management and services, the people's livelihood service, community environment, community infrastructure, community economy, such as information, based on cloud computing, big data, such as the Internet of things a new generation of information technology, data analysis, data mining, data security, and information optimize the community security service and enhance disaster prevention and mitigation capacity.

(b) Scientific allocation of resources: after an emergency occurs, scientific management of community emergency resources such as community managers, community plans, shelter materials, social workers and volunteers, and safe places can realize timely allocation of resources, so that the three levels of district, street, and community can be interconnected and contribute resources, so as to meet the needs of disaster emergency response. We will improve our ability to ensure emergency resources.

(c) Integration of service platforms: building owners and tenants oriented community, mobile service platform and call center are a body comprehensive service platform, through the centralized, flattening intensification of community management mode, implementation, and community service centers, community health care center, community home endowment center department of resource sharing, data exchange, information flow, and data submitted. Promote safe and intelligent communities.

According to the functional requirements of system modules in detail, including emergency safety planning and management, emergency resource management, emergency capability assessment, and the safety management of the floating population four function modules, each module has a corresponding subroutine below, and the system establishes a database to store all kinds of information related to support the operation of each function module.

(a) Emergency safety planning and management: community security planning needs to consider the management objects of vulnerability and disaster in the region where the community is located. As the regional characteristics of each community are different, some common management objects of the community are selected as representatives, managing emergency safety planning. It mainly includes the management of major hazard sources, key personnel, and important infrastructure.

(b) Emergency resource library management: the plan management system digitizes the plan elements, forms the emergency plan information into data fields, stores the data fields, and establishes the plan database. It mainly includes the management of community plans, community workers and volunteers, safe places, public activity areas, and other rescue resources.

(c) Emergency response capacity assessment: the assessment of emergency response capacity is based on

preset performance objectives of emergency management. By using scientific and reasonable evaluation system and methods, the community's emergency response ability is monitored from all directions and from many angles under the principle of realizing the highest benefit at the lowest cost, so as to make a scientific and reasonable comprehensive evaluation. It mainly includes selection of assessment methods, establishment of assessment models, assessment of capability indicators, publicity of assessment results, and other related processes.

(d) Safe management of floating population: the safe management of floating population is an important part of intelligent community emergency service platform, mainly including buildings, houses, household register, residents, and other management visualization, as well as the area of the registered population register management and other related content.

2.2. Nonfunctional Requirements Analysis. In the process of system requirement analysis, nonfunctional requirement is also a very important part, which describes the performance or way of software to complete the functions required by the service platform, so that the software platform has good promotion ability and adaptability [6]. Performance requirements, reliability, response speed, external interface requirements, design constraints, and data storage requirements fall into the category of nonfunctional requirements [7]. In order to ensure that the intelligence community emergency service platform can effectively dispose emergency disaster events and ensure achieving the goal of construction of the platform, the system of the nonfunctional requirements mainly includes the following: the response time (the system can be in the process of data management, daily office, and system management system response time less than 3 seconds), use requirement (verify system parameter configuration and management authorization mode by entering system operator user name and password), system availability (the system should support 7∗24 hour service operation, with an average trouble-free time of 23.90 hours. The average maintenance time of maintainability is less than 2 hours), failure recovery (the system is deployed in dual-system hot backup mode, which is seamless in principle. When one system is faulty, the other system automatically takes over services without affecting service running), database connection (automatic database reconnection mechanism: in the case of network interruption, the application system can proactively discover and automatically reconnect the database), manageability (the system is designed according to the requirements of expansibility, and interfaces with the external system are issued to facilitate the implementation of docking with other systems), and maintainability and safety requirements (this section describes the maintenance window, log requirements, data clearing policies, and log file clearing policies).

3. Systematic Design

Intelligent community emergency service platform is a comprehensive management application software integrating front-end information, management information, data retrieval, and analysis [8]. It can not only display the emergency data collected and analyzed by front-end collection equipment, but also serve as a link between management personnel and community residents. In this paper, intelligent community emergency management is divided into three stages: emergency preparation, response handling, and recovery and adjustment; five core levels: basic network layer, system layer, application layer, product operation layer, and user layer; and three supporting systems: system security system, infrastructure guarantee, and comprehensive operation service. Figure 1 shows the flowchart of intelligent community emergency management based on big data, in which the emergency preparation stage is the key. Therefore, with the help of convolution neural network, the efficient collection, sorting, passed throughout the community emergency management information and scattered the pieces, sorting, analysis of information, for the community monitoring, early warning and emergency resources management to provide data support and scientific basis and achieve precise, agile, efficient, and comprehensive coverage of community emergency services.

The user roles of the intelligent community emergency service platform mainly include community worker and administrator [9], where community workers are the main users of the platform, completing the responsibilities of planning management, resource database management, emergency response capacity platform, and migrant management. An administrator is a specific type of user who has the maximum management rights for the platform. The system example is shown in Figure 2.

As an important view of system architecture design, development view is mainly used to describe the organization and management of software modules [10]. As shown in Figure 3, this paper divides each module of the platform from three levels: presentation layer, business logic layer, and foundation construction layer. The presentation layer consists of Windows UI user interface module. The middle layer is the business application layer, mainly including user management, information management, ability assessment, and other system service modules, through the form of service access interaction with the presentation layer; the bottom layer is database layer, which is the data source of platform operation, including spatial database, plan database, planning database, and population database. The business application layer invokes data from the data layer to respond to requests from business applications.

The scene view can be used to describe the user's activity interaction process of the same operation or the activity interaction process of different scenarios, which is the dynamic description of user activity components [11]. Taking the preparation of emergency plan as an example, the interactive process of each activity is described, as shown in Figure 4. Users enter the information page, fill in the plan preparation information, save and submit to the review stage, and return to the system to process the information; users review the basic information of the plan, fill in the review

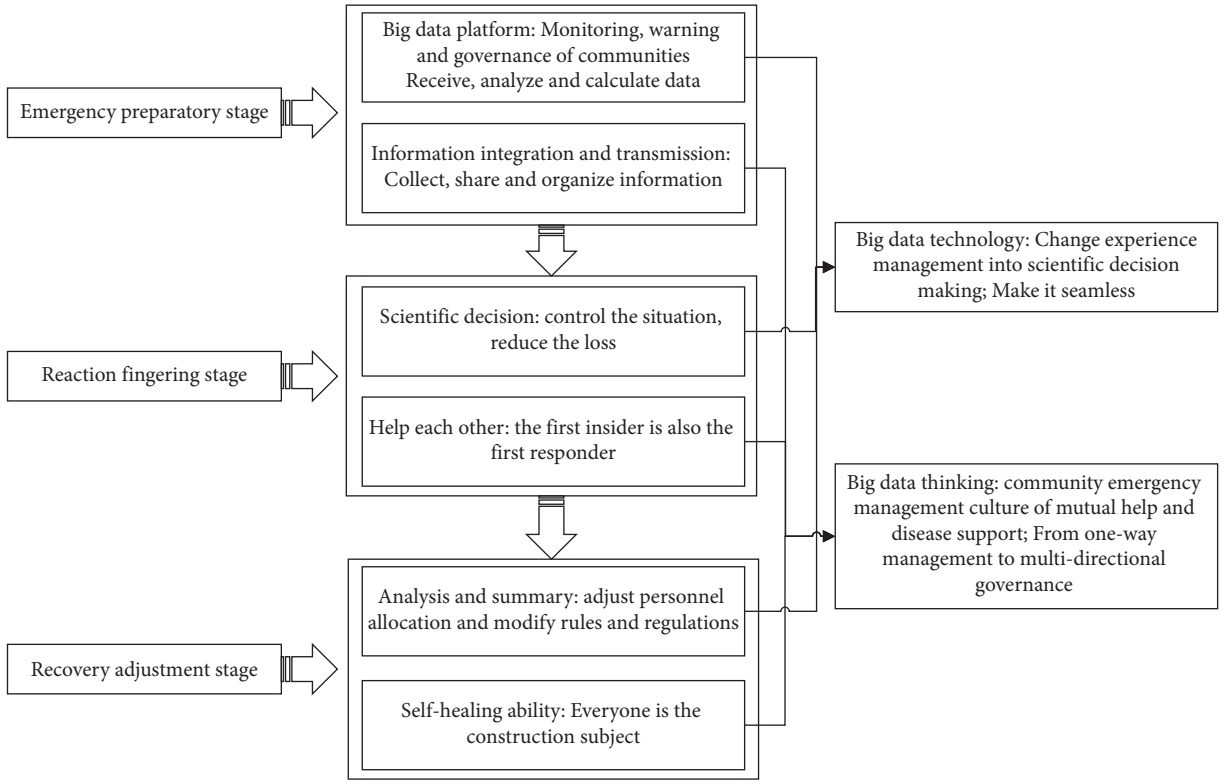


Figure 1: Intelligent community emergency management process.

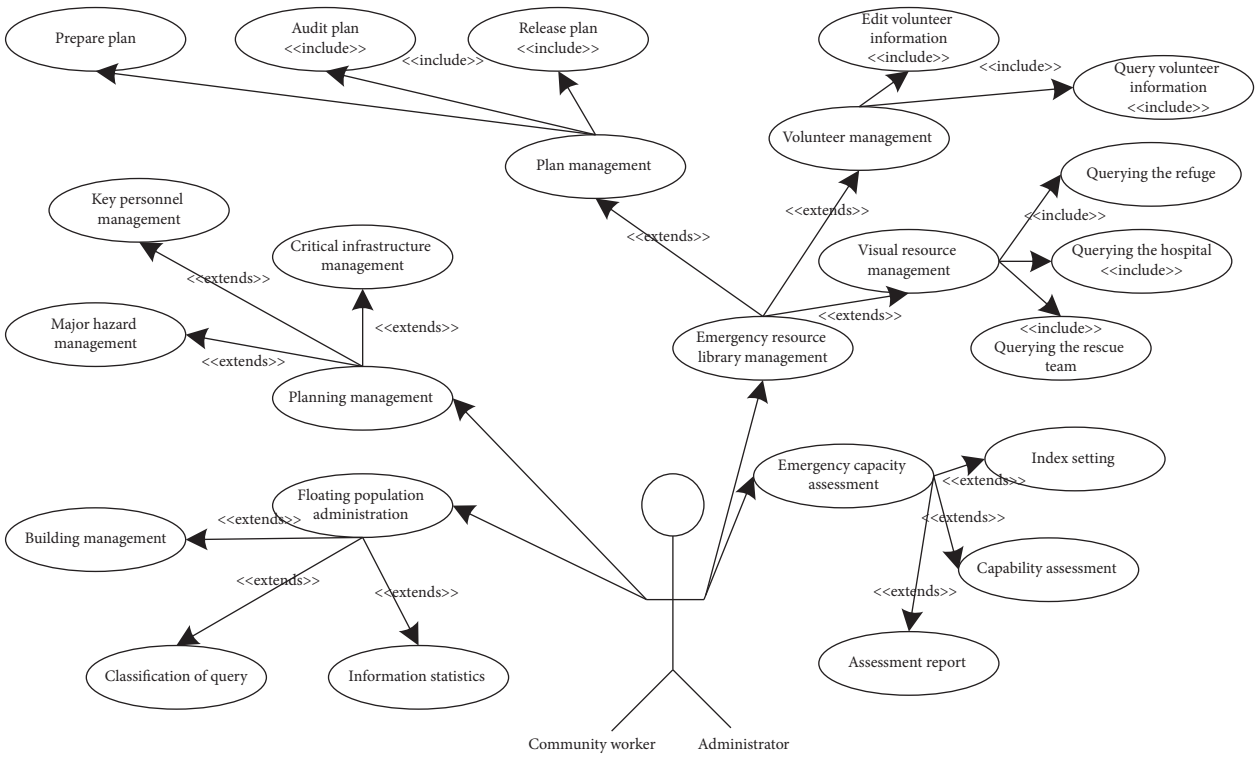


Figure 2: Intelligent community emergency service platform use case diagram.

opinion, and return the processing message, so as to realize the in-depth analysis and comprehensive application of massive data by the intelligent community.

Through the research and development of community emergency service platform, a perfect community emergency coordination network system has been built, as shown

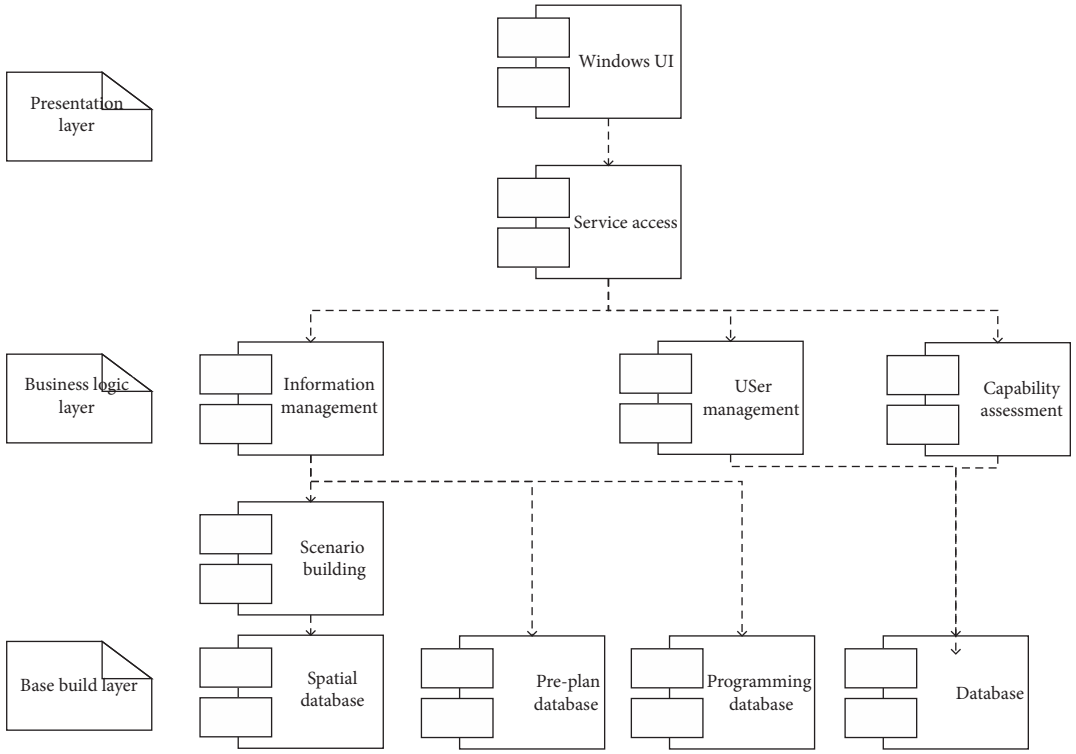


Figure 3: Community emergency service platform development diagram.

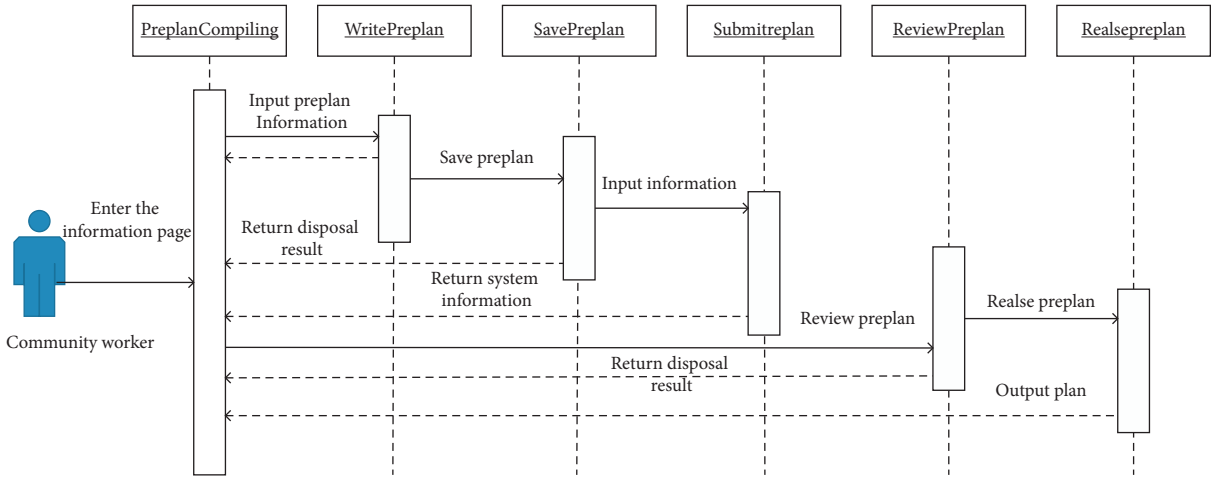


Figure 4: Community emergency service platform scenario diagram.

in Figure 5. A wide coordination governance platform has been built, a real-time sharing interactive relationship has been formed, and the centralized and unified leadership model has been followed [12].

4. Design of Emergency Q&A System Based on CNN

4.1. Basic Theory of CNN. A typical CNN model is composed of multiple convolutional layers and pooling layers, followed by one or more full connection layers. There are four core ideas behind it: local connection, weight sharing, pooling, and multilevel [13]. The overall structure of CNN is shown in Figure 5. Each convolution operation will be followed by a pooling operation until the size of the feature graph is reduced to 1, followed by multiple fully connected layers.

4.1.1. Input Layer. Figure 6 shows the shape of the input layer. Consider the input of a channel number of data F_0, the standard image processing $F_0 = 3$, exactly corresponding to the computer red, green, and blue three color channels. The size of each graph of a single channel is defined as $N_0 \times T_0$ (width × height), and the input data can be written as $N_0 \times T_0$, where i represents the index of each data, $i \in [0, I-1]$. Set the input after processing as $\widetilde{X}$, and the specific calculation process can be expressed as

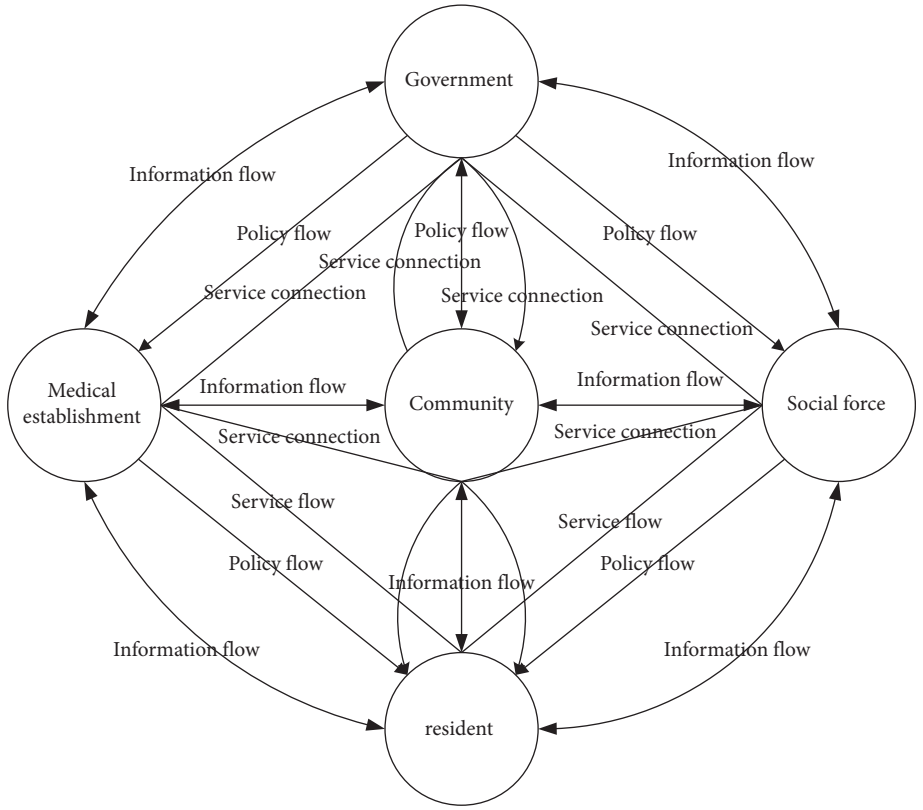


Figure 5: Collaborative network of community emergency service platform.

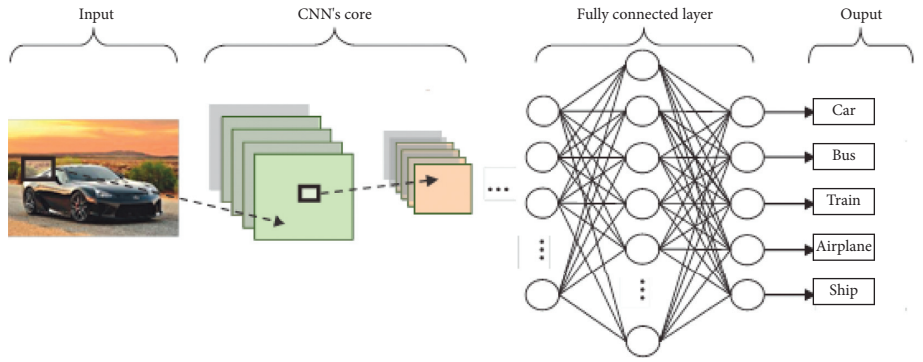


Figure 6: Typical structure of CNN.

$$\widetilde{X}_{f,j,k}^{(i)} = \frac{X_{f,j,k}^{(i)} - \widehat{X}_{f,j,k}^{(i)}}{\sigma_{f,j,k}}, \tag{1}$$

where $\widehat{X}_{f,j,k}^{(i)} = 1/l\sum_{i=0}^{l-1} X_{f,j,k}^{(i)}$ and $\sigma_{f,j,k} = \sqrt{1/l\sum_{i=0}^{l-1}(X_{f,j,k}^{(i)} - \widehat{X}_{f,j,k}^{(i)})^2}$.

4.1.2. Padding. In the convolution operation of each layer, in order to keep the width and height of the feature graph unchanged, a very convenient way is to complete the operation, which is essentially to add 0 around the original image [14]. In the completion operation of size , 1 is added to the beginning and end of each row and column of the feature graph. As shown in Figure 7, the red part represents the part of completion, where $P = 1$.

4.1.3. Convolution. The output feature graph is obtained by convolution operation of input data and weight matrix [15]. The weight matrix is actually a tensor with dimension 4, where one dimension (size is F) represents the number of channels for input data in a certain layer, and the other dimension (size is F_p) represents the number of channels for output data in this layer. The remaining two dimensions define the size (width and height) of the Receptive Field (mathematically known as the convolution kernel), and in practice, the width and height are usually equal [16]. The receptive field enables the input image to be divided into multiple subsets for sequential processing, so

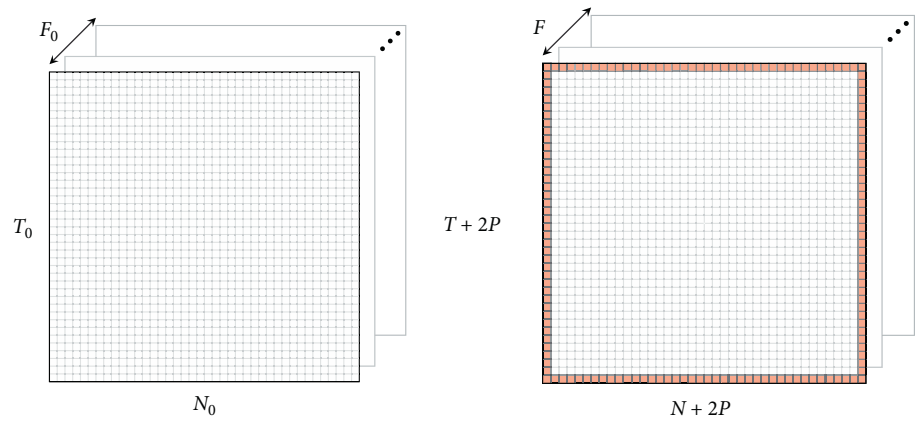


Figure 7: Padding of the feature map.

similar features can be searched in the input image without considering the location distribution of these features. The width and height of the output image are also determined by the step size. Figure 8 illustrates the above process graphically.

R_c represents the size of the convolution receptive field, and S_c represents the convolution step. Set the width of the input data as N and the height as T, and then the width and height of the output image can be calculated by the following formula:

$$N_p = \frac{N + 2P - R_c}{S_c} + 1, T_p = \frac{T + 2P - R_c}{S_c} + 1. \tag{2}$$

For the convolution operation of the first layer, input data is h, output is a, weight matrix is w, and the symbol " " is used to mark different parameters with the same meaning (for example, it represents the number of channels of the output layer and input, respectively). Then, the operation can be expressed as

$$a_{f,l,m}^{(i)} = \sum_{f'=0}^{F_v-1} \sum_{j=0}^{R_c-1} \sum_{k=0}^{R_c-1} w_{f',j,k}^{(v)f} h_{f',s_c l+j,s_c m+k}^{(i)(v)}, \tag{3}$$

where $v \in [0, N-1]$, $f \in [0, F_{v+1}-1]$, $l \in [0, N_{v+1}-1]$, $m \in [0, T_{v+1}-1]$. After each hidden layer, an activation function g is required to introduce nonlinear features. After counting completion operation, the input of the next layer is $h_{f,l+p,m+p}^{(i)(v+1)} = g(a_{f,l,m}^{(i)(v)})$.

4.1.4. Pooling. Pooling operation is essentially a dimensionality reduction process. By taking the average or maximum value of each small area (pooling receptive field, size R_p and step size S_p) of the input feature graph, an output with the same number of channels, but smaller width and height, can be obtained [17]. It is worth noting that pooling does not take into account the completed data, so the subscript $+P$ of the symbol is shown in the formula below. Figure 9 is the pooling operation diagram.

The average pooling operation can be expressed as

$$h_{f,l+p,m+p}^{(i)(v+1)} = a_{f,l,m}^{(i)(v)} = h_{f,s_p l+j+p,s_p m+k+p}^{(i)(v)} k_{(f,l,m)}^{(i)(p)} + p. \tag{4}$$

Maximum pooling is widely used in various studies, so the subsequent pooling operations in this paper are maximum pooling [18]. Setting $j_{(f,l,m)}^{(i)(p)}$, $k_{(f,l,m)}^{(i)(p)}$ represents the pixel index with the largest value in the pooled perception field, and get the input of the next layer:

$$a_{f,l,m}^{(i)(v)} = \sum_{j,k=0}^{R_p-1} h_{f,s_p l+j+p,s_p m+k+p}^{(i)(v)}. \tag{5}$$

4.1.5. Convolved to the Full Connection Layer. The last convolution layer of CNN usually makes the convolution receptive field the same size as the input feature graph [19]. In this case, the convolution operation becomes a weighted average process and transforms the two-dimensional input data into one-dimensional output. This process can be expressed as

$$a_f^{(i)(v)} = \sum_{f'=0}^{F_v-1} \sum_{l=0}^{R_c-1} \sum_{m=0}^{R_c-1} w_{f',l,m}^{(v)f} h_{f',l+p,m+p}^{(i)(v)}. \tag{6}$$

The next level of input is obtained by activating the function $h_f^{(i)(v+1)} = g(a_f^{(i)(v)})$. Figure 10 is the one-dimensional output obtained by convolution.

4.1.6. Fully Connected Layer. After the above operations, the rest of the model looks like a traditional fuzzy neural network (FNN) [20]. Its weighted average process is $a_f^{(i)(v)} = \sum_{f'=0}^{F_v-1} w_{f'}^{(v)f} h_{f'}^{(i)(v)}$

$$a_f^{(i)(v)} = \sum_{f'=0}^{F_v-1} w_{f'}^{(v)f} h_{f'}^{(i)(v)}. \tag{7}$$

4.1.7. Output Layer. The last layer of the model is the output layer:

$$a_f^{(i)(v)} = \sum_{f'=0}^{F_v-1} w_{f'}^{(v)f} h_{f'}^{(i)(v)}, \tag{8}$$

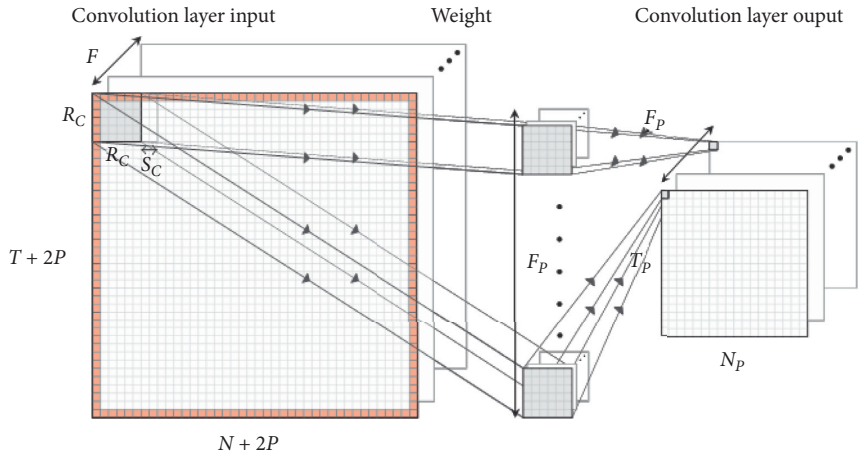


Figure 8: Convolution operation diagram.

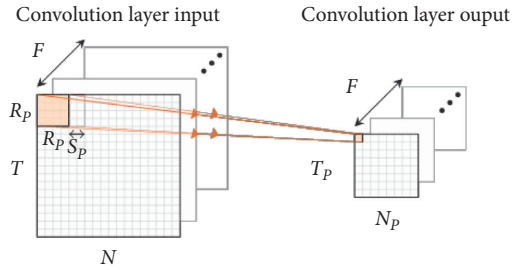


Figure 9: Pooling operation diagram.

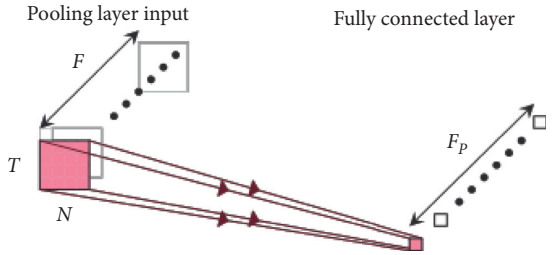


Figure 10: One-dimensional output obtained by convolution.

where V is the total number of layers of the model. o is the output function, and the following Softmax [21] function is usually used when the model is used to classify tasks:

$$\text{soft max}\left(a_f^{(i)(V-1)}\right) = \frac{e^{a_f^{(i)(V-1)}}}{\sum_{f'=0}^{F_{v-1}-1} e^{a_{f'}^{(i)(V-1)}}}. \tag{9}$$

4.2. Backpropagation Calculation Method of CNN. In traditional FNN, only two gradients from the output layer to FC and FC to the upper layer FC need to be calculated, while, in a typical CNN, four new gradients need to be calculated: FC=>Pool, Pool=>Conv, Conv=>Conv, Conv=>Pool [22, 23]. Set the CNN model of a layer, whose input layer is 0, hidden layer is v, and output layer is V.

4.2.1. Error Calculation. Define the total error rate for a layer relative to the output:

$$\varepsilon_{f,l[+p],m[+p]}^{(i)(v)} = \frac{\partial}{\partial a_{f,l,m}^{(i)(v)}} J(w), \tag{10}$$

where $[+P]$ only appears when completion operation is required, and J is the mutual entropy Loss Function of the model, expressed as

$$J(w) = -\frac{1}{l}\sum_{i=0}^{l-1}\sum_{f=0}^{F_v-1} \delta_{y^{(i)}}^{f} \ln h_f^{(i)(V)}, \tag{11}$$

where δ is the Dirac function, and y is the label value corresponding to the input data. The error calculation methods between different network layers are calculated as follows:

Output layer to full connection layer:

$$\varepsilon_f^{(i)(V-1)} = \frac{\partial}{\partial a_f^{(i)(v-1)}} J(w) = \sum_{f'=0}^{F_v-1} \frac{\partial h_{f'}^{(i)(V)}}{\partial a_f^{(i)(v-1)}} \frac{\partial}{\partial h_{f'}^{(i)(v)}} J(w). \tag{12}$$

Take the partial derivative of formula (9)

$$\frac{\partial h_{f'}^{(i)(V)}}{\partial a_f^{(i)(v-1)}} = \delta_{f'}^{f} h_f^{(i)(V)}\left(1 - h_f^{(i)(V)}\right). \tag{13}$$

Take the partial derivative of equation (11)

$$\frac{\partial}{\partial h_{f'}^{(i)(v)}} J(w) = -\frac{\delta_{y_{f'}^{(i)}}^{f'}}{\mathrm{Ih}_{f'}^{(i)(v)}}. \tag{14}$$

Substitute the above two equations into equation (12) to obtain

$$\varepsilon_f^{(i)(V-1)} = h_f^{(i)(V)}\left(1 - h_f^{(i)(V)}\right)\left(-\frac{\delta_{y_{f'}^{(i)}}^{f'}}{\mathrm{Ih}_{f'}^{(i)(v)}}\right) = \frac{1}{I}\left(h_f^{(i)(V)} - \delta_{y_f^{(i)}}^{f}\right). \tag{15}$$

The following mathematical derivation is similar to this process and is not explained step by step.

Full connection layer to full connection layer:

$$\varepsilon_f^{(i)(v)} = \sum_{f'=0}^{F_{v+1}-1} \frac{\partial a_{f'}^{(i)(v+1)}}{\partial a_f^{(i)(v)}} \varepsilon_{f'}^{(i)(v+1)} = \dot{g}\left(a_f^{(i)(v)}\right) \sum_{f'=0}^{F_{v+1}-1} w_f^{(v+1)} \varepsilon_{f'}^{(i)(v+1)}, \tag{16}$$

where $\dot{g}$ is the derivative of the activation function.

Full connection layer to pooling layer:

$$\varepsilon_{f,l,m}^{(i)(v)} = \sum_{f'=0}^{F_{v+1}-1} \frac{\partial a_{f'}^{(i)(v+1)}}{\partial a_{f,l,m}^{(i)(v)}} \varepsilon_{f'}^{(i)(v+1)} = \sum_{f'=0}^{F_{v+1}-1} w_{f,l,m}^{(v)f'} \varepsilon_{f'}^{(i)(v+1)}. \tag{17}$$

Pooling layer to convolution layer:

$$\begin{aligned} \varepsilon_{f,l+P,m+P}^{(i)(v)} &= \sum_{l'=0}^{N_{v+1}-1} \sum_{m'=0}^{T_{v+1}-1} \frac{\partial a_{f,l',m'}^{(i)(v+1)}}{\partial a_{f,l,m}^{(i)(v)}} \varepsilon_{f,l',m'}^{(i)(v+1)} \\ &= \dot{g}\left(a_{f,l,m}^{(i)(v)}\right) \sum_{l'=0}^{N_{v+1}-1} \sum_{m'=0}^{T_{v+1}-1} u_{f,l,m,l',m'}^{(i)(v)} \varepsilon_{f,l',m'}^{(i)(v+1)}. \end{aligned} \tag{18}$$

According to the sampling theorem,

$$u_{f,l,m,l',m'}^{(i)(v1)} = \begin{cases} 1, & l = j_{f',l',m'}^{(i)(p)} \\ 1, & m = k_{f',l',m'}^{(i)(p)} \end{cases} \tag{19}$$

Convolution layer to convolution layer:

$$\begin{aligned} \varepsilon_{f,l+P,m+P}^{(i)(v)} &= \sum_{f'=0}^{F_{v+1}-1} \sum_{l'=0}^{N_{v+1}-1} \sum_{m'=0}^{T_{v+1}-1} \frac{\partial a_{f',l',m'}^{(i)(v+1)}}{\partial a_{f,l,m}^{(i)(v)}} \varepsilon_{f',l'+P,m'+P}^{(i)(v+1)} \\ &= \dot{g}\left(a_{f,l,m}^{(i)(v)}\right) \sum_{f'=0}^{F_{v+1}-1} \sum_{l'=0}^{N_{v+1}-1} \sum_{m'=0}^{T_{v+1}-1} \sum_{j=0}^{R_c-1} \sum_{k=0}^{R_c-1} \frac{\partial a_{f',l',m'}^{(i)(v+1)}}{\partial a_{f,l,m}^{(i)(v)}} w_{f,j,k}^{(i)f'} \varepsilon_{f',l'+P,m'+P}^{(i)(v+1)}. \end{aligned} \tag{20}$$

Convolution layer to pooling layer:

$$\begin{aligned} \varepsilon_{f,l+P,m+P}^{(i)(v)} &= \sum_{f'=0}^{F_{v+1}-1} \sum_{l'=0}^{N_{v+1}-1} \sum_{m'=0}^{T_{v+1}-1} \frac{\partial a_{f',l',m'}^{(i)(v+1)}}{\partial a_{f,l,m}^{(i)(v)}} \varepsilon_{f',l'+P,m'+P}^{(i)(v+1)} \\ &= \sum_{f'=0}^{F_{v+1}-1} \sum_{j=0}^{R_c-1} \sum_{k=0}^{R_c-1} w_{f,j,k}^{(i)f'} \varepsilon_{f',l+2P-j,m+2P-k}^{(i)(v+1)}. \end{aligned} \tag{21}$$

4.2.2. Weights Update. For weight update calculation, let $\Delta = w_{\mathrm{new}} - w_{\mathrm{prew}}$, where w_{new} and w_{prew} respectively represent the weights before and after the update [24]. Also, consider the following situations:

Full connection layer to full connection layer: $\Delta_{f'}^{(v)f} = \sum_{i=0}^{l-1} h_{f'}^{(i)(v)} \varepsilon_f^{(i)(v)}$

Full connection layer to pool layer: $\Delta_{f',j,k}^{(v)f} = \sum_{i=0}^{l-1} h_{f',j+P,k+P}^{(i)(v)} \varepsilon_f^{(i)(v)}$

Convolution layer to convolution layer, convolution layer to pooling layer, and convolution layer to input layer: $\Delta_{f',j,k}^{(v)f} = \sum_{i=0}^{l-1}\sum_{i=0}^{l-1}\sum_{i=0}^{l-1} h_{f',l+j,m+k}^{(i)(v)} \varepsilon_{f,l+P,m+P}^{(i)(v)}$

4.3. Emergency Q&A System Based on CNN. In the intelligent community emergency service platform, intelligent question answering is very critical [25]. It is the core of data sharing and data interaction, and the key element of accurately and comprehensively obtaining external data such as comprehensive community governance, community hospitals, intelligent city operation, and management and emergency resources [26], the premise of intelligent question answering is the design of semantic matching algorithm, so as to grasp the real-time information of danger sources, the classification and management of vulnerable groups such as the old and the weak, the sick and the young, and the timely deployment of spatial information such as traffic roads, public places, and water supply and power facilities.

Studies at home and abroad show that convolutional neural network has outstanding advantages in text feature extraction, and the work of sentiment analysis, text classification, semantic matching, and intelligent recognition based on convolutional neural network has achieved remarkable results in open data sets and evaluation [27]. A typical convolutional neural network structure used for sentence modeling is shown in Figure 11.

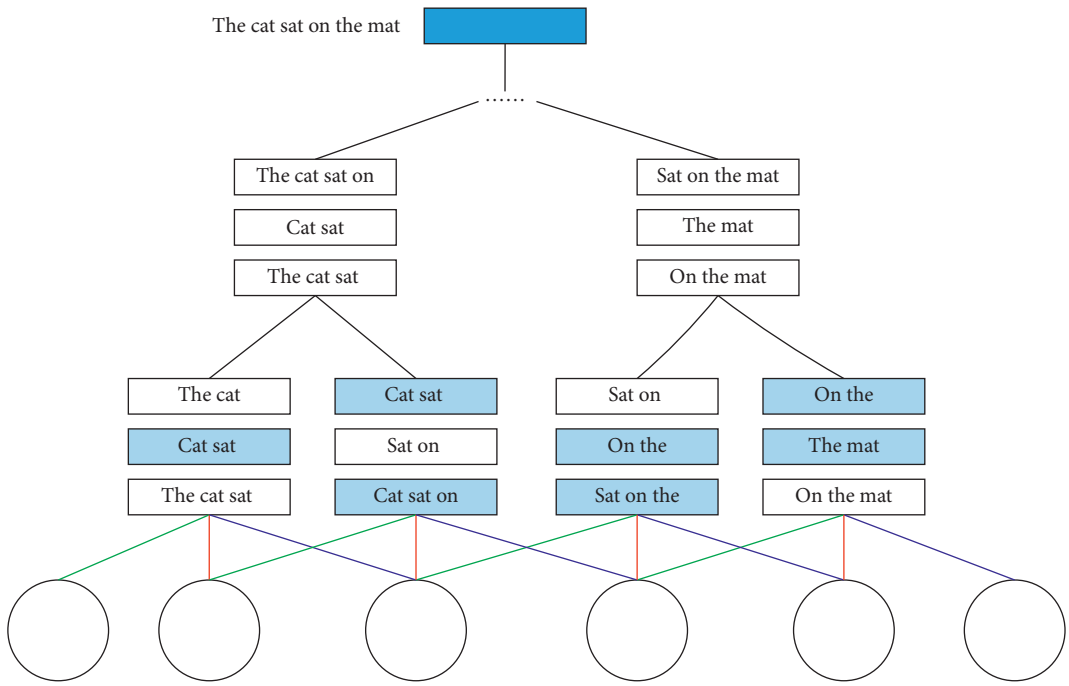


Figure 11: Sentence modeling structure diagram based on CNN.

Convolutional neural network needs to process two sentences and finally combine the two sentences for classification [28]. Therefore, this paper uses a parallel convolutional neural network architecture to process the two sentences. The overall structure of the network is shown in Figure 12. The figure shows two parallel convolution neural network frameworks to deal with two sentences, each sentence word vector of two-dimensional matrix, respectively, as two convolution of the neural network inputs, and then the convolutional neural network through the layers of convolution and pooling high-level abstract composition feature extracting operation, after the last convolution and pooling feature vectors of two sentences, Then, input to a nonlinear hidden layer for nonlinear mapping, and finally input to a logistic regression classifier for classification. The output of the network is a class standard of 0 or 1 and the confidence degree of classification.

However, in the above basic parallel convolutional neural network, there is a lack of interaction between the two sentences at the convolutional layer and the pooling layer, and the low-dimensional vector finally obtained only contains the high-level information of the sentence [29]. In this process, some relevant information of the sentence itself may be lost. In this paper, an improved attention mechanism is added after each pooling layer to modify the model structure. The overall structure of convolutional neural network with attention mechanism is shown in Figure 13. The figure shows that, after each convolution operation to join focus attention mechanism, and the network has been unearthed in the underlying computation of the semantic relationship between two sentences, through constant weight highlighting the correlation and difference between the two sentences to get an effective sentence said after the information, so in the actual experiments, it will also obtain a good effect.

Question answering engine is a control module of the whole question answering system, which is mainly responsible for the system input processing and some external information application. The semantic matching module is mainly CNN algorithm [30]. In this paper, the offline way is used to train the semantic matching model first, and the model code is designed abstractly. A matching interface is directly provided to the question and answer engine module. The system construction process is divided into system architecture and online deployment. Figure 14 shows the overall system architecture. The automatic question answering system based on CNN can digitally decompose the elements of emergency plan, form the information of emergency plan into data fields, store the data fields, and establish the plan database. The system provides personnel database input for emergency volunteers by function, profession, industry, and region. At the same time, maintain the personnel, with the function of adding, deleting, modifying, and querying personnel information, and train emergency volunteers. The system provides the functions of emergency training plan and training knowledge base. When people need to evacuate and take refuge, the system can take measures to arrange people's refuge according to the evacuation route provided by the system, forming a linkage whole.

Four different speech signals are selected to realize effective speech classification and recognition through CNN. The parameters of model training are set as follows: Learning Rate: 0.0001; Batch Size: 420, that is, full data set input; iteration times (Epoch): 8 0. After each iteration, the model parameters were updated, and the loss value and accuracy of the model on the test set and the accuracy of the model on the verification set were calculated. In the training process, the training loss, training accuracy, and test accuracy of the model are shown in Figure 15.

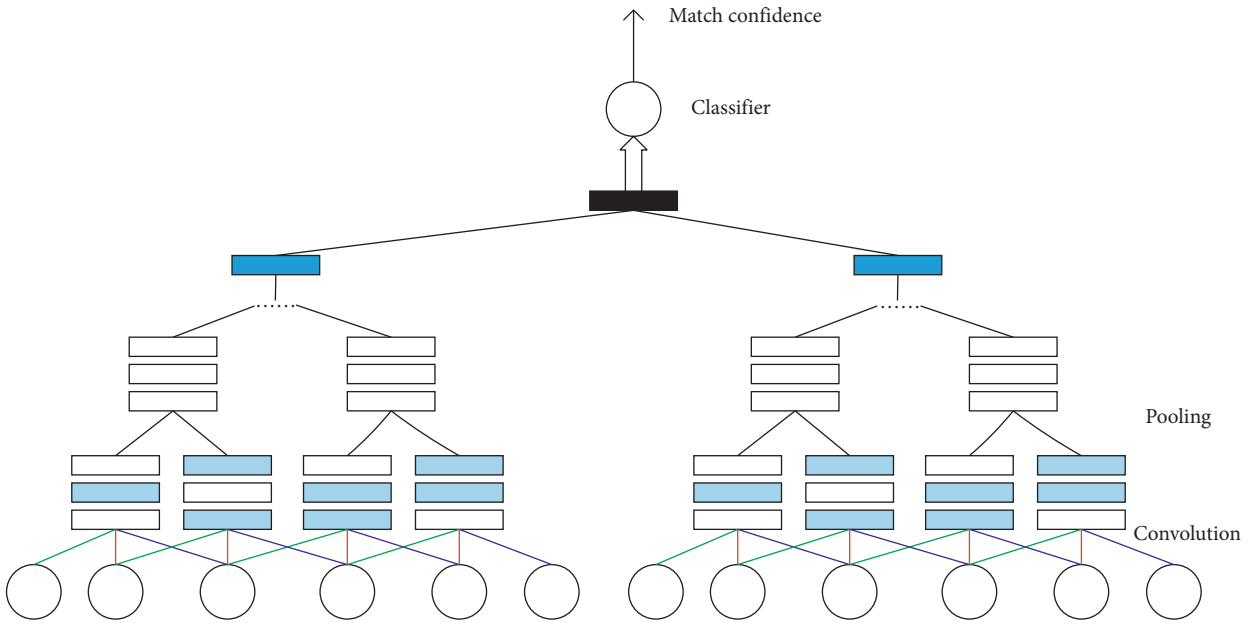


Figure 12: Parallel CNN structure diagram.

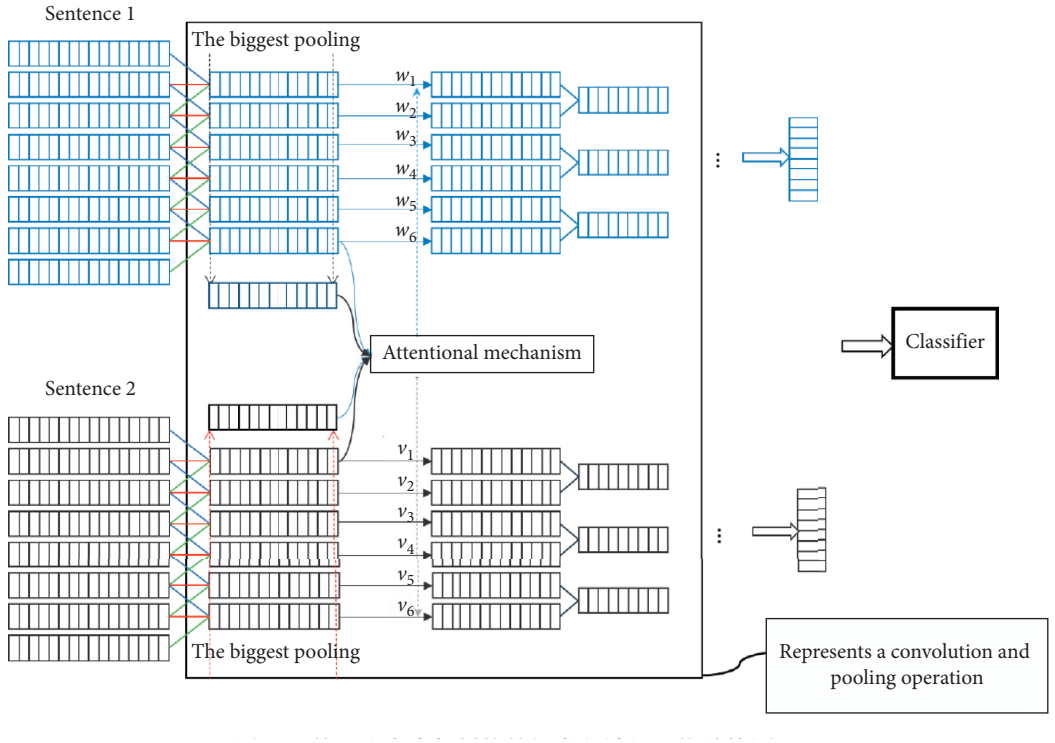


Figure 13: Structure diagram of parallel CNN based on attention mechanism.

It can be seen from the figure that the loss function decreases with the increase of iteration times and finally approaches 0 after 80 iterations. The classification accuracy of training set and verification set fluctuates with the increase of iterations and converges stably to 100% after the number of iterations exceeds 60. The samples in the test set were used to test the trained model, and the classification accuracy was shown in Table 1.

It can be seen from the classification results of CNN that the speech recognition algorithm has high accuracy and can

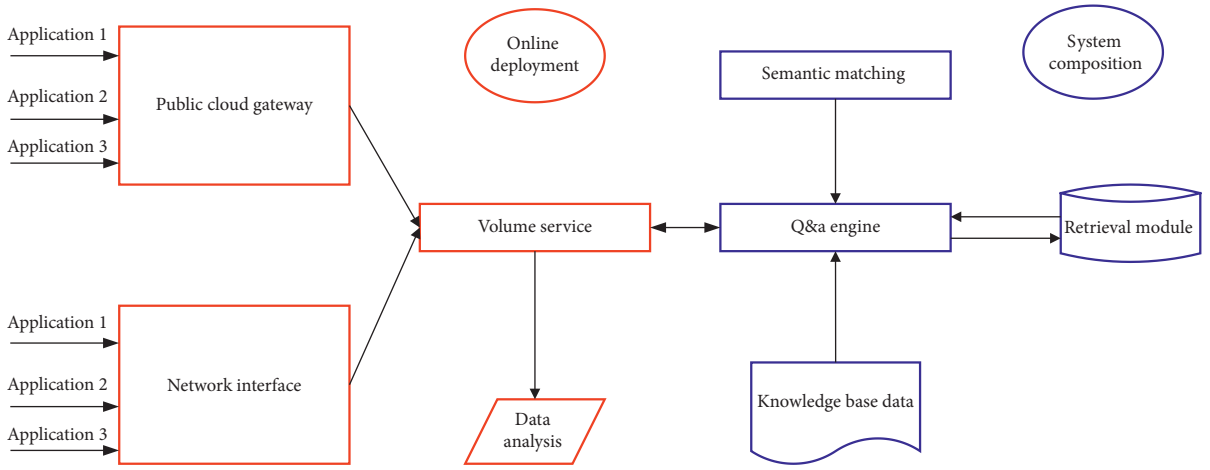


Figure 14: Overall architecture of question answering system.

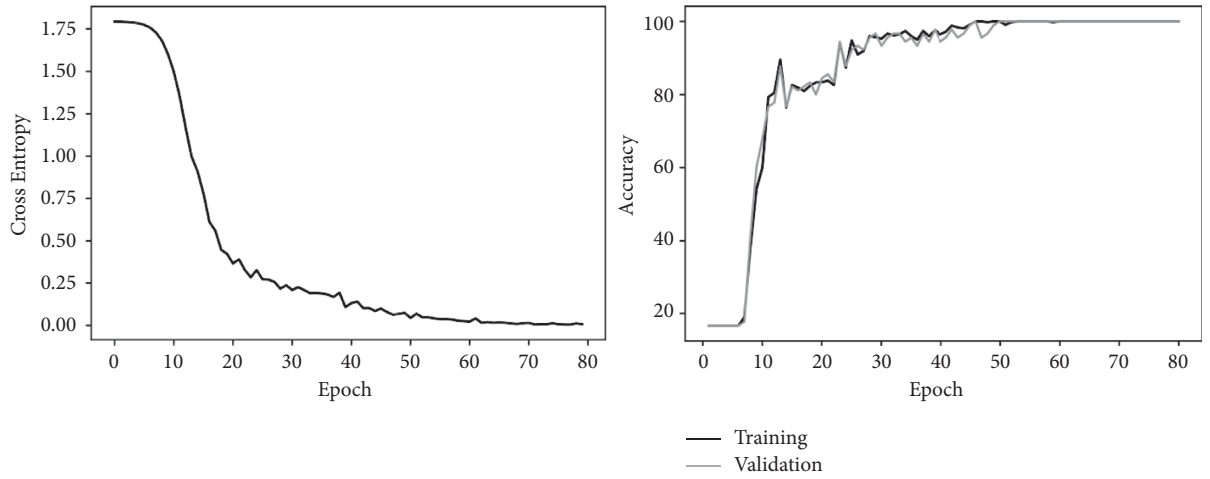


Figure 15: Change of loss function and classification accuracy with the number of iterations.

Table 1: Recognition accuracy of CNN.

Class of signal	Type 1	Type 2	Type 3	Type 4
Recognition accuracy	0.7293	1.0000	0.8770	0.9569

accurately identify the categories of speech signals, effectively avoiding artificial subjective arbitrariness.

5. Conclusion

This paper developed the wisdom of the community emergency service platform based on CNN, fully used the Internet, cloud computing and big data, artificial intelligence cutting-edge technology, and had realized the display of all kinds of sensors, running state, and the distribution of testing data, based on the analysis of speech recognition, the wisdom, creating various warning strategies, as well as the trend of changes of unexpected events of the community to make predictions, and the factors affecting the health of community residents would be eliminated as soon as possible. The main work and conclusions are summarized as follows:

(1) Through the analysis of the requirements of the emergency service platform of intelligent community, the design principles of the emergency service platform are clarified. The used case view, development view, and scene view are designed to solve the problem of data interaction and provide a basis for the subsequent module development.

(2) Through the emergency question answering system designed by CNN, the semantic matching algorithm has high precision, the automatic question answering system construction is reasonable and feasible, and the system speech recognition and return accuracy is high, thus providing strong technical support for the community emergency response.

(3) The developed intelligent community emergency service platform has the characteristics of high security, wide scalability, and strong stability. It is of great significance to optimize and improve the community emergency management ability and promote the construction of safe intelligent community. It provides a refined, easy to operate,

efficient, and all-round information tool for community emergency management.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by Health and Economic and Social Development Research Center of Guangxi University Humanities and Social Sciences Key Research Base: The practical mechanism of social work in public health emergencies in Guangxi (no. 2021RWB05).

References

[1] S. Yang, F. R. Jiménez, J. Hadjimarcou, and G. L. Frankwick, "Functional and social value of Chinese brands," *Journal of Global Marketing*, vol. 32, no. 3, pp. 200–215, 2019.

[2] A. Meijer and M. P. R. Bolívar, "Governing the smart city: a review of the literature on smart urban governance," *International Review of Administrative Sciences*, vol. 82, no. 2, pp. 392–408, 2016.

[3] I. Kanako, J.-i. Aoki, N. Ikeda et al., "Use of a human-type communication robot to evaluate the categorized communicative ability of older adults with dementia[J]," *Geriatrics and Gerontology International*, vol. 18, no. 1, pp. 188–190, 2018.

[4] H. Hyun and E. W. Johnston, "A framework for analyzing digital volunteer contributions in emergent crisis response efforts," *New Media & Society*, vol. 19, no. 8, pp. 1308–1327, 2017.

[5] A. Plough, J. E. Fielding, A. Chandra et al., "Building community disaster resilience: perspectives from a large urban county department of public health," *American Journal of Public Health*, vol. 103, no. 7, pp. 1190–1197, 2013.

[6] B. D. Eric, P. K. Kannan, and J. Slotegraaf Rebecca, "Branded apps and their impact on firm value: a design perspective [J]," *Journal of Marketing Research*, vol. 56, no. 1, pp. 76–88, 2019.

[7] Y. Zhao, Y. Chen, R. Zhou, and Y. Ci, ""Factors influencing customers' willingness to participate in virtual brand community's value co-creation," *Online Information Review*, vol. 43, no. 3, pp. 440–461, 2019.

[8] X. Yang, Y. Ma, and J. Li, "Quality problems and countermeasures in 2012 edition of medical record homepage," *Chinese Medical Record English Edition*, vol. 1, no. 5, pp. 206–209, 2013.

[9] P. Shi, "On the role of government in integrated disaster risk governance-Based on practices in China," *International Journal of Disaster Risk Science*, vol. 3, no. 3, pp. 139–146, 2012.

[10] R. Agranoff and M. McGuire, "Big questions in public network management research," *Journal of Public Administration Research and Theory*, vol. 11, no. 3, pp. 295–326, 2001.

[11] K. R. Isett, I. A. Mergel, K. Leroux, P. A. Mischen, and R. K. Rethemeyer, "Networks in public administration scholarship: understanding where we are and where we need to go [J]," *Journal of Public Administration Research and Theory*, vol. 21, pp. 157–173, 2011.

[12] Yi Peng, Y. Wu, J. Shen, W. Shang, and Z. Mo, "Smart city with Chinese characteristics against the background of big data: idea, action and risk [J]," *Journal of Cleaner Production*, vol. 173, pp. 60–66, 2018.

[13] H. Yu, M. Matsumoto, M. Okita et al., "The vanguard of community-based integrated care in Japan: the effect of a rural town on national policy [J]," *International Journal of Integrated Care*, vol. 17, no. 2, pp. 1–9, 2017.

[14] M. Alonso Jose, R. Andrews, C. Judith, and D. Diaz-Fuentes, "Factors influencing citizens' co-production of environmental outcomes: a multi-level analysis [J]," *Public Management Review*, vol. 21, no. 12, pp. 1620–1645, 2019.

[15] M. Olmedilla and R. Martinez-Torres, "Identification of innovation solvers in open innovation communities using swarm intelligence[J]," *Technological Forecasting and Social Change*, vol. 109, pp. 15–24, 2016.

[16] S. L. Vargo and R. F. Lusch, "Institutions and axioms: an extension and update of service-dominant logic," *Journal of the Academy of Marketing Science*, vol. 44, no. 1, pp. 5–23, 2016.

[17] X. Li and X. Li, "Big data and its key technology in the future," *Computing in Science & Engineering*, vol. 20, no. 4, pp. 75–88, 2018.

[18] D. Silver, A. Huang, C. J. Maddison et al., "Mastering the game of Go with deep neural networks and tree search," *Nature*, vol. 529, no. 7587, pp. 484–489, 2016.

[19] Z. E. Ross, M. A. Meier, and E. Hauksson, "P wave arrival picking and first-motion polarity determination with deep learning," *Journal of Geophysical Research: Solid Earth*, vol. 123, no. 6, pp. 5120–5129, 2018.

[20] W. Zhu and C. Beroza Gregory, "Phase Net: a deep-neural-network-based seismic arrival-time picking method [J]," *Geophysical Journal International*, vol. 216, no. 1, pp. 261–273, 2019.

[21] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: the KITTI dataset," *The International Journal of Robotics Research*, vol. 32, no. 11, pp. 1231–1237, 2013.

[22] K. Jia, B. Cai, C. Qing, X. Xu, and D. Tao, "Dehaze Net: an end-to-end system for single image haze removal [J]," *IEEE Transactions on Image Processing*, vol. 25, no. 11, pp. 5187–5198, 2016.

[23] X.-z. Zhao, D. Cheng, Y.-f. Zhang, and M.-y. Li, "Experimental and numerical study on the hydrodynamic characteristics of solitary waves passing over A submerged breakwater," *China Ocean Engineering*, vol. 33, no. 3, pp. 253–267, 2019.

[24] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," *IEEE Signal Processing Letters*, vol. 23, no. 10, pp. 1499–1503, 2016.

[25] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2017.

[26] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: a deep convolutional encoder-decoder architecture for image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.

[27] K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask R-CNN," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 2, pp. 386–397, 2020.

[28] A. S. Qureshi, A. Khan, A. Zameer, and A. Usman, "Wind power prediction using deep neural network based meta

regression and transfer learning," *Applied Soft Computing*, vol. 58, no. 1, pp. 742–755, 2017.

[29] P. Song, "Transfer linear subspace learning for cross-corpus speech emotion recognition," *IEEE transactions on affective computing*, vol. 10, no. 2, pp. 265–275, 2019.

[30] C. Wang, L. Gong, Y. Qi, X. Li, X. Yuan, and X. Zhou, "DLAU: a scalable deep learning accelerator unit on FPGA[J]," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems: A publication of the IEEE Circuits & Systems Society*, vol. 36, no. 5, pp. 513–517, 2017.

Hindawi
Scientific Programming
Volume 2021, Article ID 7607623, 8 pages
https://doi.org/10.1155/2021/7607623

Research Article

Prediction of Evolution and Development Trend in Sports Industry Cluster Based on Particle Swarm Optimization

Rui Cong and Hailong Wang

School of Sports Institute, East China University of Technology, JiangXi, NanChang 330022, China

Correspondence should be addressed to Hailong Wang; 201962029@ecut.edu.cn

Received 15 November 2021; Revised 6 December 2021; Accepted 14 December 2021; Published 26 December 2021

Academic Editor: Baiyuan Ding

Sports industry cluster refers to the economic phenomenon that sports related enterprises gather in a large number in a specific area. For the sports enterprises in the cluster, they can obtain huge competitive advantages through enterprise agglomeration, thus obtaining better development and rich economic benefits. The optimization of particle swarm optimization is interlinked with the agglomeration of industrial clusters. Therefore, in view of the limitation of the standard particle swarm optimization (PSO) algorithm, an improved particle swarm optimization algorithm-diaphragm particle swarm optimization (D-PSO) was proposed and used to simulate the formation of sports industry clusters. D-PSO introduces the cell membrane processing mechanism of the biological system into the PSO algorithm, which improves the ability of the PSO algorithm to get rid of local extremum points. The competitiveness value of the sports industry cluster is the value of the objective function solved by the D-PSO algorithm. The geographical coordinates of the industrial cluster were the locations in the particle search space of the D-PSO algorithm. The D-PSO algorithm is used to simulate the aggregation process of enterprises in the cluster. Compared with the standard PSO, the D-PSO algorithm has better convergence performance and optimal rate. The results of case analysis show that the proposed method can effectively predict the development trend of sports industrial clusters.

1. Introduction

Industrial cluster is the mainstream trend of modern industrial development and the main driving force of industrial upgrading. Industrial clusters play an important role in promoting regional industrial development and enhancing regional competitiveness. By reducing transaction costs and sharing facilities, industrial clusters can significantly reduce cost advantages [1–5]. At the same time, knowledge spillovers and diffusion among enterprises in industrial clusters can provide a source for industrial upgrading. Industrial clusters can promote the expansion of the scale of industrial fields and the continuous and rapid development of industries. Industrial cluster is an economic phenomenon based on self-organization structure. Self-organization structure is characterized by self-adaptation and self-organization, which coincides with industrial clusters. Similarly, if the cluster is regarded as a system composed of many enterprises and institutions, it is also a self-organizing system [6–8]. The formation of industrial clusters is also evolving through the open dissipative structure.

Sports industry cluster is the new highlight and main melody of today's sports industry development, and it is also a distinctive organizational form in social and economic development, which plays an important role in promoting regional economic progress and promoting local industrial competitiveness [9–11]. Sports industrial cluster is the subordinate concept of industrial cluster. Although the research on sports industrial cluster at home and abroad has been gradually enriched in recent years, there is no unified view on the definition of this concept so far.

In recent years, sports industry cluster has become a hot field of academic research; many researchers will be full of research enthusiasm into it and has made many scientific research results. However, there are still the following deficiencies in the current research on sports industrial cluster

[12–16]: (1) the research on the generation process and evolution law of this economic phenomenon, sports industrial cluster, is still in the stage of qualitative discussion, and the research depth is not enough; (2) the current research mainly aims at the existing industrial clusters, carries on the case analysis, carries on the conceptual elaboration to be quite many, and rarely simulates the industrial cluster formation process through the computer.

Particle swarm optimization is interlinked with the agglomeration of sports industry clusters. The formation of sports cluster is actually a self-organizing process. The particle swarm optimization algorithm is a self-organizing algorithm [17–19], and its optimization process is also self-organizing. If the enterprises in the cluster are regarded as the particles in the particle swarm optimization algorithm and the location of the cluster is the location where the cluster has the greatest competitiveness and it is regarded as the location of the optimal solution in the particle swarm optimization algorithm [20], then the aggregation process of industrial clusters can be regarded as the optimization process of particle swarm optimization. It can be seen that the optimization of particle swarm and the agglomeration of sports industry cluster are interlinked.

Therefore, this paper makes a detailed analysis of industrial clusters based on self-organization and PSO algorithm. An improved PSO algorithm-diaphragm particle swarm optimization (D-PSO) is proposed to accurately solve the evolution and formation mechanism of industrial clusters. Studying industrial clusters from the perspective of self-organization provides a new analytical thinking for academic research of industrial clusters.

2. Analysis of Self-Organization Characteristics of Sports Industrial Clusters

2.1. Formation Mechanism of Sports Industry Cluster. Industrial cluster is an important carrier of regional economic development. The emergence of industrial cluster can effectively enhance the competitiveness of regional economy and greatly promote the development of regional economy. Therefore, the evolution process and formation law of industrial clusters are topics worthy of study.

At present, there are two main theories about industrial clusters [17, 18]: one is Marshall's industrial zone theory and the other is Porter's cluster theory. The concept of industrial cluster can be defined in broad sense and narrow sense. He et al. [19] on the basis of Porter's research results on industrial clusters, combined with the characteristics of sports industry, defined sports industrial clusters as a group of geographically adjacent and interrelated enterprises and institutions, which is a broad concept of industrial clusters. Tao et al. [20] mainly studied the main feature of industrial clusters-spatial agglomeration from the perspective of economics and geography. Wei [21] pointed out the important role of innovation network in cluster development. Coordination among enterprises in the network will effectively enhance the overall competitive advantage. The development of network innovation will be an important driving force for the future growth of industrial clusters.

Researchers have analyzed and discussed the formation mechanism of industrial clusters from different perspectives, and their representative views mainly include four kinds: factor theory, model theory, dynamic theory, and system theory. The formation mechanism of industrial clusters in "Factor Theory" is shown in Figure 1.

The view of industrial cluster formation mechanism of "model theory" is a further extension of "factor theory" [22]. This view not only lists the essential conditions for the formation of industrial mechanism but also logically analyzes in detail how different elements interact to form clusters. However, scholars who hold this view mostly analyze how different elements work from a static point of view. The formation and development of industrial clusters is a dynamic change process, and it is incomplete to analyze the formation mechanism of industrial clusters from a static point of view. "Dynamic theory" is a new viewpoint formed with the rapid development of industrial clusters in recent years, and it is the most advanced theory to analyze the formation mechanism of industrial clusters at present. This view regards the formation of industrial clusters as a dynamic change process and analyzes the formation mechanism of each stage from different stages of industrial clusters formation. The viewpoint of industrial cluster formation mechanism of "dynamic theory" accords with the basic concept of particle swarm.

2.2. Self-Organizing Structure Analysis. Since 1960s, people began to study the theory of self-organization. Self-organization theory plays an important role in the research of complexity science. From the perspective of system science, self-organization can be understood as an orderly system structure in the system environment, in which internal members are orderly without unified leadership and command [23, 24]. The self-organizing system is not formed by external forces. During its birth, internal members interact with each other, which is a spontaneous behavior. On the contrary, in the process of organization formation, the organization formed by external driving force is generally called other organization.

Self-organization structure is characterized by self-adaptation and self-organization, which coincides with industrial clusters. Actually, industrial cluster is a self-organizing system. The formation of industrial clusters is also evolving through the open dissipative structure. The whole process of industrial cluster from the agglomeration of small-scale enterprises to its development and growth to its decline is accompanied by the influence of "fluctuation" factors. In industrial clusters, the emergence of new orderly structure is also completed under the action of fluctuation, which is the expression of catastrophe theory in clusters. Thus, industrial clusters are self-organizing structures, and their evolution process is also self-organizing.

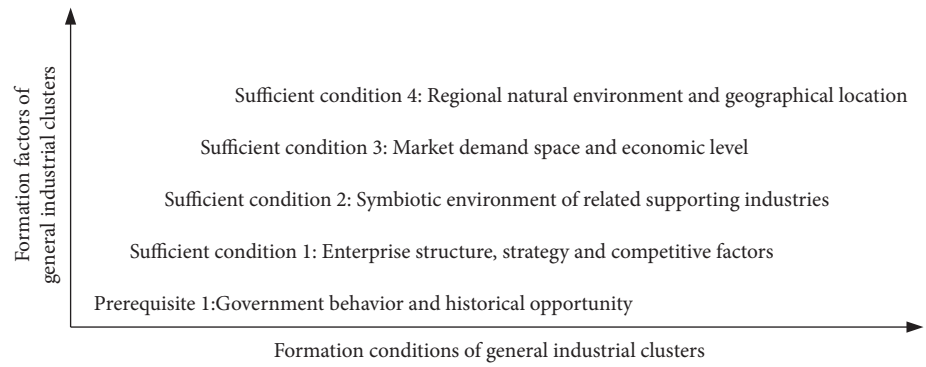


Figure 1: The formation mechanism of general industrial clusters.

3. Diaphragm Particle Swarm Optimization (D-PSO) Algorithm

3.1. Principle of PSO Algorithm. In the PSO algorithm, the potential solution of each optimization problem corresponds to the position of a bird in the search space [25–28], which we call "Particle." Each particle will move in the solution space, and the direction and distance of their flight will be determined by a velocity. There is also a fitness value determined by the optimized function. Then, the particles follow the current optimal particles and search in the solution space.

Suppose that in a D-dimensional target search space, there are N particles representing the potential solutions of the optimization problem to form a group. The position of the particle i is X_i, and the "flying" speed is V_i. The fitness value of X_i can be calculated by substituting it into the objective function, and the advantages and disadvantages can be measured according to its fitness value. The i-th particle determines the next movement through the individual extreme value pbest and the global extreme value gbest. The particle updates its own speed and new position according to the following formula:

$$\begin{aligned} v_{i\,d}^{k+1} &= v_{i\,d}^{k} + c_1 \mathrm{rand}_1^k\left(\mathrm{pbest}_{i\,d}^{k} - x_{i\,d}^{k}\right) + c_2 \mathrm{rand}_2^k\left(\mathrm{gbest}_{d}^{k} - x_{i\,d}^{k}\right), \\ x_{i\,d}^{k+1} &= x_{i\,d}^{k} + v_{i\,d}^{k+1}, \end{aligned} \tag{1}$$

where $v_{i\,d}^{k}$ is the velocity of the D dimension of particle i in the k-th iteration, c_1 and c_2 are learning factors (nonnegative constants), $\mathrm{rand}_{1,2}$ is a random number between (0,1), and $x_{i\,d}^{k}$ is the position of the D dimension of particle i in the k-th iteration.

3.2. D-PSO Algorithm Design. Cell diaphragm principle [29] is a bionic technology that simulates the intelligent behavior of the biological cell system, and it is a heuristic random search algorithm that combines deterministic and random selection. The diaphragm algorithm is inspired by somatic cell theory and network theory and realizes the self-regulation function similar to biological cells and the function of generating different diaphragms. The diaphragm particle swarm optimization (D-PSO) model is formed by introducing cell membrane processing mechanism (diversity, self-regulation, diaphragm memory, and so on) into the PSO algorithm. This optimization model combines the advantages of the PSO algorithm and diaphragm algorithm, thus avoiding the disadvantage that the PSO algorithm is easy to fall into local extremum and improving the convergence speed and accuracy of the later evolutionary algorithm. The diversity of septum particles is ensured while keeping individuals with high fitness, thus avoiding the premature phenomenon.

In the process of updating the particle population, the strategy of maintaining diversity based on the concentration mechanism is adopted, which makes the particles (diaphragms) of each fitness level maintain a certain concentration in the new generation particle population. The concentration of the i-th particle (diaphragm) is defined as follows:

$$D(x_i) = \frac{1}{\sum_{j=1}^{N+M} f(x_i) - f\left(x_j\right)}, \quad i = 1, 2, \ldots, N + M. \tag{2}$$

The probability selection method based on particle concentration is as follows [30]:

$$\begin{aligned} P(x_i) &= \frac{1/D(x_i)}{\sum_{i=1}^{N+M}\left(1/D(x_i)\right)}, \\ &= \frac{\sum_{j=1}^{N+M} f(x_i) - f\left(x_j\right)}{\sum_{i=1}^{N+M}\sum_{j=1}^{N+M} f(x_i) - f\left(x_j\right)}, \quad i = 1, 2, \ldots, N + M, \end{aligned} \tag{3}$$

where x represents the i-th diaphragm (particle) and $f(x_i)$ represents the fitness value of the i-th diaphragm (particle). It can be seen that the fewer the diaphragms similar to diaphragm i, the greater the probability that diaphragm i will be selected. On the contrary, the more the diaphragms similar to diaphragms, the smaller the probability of diaphragm I being selected. This makes individuals with low fitness also get the opportunity of evolution. Therefore, in theory, the probability selection formula based on antibody concentration can ensure the diversity of diaphragm.

The flowchart of D-PSO is shown in Figure 2.

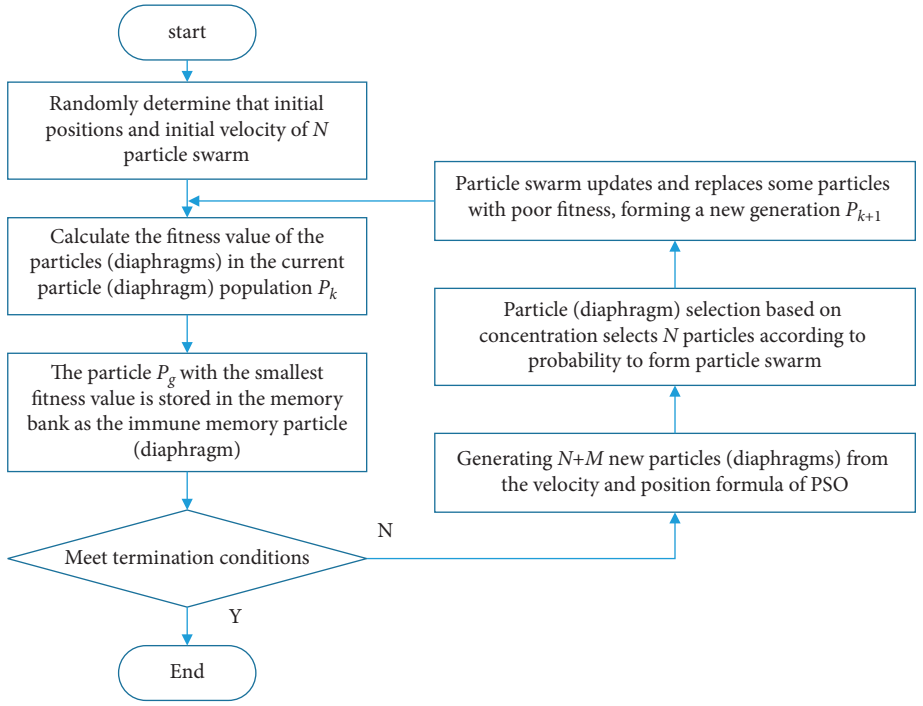


Figure 2: D-PSO algorithm flowchart.

4. Industrial Cluster Evolution Method Based on D-PSO

4.1. Similarity between Industrial Cluster Evolution Mechanism and PSO Algorithm. Through the above discussion, we know that the formation of industrial clusters is actually a self-organization process. However, the PSO algorithm is self-organizing, and its optimization process is also self-organizing. If the enterprises in the cluster are regarded as the particles in the PSO algorithm, the location of the cluster is the most competitive position of the cluster. Considering it as the location of the optimal solution in the PSO algorithm, the clustering process of industrial clusters can be regarded as the optimization process of the particle swarm optimization. It can be seen that the optimization of particle swarm is interlinked with the agglomeration of industrial clusters.

For particle swarm, in the evolution stage, particle swarm performs organized evolution under the domination of "self-consciousness" and "social consciousness" according to certain rules, from an unstable nonlinear system to a stable state mutation. For industrial clusters, enterprises spontaneously move closer to the cluster group under the impetus of self-interest and various external forces.

For particle swarm, particle swarm can find the optimal value under the guidance of specific rules through a lot of iterative learning. For industrial clusters, in the process of evolution, the same industry within the cluster through continuous cooperation and competition to eliminate the lack of competitive enterprises retains the competitive enterprises. Economies of scale of clusters are reflected. And finally, the cluster achieves a dynamic and stable state.

Therefore, in order to use the PSO algorithm to simulate the formation of industrial clusters, we use particle swarm optimization to solve the problem. The competitiveness value of the industrial cluster is the value of the objective function solved by the PSO algorithm. The geographical coordinates of the industrial cluster are the position in the particle search space in the PSO algorithm. There must be "cooperation" and "competition" between enterprises within the cluster, which can be achieved through the "self-cognition" part and "society" part of the PSO algorithm. Finally, taking the sports industrial cluster as an example, we use the PSO algorithm to simulate the aggregation process of enterprises in the cluster and thus predict the development of the industrial cluster.

4.2. Industrial Cluster Evolution Model Based on D-PSO Algorithm. In order to simulate the formation of industrial clusters with the particle swarm optimization algorithm, we cluster industrial clusters into particles. The competitiveness of industrial clusters is the value of the objective function solved in the particle swarm optimization algorithm. The geographical coordinates of industrial clusters are the positions in particle search space in the D-PSO algorithm. There must be "cooperation" and "competition" among enterprises in the cluster, which can be realized by the "self-awareness" part and "society" part of the D-PSO algorithm.

To sum up, we have established the evolution model of industrial clusters based on particle swarm optimization. The basic speed and location updating formulas in the model are as follows, in which each variable endows the industrial clusters with characteristics. The iterative process is carried out according to the proposed D-PSO algorithm. The following is the update expression:

$$\begin{aligned} v_{i\,d}^{t+1} &= w v_{i\,d}^{t} + c_1 r_1 \left(P_{i\,d}^{t} - X_{i\,d}^{t}\right) + c_2 r_2 \left(P_{g\,d}^{t} - X_{i\,d}^{t}\right), \\ X_{i\,d}^{t+1} &= X_{i\,d}^{t} + v_{i\,d}^{t+1}. \end{aligned} \tag{4}$$

In this paper, we give new meanings to the quantities in formula (4):

(1) w represents the promotion role of the government in the process of cluster evolution, such as the guidance of relevant policies

(2) X represents the geographical coordinate position of the cluster, here is the position on the coordinate axis

(3) c_1 and c_2, respectively, represent the competition and cooperation of industrial clusters in economic activities

(4) P_i^t and P_g^t, respectively, represent the best cluster positions searched by particle individuals and particle populations

(5) $Z = f(x, y)$ is the objective function, Z is the value of cluster competitiveness, and (x, y) is the coordinate value

5. Experiment and Result Analysis

5.1. D-PSO Performance Verification

5.1.1. Selection of Test Function. In order to test the performance of the D-PSO algorithm, we use three commonly used test functions, Rastrigin function, Rosenbrock function, and Griewank function, to carry out the experimental test. The three test function expressions are as follows:

$$f_1(x) = \frac{1}{4000}\sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n} \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1, f_2(x) = \sum_{i=1}^{n-1}\left(100\left(x_{i+1} - x_i^2\right)^2 + \left(x_i - 1\right)^2\right), f_3(x) = \sum_{i=1}^{n}\left(x_i^2 - 10\cos\left(2\pi x_i\right) + 10\right). \tag{5}$$

The simulation experiment is realized with the help of MATLAB 2019 b. The CPU of PC is i5-7200U@2.50 GHz, and the memory is 8G. The operating system is Window 10(64 bit), and the software is MATLAB 2019 b. In order to test the performance of the D-PSO algorithm, we choose the standard PSO algorithm for the comparative test. Two algorithms are used to solve the above three test functions, respectively. We set the evolution algebra as 500, the inertia weight w of the standard PSO as 0.729, the learning factor $c_1 = c_2 = 1.429$, the total number of particles in the population as 100, and the dimensions of the three main test functions as 10.

5.1.2. Analysis of Convergence Results. When testing the performance of the D-PSO algorithm, it is necessary to exclude the influence of random factors on the experiment. Therefore, we have repeated experiments for each algorithm for many times, where the number of times is 10. The average of 10 independent repeated experiments is taken as the final result of the experiment. Griewank function, Rosenbrock function, and Rastrigin function are solved by two algorithms, respectively. The evolution curve when solving the function is shown in Figure 3.

According to the above experiments, the optimal fitness logarithm values of Griewank function, Rosenbrock function, and Rastrigin function corresponding to D-PSO are −0.13, 0.92, and 0.83, respectively. When the standard PSO algorithm is used to solve the above three functions, the logarithmic values of the optimal fitness are 3.19, 1.42, and 1.05 in turn, and the results of the two algorithms are shown in Table 1.

According to the trend of solving curves, the D-PSO algorithm has achieved good convergence effect within 50 iterations and found the optimal solution. It can be seen that the D-PSO algorithm converges faster. Therefore, compared with the standard PSO algorithm, the D-PSO algorithm has faster convergence speed and higher solution accuracy.

5.2. Example Analysis of Sports Industry Cluster

5.2.1. Evaluation of Cluster Competitiveness. According to the AHP method, the weight distribution table of the competitiveness evaluation index of sports industry cluster can be obtained, as shown in Table 2.

According to the competitiveness values, the competitiveness fitting topographic maps of sports industry clusters in different regions are established, as shown in Figure 4. The larger the value of z in the graph, that is, the higher the protruding part, the greater the cluster competitiveness of this coordinate. If the coordinates of a certain area are (x, y), the function $Z = f(x, y)$ is a three-dimensional fitting function of geographical coordinates and competitiveness. In order to interpolate the three-dimensional function, the fitting function Z based on 1stopt is adopted. The main parameter values of the fitting function are shown in Table 3.

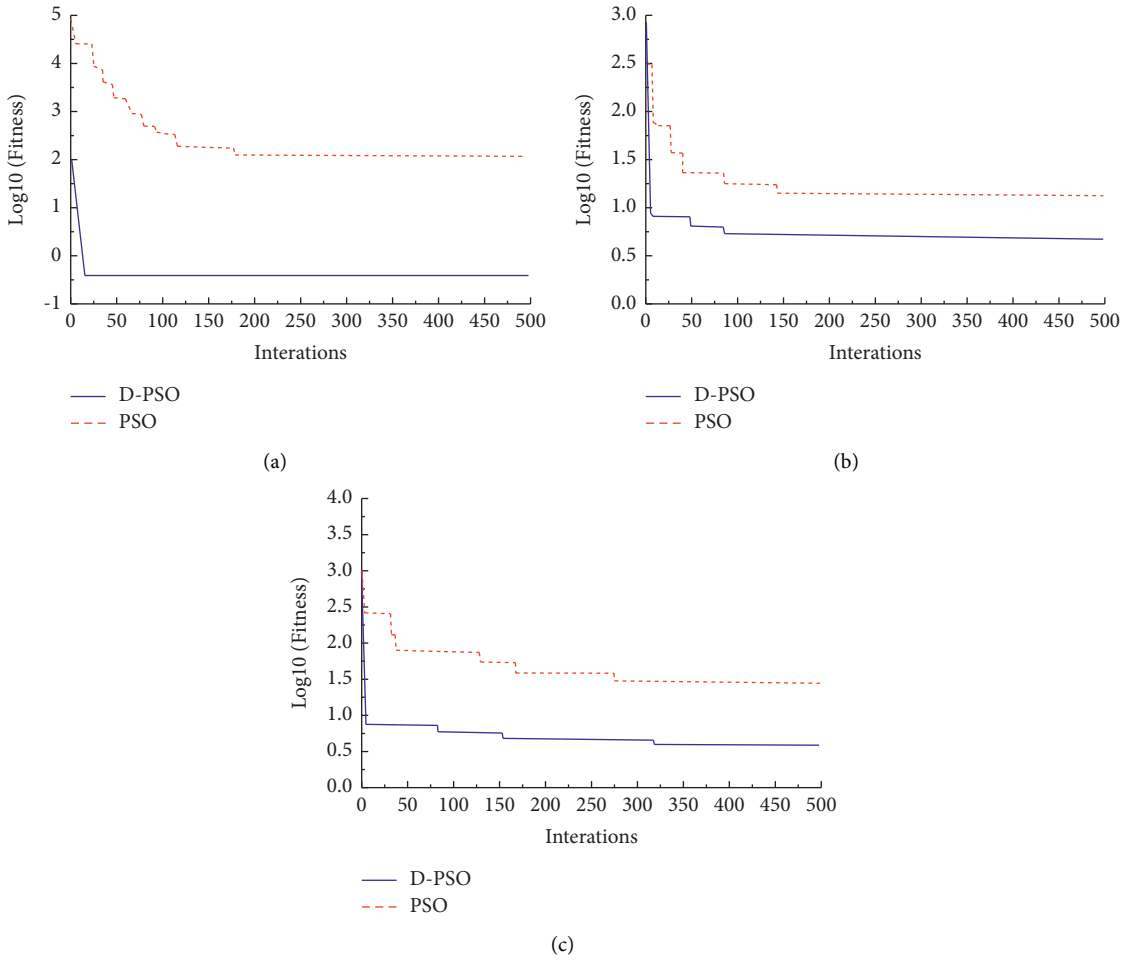


Figure 3: Evolutionary curve when solving function: (a) Griewank; (b) Rosenbrock; (c) Rastrigin.

Table 1: Average fitness of two algorithms.

Function	D-PSO-fitness	PSO-fitness
Griewank	1.53	136.07
Rosenbrock	7.64	23.52
Rastrigin	6.28	10.41

Table 2: Weight distribution table of the competitiveness evaluation index.

Criterion layer	Index layer	Index weight	
Competitiveness of sports industry cluster	Cluster scale	Number of enterprises (units)	0.0391
		Total industrial output value (ten thousand yuan)	0.0966
		Sales revenue (ten thousand yuan)	0.0966
		Total investment (ten thousand yuan)	0.1744
		Total investment (ten thousand yuan)	0.0326
		Total assets at the end of the year (ten thousand yuan)	0.0966
		Main equipment ownership (units)	0.0241
		Production area (100 m^2)	0.0241
	Economic benefits	Sales profit (ten thousand yuan)	0.0966
		Total net profit (ten thousand yuan)	0.1866
		Industrial added value (ten thousand yuan)	0.1744

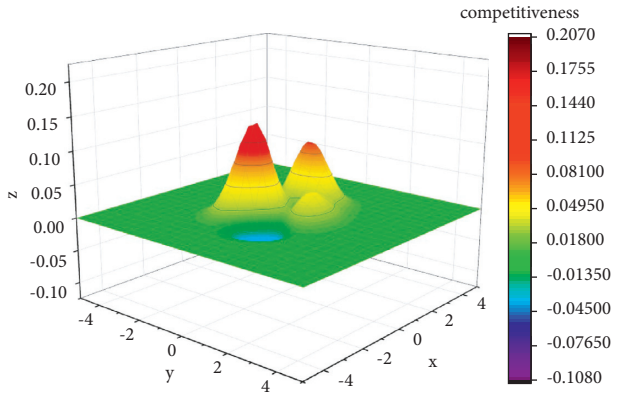


Figure 4: Fitting topographic map of competitiveness of sports industrial clusters in different regions.

Table 3: Main parameter values of fitting function Z.

Parameter	Value
P_1	−0.0001808310780035
P_2	0.00097865676360172
P_3	−0.00012545761310271
P_4	0.000312243741375035
P_5	0.073322666527124324
P_6	−0.0301464618338314
P_7	−0.16584377332312
P_8	1.82335544721859
P_9	0.731132146105326

$$Z = \frac{\left(p_1 + p_2 \ln(x) + p_3 (\ln(x))^2 + p_4 y\right)}{1 + p_5 \ln(x) + p_6 (\ln(x))^2 + p_7 (\ln(x))^3 + p_8 y + p_9 y^2}. \tag{6}$$

5.2.2. D-PSO Solution and Analysis. With the help of MATLAB 2019b, the proposed D-PSO algorithm is used to solve the objective function (formula (4)). In the experiment, the evolution algebra is set to 500 times, the number of particles is 100, and the learning factor $c_1 = c_2 = 1.429$ and $w = 0.729$. The result is (−2.3, −1.1). The relative city is a provincial capital city in China, and it is very close to the Yangtze River Delta. The forecast shows that with the development and upgrading of the industry, a super-large-scale sports industry cluster will appear in (−2.3, −1.1) several years later.

Therefore, appropriate suggestions for cluster development can be given according to the predicted coordinates. (1) From the government's point of view, the government plays a role in promoting and assisting the cluster development. If the local government appropriately adjusts various policies according to the trend of cluster development and continuously improves the symbiotic environment within the cluster, it will continuously promote the upgrading of the cluster and effectively prevent the cluster from declining. (2) At the enterprise level, enterprises are the most important members in the cluster and the largest number. The overall development of clusters is good, which will inevitably promote the development of internal enterprises. Similarly, the good operation of internal enterprises will inevitably promote the development of the whole cluster, which is a virtuous circle.

6. Conclusions

Because particle swarm optimization is interlinked with the agglomeration of industrial clusters, this paper proposes to use the particle swarm optimization algorithm to predict the formation position of industrial clusters. An improved PSO algorithm-diaphragm particle swarm optimization (D-PSO) is proposed to accurately solve the evolution and formation mechanism of industrial clusters. The competitiveness value of industrial clusters is the value of the objective function solved in the PSO algorithm. The geographical coordinates of industrial clusters are the positions in particle search space in the PSO algorithm. Finally, taking the sports industrial cluster as an example, the D-PSO algorithm is used to simulate the aggregation process of enterprises in the cluster, and the development of the industrial cluster is predicted. We will try to use other swarm intelligence algorithms to simulate the aggregation process of enterprises in the cluster such as artificial fish swarm optimization algorithm.

Data Availability

The experimental data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest to report regarding the present study.

Acknowledgments

This work was supported by Open Fund Project of Jiangxi Drama Resource Research Center, East China University of Technology: Research on the amalgamation of culture tourism with the cultural heritage of opera and the tourism branding of the Chinese opera capital of Fuzhou (21XJ04); East China University of Technology Ph.D. Startup Foundation: Research on the constraints of recreational sports and leisure of urban residents in the context of "Healthy China" and their willingness to continue to participate (DHBK2019401).

References

[1] Y.-L. Lai, M.-S. Hsu, F.-J. Lin, C. Yi-Min, and L. Yi-Hsin, "The effects of industry cluster knowledge management on innovation performance," *Journal of Business Research*, vol. 67, no. 5, pp. 734–739, 2014.

[2] E. Bhaskaran, "Technical efficiency of automotive industry cluster in Chennai," *Journal of the Institution of Engineers*, vol. 93, no. 3, pp. 243–249, 2012.

[3] T. Zhou and J. S. Liu, "The distribution and developmental patterns of agricultural products processing industry cluster

in Jilin province," *Scientia Geographica Sinica*, vol. 33, no. 7, pp. 815–823, 2013.

[4] Z. Xia, X. Wang, L. Zhang, Z. Qin, X, Sun, and K, Ren, "A privacy-preserving and copy-deterrence content-based image retrieval scheme in cloud computing," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 11, pp. 2594–2608, 2017.

[5] P. Wang and Q. Liu, "Game analysis on generic technology diffusion process in industry cluster," *International Business and Management*, vol. 4, no. 2, pp. 2868–2877, 2012.

[6] I. D. Hernandez, F. Alemàn, and J. Taborda, ""Parquesoft": a study of social entrepreneurship in software industry cluster in Cali, Colombia," *Revista Facultad De Ciencias Económicas*, vol. 14, no. 2, pp. 11–20, 2015.

[7] J. Wu, "SWOT analysis and research on the informatization application model of the agricultural industry cluster," *Applied Mechanics and Materials*, vol. 571, pp. 1105–1109, 2014.

[8] X. Chen, "Ten years' development & achievements the tenth anniversary of national textile industry cluster pilot work conference was held in Beijing," *China Textile*, vol. 18, no. 1, pp. 46-47, 2013.

[9] J. Zeng and N. Ying, "Panel data identification of convergent and divergent power in regional furniture industry cluster: empirical analysis on second natural factor," *World Forestry Research*, vol. 9, no. 5, pp. 489–495, 2012.

[10] L. Kiminami and S. Furuzawa, "Theoretical and empirical study on regional and local innovation::focusing on the health-related industry cluster in Niigata, Japan," *Studies in Regional Science*, vol. 44, no. 4, pp. 495–515, 2014.

[11] A. Verdu, J. M. Gómez-Gras, and J. Martínez-Mateo, "Value creation through production offshore–inshore strategies in a footwear industry cluster: a coevolutionary perspective," *International Business Review*, vol. 21, no. 3, pp. 342–356, 2012.

[12] F. Røyne, J. Berlin, and E. Ringström, "Life cycle perspective in environmental strategy development on the industry cluster level: a case study of five chemical companies," *Journal of Cleaner Production*, vol. 86, no. 6, pp. 125–131, 2015.

[13] A. R. Jordehi, "Particle swarm optimisation for discrete optimisation problems: a review," *Artificial Intelligence Review*, vol. 43, no. 2, pp. 243–258, 2015.

[14] Y. Zhang, D. Gong, Y. Hu, and W. Zhang, "Feature selection algorithm based on bare bones particle swarm optimization," *Neurocomputing*, vol. 148, pp. 150–157, 2015.

[15] G. J. Osório, J. C. O. Matias, and J. P. S. Catalão, "Short-term wind power forecasting using adaptive neuro-fuzzy inference system combined with evolutionary particle swarm optimization, wavelet transform and mutual information," *Renewable Energy*, vol. 75, no. 5, pp. 301–307, 2015.

[16] M. R. Bonyadi and Z. Michalewicz, "Particle swarm optimization for single objective continuous space problems: a review," *Evolutionary Computation*, vol. 25, no. 1, pp. 1–54, 2017.

[17] C. H. Ai and H. C. Wu, "Cross-regional corporations and learning effects in a local telecommunications industry cluster of China," *Journal of the Knowledge Economy*, vol. 8, no. 1, pp. 1–19, 2015.

[18] Z. Wang, C. Liu, and K. Mao, "Industry cluster: spatial density and optimal scale," *The Annals of Regional Science*, vol. 49, no. 3, pp. 719–731, 2012.

[19] Z. He, S. Xu, and W. Shen, "Overview of the development of the Chinese Jiangsu coastal wind-power industry cluster," *Renewable and Sustainable Energy Reviews*, vol. 15, no. 6, pp. 580–587, 2016.

[20] Z. Tao, "Operation mechanism of industrial cluster: based on the synergetios theory," *Lecture Notes in Electrical Engineering*, vol. 163, pp. 127–135, 2013.

[21] W. Qi, "The analysis of network structure model and innovation network characteristics for industrial cluster: using complicated network perspective," *Journal of Digital Information Management*, vol. 11, no. 4, pp. 277–283, 2013.

[22] Y. Xun, L. V. Yuping, and Z. S. Univ, "Transfer of sporting goods industry cluster theory in uiew of global value chain," *Journal of Wuhan Institute of Physical Education*, vol. 23, no. 21, Article ID 27558, 2015.

[23] D. Han, K. Doya, and J. Tani, "Self-organization of action hierarchy and compositionality by reinforcement learning with recurrent neural networks," *Neural Networks*, vol. 129, pp. 149–162, 2020.

[24] Q. Ma, S. Zhang, and Y. Qiao, "Traffic guidance self-organization method for neighbor weaving segments based on self-organized critical state," *IEEE Access*, vol. 8, Article ID 171784, 2020.

[25] J. Li, J. Zhang, C. Jiang, and M. Zhou, "Composite particle swarm optimizer with historical memory for function optimization," *IEEE Transactions on Cybernetics*, vol. 45, no. 10, pp. 2350–2363, 2015.

[26] P. Li, D. Xu, Z. Zhou, W.-J. Lee, and B. Zhao, "Stochastic optimal operation of microgrid based on chaotic binary particle swarm optimization," *IEEE Transactions on Smart Grid*, vol. 7, no. 1, pp. 66–73, 2016.

[27] H. Duan, L. Pei, and Y. Yu, "A predator-prey particle swarm optimization approach to multiple UCAV air combat modeled by dynamic game theory," *IEEE/CAA Journal of Automatica Sinica*, vol. 2, no. 1, pp. 11–18, 2015.

[28] M. H. Aghdam and S. Heidari, "Feature selection using particle swarm optimization in text categorization," *Journal of Artificial Intelligence and Soft Computing Research*, vol. 5, no. 4, pp. 38–43, 2015.

[29] T. Xu, L. Zhao, Z. Jiang, and X. Guo, "A high sensitive pressure sensor with the novel bossed diaphragm combined with peninsula-island structure," *Sensors & Actuators A Physical*, vol. 244, pp. 66–76, 2016.

[30] S. Maihemuti, W. Wang, and H. Wang, "Voltage security operation region calculation based on improved particle swarm optimization and recursive least square hybrid algorithm," *Journal of Modern Power Systems and Clean Energy*, vol. 9, no. 1, pp. 138–147, 2021.

Hindawi
Scientific Programming
Volume 2021, Article ID 5862037, 10 pages
https://doi.org/10.1155/2021/5862037

Research Article

Research on the Basketball Goal Recognition Method Based on Improved MobileNet

Ke Yang

Physical Education Department, Xihua University, Chengdu, Sichuan, China

Correspondence should be addressed to Ke Yang; 0120030089@mail.xhu.edu.cn

Received 9 November 2021; Revised 22 November 2021; Accepted 30 November 2021; Published 26 December 2021

Academic Editor: Baiyuan Ding

Moving target detection is involved in many engineering projects, but it is difficult because of the strong time-varying speed and uncertain path. Goal recognition is the key technology of the basketball goal automatic test. Also, accurate and timely judgment of basketball goals has important practical value. Therefore, a basketball goal recognition method based on an improved lightweight deep learning network model (L-MobileNet) is proposed. First of all, the basket detection is carried out by the Hough circle transform algorithm. Then, in order to further improve the detection speed of basketball goals, based on the lightweight network MobileNet, an improved lightweight network (L-MobileNet) is proposed. First of all, for deeply separable convolution, channel compression and block convolution reduce the parameters and computational complexity of the module. At the same time, because block convolution will hinder the information exchange between characteristic channels, an improved channel shuffling method, IShuffle, is introduced. Then, combined with the residual structure to improve the generalization ability of the network, the RLDWS module is constructed. Finally, a more lightweight network L-MobileNet is constructed by using the RLDWS module. The experimental results show that the proposed method can effectively realize the judgment of basketball goals, and the judgment accuracy is improved by 8.35%. At the same time, the amount of parameters and computation is only 29.7% and 53.2% of the original, and it also has certain advantages compared with other lightweight networks.

1. Introduction

The NBA (National Basketball Association) and CBA (China Basketball Association) are popular sports in current ball games. Through the understanding of the NBA and CBA, it is found that they use artificial and intelligent devices to realize timing and scoring. The method of combining artificial and intelligent devices is used to realize timing and scoring. There is a camera on the backboard of the NBA, which automatically takes photos whenever players are under the basket or on the layup. When the basketball falls from the basket, it will touch the net, which will drive the sensor. The controller will receive the goal information, and the referee will update the score and the time. The competition between the NBA and CBA is dominated by manpower and supplemented by equipment.

At present, the physical education professional examination includes special items and supplementary items. Special items and auxiliary items should be completed in a short time, and there are many examination items. Therefore, the task for invigilators is very heavy. However, the fairness and justice of the college entrance examination must be observed by every examinee and invigilator. At present, many places still use some manual methods to score basketball exams, so too much workload will lead to unfairness. This method is now slowly being replaced by some smart devices. Nowadays, infrared detection [1–5] is used in many exams to judge basketball goals. Through the two intelligent detection methods of infrared detection and microswitch, although some problems existing in traditional manual work have been effectively solved, these two methods also have some shortcomings, such as easy damage and high maintenance cost. Then, finding a better alternative method is an urgent research work at present.

With the rapid development of image processing technology, the detection technology of moving objects in a

video has been more widely used [6–10]. With the rapid development of the same era, more and more requirements for video processing applications have been put forward. For example, the fixed-point shooting test is conducted in the college entrance examination for physical education. Moving object detection technology is a particularly important branch of vision technology. Because of the strong scientific research value of moving target detection technology, researchers have devoted a lot of energy to research it and achieved good research results.

2. Target Detection Analysis Based on the Deep Learning Network

At present, the target detection methods based on deep learning are mainly divided into two categories [11–14]: the two-stage detection framework and single-stage detection framework. The two-stage detection framework mainly divides the detection task into two stages. Firstly, the candidate region is generated, and then, the region is regressed and classified by using the deep network model. This class detection algorithm mainly has high detection accuracy, but the detection speed is slow. Girshick et al. [15] put forward the application of RCNN in the field of target detection and improved the map of the algorithm to 53.3% in the Pascal VOC 2012 dataset [16], which is far superior to the traditional target detection algorithm. However, there are also some problems such as tedious training steps, slow detection speed, and the need to input fixed-size images. He et al. [17] proposed an SPPNet algorithm to solve the problem that the fixed-size image must be input to extract features by the CNN. The innovation of this algorithm is that a Spatial Pyramid Pooling (SPP) layer is added between the convolution layer and fully connected layer.

The single-stage detection framework can classify and regress the targets in the image at the same time, without the operation of generating candidate regions. This class detection algorithm has a fast detection speed, but the detection accuracy is usually not as good as that of the two-stage detection algorithm. Redmon et al. [18] proposed the classic YOLO algorithm. In 2016, Liu et al. [19] proposed the SSD algorithm. This algorithm combines the advantages of YOLO's fast speed and RPN's accurate positioning [20] so as to achieve the effect of detecting the target at different scales. Compared with YOLO, the SSD algorithm can predict more candidate regions, and the detection effect is better, but the disadvantage is that the speed is slower than that of YOLO. The lightweight network model MobileNet [21] proposed by Google focuses on devices with limited resources, such as mobile or embedded devices, to maximize classification accuracy. The main innovation of this network lies in the proposal of the Depth-Wise Convolution (DWC) module to reduce the parameters and computation and the effective compromise between classification accuracy and speed by using two superparameters of width multiplier and resolution multiplier.

Therefore, in order to better solve the problems of easy damage, high cost, and misjudgment in traditional fixed-point shooting devices, this paper proposes a basketball goal recognition method based on image analysis and uses deep learning technology to solve the abovementioned problems. Firstly, the problems of the existing lightweight network model MobileNet are analyzed theoretically, and the improved strategies are put forward to solve these problems, and the RLDWS module is gradually constructed. Then, the improved L-MobileNet model is constructed by using this module instead of the original deep separable convolution. Finally, a comparative experiment with other lightweight networks is carried out to further verify the rationality and effectiveness of the proposed improved network.

3. Basketball Goal Recognition Based on the Improved MobileNet Model

3.1. Detection of Basket. The detection of the basket is that the original color image is processed into a gray image, the gray image is subjected to median filtering and mathematical morphology processing, and then, the Hough circle transformation algorithm [22] is used to extract the basket. Finally, the extracted basket circle is added to the original color image, and the position and size of the basket are marked. The essence of the Hough circle transformation is to transform the coordinates of the image and to transform the plane coordinates with the parameter coordinates so that the transformed results are easier to identify and detect. The general equation of a circle is as follows:

$$(x-a)^2 + (y-b)^2 = r^2, \tag{1}$$

where (a, b) is the center of the circle and r is the radius of the circle.

When the circle on the X-Y plane in the image space is transformed into the a-b-r parameter space, a three-dimensional cone will be formed in the parameter space corresponding to the circle containing (x, y) points. The Hough transform principle is shown in Figure 1. The parameters of the circle can be obtained from the detected point so as to determine the circle. The parameter image of the circle is shown in Figure 2.

3.2. Problem Analysis of the MobileNet Model. MobileNet has the following three problems [23–25]:

(1) 1×1 convolution has a large amount of computation: by deducing the computation and parameters of the network structure, it is found from the perspective of the layer type that the computation and parameters are mainly concentrated on 1×1 point-by-point convolution operation, in which the computation accounts for about 95% of the whole network and the parameters account for 75%, as shown in Table 1.

In the depth separable convolution operation, the calculation amount N_{DW} of the depth convolution and the calculation amount N_{PW} of the point-by-point convolution are shown in equations (2) and (3), respectively.

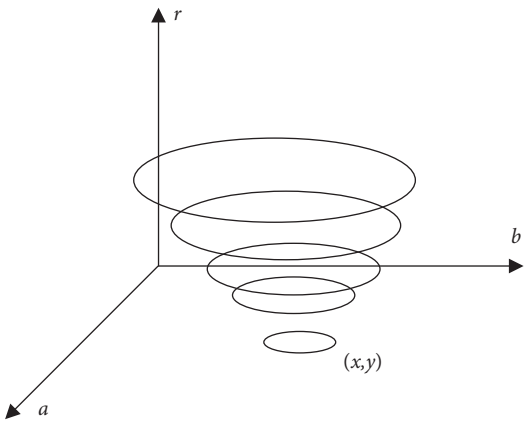


Figure 1: Hough transform principle.

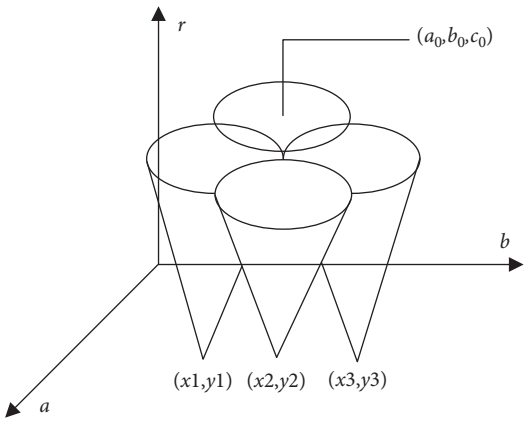


Figure 2: Parametric image of circle.

Table 1: Computing resources at different levels.

Layer Type	Mult-Adds (%)	Parameters (%)
Conv 1 × 1	94.86	74.59
Conv DW 3 × 3	3.06	1.06
Conv 3 × 3	1.19	0.02
Fully Connected	0.18	24.33

$$N_{DW} = H \times W \times D_K \times D_K \times M, \quad (2)$$

$$N_{PW} = H \times W \times M \times N, \quad (3)$$

where D_K is the size of the convolution kernel, $H \times W$ is the size of the input feature, M is the dimension of the input feature, and N is the dimension of the output feature. It can be seen that the calculation amount of point-by-point convolution is positively correlated with N, and the value of N will gradually increase with the deepening of network layers, resulting in an increase in the proportion of calculation amount of point-by-point convolution operation. Subsequently, this problem is mainly improved by improving the depth separable convolution structure.

(2) Low-dimensional data collapse caused by ReLU: when low-dimensional (n-dimensional) data are mapped to high-dimensional (m-dimensional) by the random matrix for ReLU operation and then mapped back to this dimension by the generalized inverse matrix, some information will be lost, and the smaller the m is, the more the information will be lost.

To solve this problem, in the subsequent module design, it is considered to use the Mish activation function [26, 27] instead of ReLU after feature mapping with few channels; otherwise, the information will be destroyed.

(3) No reuse feature: MobileNet is a very simple straight-cylindrical structure. In the training process of the network model, if the weight of a convolution node becomes 0, the output of the node will be 0 for any input. However, the gradient of ReLU operation to 0 value is 0, so the value of this node will not be recovered no matter how much the iteration is, and the residual module will be added to improve it later.

3.3. Improved MobileNet Model. In response to problems 1 and 2, in order to maximize the use of packet convolution, we modified the improved channel shuffle (IShuffle), as shown in Figure 3.

First of all, it is still uniform recombination of different groups of features, but there is a group of recombined features that are different, and this group of features is obtained by merging and combining the recombined features of each group, respectively. Specifically, it is assumed that the number of input features m is 9 and the number of scores is 3. The first six features are still the same as those in channel shuffling. The remaining three features are obtained by intergroup feature fusion of these nine feature channels. The last six feature channels are spliced with these three features in the channel dimension, and the final output features are obtained. As can be seen from Figure 3, after the uniform recombination of features, IShuffle fused features between groups to improve the information exchange between groups.

In addition, because ReLU is prone to data collapse in low dimensions, we consider changing ReLU nonlinearity to the Mish activation function after compressing the 1 × 1 convolution kernel of dimension, and its expression and graph are shown in Figure 4.

Compared with ReLU, the Mish activation function has better smoothness. When the value is negative, a smaller gradient flow is allowed, which makes the information better penetrate into the network, thus improving the accuracy and generalization ability of the network while still being borderless.

Because the block convolution reduces the information flow between channels, the modified IShuffle strategy is subsequently added. Finally, the outputs from the two parts are spliced to obtain the final output features. As shown in Figure 5, the module is named the "LDWS module."

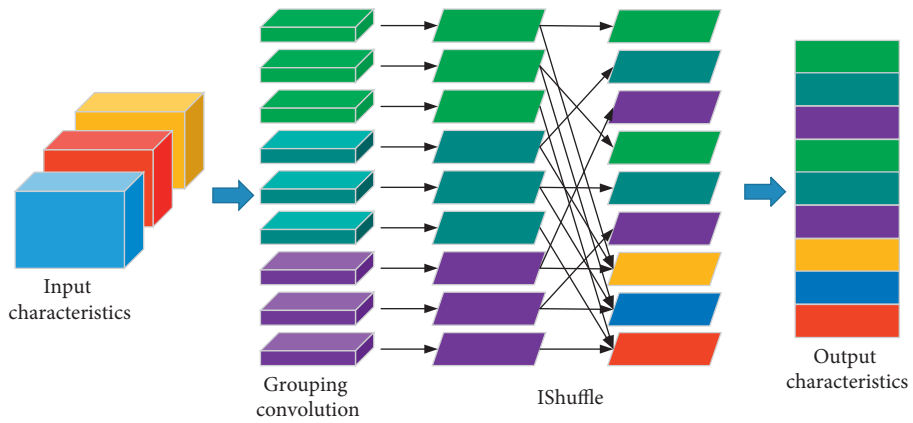


Figure 3: Modified IShuffle operation.

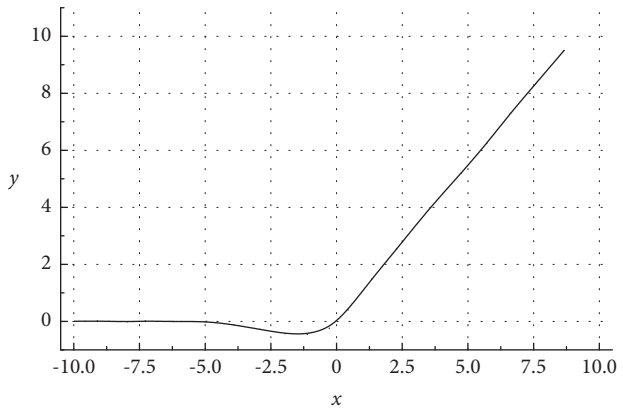


Figure 4: Mish activation function.

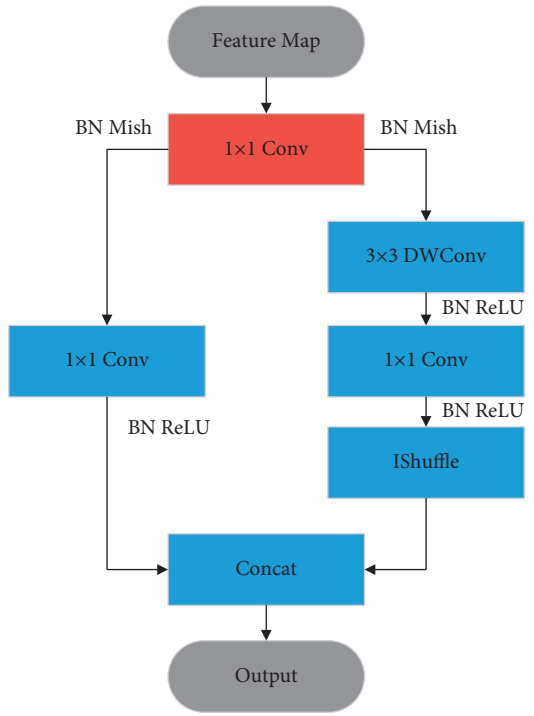


Figure 5: LDWS module.

The LDWS module is mainly composed of a compression layer and expansion layer. The compression factor is t, which represents the dimension reduction ratio of the compression layer, and its calculation is shown in the following equation:

$$t = \frac{s_1}{m}, \tag{4}$$

where m represents the number of input channels of the LDWS module and s_1 represents the number of convolution kernels of the compression layer.

The calculation amount for the LDWS module is shown in the following equation:

$$\begin{aligned} N_{\mathrm{LDWS}} &= H \times W \times M \times tM \\ &\quad + H \times W \times tM \times e_1 + H \times W \times tM \times D_K \times D_K \\ &\quad + \frac{1}{x} \times H \times W \times tM \times e_2 + N, \end{aligned} \tag{5}$$

where e_1 represents the convolution kernel number of the 1×1 part in the expansion layer and e_2 represents the convolution kernel number of the depth separable convolution part in the expansion layer.

The comparison between the calculation amount of the LDWS module and that of depth separable convolution is as follows:

$$\frac{N_{\mathrm{LDWS}}}{N_{\mathrm{DWS}}} = \frac{t \times \left(D_K^2 + M + 4 \times tM + 4/g_1 \times tM\right) + N}{D_K^2 + 8 \times tM}, \tag{6}$$

where g_1 represents the number of packet groups in point-by-point convolution. It can be seen that by reducing the compression factor t and increasing the number of packets in packet convolution, the amount of parameters and computation can be greatly reduced. In the experiment, the value of t is 0.125.

In view of problem 3, the RLDWS module is designed by introducing residual connection based on the LDWS module, as shown in Figure 6, in which the blue and red parts are LDWS1 modules, and the purple part is an improved residual structure. Before ResNet appeared, in order to improve the recognition accuracy of the neural network model, a deeper network was often built by simply stacking layers. However, due to the back propagation process of the gradient, the deeper network may make the parameters of the shallow layer unable to be updated, and the gradient disappears, thus leading to the saturation or even decline of the network performance. Therefore, a residual connection structure is proposed in ResNet.

Using the RLDWS module and the LDWS module to replace some DWS modules in MobileNet and increase the number of RLDWS modules, the final structure of L-MobileNet is constructed, as shown in Table 2, where s is the step size in deep convolution, c is the number of output channels, and k is the number of final categories.

L-MobileNet is mainly composed of one standard convolution, five DWS modules, and nine RLDWS modules. Like MobileNet, in L-MobileNet, the standard convolution operation of 3×3 is the first step, followed by five DWS modules. Then, the remaining eight DWS modules are replaced with the improved RLDWS module in this paper, and an additional RLDWS module is added after the last RLDWS module. Finally, through the average pooling and full connection layer of 7×7 sizes and multiclassification with softmax, the output of the network is obtained.

4. Experiment and Result Analysis

4.1. System Construction. The basketball goal recognition system is mainly built by hardware and software environment. The hardware environment mainly provides video image data and running environment for the system, mainly including cameras and PCs. The software mainly processes video image data. The model of the camera is M30A, and the frame rate is 60 frames per second when the resolution is 640∗480. The image data format is YUV422. The data transmission protocol is USB2.0. The CPU of the PC is i5-7200U@2.50 GHz, the memory is 8G, and the graphics card is GTX 2060s. The operating system is Windows 10 (64-bit), and the software is MATLAB 2019 b.

FLOPs are used to measure the complexity of the model, specifically the number of multiply-add. The smaller the value of this index, the less the amount of calculation required by the model, that is, the faster the speed.

4.2. Selection of Grouping Quantity. FLOPs can be greatly reduced by using block convolution operation. However, the fact that FLOPs do not increase does not mean that the speed becomes faster. Grouping too much will increase the memory access consumption, and it will also reduce the speed. Therefore, it is necessary to weigh the selection of the number of packets through experiments. Table 3 shows the comparison of the effects of setting different packet convolutions on the network model on L-MobileNet, and g represents the number of packets.

With the number of packets increasing from 1 to 16, the number of parameters and calculation of the network also decreases, and the judgment accuracy first increases and then decreases. Therefore, it is necessary to make a certain tradeoff between the judgment accuracy and the speed, so the convolution number of the grouping is selected as 4 to continue the subsequent experiments.

4.3. Verification of Residual Structure. Table 4 shows the LDWS results of the influence of the residual structure on the judgment accuracy and speed of the network. When the residual structure is added, the error of the network is reduced by 6.84%, respectively.

4.4. Analysis of the Experimental Process. The basket test results are shown in Figure 7. In addition, through the comprehensive analysis of three simulated test videos collected from the left, middle, and right shots, the

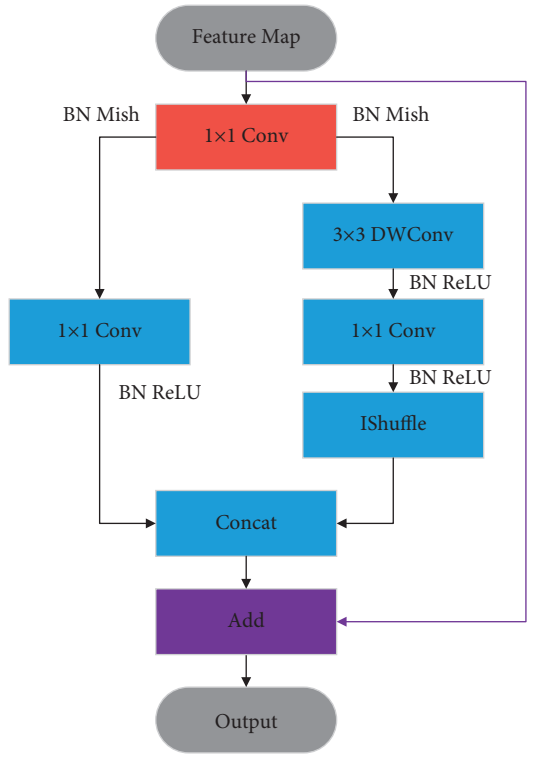


Figure 6: RLDWS module.

Table 2: Network structure of L-MobileNet

Input dimensions	Module	c	n	s
$224 \times 224 \times 3$	3×3 conv	32	1	2
$112 \times 112 \times 32$	DWS	64	1	1
$112 \times 112 \times 64$	DWS	128	2	2
$56 \times 56 \times 128$	DWS	256	2	2
$28 \times 28 \times 256$	RLDWS	512	6	2
$14 \times 14 \times 512$	RLDWS	1024	3	2
$7 \times 7 \times 1024$	Avg pool	1024	1	1
$1 \times 1 \times 1024$	FC	1000	1	1
$1 \times 1 \times 1000$	Softmax	k	1	1

Table 3: Comparison of group convolution quantity and effect.

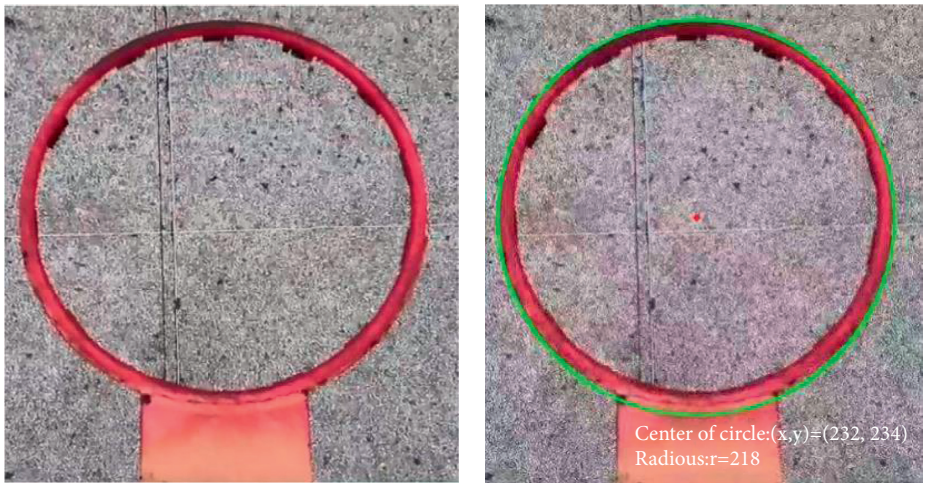
Number of packets	Error (%)	Parameter (MB)	GFLOPS
$g = 1$	30.66	1.240	1.299
$g = 2$	31.35	1.113	1.248
$g = 4$	30.11	1.050	1.222
$g = 8$	30.97	1.018	1.209
$g = 16$	30.87	1.002	1.203

Table 4: Comparison of the residual structure effect.

	Error (%)	Parameter (MB)	GFLOPS
Without residual	36.95	1.050	1.222
With residual (ours)	30.11	1.050	1.222

basketball was detected from the video sequence frame images. The test sequence of shooting the video is shown in Figure 8.

After testing 30 groups of data (10 groups of data on the left, middle, and right), the test results are as shown in Tables 5–7, respectively.



(a) (b)

Figure 7: Basket detection image (a) Original image. (b) Basket detection result.

(a) (b)

Figure 8: Basketball goal detection image. (a) 89-th detection image; (b) 147-th detection image.

Table 5: Left shot statistics table.

No.	Left layup times	Actual goal	Judgement goal
1	22	8	8
2	19	11	11
3	17	10	10
4	14	8	8
5	19	9	9
6	22	12	12
7	17	10	10
8	17	7	7
9	23	8	8
10	18	9	9

In this experiment, the performance of L-MobileNet is compared with the commonly used lightweight models MobileNet, MobileNet V2 [28], SqueezeNet [29], and ShuffleNet [30]. Table 8 shows the performance comparison of each network model.

The L-MobileNet model is better than the other network models in every index. Compared with MobileNet, the judgment accuracy of L-MobileNet is improved by 8.35%, and the amount of parameters and calculations is only 29.7% and 53.2% of the original ones. SqueezeNet greatly reduces the number of parameters due to the compression of parameters, but the amount of calculation is still very large. L-MobileNet is superior to SqueezeNet except for a few more parameters. Compared with ShuffleNet, L-MobileNet's judgment accuracy and parameter quantity are almost the same, but the amount of calculation is almost half that of ShuffleNet.

The change curve of judgment accuracy of each network model during dataset training is shown in Figure 9.

Table 6: Middle shot statistics table.

No.	Left layup times	Actual goal	Judgement goal
1	18	7	7
2	19	11	11
3	21	12	11
4	18	12	12
5	15	9	9
6	13	6	6
7	19	10	10
8	20	12	12
9	21	13	13
10	15	12	12

Table 7: Right shot statistics table.

No.	Left layup times	Actual goal	Judgement goal
1	13	6	6
2	16	10	10
3	17	11	11
4	12	8	8
5	18	10	10
6	22	13	13
7	23	14	14
8	11	7	7
9	19	10	10
10	17	12	12

Table 8: Performance comparison of various network models.

Network models	Error (%)	Parameter (MB)	GFLOPS
L-MobileNet (ours)	30.11	1.050	1.222
MobileNet	33.94	3.315	2.322
MobileNet V2	32.66	2.369	2.397
SqueezeNet	31.07	0.781	2.664
ShuffleNet	29.68	1.012	2.174

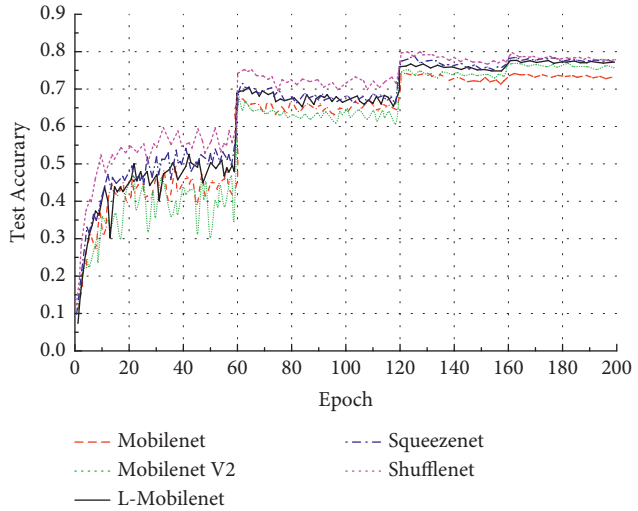


Figure 9: Judgment accuracy of each network model.

All the network models tend to converge after 160 epochs, and the judgment accuracy from low to high is MobileNet, MobileNet V2, L-MobileNet, SqueezeNet, and ShuffleNet. L-MobileNet has higher judgment accuracy than MobileNet and MobileNet V2. On the other hand, the training time of L-MobileNet is less than that of the other models, which is due to the advantages of lighter models.

5. Conclusions

In this paper, an improved lightweight neural network model (L-MobileNet) based on MobileNet is proposed and applied to basketball goal recognition. Through the construction experiment of the improved network model, the improved network with a compromise between accuracy and speed is selected. Compared with MobileNet, the judgment accuracy of the improved network is increased by 8.35%, while the amount of parameters and calculations are only 29.7% and 53.2% of the original ones. Finally, the improved network is compared with some commonly used lightweight network models, and it is concluded that the improved network has a good effect on the accuracy and speed of goal judgment. Then, the improved RLDWS module is tried to combine with VGGNet so as to further reduce the complexity of the model. At the same time, neural structure search will try to use data-driven and intelligent methods to automatically build a better network.

Data Availability

The experimental data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares no conflicts of interest regarding the present study.

References

[1] A. Chatterjee, A. Abhale, N. Pendyala, and K. S. R. K. Rao, "Group II-VI semiconductor quantum dot heterojunction photodiode for mid wave infrared detection," *Optoelectronics Letters*, vol. 16, no. 4, pp. 290–292, 2020.

[2] L. V. Bao-Lin, S. F. Tong, and X. U. Wei, "Research on non-uniformity correction of airborne infrared detection system," *Chinese Optics*, vol. 13, no. 5, pp. 1–14, 2020.

[3] G. Amato, F. Fa Lchi, and L. Vadicamo, "Image retrieval," *Multimedia Tools and Applications*, vol. 77, no. 5, pp. 5385–5415, 2016.

[4] Z. Xia, X. Wang, and L. Zhang, "A privacy-preserving and copy-deterrence content-based image retrieval scheme in cloud computing," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 11, pp. 2594–2608, 2017.

[5] X.-S. Wei, J.-H. Luo, J. Wu, and Z.-H. Zhou, "Selective convolutional descriptor aggregation for fine-grained image retrieval," *IEEE Transactions on Image Processing*, vol. 26, no. 6, pp. 2868–2881, 2017.

[6] Z. Xia, N. N. Xiong, A. V. Vasilakos, and X. Sun, "EPCBIR: an efficient and privacy-preserving content-based image retrieval scheme in cloud computing," *Information Sciences*, vol. 387, pp. 195–204, 2017.

[7] M. H. Memon, J. P. Li, and I. Memon, "GEO matching regions: multiple regions of interests using content based image retrieval based on relative locations," *Multimedia Tools and Applications*, vol. 76, no. 14, pp. 1–35, 2017.

[8] Z. Zeng, "A novel local structure descriptor for color image retrieval," *Information*, vol. 7, no. 1, p. 9, 2016.

[9] W. Zhou, S. Newsam, and C Li, "Learning low dimensional convolutional neural networks for high-resolution remote sensing image retrieval," *Remote Sensing*, vol. 9, no. 5, p. 489, 2016.

[10] M. Paulin, J. Mairal, and M. Douze, "Convolutional patch representations for image retrieval: an unsupervised approach," *International Journal of Computer Vision*, vol. 121, no. 1, pp. 1–20, 2016.

[11] K. Waldhoer, "Artificial intelligence engines: a tutorial introduction to the mathematics of deep learning," *Computing Reviews*, vol. 61, no. 2, p. 63, 2020.

[12] J. Gao, P. Liu, G.-D. Liu, and L. Zhang, "Robust needle localization and enhancement algorithm for ultrasound by deep learning and beam steering methods," *Journal of Computer Science and Technology*, vol. 36, no. 2, pp. 334–346, 2021.

[13] B. B. Nair, S. Krishnamoorthy, and M. Geetha, "Machine vision based flood monitoring system using deep learning techniques and fuzzy logic on crowdsourced image data," *Intelligent Decision Technologies*, vol. 21, no. 9, pp. 1–14, 2021.

[14] A. S M. Torres, S. M. Marinho, S. D. Dominguez, R. D. Silva, and H. C. Almeida, "Basic food basket monthly price of southern bahia cities: a time series forecasting with deep learning using a recurrent neural network approach," *Revista Cereus*, vol. 12, no. 4, pp. 146–159, 2020.

[15] R. Girshick, J. Donahue, and T. Darrell, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 580–587, Columbus, OH, USA, June 2014.

[16] X. H. Han, Y. W. Chen, and X. Ruan, "Multi-class object recognition by fusion of image descriptors: classification evaluation of PASCAL VOC Challenge database," *Ieice Technical Report*, vol. 109, pp. 103–108, 2009.

[17] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 9, pp. 1904–1916, 2015.

[18] J. Redmon, S. Divvala, and R. Girshick, "You only look once: unified, real-time object detection," in *Proceedings of the Computer Vision & Pattern Recognition*, pp. 779–788, IEEE, Las Vegas, NV, USA, June 2016.

[19] W. Liu, D. Anguelov, D. Erhan et al., "SSD:Single Shot Multibox detector," in *Proceedings of the European Conference on Computer Vision*, pp. 21–37, Amsterdam, The Netherlands, October 2016.

[20] J. A. Lin, C. T. Chiu, and Y. Y. Cheng, "Object detection in RGB-D images via anchor box with multi-reduced region proposal network and multi-pooling," *Journal of Signal Processing Systems*, vol. 93, pp. 1–15, 2021.

[21] I. Darma, N. Suciati, N. Suciati, and D. Siahaan, "Neural style transfer and geometric transformations for data augmentation on Balinese carving recognition using MobileNet," *International Journal of Intelligent Engineering and Systems*, vol. 13, no. 6, pp. 349–363, 2020.

[22] R. O. Duda and P. E. Hart, "Use of the Hough transformation to detect lines and curves in pictures," *Communications of the ACM*, vol. 15, no. 1, pp. 11–15, 1972.

[23] T. Zhao, X. Yi, and Z. Zeng, "Mobile Net-Yolo based wildlife detection model: a case study in yunnan tongbiguan nature reserve, China," *Journal of Intelligent and Fuzzy Systems*, vol. 41, no. 2, pp. 1–11, 2021.

[24] R. Sadik, S. Anwar, and M. L. Reza, "AutismNet: recognition of autism spectrum disorder from facial expressions using MobileNet architecture," *International Journal of Advanced*

Trends in Computer Science and Engineering, vol. 10, no. 1, pp. 327–334, 2021.

[25] M.-H. So, C.-S. Han, and H.-Y. Kim, "Defect classification algorithm of fruits using modified MobileNet," *The Journal of Korean Institute of Information Technology*, vol. 18, no. 7, pp. 81–89, 2020.

[26] G. Renith and A. Senthilselvi, "Accuracy improvement in diabetic retinopathy detection using DLIA," *Journal of Advanced Research in Dynamical and Control Systems*, vol. 24, no. 4, pp. 133–149, 2020.

[27] H. Sun, A. Wang, and W. Wang, "An improved deep residual network prediction model for the early diagnosis of alzheimer's disease," *Sensors*, vol. 21, no. 12, pp. 54–65, 2021.

[28] M. Arora, S. Garg, and A. Srivani, "Face mask detection system using Mobilenetv2," *International Journal of Engineering and Advanced Technology*, vol. 10, no. 4, pp. 127–129, 2021.

[29] B. Sra, B. Ns, and C. Yt, "EEG based multi-class seizure type classification using convolutional neural network and transfer learning," *Neural Networks*, vol. 124, pp. 202–212, 2020.

[30] S.-H. G. Chan and H. Kobayashi, "Packet scheduling algorithms and performance of a buffered shufflenet with deflection routing," *Journal of Lightwave Technology*, vol. 18, no. 4, pp. 490–501, 2000.

Hindawi
Scientific Programming
Volume 2021, Article ID 6052182, 13 pages
https://doi.org/10.1155/2021/6052182

Research Article

An Empirical Study on the Relationship between Education and Economic Development Based on PVAR Model

Zhenzi Sun

Institute of Higher Education, Liaoning University, Shenyang 110136, China

Correspondence should be addressed to Zhenzi Sun; zsun@mail.umhb.edu

Received 9 November 2021; Revised 24 November 2021; Accepted 3 December 2021; Published 24 December 2021

Academic Editor: Baiyuan Ding

In order to effectively analyze the dynamic relationship between education and economic development, an empirical study on the relationship between education and economic development based on the PVAR model is proposed. This article expounds on the principles, assumptions, identification, and estimation methods of the PVAR model, takes the education level and economic development level as the research object, explains the corresponding variables, and selects the indicators. Using Cobb–Douglas production function as a theoretical model, this article analyzes the theoretical relationship between the education level and economic development level. Based on this theoretical relationship, this article makes an empirical analysis on the relationship between education level and economic development level using the PVAR model. The results show that, on the whole, economic development can drive the development of education, but there are obvious regional differences in the impact of economic growth on the development of education. The impact of economic growth in the eastern region on education is significantly higher than that in the central and western regions. There is an interactive relationship between education level and economic development level, and there is a certain incubation period for education level and economic development level to play their role.

1. Introduction

As a major undertaking involving the national economy and the people's livelihood, education affects not only the inheritance of a country's national culture and spiritual wealth but also the sustainable and healthy development of a country's economy and society [1–3]. Economic development is one of the most indispensable links in the development of all countries in the world and one of the most basic parts in the history of human development. Economic development provides power and guarantee for the activities of all countries and other human beings [4–6]. Education and economic development are closely related. Education is the technical means of economic development, and economic development is the driving force and guarantee of educational development. In order to achieve good economic development, a country must improve the quality of its human resources and do a good job in the development of human resources and scientific management, and education is the key to the work [7–9].

The innovation ability of a country's knowledge, technology, and management is determined by the country's education level. In turn, the improvement of economic level urges people to improve their demand for educational activities and improve their value and taste through more educational activities. With the continuous improvement of ability and values, people will put forward higher requirements for the current educational system and form to meet the future learning needs. However, the existing research ignores the research on the direct contribution of education investment or education expenditure to economic growth. Most of the existing quantitative studies focus on the long-term impact and indirect contribution of education to economic growth and follow the idea of education ⟶ improvement of labor quality⟶economic growth. However, the education sector is an important part of the tertiary industry. Theoretically, the investment in education and increasing education expenditure will inevitably directly drive the development of the tertiary industry. However, there is a lack of quantitative research in this field.

Although China has made great achievements in education and economic development, on the whole, China has a vast territory, the development situation and relevant policies are different in various regions, and there are great differences in the education level and economic development in various regions [10–12]. Moreover, the education level in most areas lags behind the local economic development level, which not only does not give full play to the role of education in promoting economic development but also seriously hinders the further development of the local economy [13–15]. Therefore, this article proposes an empirical study on the relationship between education and economic development based on the PVAR model. Firstly, based on the theoretical analysis results of the PVAR model, then the two variables of education level and economic development level are explained and the indicators are selected. Finally, the Cobb–Douglas production function is used to analyze the theoretical relationship between education level and economic development level so as to provide a theoretical basis for empirical analysis.

2. Theoretical Basis of PVAR Model

Holtz Eakin (1987) first used the PVAR model to analyze the interaction between endogenous variables of panel data. He studied the vector autoregressive model of panel data; that is, all variables are regarded as endogenous variables to analyze the relationship between each variable and its lag term. Using panel data, the PVAR model can not only effectively solve the problem of individual heterogeneity but also fully consider individual and time effects.

2.1. Principle of PVAR Model. The PVAR model is a problem of establishing VAR based on panel data [16–18]. The general model of PVAR can be expressed as follows:

$$Y_{i,t} = \tau_i + \sum_{k=1}^{m} \Phi_{l,k} Y_{i,t-k} + \sum_{j=1}^{m} \psi_{l,i} X_{i,t-j} + \gamma_i + u_{i,t}. \quad (1)$$

In formula (1), $Y_{i,t}$ is the $M \times 1$ vector of M observable variables of section individual i at time point t, $X_{i,t}$ is the $M \times 1$ vector of observable deterministic strictly exogenous variables, $\Phi_{l,k}$, $\psi_{l,i}$ is the coefficient matrix to be estimated of $M \times M$, γ_i is the unobservable M individual fixed effect matrix of individual i, and $u_{i,t}$ is the random error term [19–21].

In practical application, it is often the case that the coefficient matrix to be estimated of the lag endogenous and exogenous variables is not time-varying; that is,

$$Y_{i,t} = \tau_i + \sum_{k=1}^{m} \Phi_k Y_{i,t-k} + \sum_{j=1}^{m} \psi_j X_{i,t-j} + \gamma_i + u_{i,t}. \quad (2)$$

2.2. PVAR Model Assumptions.

Hypothesis 1: for any number of individuals N and period length T, $Y_{1,t}, Y_{2,t}, \ldots, Y_{N,T}$ is an observable variable.

Hypothesis 2: for any $i = 1, \ldots, N$, $t = 1, \ldots, T$, $u_{i,t}$ is an independent and identically distributed random variable whose random error term satisfies zero expectations and the covariance matrix is Ω; that is, $u_{i,t} \sim i.i.d(0, \Omega)$.

Hypothesis 3: when $s < t$, $Y_{i,t}$, $X_{i,t}$, and γ_i are orthogonal to the random error term; that is,

$$\begin{aligned} E\left[Y_{i,s}\right] &= E\left[X_{i,s}\right] \\ &= E\left[\gamma_i\right] = 0, \ (s < t). \end{aligned} \quad (3)$$

According to the above assumptions, a foundation is laid for the identification of coefficients, lag period, and other parameters of the PVAR model in the next step.

2.3. PVAR Model Identification. The so-called model identification refers to the estimation and judgment of parameters such as coefficients and lag period in the model [22–24].

The first-order difference of formula (2) can be obtained:

$$\Delta Y_{i,t} = \Delta \sum_{k=1}^{m} \Phi_k Y_{i,t-k} + \Delta \sum_{j=1}^{m} \psi_j X_{i,t-j} + \Delta u_{i,t}. \quad (4)$$

From Hypothesis 3, we know that when $s < t - 1$, the following is

$$\begin{aligned} E\left[\Delta Y_{i,s}\right] &= E\left[\Delta X_{i,s}\right] \\ &= 0, \ (s < t - 1). \end{aligned} \quad (5)$$

Assuming that $y_{i,t}^{j}$ is the j variable in the economic variable vector $Y_{i,t}$, the first-order difference model of $y_{i,t}^{j}$ is expressed as a vector:

$$\Delta y_{i,t}^{j} = \sum_{k=1}^{m} \Phi_k^{j} \Delta Y_{i,t-k} + \sum_{l=1}^{m} \psi_l^{j} \Delta X_{i,t-l} + v_{i,t}^{j}. \quad (6)$$

In formula (6), $v_{i,t}^{j}$ is the random error term of the single-equation first-order difference model with the endogenous variable $y_{i,t}^{j}$ [25–27]. Therefore, from the orthogonality condition of Hypothesis 3, formula (6) has a vector of instrumental variables:

$$Z_{i,J} = \left[1, \Delta Y_{i,t-2}', \Delta Y_{i,t-3}', \ldots, \Delta Y_{i,2}', \Delta X_{i,t-2}', \Delta X_{i,t-3}', \ldots, \Delta X_{i,2}'\right]. \quad (7)$$

That is, the number of instrumental variables in formula (6) is $2t - 3$.

2.4. PVAR Model Parameter Estimation. This article mainly studies the fixed effect PVAR (1) model with a common deterministic event trend. The conclusion of the PVAR (1) model can be simply deduced and can be applied to PVAR (m) model. Therefore, this article focuses on the GMM estimation of the fixed effect PVAR (1) model. The fixed effect PVAR (1) model is as follows:

$$\left(I_m - \phi L\right)\left(Y_{i,t} - \gamma_i - \delta\right) = u_{i,t}. \quad (8)$$

The GMM estimation process of formula (8) is as follows:

$$\Delta Y_{i,t} - \delta = \phi\left(\Delta Y_{i,t-1} - \delta\right) + \Delta u_{i,t}, \ (t = 2, 3, \ldots, T). \quad (9)$$

Thus, the moment condition is obtained:

$$E\left\{\left[\left(\Delta Y_{i,t} - \delta\right) - \phi\left(\Delta Y_{i,t-1} - \delta\right)\right]Q_{i,t}^{T}\right\} = 0, \ (t = 2, 3, \ldots, T), \quad (10)$$

where $Q_{i,t}^{T}(1, Y_{i,0}^{T}, Y_{i,1}^{T}, \ldots, Y_{i,t-2,}^{T})^{T}$.

Let $\Delta Y_i = (\Delta Y_{i,2}, \ldots, \Delta Y_{i,T})^T$, $\Delta u_i = (\Delta u_{i,2}, \ldots, \Delta u_{i,T})^T$, $\Delta Y_{i-1} = (\Delta Y_{i,1}, \ldots, \Delta Y_{i,T-1})^T$, and $R_i = [\Delta Y_{i,-1}, l_{T-1}]$, $\Lambda = [\phi, \psi]$, $\psi = [I_m - \phi]\delta$, "stack" $T - 1$ formula (9) to get the following:

$$\Delta Y_i = R_i\Lambda + \Delta u_i, \quad i = 1, 2, \ldots, N. \quad (11)$$

Meanwhile, formula (11) multiplies the left tool variable matrix:

$$Q_i = \begin{bmatrix} Q_{i2} & 0 & \cdots & 0 \\ 0 & Q_{i3} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & Q_{iT} \end{bmatrix}, \quad (12)$$

$$Q_i\Delta Y = Q_iR_i\Lambda + Q_i\Delta u_i.$$

By solving the minimization problem [28–30],

$$\min_{\lambda} \sum_{i=1}^{N} \left(Q_i \otimes I_m\right)\mathrm{Vec}\left(\Delta Y_i\right) - \left(Q_iR_i \otimes I_m\right)\mathrm{Vec}\left(\Lambda\right). \quad (13)$$

Get Λ's estimate. The following is

$$\Sigma = \begin{bmatrix} 2\Omega & -\Omega & \cdots & 0 \\ -\Omega & 2\Omega & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 2\Omega \end{bmatrix}. \quad (14)$$

Moreover, by solving the moment condition equation,

$$E\left[\left(\Delta Y_{i,t} - \delta\right) - \phi\left(\Delta Y_{i,t-1} - \delta\right)\right] - 2\Omega = 0, \quad (t = 2, 3, \ldots, T). \quad (15)$$

The covariance matrix Ω of $u_{i,t}$ can be estimated. If all the roots of the characteristic formula (11) are outside the unit circle, that is, the model formula (11) is a trend-stationary PVAR process, then when $N \longrightarrow \infty$, the above GMM estimation is consistent and obeys the asymptotic normal distribution.

3. Theoretical Model Analysis of the Relationship between Education and Economic Development

Firstly, the two variables of education level and economic development level are explained and the indicators are selected; then, the theoretical relationship between education level and economic development level is analyzed based on Cobb–Douglas production function [31–33], which puts forward a theoretical basis for the empirical analysis of the interaction between the two.

3.1. Model Description. Due to the great limitations of the VAR model in the application process, the results obtained can be true and reliable only when the number of variables is small. The advantage of panel data is that it can collect a large number of sample observations. Therefore, the effective measure to solve the limitations of the VAR model is to combine the panel and VAR model, which gives birth to the panel data vector autoregressive model (PVAR).

In recent years, the PVAR model has been widely used, especially in studying the impact of economic fluctuations on the world, countries, and industries. Therefore, this article analyzes the relationship between education and economic development with the help of this model.

The PVAR model is expressed by the following formula:

$$y_{\mathrm{it}} = \alpha_{i,0} + \sum_{j=1}^{p} \alpha_{i,0}y_{i,t-j} + \gamma_i + \theta_t + \varepsilon_{\mathrm{it}}. \quad (16)$$

In formula (16), i represents region, t represents time, j represents lag period, γ_i represents individual effect, θ_t represents time effect, and ε_{it} represents random disturbance term.

According to the above theoretical model, PVAR studies the relationship between education and economic development. PVAR inherits the advantages of the VAR model, regards the research variables as endogenous variables, and takes each endogenous variable as a function of the lag value of all endogenous variables in the system so as to provide a rich structure and capture more characteristics of the data. In addition, the PVAR model allows individual effects and heteroscedasticity in the data. Due to the existence of a large number of cross-sectional data, the model allows the lag coefficient to change with time, relaxing the requirements of time stationarity of the data.

3.2. Index Selection. According to the above model, the following two indicators are selected to analyze the relationship between education and economic development.

(1) Education level: the development of education can also reflect a country's comprehensive national strength. In order to eliminate the influence of population and synthesize the topics and content involved in this article, the number of college students per 100000 population is selected as the index to measure the level of education.

(2) Economic development level: economic development refers to the improvement of a country or region's economic structure, the innovation of social structure, the quality of social life, the improvement of consumption capacity, and input-output efficiency [34–36]. China's population base is large, and the population distribution in various regions is also inconsistent. In order to eliminate the impact of the

population, combined with the research theme and content of this articke, per capita GDP is selected to measure the level of economic development.

3.3. Theoretical Relationship Analysis. In order to better explore the theoretical relationship between education level and economic development level, the Cobb–Douglas production function model is used as the analysis basis, education is incorporated into labor factors, economic development is used as output, and an economic growth model is constructed. Construct a production function for the i city:

$$P_i = A_i\left(M_i \cdot M_i'\right)^{\beta} v. \tag{17}$$

In formula (17), P_i represents the economic output, A_i represents the technical level, M_i represents the education factors in labor factors, M_i' represents the other factors in labor factors, and v represents the disturbance terms. Taking the natural logarithm on both sides of formula (17), we have the following:

$$\ln P_i = \ln A_i + \beta\left(\ln M_i \cdot \ln M_i'\right) + \ln v. \tag{18}$$

Using lowercase letters to replace formula (18) and calculating the growth rate between t period and $t-1$ period, you can get the following:

$$\dot{p} = \dot{a} + \dot{\beta}\left(\dot{m} + \dot{m}\prime\right). \tag{19}$$

Then, dividing both sides of formula (19) by p_{t-1} at the same time to calculate the EG_t path value of the equilibrium economic growth rate in period t, we can get the following:

$$EG_t^{\delta} = \frac{\dot{p}_t}{p_{t-1}} = \frac{1}{p_{t-1}}\left[\dot{a} + \dot{\beta}\left(\dot{m} + \dot{m}\prime\right)\right]. \tag{20}$$

The deviation of economic growth is as follows:

$$\begin{aligned} \mathrm{res}_t &= ERG_t - EG_t^{\delta} = ERG_t - \frac{\dot{p}_t}{p_{t-1}}, \\ &= ERG_t - \frac{1}{p_{t-1}}\left[\dot{a} + \dot{\beta}\left(\dot{m} + \dot{m}\prime\right)\right]. \end{aligned} \tag{21}$$

Based on the principle of production function and the above inference, we can know that economic growth benefits from investment in education, which can explain the relationship between economic development level and education level.

4. Empirical Research

In order to further analyze the dynamic relationship between China's education level and economic development level, a PVAR model is established for empirical research. At the same time, Stata11.0 and Eviews9.0 software are used to estimate the parameters of the system, obtain variance decomposition and impulse response function, so as to accurately analyze the impact force and action direction between them.

4.1. Stability Test. Before establishing the PVAR model, it is necessary to judge whether each variable is stable, that is, whether the collected data has a unit root. Use edu and eco to represent the level of education and economic development after processing. This article chooses the LLC test and PP-Fisher test in the unit root test method. Using the software Eviews9.0, the results of the unit root test for each variable or its difference value are shown in Table 1.

In Table 1, d (.) represents the first-order difference of variables. The numerical value represents the t value obtained by the unit root test of the variable. $***$, $**$, $*$ all mean that the t value is significant at the level of 1%, 5%, and 10%, respectively.

According to Table 1, the two variables in the whole country are stationary series, so the two variables in the whole country are $I(0)$ processes. The first-order differences of the two variables in the eastern, central, and western regions are stationary sequences, so each variable in these three regions is an $I(1)$ process.

4.2. Determination of Lag Period of PVAR Model. After determining the stationarity of the model variables, the next step is to determine the lag period suitable for constructing the PVAR model. This article selects the AIC information criterion, BIC information criterion, and HQIC information as the method of determining the lag period. The lag period of the model is tested using Stata11.0, and its specific procedures and operating steps are borrowed from the Stata package of the PVAR model. The specific test results are shown in Table 2.

According to Table 2, according to the criterion of minimum information value, the lag of the PVAR model suitable for the whole country is 4, and the lag of the PVAR model suitable for the eastern region is 4. The lag of the PVAR model suitable for the central region is 3, and the lag of the PVAR model suitable for the western region is 4.

4.3. Estimation of the PVAR Model. In order to ensure the stability of the model and the accuracy of parameters, the mean difference method and the previous difference method on the cross-section need to be used to eliminate the time effect and individual fixed effect of variables, respectively, before estimating the model. In this article, the method of estimating the model is the generalized moment estimation method (GMM estimation Method). The specific GMM estimation results of the PVAR model are shown in Table 3.

According to Table 3, (1) taking the level of economic development as the explanatory variable, on the whole, the level of economic development has a positive effect on itself. This positive effect is more obvious in the whole country and the eastern region. That is, the estimated coefficients of lag periods 1 to 4 are greater than 0 and significant at the significance level of 5%, which also shows that the economic development of the whole country and the eastern region in the previous period has a strong role in promoting its own later development. In terms of the four levels of education, its overall effect on the level of economic development shows positive and negative fluctuations. That is, the lagging education level sometimes has a

Table 1: Unit root test of each variable.

Area	Variable	LLC	PP-Fisher
Whole country	eco	−12.4798 ***	211.987 ***
	edu	−12.9384 ***	389.478 ***
East	d (eco)	−6.12987 ***	54.5919 ***
	d (edu)	−10.1687 ***	105.857 ***
Central	d (eco)	−4.38465 ***	30.6987 **
	d (edu)	−11.4721 ***	31.9172 **
West	d (eco)	−2.18765 **	47.7794 ***
	d (edu)	−7.57986 ***	81.3957 ***

Table 2: PVAR model lag period test.

Area	Lag order (lag)	AIC	BIC	HQIC
Whole country	1	−5.74972	−5.24832	−5.54873
	2	−6.82687	−6.21346	−6.58397
	3	−6.84425	−6.11079	−6.55247
	4	−7.10294 *	−6.23894 *	−6.75793 *
East	1	−5.47135	−4.3569	−5.02183
	2	−4.9148	−3.47178	−4.33353
	3	−6.83473	−5.01248	−6.10198
	4	−9.39398 *	−7.1241 *	−8.48669 *
Central	1	−7.61029	−6.1839	−7.04317
	2	−8.09864	−6.21294	−7.36417
	3	−11.1023 *	−8.64367 *	−10.1748 *
	4	−9.72284	−6.53641	−8.67687
West	1	−4.29124	−3.24568	−3.85558
	2	−3.42365	−2.08721	−2.88471
	3	1.47024	3.15289	2.15097
	4	−7.52485 *	−5.45487 *	−6.70127 *

Table 3: GMM estimation results of the PVAR model.

Area	Variable	h_eco		h_edu	
		Coef.	t	Coef.	t
Whole country	L1.h_eco	1.358 ***	−5.04	0.682 ***	−3.26
	L1.h_edu	0.12	−1.06	0.372 ***	−2.62
	L2.h_eco	0.968 ***	−4.28	0.624 ***	−3.12
	L2.h_edu	−0.329 ***	−2.67	0.186 ***	−3.68
	L3.h_eco	0.657 ***	−3.18	0.274	−2.68
	L3.h_edu	0.682 ***	−4.34	0.319 ***	−1.57
	L4.h_eco	0.828 ***	−3.34	0.457 **	−2.36
	L4.h_edu	−0.003	−0.02	0.081	−0.79
East	L1.h_eco	0.748 ***	−4.02	0.717 ***	−3.52
	L1.h_edu	−0.249	−1.05	0.293	−0.93
	L2.h_eco	0.379 **	−2.13	0.589 **	−2.3
	L2.h_edu	0.338	−1.45	0.212	−0.78
	L3.h_eco	1.429 ***	−5.87	1.329 ***	−4.26
	L3.h_edu	−0.710 ***	−4.02	−0.639 ***	−3.15
	L4.h_eco	1.564 ***	−6.68	1.576 ***	−5.38
	L4.h_edu	−0.177	−1.35	0.116	−0.64
Central	L1.h_eco	0.429 ***	−2.82	−0.32	−0.56
	L1.h_edu	0.812 ***	−2.7	−2.412 **	−2.38
	L2.h_eco	−2.312 ***	−3.24	2.977	−1.03
	L2.h_edu	0.131	−1.08	−0.052	−0.26
	L3.h_eco	−0.055	−0.28	−0.209	−0.45
	L3.h_edu	−0.130	−0.54	−0.35	−0.43
	L4.h_eco	0.612 ***	−3.67	−1.716 **	−2.38
	L4.h_edu	−0.366 ***	−2.69	1.245 **	−2.52

Table 3: Continued.

Area	Variable	h_eco		h_edu	
		Coef.	t	Coef.	t
West	L1.h_eco	0.539 **	−2.6	−0.454	−0.99
	L1.h_edu	0.031	−0.18	0.781 **	−2.42
	L2.h_eco	0.716 ***	−8.83	0.314 **	−2.24
	L2.h_edu	−0.054	−0.36	0.782 ***	−2.59
	L3.h_eco	−0.256	−1.43	0.845∗	−1.96
	L3.h_edu	1.167 ***	−12.33	0.492 ***	−3.32
	L4.h_eco	0.798 **	−2.45	−2.433 ***	−3.64
	L4.h_edu	−0.186	−0.81	1.069 **	−2.14

positive effect on the level of economic development and sometimes shows a negative effect. This shows that the level of education and the level of economic development have strong sensitivity and force. At the same time, it can also show that the driving effect of education level on economic development level is not easy to appear in a short time. Based on the analysis of the above GMM estimation results, we can know that the economy plays a strong role in promoting its own development, and this role is more obvious in the eastern region. (2) Taking the education level as the explanatory variable, considering the economic development level of each region at the four levels, although several periods of economic development level have a negative effect on the education level, the overall economic development level still has a positive effect on the education level. Especially in the eastern region, the lagging economic development level of phases I, III, and IV has a significant positive effect on the level of education, which also shows that the rapid economic development on the whole can promote the development of education. As far as the education level itself is concerned, the education level at the national level and the western region has an obvious positive effect itself, which also shows that the current development of education in the whole country and the western region will play a strong role in promoting the development of education in the future. The positive effect of the education level in the eastern and central regions on themselves is not obvious and sometimes even inhibits their own development. Based on the analysis of this part, we can know that the development of education is largely due to rapid economic growth. There are regional differences in the role of education in promoting itself. Education in the western region can promote its own development positively, while the eastern and central regions are just the opposite.

4.4. Result Analysis of PVAR Model

4.4.1. Impulse Response Function Analysis. The PVAR model usually discussed is essentially a system, which is composed of variables. Therefore, when estimating the model, the focus of the analysis is not only to analyze the impact of the change of a variable on another variable but also to analyze the dynamic response of the whole model system when a variable changes. Specifically, it is first to give the impact of a unit positive standard deviation to the error term and then observe the changes of endogenous variables in subsequent periods. This method is called impulse response function (IRF).

Analysis of Impulse Response Function in the Whole Country. After 500 runs of Monte Carlo simulation, when the level of economic development is impacted by a unit of positive standard deviation, the level of economic development and education will respond. The IRF of the whole country under the level of economic development is shown in Figure 1.

As can be seen from Figure 1, for the whole country and region, when the economic development level is impacted by a unit positive standard deviation, the response of economic development level and education level generally reflects a positive impact, and this impact has a long-lasting effect. This also shows that economic growth can promote the development of its own economy and education.

When the education level is impacted by a unit's positive standard deviation, the economic development level and education level respond, and the IRF of the whole country under the education level is shown in Figure 2.

As can be seen from Figure 2, for the whole country and regions, when the education level is impacted by the positive standard deviation of a unit, the response of the education level itself and the economic development level is in a fluctuating state. Still, it generally shows a negative impact, which also reflects that the development of education is difficult to promote economic development in the short term.

Analysis of Impulse Response Function in the Eastern Region. After 500 Monte Carlo simulations, when the economic development level is impacted by a unit positive standard deviation, the economic development level and education level respond. The IRF of the eastern region under the economic development level is shown in Figure 3.

As can be seen from Figure 3, for the eastern region, when the economic development level is impacted by a unit positive standard deviation, the response of economic development level and education level generally shows a positive impact, and this positive impact reaches the peak in phase IV. It can be seen that the impact of economic growth in the eastern region on its own economy and education is sustainable.

When the education level is impacted by a unit positive standard deviation, the economic development level and education level respond. The IRF of the eastern region under the education level is shown in Figure 4.

As can be seen from Figure 4, for the eastern region, when the education level is impacted by a unit positive

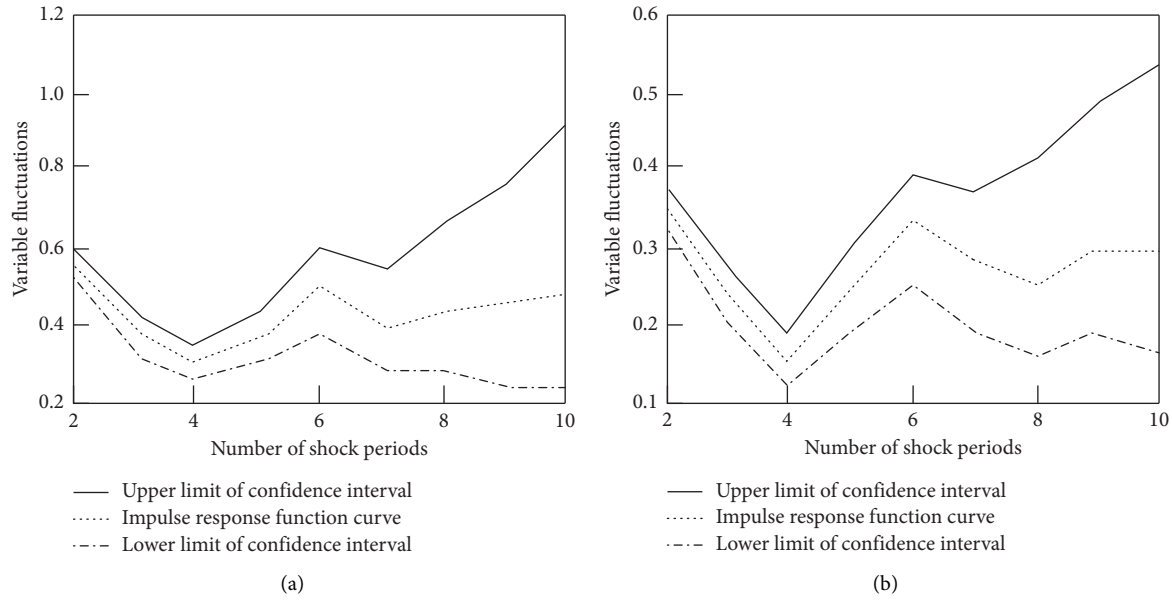


Figure 1: Impulse response function of the whole country under the level of economic development. (a) Response to the level of economic development. (b) Education level response.

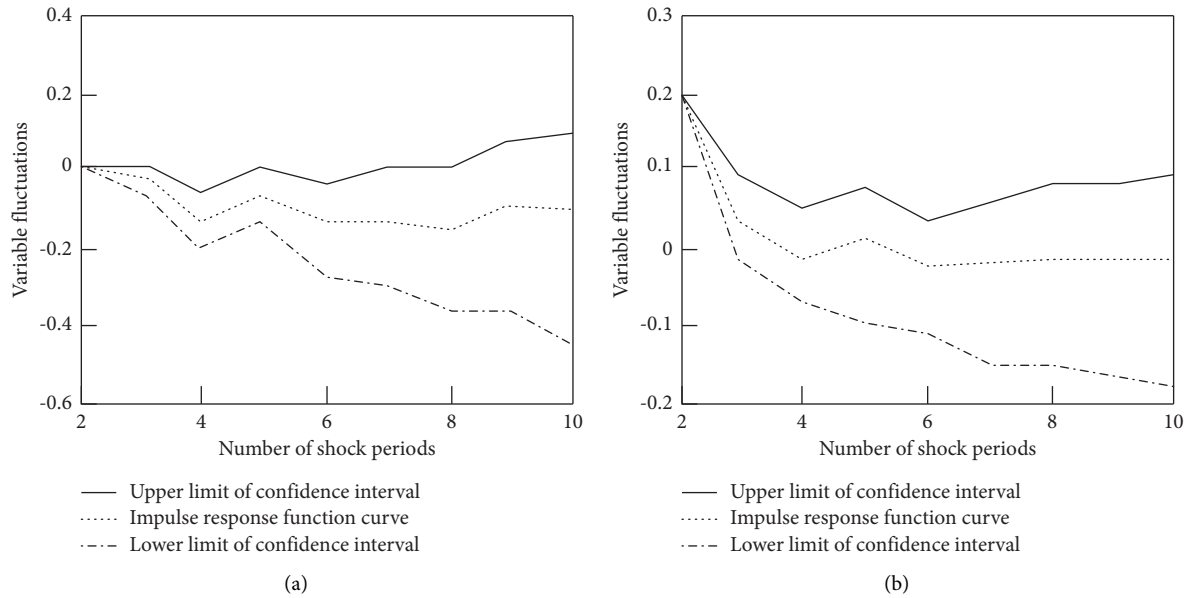


Figure 2: Impulse response function of the whole country under the educational level. (a) Response to the level of economic development. (b) Education level response.

standard deviation, the response of economic development level and education level shows a negative impact in the first four periods, while the response of the following three shows a positive impact. This shows that education in the eastern region has a certain incubation period to promote education itself and the economy.

Analysis of Impulse Response Function in the Central Region. After 500 Monte Carlo simulations, when the economic development level is impacted by a unit positive standard deviation, the economic development level and education level respond. The IRF of the central region under the economic development level is shown in Figure 5.

As can be seen from Figure 5, for the central region, when the economic development level is impacted by a unit positive standard deviation, the response of the economic development level itself shows a positive impact, while the response of the education level fluctuates between positive and negative effects. This also shows that the economic growth of the central region can drive the development of its own economy and education to a certain extent.

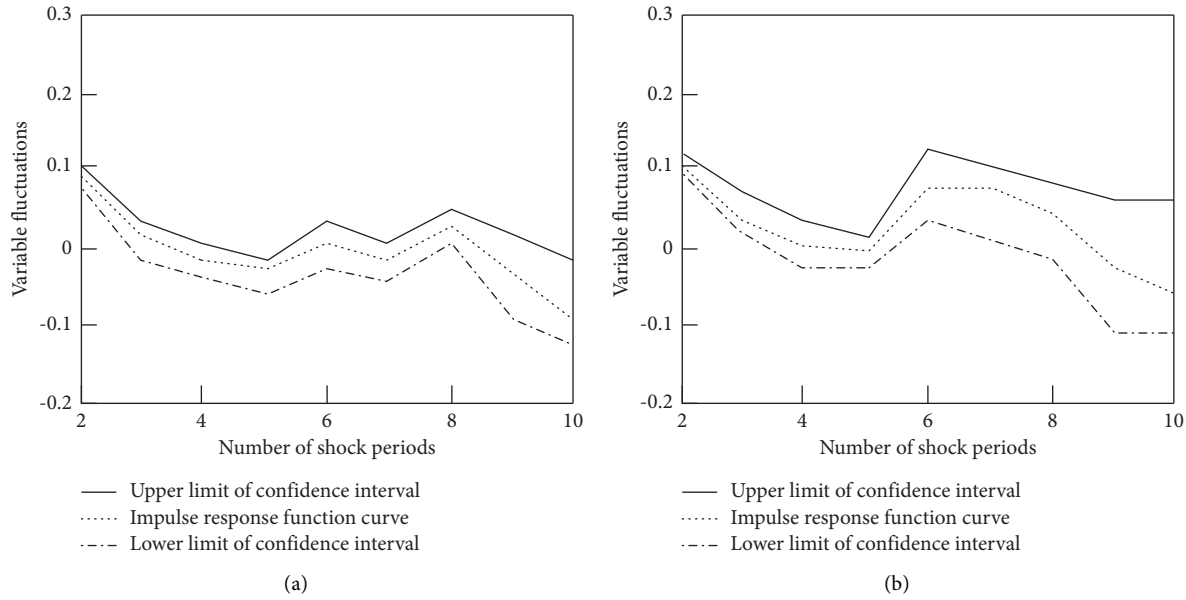


Figure 3: Impulse response function of the eastern region under the level of economic development. (a) Response to the level of economic development. (b) Education level response.

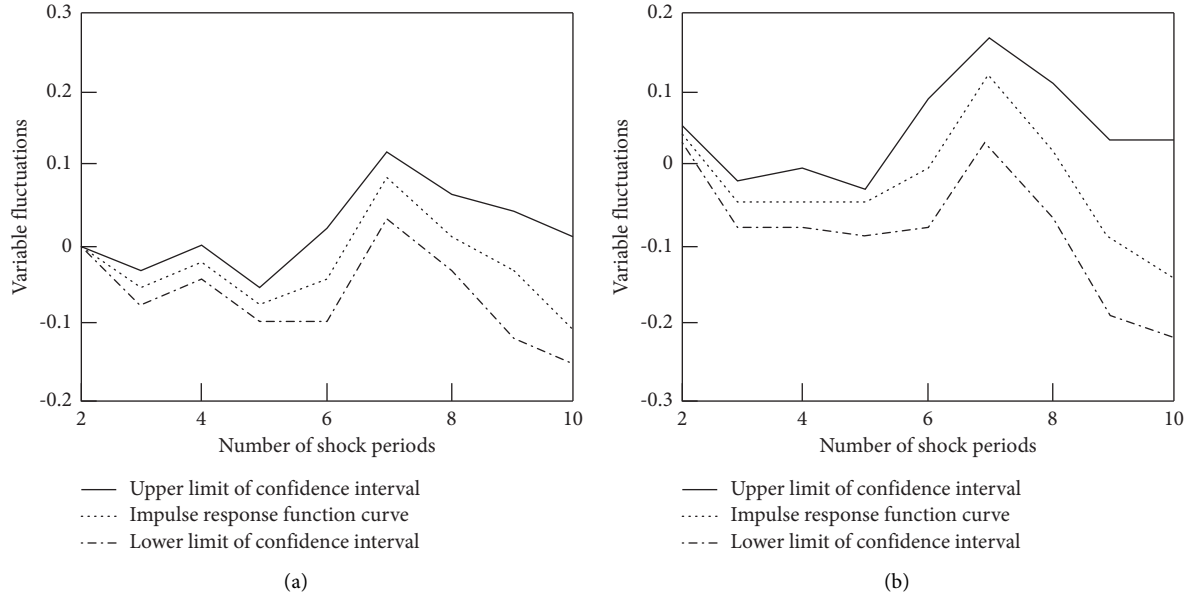


Figure 4: Impulse response function of the eastern region under education level. (a) Response to the level of economic development. (b) Education level response.

When the education level is impacted by a unit positive standard deviation, the economic development level and education level respond. The IRF of the central region under the education level is shown in Figure 6.

As can be seen from Figure 6, for the central region, when the education level is impacted by a unit positive standard deviation, the responses of economic development level and education level fluctuate between positive and negative effects, and there will be a relatively large positive response in the latter three. This also shows that the development of education in the central region has a certain incubation period to promote its own economy and education.

Analysis of Impulse Response Function in the Western Region. After 500 Monte Carlo simulations, when the economic development level is impacted by a unit positive standard deviation, the economic development level and education level respond. The IRF of the western region under the economic development level is shown in Figure 7.

As can be seen from Figure 7, for the western region, when the economic development level is impacted by a unit positive standard deviation, the response of the economic

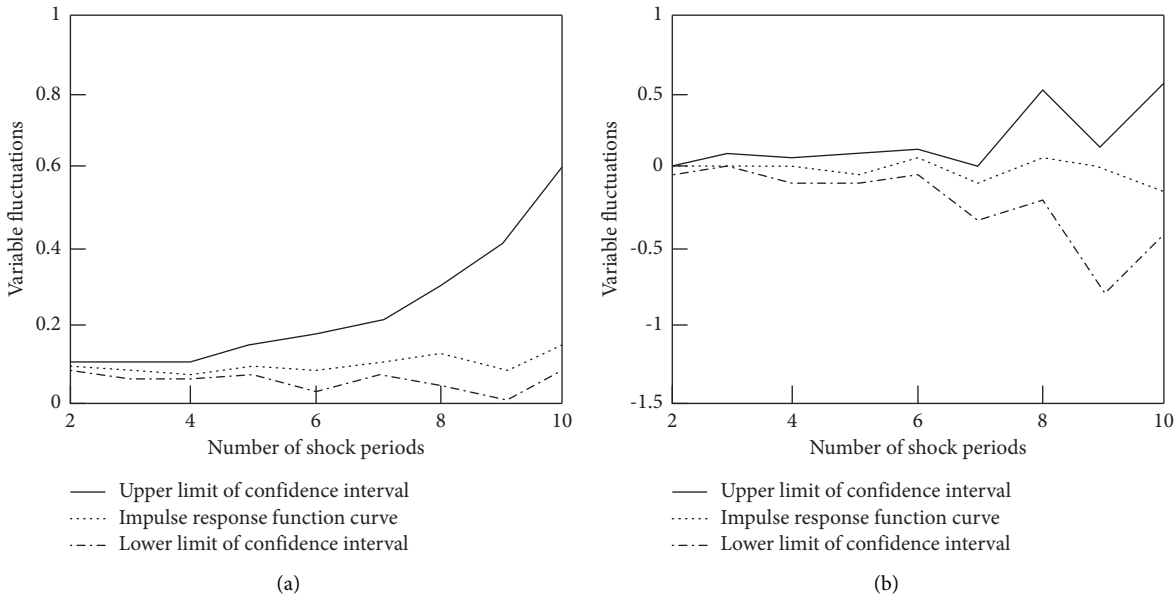


Figure 5: Impulse response function of the central region under the level of economic development. (a) Response to the level of economic development. (b) Education level response.

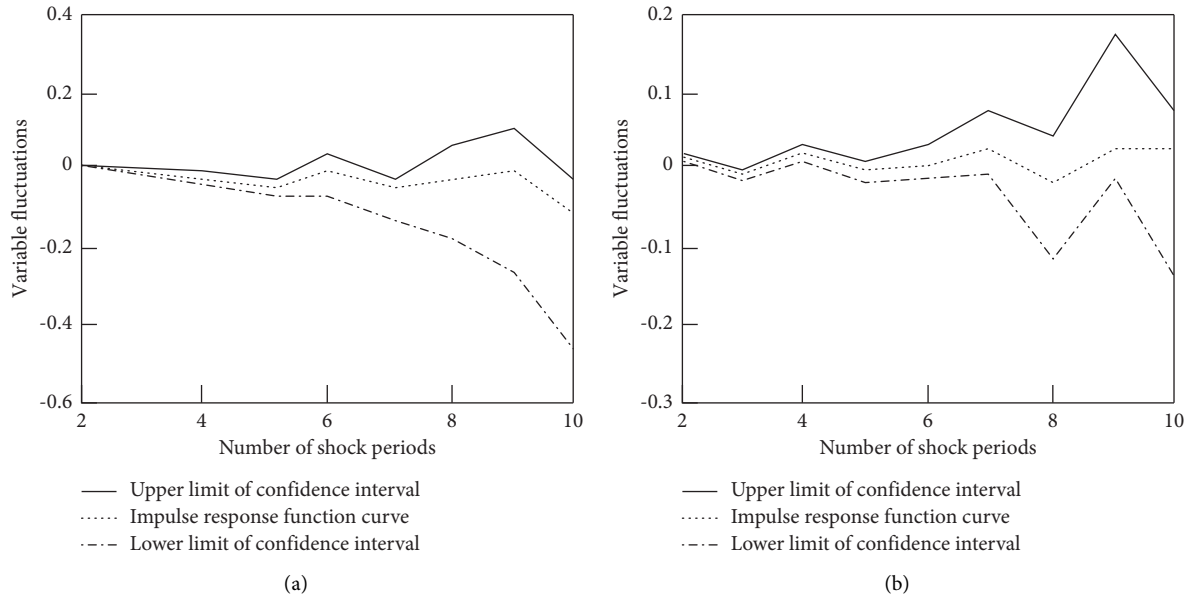


Figure 6: Impulse response function of the central region under education level. (a) Response to the level of economic development. (b) Education level response.

development level itself in the first two periods shows a positive impact, while the response of the education level basically shows a negative impact, but the response range is relatively small.

When the education level is impacted by a unit positive standard deviation, the economic development level and education level respond. The IRF of the western region under the education level is shown in Figure 8.

As can be seen from Figure 8, for the western region, when the education level is impacted by a unit positive standard deviation, the responses of economic development level and education level are roughly the same. In the previous periods, neither of them will be affected, and the response values of the latter two have a slow upward trend. This shows that the promotion of the development of education in the west to its own economy and education is not easy in a short time.

Combining the results of IRF analysis at four levels, we can get the results similar to GMM estimation analysis: economic growth can not only drive its own development

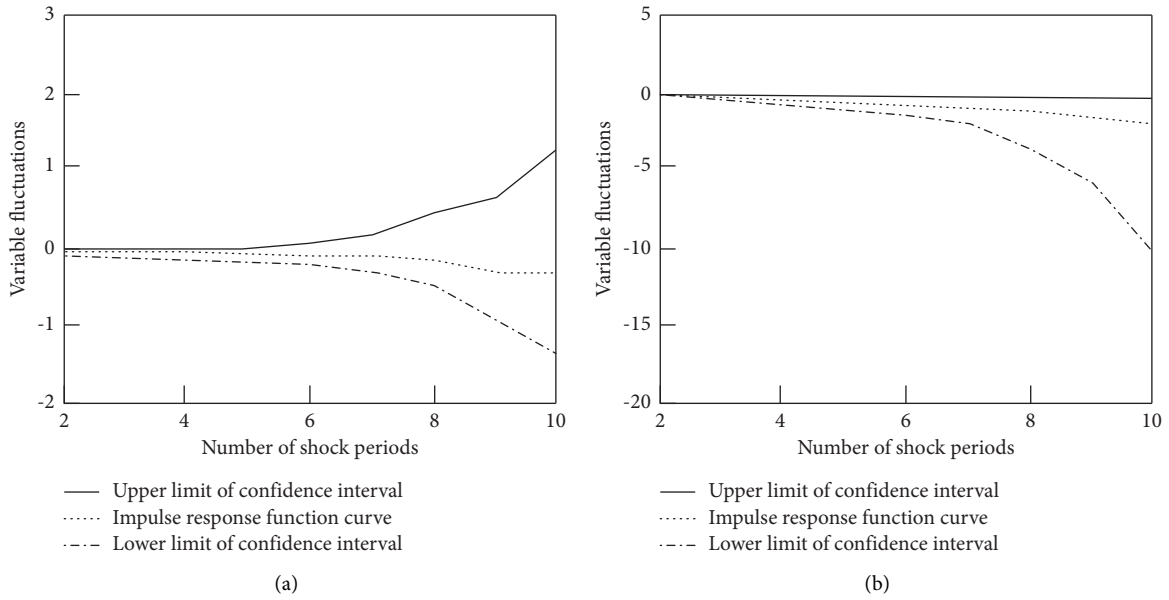


Figure 7: Impulse response function of the western region under the level of economic development. (a) Response to the level of economic development. (b) Education level response.

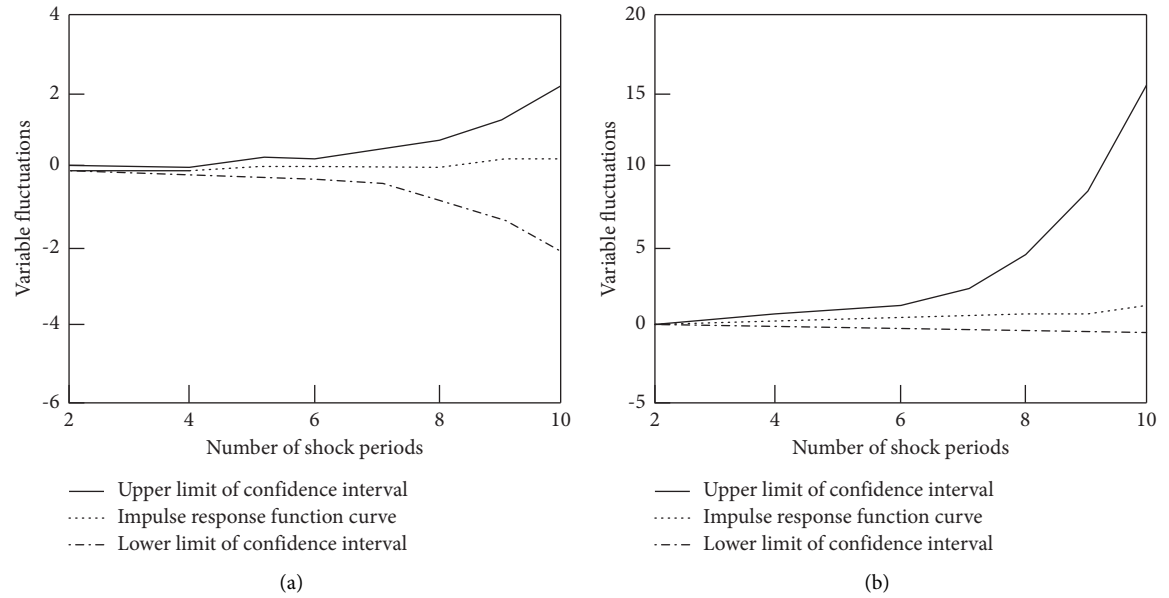


Figure 8: Impulse response function of the western region under education level. (a) Response to the level of economic development. (b) Education level response.

but also promote the improvement of education level. The role of education in promoting both is difficult to show in the short term.

4.4.2. Variance Decomposition Analysis. Similar to IRF, variance decomposition is another method to analyze the dynamic change of the whole model system. This method is to analyze the proportion of mean square error borne by each variable in the whole model and use it to measure the importance of the change of the variable to the whole model. The results of the method decomposition of the national, eastern, central and western PVAR models made by the software Stata11.0 are shown in Table 4.

According to Table 4, (1) the interpretation degree of 20 periods in China and central regions is basically consistent with that of 10 periods; that is, the interpretation degree of variables to all variable error terms remains basically stable for a long time. (2) Most of the changes in the economic development level of the whole country, the eastern, central, and western regions, are caused by themselves, and all of them exceed 50%. This shows that the level of economic

Table 4: Variance decomposition of the PVAR model.

Area	Variable	Number of shock periods	eco	edu
Whole country	eco	10	0.579	0.006
	edu	10	0.56	0.058
	eco	20	0.59	0.008
	edu	20	0.541	0.07
East	eco	10	0.858	0.001
	edu	10	0.858	0.001
	eco	20	0.829	0.004
	edu	20	0.829	0.004
Central	eco	10	0.719	0.286
	edu	10	0.719	0.286
	eco	20	0.719	0.286
	edu	20	0.719	0.286
West	eco	10	0.668	0.002
	edu	10	0.415	0.419
	eco	20	0.658	0.002
	edu	20	0.623	0.08

development has strong inertia. At the same time, it can be seen that the economic development of the eastern and central regions depends more on their own role. (3) The development of education in the whole country, eastern, central, and western regions, depends more on promoting the economy. Moreover, from the 10th to 20th period, the influence of the economic development of the western region on the development of education has gradually increased, and the influence on the development of education has changed more greatly. (4) Compared with other regions, the influence of education in Central China on the economy and education is stronger, and the dependence of education in Central China on itself is also higher than that in other regions.

5. Conclusion

Using the Cobb–Douglas production function as the theoretical model and PVAR model, this article makes an empirical analysis on the relationship between education level and economic development level. The following conclusions can be obtained:

(1) On the whole, economic development can drive the development of education, but there are obvious regional differences in the impact of economic growth on the development of education. The impact of economic growth in the eastern region on education is significantly higher than that in the central and western regions. This also shows that there is a regional imbalance in China's education and economic development, and the development degree of the East is obviously better than that of the middle and west.

(2) There is an interactive relationship between the level of education and the level of economic development, but these interactive relationships are not equal and there are great differences. The model analysis shows that economic growth can promote the development of education, but on the contrary, the improvement of education level does not promote economic growth very significantly, and some even inhibit economic growth in the short term.

(3) There is a certain incubation period for the interaction between the level of education and the level of economic development. Through empirical analysis, we can know that in the short term, educational development in the areas with rapid economic development can also develop rapidly. It is difficult for education to play its role in promoting economic development in the short term, but with the passage of time, the influence of education is gradually highlighted.

Data Availability

The raw data supporting the conclusions of this article will be made available by the author, without undue reservation.

Conflicts of Interest

The author declares no conflicts of interest regarding this work.

Acknowledgments

This work was sponsored in part by 2020 Education Sciences "The Thirteenth Five-Year" Planning Foundation of Liaoning Province "Research on the Optimization of Financial Resource Allocation to Local Higher Education in China" (JG20DB193).

References

[1] S. Heaney, H. O'Connor, S. Michael, J. Gifford, and G. Naughton, "Effectiveness of education interventions designed to improve nutrition knowledge in athletes: a systematic review," *Sports Medicine*, vol. 49, no. 3, pp. 1769–1786, 2019.

[2] S. Lindgren, "Cookstove implementation and Education for Sustainable Development: a review of the field and proposed

research agenda," *Renewable and Sustainable Energy Reviews*, vol. 146, no. 3, Article ID 111184, 2021.

[3] M. M. Gomes, D. Driman, Y. S. Park, T. J. Wood, R. Yudkowsky, and N. L. Dudek, "Teaching and assessing intra-operative consultations in competency-based medical education: development of a workplace-based assessment instrument," *Virchows Archiv*, pp. 1–11, 2021.

[4] S. S. Asumadu and S. Vladimir, "Effect of foreign direct investments, economic development and energy consumption on greenhouse gas emissions in developing countries," *The Science of the Total Environment*, vol. 646, pp. 862–871, 2019.

[5] L. Wang, H. Yan, X. W. Wang et al., "The potential for soil erosion control associated with socio-economic development in the hilly red soil region, southern China," *Catena*, vol. 194, Article ID 104678, 2020.

[6] A. K. Magnusson and W. Dekker, "Economic development in times of population decline—a century of European eel fishing on the Swedish west coast," *ICES Journal of Marine Science*, vol. 78, no. 1, pp. 185–198, 2020.

[7] W. Widaryanti, A. Putra, and E. Timotius, "Implications of digital transformation on developing human resources in business practice in Indonesian: analysis of the publication," *International Journal of Business, Economics & Management*, vol. 4, no. 1, pp. 157–164, 2021.

[8] H. Siti, H. Sumardi Wardhani, and A. H. P. Kusuma, "Structural model of developing human resources performance: empirical study of Indonesia states owned enterprises," *Journal of Asian Finance Economics and Business*, vol. 7, no. 3, pp. 211–221, 2020.

[9] F. E. Hardiyanto, "Revitalizing the prophetic teacher ethic in developing human resources for education," *Universal Journal of Educational Research*, vol. 8, no. 6, pp. 2686–2692, 2020.

[10] E. Soukiazis, S. Proenca, and P. A. Cerqueira, "The interconnections between renewable energy, economic development and environmental pollution: a simultaneous equation system Approach," *Energy Journal*, vol. 40, no. 4, pp. 1–23, 2019.

[11] S. Hossain, M. M. Khudri, and R. Banik, "Regional education and wealth-related inequalities in malnutrition among women in Bangladesh," *Public Health Nutrition*, vol. 18, no. 8, pp. 1–19, 2021.

[12] T. Agasisti, A. Egorov, D. Zinchenko, and O. Leshukov, "Efficiency of regional higher education systems and regional economic short-run growth: empirical evidence from Russia," *Industry & Innovation*, vol. 28, no. 4, pp. 507–534, 2021.

[13] S. A. Solarin, M. Shahbaz, and S. M. Hammoudeh, "Sustainable economic development in China: modelling the role of hydroelectricity consumption in a multivariate framework–ScienceDirect," *Energy*, vol. 168, pp. 516–531, 2019.

[14] B. Liu, Q. Gao, L. Liang, J. Sun, C. Liu, and Y. Xu, "Ecological relationships of global construction industries in sustainable economic and energy development," *Energy*, vol. 234, no. 8, Article ID 121249, 2021.

[15] L. Miao and Y. Sun, "Quantitative analysis of regional economic balance and sustainable development in yangtze river delta and pearl river delta," *Journal of Coastal Research*, vol. 115, no. 1, pp. 570–574, 2020.

[16] W. Wang, "The correlation effect between transportation and economy in beijing-tianjin-heibei area: an analysis based on PVAR model," *E3S Web of Conferences*, vol. 253, no. 4, Article ID 01009, 2021.

[17] F. Xu, L. Ma, and N. Iqbal, "Interaction mechanism between sustainable innovation capability and capital stock: based on PVAR model," *Journal of Intelligent and Fuzzy Systems*, vol. 38, no. 6, pp. 1–17, 2020.

[18] D. Guo, P. N. Hu, F. Li, H. Wang, and G. Chen, "An empirical study on the impact of seaports on urban economic development based on VAR model: the case of shenzhen and shenzhen port in the context of industrial transfer," *Journal of Coastal Research*, vol. 111, no. 1, pp. 35–39, 2020.

[19] T. T. Nguyen, M. T. Dang, A. W. Liew, and J. C. Bezdek, "A weighted multiple classifier framework based on Random Projection," *Information Sciences*, vol. 490, pp. 36–58, 2019.

[20] M. S. Tenan, A. D. Vigotsky, and A. R. Caldwell, "Comment on: "A method to stop analyzing random error and start analyzing differential responders to exercise"," *Sports Medicine*, vol. 50, no. 2, pp. 431–434, 2020.

[21] S. J. Dankel and J. P. Loenneke, "A method to stop analyzing random error and start analyzing differential responders to exercise," *Sports Medicine (Auckland, N.Z.)*, vol. 50, no. 103–7, pp. 231–238, 2020.

[22] H. Hong, G. Song, and S. Du, "Collective causal inference with lag estimation," *Neurocomputing*, vol. 323, pp. 299–310, 2019.

[23] T. J. Ludeña Cervantes, S. H. Choi, and B. S. Kim, "Flight control design using incremental nonlinear dynamic inversion with fixed-lag smoothing estimation," *International Journal of Aeronautical and Space Sciences*, vol. 21, no. 4, pp. 1047–1058, 2020.

[24] R. Biswas and P. A. Durbin, "Assessment of viscosity models that incorporate lag parameter scaling," *International Journal of Heat and Fluid Flow*, vol. 78, Article ID 108427.1, 2019.

[25] Y. W. Nam, "Hyers-ulam stability of the first order difference equation generated by linear maps," 2021, https://arxiv.org/abs/2101.02364, Article ID 02364.

[26] Y. Liu, R. Xu, T. Feng, and H. Yao, "Analysis of phase synchronizaiton error of bistatic SAR," *Computer Simulation*, vol. 36, no. 9, pp. 15–19, 2019.

[27] A.-R. Baias, F. Blaga, and D. Popa, "On the best Ulam constant of a first order linear difference equation in Banach spaces," *Acta Mathematica Hungarica*, vol. 163, no. 2, pp. 563–575, 2021.

[28] A. Repetti and Y. Wiaux, "Variable metric forward-backward algorithm for composite minimization problems," *SIAM Journal on Optimization*, vol. 31, no. 2, pp. 1215–1241, 2021.

[29] Y. Liu, S. Bi, and S. Pan, "Several classes of stationary points for rank regularized minimization problems," *SIAM Journal on Optimization*, vol. 30, no. 2, pp. 1756–1775, 2020.

[30] S. Bartels and M. Milicevic, "Efficient iterative solution of finite element discretized nonsmooth minimization problems," *Computers & Mathematics with Applications*, vol. 80, no. 5, pp. 588–603, 2020.

[31] Q. Zhang, W. Dong, C. Wen, and T. Li, "Study on factors affecting corn yield based on the Cobb-Douglas production function," *Agricultural Water Management*, vol. 228, Article ID 105869, 2020.

[32] L. Dugulean, "Forecasting the economic growth in Romania with the cobb–douglas production function," *SERIES V - ECONOMIC SCIENCES*, vol. 12, pp. 157–174, 2020.

[33] A. Matsumoto and F. Szidarovszky, "Delay two-sector economic growth model with a Cobb–Douglas production function," *Rivista di Matematica per le Scienze Economiche e Sociali*, vol. 44, no. 1, pp. 341–358, 2021.

[34] H. Wu and C. Fu, "The influence of marine port finance on port economic development," *Journal of Coastal Research*, vol. 103, no. 1, pp. 163–167, 2020.

[35] M. Ahmad, C. Işık, G. Jabeen, T. Ali, I. Ozturk, and D. W. Atchike, "Heterogeneous links among urban

concentration, non-renewable energy use intensity, economic development, and environmental emissions across regional development levels," *The Science of the Total Environment*, vol. 765, Article ID 144527, 2021.

[36] Y. Zhao, K. Li, and L. Zhang, "A meta-analysis of online health adoption and the moderating effect of economic development level," *International Journal of Medical Informatics*, vol. 127, pp. 68–79, 2019.

Hindawi
Scientific Programming
Volume 2021, Article ID 9948800, 11 pages
https://doi.org/10.1155/2021/9948800

Research Article

Construction of Multimedia Assisted Legal Classroom Teaching Model Based on Data Mining Algorithm

Yu Lu[1] **and Wang Lizhi**[2]

[1]*Faculty of Law, Changchun University of Finance and Economics, Changchun 130122, China*
[2]*School of Political Science and Law, Northeast Normal University, Changchun 130117, China*

Correspondence should be addressed to Wang Lizhi; wanglz275@nenu.edu.cn

Received 9 November 2021; Accepted 2 December 2021; Published 24 December 2021

Academic Editor: Baiyuan Ding

In order to quickly and accurately retrieve a required part from massive multimedia educational resources and improve the utilization of educational resources, a multimedia assisted legal classroom teaching model based on data mining algorithm is designed. Firstly, the attributes of multimedia assisted legal classroom teaching resources are judged, and the numerical resources are standardized and discretized. Then, the B+ tree is used to establish the model's indexes and index library, and the corresponding retrieval algorithm is designed to complete the resource search, establish the data distribution structure model of the multimedia assisted legal classroom teaching system, mine the data, reconstruct the phase space of the fused data information flow, extract the high-order moment features of the specific data in the multimedia assisted legal classroom teaching system in the reconstructed high-dimensional phase space, and realize the accurate mining of the feature data. The experimental results show that the teaching effect of the designed model has more advantages and can promote the improvement of students' performance.

1. Introduction

With the development of modern information technology, the use of multimedia technology to intuitively display legal practice problems in law classroom teaching has become a regular auxiliary means of law teaching in colleges and universities [1, 2]. Multimedia assisted law classroom teaching not only helps to increase the amount of information in law classroom teaching and improve the efficiency of teaching and learning, but also helps to change the traditional mode of law teaching emphasizing theory and neglecting practice, which plays an important role in promoting the reform of law classroom teaching.

However, in the process of applying multimedia technology to assist law teaching, there are also many problems, which affect the effect of multimedia assisted teaching and need to be improved [3, 4]. The traditional law classroom teaching model no longer meets the needs of modern teaching, and great changes have taken place in teaching and learning model. In law classroom teaching, the use of computer multimedia hardware has developed the teaching form of law learning [5]. Multimedia network technology shows the legal teaching content to students through text, image, sound, and other ways; enriches the legal teaching content; and can mobilize students' interest in learning law [6, 7]. Therefore, many scholars have carried out a lot of research on multimedia network assisted legal classroom learning and teaching methods, which makes people have a further understanding of multimedia assisted legal classroom learning and teaching methods. Chen [8] proposed the application of computer information technology in the reform of law classroom teaching methods, and by establishing a legal resource database, creating a virtual legal classroom, and making suggestions on the reform of legal education through electronic bulletin boards, we can strengthen the application of computer information technology in the teaching reform of law courses. However, this method is inefficient because it cannot quickly retrieve the required part from the massive multimedia educational resources. Reference

[9] proposes the introduction of classroom teaching methods of moral education and rule of law. According to different introduction media, teachers can adopt different introduction methods, such as video introduction, graphic introduction, and case introduction. Each introduction method pays attention to specific operation skills, which are student-centered and problem-oriented and can quickly guide students into the predetermined learning track. However, there are still some deficiencies, such as low utilization of educational resources and poor teaching effect. Further research is needed to improve the level of law classroom learning and teaching. Therefore, a multimedia assisted law classroom teaching model based on data mining algorithm is proposed. Data mining algorithm is a group of heuristics and calculations to create data mining model based on data. In order to build a multimedia assisted legal classroom teaching model, the algorithm will first analyze the data and find specific types of patterns and trends. The algorithm uses the results of this analysis to define the best parameters for creating and mining multimedia assisted legal classroom teaching model. These parameters are then applied to the entire data set to extract feasible patterns and detailed statistics, reconstruct the phase space of the fused data information flow, extract the high-order moment features of specific data in the multimedia assisted legal classroom teaching system from the reconstructed high-dimensional phase space, and realize the accurate mining of feature data. The research shows that the model has good teaching effect.

2. Multimedia Assisted Legal Classroom Teaching Resource B+ Tree Hierarchical Index

Index is an important part of relational database, as shown in Figure 1.

As can be seen from Figure 1, the index occupies a central position and plays an important role in the multimedia assisted legal classroom teaching resource database system. The role of index in multimedia assisted legal classroom teaching resource database is to quickly locate the resource data required by users from the resource database and achieve rapid acquisition [10]. The index establishment process is shown in Figure 2.

It can be seen from Figure 2 that the quality of database index establishment is related to the operation performance of the database. Therefore, using the B+ tree index structure, this paper establishes a hierarchical index model for multimedia assisted legal classroom teaching resources [11, 12]. B+ tree index has the advantages of lower disk reading and writing cost, more convenient data information traversal, more stable query efficiency, and less memory space.

2.1. Multimedia Assisted Legal Classroom Teaching Resource Attribute Judgment and Processing. Before giving the data to be indexed, it is necessary to judge the attributes of multimedia assisted legal classroom teaching resources. These

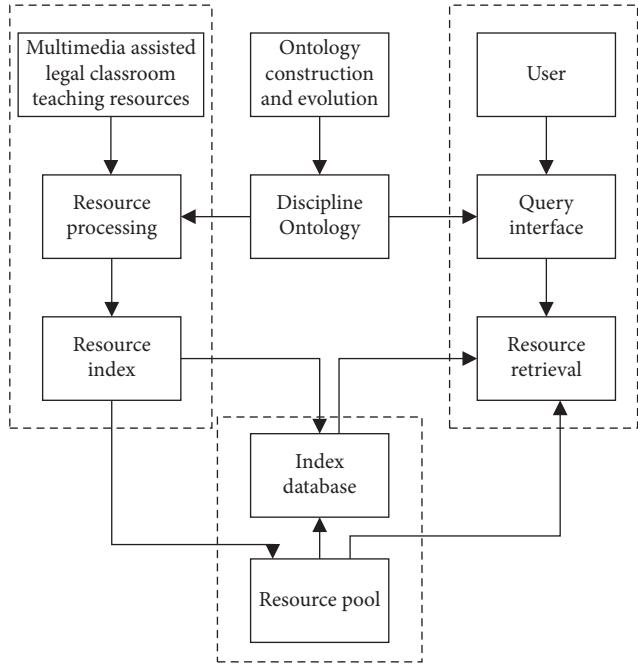


Figure 1: Multimedia assisted legal classroom teaching resource relational database system.

attributes specifically include value attribute, time attribute, and quasi-public goods attribute.

Value attribute: The development of multimedia assisted legal classroom teaching resources requires the builders to continuously study, research, collect, summarize, improve, practice, integrate the builders' high sense of responsibility and selfless professionalism, bring enlightenment to the majority of teachers and students, and obtain universal recognition. Therefore, the value attribute of multimedia assisted legal classroom teaching resources is obvious.

Time attribute: The time attribute of multimedia assisted legal classroom teaching resources includes long development cycle and more time. The resources developed in a specific time interval must be updated with the continuous development of social economy, new technologies emerging endlessly, and upgraded standards and specifications. In addition, the time attribute is also reflected in the effectiveness and quickness of teachers' and students' access to teaching resources.

Quasi-public goods attribute: The open sharing of multimedia assisted legal classroom teaching resources prevents mutual competition and exclusion among people. However, due to the high cost and limited number of multimedia assisted legal classroom teaching resources and the individualization of teaching in various colleges and universities, the sharing mechanism is not perfect. The attribute of quasi-public goods of multimedia assisted legal classroom teaching resources is more obvious, making it difficult to strictly comply with the nonexclusivity of public goods.

The function of judging attributes is to determine whether resources are character resources or numerical resources. The index forms of the two are different. The

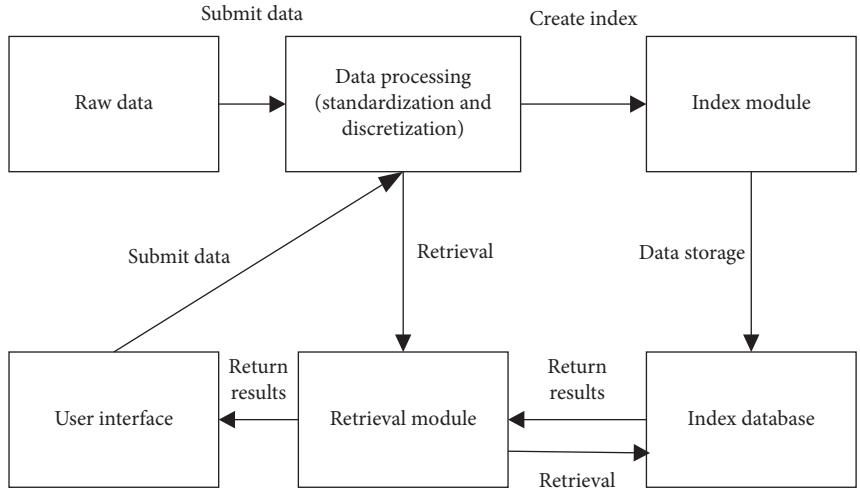


Figure 2: Database index establishment process.

former establishes inverted table index and the latter establishes B+ tree index [13]. The core of the attributes of multimedia assisted legal classroom teaching resources is the B+ tree index model. Therefore, only the subsequent processing of numerical resources is analyzed, including the standardization and discretization of numerical multimedia assisted legal classroom teaching resources [14, 15].

2.1.1. Standardization of Numerical Multimedia Assisted Legal Classroom Teaching Resources. Standardization refers to the form and dimension of statistical numerical multimedia assisted legal classroom teaching resources for subsequent processing and analysis. There are three main standardization methods:

(1) Standardization method

The sequence $\{x_1, x_2, ..., x_n\}$ is transformed to generate the new sequence $\{y_1, y_2, ..., y_n\} \in [0, 1]$.

$$y_i = \frac{x_i - \min_{1 \le j \le n} \{x_j\}}{\max_{1 \le j \le n} \{x_j\} - \min_{1 \le j \le n} \{x_j\}}. \quad (1)$$

(2) Formalization method

The sequence $\{x_1, x_2, ..., x_n\}$ is transformed to generate the new sequence $\{y_1, y_2, ..., y_n\}$. The mean value of the new sequence is 0, and the variance is 1.

$$y_i = \frac{x_i - \left(\sum_{i=1}^{n} x_i / n\right)}{\sqrt{\sum_{i=1}^{n} \left(x_i - \sum_{i=1}^{n} x_i / n\right)^2 / n - 1}}. \quad (2)$$

(3) Normalization method

The sequence $\{x_1, x_2, ..., x_n\}$ is transformed to generate the new sequence $\{y_1, y_2, ..., y_n\} \in [0, 1]$, and $\sum_{i=1}^{n} y_i = 1$.

$$y_i = \frac{x_i}{\sum_{i=1}^{n} x_i}. \quad (3)$$

2.1.2. Discretization of Numerical Multimedia Assisted Legal Classroom Teaching Resources. In order to easily extract the key information in the resources and establish the index, it is necessary to discretize the numerical multimedia assisted legal classroom teaching resources with continuous attributes [16, 17]. The basic process is as follows:

Step 1: Sort the continuous attribute values of multimedia assisted legal classroom teaching resources according to a specified rule.

Step 2: Preliminarily determine the division breakpoint of continuous attributes.

Step 3: Continue to split or merge breakpoints according to the given judgment criteria [18].

Step 4: Judge whether the number of breakpoints meets the number specified by the user. If yes, terminate the whole continuous attribute discretization process; otherwise, continue to follow the third step until the conditions are met.

At present, discretization is mainly divided into unsupervised discretization and supervised discretization, as shown in Table 1.

Among them, unsupervised discretization refers to the method that does not use class information and does not consider class attributes in the discretization process and whose input data set only contains the values of attributes to be discretized. Supervised discretization refers to the method of using class information in the discretization process. It can be seen from Table 1 that unsupervised discretization has the characteristics of simple method and easy operation, but it needs to manually specify the number of intervals. Supervised discretization does not need to specify the number of intervals artificially, but the operation is complex. The number of data types in the discrete set is the number of data discrete classes. As an effective data preprocessing method, the results of discretization have an essential impact on the results of data analysis. In data processing, discretization should be carried out according to the data characteristics. In discretization, different strategies such as static classification

Table 1: Discretization method.

Type	Definition	Characteristic	Specific representative method
Unsupervised discretization	A method that does not use class information in the discretization process	The method is simple and easy to operate, but it needs to specify the number of intervals artificially	Equal width box method Equal frequency box method String analysis method
Supervised discretization	A method that uses class information in the discretization process	There is no need to specify the number of intervals artificially, but the operation is complex	IR method Entropy based discretization method Discretization method based on chi square

or dynamic combination can be adopted. Reasonable strategies are helpful to effectively mine data characteristics.

2.2. Index and Its Database Creation. Based on the above processed data, research on index and database creation was conducted [19, 20]. The essence of an index is a special table or graph, which is composed of the key words of the index and the address where the key words are stored, and the index library is the collection of all indexes.

2.2.1. Index Creation. B+ tree is mainly used to index each multimedia assisted legal classroom teaching resource. B+ tree is developed on the basis of B-tree, and it solves the problem that B-tree cannot search in order. The comparison between B-tree and B+ tree is shown in Table 2.

Based on the basic structure of B+ tree, this paper constructs the index structure of multimedia assisted legal classroom teaching resources, including B+ tree skeleton construction, B+ tree loading, and B+ tree assignment, as follows.

Step 1: B+ tree skeleton. B+ tree is a typical tree structure, which is mainly composed of leaf nodes, branch nodes, and root nodes. Firstly, determine the parameters in the B+ tree index, including the number of layers of B+ tree, the number of subnodes of nodes in B+ tree, the number of nodes in B+ tree, the number of B+ leaf nodes, the number of (key, value) pairs in B+ leaf nodes, the number of B+ tree nodes, and node array. Then, the parent-child or sibling relationship between each node is established, and finally a complete B+ tree index framework is formed [21, 22].

Step 2: B+ tree loading. The previous step completes the association between the B+ tree nodes. In this step, each node in the B+ tree needs to be loaded with the corresponding multimedia assisted legal classroom teaching resource "key." The loading sequence is the same as the B+ tree skeleton creation sequence, from left to right and from bottom to top.

Step 3: B+ tree assignment. Assign the "key" of multimedia assisted legal classroom teaching resources value. The assignment is related to the position in the index and the sorting in later user retrieval.

2.2.2. Index Library Creation. If the indexes of each multimedia assisted legal classroom teaching resource are stored, it will occupy a lot of content space. Therefore, it is necessary to establish a comprehensive index library to collect these indexes [23–25]. A word table is saved for each different character in the index library, and the word table corresponding to each character contains all the positions of the character in the resource document. The word list is shown in Table 3.

2.3. Search Method. After the indexes and index database are established, a matching retrieval method is also needed to help users complete the retrieval. The retrieval process is as follows:

Step 1: Select the literature search tool and confirm the search field.

Step 2: The user inputs a search term or sentence.

Step 3: Process the search words or sentences, including query segmentation and keyword extraction. The former refers to dividing a search word sequence into several query blocks to judge the relationship between these blocks" for search engine analysis. Retrieval word segmentation methods mainly include word segmentation method based on preset dictionary, word segmentation method based on natural language understanding, and word segmentation method based on statistics [26]. The latter mainly extracts the most critical part from the divided query blocks, that is, the retrieval focus. This is directly related to the query accuracy.

The key query block extraction process is as follows: Based on the correlation between some statistical characteristics of words and long query sentences, the word block with the highest correlation is used as the query keyword [27, 28].

Step 4: Query extension. In Chinese, there are a large number of synonyms and polysemy words, which make a great difference between the user search words and the "key" words used in the multimedia assisted legal classroom teaching resource index. Here, the query expansion technology is used to solve the problem; that is, the computer is used to find the word blocks with similar meanings of key query blocks and then take

Table 2: Comparison of B-tree and B+ tree.

Indexes	B-tree	B+ tree
Difference	A leaf node does not contain any information and does not point to its sibling node	All leaf nodes contain the information of all keywords and pointers to the records containing these keywords, and the leaf nodes themselves are linked in order of the size of the keywords from small to large
Advantage	The efficiency of interpartition query is high, the space utilization is high, and the average time is small	The whole index tree structure has few layers and stable performance
Inferiority	The search range is limited	There is an additional burden of insertion and deletion, and the space utilization is low

Table 3: Word list.

Resources	Character
Of	P_{11} P_{12} P_{13}.....
No	P_{21} P_{22} P_{23}.....
:	:
In	P_{i1} P_{i2} P_{i3}.....
Number	P_{j1} P_{j2} P_{j3}.....

them as new query word blocks for information retrieval [29–31].

Step 5: The query word block is matched with the multimedia assisted legal classroom teaching resource index to realize information retrieval. The matching method is realized by similarity calculation. There are four similarity calculation methods: inner product measurement method, cosine measurement method, Jaccard measurement method, and Dice coefficient method.

2.4. Establishment of Data Distribution Structure Model of Multimedia Assisted Legal Classroom Teaching System. In order to realize the specific data mining in the multimedia assisted legal classroom teaching system, data structure analysis and data fusion preprocessing are needed first. The research object is the multimedia distance teaching system in the cross platform network environment. Based on the high-speed operation and data management results of the system users in all terminals and servers, the data structure in the model is analyzed [32]. Suppose v represents any node in the data perception layer of the multimedia assisted legal classroom teaching system in the model (i.e., $v \in V$), e represents any edge of the application business layer of the multimedia teaching system, and the specific data sources in the multimedia assisted legal classroom teaching system are distributed in the data management array of the system in the form of edge node [33, 34]. The node frame of data transmission is divided into $N \times N$ unit line array, and the channel bandwidth is $T_s = K_b T_f$. Suppose V_{M_j} represents the length of information flow vector in the database of embedded multimedia assisted legal classroom teaching system; $E\{[X - E(X)][Y - E(Y)]\}$ is the covariance of fuzzy classification random variables X and Y, recorded as $\mathrm{Cov}(X, Y)$; the attribute weights of the data storage model in the multimedia assisted legal classroom teaching system are adaptively learned; and the vector field X of specific data distribution in the database is constructed in the multilayer vector autoregressive feature space. In the multimedia assisted legal classroom teaching system, the local outliers will produce cross data sets through (4) to establish local outlier information flow model in specific data mining:

$$D_w(S) = \frac{\sum_{u,v \in S} V_{M_j}}{|S| \times (|S| - 1)/2}. \tag{4}$$

In the formula, $w_{u,v}$ represents the weight between nodes u and v in the network cluster of the multimedia assisted legal classroom teaching system, and $|S|$ refers to the number of intersection nodes of data characteristics in undirected graph S. Using the above outlier information flow model, this paper constructs a group of dense subgraphs with intersection, to reflect the characteristics of cross data sets in the multimedia assisted legal classroom teaching system, and extracts the time sequence of specific big data information flow in the multimedia assisted legal classroom teaching system. Assuming that y represents the fuzzy clustering center of the specific data feature vector of the multimedia assisted legal classroom teaching system, we use the fuzzy clustering center to classify the difference attributes of the specific data, and an undirected graph model $G = (V, E)$ of the data distribution structure in the multimedia assisted legal classroom teaching system under the cross platform network application support layer is established, that is, the data distribution structure diagram model of multimedia assisted legal classroom teaching system, as shown in Figure 3.

2.5. Data Mining Processing of Information. Based on the above data structure analysis, the scale affine transformation is used to mine the information. The specific steps are as follows:

The continuous wavelet transform is used to decompose the empirical modal features of the specific data distributed in the above model, and a set of two-dimensional functions representing the time scale a and time translation b of the internal details of the specific data in the multimedia assisted legal classroom teaching system are obtained; formula (5) is used to perform probabilistic feature decomposition of specific data in the teaching system on the two-dimensional projection plane. The decomposition formula is as follows:

$$b_j = r_k(v + 1) \times \mathrm{parity}(z). \tag{5}$$

It is assumed that $r_k(v + 1)$ is the nonlinear time series of specific data in the collected multimedia assisted legal

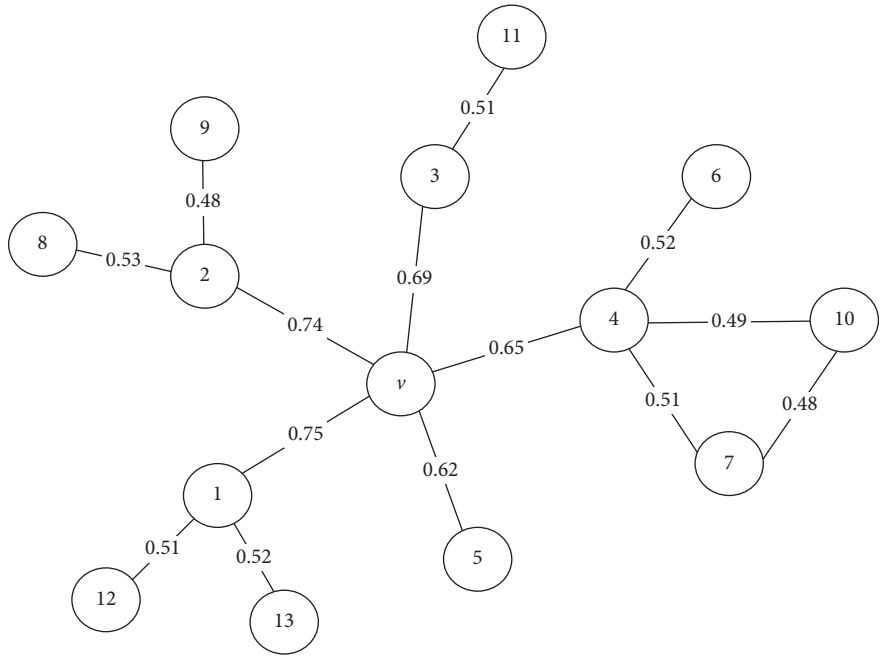


Figure 3: Data distribution structure model of multimedia assisted legal classroom teaching system.

classroom teaching system and that parity (z) represents the mother wavelet of data sampling. When the specific data $x(k) = s(k) + w(k)$ in the collected multimedia assisted legal classroom teaching system is a quasi-stationary random information flow, the edge subband level information of characteristic data distributed in the teaching system database is obtained by the following formula:

$$\Phi(\omega) = \int_{-\infty}^{+\infty} f(x) e^{(j\omega x)} \mathrm{d}x = \mathrm{parity}\left(q_p(z)\right) r_k(v+1). \tag{6}$$

According to the distribution probability of specific multimedia data, the data with large distribution probability is subject to scale affine transformation, and the information fusion model of data mining under limited signal-to-noise ratio is established:

$$\psi_{a,b}(t) = [U(a,b)\psi(t)] = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right). \tag{7}$$

In the formula, $U(a,b)$ is the fractional Fourier transform of the time domain part of the specific data in the multimedia assisted legal classroom teaching system, and the factor $1/\sqrt{|a|}$ ensures the energy normalization of the fractional Fourier transform. Because the students' classroom learning behavior is most directly related to the learning effect, in this case, the relationship between the learning behavior of a certain type of students and the performance data of the test results is processed through information data mining, so as to provide data analysis and reference for teachers, so that teachers can understand the close relationship between students' learning habits and learning performance, fully understand the characteristics of individual students, and realize the information fusion of specific data in the multimedia assisted legal classroom teaching system through the above processing.

2.6. Implementation of Multimedia Assisted Legal Classroom Teaching Model. Based on the above analysis of the data distribution structure model of the multimedia assisted legal classroom teaching system and the data mining processing of information by scale affine transformation, an improved design of data mining algorithm is carried out, and accurate information feature extraction and mining are carried out for the specific key data in the multimedia teaching system [35], and the information compatibility and data access ability of multimedia assisted legal classroom teaching system can be improved. To realize the construction of multimedia assisted legal classroom teaching model, the implementation steps are as follows:

Assuming the specific data fusion optimal weight coefficient β and penalty factor C in the multimedia assisted legal classroom teaching system, we construct the specific data feedforward gain adjustment mean square deviation function by using the following formula:

$$F_{\mathrm{fitness}} = \frac{1}{m} \sum_{i=1}^{m} (\beta - C)^2. \tag{8}$$

The data management in the multimedia assisted legal classroom teaching system is realized through the distributed database model. A large amount of cloud data is integrated into the distributed database. It is necessary to use the feedforward gain to adjust the mean square deviation function to optimize the structure of the cloud data in the distributed database in the multimedia assisted legal classroom teaching system, so as to improve the accuracy and

efficiency of data information mining. The phase space reconstruction method of (8) is used to decompose the attenuation characteristics of high-frequency components of specific data stream information:

$$F(\tau) = b_k \times \phi \times x(t) \times F_{\text{fitness}}. \quad (9)$$

In the formula, τ represents the sampling delay of phase space reconstruction for a specific data information stream, b_k is the embedding dimension of phase space reconstruction, ϕ is the phase difference of sampling interval, and $x(t)$ is the spatial directivity gain adjustment coefficient. In the nonuniform sampling output results, it is assumed that $y(k)$ is the approximate statistical average of specific data in the multimedia assisted legal classroom teaching system obtained after multidimensional parameter mixed estimation. According to the invariant characteristics of Gaussian random linear separation of specific data [36], the invariant features of specific data are mined in the high-dimensional phase space to improve the frequency domain focusing ability of the signal. We use the above process to obtain the multimedia assisted legal classroom teaching model:

$$y(t) = x(t - t_0) \Rightarrow W_y(t, v) = W_x(t_0, v - v_0). \quad (10)$$

In the formula, $x(t)$ is the singular value decomposition result of the original data, t_0 is the initial sampling time point, $W_y(t, v)$ is the observation vector, v_0 is the power spectral density of the specific data, and v is the interference intensity of the data in the teaching system. The high-order moment characteristics of specific data in the multimedia assisted legal classroom teaching system can reflect the specific data characteristics to a great extent. Through the positioning of the high-order moment characteristics, the specific data mining in the multimedia assisted legal classroom teaching system can be realized, so as to complete the construction of the multimedia assisted legal classroom teaching model based on the data mining algorithm.

3. Experimental Analysis

In order to verify the effectiveness and feasibility of the multimedia assisted legal classroom teaching model based on data mining algorithm proposed in this paper, an experiment is carried out. During the experiment, MATLAB software is used as the experimental simulation software. The experimental topology is shown in Figure 4. Through the experiment, the results are as follows.

Taking the law classroom as an example, we select 2021 students majoring in law in a school as the experimental object, with a total of 240 students, and the number of students in each class is 40. Among these six classes, two classes adopt the proposed multimedia assisted legal classroom teaching model based on data mining algorithm, recorded as group A; two classes were recorded as group B by the method of [8]; and the other two classes were recorded as group C by the method of [9]. The proposed teaching mode of multimedia assisted legal classroom teaching model, based on data mining algorithm, is based on the data distribution structure model of the established multimedia assisted legal classroom teaching system. The teaching mode flow is shown in Figure 5.

According to the application of the teaching mode process of the model proposed in Figure 5, the majority of teachers have more modes to use for reference in teaching, so that the teaching theory can be applied quickly and successfully. On the other hand, learning, mastering, and applying the teaching mode enable teachers to learn, understand, and master the multimedia assisted legal classroom teaching theory. Teaching mode plays a bridge and link role between teaching theory and teaching practice. The research and application of teaching model process will help to change the separation between teaching theory and practice and promote the research of teaching theory.

The teaching materials used by the three groups of students are the same. The teaching places of the first two teaching methods are multimedia classrooms, and each student is equipped with a computer. In the process of law teaching, teachers grade each student through grading test.

Legal tests were conducted for the three groups of students. In the legal test part, the real questions of 2019 legal examination are selected as the midterm test questions, and the real questions of 2020 legal examination are selected as the final test questions. The total score of each set of test papers is set to 20 points. After marking the test paper, the scores of all students are input into the statistical software for statistical analysis, and the standard deviation of students' scores is calculated using the standard deviation calculation formula as follows:

$$s = \sqrt{\left(\frac{1}{n}\right)\left[(x_1 - \overline{x})^2 + (x_2 - \overline{x})^2 + \cdots + (x_n - \overline{x})^2\right]}. \quad (11)$$

In formula (1), s represents the standard deviation of student performance, n represents the number of subjects, $\overline{x}$ represents the average of student performance, and x_n represents the n student. Through statistical analysis and calculation, the test results are shown in Table 4.

As shown in Table 4, through statistical analysis, in the midterm test, the average scores of groups A, B, and C students in the legal test are 3.55, 3.53, and 3.53, respectively, and the standard deviations are 0.24, 0.28, and 0.23 respectively. Through statistical analysis, there is no significant difference among the three groups of students. In the final test, the average score of the legal test of group A students is 9.24, while the average score of the legal test of group B students is only 7.15. The average score of group A students is significantly higher than that of group B students. The average score of group C students was only 3.96, which was not significantly improved compared with the previous one. Through comparison, it is found that the scores of group A and group B students have improved after adopting two multimedia network law learning and teaching methods. However, the students who adopt the proposed multimedia assisted legal classroom teaching model based on data mining algorithm have made more obvious progress than those who adopt the methods of [8] and [9]. The experimental results show that compared with the other two legal learning teaching methods, the proposed model has more advantages in terms of teaching effect. The reason is that, by

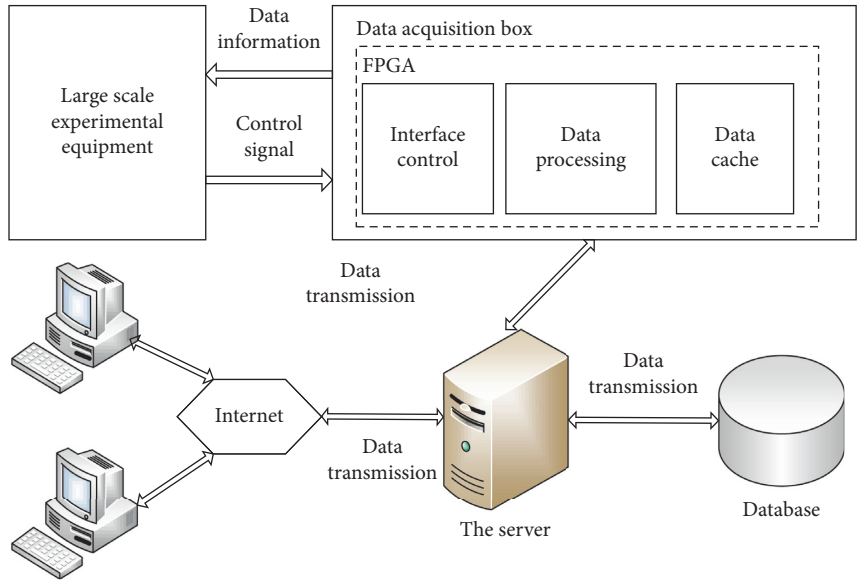


Figure 4: Experimental topology.

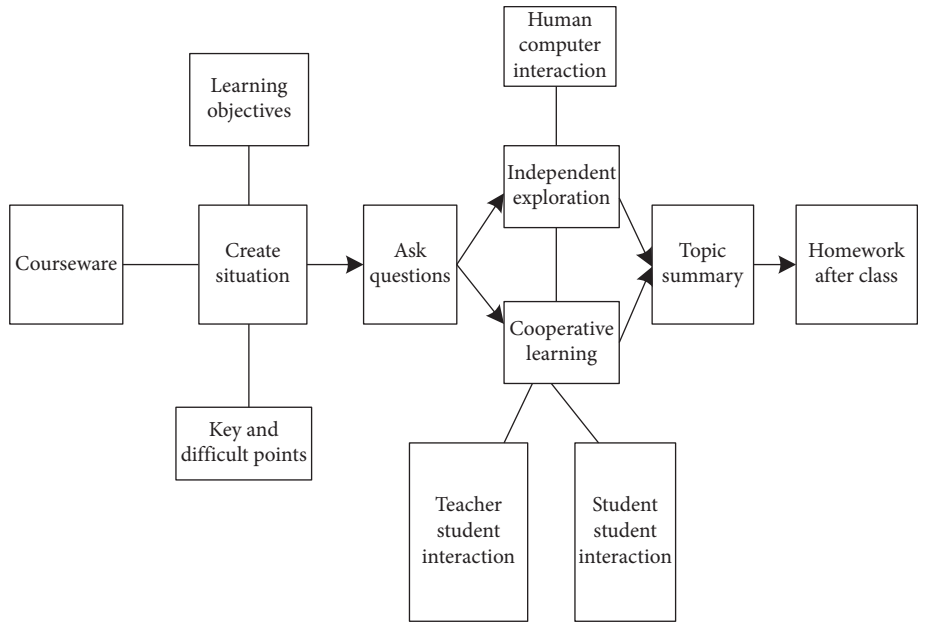


Figure 5: Flow chart of teaching mode of the proposed model.

Table 4: Test results of legal examination.

Law examination test	Student group	Average	Standard deviation
Interim test	Group A	3.55	0.24
	Group B	3.53	0.28
	Group C	3.53	0.23
Final test	Group A	9.24	0.28
	Group B	7.15	0.25
	Group C	3.96	0.24

focusing on the students themselves, it improves their legal learning effect and contributes to the improvement of students' performance.

In order to further verify the comprehensive effectiveness of the proposed multimedia assisted legal classroom teaching model, the experiment selects 30 law students (15 men and 15 women) with similar academic achievements in a university as the research object and makes an empirical analysis from the two aspects of students' learning efficiency and teaching quality.

The 30 students selected in the experiment use different multimedia assisted legal classroom teaching models to study the legal practice course. The experiment divides the 30 students into three groups, with five men and five women in each group. After the model learning, the students are tested in law, and the learning score is used to determine whether the

Table 5: Comparison results of teaching quality of different models.

Test object	Student test scores (points)		
	Proposed model	Reference [8] method	Reference [9] method
Male 1	98	87	78
Male 2	97	82	85
Male 3	89	91	89
Male 4	93	85	81
Male 5	92	83	87
Female 1	94	83	92
Female 2	93	82	87
Female 3	92	82	79
Female 4	94	83	80
Female 5	92	83	83

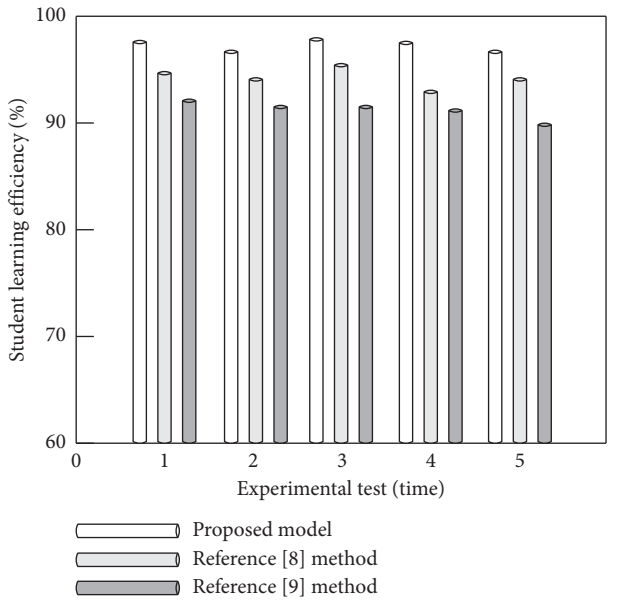


Figure 6: Comparison of students' learning efficiency of different models.

model can effectively improve the teaching quality, where 100–85 points mean excellent, 84–60 points mean qualified, and points below 60 mean a failure. The specific experimental comparison results are given in Table 5.

By analyzing the experimental data in Table 5, it can be seen that after the legal practice course teaching of the proposed model, the students' legal test scores are significantly higher than those of the other two methods, and all students' scores are excellent, which shows that the proposed model can obtain more ideal teaching quality. At the same time, in the process of practical research, the proposed model also effectively solves the problem of information loss of legal practice course and lays a solid foundation for further improving the teaching quality.

Because the teaching assistance effects of different multimedia assisted legal classroom teaching models are inconsistent, the experiment focuses on comparing the changes of students' learning efficiency after learning each multimedia assisted legal classroom teaching model. The specific experimental comparison results are shown in Figure 6.

By analyzing the experimental data in Figure 6, it can be seen that the students' learning efficiency of the proposed model is significantly higher than that of the other two methods, because the proposed model effectively solves the problem of information loss of legal practice course and ensures the integrity of teaching scheme, so as to improve students' learning efficiency.

To sum up, after adopting the multimedia assisted legal classroom teaching model based on data mining algorithm, the scores of group A and group B students have improved, and the scores of all students are excellent. The students' learning efficiency of the proposed model is significantly higher, indicating that the proposed model can obtain more ideal teaching quality.

4. Conclusion

(1) The students who use the multimedia assisted legal classroom teaching model based on data mining algorithm have more obvious progress than the students who use the literature method, and the teaching effect has more advantages, which can promote the improvement of students' grades.

(2) After teaching the legal practice course with the proposed model, all students have excellent grades, and the proposed model can obtain ideal teaching quality.

(3) Because the proposed model effectively solves the problem of information loss of legal practice course, ensures the integrity of teaching scheme, and achieves the purpose of improving students' learning efficiency, the students' learning efficiency of the proposed model is obviously higher.

Building a multimedia assisted legal classroom teaching model based on data mining algorithm can provide teaching information services for teachers and students. At the same time, using data mining algorithm to collect and analyze the behavior data of teachers and students can help to better find out the key materials affecting the practicability and expansibility of classroom teaching, better carry out multimedia assisted legal classroom teaching, promote the sound development of multimedia platform, contribute positive influence value to legal classroom teaching, and promote the

high-quality and high-level development of multimedia assisted legal classroom teaching.

Data Availability

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding this work.

References

[1] D. Simon, "Laughing your way to academic success: can laughter impact learning and well-being in the law school classroom and are there cross-cultural differences?" *SSRN Electronic Journal*, vol. 1, pp. 1–10, 2020.

[2] K. S. Odeku, "Conducting law pedagogy using virtual classroom in the era of COVID-19 pandemic: opportunities and existing obstacles," *Journal of Educational and Social Research*, vol. 11, no. 1, pp. 101–111, 2021.

[3] A. M. Kupenda, "Collaborative learning in the constitutional law classroom: adapting the concept of inevitable disagreement in seven steps," *Journal of Legal Education*, vol. 68, no. 2, pp. 6–16, 2019.

[4] S. Bancroft-Billings, "Identifying spoken technical legal vocabulary in a law school classroom," *English for Specific Purposes*, vol. 60, pp. 9–25, 2020.

[5] S. Lanham, "Analyzing big data with benford's law: a lesson for the classroom," *American Journal of Business Education*, vol. 12, no. 2, pp. 33–42, 2019.

[6] Y. Vasylvna Yulnetska, O. Yurvna Babi, R Vasilvna Glovyuk, A. Sergeevich Stepanenko, and M. Ivanovich Pashkovsky, "Implementing ict into language and law classroom to develop law students' communicative competence," *Information Technologies and Learning Tools*, vol. 81, no. 1, pp. 310–326, 2021.

[7] A. A. Periola and O. A. Osanaiye, "Low cost intent driven future multi-media content access system and network," *Multimedia Systems*, vol. 14, no. 4, pp. 12–20, 2021.

[8] S. Chen, "Application of computer information technology in the reform of law classroom teaching method," *IOP Conference Series: Materials Science and Engineering*, vol. 750, no. 1, Article ID 012061, 2020.

[9] Y. Wei, "Leading-in methods of junior middle school moral and rule of law classroom teaching," *Jiangsu Education Research*, vol. 17, pp. 55–57, 2019.

[10] F. Batlan, "Beyond greed is good: pop culture in the business law classroom," *Journal of Legal Education*, vol. 68, no. 1, pp. 5–15, 2019.

[11] D. B. Kraft, "Contrastive analysis and contrastive rhetoric in the legal writing classroom," *New Mexico Law Review*, vol. 12, no. 6, pp. 15–21, 2019.

[12] X. Zuo, Z. Chen, L. Dong, J. Chang, and B. Hou, "Power information network intrusion detection based on data mining algorithm," *The Journal of Supercomputing*, vol. 76, no. 7, pp. 5521–5539, 2020.

[13] H. M. van Loo, T. B. Bigdeli, Y. Milaneschi, S. H. Aggen, and K. S. Kendler, "Data mining algorithm predicts a range of adverse outcomes in major depression," *Journal of Affective Disorders*, vol. 276, no. 11, pp. 945–953, 2020.

[14] J. Hu, J. Fang, Y. Du, Z. Liu, and P. Ji, "Application of PLS algorithm in discriminant analysis in multidimensional data mining," *The Journal of Supercomputing*, vol. 75, no. 9, pp. 6004–6020, 2019.

[15] M.-R. Kazemzadeh, A. Amjadian, and T. Amraee, "A hybrid data mining driven algorithm for long term electric peak load and energy demand forecasting," *Energy*, vol. 204, no. 8, Article ID 117948, 2020.

[16] C. Yuan, X. Yu, D. Li, and Y. Xi, "Overall traffic mode prediction by VOMM approach and AR mining algorithm with large-scale data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 4, pp. 1508–1516, 2019.

[17] S. Kawtar, J. Damien, G. Roger et al., "A data mining approach for improved interpretation of ERT inverted sections using the DBSCAN clustering algorithm," *Geophysical Journal International*, vol. 225, no. 2, pp. 1304–1318, 2021.

[18] X. Zhang and B. Wang, "Design of estimation algorithm of island intelligent tourist volume based on data mining," *Journal of Coastal Research*, vol. 95, no. 1, pp. 985–996, 2020.

[19] L. T. T. Nguyen, V. V. Vu, M. T. H. Lam et al., "An efficient method for mining high utility closed itemsets," *Information Sciences*, vol. 495, no. 8, pp. 78–99, 2019.

[20] Z. Fang, X. Yang, L. Han, and X. Liu, "A sequentially truncated higher order singular value decomposition-based algorithm for tensor completion," *IEEE Transactions on Cybernetics*, vol. 49, no. 5, pp. 1956–1967, 2019.

[21] Q.-X. Zhu, C. Jin, Y.-L. He, and Y. Xu, "Pattern mining of alarm flood sequences using an improved prefixspan algorithm with tolerance to short-term order ambiguity," *Industrial & Engineering Chemistry Research*, vol. 60, no. 11, pp. 4375–4384, 2021.

[22] M. Fahimeh, H. Pérez-Sánchez, M. Alireza, F. Afshin, and G. Fahimeh, "Accelerating big data quantitative structure-activity prediction through lasso-random forest algorithm," *Bioinformatics*, vol. 2, no. 10, pp. 1–10, 2021.

[23] W. Lin, Z. Wu, and H. Zhou, "Research on mining association between weather and power grid based on Apriori algorithm," *Electrical Engineering*, vol. 20, no. 6, pp. 44–49, 2019.

[24] Y. Tang, "A mnotone data augmentation algorithm for longitudinal data analysis via multivariate skew-t, skew-normal or t distributions," *Statistical Methods in Medical Research*, vol. 29, no. 6, Article ID 096228021986557, 2019.

[25] W. Xiao and J. Hu, "SWEclat: a frequent itemset mining algorithm over streaming data using Spark Streaming," *The Journal of Supercomputing*, vol. 76, no. 5, pp. 7619–7634, 2020.

[26] L. Ribeiro, P. Rosário, I. Moreira, and R. Serrão Cunha, "First-year law students' and teacher's questioning in class," *Frontiers in Psychology*, vol. 21, no. 3, pp. 1–10, 2019.

[27] Y. Levin-Schwartz, C. Gennings, B. C. Henn et al., "Multimedia biomarkers: integrating information to improve lead exposure assessment," *Environmental Research*, vol. 183, no. 4, Article ID 109148, 2020.

[28] A. Dyevre and M. Ovádek, "Experimental legal methods in the classroom," *SSRN Electronic Journal*, vol. 16, no. 1, pp. 1–12, 2019.

[29] C. Jomdecha, X. Li, W. Cai, and X. Shejuan, "Advanced multimedia element for simulating distribution of magnetic flux density influenced by narrow crack," *IEEE Transactions on Magnetics*, vol. 55, no. 99, pp. 1–4, 2019.

[30] J. Torres Gomez, A. Rodriguez-Hidalgo, Y. V. Jerez Naranjo, and C. Pelaez-Moreno, "Teaching differently: the digital signal processing of multimedia content through the use of liberal arts," *IEEE Signal Processing Magazine*, vol. 38, no. 3, pp. 94–104, 2021.

[31] D. Wu, H. Shen, and Z. Lv, "An artificial intelligence and multimedia teaching platform based integration path of IPE and IEE in colleges and universities1," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 115, pp. 1–10, 2020.
[32] Y. Liu, "Design and implementation of multimedia teaching platform based on SOA architecture," *Multimedia Tools and Applications*, vol. 79, no. 3, pp. 10899–10914, 2020.
[33] H. Fu and W. Fu, "Research on the influence of multimedia on Chinese teaching in senior high school," *World Scientific Research Journal*, vol. 6, no. 5, pp. 86–94, 2020.
[34] X. U. Zhe, "Research on strategies of improving the experience of college teachers for multimedia teaching equipment," *Journal of Guizhou Radio & Television University*, vol. 27, no. 1, pp. 62–67, 2019.
[35] L. Tong, "The probe on the strategy of multimedia teaching in Post-MOOC Times," *Sino American English Teaching*, vol. 16, no. 7, pp. 306–309, 2019.
[36] Z. Q. Xu and T. Zhang, "Simulation of a ccurate classification method for digital multimedia scene images," *Computer Simulation*, vol. 36, no. 7, pp. 385–388, 2019.

Hindawi
Scientific Programming
Volume 2021, Article ID 3429463, 10 pages
https://doi.org/10.1155/2021/3429463

Research Article

Research and Design of Automatic Scoring Algorithm for English Composition Based on Machine Learning

Yu Zhao

School of International Studies, University of Science and Technology Liaoning, Anshan City 114000, Liaoning Province, China

Correspondence should be addressed to Yu Zhao; zhaoyu963@ustl.edu.cn

Received 15 November 2021; Revised 29 November 2021; Accepted 6 December 2021; Published 23 December 2021

Academic Editor: Baiyuan Ding

With the development of artificial intelligence and big data, the concept of "Internet plus education" has gradually become popular, including automatic scoring system based on machine learning. Countries all over the world vigorously promote the deep integration of information technology and discipline teaching in various fields. English is a medium of communication in the current era of education information development trend. English composition automatic scoring mode is gradually accepted by the majority of educators and applied in the actual classroom teaching. However, the research of English composition automatic grading in teaching space is not perfect. Most systems have used traditional algorithms. Therefore, this paper constructs the automatic scoring algorithm and sentence elegance feature scoring algorithm of English composition based on machine learning, explores the influence of the algorithm on English writing teaching, and proves the correctness of the design idea and algorithm of this paper through a lot of experiments.

1. Introduction

English is one of the main international languages and the most widely used language in the world. Nowadays, the process of internationalization is accelerating. As an international lingua franca, English is a basic skill for students to go to the world [1]. Learning English can not only enrich their knowledge but also broaden their horizons and improve personal advantages. In the process of English learning, English composition writing is indispensable. Writing can not only reflect learners' grasp of basic knowledge, such as vocabulary and grammar, but also reflect learners' overall grasp of sentence structure, discourse structure, and logical reasoning ability [2]. As English composition can directly reflect the comprehensive language ability of English learners; writing ability is the key item of the test organizers to investigate the language ability of students. Thus, more and more learners pay attention to it [3]. At present, writing is still a weak link in students' English learning, and it is also a link that teaching staff need to spend a lot of time preparing. In usual learning, teachers generally manually review English compositions, and it is suitable for a small number of short ones [4]. However, when the intensity of marking is high, and the length of essays is long, this method will lead to low efficiency and high rate of misjudgments. More importantly, the subjectivity of manual assessment is too strong, and people with different knowledge and regions may have a great bias on the same essay.

In recent years, the rapid development of computer technology has led to the prosperity of other industries, which has promoted the research and application of Automated Essay Scoring (AES) technology in education [5]. On the one hand, this technology can intelligently analyze and grade compositions. Compared with manual evaluation, computer evaluation has low cost and high review efficiency. The automatic scoring system gives full play to the characteristics that computers are good at repetitive work, which can liberate teachers' physical and mental strength to a great extent and allow teachers to devote more energy to teaching and research. On the other hand, the analysis results can provide feedback regarding more evaluation information, such as spelling and grammatical errors of each composition, so that students can initially modify their own articles according to systematic suggestions. The analysis results can

also recommend excellent words and sentences and composition materials to provide more scientific writing guidance for learners. At present, the construction and evaluation of AES system are mostly based on the statistical information of the content of the paper. The research content is relatively simple, and the evaluation of the logic of the composition and the quality of words and sentences is still not in-depth and accurate. Therefore, while improving the accuracy of the predicted score, we should also comprehensively evaluate the composition content, so that the AES system can be better applied to the actual composition correction.

2. Related Work

The research on automatic scoring in education industry started earlier, and there were many systematic studies on different subjects and languages. The research on the scoring system related to composition can be traced back to the 1960s. The earliest scoring system Project Essay Grader (PEG) [6] developed by Professor Ellis Page took the features extracted from shallow linguistic methods (such as article length, word length, punctuation marks, grammar, etc.) as independent variables. The AES model is trained by multiple linear regression with the score due to the composition as the dependent variable. This method does not involve the language content and composition structure, so the evaluation results are biased. Subsequently, Landauer Thomas et al. developed an automatic Essay scoring system named Intelligent Essay Analysis (IEA) [7] based on Latent Semantic Analysis (LSA) [8]. When the model is used to score students' English compositions, the compositions to be graded and excellent examples are mapped to the vector space, and the scores of compositions are predicted by comparing the similarity values. This approach takes the overall content of the article into account and is more accurate than PEG. It can also be used to detect plagiarism in the article, which has greatly advanced the field of automatic grading.

With the deepening of the research on automatic scoring system, American Educational Examination Institute developed the E-rate system in the 1990s [9]. It used natural language processing technology and statistical methods to mark the part of speech of each word through a part-of-speech tagger and analyzed the syntactic structure of the text through syntactic analysis. These methods can be used to evaluate the quality of writing language, content, and text structure and have been applied to the automatic scoring of GMAT and GRE. Although the design of E-Rater is more comprehensive, it is not as comprehensive as PEG in language analysis and not as in-depth as IEA in content analysis, so there is a lot of room for improvement. The AES system developed by Professor Liang's team in China [10] is based on superficial linguistic features and linear regression model training to analyze the accuracy of word spelling and grammar usage in each sentence. However, it failed to give students more evaluation results in terms of discourse quality, sentence quality, and relevance. Subsequently, Wang et al. [11] improved the automatic scoring effect from the perspective of semantic dispersion and introduced the convolutional neural network training model, which showed good performance in composition prediction ability. Qiu [12] evaluated the fluency of the composition and quantified it into AES model to improve the scoring effect of the system. Liu et al. [13] considered figurative parallelism and other rhetorical devices in Chinese compositions and constructed a corpus of ancient poems to automatically identify ancient poems in compositions, which has better accuracy compared with the benchmark system. Yangwei and Huang used Auto-Encoder (AE) [14] to reconstruct linguistic features and then input the reconstructed feature vectors into SVM for regression training, which improved the performance compared with its previous reconstruction.

With the rapid development of intelligent hardware, artificial intelligence has made rapid progress, including natural language processing. Natural language processing based on deep learning can be summarized as the problem of original data feature representation in application field and the problem of selecting appropriate deep learning algorithm to construct application model. For the former problem, mature models include bag-of-words (BOW) [15] and Vector Space Model (VSM) [16]. These methods have certain defects. The word bag model, such as one-hot Encoding, will also become very large and sparse when the number of categories is large. Vector space models such as Term Frequency-Inverse Document Frequency (TF-IDF) [17] characterize text features by calculating the probability of words becoming keywords in the text. However, this method is greatly influenced by the global text set and only makes use of the statistical information of words. The location information and context information are not utilized, so the text features cannot be fully represented. Bengio et al. [18] used deep neural network to construct a language model, which could map words into a vector space of fixed dimensions, solving the problems of sparse features and large dimensions caused by one-hot coding. By training neural network language model through unsupervised learning, semantic information contained in the text could be obtained. Its disadvantage is that it involves a lot of parameter training, which leads to a long training cycle.

Mikolov et al. proposed word2vec [19], a word vector training model in 2013. This model is based on Continuous bag-of-words (CBOW) and Skip-Gram models. The former can predict the probability of the occurrence of the current word based on the semantic information before and after the word. The latter is the most widely used word vector representation model, which uses the current word to predict the probability of the occurrence of the preceding word. Mikolov then proposed the Doc2veC model based on paragraph vector, which mainly added paragraph vector to the word2vec model and also included two model structures: Distributed Memory Model and Distributed Bag-of-Words, which can represent sentences and text. Jeffrey proposed the Glove word vector representation model in 2014 [20], which accelerated the training of word vector and enriched the semantic information of word vector. The Embedding from Language Model (ELMO) was proposed by Peters in March 2018 [21], which adopted the double-layer bidirectional

LSTM structure pretraining model. Then, Word Embedding dynamically adjusts the representation of corresponding words according to the context of the words in the input sentence, which can solve the problem of polysemy. In October 2018, Jacob Devlin et al. proposed the Bidirectional Encoder Representations from Transformers (BERT) representations from Transformers language representation model [22]. The bidirectional encoder of Transformer is used to pretrain the model based on the context of all layers. Fine-tuning an output layer can create an optimal model for downstream tasks, which is the best language representation model at present.

Researchers generally believe that natural language has certain logical and recursive characteristics, and language contexts are closely related. For example, sentences in natural language are actually composed of words and phrases recursively. Therefore, recursion is an important feature of natural language. Therefore, models can be selected according to the characteristics of natural language, such as recurrent neural network, convolutional neural networks, and a series of improved models [22]. KIM proposed to use text-CNN model for sentence classification in 2014 [23] and proved that convolutional neural network plays an important role in extracting text features. In 2017, Liu et al. [24] proposed a combination model of deep neural network DC-NN, which used an improved recursive neural network to generate phrase pair semantic vectors suitable for phrase generation process and used an auto-encoder to improve the performance of phrase generation process. It performs better than baseline models on machine translation tasks. Hassan and Mahmood [25] proposed the training method of integrating CNN and RNN and obtained the convolution layer of long-term dependence through joint training, achieving an accuracy of 93.3% in emotion analysis. LSTM, first proposed by Schuster and Paliwal [26], was an improved model based on cyclic neural network, which selectively stored and forgot long-term memories through memory units. Wang et al. [27] applied the attention mechanism to convolutional neural network to extract features and integrated them into the LSTM model, which showed excellent features in Tweet sentiment analysis. Hochreiter and Schmidhuber [28] completed the multilabel classification task of books based on long-short-term memory, with good performance in various indicators. In view of the fact that LSTM model can only remember the information above and does not make full use of the information below, Schuster and Paliwal [26] proposed bidirectional LSTM model (BiLSTM), which added a reverse layer to LSTM model, and allowed it to use context information to improve model performance. Zeng et al. [29] used the bidirectional LSTM model to obtain the bidirectional semantic information of comments to achieve the task of sentiment classification of comment texts, which has a better effect compared with convolutional neural network. Although many researches have been done in the field of automatic scoring, the accuracy and comprehensiveness of scoring are not thorough enough. In this paper, composition features are extracted from multiple dimensions, and appropriate technical models are selected through specific analysis to extract features and train models, so as to obtain the best performance of the scoring model.

3. Network Framework

3.1. Basic Feature Extraction

3.1.1. Based on Linguistic Features. Feature extraction is an important process of machine learning modeling, which determines the quality of model results. Linguistic features refer to the use of statistical methods to extract shallow features of composition words and sentences without considering the meaning of the text, such as the proportion of part of speech and number and length of words and sentences. Simply speaking, linguistic features are the direct application of statistics. The early AES models, such as PEG scoring systems, basically used superficial linguistic features such as text length, number of sentences, part of speech, and other surface features to make models and then used regression equations to train the model and build AES system. These linguistic features have some effect. These linguistic features have some effect. The statistical results based on words and sentences can well reflect the complexity of words and sentences, so as to reflect the author's ability to use words and sentences. For example, the total number of words, the number of words after repetition, and the number of high-frequency words reflect the author's vocabulary. The number of clauses, phrases, etc. reflect the author's ability to use complex sentence patterns, so we use morphological and syntactic statistical features, and the details are shown in Table 1.

3.1.2. Features Based on Semantic Expression. Features of semantic expression are not only to consider statistical features but also to consider deep meanings of language texts. The common ones are to get the word vector clustering features according to composition words vector set under different clustering number distribution, to get text vector features of text vector from the text level, and to get theme distribution characteristics of the theme probability distribution of compositions through training the LDA model.

(1) Word Vector Clustering Model Based on word2vec. The word vector clustering model based on word2vec feature extraction is one of the most classic maturity models. The first step is to build a word vector by using word2vec tool based on the fixed code words, which means that each word is mapped to a fixed dimension of feature space, thus obtaining term vectors. This process needs to set the word vector length and receive a word vector corresponding to the whole composition. For the second time, k-means clustering method is used to cluster the word vectors, and "optimal fitness" is adopted to select the number of clustering clusters. The main idea of "optimal fitness" is to minimize the distance between word vectors in the cluster and maximize the distance between cluster centroids [30, 31]. The number distribution of word vectors under different clustering clusters is counted, and the returned number distribution is regarded as the clustering feature of word vectors. The

Table 1: Text feature description.

Overview of lexical features	Overview of syntactic features
The proportion of word list size to composition Length (not repeated)	Average sentence length and variance
Number of sentences whose length is greater than a fixed value (e.g., 4, 8, 12)	Number of sentences whose length is greater than a fixed value (e.g., 4, 8, 12)
Statistical characteristics such as average character length (mean word length, median, standard deviation)	Average number of verbs, nouns, modal verbs, prepositions in sentences
The proportion of nouns, adjectives, verbs and prepositions	Average number of punctuation marks in a sentence
Number of high-frequency words	The number of sentences that fully express the meaning
The size of the word list after removing the stop word	The number of clauses and the average length of clauses

clustering model trained based on this idea can show better robustness and higher performance. The core formula is as follows:

$$\alpha = \frac{1}{k}\sum_{1\le p<q\le k}^{k}\cos\left(s_p, s_q\right), \beta = \frac{1}{k}\sum_{b=1}^{k}\frac{1}{t_b}\sum_{1\le i<j\le t_b}^{t_b}\cos\left(w_i, w_j\right), f(x)\frac{\alpha}{\beta}, \tag{1}$$

where s_p, s_q are the centroid of clustering clusters p and g of composition word package. k represents clusters number of word package, α is the average distance between cluster, t_b is the first b clustering clusters to the total number of words, w_i, w_j are for the i and j word vectors, β is for all the average distance between cluster-heads clustering vector, $f(x)$ is related to the number of clustering cluster of fitness function, and adjust the k values to make the value of fitness function be the largest number. At this point, the average distance between clusters should be as large as possible, and the average distance between samples in clusters should be as small as possible. After such a series of operations, the number distribution of words belonging to each cluster in the composition can be counted, so as to characterize the word vector features of the composition.

(2) Text Vector Features Based on Doc2vec. Although word2vec has greatly promoted the development of natural language processing, the common sentence representation methods based on word2vec include joining, adding, averaging, local maximum, and minimum values. These methods use simple superposition averaging and other relatively rough methods, without considering the influence of context information on vocabulary. Mikolov et al. proposed and improved the Doc2vec text vector features proposed. Doc2vec is basically similar to the word2vec model, except that the text vector is added in the training process, so the text vector of the context can be encoded. The model also includes two structures: Distributed Memory (DM) Model and Distributed Bag-of-Words (DBOW). It also has the same functionality as the Skip-Gram model and the CBOW model. By encoding the text with considering the context, text vectors can be finally got.

(3) LDA Feature Extraction. The Latent Dirichlet Allocation (LDA) topic model was proposed by Friedman et al [32]. They believed that the topic of an article is in line with the Dirichlet distribution, so as to obtain the relationship between texts, and the VSM is compared to increase probability information. LDA model is composed of a three-layer generative Bayesian network structure [32], including documents, topics, and words. The core probability calculation is shown in Formula (4).

$$p\left(w_i \mid d_j\right) = \sum_{s=1}^{k} p\left(w_i \mid z = s\right) p\left(z = s \mid d_j\right), \tag{2}$$

where $p(w_i \mid z = s)$ represents the probability that the word w_i belongs to s topics, and $p(z = s \mid d_j)$ is the probability of the s topic in the short text d_j. Based on the LDA topic model, the topic probability distribution of the text can be obtained, which can be extracted as the topic features of the text.

3.2. AES Model Based on BiLSTM. Short-long-term memory network (LSTM) is a special kind of recurrent neural network (RNN). RNN network takes the output of the previous moment as a part of the input and inputs it into the neural network together with the external input at the moment, but there are problems of short-term memory and gradient disappearance. LSTM improves the RNN model by replacing the hidden layer nodes of RNN with Memory units and protects or controls the node states of LSTM neural network through the gate structure, so that the closed loop is formed between the hidden layer, and the weight of the hidden layer is responsible for controlling scheduling Memory. The state of the hidden layer participates in the next prediction as the memory state of the current moment, which solves the problems of short-term memory and gradient disappearance.

BiLSTM allows bidirectional information flow in LSTM. Two hidden states are used from backward to forward and from forward to backward; the internal structure is shown in Figure 1. The calculation formula of BiLSTM at the current moment is shown as follows:

$$\begin{aligned} f_t &= \sigma\left(w_f x + b_f\right), \\ i_t &= \sigma\left(w_i + b_i\right), \\ o_t &= \sigma\left(w_o + b_o\right), \\ c_t &= f_t \otimes c_{t-1} + i_t \otimes \tanh\left(w_c x + b_c\right), \\ h_t &= o_t \otimes \tanh\left(c_t\right), \end{aligned} \tag{3}$$

where (f_t, i_t, o_t) is the output value of the forgetting gate component, and c_t is the current state at moment c. The unit state matrix and neuronal paranoia of forgetting gates, memory gates, are represented by $W_i, W_f, W_o \in R$ and $b_i, b_f, b_o \in R$.

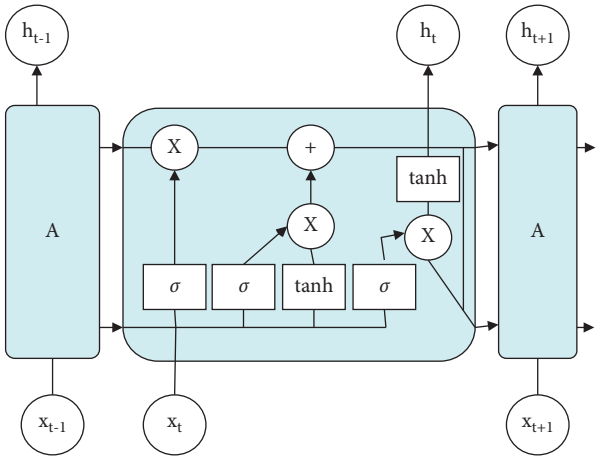


Figure 1: Internal structure of neuron.

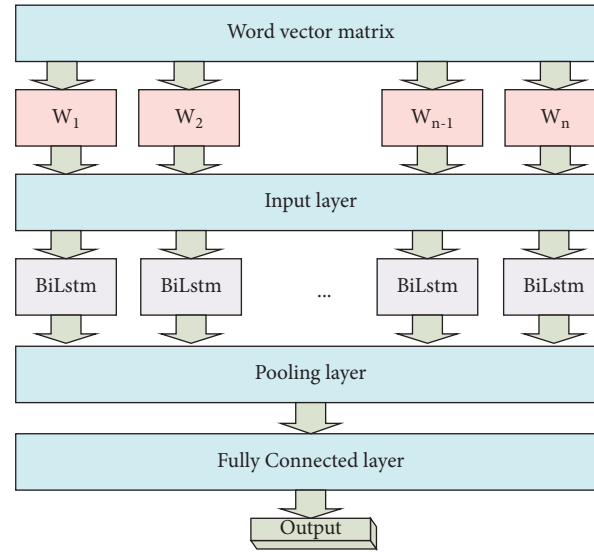


Figure 2: AES network framework based on BiLSTM.

This method of memory selection is consistent with the expression of the relation between words before and after in English composition, in order to mine the contextual timing information in the composition, so as to extract the characteristics of the logic before and after the composition. We take the word vector matrix as the feature input to represent the article and use the LSTM model, which is suitable for dealing with timing problems to build the model. Compared with other network models, LSTM can make full use of context information of composition. The structure of AES model based on LSTM is shown in Figure 2.

3.3. Beautiful Sentence Recognition Based on CNN

3.3.1. Network Overview. Composition morphology and grammar are the basic requirements of writing; really judging the level of a composition is advanced expression of beautiful sentences; these beautiful sentences greatly improve the appreciation of the composition. These sentences usually include fusion advanced vocabulary, clever sentences of English grammar, and some contain rhetoric devices. Therefore, to quantify the beautiful degree and distribution of all writing sentences and fused related characteristics at the same time by building beautifully set identification model can help build the model of AES as well as improving the score prediction efficiency. And the result will not mention mechanization.

The problem of sentence elegance recognition can be regarded as a text classification problem. The main task is to learn the text content by computer and train the classification model according to the given text label, so as to obtain the classification results of the new input text. Traditional statistics-based machine learning methods usually extract features from text manually and then train the classification model with machine learning classifier. However, it is difficult to grasp the beauty features of language and make perfect features artificially, while the biggest advantage of deep learning method is that automatic selection and combination of features can better reflect the features of text information. Traditional methods based on statistics and rules use manual sentence features, which make it difficult to extract the core information of good sentences containing advanced grammar (parallelism and inversion have fixed structures that are easy to extract linguistic features, but metaphors, personification, and other figures of speech have unique linguistic structures). The neural network model can automatically learn the semantic vector that can represent the sentence feature from a large amount of data, so as to complete the task of binary classification.

Convolutional Neural Network is a commonly used artificial Neural Network structure, which uses convolutional kernel arithmetic to obtain local information and then obtains global information through the pooling layer. Its main structure includes input layer, convolution layer, and pooling layer. In the task of automatic composition scoring, the input layer is the text representation matrix of word vector, and the convolution layer can set convolution kernels of different sizes. This method can obtain certain context timing information. In addition, compared with the traditional neural network model, CNN introduces weight sharing, which reduces the complexity of the network model and improves the training speed. In the pooling layer, the piecewise maximum pooling method is adopted, which can retain the relative location information of multiple local maximum eigenvalues. If strong features appear repeatedly, this method can also obtain feature intensity. However, absolute position information is lost, and only coarse position information is retained. After convolution and pooling, sentence level representation is obtained.

3.3.2. AES Network Model Based on CNN

(1) Input Layer. This section attempts to use convolutional neural network to extract sentence elegance features, so it is necessary to complete the transformation from word granularity features to sentence granularity features. In the input layer of convolutional neural network, the word vector is obtained by word embedding method, and the two-

dimensional vector matrix that can represent the sentence is obtained $m * k$ (m is the number of words in the sentence, and k is the length of the word vector). The specific operation is as follows: firstly, word2vec model is trained based on Wikipedia corpus, and then word vector of each word is obtained by incremental training based on the corpus obtained in this paper. The vector length is set to 128, and then word vector matrix is obtained by splicing word vectors in sentences. Since the longest sentence in the training set has 123 words, this paper adopts zero-complement processing to make the input matrix of each sentence the same size for sentences with a length less than 47, that is, 123×128, and the model structure is shown in Figure 3.

(2) Convolution Layer. Convolution is a mathematical operator that generates a third function through two functions f and g and represents the integral of the overlapping length of the product of the overlapping function values of f and g through inversion and translation. For neural networks, it is actually to search for more representative features through operation. The word vector matrix is input into the convolutional neural network, and three types of convolution kernels, 3×128, 4×128, and 5×128, are designed, respectively, in the convolution layer, each of which has 50 convolution kernels. For the operation with sentence length m, the jth convolution kernel computes convolution of the words in the window with length a, and the output result is c_j and the sentence length is expressed as follows:

$$\begin{aligned} x_{i:i+h-1} &= x_i \oplus x_{i+1} \oplus \cdots \oplus x_{i+h-1}, \\ c_j &= f\left(w * x_{i:i+h-1} + b_j\right), \end{aligned} \tag{4}$$

where x_i is the i vector in the two-dimensional matrix composed of adjacent a word vectors, w is the weight parameter of the convolution kernel, b_j is the bias element, and f is the activation function. In order to consider the contextual information of each word as much as possible, this paper extracts local features of different dynamics by changing the size of window length a. By convolution operation on the input data of convolution check, a feature graph c can be obtained, and finally 150 feature graphs can be obtained. In order to accelerate the convergence speed of network training, ReLu function is used as the activation function of each neuron.

$$\begin{aligned} c &= \left[c_1, c_2, \ldots, c_{m-a+1}\right], \\ f(x) &= \max(0, x). \end{aligned} \tag{5}$$

(3) Pooling Layer. In order to obtain more accurate feature expression, local optimal features need to be obtained from feature maps; that is, chunk-max-pooling operation is performed on feature maps extracted from the convolutional layer. Firstly, the feature graph is divided into parts, then the maximum value of each one is extracted and combined into a vector composed of maximum value, and then the maximum value vector of all feature graphs is spliced to complete the whole process of extracting a local feature from the input data. The feature maps were input to the pooling layer, and each feature map was divided into three parts by local maximum pooling operation, and the features were combined in sequence. Local maximum pooling has better performance than global maximum pooling, because it preserves relative order information as much as possible while capturing the strength of features.

(4) All Connections. The combination and artificial features vectors are joined together into one dimensional vector and connected to the whole connection layer, finally through the beautiful output node to output value. The output node is the sigmoid function. In the frame of the Keras neural network, we adopt the method of stochastic gradient descent to adjust parameters reversely through iterations. Thus, the final training meets the requirements of model structure.

4. Experimental Analyses

4.1. Experimental Preparation and Evaluation Indicators

4.1.1. The Experimental Data. In order to ensure the accuracy of the test results, the "Composition Scoring Competition" on the international data mining platform Kaggle is selected [33]. The composition data set includes 8 data subsets, and each subset has corresponding composition questions. Students write according to the requirements of the questions. Under each training data subset, there are more than one thousand compositions that students have learned and the corresponding manual grading scores, and each composition is usually between 150 and 650 words. The score range of each data subset is also different. For example, the score range of data set 3 is 0–3, while the score range of data set 7 is 0–30, and the score range of data set 8 is 0–60.

In order to further verify the reliability of the beautiful essay model, we collected the beautiful sentences of IELTS and TOEFL on Douban and "English beautiful essay" on other English websites. We cut the beautiful essays into sentences and defined the above sentences as beautiful sentences. Then, primary school English composition and junior high school English composition modules are collected from the English network. We cut the composition into sentences, respectively, and define such sentences as not beautiful sentences. Among them, 3340 original beautiful sentences and 3304 unbeautiful sentences were selected, and the sentences with more than 6 words and less than 50 words were selected as valid samples. In the end, there were 3200 beautiful sentences and 2900 unbeautiful sentences.

4.1.2. The Evaluation Indicators. Quadratic Weight Kappa (QWK) is used to evaluate the performance of the model. QWK is a consistency test method used to evaluate whether the results of the model are consistent with the actual results. Let the composition score have N natural number grades, and there are two markers (A: manual grading, and B: algorithm grading). The score of each composition E can be represented by an array (e_a, e_b), which represents the score of composition e by A and B markers, respectively. First, A histogram matrix O of order n is constructed, which represents the number of essays that marker A typed with A score of i and marker B typed with A score of j. An n-order

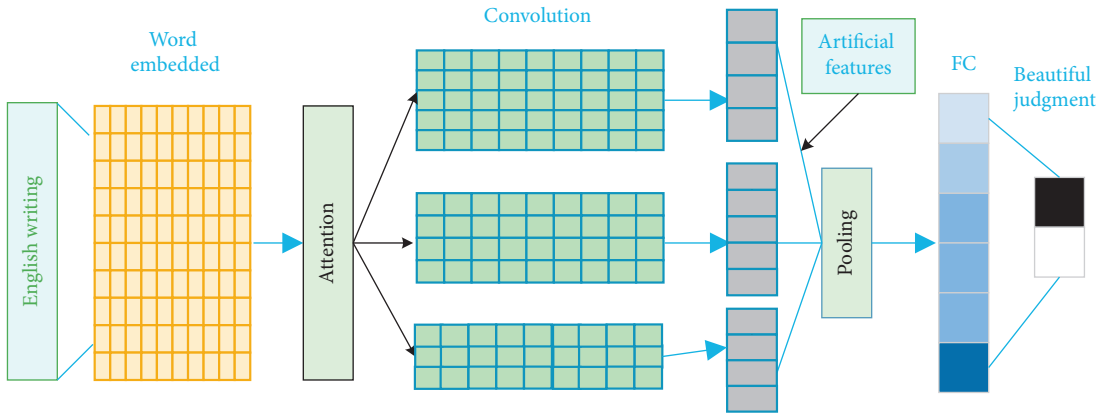


Figure 3: AES network model based on CNN.

weighting matrix w is then calculated based on the difference between the marks scored by the two markers. After calculating the cross product of the marker's scoring histogram vector to obtain the matrix E, the quadratic weighted Kappa value K is solved in the following way:

$$w_{i,j} = \frac{(i-1)^2}{(N-1)^2},$$
$$K = 1 - \frac{\sum_{i,j} w_{i,j} O_{i,j}}{\sum_{i,j} w_{i,j} E_{i,j}}. \tag{6}$$

Then, the approximate variance stable Fischer transformation is carried out to obtain Z. Finally, after taking the mean value of Z, the inverse Fischer transformation is carried out to obtain the final average quadratic weighted Kappa value.

$$Z = \frac{1}{2}\ln\frac{1+k}{1-k},$$
$$K = 1 - \frac{\sum_{i,j} w_{i,j} O_{i,j}}{\sum_{i,j} w_{i,j} E_{i,j}}. \tag{7}$$

The evaluation model still adopts the conventional neural network evaluation indexes: accuracy rate, recall rate, and $F1$-score.

$$AR = \frac{TP + TN}{TP + TN + FP + FN},$$
$$R = \frac{TP}{TP + FN} \times 100\%, \tag{8}$$
$$F1 = \frac{2(TP + FP) + (TP + FN)}{TP} \times 100\%.$$

4.2. The Experimental Contrast

4.2.1. Comparison of Basic Experiments. The baseline methods are GBRT [34] model based on linguistic features and LSTM model based on word2vec word vector features, denoted as A1 model and A2 model, respectively. The former is an AES model based on machine learning and manual feature extraction, while the latter uses neural network to extract features and train the model. GBRT is an integrated learning method whose base model is regression tree. Gradient descent approximation method is used to fit the residual term of the lifting tree, which is considered to be one of the machine learning algorithms with strong generalization ability. The average quadratic weighted Kappa value of GBRT model A1 based on linguistic features is 0.63, and that of LSTM model A2 based on word vector features is 0.7.

In order to verify the reliability of the machine learning English composition automatic scoring algorithm proposed in this paper, we carry out eight feature combination experiments. The purpose is to find the best combination of features suitable for the model through the combination of multiple groups of features, so that the final result is consistent with the actual situation. It is obvious from Figure 4 that the evaluation results of incorporating semantic or linguistic features into the baseline model are improved to some extent compared with the baseline model. The word vector clustering feature (C1) has the most obvious improvement effect, while the improvement effect of feature combination 2 is not obvious after the addition of theme feature (C3). The comparison of feature combination 5 and 7 shows that the model effect even decreases after the addition of theme feature, and the average quadratic weighted Kappa value is the lowest in the whole experiment. This is because this paper did not make a careful analysis and comparative experiment on the selection of the number of topics when extracting the theme features. In experimental group 5, the performance of the model improved greatly after the inclusion of linguistic features, reflecting the important role of lexical features and syntactic features in AES.

4.2.2. Experimental Comparison of Beauty Sentence Evaluation. Convolutional neural network can extract deep sentence features well, so three groups of feature type experiments are carried out (artificial feature extraction A, CNN feature extraction, and artificial + CNN feature extraction ACNN). The trailing + - indicates whether it is a beautiful sentence. From the overall trend of accuracy and recall rate in Figure 5, the experimental results of these three

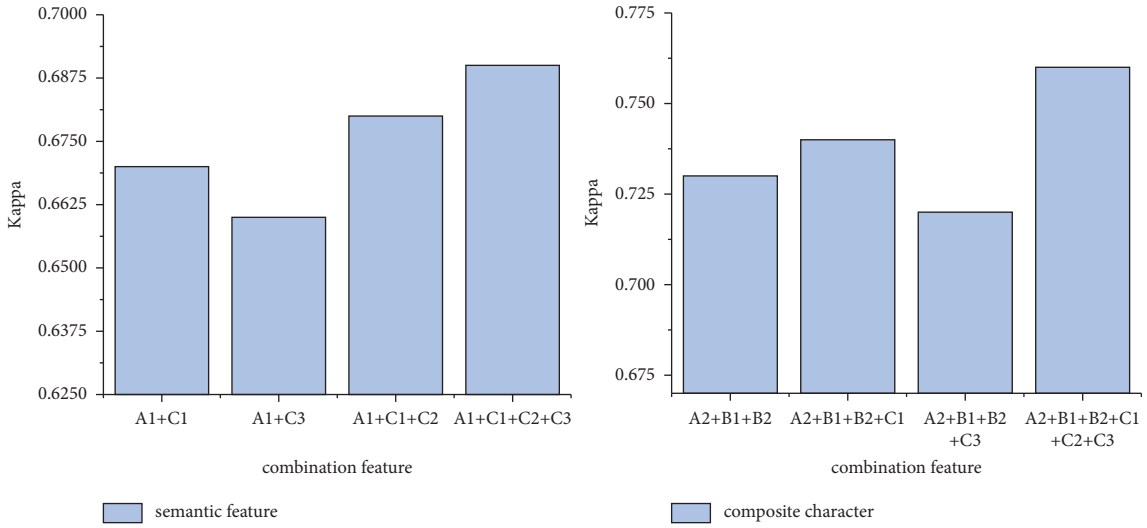


Figure 4: Feature fusion experiment. *Note*. Linguistic features: B1 lexical features and B2 syntactic features; semantic features: C1 word vector clustering feature, C2 text vector feature, and C3 topic feature.

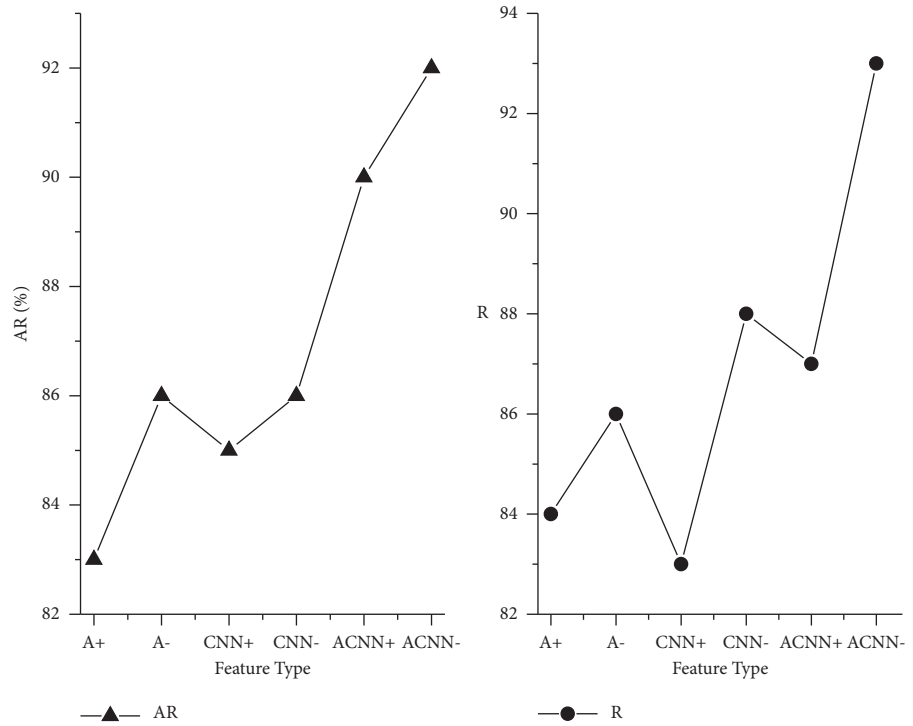


Figure 5: Sentence elegance judgment based on different feature combinations.

groups are basically similar; what is different is that the experiment accuracy rate of artificial feature extraction of beautiful sentence is the lowest, that is, only 83%. And the lowest recall rate is CNN feature extraction for the judgment of the beautiful sentences, and the experimental results and characteristics of the same type of beautiful sentence are always lower than the experimental results of unbeautiful sentences. It can be concluded that there is still a bias in the algorithm for the definition of beautiful sentences, and sometimes, there is a certain difference in people's judgment of beautiful sentences in practical application.

According to the F1 evaluation index in Figure 6, the graceful sentence is still lower than the ungraceful sentence, which further verifies our previous test and analysis. It can be seen from the experimental data that, compared with the single manual definition of features and the method of extracting sentence features only using CNN, combining the two feature extraction methods can greatly improve the classification effect of beautiful sentences and unbeautiful sentences. At the same time, the evaluation index of beautiful sentences in each category is smaller than that of nonbeautiful sentences. It can be seen that the judgment of

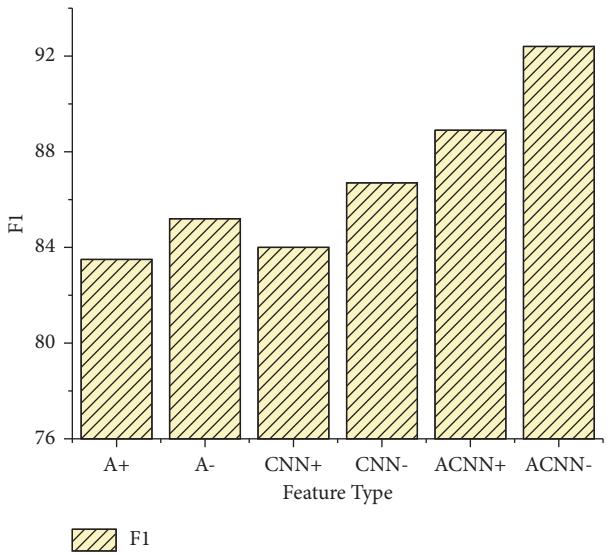


Figure 6: Sentence elegance accuracy and recall rate under different feature combinations.

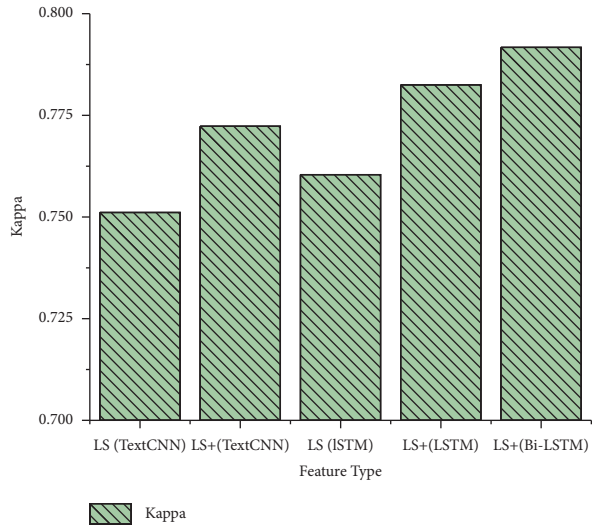


Figure 7: F1 value of sentence elegance under different feature combinations.

sentence beauty is a difficult point, which is also the key point to distinguish the writing level.

In order to further improve the experiment, we not only consider the combination of machine learning feature extraction but also try to analyze the experimental results of different network classification for different feature combinations from the perspective of network learning (LS represents the combination of linguistic and semantic features, and + represents the introduction of sentence elegance features). It can be clearly seen from the experimental results in Figure 7 that the model incorporating sentence elegance features has better performance than the text-CNN and LSTM models before inclusion. Meanwhile, when the feature types are the same, the performance of LSTM model is better than that of Text-CNN, showing the former's excellent memory learning ability in Text mining. The improved BiLSTM based on LSTM has achieved the best result of the second experiment, mainly because it can better discover the time influence of different features from different directions and get better sentence features with time validity.

5. Conclusions

Aiming at the problems existing in automatic grading of English compositions in English learning, this paper studies AES algorithm based on machine learning feature extraction. Based on linguistic features, semantic features, and their extraction methods, the paper studies the effect of linguistic features and semantic features used in AES on the model. Based on this, the BiLSTM basic automatic scoring algorithm and the improved CNN automatic scoring algorithm for beautiful sentence judgment are discussed. The experimental results show that the combination of linguistic features and semantic features is better, and the combination of a variety of features is used to analyze the performance of automatic composition scoring to select the best model through the combination of features. This paper aims to reduce the teacher's burden of assessment and uses it as the starting point, and also to improve the quality of teaching and students' English writing level; a complete set of improved English composition automatic scoring system has been designed through the study and discussion of English composition scoring. [35].

Data Availability

The datasets used in this paper are available from the author upon request.

Conflicts of Interest

The author declares no conflicts of interest regarding this work.

Acknowledgments

The study is the phased achievement of "the curriculum ideological and political demonstration class construction project" of Department of Education of Liaoning Province, China (Grant no. 2021-01380).

References

[1] F. Sharifian, "Globalization and the development of metacultural competence in English as an international language learning," *Multilingual education*, vol. 3, no. 1, pp. 1–11, 2013.

[2] Y. Guo, "Cultivation of learners' subjective awareness in English writing teaching," *Teaching and Management: Theory*, vol. 6, no. 7, pp. 100-101, 2009.

[3] Z. Zhu, "Good English composition is an important embodiment of improving students' language application ability," *North-south Bridge*, vol. 9, no. 4, p. 184, 2015.

[4] Y. Ye, "Weak links in writing teaching from the perspective of language errors in students' exercises," *Foreign Language Teaching,*, vol. 12, no. 4, pp. 77–81.

[5] D. Boulanger and V. Kumar, *Deep Learning in Automatic Paper Grading*, Springer, Berlin, Germany, 2018.
[6] E. B. Page, "The imminence of grading essays by computer," *Phi Delta Kappan*, vol. 47, no. 5, pp. 238–243, 1966.
[7] K. Landauer Thomas, L. Darrell, and F. Peter, "The intelligent essay assessor," *IEEE Intelligent Systems & Their Applications*, vol. 15, 2000.
[8] P. W. Foltz, "Latent semantic analysis for text-based research," *Behavior Research Methods, Instruments, & Computers*, vol. 28, pp. 197–202, 1996.
[9] R. D. Manning, M. S. Cohen, and R. L. Demichiell, "An overview of current research on automated essay grading," *Journal of Information Technology Education: Research*, vol. 2, 2003.
[10] M. Liang and Q. Wen, "Review and enlightenment of automatic composition scoring system in foreign countries," *Audio-visual Teaching of Foreign Languages*, vol. 7, no. 5, p. 7, 2007.
[11] Y. Wang, Z. Li, and Y. He, "Key techniques for automatic essay scoring based on semantic dispersion of text," *Journal Of Technology*, vol. 30, no. 6, p. 9, 2016.
[12] T. Qiu, "Research and implementation of discourse fluency assessment techniques in intelligent English composition assessment," *Beijing University of Posts and Telecommunications*, vol. 1533, no. 3.
[13] M. Liu, B. Qin, and T. Liu, *Intelligent Computers and Applications*, vol. 6, no. 1, pp. 1–4, 2016.
[14] W. Yangwei and X. Huang, "Automatic scoring of English composition based on linguistic features and self-encoder," *Application of Computer Systems*, vol. 26, no. 1, p. 8, 2017.
[15] c. Zhao and y. wang, "An image optimization classification method based on word bag model," *Journal of Electronics and Information Technology*, vol. 34, no. 9, 2012.
[16] Q. Guo, Y. Li, and Q. Tang, "Research on text similarity calculation based on VSM," *Application Research of Computers*, vol. 25, no. 11, pp. 3256–3258, 2008.
[17] P. Tao and L. Lu, "PU text classification enhanced by term frequency-inverse document frequency-improved weighting," *Concurrency and Computation: Practice and Experience*, vol. 26, no. 3, pp. 728–741, 2013.
[18] Y. Bengio, R. Ducharme, and P. Vincent, "A neural probabilistic language model," *Journal of Machine Learning Research*, vol. 3, 2003.
[19] T. Mikolov, K. Chen, and G. Corrado, "Efficient estimation of word representations in vector space," in *Proceedings of the International Conference on Learning Representations (ICLR 2013*, Scottsdale, AZ, USA, May 2013.
[20] E. Emnlp, *Empirical Methods in Natural Language Processing*, 2014.
[21] M. Peters, M. Neumann, M. Iyyer et al., "Deep contextualized word representations," in *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, New Orleans, LA, USA, June 2018.
[22] J. Devlin, M. W. Chang, and K. Lee, "BERT: pre-training of deep bidirectional, transformers for language understanding," 2018, https://arxiv.org/abs/1810.04805.
[23] Y. Kim, "Convolutional neural networks for sentence classification," 2014, https://arxiv.org/abs/1408.5882.
[24] Y. Liu, X. Qiao, and S. Zhao, "Deep Fusion of large-scale features in statistical machine translation," *Journal of Zhejiang University (Engineering Science)*, no. 1, 2017.
[25] A. Hassan and A. Mahmood, "Convolutional recurrent deep learning model for sentence classification," *IEEE Access*, vol. 6, p. 1, 2018.
[26] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Transactions on Signal Processing*, vol. 45, no. 11, pp. 2673–2681, 1997.
[27] W. Wang, L. Wang, and Y. Chai, *Application Research of Computers*, vol. 36, no. 5, p. 5, 2019.
[28] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
[29] Z. Zeng, L. Li, and J. Chen, "Bidirectional depth LSTM for emotion classification," *Journal of Computer Science*, vol. 45, no. 8, p. 6, 2018.
[30] H. Liu, S. Liu, and X. Zhang, "An optimized K-means text feature selection in clustering mode," *Computer Science*, vol. 38, no. 1, p. 3, 2011.
[31] Z. Zhou, "Machine learning," *China Civil And Commercial*, vol. 3, no. 21, p. 93, 2016.
[32] N. Friedman, D. Geiger, and M. Goldszmidt, "Bayesian network classifiers," *Machine Learning*, vol. 29, 1997.
[33] Kaggle, "The Hewlett Fewlett foundation:automated essay scoring," 2019, http://www.kaggle.com/c/asap-aes/data.2012-2-10.
[34] J. Friedman, "Greedy function approximation: a gradient boosting machine," *Annals of Statistics*, vol. 29, 2001.
[35] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," *The Annals of Applied Statistics*, vol. 1, 2001.

Hindawi
Scientific Programming
Volume 2021, Article ID 9141339, 10 pages
https://doi.org/10.1155/2021/9141339

Research Article

Artificial Intelligence Technology Assisted Music Teaching Design

Dan Dan Dai

NanChang JiaoTong Institute, NanChang, Jiangxi 330100, China

Correspondence should be addressed to Dan Dan Dai; 06054@ncjti.edu.cn

Received 8 November 2021; Accepted 3 December 2021; Published 21 December 2021

Academic Editor: Baiyuan Ding

With the continuous improvement of the global economic level and scientific level, information technology has penetrated into all fields of people's life. Today, the strategy of vigorously promoting educational modernization has created conditions for intelligent music teaching and promoted the in-depth application of artificial intelligence technology in education. Intelligent instructional design supported by artificial intelligence technology is the deep integration of information technology and music teaching. Through the intelligent perception technology, learning analysis technology, and emotional computing technology of artificial intelligence, an intelligent music teaching model is established. The online learning and education platform based on big data intelligence provides teachers with rich teaching methods, provides personalized evaluation and adaptive learning services for students' learning, and helps improve the efficiency of music teaching. The traditional teaching design model cannot effectively guide the intelligent music teaching and cannot meet the needs of students' development. Therefore, on the basis of previous studies, the author studies the intelligent teaching design supported by artificial intelligence technology. This study uses new generation information technologies such as big data, Internet of things, mobile Internet, and artificial intelligence to build a complete set of scientific intelligent music teaching design model. Wisdom teaching provides a reference for the whole process before, during, and after class, helps guide teachers to better carry out wisdom teaching, helps students explore cooperative and autonomous learning, and promotes the wisdom transformation of teaching methods and learning methods to a certain extent. Music classroom teaching has become more targeted and effective. Therefore, it is of great significance to cultivate intelligent music talents.

1. Introduction

Technology changes life; innovation leads the future. Traditional music teaching methods and learning methods can no longer satisfy the increasingly updated technology and culture, and the new ecology and new model of intelligent teaching must be reconstructed [1–5]. Intelligent music teaching empowered by artificial intelligence technology is the key goal of educational research in the new era, which is the organic integration of high technology and educational science. In order to keep up with the pace of today's era, it is very important to conduct intelligent teaching. Therefore, we must bravely shoulder the important historical task, put teaching reform at the top of education, gradually transform traditional music education to intelligent music education, and use new technologies such as artificial intelligence to cultivate smart talents and construct scientific. The intelligent teaching design model opens a new mode of intelligent teaching to guide the implementation of intelligent teaching and lead students to realize interdisciplinary learning, deep learning, and unbounded learning [6, 7]. Artificial intelligence technology drives the development of smart teaching. The "father of cybernetics" Wiener said: "We have transformed our environment so thoroughly that we must now transform ourselves in order to survive in this new environment." Education does not transfer knowledge from the outside to the inside, but to realize wisdom from the inside out. It is necessary to break the uniform traditional education form, build a smart education system that is compatible with the artificial intelligence era, and use artificial intelligence technology to systematically upgrade the learning environment, learning content, teaching methods, and management models, so as to provide students with rich choices and more personalized and more precise wisdom

teaching. Without intelligent education, it is difficult to realize modern education. The use of artificial intelligence technology is not only a demonstration of teaching tools, but also a new teaching ecology that is people-oriented, based on symbiosis, mutual building, and shared sharing. Intelligent education is a fair education for everyone, an education that extends to everyone's life, an education that suits everyone's personality, and an education that is more open and flexible [8], applying artificial intelligence and other modern information technologies to smart teaching, transforming the traditional teaching model of teachers plus textbooks plus students plus books to using high-tech big data analysis, cloud desktop cloud classrooms, and smart teaching using human-computer interaction form, with rich content and novel methods to guide students to actively participate in learning activities. The intelligent teaching environment is constructed by the extremely fast mobile 5G network, various big data, and new artificial intelligence technology products. Intelligent teaching is teaching aided by a new generation of information technology. Its core is to change the way teaching materials are used [9–13]. Informatization helps teachers prepare lessons intelligently, analyze and master their studies, and finally display teacher-student activities through tablets and computer devices. The teaching concept of "student-oriented, learning-based teaching" should be run through the entire learning activities, and artificial intelligence technology and equipment should be used to serve students' learning. Through the feedback and analysis of learning big data, teachers can quickly and accurately know which knowledge points of this class are difficult for students, so that they can accurately teach students in accordance with their aptitude and accurately make up for the difference. The biggest change brought about by the application of artificial intelligence technology is to empower teaching. The use of electronic teaching plans can avoid teachers' repeated copying work. Teachers can reintegrate high-quality resources at any time, which reduces the workload of teachers and can use more time and resources, energetic to provide students with personalized learning micro-tutoring to improve teaching efficiency. Students actively choose their favorite subjects and chapters for autonomous learning through computers or mobile phones or perform targeted analysis and extended exercises of wrong questions, which fully stimulates students' enthusiasm and initiative in learning. It is no longer necessary for teachers to ask questions one by one when answering questions. Students can use the tablet in their hands to answer. Teachers can see the feedback of students' answers in real time through the electronic whiteboard. Teaching is efficient and targeted; students have high interest, automatic participation, and strong interaction. This make teaching more dynamic and scientific. Not only that, the system can also automatically count the correct rate of answers to help teachers analyze the teaching situation. Teachers use various intelligent tools to carry out teaching work, which greatly reduces teachers' repetitive work and greatly improves classroom efficiency[13–15].

There are many information platforms on the Internet, and many excellent teachers selflessly upload their own exquisite courseware and other teaching resources to the platform, providing a wealth of cloud classroom, cloud courseware, and other course resources for everyone to share, learn from each other, and truly achieve mutual aid and mutual learning in education, cooperation, and win-win. The way of class has also changed a lot. Some platforms can realize the same class in different places around the world. Starting from the original intention, selflessly conveying the infinite power of knowledge to others, it reflects the respect for knowledge to the greatest extent and once again proves that knowledge has no boundaries and learning has no framework. The way students practice has also changed. Students can use mobile phones or tablets to practice online, and teachers can get feedback in time [16]. Teacher evaluation has also been greatly changed. For example, some software can easily evaluate and provide feedback regarding the situation of students. Examination methods have also changed. For example, with some answering software, teachers can use the software to automatically collect statistics on students' performance after students have completed them and provide fast and accurate data for teachers to analyze test papers and quality after the exam.

The integration of artificial intelligence technology into the smart teaching design into the classroom is not only an update of teaching methods and technology, but also an enrichment of knowledge and a transformation of teaching concepts. The development of smart teaching from the three-dimensional dimensions of time, space, and equality education provides both for top students. The broader space for progress and the precise support of intelligence have realized that, in all stages of talent training, in different situations, the most efficient and reasonable methods are used to train future highly intelligent talents. The deep integration of artificial intelligence technology and smart teaching design can effectively make up for the shortcomings of traditional education, focusing on cultivating students' abilities, improving students' information literacy, meeting students' individualized learning needs, and achieving the goals of smart education [17].

2. Artificial Intelligence Technology and Intelligent Teaching

2.1. Artificial Intelligence. Artificial intelligence itself is an interdisciplinary subject that simulates human abilities and intelligent behaviors. It has a wide range of fields, including mathematics, linguistics, psychology, brain science, and philosophy [18]. Artificial intelligence technology is the core technology of information technology in the new era. It mainly uses computers to imitate human thinking and some intelligent behaviors, such as search, reasoning, memory, speech recognition, knowledge expression, and processing of fuzzy information, so that it can behave like humans, as well as advanced intelligence and thinking to achieve higher-level computer applications. It is an extension of human's existing ability and has high intelligent technology beyond the scope of manpower. Since entering the twenty-first century, artificial intelligence technology has made great progress in big

data processing, algorithm, and deep learning and solved many challenging problems such as reasoning, speech recognition, image recognition, natural language processing, perception, and moving and manipulating objects [19]. At present, relevant artificial intelligence technologies include machine learning, deep learning, natural language processing, artificial intelligence algorithm, inference engine, learning calculation, image recognition, and other technologies [20]. With the strength of our country and the progress of science and technology, artificial intelligence technology has also developed rapidly. Today, it has been widely used in various important fields, such as transportation, logistics, finance, communication, medical treatment, education, and even into families and individuals. It is closely related to the lives of ordinary people, such as mobile phone voice, eye iris and face recognition, and fingerprint recognition. It has formed a series of rich industrial fields such as new forms and models, and education is no exception, which has brought revolutionary influence and endogenous driving force to the society. With the help of the Internet and the high development of science and technology, artificial intelligence technology has the advantages of repeatability, digitization, stylization, accuracy, and stability beyond human ability [21–23]. Only human unique consciousness and high-level thinking activities are realized through search, reasoning, induction, and association. Applying the integration of artificial intelligence technology and mobile 5G to the field of education, giving advanced power and energy to educational development and classroom reform, and improving some existing technical problems in current education, plays a very important role in promoting educational reform.

2.2. Intelligent Music Teaching. At present, the academic community has a variety of understandings on wisdom Teaching: Qian Xuesen, a famous Chinese scientist, from the perspective of system science, proposed to use Dacheng wisdom teaching to cultivate high-end innovative talents, eliminate the differences between various science and technology, integrate science and art, and integrate natural science and philosophy and social science, so as to achieve a large-span analogy and complete innovation. Professor Zhu Zhiting believes that intelligent teaching is to use various intelligent technologies to optimize the learning environment, let teachers and students jointly display novel teaching methods, and provide students with new and simple learning methods, so as to obtain efficient learning results and cultivate students into quick thinking, straight and intelligent talents. Pang Jingwen, a domestic researcher, believes that intelligent teaching is to use the intelligent teaching design model to complete the teaching goal of cultivating students' wisdom through pleasant and personalized classroom teaching in the environment of artificial intelligence technology. It is an intelligence based education to cultivate students into intelligent talents. Some scholars believe that intelligent teaching is to set up all links of intelligent learning teaching with the support of a new generation of information technology such as artificial intelligence, with the goal of students' intelligent learning and development. That is, under the intelligent environment such as artificial intelligence, teachers use various new generation information technologies and a variety of teaching resources to carry out teaching activities, aiming at stimulating teachers' teaching wit, improving teachers' professional technology, and cultivating intelligent talents. It has the basic characteristics of activity, efficiency, sharing, and interaction, so as to further improve the students' laziness and laziness in traditional classroom teaching negative and other phenomena [24]. Intelligent teaching is not only the informatization and intellectualization of educational infrastructure, but also the transformation and upgrading of educational ideas and methods, from focusing on the construction of "things" to meeting the diversified needs and services of "people." Only by focusing on students, tapping students' potential, awakening students' self-worth and enlightening students' wisdom, can we calmly deal with the challenges brought by the era of artificial intelligence technology.

2.3. Instructional Design. In the 1950s and 1960s, the systematic method of system science (including system theory, information theory, and cybernetics, also known as the "old three theories) was applied to the field of educational technology for the first time, the new theory of "instructional design" was created, and the "old three theories" had a far-reaching impact on instructional design. Now it is generally believed that teaching design is to analyze and plan the teaching objectives, teaching difficulties, teaching process, student exercises, summary and evaluation, and other elements according to the provisions of curriculum standards and students' characteristics, make an overall and detailed plan for a class, design reasonable teaching links, and complete a detailed and operable teaching plan and complete the teaching task scientifically according to the teaching plan. According to different elements in the teaching process, classroom teaching can be divided into three categories: classroom teaching with "teacher teaching" as the main body, classroom teaching with "student learning" as the main body, and classroom teaching with "resources" as the main body.

2.4. Intelligent Music Teaching Design. In today's intelligent education environment, students' access to information has greatly increased, the traditional music classroom teaching design has been difficult to meet the development needs of students, students have no passion for backward and monotonous teaching media, and students' learning interest is not high. The research and promotion of intelligent teaching design model are particularly urgent and important. "Intelligent teaching design" is an intelligent and efficient teaching guidance scheme applied in the whole process before, during and after class based on the constructivist theory and around the student-oriented central concept and supported by a new generation of information technology such as mobile Internet, Internet of things, big data, cloud computing, and artificial intelligence technology. Intelligent instructional design is the deep integration of artificial

intelligence technology and teaching, pays attention to the mutual penetration of "teaching" and "learning," plans classroom elements such as teaching resources, problem tasks, teaching links, and teaching hands, and forms specific teaching cases. In the intelligent instructional design model, there should be the application link design of intelligent technology and sufficient resource preparation to facilitate students' personalized learning and open learning; we should also predict the learning situation and students' psychological development provided by big data and design a variety of emergency plans for temporary adjustment in the teaching process. When designing intelligent teaching, teachers should carefully design teaching situations in combination with students' own basic situation, students' social living environment, and cognitive development characteristics, so as to cultivate students' intelligent brain and stimulate students' learning enthusiasm, resulting in "I want to learn and I love learning."

3. Artificial Intelligence Technology Aided Intelligent Music Teaching Design

3.1. Design Conditions. Intelligent teaching requires music teachers to pay more attention to cultivating students' wisdom generation and practical ability while practicing the traditional three-dimensional goal. "Smart classroom" is to integrate information technology and network technology into classroom teaching, so that students and teachers can give full play to their enthusiasm and initiative and educate people in an all-round and systematic way. According to the teaching characteristics of each class, teachers should design specific learning objectives, set up teaching situations that students are interested in, provide students with a vivid and interesting virtual environment, encourage students to practice, and properly integrate the teaching objectives into each learning link of intelligent teaching. The ultimate goal of wisdom classroom is to promote the generation of students' wisdom, and wisdom is a comprehensive and multidimensional goal, integrating knowledge, ability, emotion, and will. The important mission of education in the "Internet+" era is to cultivate intelligent talents.

The traditional classroom environment is mainly composed of teachers and students, blackboard, and chalk. "Wisdom classroom" is to integrate information technology and network technology into classroom teaching. This study mainly uses Shivo whiteboard, Shivo easy classroom, and Shivo class optimization masters to build a powerful interactive teaching platform, which takes multimedia interactive whiteboard tools as the application core, intelligent teaching under the support of the Internet: the massive resources and rich teaching tools of the electronic whiteboard, as well as infrared sensing technology and multitouch, intelligent recognition of gesture actions, can enlarge, shrink, and erase the whiteboard content, make the interaction between teachers and students more smooth, and make the classroom more interesting. Easy classroom intelligently pushes learning materials, which facilitates the interaction between teachers and students. It summarizes and generates individual evaluation data for various teaching links such as real time answer and selective answer. Teachers can reasonably group students according to their actual learning situation to realize hierarchical teaching. The class optimization master makes the class management easier, evaluates in an all-round way, sends comments in real time, makes the classroom performance clear at a glance, and stimulates students' competitiveness. Students' terminal tablet realizes paperless learning and greatly improves learning efficiency. Multimedia, multiscreen display and other aids make the smart classroom more interesting. Intelligent learning environment creates more suitable learning conditions for students and provides effective support for students' independent construction and learning of knowledge [25, 26]. Wireless network coverage, teacher-student interaction system, and other equipment enable teachers and students to easily obtain teaching resources and make teaching achievements more efficient.

3.2. Design Elements. "Teaching framework" refers to music teachers' comparative learning based on the previous teaching design and the current intelligent teaching before teaching design, and cleaning up the elements of the teaching process composed of intelligent classroom environment, teaching objectives, teaching methods, and teaching evaluation, so as to help teachers carry out teaching design in an orderly manner. Therefore, the teaching framework is the train of thought preparation for teachers to carry out intelligent teaching design in the intelligent teaching environment, the preliminary preparation for the perfect implementation of intelligent teaching, and the overall planning for the whole process of intelligent teaching before, during and after class. It includes six design elements: getting familiar with the intelligent environment, creating teaching objectives, obtaining teaching resources, conceiving teaching methods, presupposing teaching situations, and predicting learning results. They are carried out from front to back to complete the whole teaching process. This detail is shown in Figure 1.

Familiar with the smart environment, the smart environment is the smart classroom, which is composed of mobile 5G networks, PCs, electronic whiteboards, cloud classrooms, and artificial intelligence education products. Teachers use these new generation technology products to automatically collect and analyze students' learning conditions and environmental characteristics. Integrate analysis, automatically evaluate learning effects, automatically identify students' learning needs, and formulate personalized learning plans and push resources. Intelligent classroom is not a simple accumulation of information technology products. When choosing to use it, we should take into account the students' individual chemical situation and the actual needs of teaching links, be familiar with and flexibly use new technologies, effectively apply them to teaching links, and realize technical empowerment for intelligent teaching.

3.2.1. Create Teaching Objectives. The presupposition of intelligent classroom teaching objectives should not only meet the three-dimensional objectives proposed by the new

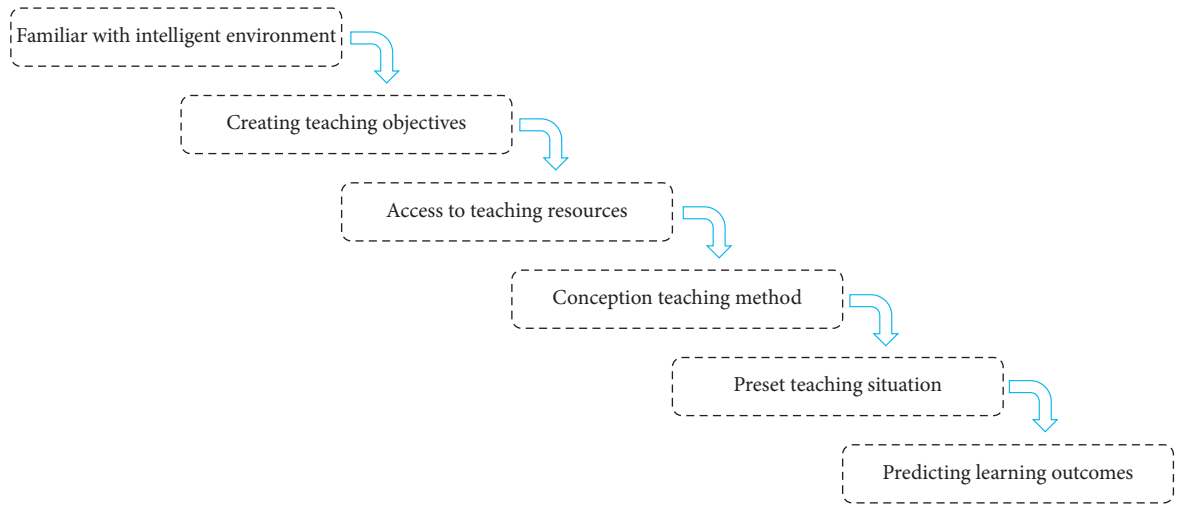


Figure 1: Design elements of intelligent teaching architecture.

curriculum reform, namely, knowledge and skills, process, and methods, emotional attitudes, and values, but also pay attention to cultivating students' wisdom generation and practical ability. The ultimate goal of intelligent teaching is to promote the generation of students' wisdom through intelligent classroom teaching and cultivate students into intelligent talents.

3.2.2. Access to Teaching Resources. The resources of the Internet are massive. Teachers should make full use of the intelligent cloud class to acquire effective resources through terminal devices, or provide students with ways and paths to acquire resources, and can access the resource library at any time. The cloud classroom can also automatically push resources through big data analysis, so that students can take the initiative to get information, scientifically collate information and process information. Explore in a cooperative way to effectively improve students' ability of intelligent learning and achieve the realization of teaching objectives. The characteristics of each discipline require teachers to collect relevant materials of the discipline as much as possible, so as to make it consistent with the reality of students' life and the characteristics of students' age development, mobilize students' enthusiasm for active participation, and learn independently. Rich network resources provide students with learning content far beyond textbooks.

3.2.3. Conceive Teaching Methods. The student-oriented teaching concept should be reflected in the link of wisdom teaching. According to the requirements of the realization of teaching objectives, teachers should conceive effective teaching methods. The teaching methods of wisdom teaching mainly include independent inquiry, task driven and cooperative discussion. Teachers should create a situation for students to perceive problems, carry out learning according to problem tasks, guide students to actively acquire knowledge, let students actively construct learning methods such as practice, exploration, and application, verify the learning process with the answers to problem tasks, and promote the generation of students' wisdom.

3.2.4. Preset Teaching Situation. In the smart classroom environment, teachers create real and interesting situational themes with the help of information technology resources and equipment according to the needs of teaching content, so as to inspire students' thinking and cultivate students' problem-solving ability. Appropriate teaching situation can arouse students' interest in learning and help improve teaching effect. Intelligent teaching requires teachers not only to skillfully operate intelligent teaching software and hardware and various advanced information technologies, but also to preset the automatic generation, diagnosis, and analysis of big data in the teaching process, as well as the difficulties that students may encounter in the learning process and the possible conditions in the interactive process, analyze and select the appropriate artificial intelligence auxiliary system, and envisage the solution strategy in advance. Through the presupposition of the situation, the whole teaching process is under the control of teachers at any time and carried out in an orderly manner.

3.2.5. Predict Learning Outcomes. Teachers release test questions to students at any time according to students' learning. After students answer questions, the system automatically calculates scores and corrects errors, then displays and analyzes them according to the feedback of big data, puts forward suggestions and guidance to students in time, analyzes students' error prone and weak links, and continues to automatically push the student's error prone question type. Systematically analyze the answers of each student, sort out the overall learning level of the class, and predict the confusion and error prone questions that students may have. Teachers can collect these data and carry out targeted teaching.

3.3. System Design. With the development of music education and technology, the traditional classroom teaching design process cannot meet the requirements of intelligent teaching. The traditional classroom teaching process usually adopts the "4 + 3″" model, that is, the four operation links of teachers (lesson preparation, teaching, assignment, and evaluation) and the three learning links of students (preview, listening, and completing homework). In the classroom process, the classroom communication between teachers "teaching" and students "learning" is relatively single, lacking in-depth interaction between teachers, students, and students. The intelligent teaching design of this study is several important intelligent classroom solutions and intelligent teaching practice in the stage of comprehensive basic education. Combined with the intelligent teaching experimental results in the intelligent education demonstration area, based on the teaching concept of "learning first," and according to the use of artificial intelligence technology and information technology, the teaching process is divided into three modules and seven links of "before class, during class, and after class." The teaching cycle of "7 + 7″" sustainable development has been formed. In smart teaching, teachers' "teaching" has become seven steps (resource release, goal setting, sensory introduction, task distribution, guidance and explanation, detection and evaluation, and extension and push), and students' "learning" has also become seven steps (independent preview, learning expectation, situational experience, cooperative learning, onstage explanation, consolidation of quiz, and breakthrough points), and the interaction between teachers and students is more vivid and rich.

From the "4 + 3 model" of traditional classroom teaching process structure to the "7 + 7″" model of intelligent teaching, it fully reflects the characteristics of intelligent teaching supported by artificial intelligence technology. Therefore, the ideal intelligent teaching process structure is the "7 + 7″" model, which focuses on "student-centered," pays attention to the mutual penetration and integration of "teaching" and "learning," no longer constructs the teaching process by taking "teaching" and "learning" as separate elements, and is a complete large cycle model based on "before, during, and after class." It is a new model of using artificial intelligence technology to achieve teaching objectives and finally promote the improvement and development of intelligent teaching in practice. The specific steps are shown in Figure 2.

Push resources, autonomous learning. Based on the psychological needs of students' self-realization based on humanistic learning theory, the electronic whiteboard is used to push preview resources to students before class, such as learning task lists such as Mu class, micro class, learning courseware, and preview test. Driven by psychological needs, students can study independently with the help of smart phones or PC terminals and submit preclass preview. Let students have a sense of expectation for learning and mobilize students' enthusiasm. Through data feedback, teachers analyze and study students' preview, so as to select and design effective teaching methods and educational technology.

Analyze the learning situation and clarify the goal. Since the grouping of ability stratification is a simple classification of students according to problem answers and learning conditions, on the basis of this classification, students' answer data are obtained with the help of easy classroom for learning situation analysis, and prominent cases are extracted for research, so as to clarify the tasks, interactive links, personality push, and technology use in the learning process: means, methods, and other aspects of personality guidance and guidance. The preclass stage of wisdom teaching is a complete learning process. First, teachers check students' historical achievements and knowledge points through electronic whiteboard and easy classroom to determine teaching objectives. Secondly, students independently complete the preview test questions pushed by teachers and submit them to the platform to form big data for teachers' analysis. Students can also discuss the problems encountered in the preview process based on the class learning exchange group. Thirdly, according to the feedback of goal presupposition, preview, and test, teachers conduct a comprehensive analysis of students' learning situation, accurately understand the learning situation information, and formulate an appropriate teaching design scheme to realize teaching based on learning. Teaching materials are unified, but there are various differences between schools, classes, students, and survival. Therefore, learning situation analysis is particularly important and is the basic link of teaching design. For a class (50 people) in a medium development area, the learning level of class students is divided into three levels: excellent students, medium students, and backward students. The hierarchical model of "7 (middle) + 2 (excellent) + 1 (back)" can be introduced to study and judge the learning situation. This simple learning situation classification, combined with the automatically generated learning big data, analyzes the students' preview situation. It is helpful for teachers to make a basic judgment on students and adopt reasonable teaching strategies and carry out classroom teaching design for all students.

Situational introduction to stimulate interest. As the saying goes, a good beginning is half success, and interest is the best teacher. Teachers put forward questions through the process description of the problem generation, stimulate students' motivation, interest, and curiosity of independent discovery and cooperative exploration by creating a problem scene, and guide students into an immersive learning state of continuous exploration experience by creating a real environment. In actual teaching, make full use of the playing function of electronic whiteboard for images and videos and the game design function of classroom activities to set up teaching situations, flexibly select adaptation methods, and mobilize classroom atmosphere. Classroom atmosphere is an important psychological condition affecting teaching. How to stimulate students' learning interest and desire for knowledge at the beginning of class is discussed. Situation introduction is particularly important. According to the teaching objectives, teachers create situations that meet the characteristics of students, set suspense, arouse students' strong curiosity, and stimulate students' learning interest and learning initiative. Situational introduction is a direct

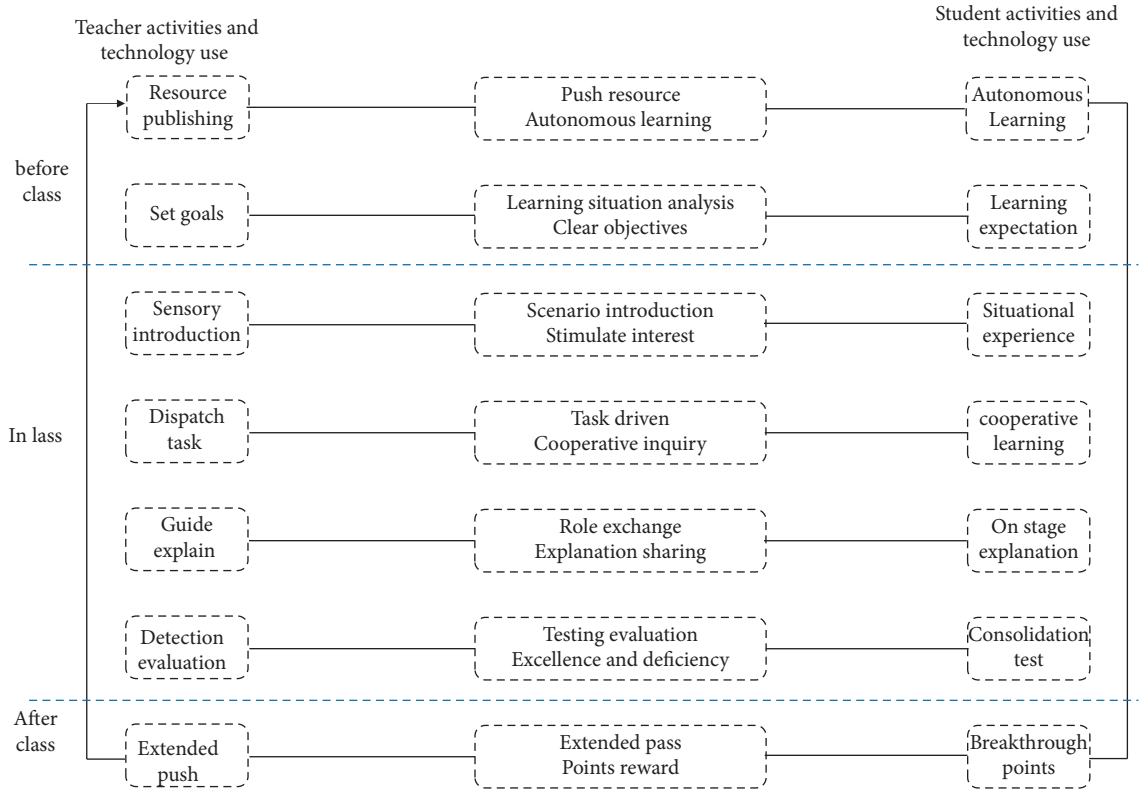


Figure 2: "7 + 7" model of intelligent music teaching.

stimulation of senses, which is deeply loved by students. Happy and novel situations can quickly attract students' attention to teaching. It is a teaching link worthy of use and praise. Situational design should vary according to materials, which requires careful design by teachers. There are the following common situational designs: first, film and television animation situation. We should make full use of multimedia and use animation to show students the situation of the problem. The second is the problem situation. The core of intelligent teaching is to analyze and solve problems. As long as teachers carefully design problems before class, problems from daily life can generally arouse their interest and make students really feel that knowledge is closely related to life. Although situational introduction is only a small part, it can activate the learning atmosphere of the whole classroom and promote students to complete teaching tasks with passion.

Task driven, cooperative exploration. Guided by the distributed cognitive theory, following the principle of "student-oriented," intelligent teaching establishes a correct view of teaching and teachers and students, emphasizes students' active learning, cooperative exploration, and interacts among groups to promote students' independent acquisition of knowledge. In the process of wisdom teaching, teachers should abandon the idea of being eager for success and instant success, always pay attention to students' learning state, slow down the rhythm of the classroom, create problem situations that help students explore independently, and let students actively participate in learning. The interactive teaching platform integrates cloud computing, situational awareness, intelligent recording and broadcasting, resource sharing, learning performance prediction, content planning, and other functions. Teachers distribute learning tasks in the form of questions to students' terminal tablets through electronic whiteboards for students to try to explore. Driven by the goal and task, students conduct cooperative discussion and explore problems through groups in the easy classroom. Teachers should create an inquiry atmosphere as much as possible and introduce learning resources according to the difficulty of learning content. Teachers should create conditions and help for students' attempt and inquiry and provide intelligent search, intelligent simulation teaching system and intelligent partners through understanding students' learning situation and according to specific learning tasks and students' problem-solving needs. AI assistant explains intelligent exercises, guides students to solve problems, and helps students easily master knowledge difficulties, use a variety of intelligent technologies to verify and summarize, and use the favorable environment of intelligent classroom to judge the learning results. In this process, even if students make some mistakes, it is desirable. At least they dare to practice. More importantly, they can cultivate students' ability of cooperative exploration and intelligent learning.

Exchange roles, explain, and share. Based on the constructivist theory, the cognition that students have generated is not fixed. It follows that students continue to blend with new knowledge in the later learning process. The understanding of new knowledge will be gradually updated and

strengthened with the increase of time and experience, and different students will have different understanding and understanding when learning the same knowledge. Therefore, it is necessary to teach students according to their aptitude and create a personalized learning environment for students. By trying to explore, students have a preliminary understanding of the learning tasks of this class, and some top students have obtained the correct methods and results. At this time, if the new course is explained according to the traditional teachers, and the students passively accept learning, it will weaken the importance and sense of achievement of the students' inquiry process. On the contrary, teachers choose better students to act as teachers. Each student is responsible for teaching the content of a knowledge point or a topic, explaining and demonstrating with an electronic whiteboard, and sharing how to explore the process and methods of learning and acquiring knowledge through the application of a variety of intelligent technologies. The knowledge and methods spoken by peers with their familiar and friendly language and synchronous adaptation rhythm are closer to their age characteristics and more acceptable and understandable. Teachers make timely comments, guide, summarize, and supplement the key and difficult points, pay attention to the learning status of backward students, stimulate students and teachers to ask questions, or carry out online communication and discussion between teachers and students, and students think about each other, exchange ideas and understanding, promote telepathy and thinking consensus, and break through the limitations of time and space of traditional classroom teaching, to make classroom teaching always maintain an active life state, so as to effectively develop and enhance wisdom. According to the interaction of students, teachers should timely provide feedback regarding the results obtained and then summarize and comment on the key, difficult, and doubtful problems, so that students can complete the transfer of knowledge and achieve their goals in practice. Let students play the role of teachers to show and explain onstage. Students are highly motivated. At the same time, they also encourage other students to strive for the opportunity to go on-stage. This learning process has strongly exercised students' knowledge inquiry ability, organization and planning ability, and oral expression ability, which is of great significance to their future growth and truly reflects the intelligent teaching concept with students as the main body and teachers as the leading role.

Test and evaluate, cultivate excellence, and make up for deficiency. Through cooperative exploration, peer explanation and sharing, and teacher guidance, students have a preliminary understanding of the knowledge of this class. In order to consolidate and extend the classroom learning effect, according to the students' learning rules, teachers should provide students with the opportunity to consolidate knowledge and complete the transfer in real time and conduct in class quiz in time. From the easy classroom test paper system, through automatic questions, manual questions, and self-questions, each student is matched with the exercises and tests of "tiptoe is enough," and the task list is intelligently pushed to students, so that students at different levels can get the most suitable test content, promoting students to complete actively and complete the understanding and digestion of knowledge in the shortest time. Students complete and submit tasks through the intelligent terminal. The system automatically calculates the total score and analyzes the test paper through manual or automatic marking by the teacher. The teacher makes targeted explanation and diversified evaluation according to the results of big data, uses the class optimization master to reward students with medals or points, updates the ranking of the honor list in real time, and carries out student-centered teacher-student interaction, Encourage top students to help poor students and let them play the role of "leader"; pay attention to secondary students and try to give full play to their learning potential; help backward students more, enhance their confidence in learning, and strive to minimize polarization and improve classroom teaching efficiency. Therefore, the dynamic evaluation data on students' learning effect can be recorded and analyzed through learning history data, data mining, and in-depth analysis, and intervene according to the data analysis results, so as to evaluate the learning process, predict future performance, and find potential problems. The intelligent auxiliary system dynamically monitors students. Every time students complete a learning task, the system will continuously identify and correct the individual knowledge ability value and the overall learning ability value of students, locate the loopholes of students' knowledge points, and check and make up for deficiencies, and the personalized chemistry learning plan determined for students will become more and more effective, providing a basis for the adjustment and improvement of teaching design.

Extend to break through the hurdle and earn points. It is students' nature to love games. The breakthrough mechanism of games is introduced into intelligent teaching. Through the setting of breakthrough level, students' desire for challenge is stimulated. After breakthrough, students experience the joy of reward and the feeling of success, so that students can place their learning in play, which not only meets students' psychological needs and emotional experience, but also helps cultivate students' learning interest and develop their potential. Class optimization Masters can set up a unique learning account for each student. After class, teachers face all students, push 1–2 time limited breakthrough questions to students through easy class, and students submit them after answering online. Teachers can set task points reward. Students can get corresponding points after submitting tasks. Teachers can get corresponding points for each unit or conduct periodic summary and reward every learning month. This periodic time should not be too long; otherwise, it will affect the stability of students' attention, but it should be continued to keep students interested in learning for a long time. The game encourages students to actively participate and develop and stimulates students' learning interest and motivation, and students' knowledge has also been consolidated and improved.

4. Conclusion

This paper discusses the design mode of intelligent music teaching from the aspect of music classroom teaching supported by artificial intelligence technology. This paper discusses the current research background of intelligent music teaching design, analyzes the research status of artificial intelligence technology and intelligent music teaching design, puts forward the research problems and difficulties, combines the research ideas and methods, and obtains the research value and significance of this paper. This paper defines the related concepts of artificial intelligence technology and intelligent music teaching design, expounds the cognitive theory and learning theory, and lays a theoretical foundation for intelligent music teaching design. This paper analyzes the six characteristics of intelligent music teaching design realized by artificial intelligence technology, which provides conditions for intelligent music teaching design; This paper summarizes the basic principle of intelligent music teaching design and expounds the important links of intelligent music teaching design. This paper analyzes the feasible conditions of intelligent music teaching and lists the design elements of intelligent music teaching framework. This paper designs the intelligent music teaching design model and process supported by artificial intelligence technology, so as to realize the continuous improvement and development of intelligent music teaching on the basis of the virtuous cycle before, during, and after class.

Data Availability

The dataset can be accessed upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

The authors thank Jiangxi Province Department of Education Science and Technology Project, Based on b/s Structure University Music Distance Learning System Development and application research (no. GJJ209311).

References

[1] I. Roll and R. Wylie, "Evolution and revolution in artificial intelligence in education," *International Journal of Artificial Intelligence in Education*, vol. 26, no. 2, pp. 582–599, 2016.

[2] J. Jia and Y. Yu, "Online learning activity index (OLAI) and its application for adaptive learning," in *Proceedings of the International Conference on Blended Learning*, pp. 213–224, Springer-Verlag, Hong Kong, China, June 2017.

[3] J. Tang and Y. Chen, "Reshaping the humanist position in the era of smart education," *World Scientific Research Journal*, vol. 6, no. 9, 2020.

[4] S. Liu, J. Wang, and S. Vijayalakshmi, "Ice and snow talent training based onconstruction and analysis of artificial intelligence education informatization teaching model," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 2, 2021.

[5] Z. Ma and Y. Tian, "The reform and influence of "artificial Intelligence+ education" on education and teaching design," *Advances in Higher Education*, vol. 4, no. 1, 2020.

[6] M. Ken, "Artificial intelligence in medical education," *Medical Teacher*, vol. 33, no. 2, pp. 23–33, 2019.

[7] R. Clariso, C. A. Gonzalez, and J. Cabot, "Smart bound selection for the verification of UML/OCL class diagrams," *IEEE Transactions on Software Engineering*, vol. 45, no. 4, pp. 45–50, 2019.

[8] M. J. Tis, "Letting artificial intelligence in education out of the box:educational cobots and smart classrooms," *International Journal of Artificial Intelligence in Education*, vol. 26, no. 2, pp. 701–712, 2016.

[9] Z. Wang, M. Li, J. Lu, and X. Cheng, "Business Innovation based on artificial intelligence and Blockchain technology," *Information Processing & Management*, vol. 59, no. 1, Article ID 102759, 2022.

[10] A. K. TeemuLeinonen, M. Veermans, and T. Toikkanen, "Mobileapps for reflection in learning: a design research in K-12 education," *British Journal of Educational Technology*, vol. 47, no. 1, pp. 15–18, 2016.

[11] T. An and M. Oliver, "What in the world is educa-tional technology ? rethinking the field from the perspective of the Philosophy of technology," *Learning, Media and Technology*, vol. 46, no. 1, pp. 6–19, 2021.

[12] R., J. S. Beichner and M. Jeffery, "Introduction to the SCALE -up (student-centered activities for large enrollment undergraduate programs) Project," *The International School of Physics Enrico Fermi,Varenna,Italy*, no. 6, pp. 1–17, 2003.

[13] G. Porter, K. Hampshire, and J. Milner, "Mobile phones and education inSub-saharan Africa: from youth practice to public policy," *Journal of International Development*, vol. 28, no. 1, pp. 2–5, 2016.

[14] A. K. Jena, "Does smart classroom an effective technology for teaching: a research analysis," *Journal Of Educational Technology*, vol. 10, no. 1, pp. 55–64, 2016.

[15] X. Lv, Y. Yang, D. Qin, X. Cao, and H. Xu, "'Artificial intelligence service recovery: the role of empathic response in hospitality 'customers' continuous usage intention," *Computers in Human Behavior*, vol. 126, Article ID 106993, 2022.

[16] L. Bornmann, "The problem of citation impact assessments for recent publication years in institutional evaluations," *Journal of Informetrics*, vol. 7, no. 3, pp. 722–729, 2013.

[17] L. Bornmann, F. de Moya Anegón, and L. Leydesdorff, "The new excellence indicator in the world report of the SCImago Institutions Rankings 2011," *Journal of Informetrics*, vol. 6, no. 2, pp. 333–335, 2012.

[18] J. Guan and N. Liu, "Measuring scientific research in emerging nano-energy field," *Journal of Nanoparticle Research*, vol. 16, no. 4, Article ID 2356, 2014.

[19] J. Huang, N. A. Shlobin, S. K. Lam, and M. DeCuypere, "Artificial intelligence applications in pediatric brain tumor imaging: a systematic review," *World Neurosurgery*, vol. 157, pp. 99–105, 2022.

[20] C. Huang, A. Notten, and N. Rasters, "Nanoscience and technology publications and patents: a review of social science studies and search strategies," *The Journal of Technology Transfer*, vol. 36, no. 2, pp. 145–172, 2011.

[21] Y. Huang, J. Schuehle, A. L. Porter, and J. Youtie, "'A systematic method to create search strategies for emerging technologies based on the Web of Science: illustrated for 'Big Data'," *'Scientometrics*, vol. 105, no. 3, pp. 2005–2022, 2015.

[22] S. Russell and J. Bohannon, "Artifcial intelligence. Fears of an AI pioneer," *Science*, vol. 349, no. 6245, p. 252, 2015.

[23] P. Shapira, S. Kwon, and J. Youtie, "Tracking the emergence of synthetic biology," *Scientometrics*, vol. 112, no. 3, pp. 1439–1469, 2017.

[24] Z. Wang, A. L. Porter, S. Kwon et al., "Updating a search strategy to track emerging nanotechnologies," *Journal of Nanoparticle Research*, vol. 21, no. 9, p. 199, 2019.

[25] X. Zhang, X. Ming, Z. Liu, D. Yin, Z. Chen, and Y. Chang, "A reference framework and overall planning of industrial artifcial intelligence (I-AI) for new application scenarios," *International Journal of Advanced Manufacturing Technology*, vol. 101, no. 9–12, pp. 2367–2389, 2019.

[26] G. Chartrand, P. M. Cheng, E. Vorontsov et al., "Deep learning: a primer for radiologists," *RadioGraphics*, vol. 37, no. 7, pp. 2113–2131, 2017.

Hindawi
Scientific Programming
Volume 2023, Article ID 9825727, 1 page
https://doi.org/10.1155/2023/9825727

Retraction

Retracted: Construction of the 3D Reconstruction System of Building Construction Scene Based on Deep Learning

Scientific Programming

Received 15 August 2023; Accepted 15 August 2023; Published 16 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Z. Lu, "Construction of the 3D Reconstruction System of Building Construction Scene Based on Deep Learning," *Scientific Programming*, vol. 2021, Article ID 5839391, 9 pages, 2021.

Hindawi
Scientific Programming
Volume 2021, Article ID 5839391, 9 pages
https://doi.org/10.1155/2021/5839391

Research Article

Construction of the 3D Reconstruction System of Building Construction Scene Based on Deep Learning

Zhou Lu

College of Architecture, South China University of Technology, Guangzhou 510641, Guangdong, China

Correspondence should be addressed to Zhou Lu; 201821005113@mail.scut.edu.cn

Received 11 November 2021; Revised 24 November 2021; Accepted 30 November 2021; Published 14 December 2021

Academic Editor: Baiyuan Ding

The increasing complexity and enormity of construction projects, as well as the fact that the actual operation of construction schedule management still mainly relies on traditional manual management methods, have led to low efficiency of construction schedule management and caused many construction projects to have cost overruns and legal disputes due to schedule delays. Existing 3D reconstruction algorithms often lead to significant voids, distortions, or blurred parts in the reconstructed 3D models, while the machine learning-based 3D reconstruction algorithms are often only to reconstruct simple separated objects and represent them as 3D boxes. A novel architecture of semisupervised 3D reconstruction algorithm is proposed. The algorithm iteratively improves the quality of the original 3D reconstruction model by training a generative adversarial network model to a converged state. Only the prior observed 2D images are required as weakly supervised samples, without any dependence on prior knowledge of the 3D structure shape or reference observations. Experimental results show that this algorithmic framework has significant advantages over the current state-of-the-art 3D reconstruction methods on the standard 3D reconstruction test set.

1. Introduction

A lot of research has been conducted on the topic of automated building construction schedule management with various technologies, but the existing research is hardly applicable to the complex building construction management practices [1–3]. These existing researches mainly focus on three aspects: management based on BIM (Building Information Modeling) technology [4–6], management based on RFID technology combined with BIM [7–9], and management based on Scan to BIM technology combined with 3D reconstruction technology [10–12]. For example, in schedule management, [13] conducted a study on building construction progress based on UAVs carrying Li DAR technology combined with BIM technology to achieve automatic monitoring of outdoor progress at building construction sites [14]. However, the existing automated construction schedule management approach suffers from two drawbacks.

For one, the high equipment dependence causes high management costs, such as Li DAR equipment generally costing tens of thousands of dollars, and the high cost of the UAV equipment required for tilt photography and the high maintenance costs during use make it difficult to apply in the actual management process [15–17].

Second, the poor operability of automation leads to a low level of automation, such as the use of Li DAR equipment has high requirements for the field environment [18], while the tilt photography method requires trained UAV professionals to operate and requires the implementation of work in specific airways and in practice to consider complex issues such as obstacle avoidance [19–22], requiring a high degree of human involvement.

Artificial intelligence technologies such as deep learning have gradually demonstrated strong productivity in the field of construction engineering in recent years [23–25], while a low-cost, automated, and intelligent construction schedule management method that can be applied to the construction site environment in combination with artificial intelligence technologies has yet to be studied [26]. In the field of computer vision and computer graphics, 3D reconstruction is a technique for recovering the shape, structure, and

appearance of real objects. Due to its rich and intuitive expressiveness, in this paper, we propose a 3D reconstruction algorithm based on semisupervised generative adversarial networks, which combine the advantages of traditional 3D reconstruction techniques with the latest machine learning principles of generative adversarial networks. By fine-tuning the adversarial training process of the 3D generative model and the 3D discriminative model simultaneously, the framework proposed in this paper can steadily refine the reconstruction quality of the reconstructed 3D objects in a semisupervised learning manner. On the basis of this algorithm, a 3D reconstruction cloud studio is also built to provide a convenient and accessible 3D reconstruction cloud service system to a wide range of users.

2. Related Work

The targets of 3D reconstruction can be some detached objects [27] or large scale scenes [28, 29]. For different reconstruction targets, researchers will try to present the reconstructed 3D models in different ways. Common forms of presentation include stereo body elements [6], point clouds [8], and a combination of mesh skeleton and surface textures [30]. In recent years, researchers have made great progress in the research of new methods for 3D reconstruction techniques.

This class of algorithms first performs feature matching based on two images, then uses the obtained dual-view reconstruction results to initialize the 3D model, adds new matching images and iterates repeatedly to perform triangular feature matching, and uses the beam leveling method to recover the motion structure. The time complexity of this class of algorithms is $O(n4)$, where n represents the number of observed cameras. The most representative algorithm in this class is VisualSFM [31], which further improves the computational performance and optimizes a large number of time-consuming steps including the beam leveling method.

However, such algorithms also have obvious limitations; they are all based on the important assumption that feature information is perfectly watchable across multiple viewpoints. If the spatial distance between the views is large, feature matching becomes extremely difficult due to local appearance changes or mutual occlusion. Another limitation is that if the surface of the object to be reconstructed lacks texture information, or if there are specular reflections on the surface, the feature matching process is likely to fail completely.

The most famous algorithm of this class is Kinect Fusion [32], which is able to continuously track and solve the pose information of the depth camera in 6 degrees of freedom by means of the detected depth information. The tracking accuracy of this method is significantly better than the 3D reconstruction method based on motion structure recovery (since this approach can only track the camera poses by matching features from frame to frame of color pictures). By iteratively fusing the depth and pose information into a dense global stereo model, the final output of the constructed 3D model is achieved. In Whelan's work [8], he further improves the tracking accuracy, tenacious robustness, and reconstruction quality based on KinectFusion. The improved algorithm uses techniques such as dense image frames corresponding to camera tracking of the model, sliding window point element fusion, and nonrigid surface deformation to obtain a higher quality 3D reconstructed model.

The limitations of this type of algorithm mainly lie in the existence of self-obscuration, light reflection, and depth sensor fusion errors, which can lead to significant voids, distortions, or blurred parts of the reconstructed 3D model.

The representative algorithm of this class is the 3D Recurrent Reconstruction Neural Network (3D-R2N2) [33] algorithm, which uses a deep CNN to learn the mapping relationship between the observed 2D image and the corresponding 3D shape of the target object from a large training dataset.

The most representative algorithm in this class is the 3D-GAN [7] algorithm. 3D-GAN algorithm introduces generative adversarial loss and uses it as a judging criterion to distinguish whether an object is real or reconstructed. Because 3D objects are highly structured, the use of generative adversarial loss is more effective than the traditional voxel-level independent heuristic judging criterion, which can capture the subtle differences of the 3D structure of the target object more accurately.

3. 3D Reconstruction Algorithm Is Based on Semisupervised Generative Adversarial Network

3.1. Algorithm Principle. Imagine an example where an observer wants to distinguish a real scene from an artificially reconstructed model of the scene. First, he would observe in the real 3D scene, and then he would also observe in the reconstructed 3D scene model, and the position and perspective of each observation would be the same as when he was in the real scene. If he observes a series of two-dimensional pictures in the reconstructed 3D scene model, which are exactly the same as what he observes in the real 3D scene, then it is actually extremely difficult for the observer himself to distinguish which is the real 3D scene and which is the reconstructed 3D scene model. In order to construct a 3D reconstruction algorithm, the differences between each set of 2D pictures were observed in the real scene, and the 2D pictures observed in the reconstructed scene model can be accumulated. If such differences are small enough for each observation position and viewpoint, this reconstructed 3D model can be considered to be of high quality. And from a quantitative point of view, the smaller the accumulated differences, the higher the quality of the reconstructed 3D model. This can be taken as the final criterion for judging the 3D reconstructed model. A more intuitive representation of this concept is shown in Figure 1.

From Figure 1, it can be seen that the improved algorithm runs significantly more efficiently than the original Apriori algorithm [34], especially when the support elucidation value is low.

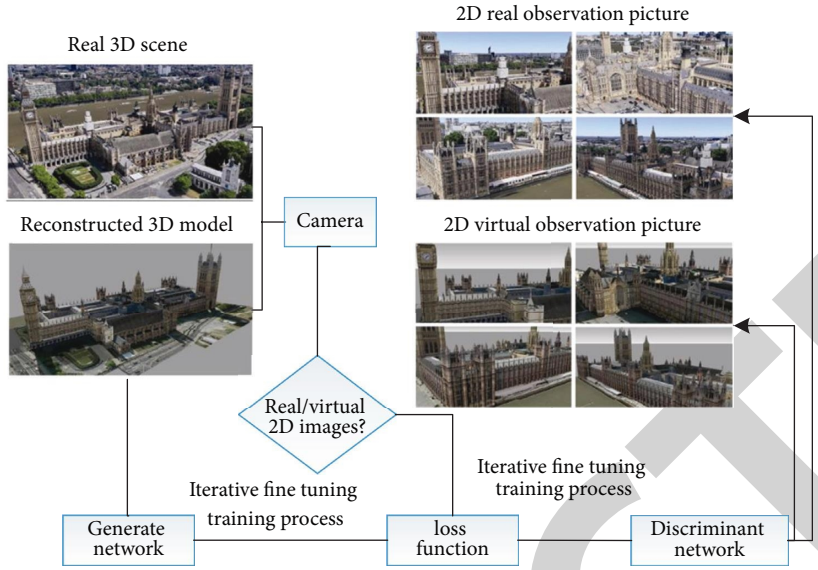


Figure 1: Principle and algorithm flow of 3D reconstruction algorithm based on GAN.

The discriminant network also estimates how likely it is that a particular sample is a sample synthesized by the generative network. When the entire generative adversarial network model reaches Nash equilibrium, i.e., the generative network can produce new samples with characteristics and distributions identical to those of the real samples, and the discriminant network outputs a discriminant probability of 0.5 for each pair of real and generative sample sets, the entire generative adversarial network model completes training and reaches convergence.

Combining the goal of 3D reconstruction with a generative adversarial network model, a new 3D reconstruction architecture is developed in this paper: a 3D reconstruction network based on semisupervised generative adversarial network (SS-GAN-3D). SS-GAN-3D is composed of a 3D model generation network and a 3D model discriminator network together. Here, the discriminator network can be imagined as the observer mentioned in the above example. In this way, the goal of the generative network is to reconstruct a 3D model that is extremely similar to the real 3D scene and try to confuse the discriminator network with this 3D scene model [35]. The goal of the discriminative network is to clearly distinguish the difference between the real 3D scene and the reconstructed 3D model. In this way, it also meets the measure of 3D reconstructed model quality given above. In conclusion, the new proposed architecture in this paper equivalently transforms the traditional 3D reconstruction solution problem into a machine learning problem that trains SS-GAN-3D and achieves convergence.

3.2. Algorithm Flow. When SS-GAN-3D is trained, an extremely rough 3D model is first generated as an initialization of the 3D model generation network. Here, the rough 3D model is represented in ".ply" format. The vertex, edge, and color information are stored in a triplet format [13]. The spatial stereo matching method estimates the depth information of each point on the image of space by comparing the differences between adjacent observed image frames. Also, two-dimensional observation images truncated from the video stream are used to from the truth-value image dataset.

Since SS-GAN-3D requires 2D observation images from the reconstructed 3D model, the reconstructed 3D model is imported into the professional open-source 3D engine software blender and OpenDR [14]. OpenDR is a differentiable renderer that realistically approximates the realistic rendering from the 3D model to the 2D image and, at the same time, can provide the gradient change from the 2D image to the 3D model required by the backpropagation algorithm. It is a differentiable renderer that gives a realistic approximation of the rendering of the 3D model to the 2D image, while providing the gradient change from the 2D image to the 3D model required by the backpropagation algorithm. The differentiability of the renderer is essential, because the structure of the generative adversarial network needs the entire network to be fully differentiable, so that the gradient changes of the discriminative network can be passed back to update the generative network and form a complete circular iterative structure.

In Blender, a virtual camera can be set up with the exact same optical parameters as the real camera used to capture the video stream in the real 3D scene. When processing the real video stream, the camera's trajectory is already calculated. So, in Blender, the virtual camera is made to move along this trajectory and, using the OpenDR renderer, is observed at the same position and perspective as in the real scene and rendered to generate a 2D image. In this way, the same number of 2D virtual and real observation images can be obtained from the reconstructed 3D model and the real 3D scene, respectively.

With a collection of 2D virtual and real observation images, a discriminant network is used to distinguish whether they come from observations of the real 3D scene or

of the reconstructed 3D model. The loss value of the whole network is also calculated based on the loss function. With the network loss values, SS-GAN-3D can continue to fine-tune the training process to generate new 3D generative and 3D discriminative networks. The newly trained 3D generative network will reconstruct a new 3D model for the virtual camera to make observations. The virtual observation images from the new observations are fed together with the real original observation that the SS-GAN-3D is trained iteratively and continues to generate new 3D generative and discriminative networks until the overall loss value converges to a desired threshold.

3.3. Definition of Loss Function. The overall loss function of SS-GAN-3D contains two parts: the reconstruction loss ReconsL and the cross-entropy loss SS-GAN-3DL. So, the loss function can be written as

$$L_{\text{Overall}} = L_{\text{Recons}} + \lambda L_{\text{SS-GAN-3D}}, \tag{1}$$

where λ is the parameter value that regulates the reconstruction loss and cross-entropy loss weights.

In this paper, three quantitative measures of image quality [15] are selected for calculating the differences. Peak signal-to-noise ratio (PSNR) quantifies the picture different from the perspective of gray value fidelity. Structural similarity (SSIM) [16] quantitatively measures picture differences from the perspective of structural-level fidelity, while this metric refers to and simulates the judgment criteria of the human eye system for structural patterns. Normalized correlation (NC) [36], on the other hand, indicates the matrix similarity of pictures with the same dimension. Expressions of these 3 evaluation quantitative metrics are shown as follows:

$$\text{PSNR}(\mathbf{x}, \mathbf{y}) = 10\lg\left(\frac{(\text{MAX}_I)^2}{\text{MSE}(\mathbf{x}, \mathbf{y})}\right), \tag{2}$$

where MAX_I represents the maximum value that can be obtained for each pixel in images x and y. $\text{MSE}(\mathbf{x}, \mathbf{y})$ represents the mean square error of picture x and y.

$$\text{SSIM}(\mathbf{x}, \mathbf{y}) = \frac{\left(2\mu_x\mu_y + C_1\right)\left(2\sigma_{xy} + C_2\right)}{\left(\mu_x^2 + \mu_y^2 + C_1\right)\left(\sigma_x^2 + \sigma_y^2 + C_2\right)}. \tag{3}$$

Among them, $\mu_x = 1/N\sum_{i=1}^{N} x_i$ and $\mu_y = 1/N\sum_{i=1}^{N} y_i$ represent the average gray values of pictures X and y. σ_x and σ_y represent the variance of pictures X and y. σ_{xy} represents the covariance of pictures X and y. Parameters C_1 and C_2 are two constants. When $\mu_x^2 + \mu_y^2$ or $\mu_x^2 + \mu_y^2$ is very close to 0, C_1 and C_2 can prevent divergent results from the final SSIM.

$$\text{NC}(\mathbf{x}, \mathbf{y}) = \frac{(\mathbf{x} \cdot \mathbf{y})}{\|\mathbf{x}\|\|\mathbf{y}\|}, \tag{4}$$

where $\mathbf{x} \cdot \mathbf{y}$ represents the inner product of matrices X and y, and operator$|.|$ represents the Euclidean norm of the vector.

Obviously, the structural similarity index of the two pictures is 0~1, and the normalized correlation index is −1~1. If SSIM index or NC index is very close to 1, the gap between X and Y is very small. For the peak signal-to-noise ratio index, the value of common pictures is 20~70 dB, which needs to be normalized by generalized sigmoid function.

$$E\_\text{Sig}(\text{PSNR}(\mathbf{x}, \mathbf{y})) = \frac{1}{1 + e^{-0.1(\text{PSNR}(x,y)-45)}}. \tag{5}$$

Therefore, the final reconstruction loss can be written in the following form:

$$L_{\text{Recons}} = \sum_{j=1}^{N}\left\{\alpha \cdot \left[1 - E\_\text{Sig}\left(\text{PSNR}_{G_jF_j}\right)\right] + \beta\left(1 - \text{SSIM}_{G_jF_j}\right) + \gamma \cdot \left(1 - \text{NC}_{G_jF_j}\right)\right\}. \tag{6}$$

Among them, α, β, γ are parameters to adjust the proportion of PSNR, SSIM, and NC indicators in the overall loss value. Subscript G_jF represents a pair of real and virtual observation two-dimensional pictures. The superscript n indicates the total number of such picture pairs in the picture set. Section 3.4 will discuss the cross-entropy loss of ss-gan-3d in detail in combination with the network structure.

3.4. Network Structure of SS-GAN-3D. For SS-GAN-3D, the discriminant network needs strong classification performance to deal with the complex two-dimensional slices generated by three-dimensional spatial projection. Therefore, this paper adopts ResNet-101 network [37] as the main structure of discrimination network. Typical ResNet networks adopt back normalization, which makes the whole training process more stable. However, the introduction of batch normalization operation makes the discriminant network judge the mapping relationship between a batch of input and a batch of outputs. In ss-gan-3d, it is hoped to ensure the mapping relationship between single input and single output in the training process. In order to improve the training effect, the ReLu layer is also replaced with a parametritis ReLu layer. In order to improve the convergence performance, Adam solver is actually used to replace the random gradient descent (SGD) solver. In practical application, Adam solver can make ss-gan-3d train at a large learning rate. The detailed network hierarchy is shown in Figure 2.

According to the researchers' experiments, as of now, only the Wasserstein GAN (WGAN) [18] structures with the addition of gradient penalty restriction can successfully train complex generative and discriminative networks similar to the ResNet structure. Therefore, in this paper, we borrow the improved training algorithm of WGAN and apply it to the training process of SS-GAN-3D. The objective functions for

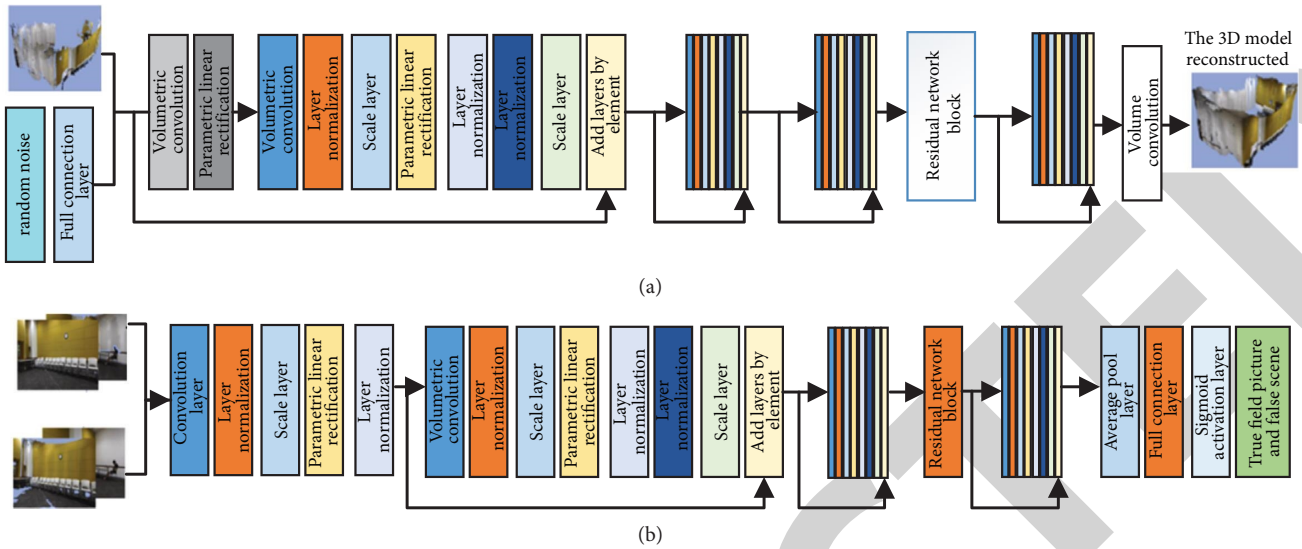


FIGURE 2: Structure of 3D generative and discriminative network. (a) 3D generation network structure. (b) Three-dimensional discriminant network structure.

training the generative network G and the discriminative network D are as follows:

$$\min_{G}\max_{D} \mathop{\mathbb{E}}_{x\sim\mathbb{P}_f}[D(x)] - \mathop{\mathbb{E}}_{i\sim\sim_g}[D(\tilde{x})], \tag{7}$$

where $\mathbb{P}_r$ represents the distribution of the real images, and $\mathbb{P}_g$ represents the distribution of the generated images. $\tilde{x}$ is the implicit output result of the generative network G. During the training of the original version of WGAN, the clipping of the weight values could easily lead to optimization failure, including network performance degradation, gradient explosion, or gradient disappearance. In the improved version of the gradient penalty, it is used as a looser constraint instead of simple weight cropping. So, the final cross-entropy loss of SS-GAN-3D is

$$\begin{aligned} L_{\text{SS-GAN-3D}} = {} & \mathop{\mathbb{E}}_{x\sim\mathbb{P}_r}[D(x)] - \mathop{\mathbb{E}}_{\tilde{x}\sim\mathbb{P}_g}[D(\tilde{x})] \\ & - \theta \mathop{\mathbb{E}}_{\hat{x}\sim\mathbb{P}_i}\left[\left(\left\|\nabla_{\hat{x}} D(\tilde{x})\right\|_2 - 1\right)^2\right], \end{aligned} \tag{8}$$

where θ is the parameter that regulates the percentage of the gradient penalty in the cross-entropy loss. $\mathbb{P}_{\hat{x}}$ indicates that the value of the cross-entropy loss of the dataset formed by uniform sampling on the straight line between pairs of sample points obtained from the $\mathbb{P}_r$ and $\mathbb{P}_g$ distribution can quantitatively reflect the training process of SS-GAN-3D. The smaller this value, the smaller the Wasserstein distance between the real and virtual 2D observed images.

4. Simulation Results

4.1. Modeling Effect. Based on the system's high-speed camera's acquisition of image data from all angles of the real-time scenes of the construction site (as shown in Figure 3), the DLR-P system automatically analyzes the real-time scenes of the project construction site and obtains the actual progress in the form of point cloud models for each of the three construction processes as shown in Figure 4. By cross-referencing the point cloud model with the idea BIM point cloud, the difference between the actual construction progress and the expected ideal progress is automatically calculated [33].

The read position in the leftmost figure indicates the comparison of the progress of sensor acquisition angle. By comparing the ideal BIM model converted into point cloud format (containing 3D information, construction schedule, and cost plan information) with the actual 3D point cloud model of the project site automatically identified by the 3D reconstruction technology based on deep learning, the difference of the construction site schedule relative to each plan is derived (as shown in Table 1). On this basis, the DLR-P system automatically adjusts the construction site plan to meet the total construction schedule and automatically provides on-site labor, material, and machinery resource responses according to the project volume and duration.

Operation speed: in order to realize the real-time automated management of the DLR-P system for construction projects, the time consumed by the 3D reconstruction process of various scenes was recorded, as shown in Table 2. The running speed is the time required from the moment when the high-speed camera acquires the image until the moment when the system outputs the final point cloud model. However, since the 3D reconstruction process mainly involves two parts, sparse reconstruction and dense reconstruction, the respective time consumed is related to many factors such as the number of reconstructed relevant images, image resolution, system background computing power, and the complexity of the images. Therefore, system operation speed recorded in the case study only represents the average speed required for the 3D reconstruction of the relevant scenes.

Figure 3: Real-time image data of the construction site from various angles: (a) construction of scaffolding on an exterior wall; (b) construction of protective net on an exterior wall; (c) construction of paint on the exterior wall.

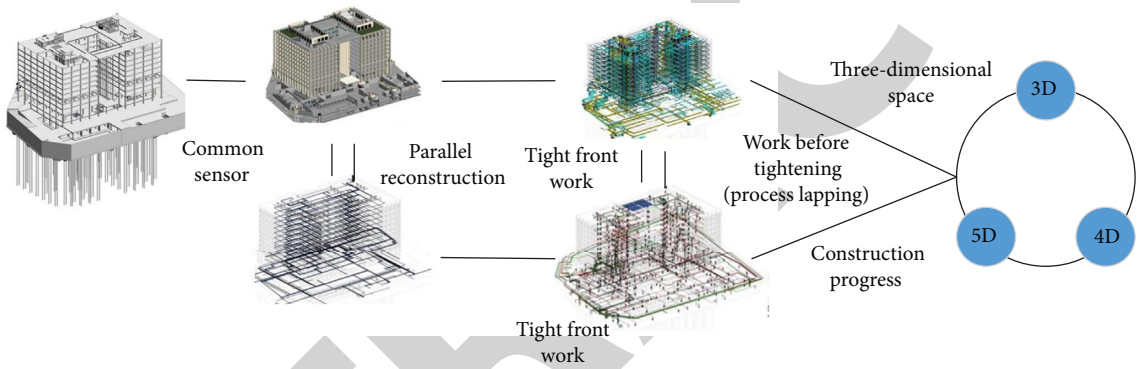


Figure 4: DLR-P system operation results.

Table 1: Progress plan variance and response.

Construction procedure		Construction of external wall scaffold	Construction of external wall protective net	Exterior wall coating construction
Schedule		4th floor	3th floor	4th floor
Actual progress		4th floor	4th floor	3th floor
Schedule variance		0	+1	−1
Plan adjustment		0	−1	+1
Resource response	Artificial	0	Workers—2	+3
	Material science	0	0	Paint +10 barrels
	Mechanics	0	0	0

4.2. System Operation. Operating costs: as shown in Table 2, the DLR-P system achieves fully automated construction schedule control without manual labor, and its main operating cost consists of the hardware cost of both the system backend and the system sensors, and its hardware equipment cost is only $33,000. The hardware cost is only $33,000, while the hardware cost of the management method based on the UAV method is about $370,000, and the cost of the handheld Li DAR equipment-based method is higher, about $820,000. During the case study of this project, only part of the construction of the project was studied, so if the whole project is controlled, the deployment cost of the DLR-P system should be higher than the above data, mainly due to the increase in the number of camera sensors. However, the DLR-P system proposed in this paper still has significant cost advantages compared to the other two schedule management implementations.

As depicted in Figure 5, the accuracy of calculation of construction volume is low, and the phenomena such as underinvestment or waste can occur. Based on 3D design and collaborative design technology, a more feasible and accurate construction plan can be simulated by the construction profession from the process design, thus providing the estimation profession with a relatively accurate base

Table 2: The speed of 3D reconstruction of our system.

Construction procedure	Enter the number of images	Image fraction	Average operating speed (S)
Construction of external wall scaffold	10	1600 × 1200	31
		2400 × 1800	38
		3200 × 2400	55
	20	1600 × 1200	45
		2400 × 1800	68
		3200 × 2400	75
	30	1600 × 1200	69
		2400 × 1800	105
		3200 × 2400	152
Construction of external wall protective net	10	1600 × 1200	31
		2400 × 1800	43
		3200 × 2400	25
	20	1600 × 1200	40
		2400 × 1800	54
		3200 × 2400	59
	30	1600 × 1200	62
		2400 × 1800	92
		3200 × 2400	101
Rice square fish	10	1600 × 1200	23
		2400 × 1800	35
		3200 × 2400	40
	20	1600 × 1200	44
		2400 × 1800	59
		3200 × 2400	58
	30	1600 × 1200	53
		2400 × 1800	88
		3200 × 2400	105
Average value			61

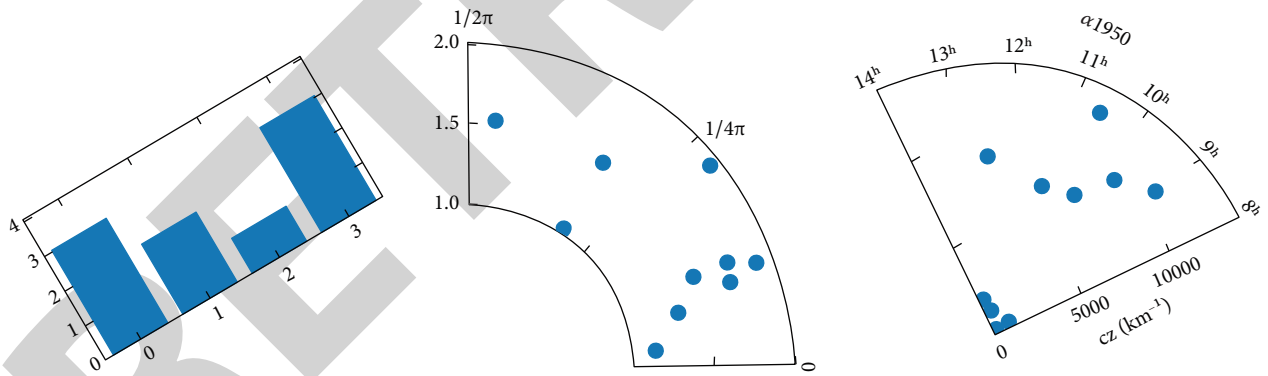


Figure 5: Different pattern generations.

information for the preparation of the estimated unit price and making the basis for the calculation of the whole project investment relatively accurate and reliable. In addition, the shared collaborative design platform can make use of the linkage of its various specialties to update the information on the changes in construction volume generated by design modifications at any time and link it with its own estimates. It can not only greatly improve the work efficiency, but also effectively reduce the unnecessary design errors due to the coordination of various professions [4]. By integrating more nongeometric information such as price parameters, market information, and price change factors into the 3D model, the construction process or project plan can be compared from the perspective of engineering cost, effectively reducing design changes and making engineering investment more accurate and reasonable.

By integrating more nongeometric information such as price parameters, market information, and price change factors into the 3D model, the construction process or project plan can be compared from the perspective of engineering cost, effectively reducing design changes and making the project investment more accurate and reasonable. The construction camp layout is shown in Figure 6, which utilizes Infraworks' intuitive and concise 3D dynamic display function through terrain analysis.

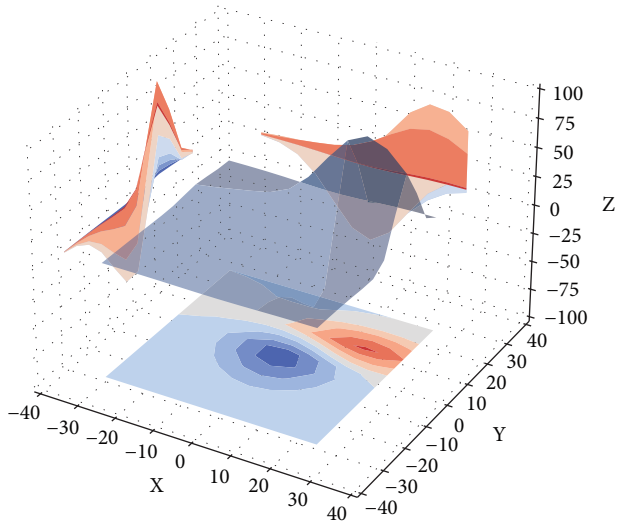


Figure 6: Different site terrain generations.

5. Conclusions

Existing 3D reconstruction algorithms often lead to the existence of obvious voids, distorted distortion, or blurred parts on the reconstructed 3D models, while the machine learning-based 3D reconstruction algorithms can often only reconstruct simple separated objects and represent them in the form of 3D boxes. So, all these algorithmic frameworks are far from sufficient for practical applications. Therefore, the focus of this paper is to use the production adversarial network principle to obtain high-quality 3D reconstruction results. Only the prior observed 2D images are required as weakly supervised samples, and there is no dependence on the prior knowledge of the 3D structure shape or the reference observation. Experimental results show that this algorithmic framework has significant advantages over the current state-of-the-art 3D reconstruction methods on the standard 3D reconstruction test set.

Data Availability

The datasets used in this paper are available upon request to the author.

Conflicts of Interest

The author declares no conflicts of interest regarding this work.

References

[1] L. Xin, X. Liu, Z. Yang, X. Zhang, Z. Gao, and Z. Liu, "Three-dimensional reconstruction of super-resolved white-light interferograms based on deep learning," *Optics and Lasers in Engineering*, vol. 145, no. 12, Article ID 106663, 2021.

[2] H. S. Cha and J. Kim, "A study on 3D/BIM-based on-site performance measurement system for building construction," *Journal of Asian Architecture and Building Engineering*, vol. 19, no. 6, pp. 574–585, 2020.

[3] X. D. Yang, B. H. Liu, and Y. Wang, "3D reconstruction of CT images based on isosurface construction," *Applied Mechanics and Materials*, vol. 10-12, pp. 503–507, 2008.

[4] M. R. Jahanshahi and S. F. Masri, "Adaptive vision-based crack detection using 3D scene reconstruction for condition assessment of structures," *Automation in Construction*, vol. 22, pp. 567–576, 2012.

[5] R. Li, T. Zeng, H. Peng, and S. Ji, "Deep learning segmentation of optical microscopy images improves 3-D neuron reconstruction," *IEEE Transactions on Medical Imaging*, vol. 36, no. 7, pp. 1533–1541, 2017.

[6] M. O. Malavé, C. A. Baron, S. P. Koundinyan et al., "Reconstruction of undersampled 3D non-Cartesian image-based navigators for coronary MRA using an unrolled deep learning model," *Magnetic Resonance in Medicine*, vol. 84, no. 2, pp. 800–812, 2020.

[7] J. Chen, Z. Kira, and Y. K. Cho, "Deep learning approach to point cloud scene understanding for automated scan to 3D reconstruction," *Journal of Computing in Civil Engineering*, vol. 33, no. 4, pp. 04019027.1–04019027.10, 2019.

[8] Y. Liu, S. Chen, and B. Guan, "Layout optimization of oil-gas gathering and transportation system in constrained three-dimensional space," *Chinese Science Bulletin*, vol. 65, no. 9, pp. 834–846, 2020.

[9] Y. Liu, S. Chen, B. Guan, and P. Xu, "Layout optimization of large-scale oil-gas gathering system based on combined optimization strategy," *Neurocomputing*, vol. 332, no. 7, pp. 159–183, 2019.

[10] B. Guan, S. Chen, Y. Liu, X. Wang, and J. Zhao, "Wave patterns of (2+1)-dimensional nonlinear Heisenberg ferromagnetic spin chains in the semiclassical limit," *Results in Physics*, vol. 16, Article ID 102834, 2019.

[11] X. Zhang, J. Yao, L. Dong, and N. Ye, "Research on 3D architectural scenes construction technology based on augmented reality," *Journal of Computational Methods in Science and Engineering*, vol. 21, no. 1, pp. 1–17, 2020.

[12] F. Chen, Mm Ad K. Muha, and S. H. Wang, "Three-dimensional reconstruction of CT image features based on multi-threaded deep learning calculation," *Pattern Recognition Letters*, vol. 136, pp. 309–315, 2020.

[13] D. Yu, S. Ji, J. Liu, and S. Wei, "Automatic 3D building reconstruction from multi-view aerial images with deep learning," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 171, no. 2021, pp. 155–170, 2021.

[14] G. Fahim, K. Amin, and S. Zarif, "Single-view 3D reconstruction: a survey of deep learning methods," *Computers & Graphics*, vol. 94, no. 1, 2021.

[15] F. Kurz, S. Azimi, C. Y. Sheu, and P. D. Angelo, "Deep learning segmentation and 3D reconstruction of road markings using multiview aerial imagery," *International Journal of Geo-Information*, vol. 8, no. 1, 2019.

[16] G. Stuebl, C. Heindl, H. Bauer, and A. Pichler, "Deep learning based aesthetic evaluation of state-of-the-art 3D reconstruction techniques," in *Proceedings of the 3DBODY.TECH 2017 - 8th International Conference and Exhibition on 3D Body Scanning and Processing Technologies*, Montreal, Canada, Octobar, 2017.

[17] Q. Xie, X. Chen, L. Shen, and G. Li, "Brain microstructure reconstruction based on deep learning," *Systems Engineering-Theory & Practice*, vol. 38, no. 2, pp. 482–491, 2018.

[18] J. Teuwen, N. Moriakov, C. Fedon et al., "Deep learning reconstruction of digital breast tomosynthesis images for accurate breast density and patient-specific radiation dose estimation," *Medical Image Analysis*, vol. 71, no. 22, Article ID 102061, 2021.

[19] S. Sharma and V. Kumar, "Voxel-based 3D face reconstruction and its application to face recognition using