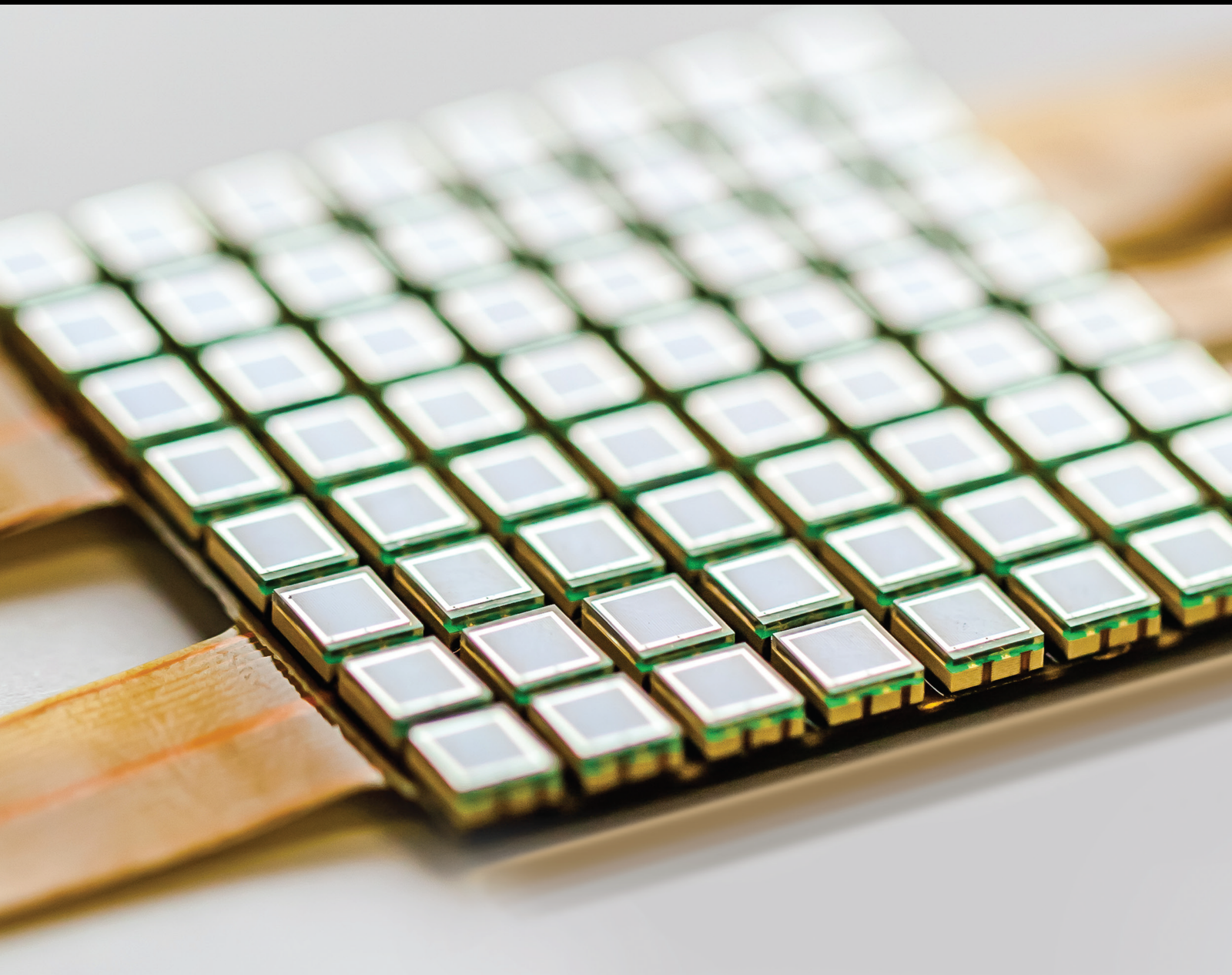# Information Fusion and Its Applications for Smart Sensing 2021

Lead Guest Editor: Mu Zhou
Guest Editors: Ying-Ren Chien and Qiao Zhang

Information Fusion and Its Applications for Smart Sensing 2021

Information Fusion and Its Applications for Smart Sensing 2021

Lead Guest Editor: Mu Zhou
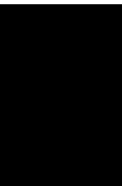Guest Editors: Ying-Ren Chien and Qiao Zhang

Ehsan Namaziandost (iD), Iran
Heinz C. Neitzert (iD), Italy
Sing Kiong Nguang (iD), New Zealand
Calogero M. Oddo (iD), Italy
Tinghui Ouyang, Japan
SANDEEP KUMAR PALANISWAMY (iD), India
Alberto J. Palma (iD), Spain
Davide Palumbo (iD), Italy
Abinash Panda (iD), India
Roberto Paolesse (iD), Italy
Akhilesh Pathak (iD), Thailand
Giovanni Pau (iD), Italy
Giorgio Pennazza (iD), Italy
Michele Penza (iD), Italy
Sivakumar Poruran, India
Stelios Potirakis (iD), Greece
Biswajeet Pradhan (iD), Malaysia
Giuseppe Quero (iD), Italy
Linesh Raja (iD), India
Maheswar Rajagopal (iD), India
Valerie Renaudin (iD), France
Armando Ricciardi (iD), Italy
Christos Riziotis (iD), Greece
Ruthber Rodriguez Serrezuela (iD), Colombia
Maria Luz Rodriguez-Mendez (iD), Spain
Jerome Rossignol (iD), France
Maheswaran S, India
Ylias Sabri (iD), Australia
Sourabh Sahu (iD), India
José P. Santos (iD), Spain
Sina Sareh, United Kingdom
Isabel Sayago (iD), Spain
Andreas Schütze (iD), Germany
Praveen K. Sekhar (iD), USA
Sandra Sendra, Spain
Sandeep Sharma , India
Sunil Kumar Singh Singh (iD), India
Yadvendra Singh (iD), USA
Afaque Manzoor Soomro (iD), Pakistan
Vincenzo Spagnolo, Italy
Kathiravan Srinivasan (iD), India
Sachin K. Srivastava (iD), India
Stefano Stassi (iD), Italy

Danfeng Sun, China
Ashok Sundramoorthy, India
Salvatore Surdo (iD), Italy
Roshan Thotagamuge (iD), Sri Lanka
Guiyun Tian (iD), United Kingdom
Sri Ramulu Torati (iD), USA
Abdellah Touhafi (iD), Belgium
Hoang Vinh Tran (iD), Vietnam
Aitor Urrutia (iD), Spain
Hana Vaisocherova - Lisalova (iD), Czech Republic
Everardo Vargas-Rodriguez (iD), Mexico
Xavier Vilanova (iD), Spain
Stanislav Vítek (iD), Czech Republic
Luca Vollero (iD), Italy
Tomasz Wandowski (iD), Poland
Bohui Wang, China
Qihao Weng, USA
Penghai Wu (iD), China
Qiang Wu, United Kingdom
Yuedong Xie (iD), China
Chen Yang (iD), China
Jiachen Yang (iD), China
Nitesh Yelve (iD), India
Aijun Yin, China
Chouki Zerrouki (iD), France

Contents

Research Article

Multimedia Analysis of Digital Museum User Interface Based on Goal-Oriented Theory and Information Fusion and Intelligent Sensing

Qingbin Zhuang [1], **Wanni Xu** [2], **Danyang Yang**,[1] and **Ning Wei** [3]

¹*Fine Art and Design College, Quanzhou Normal University, Quanzhou 362000, China*
²*Department of Fine Arts, Xiamen Academy of Arts Design, Fuzhou University, Xiamen 361024, China*
³*Department of Cultural Industry, Fujian Normal University, Fuzhou 350117, China*

Correspondence should be addressed to Qingbin Zhuang; zqb2021@qztc.edu.cn and Wanni Xu; 80760017t@ntnu.edu.tw

Introduction. The development of network technology is promoting the process of digitalization. The digital museum, an emerging museum display mode, is gradually highlighting its great value in the wave of digitalization. In the context of the rapid development of digital museums and mobile applications, the user interface of digital museums, as the integration point of human-computer interaction, the artistic expression of its visual design is also more important. This paper takes the digital museum application (APP) user interface design as the research direction. By optimizing the visual design of the interface, the user's operating experience is improved, so that users can enjoy an orderly, time-saving, efficient, comfortable, and interesting interactive experience. *Methods.* User interface design to analyze related theories such as target-oriented design and visual hierarchy design and dig out the manifestation of user goals in the museum's APP visual hierarchy and clarify the design content. Use qualitative and quantitative user research methods to conduct demand research for museum users and build user role models. Build user role target task model. Determine the function settings of the museum's APP and establish an information structure. Secondly, build a low-fidelity model through the level analysis of the visual elements in the interface and use the low-fidelity prototype test to guide the museum's APP visual level design. The interface visual level elements are integrated into the museum's APP visual level framework and integrated with the museum's characteristics Together, put forward the visual hierarchical design strategy of the museum APP. *Results.* A new model for the dissemination and display of cultural information resources by transposing the display of museum information content fusion to the works of the digital platform. The ease of use, fun, and artistic quality of the user interface of the digital museum are the keys to attracting users. Through the summary of the theory and the construction of the design strategy, we build and design a museum APP that meets the user's experience, meets the user's goals, and has a good visual hierarchy. The experimental results show that the largest number of people, 62.6%, wanted to learn about local culture by visiting museums. The number of people whose purpose was to travel and the study was the next highest, and they accounted for an equal share, 43.1% and 40.6%, respectively. A smaller number, 23.4%, attended museums for hobbies and interests. The number of people who visit museums for research purposes is even lower, at 11.9%. Through usability testing and user satisfaction analysis of the design model, the rationality and effectiveness of the design strategy are verified, which can provide appropriate guidance for museum app design. *Conclusion.* The problems encountered in the visual design and production of digital museums and the solutions were discussed, and through a comparative analysis with other digital museum visual designs, it provided theoretical and practical basis for the construction and research of digital museum visual design in the future. Our digital museum user interface design is based on user needs and solves the previous problem of users not understanding how to use the museum app effectively.

1. Introduction

As a window for displaying culture and disseminating knowledge to the audience, the digital museum makes full use of the advantages of the "fast communication" of the Internet, extracts cultural information points in the collection, and selects 720-degree panoramic roaming and 360-degree surrounding objects to promote China's excellent traditional culture and spread the works of Chinese civilization. Digital museums enable the information fusion and the need for audiences to use digital technology to view exhibitions and learn about heritage wherever they are, whenever they are [1]. Han [2] suggests that user interfaces can meet the basic needs of users by integrating the disciplines of technology, art, and psychology. Text images [3], interface colors [4], interactive buttons [5], and overall layout [6] are important design elements. Designers must rationally design these elements according to the characteristics of different museums to form an organic whole, so as to achieve the unity of ease of use and interest, thereby inspiring the audience [7].

Li [8] puts forward suggestions for establishing a brand image, strengthening the publicity of museums, and enhancing the interest of museum app through the problems of existing museum app. Due to the imperfect information management of the collections and the outdated visiting mode of traditional museums, under such circumstances, digital museums have emerged [9]. Digital museums use text, images, audio, video, and other media forms for information transmission. Virtual reality technology and three-dimensional technology promote the growth of digital museums, making them look like physical museums [10–12]. Its basic advantages are as follows: it can break through the limitation of exhibition space, making the number of exhibitions and exhibition collections unrestricted; it can break through geographical restrictions, so that more people can visit and browse without visiting physical museums; and it can use museum resources as content to pass establishing an educational network connection can assist students in classroom learning and enable audiences to get a good learning experience through a new interactive form [13, 14].

Goal-oriented design is a design that starts with the specific needs of the user [15, 16], Chu [17] found that a simple and clean interface is good for the user to read and a moderate interaction experience can bring a good feeling to the user. The user interface is also known as the human-machine interface or human-computer interface. It consists of the intersection of human and computer hardware and software, between the user and the computer system, is the medium for the transmission and exchange of information between the human and the computer, and is a comprehensive operating environment for the user to use the computer system [18–20]. User-friendliness refers to the complexity of the subjective operation of the user operating system. For example, the lower the complexity of the subjective operation, the easier the system is to use, the higher the user-friendliness of the system [21]. The user interface's ease of use, fun, and artistry are all key to attracting readers [22]. Liu [23] suggests in his goal-oriented interaction design about users that it is important for designers to understand users' psychological perceptions and usage habits. Sun [24] argues that a visual interface that meets users' needs is not limited to local aesthetics, but rather the comfort of the overall interface. As shown in Figures 1(a)–1(c), an example of a successful interface design for the navigation function, the uniform interface style is designed to make it more accessible to users.

The museum's first attempt to introduce mobile phone APP applications was in Japan, and there are many museum APP applications researched and developed in European and American countries. The APP applications of excellent and mature museums abroad include the following: The British Museum, Musée du Louvre, Metropolitan Museum of Art, American Museum of Natural History, The Museum of Modern Art (MoMA), Miraikan, and Tokyo National Museum.

This research field involves knowledge of museology, mass communication, and mobile technology, as well as comprehensive research in the field of design including user experience design, user interface design, design psychology, and visual communication design. Regarding the research innovations of this thesis, from the perspective of visual communication, through the sequence and orderly design method analysis of the layout, text, color, and other elements in the interface, as well as the combing of the interrelationships, the author can design the interface information level. There is a certain degree of in-depth research on the specific visual elements.

Through sensing technology, in real life, such as touch-sensitive elevator buttons (tactile sensors) and lights that are dimmed or brightened by touching the base, there are also many applications that most people have never realized. With the development of micromachines and easy-to-use microcontroller platforms, the application of sensors has surpassed the traditional temperature, pressure, or flow measurement fields. In this research, the sensing technology will be combined with the interface design of the digital museum, through vision and touch to optimize the navigation of the museum, making the research more innovative.

Most of the existing museum APP designs are based on the stylistic features of the museum for interactivity. Our paper, however, takes into account the object of use of the museum APP and starts with the user's target needs. Based on the goal-oriented theory, the information fusion and the combination of museum APP with modern intelligent sensing technology, based on the user perspective model, multimedia analysis of the digital museum user interface, clear user demand for the functionality of the museum APP, complete the construction of the functional visual hierarchy. To achieve a more convenient to use and meet the needs of users' museum APP. Combined intelligent sensing technology with computer communication technology, the networked intelligent sensor is widely used in process control field.

2. Methodology

The digital museum APP uses the Internet as a carrier. As an Internet product, it allows users to "participate" in product
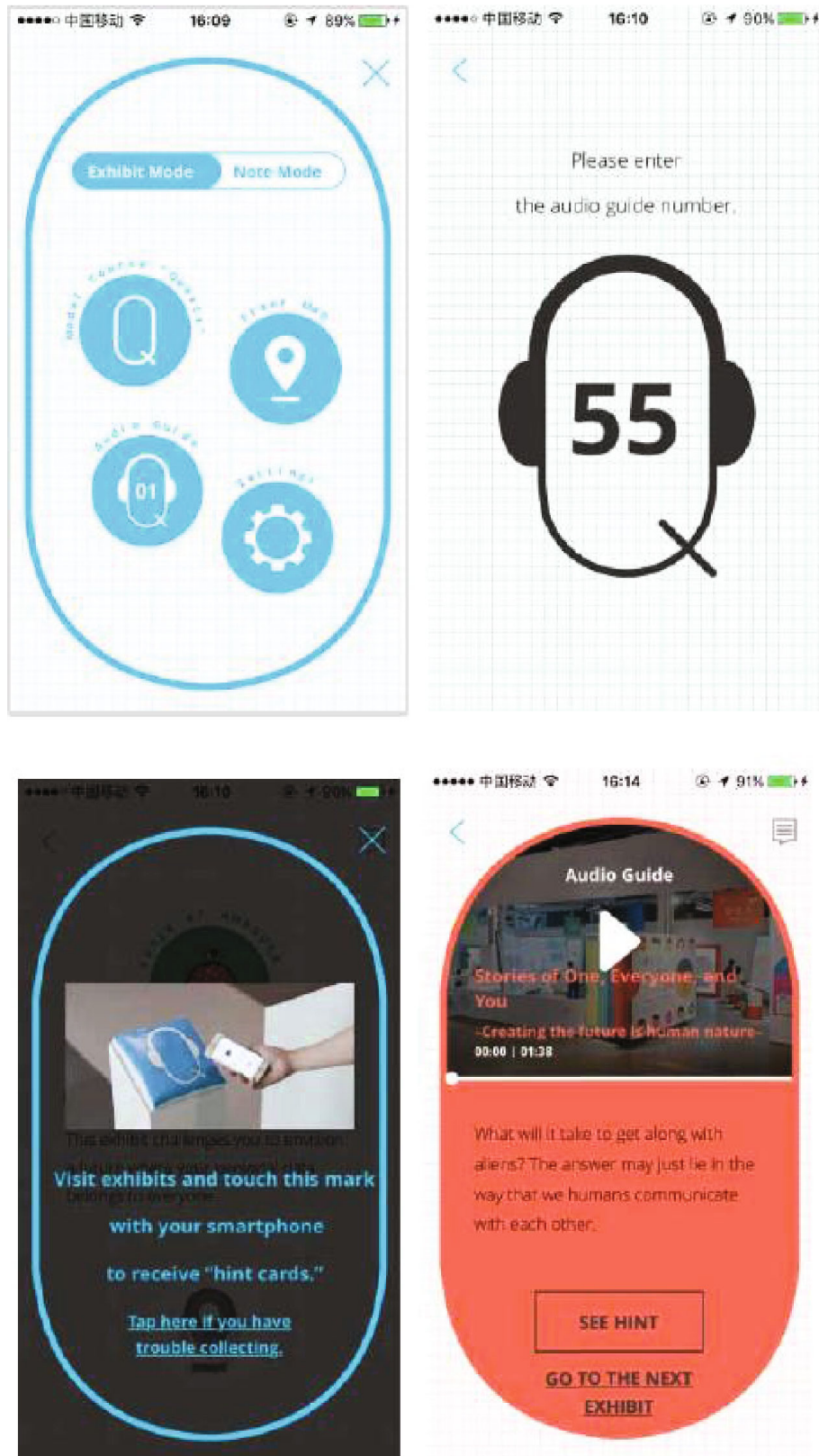
FIGURE 1: (a) The interface design for the navigation function. (b) Interface design for the "MOMA APP" guide positioning of the Museum of Modern Art in New York. (c) The "Explorer APP" developed by the Smithsonian Museum of Art.
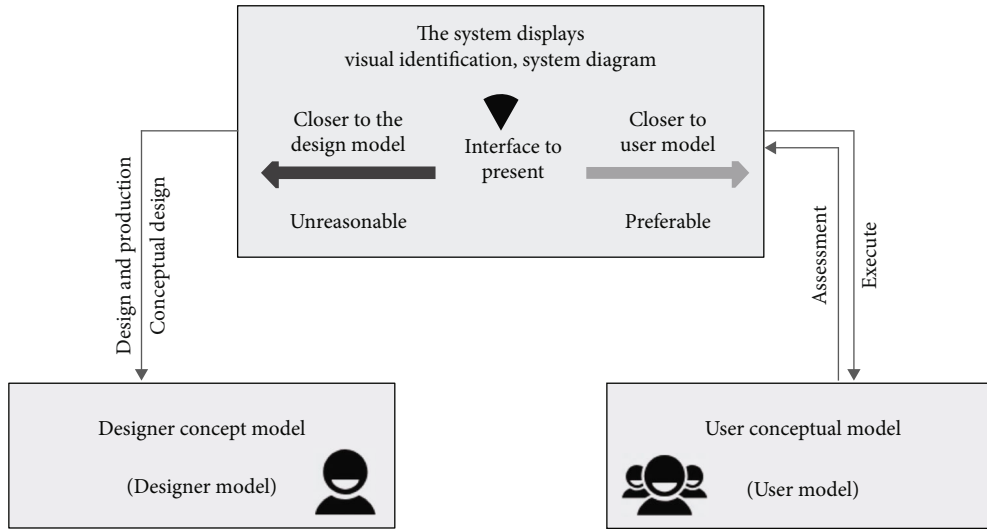
FIGURE 2: The relationship between the designer's conceptual model, the system image, and the user's conceptual model.

design. Users drive product designers and quickly iteratively update, transforming designer-centric thinking into user-centric product development [25]. The relationship between designers and users is no longer a simple supply-demand relationship, users have become "designers" in a certain sense, and real designers are more from an observation and coordination standpoint. The user interface is the platform for human-computer interaction in the digital museum. The visual design of the digital museum user interface is a comprehensive study that combines multiple disciplines. This paper proposes a quantitative research-based approach to intuitively distil users' target needs and expectations for museum APPs from their APP usage habits, and to study the visual design of digital museums based on the user role model construction method and low-fidelity prototype testing method to help guide the establishment of visual hierarchy design strategies.

2.1. Designer Conceptual Model. A concept model is a clear idea of how a product will be made and presented by the designer when designing the product. The system image is all the visual cues, system illustrations, and other content presented in the product, which the designer can present through the concept model. It also determines the interface visuals that the designer will present to the user after integrating the museum information [26]. This is the most intuitive way of presenting products such as functions and interfaces in a museum app. Its most important purpose is to serve the user, so it is important to design a conceptual model that will meet the user's experience needs.

As shown in Figure 2, there is a strong link between the designer's conceptual model, the system image, and the user's conceptual model. The system image contains the interface functions, information integration, and visual design of the digital museum, which are derived from the feedback of the designer's conceptual model. The system image is the bridge between the designer's conceptual model and the user's conceptual model, and the designer's ideas are

presented directly to the user through the system image. Only when the system image fully demonstrates the designer's conceptual model can the designed museum app meet the user's needs, and the designer's conceptual model and the user's conceptual model will then be perfectly matched.

2.1.1. Web Builds a Five-Layer Model. In order to better analyze the needs of users, a five-layer model needs to be established when building a website, namely, the strategy layer, the scope layer, the structure layer, the framework layer, and the presentation layer. The structure is shown in Figure 3. The user experience elements include five layers. The content of each layer forms a complete system from bottom to top. Each layer influences each other, and each part guides each other. The five-layer structure and establishment takes user needs as the most important consideration of the product and integrates user experience into the design.

The strategic layer contains the product objectives and user needs. In the interface design of a museum app, the product objectives refer to the brand image that the museum wants to convey. The scope layer contains the functional planning and content requirements; here, we need to clearly know the characteristics that we want to present to the user of the museum for which we are designing the app. The structural layer is a translation of the user's needs, identifying further the elements to be presented to the user. It is necessary to examine whether the information presented to the user meets the user's needs. The framework layer is the integration of the content we want to present and the layout through the page design. The presentation layer is the visualization of the results of the preliminary research and the final interface presented to the user. The presentation layer integrates functionality, text, images, etc. and makes it easier for the user to operate through the visual presentation. As shown in Figure 4(a), if we further focus on the accuracy of functional icon design, it will facilitate the operation and recognition of the user.
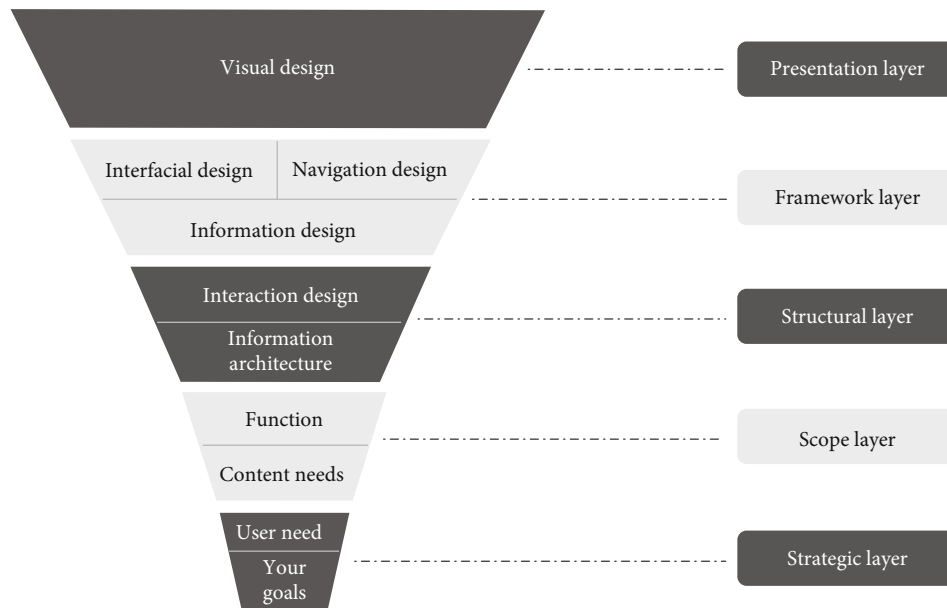
FIGURE 3: Five-layer model established by the web.

2.1.2. User Role Model Construction Method. The concept of the user role model was proposed in the second step of modeling in the target guide design. The user role model is a virtual representative of real users. The product design solution is established according to the target needs of real typical users. It can show most of the user needs and user goals. It can also extract and refine typical users and use the user role model. It can better make user goals and needs become the center of subsequent design and production. Clear and representative users can make the design more targeted and avoid the addition of too many functions. The user role model is generated by the user himself, and the consistency of the user's thinking and behavior is also ensured. The structure is shown in Figure 5.

In order to determine the different goals of different user types, the user's purpose of visiting the museum, the user's purpose of downloading and using the APP, the basic functions that the user needs most, the characteristics that the museum APP should have, and the current users are obtained from the previous user research. The data analysis of the reasons why certain functions are not needed or rarely used and the relevant factors that affect the user's use of APP are clearly identified.

2.2. Quantitative and Qualitative Research. As the audience of the product, the user is the core service object of the designer in the whole process of designing the product. Defining product audiences and understanding user groups are the most important part of all project development and APP design. After determining the user group and its positioning, the research on the user's target direction can be advanced more accurately. Understanding the user group is also the initial stage in the goal-oriented design, which lays the foundation for the subsequent analysis of user goals and user needs. The usual user research methods are divided into qualitative research and quantitative research [27, 28]. The

combination of qualitative and quantitative research can deeply explore user goals and user behaviors from multiple angles and provide reliable user analysis.

2.2.1. Two-Dimensional Matrix Diagram of User Research Methods. As the audience of the product, the user is the most core service target of the designer in the whole process of designing the product. Identifying the product audience and understanding the user group are the first considerations in all project development and APP design. After determining the user group and its positioning, the research on the user's target direction can be advanced more accurately. Understanding the user group is also the initial stage in the goal-oriented design, which lays the foundation for the subsequent analysis of user goals and user needs. The structure is shown in Figure 6.

We distinguish between qualitative and quantitative user research methods. By using these two research methods, we can gain a more comprehensive understanding of the target needs of the user. By combining the two research methods, we can better integrate user needs into the design of the digital museum pages.

2.2.2. Ease of Use of Information Architecture. After clarifying user needs, there is roughly a prototype for the functions that the APP should have, but the prototype also needs specific support, which requires the creation of a structural layer of information architecture for the APP. In the digital museum APP application, it involves the concept of information architecture and the transformation of the digital museum APP information architecture. These topics will occur in the creative and conceptual stage before the introduction of the framework layer and the presentation layer. The information architecture is designed to design the specific components of system classification and navigation, so that users can quickly and effectively read the information

首页　　服务　　资讯　　我的　　查看展品　　展区介绍

地图导览　商店　　参观指南　票务预定　藏医药咨询　细节

定位　　语音　　参观轨迹　我的预定　参观消息　设置

博物馆APP主界面　　　　文物探索界面　　　　地图语音导览界面

FIGURE 4: (a) Functional icon design. (b) APP interface display.

in the APP [29]. The perfection of the information architecture can play a key role in the subsequent provision of perfect intelligent sensing. Information architecture can organize and program huge data, design a platform for users to clearly communicate information, and bring a good user experience. In the information architecture of the museum app, the functions are divided by the goal-oriented theory, and the user's needs are sorted, some complicated functions are integrated and streamlined, and the visual presentation of the core content is highlighted, so that the overall information architecture of the museum app is clearer and more explicit.

In the construction structure, the basic unit of measurement is the node, and it can be distributed freely. Here are the main types of information architecture: tree diagram, which is the most widely used hierarchical structure; matrix structure; and natural structure, with freedom. This structure does not have a strong concept of "classification." The structure is shown in Figure 7. It is more suitable for exploring a series of apps, such as entertainment, education [30], or interactive games [31]. This structure will turn the user's experience into a challenge. This structure is suitable for functions such as enhanced exhibition experience of museum APP and virtual browsing of museums. The structure is shown in Figure 7.

For the museum app, the visit guide, booking the ticket, exhibition information, and exhibit introduction should be presented in the main interface. By improving the previous museum app interface as shown in Figure 7(a), we have reengineered the functions according to the user's objectives and obtained a more user-friendly digital museum interface as shown in Figure 7(b).

2.3. Low-Fidelity Prototyping of the Interface. In order to better understand the visual hierarchical design strategy corresponding to the user behavior level and the target level, the low-fidelity prototype design of the museum's APP is carried out based on the APP information framework of the museum. For this low-fidelity prototype test, we used a mobile phone user terminal as the test device and imported the preset low-fidelity interface to form a simulated interface for interaction with the Digital Museum App. The low-fidelity prototype is the information of the visual expression framework. Because the main interface requires more display functions and has a certain amount of text and pictures interspersed, the tabbed interface board layout is used in the presentation of important functions, and the card design is used in the collection and exhibition information display, which can clearly distinguish the two functions. In the case
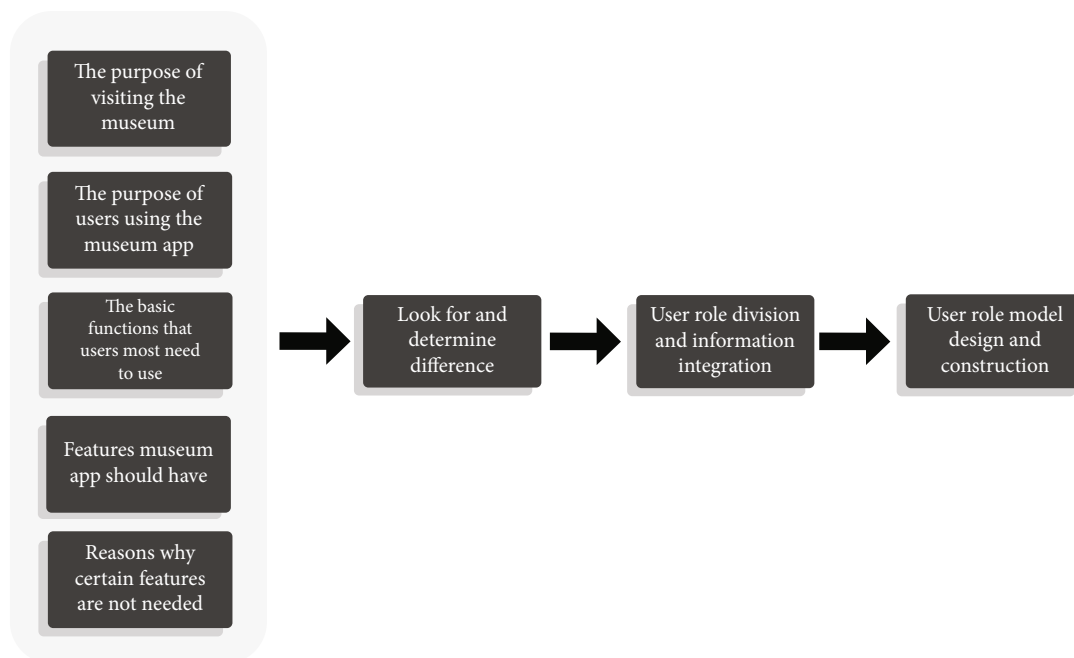
Figure 5: User role model construction method.

of the theme, the picture and the text are effectively combined, and the structure is shown in Figure 8.

The integration of information in the interactive interface is very important in order to present the user with an at-a-glance operating interface that perfectly matches the intelligent sensing of the phone. We clarify the sensory needs of the majority of users, based on user-orientation theory, in order to show them a good smart-sensing experience when operating the phone. As shown in Figure 8(a), in the first level of the interface, we present the main operations, such as Booking the ticket, Museum introduction, and Map guide. The Tab Category at the bottom of the page allows users to access the secondary interface. The layout of the area is designed to differentiate the content in order to meet the visual and aesthetic needs of the user.

As shown in Figure 8(b), we have made Make an appointment to buy tickets, Activity appointment, Cultural relics display function, Details of cultural relics, and Museum visitor guide interface according to the results of the information fusion provided by the user's needs. The visual layout of the secondary interface was based on the user requirements of Make an appointment to buy tickets, Activity appointment, Cultural relics display function, Details of cultural relics, and Museum visitor guide interface. It is understood that users have a greater need for the presentation of museum app exhibition information when buying tickets. By clicking on the exhibition information in the primary interface, we will automatically jump to the exhibition information details interface in the secondary interface. In the exhibition details, we can also learn about the specific information of each exhibit. By understanding the exhibition information, the user's ticketing needs are facilitated. Clicking on the ticketing option in the primary interface automatically jumps to the ticketing interface in the secondary interface, which substantially satisfies the user's target needs.

3. Results

3.1. Goal Guides to Assist Low-Fidelity Prototype Testing at the Design Framework Stage. Museum cultural relic exploration is one of the main functions designed for secondary users of "exploration and research." The museum's focus on cultural relics on the main page is only part of the collection, while the cultural relic information in the cultural relic exploration area is more comprehensive. This facilitates the operation of the user under intelligent sensing. The first-level interface of this function displays cultural relics in two parts: cultural relics and cultural relics. The navigation bar of the first-level interface is at the top of the interface, and the cultural species are designed in waterfall flow format to ensure cultural relics. The visual distinction on the variety display allows more cultural relic information to be presented. When designing the page hierarchy, we try to ensure that the primary functions of the Digital Museum interface can be presented in the main page, so that users can easily access the relevant functions at the first time. In order to ensure the convenience of user operation, we fuse some functions of the secondary and tertiary interfaces into a primary interface and differentiate the functions in the secondary interface. Avoiding a large number of levels of interface switching will reduce the user's APP experience and keep each function in 2-3 levels of interface as far as possible, as shown in Figure 9. We have used a timeline visual presentation to make it easier for the user to use the chronology to find the artefacts on the one hand and to enhance the user's knowledge of the chronology of the artefacts on the other. At
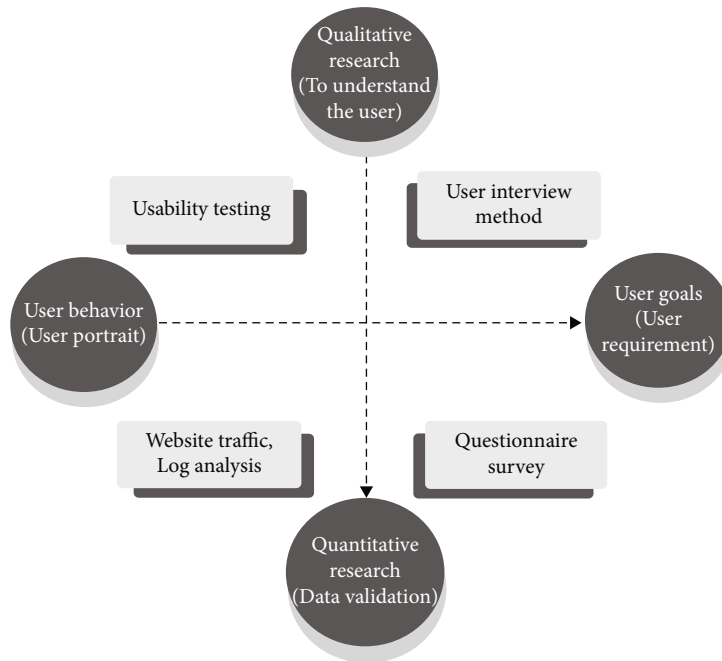
FIGURE 6: Two-dimensional matrix diagram of user research methods.

the same time, it is visually clearer and simpler. In our research, the simplicity and clarity of the interface are preferred by users.

The interactive display of some of the key functions in the low-fidelity prototype of the personal center interface. The most important ticket purchase, reservation, and collection details display functions in the personal center interface have their own branches. The ticket purchase function includes museum ticket reservation information and special exhibition purchases. Ticket information, the reservation interface contains event and lecture information, and the collection interface contains details of exhibitions and collections. The structure is shown in Figure 10.

Low-fidelity prototype testing is user testing using interactive wireframe diagrams. It can directly help designers find problems with users in a goal-oriented design, thereby helping designers to modify functions or interaction methods in the early stage. In a goal-oriented design theory, the interaction framework establishes the overall product behavior architecture, which includes the interactive form elements, functional combination, and functional levels. In the construction of the interactive framework, the elements need to be integrated and arranged. This is also the transition from the information architecture to the low-fidelity prototype design. It is proposed in the interface expansion architecture in the visual hierarchy above, through the interface, information, and navigation. After the design of the elements, a wireframe needs to be used to present the visual part of the content of this part. This is closely related to the integration of behavior, form, and content in the refinement phase of the target-oriented design. The wireframe diagram is the display of the composition of all the elements on the page, and it is also the most intuitive visual representation. The wireframe for basic user interaction is also the

basic content of the low-fidelity prototype test, as shown in Figure 11. When the number of users tested is around 5, the number of usability problems found tends to increase rapidly, indicating that the number of users tested at this time is beneficial to usability problem finding. When there are more than 9 users, the increase in usability problems tends to level off. Therefore, the number of testers can be limited to 5.

The design structure and process proposed in the definition framework part of the goal-oriented theoretical design is an information architecture based on the existing functions after previous user research, role modeling, and requirement definition. Jackob Nielsen, in a 2012 survey of usability tests conducted by the Nielsen Norman Group, found that the greater the number of users tested, the better the detection of problems. It can guide the design of low-fidelity prototypes and start testing. As shown in Figure 12, the horizontal axis represents the number of test users and the vertical axis represents the Number usability findings. When the number of testers reaches five users, almost all interface usability problems can be found through the test results. The fewer users we have, the better the usability of the questions we collect. When there are more than 5 users, it gradually becomes impossible to find more problems.

3.2. Demand Positioning. In our questionnaire, we asked survey respondents who had visited the museum in-depth questions about the "shortcomings of visiting the museum." Multiple answers were provided for respondents to choose from, and the results of the survey are shown in Figure 13, where the number of choices made by respondents for the different options was about equal.

This shows that there are many drawbacks to visiting physical museums, especially in the three options Abstract

Touch screen activation

Open page

Close page

Opening the interface

Visit guide

Booking the ticket

Image | Image

Image | Image

Map

Antique appraisals

Image

The name of the cultural relics

Classification of cultural relics

Cultural relics are introduced

Exhibition information

Video

Audio explanation

Panoramic photos

Detail drawing

Detail description

Comparision of cultural relics in different times

Image | Image

Image | Image

Brief introduction

Explanatory note

Interactive design
Advance
Retreat
Click

(a)

FIGURE 7: Continued.

(b)

FIGURE 7: Ease of use description of information architecture: (a) Early layout frame structure diagram. (b) Final layout frame structure diagram.

content, Space is too far away, and Less harvest, which were chosen by the majority of respondents. The percentage of respondents who thought that Abstract content was too far away was 32%, Space is too far away was 30% and Less harvest was 20%. There was less demand for museums to offer Less interaction, at 4%. The results show that the majority of people have the problem of not being able to visit museums independently, so we can add the function of online audio guide or tabs for exhibit annotations to the design of the museum app. Because people do not understand the content of museum exhibitions well, most people do not think they will get what they expect from a visit to a museum. We can focus on deepening this point when designing the museum app. People can miss the opportunity to visit a museum because it is too far away, so our digital museum can solve this problem. Even if visitors do not get to the museum, they can still have a remote view of the exhibition.

The user visit habits given to users of the same category in the interview volume have the same characteristics. Therefore, the interview results of user museum visit habits are summarized and summarized according to the three user categories, as shown in Table 1. In the analysis of the content related to the Museum App, we have consolidated and summarized the results based on the results obtained from different user groups. For those who have used the Museum App before, they recognize the need for the Museum App to exist.

For those who have not been exposed to the Museum App, they have certain concerns. From the results of the interviews, it is clear that most users are supportive of the existence of the Museum App and have a great deal of longing for the Museum App to be put into use. Most users want the Museum App's operating pages to be simple and to allow them to access the information they need quickly. They want to be able to see information about the museum's exhibitions on the Museum App and already have a guided tour. Then users can easily access museum-related knowledge on the Museum App through intelligent sensing.

Before conducting user interviews and user questionnaire surveys, this paper publishes multiple-choice questions about exploring museum user groups based on the reasons why tourists visit the museum. In order to clarify the categories of user groups and facilitate the user role setting in the later period, the purpose of the user's visit to the museum is set to understand the local culture, tourism, leisure, learning, hobbies, accompanying visits, and research six categories. The purpose of setting multiple-choice topics here is to be more objective and to cover as many aspects of the user group as possible and to summarize the data through the six types of user visits, and to integrate the user groups, as shown in Figure 14.

From the results of the survey, we can see that the largest number of people, 62.6%, wanted to learn about local culture by visiting museums. The number of people whose purpose
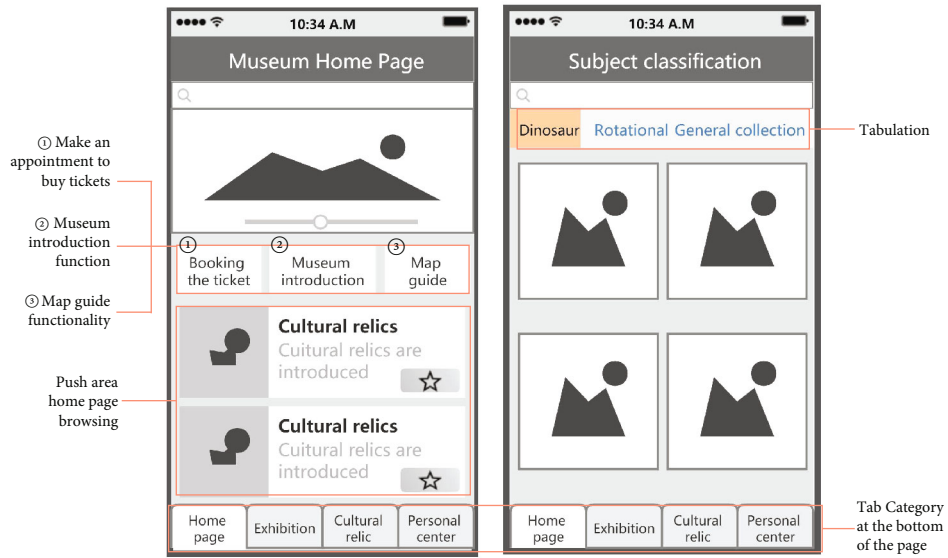
FIGURE 8: Human-machine interface interaction design depiction: (a) first-level interface interaction and (b) second-level interface interaction.

Level 1 interface: Heritage   Chronological classification   Level 2 interface: Display of
exploration category classification   of artefacts   artefacts by search category
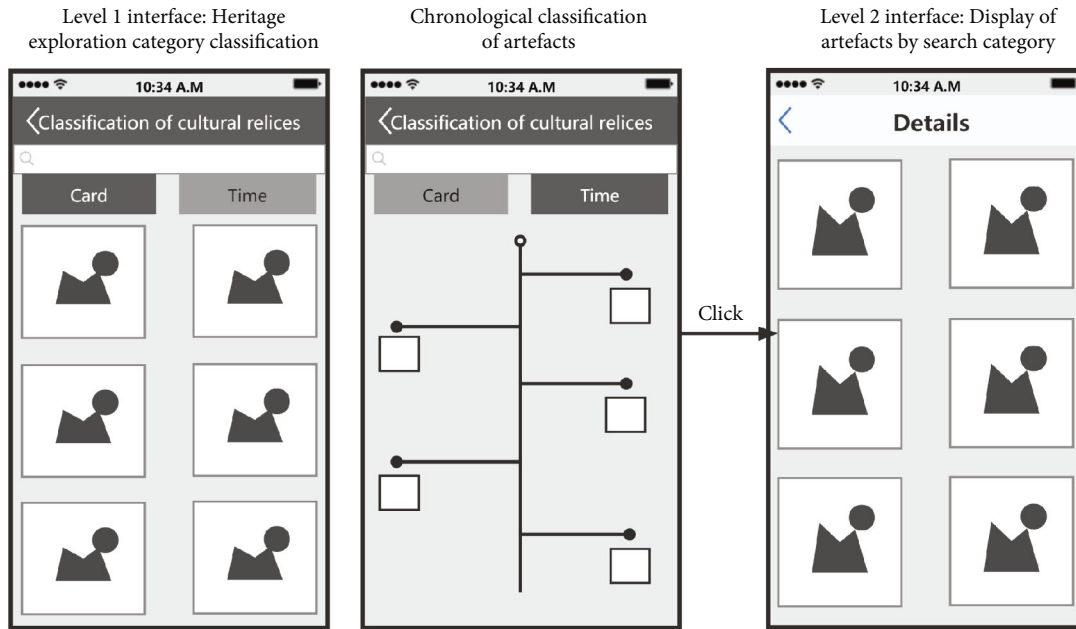
FIGURE 9: Interactive display of key functions in the low-fidelity prototype of the cultural relics' exploration interface.
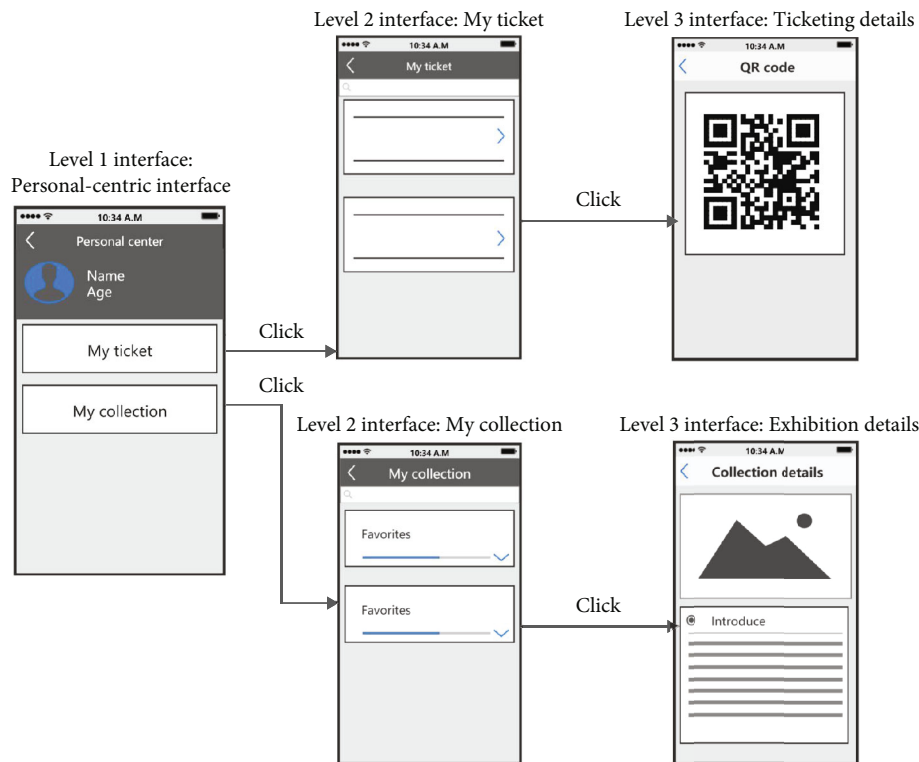
FIGURE 10: Cross-based display of some key functions in the low-fidelity prototype of the personal center interface.

was to travel and study was the next highest, and they accounted for an equal share, 43.1% and 40.6%, respectively. A smaller number, 23.4%, attended museums for hobbies and interests. The number of people who visit museums for research purposes is even lower, at 11.9%. We should

therefore weaken the academic nature of the digital museum when designing its pages. Designers need to keep the pages simple and integrate information about the museum's culture so that cumbersome and difficult knowledge can be easily understood and made accessible to a wider range of users.
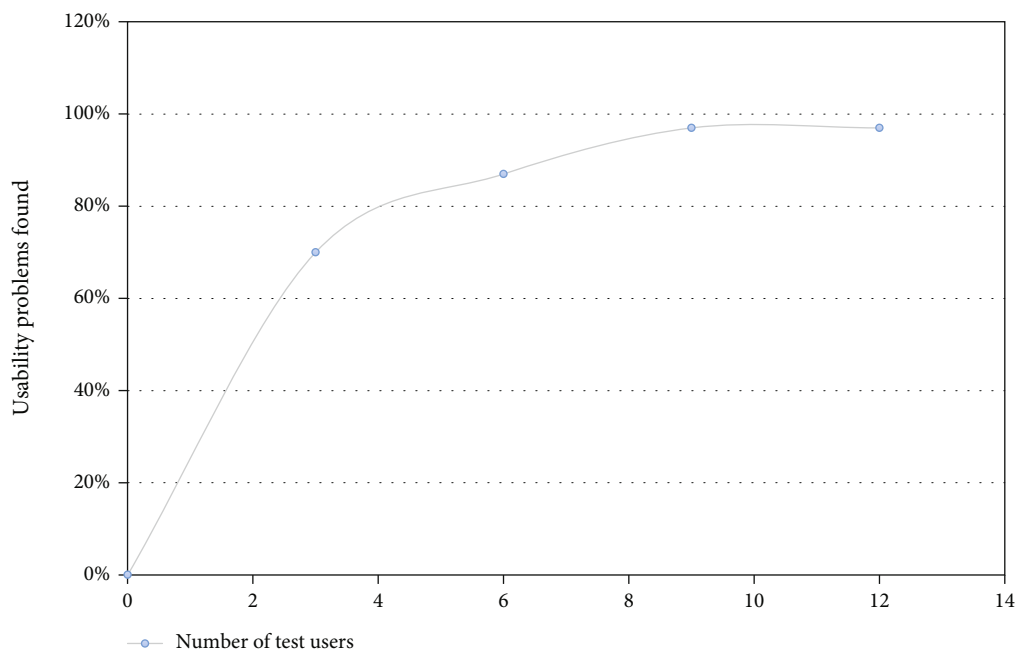
FIGURE 11: The ratio of the number of users in the low-fidelity test to the problems found.
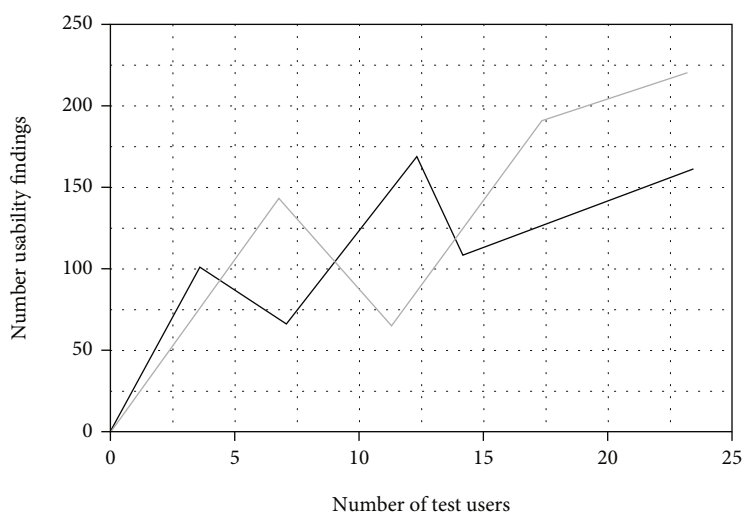
FIGURE 12: The relationship between the number of tested users and the number of problems found in the usability test.

4. Discussion

Experimental research has shown that modern digital museum user interface design is not only about the promotion and publicity of the museum but also about meeting the user's understanding and application of the museum app's functions. The user interface of the museum app is presented in a form of visual communication that is more in line with the user's preferred programming. While having the characteristic humanistic connotations of that museum, the designer should combine the modern technology of information fusion and intelligent sensing to meet the user's goal-oriented needs. The digitalization of museums follows the trend of the times, and the museum's APP, as a mobile

display for digital museums, should also be constantly updated and developed. As shown in Figure 4(b), the APP interface display case of the Nanjing Museum which provides a good reference value. This paper looks at the current situation of museums and finds that the current design of museums' APP lacks consideration for users. Problems on the visual level lead to products that do not meet the needs and experiences of users. By summing up and summarizing the problems of museum APPs, a research method that meets the needs of users is found.

The designer concept model considers the needs of the target audience and some of the target users of the museum app. The design conceptual model and the user conceptual model play an interplay through the connection of the
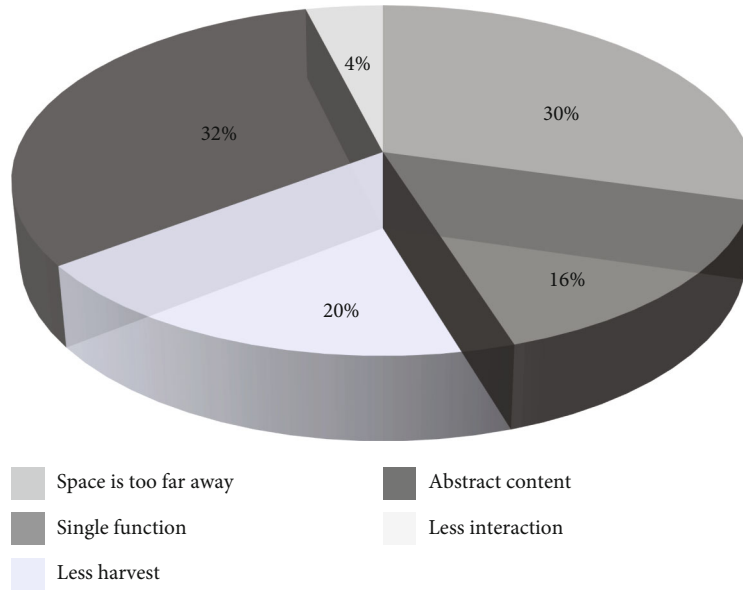
FIGURE 13: The shortcomings of visiting museums.

TABLE 1: Three types of user museum visiting habits.

Users' category	Museum visit preferences	Museum visit needs	Museum visit time	Existing problems in the museum
Tourism visit class	Visiting large provincial and municipal museums, more interested in the main museum pavilion and provincial and municipal special exhibitions.	Surrounding supporting facilities, arrival mode, venue structure, exhibition location and other information.	1-3 h	There are deficiencies in queuing, reservation and ticket purchase system, explanation service and other contents.
Learning and research	Visit niche museums with specific themes.	Relevant cultural relic text and picture's introduction, for a detailed, rich and comprehensive introduction of the museum content.	'There is no planning	It is more difficult to find the collection needed to find in the museum.
Interest and hobbies	Like the related activities and special exhibitions held by museums, museums with different themes and features.	We need to introduce more detailed and interesting exhibits, and search relevant museum information and research through the Internet.	2-3 h	Insufficient guide tour and explanation.

system image. In this paper, we first analyze the design content of the visual level through the different levels of user goals in the goal-oriented design theory and find the problems that exist in the visual level of the museum APP, where the functions do not meet the user's goal needs and the information architecture does not match the user's order of use. The Museum App combines information fusion with advanced intelligent sensing technology and we have established a web five layer model. The functionality and importance of the visual hierarchy in interactive products is analyzed, and the problems that need to be solved in the visual hierarchy are investigated in depth, providing theoretical guidance for the study of the visual hierarchy of the museum app.

The framework layer serves as the beginning of the visual design, presenting the page layout through the information architecture; presenting the product architecture in a visual hierarchy; realizing the interface, navigation, and information design; and visualizing the structure gradually. The presentation layer is the top of these five layers and is the first thing that users notice when using the interface. The presentation layer brings together function, content, and aesthetics to produce a final design that meets all the objectives of the other four layers. The visual presentation in the presentation layer needs to link all the structures more closely, adding visual elements to make product functionality and information clearer to the user and easier to recognize.

Based on the goal orientation theory, the functional, and information architecture, visual element design and stylistic features of the museum app visual hierarchy are analyzed according to the user instinctive, behavioral, and reflective levels of goals. Through the user research guided by goal-oriented design, user needs for visiting museums and for the museum app are explored in depth, user role models
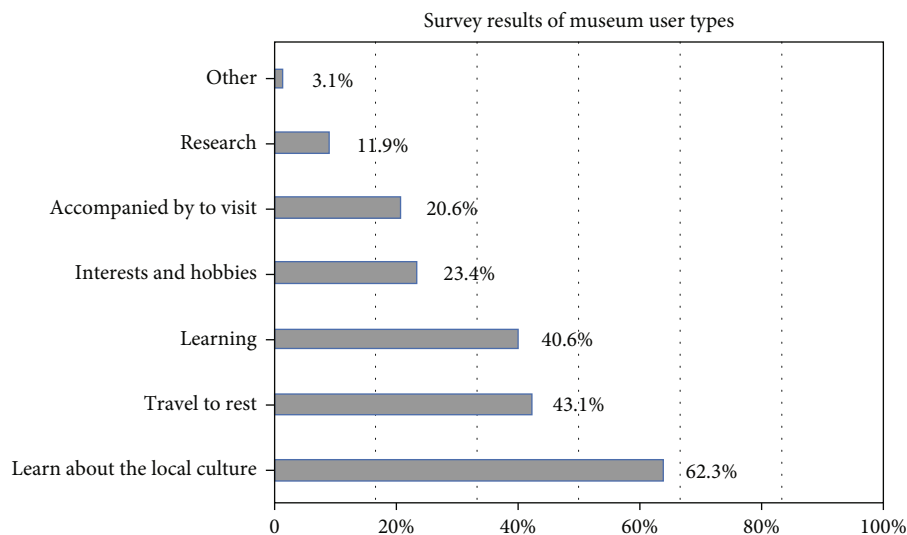
FIGURE 14: Survey results of museum user types.

are created, and user role task models are established. In the questionnaire, two main aspects were included. The first aspect was a survey of users' habits of visiting museums, the purpose of their visits, and the channels they use to learn about museums; the second aspect was a survey of users' knowledge and needs for museum APPs, which included the current popularity of museum APP downloads, the reasons for using museum APPs. The second aspect of the survey was the knowledge and needs of users of the museum app, the reasons for using the museum app, the basic functions that users need to use most, the features that the app needs to have, and the addition of more information functions. Through the questionnaire, we accurately grasped the users' needs for museum APPs in terms of functionality, information integration, smart sensor operation, and purpose of visit.

The information framework of the digital museum user interface is set by functional level and serves as the basis for the visual level. A low-fidelity prototype is created and used to understand the real goal-oriented needs of the user, and the visual layer is further guided by means of low-fidelity prototype testing. The purpose of this test is to set up user tasks for important functions in the APP. Conduct a multimedia analysis of the museum's digital user interface based on a user perspective model to clarify the user's functional requirements for the museum app and to complete the construction of a functional visual hierarchy. In our future implementations, we may use autoencoder-extreme learning techniques [32, 33] for interaction design. To achieve a greater degree of ease of use and to meet the user's needs for the Museum App.

5. Conclusion

We use quantitative user research methods, user role model construction methods, and low-fidelity prototype testing methods to study the visual design of digital museums. Through our research, we have learned that the largest num-

ber of people, 62.6%, wanted to learn about local culture by visiting museums. The number of people whose purpose was to travel and the study was the next highest, and they accounted for an equal share, 43.1% and 40.6%, respectively. A smaller number, 23.4%, attended museums for hobbies and interests. The number of people who visit museums for research purposes is even lower, at 11.9%. Starting from the perspective of the user's target needs, we understand the user's need for information. The various seemingly disorganized museum information fusion will be perfectly integrated in the museum app. After integration, it can be clearly and unambiguously presented in the museum app we designed.

The goal-oriented visual hierarchy research first obtains the audience of the product and its user behavior characteristics through quantitative research methods and then summarizes and refines the group users through the role model construction method. Using the user role model can better make the user goals and needs become the follow-up design. Finally, through the low-fidelity prototype test method, combined with the museum APP information architecture established in the early stage, low-fidelity prototype design was carried out according to the five major functional areas required by the museum APP. Use the interface low-fidelity prototype test method to list and analyze the operational errors and doubts that occurred in the user's use, and put forward the modification opinions of the museum's APP visual level based on the user's instinct and behavioral goals in response to the operational problems. It provides a basis for the museum's APP visual hierarchical design strategy.

The multimedia analysis of the user interface of digital museums is a contemporary and superior tool that combines information fusion and intelligent sensing. The digitization of museums follows the trend of the times, gathering information about the user experience in multiple dimensions and integrating feedback from users to design user interfaces that meet their needs. Starting from the perspective of the user's target needs, we understand the user's need for

information. The various seemingly disorganized museum information will be perfectly integrated in the museum APP. After integration, it can be clearly and unambiguously presented in the museum APP we have designed. The multimedia application of the digital museum interface is perfectly integrated with today's smartphones. We need to study in detail the various functions of the user interface, such as ticketing, guided tours and online visits, and combine them rationally with modern intelligent sensing to achieve a better user experience.

In summary, although the method proposed in this paper can be better applied to the development of interface applications of digital museums, the research on this topic is still at the visual level. In the future, it is necessary to continue to further improve and optimize other related auxiliary functions to explore more. A comprehensive system platform.

Data Availability

Data available on request from the authors due to privacy/ ethical restrictions.

Consent

All human subjects in this study have given their written consent for the participation of our research.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] W. Y. Zhao and Z. Y. Chen, "Internet+ era museum marketing strategy analysis," *Museum Researcher*, vol. 127, no. 2, pp. 3–7, 2019.

[2] Y. D. Han, *Research on the interaction design of virtual textile museum based on user experience*, Wuhan Textile University, 2017.

[3] R. Hoekman and I. Beauty, *How to Design Web Interfaces to Move Users*, Beijing: People's Post and Telecommunications Publishing House, 2009.

[4] Treadwell, *Interface Design Patterns*, Beijing: Electronic Industry Press, 2013.

[5] K. W. Liang and Y. L. Li, "Research on interaction design process based on user scenarios," *Packaging Engineering*, vol. 382, no. 16, pp. 209–213, 2018.

[6] C. Andujar, A. Chica, and P. Brunet, "User-interface design for the Ripoll Monastery exhibition at the National Art Museum of Catalonia," *Computers & Graphics*, vol. 36, no. 1, pp. 28–37, 2012.

[7] W. Hauser, A. Noschka-Roos, E. Reussner, and C. Zahn, "Design-based research on digital media in a museum environment," *Visitor Studies*, vol. 12, no. 2, pp. 182–198, 2009.

[8] S. Y. Li, "The current situation of museum APP application and corresponding countermeasures," *Journal of Huizhou College*, vol. 37, no. 4, pp. 89–93, 2017.

[9] R. Li, *Practice and Thinking of the Integrity Principle in the Protection of the Tombs of Tang Dynasty*, Northwestern University, Xi'an, 2008.

[10] K. B. Jones, *The transformation of the digital museum*, Museum Informatics Routledge, 2012.

[11] A. Moneerah, D. John, and N. Julianne, "Museum mobile guide preferences of different visitor personas," *Journal on Computing and Cultural Heritage (JOCCH)*, vol. 14, pp. 1–13, 2020.

[12] S. Prudhomme and J. T. Oden, "On goal-oriented error estimation for elliptic problems: application to the control of pointwise errors," *Computer Methods in Applied Mechanics and Engineering*, vol. 176, no. 1-4, pp. 313–331, 1999.

[13] A. Leshchenko, "Empowering digital museum audiences to foster museum communication," *ICOFOM Study Series*, vol. 41, no. 2012, pp. 237–244, 2012.

[14] D. Korzun, A. Voronin, and I. Shegelman, *Semantic data mining based on ranking in internet-enabled information systems*, IOS Press Ebooks, 2021.

[15] O. Goldreich, B. Juba, and M. Sudan, "A theory of goal-oriented communication," *Journal of the ACM (JACM)*, vol. 59, no. 2, pp. 1–65, 2012.

[16] T. K. Gustavsson and A. Hallin, "Goal seeking and goal oriented projects–trajectories of the temporary organisation," *International Journal of Managing Projects in Business*, vol. 8, no. 2, pp. 368–378, 2015.

[17] J. Y. Chu, "An analysis of the interface design of museum-type APP," *Art Technology*, vol. 30, no. 2, pp. 24–33, 2017.

[18] P. F. Marty, "Museum websites and museum visitors: digital museum resources and their use," *Museum Management and Curatorship*, vol. 23, no. 1, pp. 81–99, 2008.

[19] Z. Wu, "Research on the application of internet of things technology to digital museum construction," *Acta Geoscientica Sinica*, vol. 2, pp. 293–298, 2017.

[20] G. Varvin, H. Fauskerud, I. Klingvall, L. Stafne-Pfisterer, I. S. Hansen, and M. R. Johansen, "The journey as concept for digital museum design," *Digital Creativity*, vol. 25, no. 3, pp. 275–282, 2014.

[21] Y. D. Wang, X. Q. Hu, and G. M. Huang, "De- velopment and utilization research of digital information resources," Wuhan: Wuhan University Press, 2005.

[22] D. Korzun, S. Yalovitsyna, and V. Volokhova, "Smart services as cultural and historical heritage information assistance for museum visitors and personnel," *Baltic Journal of Modern Computing*, vol. 6, pp. 418–433, 2018.

[23] X. C. Liu, "A study of user models in goal-oriented interaction design," *Art Technology*, vol. 29, no. 5, pp. 16-17, 2016.

[24] D. Sun, "Research on goal-oriented visual interface design for mobile smart terminals," *Communication Power Research*, vol. 13, pp. 12–23, 2018.

[25] P. S. Soares, "New Business Models in the Digital Economy Applied to the Smart Tourism Sector-The Case of U," *Porto' s Digital Museum App*, vol. 342, no. 28, pp. 174–189, 2019.

[26] Z. Tang, G. Zhao, and T. Ouyang, "Two-phase deep learning model for short-term wind direction forecasting," *Renewable Energy*, vol. 173, pp. 1005–1016, 2021.

[27] S. Attardo, "Irony markers and functions: towards a goal-oriented theory of irony and its processing," *Rask*, vol. 12, no. 1, pp. 3–20, 2000.

[28] D. Mentor, "From app attack to goal-oriented tablet use," in *Tablets in K-12 education: integrated experiences and implications*, pp. 1–21, IGI Global, 2015.

[29] Y. C. Li, A. W. C. Liew, and W. P. Su, "The digital museum: challenges and solution," in *2012 8th International Conference on Information Science and Digital Content Technology (ICIDT2012)*, vol. 3, pp. 646–649, Jeju, June 2012.

[30] R. J. Fang, C. W. Fan, J. F. Huang, and Y. H. Wang, "The knowledge-based mobile learning system applied in printing network instructive course for science museum," in *Proceedings of the 6th Conference on WSEAS International Conference on Applied Computer Science-Volume 6*, pp. 187–193, Hangzhou, China, April 2007.

[31] G. Andritsou, A. Katifori, V. Kourtis, and Y. Ioannidis, "MoMaP-an interactive gamified app for the Museum of Mineralogy," in *2018 10th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*, pp. 1–4, Würzburg, Germany, September 2018.

[32] Z. H. Tang, S. K. Wang, X. Y. Chai, S. X. Cao, T. Ouyang, and Y. Li, "Auto-encoder-extreme learning machine model for boiler NOx emission concentration prediction," *Energy*, vol. 256, p. 124552, 2022.

[33] Y. Lu, X. Fu, F. Chen, and K. K. L. Wong, "Prediction of fetal weight at varying gestational age in the absence of ultrasound examination using ensemble learning," *Artificial Intelligence in Medicine*, vol. 102, 2020.

Hindawi

Research Article

TBR-NER: Research on COVID-19 Text Information Extraction Based on Joint Learning of Topic Recognition and Named Entity Recognition

Xin Feng (iD),[1,2,3] Yingrui Li,[4] Zhang Hang,[4] Zhang Fan,[4] Qiong Yu (iD),[1] and Ruihao Xin (iD)[4,5]

[1]*Department of Epidemiology and Biostatistics, School of Public Health, Jilin University, Changchun, Jilin 130012, China*
[2]*School of Science, Jilin Institute of Chemical Technology, Jilin 130000, China*
[3]*State Key Laboratory of Inorganic Synthesis and Preparative Chemistry, College of Chemistry, Jilin University, Changchun, Jilin 130012, China*
[4]*College of Information and Control Engineering, Jilin Institute of Chemical Technology, Jilin 130000, China*
[5]*College of Computer Science and Technology, and Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun, Jilin 130012, China*

Correspondence should be addressed to Qiong Yu; yuqiong@jlu.edu.cn and Ruihao Xin; xrh@163.com

There is a centralization of the core content in the text information of the new crown epidemic notification. This paper proposes a joint learning text information extraction method: TBR-NER (topic-based recognition named entity recognition) based on topic recognition and named entity recognition to predict the labeled risk areas and epidemic trajectory information in text information. Transfer learning and data augmentation are used to solve the problem of data scarcity caused by the initial local outbreak of the epidemic, and mutual understanding is achieved by topic self-labeling without introducing additional labeled data. Taking the epidemic cases in Hebei and Jilin provinces as examples, the reliability and effectiveness of the method are verified by five types of topic recognition and 15 types of entity information extraction. The experimental results show that, compared with the four existing NER methods, this method can achieve optimality faster through the mutual learning of each task at the early stage of training. The optimal accuracy in the independent test set can be improved by more than 20%, and the minimum loss value is significantly reduced. This also proves that the joint learning algorithm (TBR-NER) mentioned in this paper performs better in such tasks. The TBR-NER model has specific sociality and applicability and can help in epidemic prediction, prevention, and control.

1. Introduction

In December 2019, a novel coronavirus (COVID-19) was transmitted between species and spread fast worldwide in a short period. The fast spread of the disease and severe economic and social devastation are far beyond people's expectations. Utilizing data mining and natural language processing technology to anticipate epidemic development trends and carry out intelligent security early warning has piqued the interest of academics. It has progressively become one of the hotspots in natural language processing [1–3]. According to the current standardization of epidemic pre-

vention and control, the epidemic spread is characterized by random small-scale bursts, which leads to a shortage of case information data in the early stage of transmission, making the study of epidemic information more complex and problematic. With their high efficiency, standardization, and real-time features, artificial intelligence, big data, and other technologies have achieved tremendous gains in several epidemic prevention and control domains since the outbreak of novel coronavirus-infected pneumonia [4–6]. The SEIR model is one of the most extensively used epidemic prediction models in the epidemic dynamics model. It may consider transmission speed and mode and

numerous infectious disease prevention and control strategies. The trajectory location early warning model monitors and provides an early warning by mining semantic trajectory data and combining location, time, and various application circumstances. Furthermore, some researchers use multivariate cosmos algorithms with artificial intelligence algorithms, natural language processing technology, and other comprehensive considerations of different prevention and control measures and various factors to build an early warning model, which has obvious advantages. COVID-19 pandemic prediction has piqued the interest of academics from many sectors all around the world. For example, Khayyat et al. [7] provided a predictive analytic model to anticipate the spread of the epidemic in Saudi Arabia and utilized the time series correlation FB model to perform a t-test on the data, giving a foundation for the epidemic's future development. Alsunaidi et al. [8] established a dynamic prediction and transmission technique based on the SEIRD epidemic model and calculated time-varying model parameters using maximum likelihood to show the emotional influence of the infection rate, death rate, and recovery rate on COVID-19 transmission. Balaha et al. [9] extracted characteristics from CT scans and learned them using deep learning and pretrained models. A hybrid technique that superimposes several CNN models is adopted to improve prediction accuracy. Wieczorek et al. [10] developed a neural network model for epidemic spread utilizing government data and the Nadam training model to achieve good prediction accuracy; Kozio et al. [11] offered a fractional-order SIR epidemic model to forecast epidemic spread. The model parameters were estimated using the genetic method, and the simulation was validated using Spanish data. Because the critical content of the text information of the epidemic has been centralized, each phrase has its core subject, such as the basic information about the infected patients, the track information during the sickness, and the time of diagnosis. When conducting information extraction, the information extraction technology indicated in the current results frequently cannot take advantage of the peculiarities of the epidemic notice text and cannot use the topic information. Furthermore, some algorithms suffer high labeling costs and little labeled data. Due to the abovementioned issues, this work presents a cooperative learning strategy for topic recognition and named entity identification TBR-NER. This technique annotates the epidemic notification information first and then exploits the features of distinct entities belonging to different subjects to accomplish topic self-annotation without adding extra annotation data, decreasing human annotation effort.

To increase the model's generalization ability in the absence of samples, entity mention replacement and external knowledge base replacement are employed to improve data. At the same time, transfer learning enhances the model's text parsing capacity. The following are the primary contributions of this study. (1) A TBR-NER model is presented for text extraction of epidemic notification information. The text extraction accuracy of epidemic notification information is increased by integrating the notification information text training mode with topic recognition in entity recognition. (2) Propose a self-annotation approach

for cooperative learning. Topic self-annotation is done without providing new annotation data based on entity annotation. The features that distinct entities belong to various topics are utilized for topic recognition jobs. (3) Add the Chinese information corpus (Chinese Wikipedia corpus) to the training session. On the one hand, some information about entity place names is replaced by entities during the training process to achieve the goal of data enhancement. On the other hand, the pretrained model is finally transferred to the joint learning model by learning the notice text to express the learning sentence structure, word collocation information, and so on, which effectively avoids the diversity and ambiguity of the expression and improves the model's generalization ability. (4) Comparative tests were conducted on the collected datasets from the provinces of Jilin and Hebei. The findings demonstrate that the proposed topic recognition and named entity recognition (TBR-NER) combined learning system can extract critical information in the epidemic notification job. The topic recognition and entity recognition tasks will help in the cooperative learning process, resulting in higher classification results.

2. Related Work

2.1. Conditional Random Field. The conditional random field (CRF) is a probabilistic model [12] used to label and split data with a sequence structure. It combines the properties of the maximum entropy model with the hidden Markov model, and it can describe long-distance dependence. It can globally normalize the characteristics and then achieve the objective of global optimization, which better overcomes the label bias problem. When first collecting text information, the model typically pulls crucial information in the form of the character level. Even though this approach is basic and straightforward to master, it has certain flaws. Discerning things based on categorization labels is complicated when multiple continuums coexist. For example, the three geographical names of Jilin province, Changchun city, and Jilin University correspond to three-place entities in the label categorization of Jilin University in Changchun city, Jilin province. However, there is no visible border between continuous entities when using the character-level form to extract modeling and the abovementioned words will be predicted as an entity label. To address this issue, this article adds the entity label BIO (see Figure 1). Each original independent entity information annotation is labeled with one of three types of labels: the beginning label (B-), the intermediate label (I-), and the irrelevant label (O-). See Figure 1. The classification prediction results of "Qianxiguan village, Xiguan town, Gaocheng district, 'become' B-location, I-location, I-location, B-location, I-location, I-location, B-location, I-location, I-location, I-location, I-location, and I-location." The red label represents the patient's ID, the dark-green label represents the patient's hometown, the pink label represents the patient's home time, the purple label represents the patient's trajectory action event, the blue label represents the patient's opening trajectory movement time, and the light-green label represents the patient's diagnosis time; the dark-blue label represents the time when the

Case. # Confirmed case 2: Female, 36 years old, from Qianxiguan Village, Xiguan Town, Gaocheng District, daughter of the first confirmed case today. On December 28, 2020, stay at home without going out; on December 29, go to the village market in the morning, and go to the LeJia supermarket for shopping in the afternoon; December 30th, no going out at home; December 31st, at 9 o'clock in the morning, drive to the credit building in Gaocheng District for shopping, eat at the food stall on the 1st floor of the credit building at around 14:00, and drive to Beiguo Mall Gaocheng Shop shopping at 15:00, drive home after half an hour; on January 1, 2021, drive to Xinle City Credit Building Mall for shopping, exit from Credit Building at about 10 o'clock, and go to the open-air bazaar near Xinle City Cinema to buy clothes return home; from January 2 to 3 at home without going out; from January 4 to the village's daily, Zhongtong, and post courier points to pick up the express; January 5 to the village's Lejia Shopping Supermarket for shopping; from January 6 to 11th No going out at home. During the period, the test results were negative on January 5th, 7th, 9th, and the three calculations; the calculation test was positive on January 11; and it was transported by a 120 negative pressure ambulance to Shijiazhuang City People's Hospital Jianhua District on January 12; diagnosed as a confirmed case on January 13. #Confirmed case 3: Female, 55 years old, from Xiaoguozhuang Village, Zengcun Town, Gaocheng District. There will be no going out in the village from December 27, 2020 to December 31, 2020; at noon on January 1, 2021, ride an electric bike to Nanqiao Village, HaoYunLai hotel for a wedding banquet;

FIGURE 1: NER entity boundary tag BIO model display diagram.

patient was transferred, the lavender label represents the name of the hospital where the patient was transferred, and the dark-brown label represents the patient's transportation information.

Although adding BIO can help the model different continuous things, this strategy will have some drawbacks. For example, as the number of labels rises, the danger of an unlawful prediction label sequence increases. The following faults are possible in this paper's epidemic identification task: (1) The model recognizes work information labels in the case of Id, such as "B-case id, I-work information, and I-case Id." (2) The initial label "B-case Id" should appear at the beginning of entity information, and the model forecasts in the middle of entity information, such as "I-case id, B-case id, and I-case id." These incorrect labels will have an impact on the model's overall accuracy. As a result, in label prediction, it is necessary to examine the word-level classification prediction accuracy as much as possible to assure the prediction of the lawful label sequence.

The introduction of the conditional random field model (CRF) in this study may identify the part of the label sequence distribution rules, allowing the label sequence to be legalized; the formula is as follows:

$$\text{Score}(l|s) = \sum_{j=1}^{m} \sum_{i=1}^{n} \lambda_j f_j(s, i, l_i, l_{i-1}), \tag{1}$$

$$p(l|s) = \frac{\exp\left[\text{score}(l|s)\right]}{\sum_{l'} \exp\left[\text{score}\left(l'|s\right)\right]}. \tag{2}$$

Among them, formula (1) represents the score of the feature function on the label sequence I, s represents the sentence that needs to be labeled with a part of speech, i represents the ith word in sentence s, l_i represents the part of speech labeled by the label sequence to be scored for the i th word, and l_{i-1} indicates the labeling sequence that needs to be cut to mark the position of speech for the $i-1$th term.

The outer summation is the sum of the core values of each feature function f_j, and the inner outline is the sum of the feature values of the words at each position in the annotated sentence. Formula (2) is to index and standardize the score, and the probability value $p(l|s)$ of the label sequence I can be obtained. In this paper, the Viterbi algorithm [13] is used to calculate the score of legal label paths quickly and dynamic programming DP is used to solve the optimal path problem in this paper.

2.2. Transfer Learning. Transfer learning [14, 15] is characterized by a given labeled domain and an unlabeled target domain. Even though the data distribution in the two parts is different, the knowledge of D_t can be learned through the knowledge of D_s. Transfer learning is an optimization approach introduced for solving problems with a small number of samples. Transfer learning is typically appropriate when the amount of data in the source domain is sufficient and the amount in the target domain is minimal.

This work uses transfer learning to aid model training to address the underfitting problem caused by a small amount of epidemic information data. Because the quantity of notification messages is frequently insufficient in the early stages of epidemic transmission, only a small amount of data is used for model training, resulting in the undergeneralization of the model. As a result, this research gathers a Chinese information notification text corpus from Wikipedia. It employs a BERT model based on mask language (MLM) for self-supervised training to gain additional text collocation and sentence structure information.

The Bidirectional Encoder Representations from Transformer (BERT) language preprocessing paradigm can automatically extract rich word-level characteristics, grammatical and structural features, and semantic information from sequences. The BERT model utilized in this study includes 12 layers, and the output of each hidden layer after pretraining may convey word vectors to varying degrees. The hidden layer output vector of the last layer is commonly

used in the BERT model, although some tests demonstrate that the later the layer vector is, the more difficult it is to match the present job itself [16]. To improve the retention of high-dimensional abstract information and the generalization ability of the pretraining model, this paper fuses the outputs of the last four layers of the pretrained model $k_1 * a_{nt} + k_2 * a_{(n-1)t} + k_3 * a_{(n-2)t} + k_4 * a_{(n-3)t}$, where the a_{nt} is the model output vector, n is the number of layers, t is the time step, and the weight coefficients $k1$, $k2$, $k3$, and $k4$ are 0.1, 0.2, 0.3, and 0.4, respectively.

2.3. Rule Matching. Rule matching is a way of searching for matches in highly regularized data [17]. When extracting epidemic information in this research, it is discovered that some data information is very standardized, as shown in Figure 2. For example, a "certain X-year-old" statement has a greater confidence that the statement describes the age or a "XXXX year was born" statement has higher confidence that the information describes the date of birth.

The examination of the epidemic text reveals that several fundamental attribute interactions exhibit consistent patterns in the text. By evaluating the samples and removing interference factors such as manual annotation mistakes, the regular expression is utilized to extract the information attribute connection. The fusion data is searched using matching. Using the rule matching strategy to lower the cost of label annotation might focus the model on less regular label categorization, improving the model's accuracy.

2.4. Data Enhancement. Data augmentation is a technique that synthesizes new data from existing data. When training samples are sparse and labeling costs are high, it seeks to enhance the number of data points, reduce overfitting, and improve the model's overall generalization ability [18, 19].

Due to a lack of case texts in the epidemic's early stages, the model struggles to learn more detailed rules based on a few samples. Furthermore, several tags in the pandemic text, such as transit information and employment information, have a substantially lower number of occurrences than other tags. Without data augmentation, the model cannot detect the basic information of such labels. Synonym substitution (SR), random insertion (RI), spontaneous exchange (RS), random deletion (RD), and entity mention substitution are all typical data augmentation strategies used in NLP jobs [20]. The methods of synonym substitution and entity mention substitution are primarily used in this article. However, using synonymous substitution (SR) may not have the desired impact of fundamentally improving data. This is because the vector values of the interchangeable substitute terms are almost identical, resulting in enhanced subjective data. On the other hand, the trained model may perceive the augmented data as the exact phrase and there is no significant data extension.

This paper presents a solution to the above-listed challenges: first, the part of speech of each word in the original text is identified, next, comments with the same amount of speech but different are disrupted into other sentences to replace synonyms, and finally, the credibility of the data-enhanced text is checked and the data-enhanced text is

merged into the training corpus, for example, "Dining at an HSBC Hotel" and "Shopping in Joy City on X," where words like "HSBC Hotel" and "Joy City" can be identified as synonymous with different words; "Dining" and "Shopping" belong to the same entity type, and the enhanced data "dining in the Grand Hyatt" can be obtained through the improved synonym substitution scheme.

3. Construction of the Joint Learning Model Based on Topic Recognition and Named Entity Recognition

3.1. Principle of the TBR-NER Model. The corpus required for the TBR-NER joint learning model comes from two aspects: (1) the new epidemic notification corpus was extracted according to the epidemic-related information announcement issued by the National Health Commission and (2) the Chinese notification corpus was obtained from the Chinese Wikipedia corpus. After removing the text information, the two parts of the canon are data cleaned and the epidemic notification corpus is annotated. Then, the outbreak notification corpus is data augmented using entity mention replacement and external knowledge base replacement. Secondly, pretrain the Chinese information notification text corpus in Wikipedia using the MLM language model to construct the corresponding dictionary mapping. At the same time, after data enhancement of the epidemic corpus, the related word vector needs to be obtained according to the index. Thirdly, a standard learning method based on topic recognition and named entity recognition (TBR-NER) is used to train and predict the marked information. The effect of entity recognition and classification is improved with the help of the topic of text information. Finally, after the postcorrection of the results, the optimal results are obtained. The TBR-NER model framework is shown in Figure 3.

3.2. BERT Model Based on Masking Language (MLM). The BERT model based on MLM is a new language representation model released by the Google AI team in 2018, which is the bidirectional encoder representation of Transformer [21]. Unlike other language representation models, BERT is aimed at pretraining a deep bidirectional representation by jointly adjusting the context in all layers. Thus, the pretrained BERT representation can be fine-tuned by an additional output layer. It does not need to modify the architecture of specific tasks and can be applied to various fields of task model construction.

There are two commonly used BERT models:

(1) BERTBASE: $L = 12$, $H = 768$, $A = 12$, and total parameter $= 110$ M

(2) BERTLARGE: $L = 24$, $H = 1024$, $A = 16$, and total parameter $= 340$ M

BERT has excelled in several NLP tasks, including categorization, question answering, and translation [22–24]. A 12-layer BERT model is used in this work. The model's

#无症状感染者 2：男，1978 年出生，通化市人，住址为通化市东昌区东苑小区，系 1 月 15 日通报的无症状感染者 6 的儿子。1 月

#Asymptomatic infected person 2: Male, born in 1978, a native of Tonghua City, whose address is Dongyuan Community, Dongchang District, Tonghua City. He is the son of asymptomatic infected person 6 reported on January 15. January

#无症状感染者 3：男，1955 年出生，通化市如人，住址为通化市东昌区厚德载物 C 区，系 1 月 17 日通报的无症状感染者 21 的丈夫。

#Asymptomatic infected person3: Male, born in 1955 in Tonghua City, where he lives in Zone C, Houdezaiwu, Dongchang District, Tonghua City. He is the husband of asymptomatic infected person 21 notified on January 17.

FIGURE 2: Have highly normalized data.
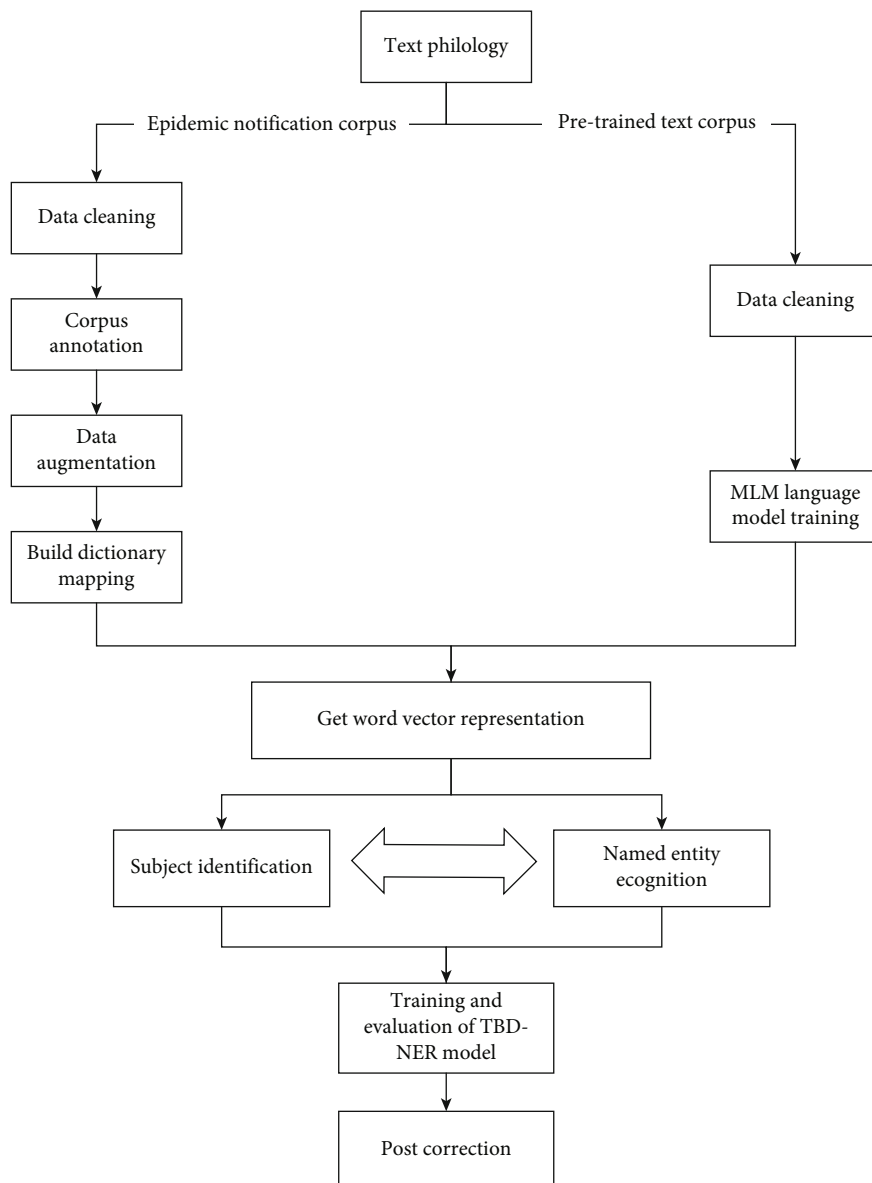
FIGURE 3: Flow chart of model construction.

primary distinction from prior models is that it recommends predicting the LOSS value of the following phrase by transforming a limited number of words into masks or randomly changing another word with a specific probability. The primary purpose of the BERT model is to acquire word vectors via massive text learning, assess the relationship between distinct token units, and then map the association to variables. The BERT model employs two-stage model training

to address the widespread occurrence of polysemy in the notification information. It starts with two-way language model training and then moves to the fine-tuning mode to address downstream objectives. As a result, the word vector taught by BERT has a high impact and flexibility.

3.3. Two-Stage Hierarchical Learning Model Based on Topic Prediction. For epidemic text information, first divide each epidemic notification text into five topics: case basic information introduction, case trajectory information during illness, home information during case illness, case diagnosis information, and other types. Then, the entity information is divided into 16 categories: patient ID, the patient's native place, the patient's residence, the patient's working class, the patient's diagnosis time, the patient's opening trajectory movement time, the patient's moving trajectory starting place, the patient's moving trajectory termination place, the patient's trajectory action event, the patient riding the transportation tool, the transportation tool information, the patient's home time, the patient's diagnosis time, the patient being transported by the hospital time, the patient being fascinated by the hospital name, and the irrelevant information. The 15-type entity information corresponds to BI's start and intermediate labels in sequence annotation and irrelevant entity information. As a result, 15 types of entity information matching 30 category labels, plus unrelated classes, must be forecasted for each letter as a category in 31 categories.

The two-stage hierarchical learning concept is as follows to improve the model's prediction of label information: first, the text statement is used to categorize the topics. Following completion of the categorization, the entity classification is based on the labels that may exist in the current statement topic. Figure 4 depicts a two-stage hierarchical learning model based on topic prediction. Two-stage hierarchical learning is utilized to scatter the titles of the original 31 categories into multiple topic categories. The label types for each topic are few, and the model is difficult to misinterpret.

Two-stage hierarchical learning based on topic prediction is mainly divided into two steps:

(1) Subject classification is carried out on a notification information text, and the text information is divided into one of five types of subjects

(2) For example, the text of the epidemic information notification is as follows in "case 1: female, 18 years old, from Shijiazhuang, Hebei, salesperson of Ping An supermarket, and living in Ping An community. On January 18, she ate at HSBC Hotel. She did not get out from home on January 19. On January 20, the nucleic acid test was positive. On January 21, after consultation with the expert group, it was diagnosed as new coronary pneumonia." In the abovementioned epidemic notification information, identifying "Ping An community" as a "patient residence" requires two steps: first, it is necessary to determine its topic. The sentence is as follows in "case 1: female, 18 years old, from Shijiazhuang, Hebei, a salesperson in a Ping An supermarket, and living in a Ping An community." It

belongs to the topic category "introduction to basic information of cases;" then, under the topic of "introduction to basic information of cases," use model 1 to identify and classify entities. The possible entity categories under this topic include patient id, patient origin, patient residence, patient workplace, and irrelevant information

3.4. Two-Stage Joint Learning Based on Topic Prediction. The suggested two-stage hierarchical learning based on topic prediction used the subject information of the epidemic notice text in the last part. However, there are still the following difficulties in the actual application scenario:

(1) For error accumulation problem due to entity label prediction after two stages, if the topic prediction model is weak in the first step, the total results are not very good no matter how accurate the succeeding entity recognition model is. The subject prediction error accumulates in the first step, and the influence of the subject prediction becomes the bottleneck of the succeeding prediction

(2) For the problem of model redundancy, in two-stage hierarchical learning, a large number of models need to be used for different tasks, which is more cumbersome in practical applications

(3) Splitting the two tasks of subject information and entity recognition of sentences in the task splitting problem prevents the two pieces of information from supplementing each other. The improvement in the entity recognition effect in the second stage will have no feedback effect on the sentence classification effect. Improved topic classification accuracy will simply enhance the bottleneck of the total classification impact, and there is no benign mutual reference between topic recognition and entity recognition

Based on the abovementioned issues, this work suggests a novel two-stage joint learning technique based on topic prediction: topic-based discrimination. TBR-NER is based on the combined learning model of topic recognition and named entity recognition. TBR-NER is a step forward based on the hierarchical learning paradigm. The original topic recognition and entity recognition are fused based on the exact usage of the subject information of the notification text, allowing the model to fulfill the two tasks of topic recognition and entity recognition simultaneously, as illustrated in Figure 5.

The TBR-NER model proposed in this paper is divided into the following steps:

(1) First, perform character-level segmentation on each sentence. This article does not use the conventional method of first-word segmentation, and then, input it into the model because word segmentation will have a priori bias. Different word segmentation algorithms will aggregate other numbers of characters into words, and these presegmented words are not necessarily suitable for the current task. BERT's
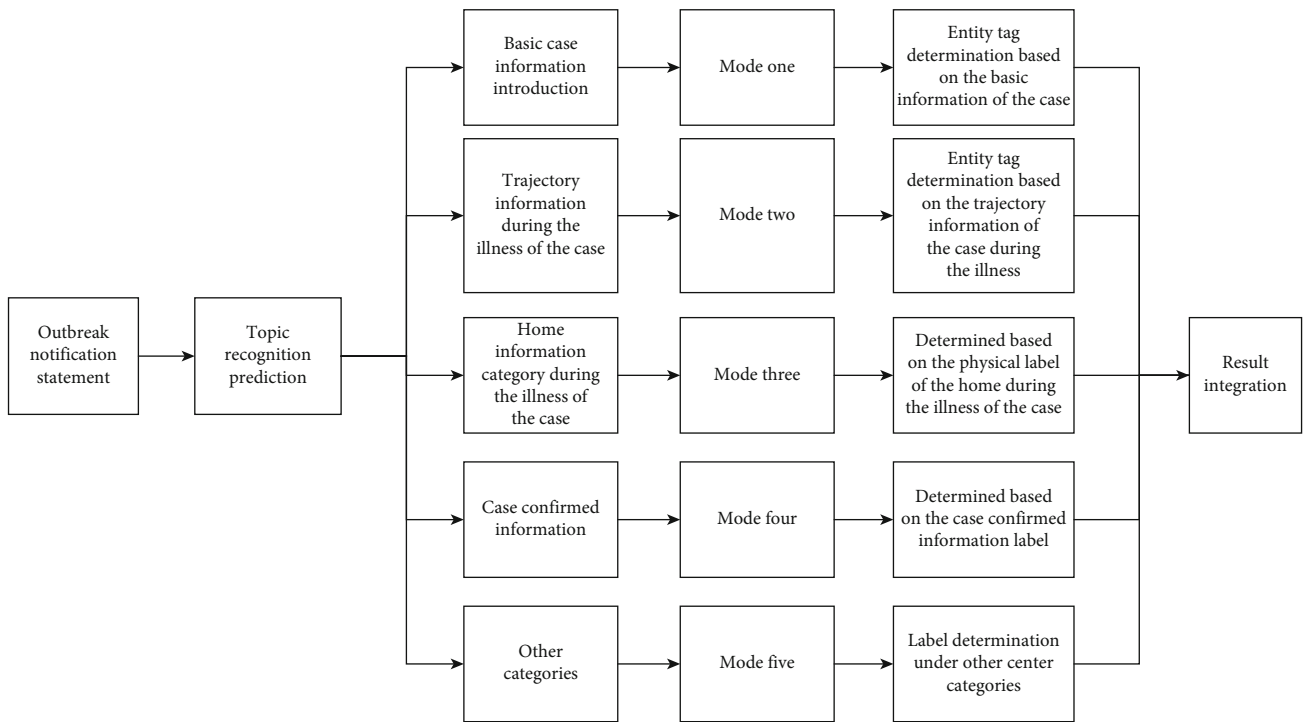
FIGURE 4: A two-stage hierarchical learning model based on topic prediction.

FIGURE 5: TBR-NRE model.

| \<TOPIC\> | 病 | 例 | 1 | ： | 女 | 18 | 岁 | ， | 河 | 北 | 石 | 家 | 庄 | 人 | 。 |

病 (disease)
例 (sample)
女 (female)
岁 (old)
河北 (Hebei province)
石家庄 (Shijiazhuang city)
人 (person)
病例1：女18岁，河北石家庄人。
(Case1: female 18 years old, from shijiazhuang, hebei.)
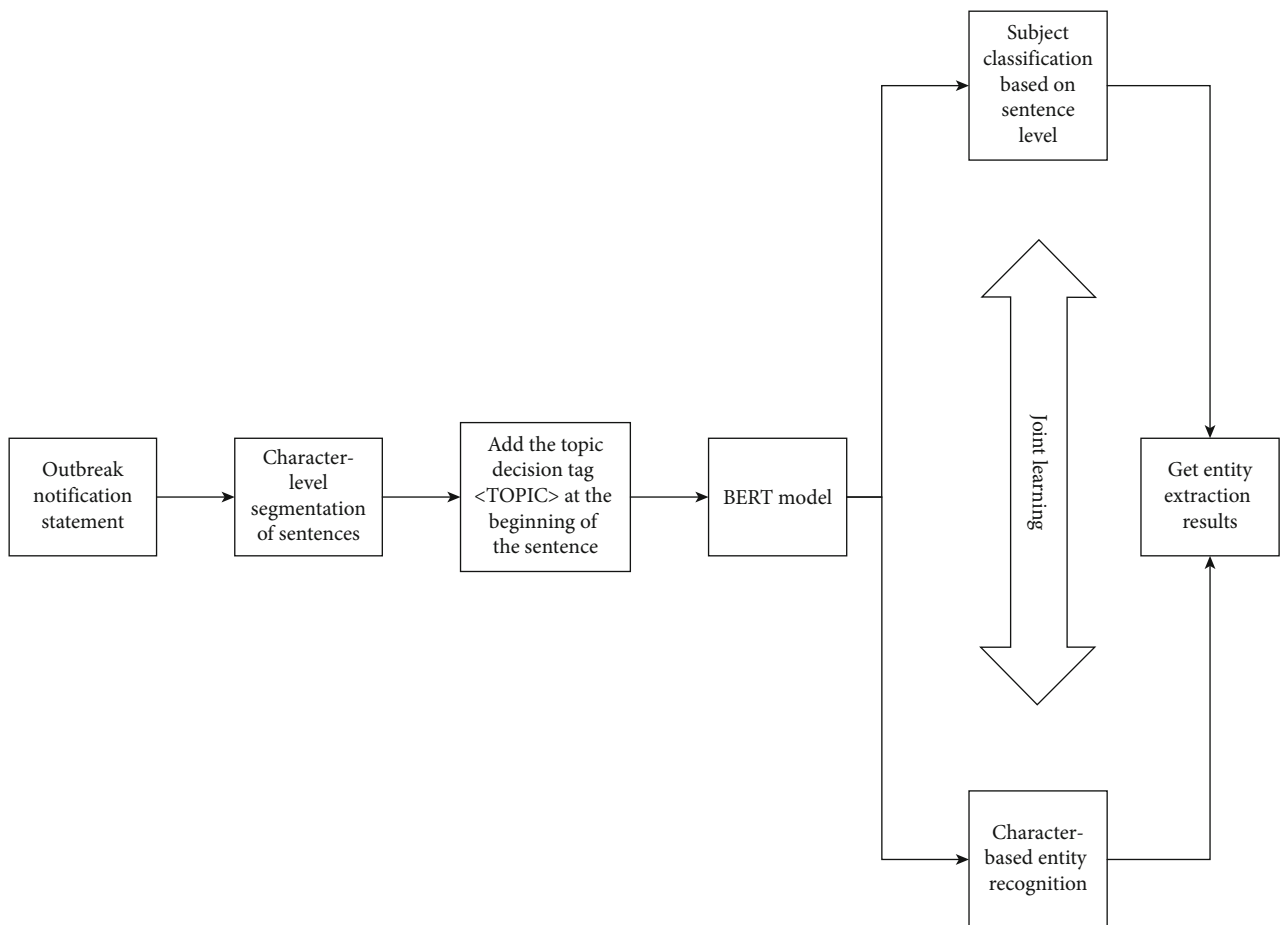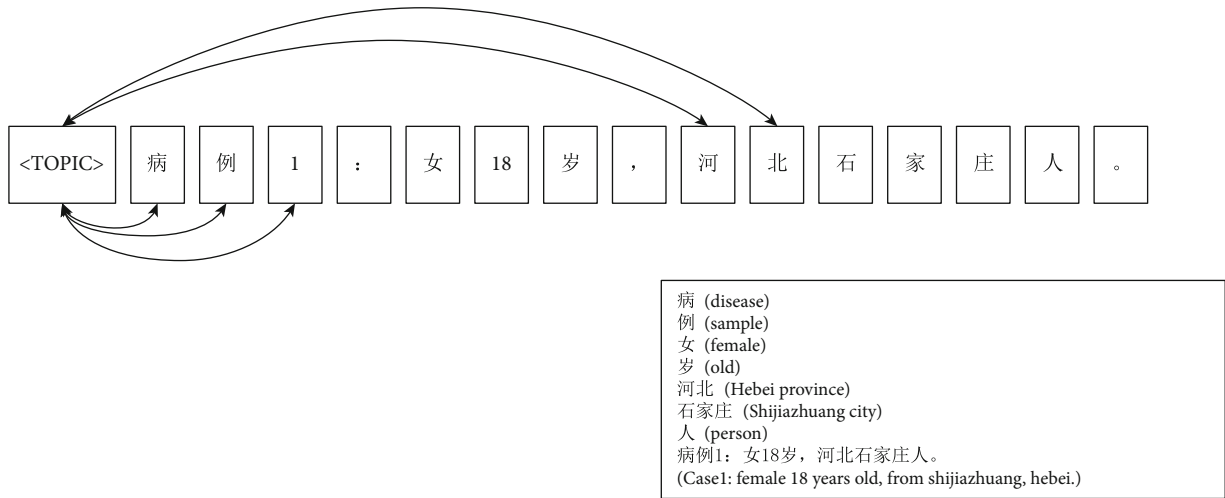
FIGURE 6: The mutual learning process of topic recognition and entity recognition.
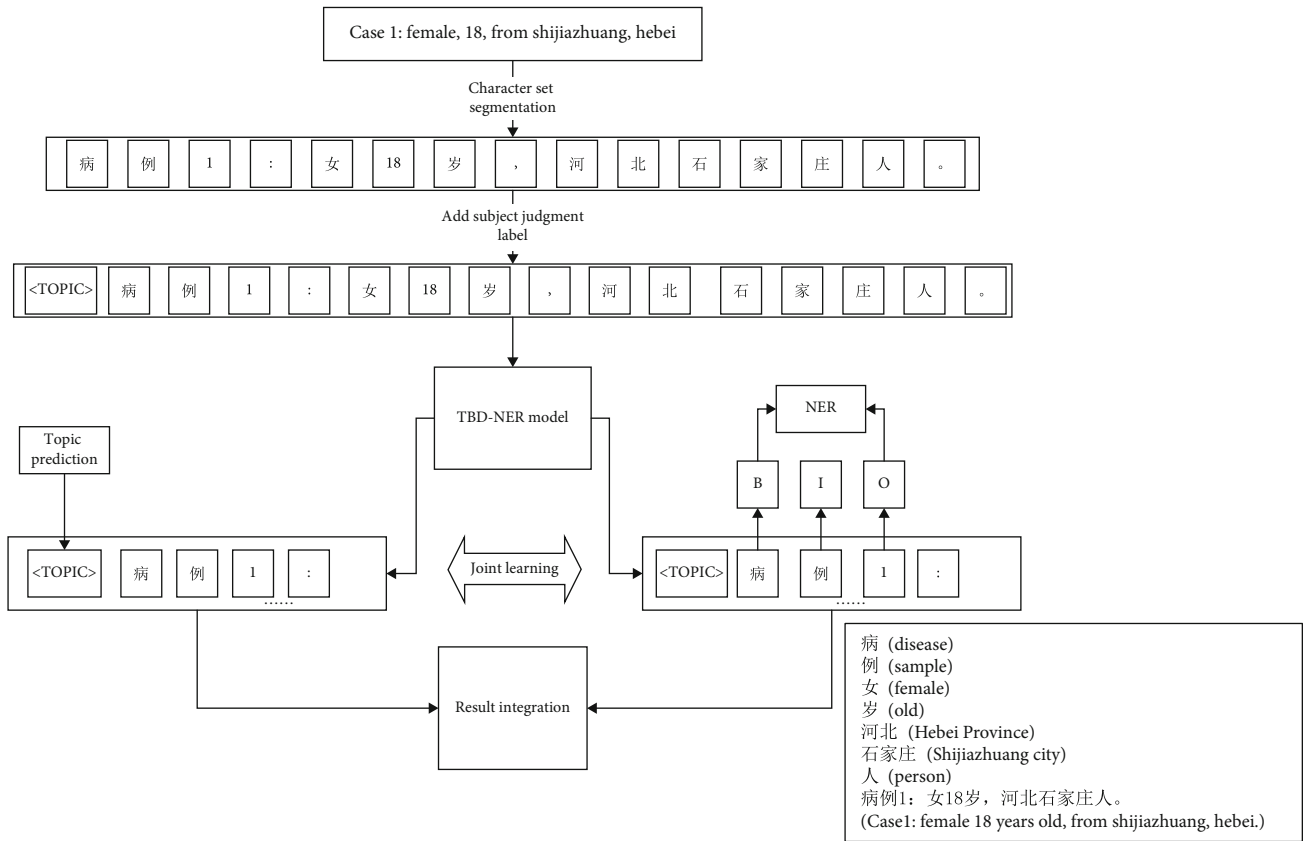
FIGURE 7: TBR-NER model learning case.

self-attention mechanism model will aggregate characters according to the current job at the low level. Therefore, this paper chooses Chinese characters or numbers as the split unit

(2) For topic self-labeling based on entity labeling results, add a "\<TOPIC\>" tag to the beginning of each sentence. This tag is used for the topic classification of the current correction, which contains the

topic information of the current sentence, as shown in Figure 6. The topic information in this article does not require additional artificial annotation. The topic is determined by the content of the entity tag marked in the previous stage, which is a "self-labeling" process. For example, when a "case id" or "case origin" entity appears in a sentence, the sentence will be automatically marked as a "case basic information introduction category." The topic self-labeling
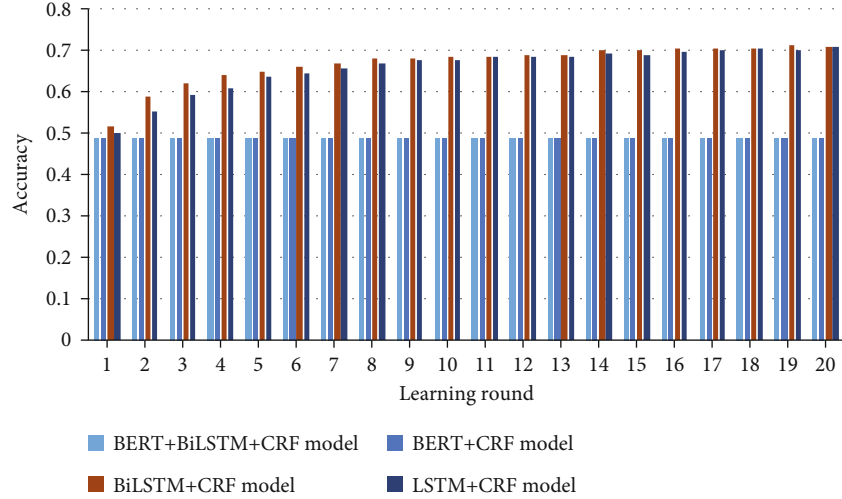
Figure 8: Four kinds of NER model entity recognition independent test set accuracy.
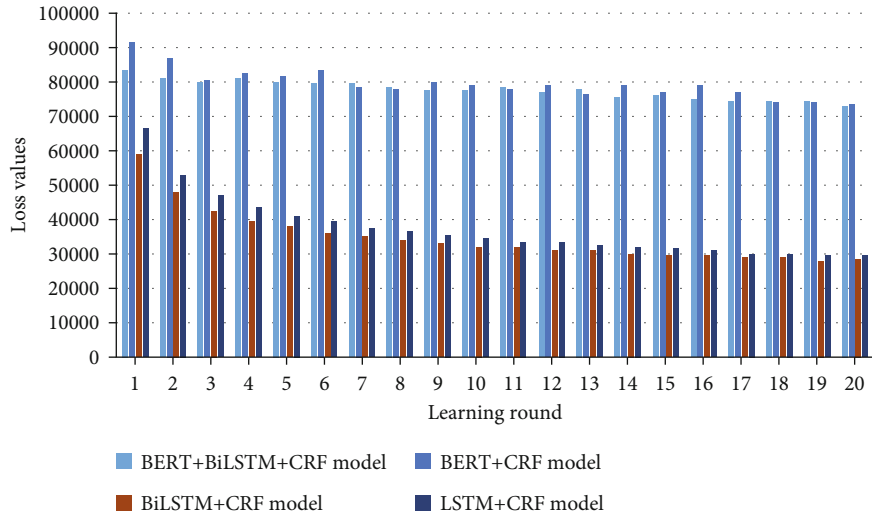
Figure 9: Loss values of four NER models.

process based on entity labeling results saves the manual labeling process and does not introduce additional knowledge

(3) Predict the topic category of the sentence where the "<TOPIC>" tag is located, and perform entity recognition on the sentence content. Since "<TOPIC>" itself is located in the sentence, the model can also observe the topic representation of "<TOPIC>" while doing entity recognition and "<TOPIC>" itself can also be used for topic classification. See the entity representation of the different characters themselves. Topic recognition and entity recognition learn from each other to achieve better prediction results. The learning process of the TBR-NER model is shown in Figure 7

3.5. TBR-NER Model Loss Function. The model loss function of TBR-NER proposed in this paper mainly consists of the following two parts:

$$\text{Loss} = \partial \text{loss}_{\text{idea}} + \beta \text{loss}_{\text{ner}}. \tag{3}$$

∂ and β correspond to their proportion weights, and by adjusting them, the model can pay more attention to a certain task. $\text{Loss}_{\text{idea}}$ represents the prediction loss of a sentence, which is called cross-entropy, and the formula is as follows:

$$\text{Loss}_{\text{idea}} = -\sum_x p(x) \log q(x). \tag{4}$$

Among them, $p(x)$ is the actual probability distribution and $q(x)$ represents the predicted probability distribution.

Loss_{ner} is the loss function of the conditional random field (CRF) optimization algorithm introduced in the summary of Section 3.1, and the formula is as follows:

TABLE 1: Analysis and comparison of four classical NER models.

	BERT + BiLSTM + CRF	BERT + CRF	BiLSTM + CRF	LSTM + CRF
Accuracy rate	0.4904	0.4904	0.7132	0.7098
Loss value	73349.1 9	73936.48	28270.13	29563.14
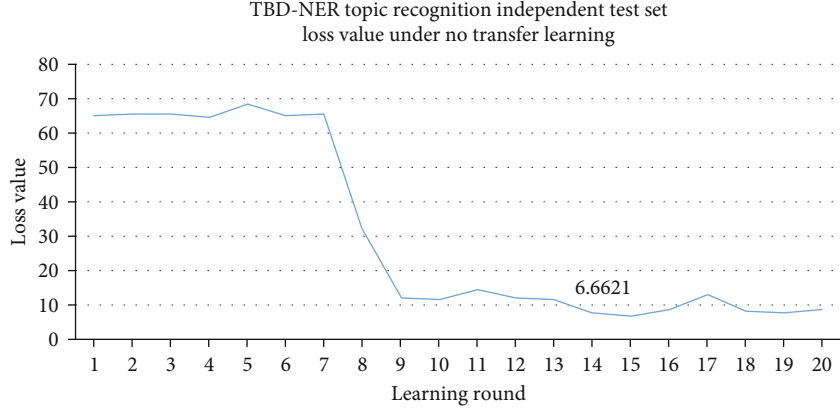
FIGURE 10: TBR-NER-independent test set accuracy without transfer learning.
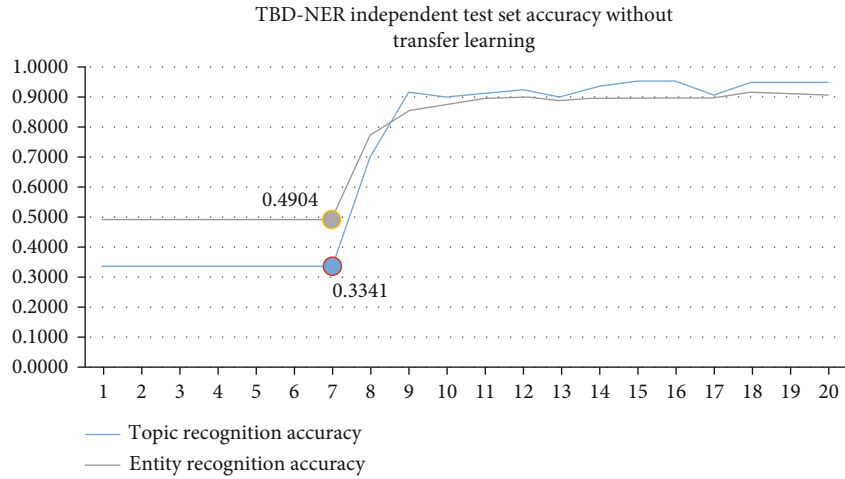
FIGURE 11: TBR-NER topic recognition with independent test set loss values without transfer learning.

$$\text{Loss}_{ner} = p(y_1, y_2, \cdots, y_n | x) = \frac{1}{z(x)} e^{h(y_1 : x) + \sum_{i=2}^{n} g(y_i : y_{i-1}) + h(y_i : x)}. \tag{5}$$

Among them, $h(y_i : X)$ represents the emission fraction, $g(y_i ; y_{i-1})$ represents the transfer score, $z(x)$ represents the normalization factor, and $P(y_1 y_2, \cdots, y_n | X)$ mark the probability of $y_1 y_2, \cdots, y_n$.

4. Experimental Results and Analysis

4.1. Analysis of Experimental Results. The text information of epidemic notifications issued by the Health Commission in Jilin province and Hebei province from January 10 to February 14, 2021, was chosen as the data source to test the validity of the suggested approach in the task of epidemic notification information. The dataset was chosen randomly, with 70% as a training set and 30% as an independent test set. The accuracy and ultimate loss of the separate test set are utilized as evaluation model indicators, with accuracy defined as follows:

$$\text{ACC}_{word} = \frac{\sum \text{token}_i}{\text{sum}(\text{word})}. \tag{6}$$

In formula (6), sum(word) represents the sum of all characters and token_i represents the sum of correctly predicted characters, which is the accuracy of the expected characters. In this study, four NER models were selected as the comparative reference experiments: BERT + BiLSTM

Loss value of independent test set of TBD-NER entity
recognition without transfer learning
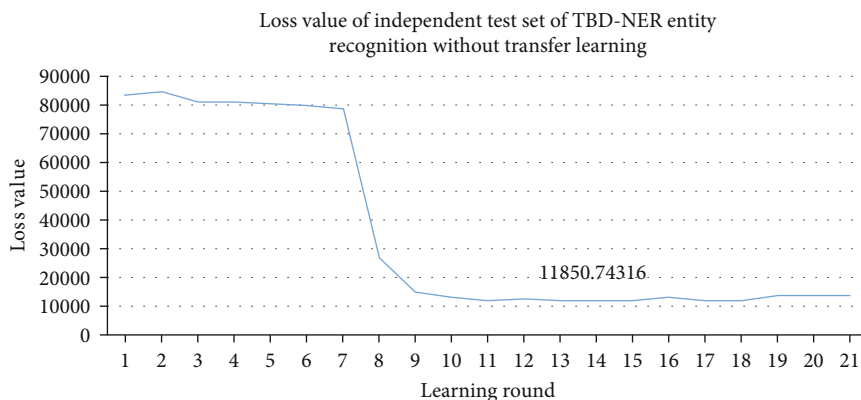
Figure 12: Loss values of independent test sets for the TBR-NER entity.

+ CRF [25], BERT + CRF [26], BiLSTM + CRF [27], and LSTM + CRF [28]. The experimental results of each model's first 20 rounds of learning were compared.

4.2. Model without Transfer Learning. The accuracy and loss values of four NER models in entity recognition are compared without transfer learning. The accuracy rates of the four NER models are relatively low, as seen in Figures 8 and 9. In 20 rounds of education, the BERT + BiLSTM + CRF model and the BERT + CRF model were both less than 0.49. They cannot converge for an extended period throughout the training process, resulting in gradient dispersion. The model cannot break past the bottleneck, resulting in a significant loss of values and an optimal local dilemma. Although the BiLSTM + CRF and LSTM + CRF models outperform the previous two models and the model loss values converge fast, the model accuracy remains poor. Table 1 compares the accuracy and loss values of the NER model for entity recognition in 4.

To assess the effect of entity recognition fairly, the TBR-NER model also performs 20 rounds of learning without incorporating transfer learning and makes entity recognition prediction and topic recognition prediction using epidemic trajectory information. The experimental findings are depicted in Figures 10–12. The TBR-NER model's fitting performance has been dramatically enhanced compared to that of the previous four NER models. Under the same settings, the accuracy of the TBR-NER model after the seventh cycle of learning. Table 2 displays the accuracy and loss values of topic recognition tasks and entity recognition tasks using TBR-NER without transfer learning.

The experimental findings show that the TBR-NER model can overcome the loss bottleneck after many learning rounds and predict with high accuracy. This is because the model entity recognition task relies heavily on the topic categorization job. The TBR-NER model's main algorithm is a two-stage joint learning technique based on topic recognition. The topic classification task of the statement is straightforward in the iterative stage of the initial model, and the topic representation vector has more excellent separability in high latitudes (see Figure 13). The entity recognition task is driven by the significant growth in the subject categorization job. Topic recognition and entity recognition continue

Table 2: Analysis and comparison of the TBR-NER model without transfer learning.

	TBR-NER model for topic recognition	TBR-NER model entity recognition
Accuracy rate	0.9709	0.9160
Loss value	6.0467	11850.74

to learn in the following learning phase, achieving synchronous convergence. The model's accuracy under entity recognition and classification reached the superior value of 91.60 percent when the TBR-NER model was taught in the 18th round. Although the TBR-NER model considerably improves prediction outcomes when compared to its four types of NER models, the model's accuracy does not achieve the optimum state in the first seven learning periods. After the model passes through the loss domain, its lost value rapidly converges and the model's accuracy eventually tends to be stable.

4.3. Model under Transfer Learning. This work employs the transfer learning auxiliary model to train and improve the model's generalization performance to address the underfitting problem caused by the small amount of epidemic information data. To maintain the fairness of the control experiment in the prior summary experiment, the above-mentioned TBR-NER model does not apply the transfer learning optimization model. In this overview, transfer learning will optimize the TBR-NER model and compare the outcomes.

Figures 14–16 exhibit the accuracy of topic recognition, entity recognition, and loss value of the TBR-NER model under transfer learning. When the TBR-NER model learns in the sixth round, the classification accuracy of entity identification may reach 95.85 percent, which is higher than the TBR-NER model's accuracy without transfer learning. Furthermore, the loss value of the TBR-NER model under transfer learning is as low as 5866, which is considerably less than the loss value of the four NER models discussed in Section 3.2. The smaller the loss value, the less the gap
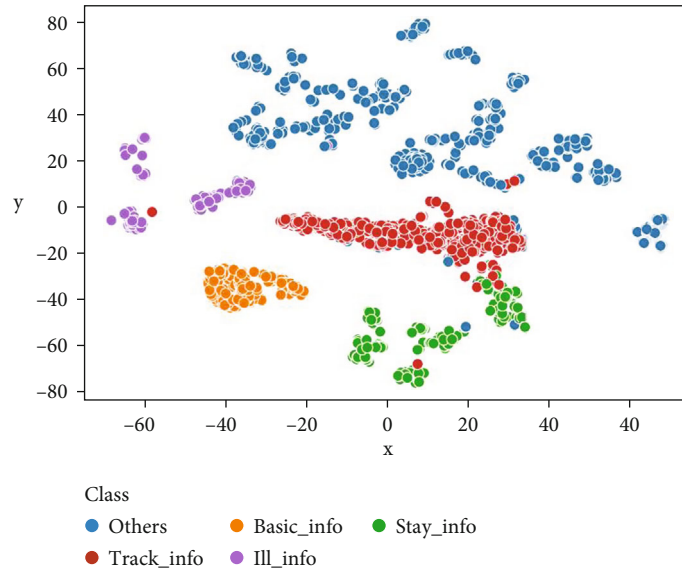
Figure 13: Visual display of 5 categories of topic t-SNE dimensionality reduction.
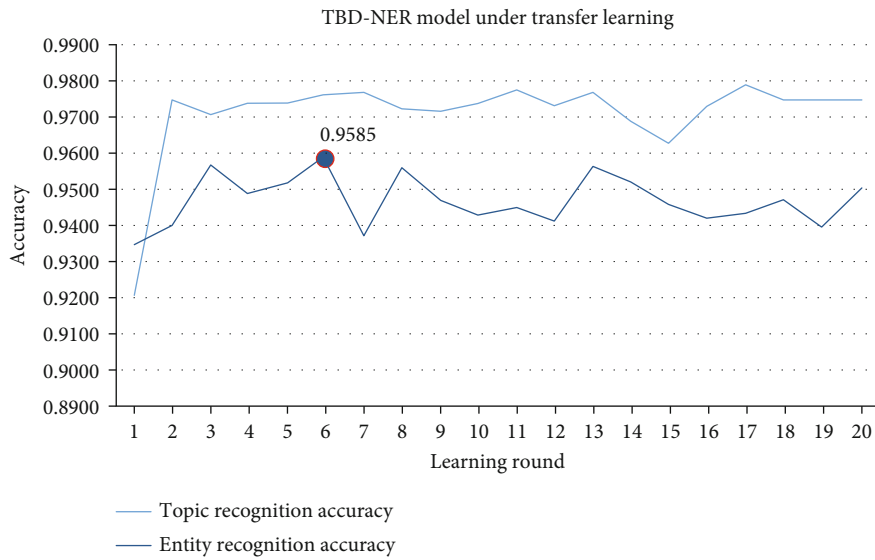
Figure 14: TBR-NER model under transfer learning.
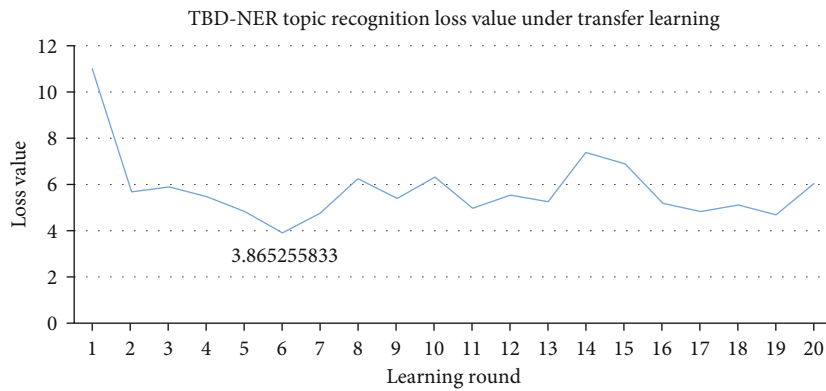
Figure 15: Loss value of TBR-NER topic recognition under transfer learning.
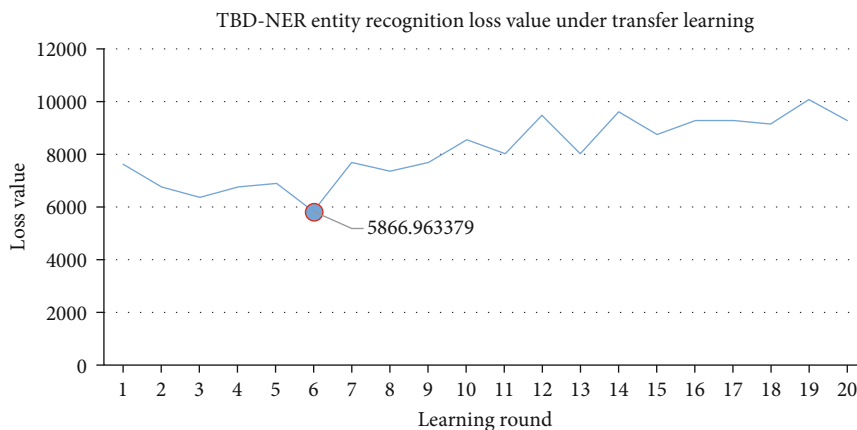
FIGURE 16: Loss value of TBR-NER entity recognition under transfer learning.

TABLE 3: Analysis and comparison of the TBR-NER model under transfer learning.

	TBR-NER model for topic recognition	TBR-NER model entity recognition
Accuracy rate	0.9791	0.9585
Loss value	3.8656	5866.963

TABLE 4: Comparison of TBR-NER under transfer learning and TBR-NER without transfer learning.

	Topic recognition of the TBR-NER model under transfer learning	Entity recognition of the TBR-NER model under transfer learning	Topic recognition of the TBR-NER model without transfer learning	Entity recognition of the TBR-NER model without transfer learning
Accuracy rate	0.9791	0.9585	0.9709	0.9160
Loss value	3.8656	5866.963	6.0467	11850.74

TABLE 5: Comparison of entity recognition between the TBR-NER model and other four NER models.

	BERT + BiLSTM + CRF	BERT + CRF	BiLSTM + CRF	LSTM + CRF	TBR-NER
Accuracy rate	0.4904	0.4904	0.7132	0.7098	0.9585
Loss value	73349.1 9	73936.48	28270.13	29563.14	5866.963

between the model's forecast and reality and the more convergent the model. Furthermore, a comparison of Figures 10 and 14 reveals that the accuracy of entity identification and topic recognition in the TBR-NER model impacts each other, an intuitive representation of the TBR-NER model's joint learning.

The analysis of topic recognition and entity recognition of the TBR-NER model under transfer learning is shown in Table 3. The comparison results of topic recognition and entity recognition between TBR-NER under transfer learning and TBR-NER without transfer learning are shown in Table 4. The comparison results of entity recognition between the TBR-NER model under transfer learning and the other four NER models are shown in Table 5.

4.4. Generalization of the TBR-NER Model. To verify the generalization of the TBR-NER model, this article uses 89 data from February 18 to March 1, 2021, in Hebei province's

TABLE 6: The prediction results were analyzed and compared.

Assessment method	Exact number	Total number	Accuracy
Word (unit)	16002	17644	0.91
Entity word (unit)	919	1052	0.87

epidemic notification information as an independent test set. The word-to-word entity recognition accuracy is used as the evaluation index. Table 6 displays the accuracy rate. Table 6 shows that the prediction accuracy of outcomes with words as labels may reach over 90%, demonstrating that the proposed TBR-NER model has strong generalization.

5. Conclusions

This paper proposes a novel text information extraction method for COVID-19 outbreak notification based on joint

learning of topic recognition and named entity recognition. The model can simultaneously complete the task of subject recognition and entity recognition, realize the subject self-labeling process based on the labeling results, and effectively improve the accuracy of the classification of epidemic notification information. This method employs rule matching and conditional random fields for word segmentation based on the regularization characteristics of epidemic information description. It uses data enhancement and transfers learning techniques to solve the model's generalization ability in the absence of samples and improve the analytical power of text information. The experiment is put up against four different natural language processing systems. The findings reveal that the suggested method's accuracy on the test set has dramatically improved and the loss function has been significantly lowered. Furthermore, the topic recognition task and the entity recognition task have many mutual promotions in the TBR-NER model's learning process, demonstrating that learning topic recognition and entity recognition together in text information extraction complement each other. The final model generated in this article is validated using datasets from Jilin and Hebei provinces. The concept will be expanded in the future to include datasets from more regions, giving important information for epidemic prediction and prevention.

In a future work, we plan to explore more countries' new crown epidemic notification information and text label information in other languages. Other countries and regions can use this model to predict their text trajectory information. Younes et al. [29] used NLP-related technologies to transcribe Arabic and Latin languages. After exploration and research, we can try to transplant the text information model to Arabic and African languages in the future to contribute to global epidemic prevention. Updating and improving the model to improve predictions' accuracy will significantly help in epidemic prevention and control. In addition, the model can be inherited in the GUI through pyqt5 and combined with the real-time map software to provide user convenience and epidemic prevention. Users can use the model to extract actual location words and path trajectories into the JSON data format and dynamically display the epidemic trajectory route map through web pages in real time.

Data Availability

The experimental data of this paper comes from the website, and the specific website is as follows: epidemic situation notification website of Hebei province, China, http://wsjkw .hebei.gov.cn/gzdt/index_43.jhtml, and epidemic situation notification website of Jilin province, China, http://wsjkw.jl .gov.cn/xwzx/xwzx/index_41.html.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

XF and QY conceived the project, designed the experiments, and drafted the manuscript. RH, YR, ZH, and ZF collected the data and conducted the experiments. RH, YR, and XF proofread and polished the manuscript and organized this project.

References

[1] J. Wu, J. Wang, S. Nicholas, E. Maitland, and Q. Fan, "Application of big data technology for COVID-19 prevention and control in China: lessons and recommendations," *Journal of Medical Internet Research*, vol. 22, no. 10, article e21980, 2020.

[2] C. A. D. Durai, A. Begum, J. Jebaseeli, and A. Sabahath, "COVID-19 pandemic, predictions and control in Saudi Arabia using SIR-F and age-structured SEIR model," *The Journal of Supercomputing*, vol. 78, no. 5, pp. 7341–7353, 2022.

[3] Q. Ye, J. Zhou, and H. Wu, "Using information technology to manage the COVID-19 pandemic: development of a technical framework based on practical experience in China," *JMIR Medical Informatics*, vol. 8, no. 6, article e19515, 2020.

[4] J. Li, S. Zhao, J. Yang et al., "WCP-RNN: a novel RNN-based approach for bio-NER in Chinese EMRs," *The Journal of Supercomputing*, vol. 76, no. 3, pp. 1450–1467, 2020.

[5] S. Bhaskar, S. Bradley, S. Sakhamuri et al., "designing futuristic telemedicine using artificial intelligence and robotics in the COVID-19 era," *Frontiers in Public Health*, vol. 8, article 556789, 2020.

[6] S. J. Alsunaidi, A. M. Almuhaideb, N. M. Ibrahim et al., "Applications of big data analytics to control COVID-19 pandemic," *Sensors*, vol. 21, no. 7, article 2282, 2021.

[7] M. Khayyat, K. Laabidi, N. Almalki, and M. Al-Zahrani, "Time series facebook prophet model and python for COVID-19 outbreak prediction," *Computers, Materials, & Continua*, vol. 67, no. 3, pp. 3781–3793, 2021.

[8] J. Song, H. Xie, B. Gao, Y. Zhong, C. Gu, and K.-S. Choi, "Maximum likelihood-based extended Kalman filter for COVID-19 prediction," *Chaos, Solitons & Fractals*, vol. 146, article 110922, 2021.

[9] H. M. Balaha, E. M. el-Gendy, and M. M. Saafan, "CovH2SD: a COVID-19 detection approach based on Harris hawks optimization and stacked deep learning," *Expert Systems with Applications*, vol. 186, article 115805, 2021.

[10] M. Wieczorek, J. Siłka, and M. Woźniak, "Neural network powered COVID-19 spread forecasting model," *Chaos, Solitons & Fractals*, vol. 140, article 110203, 2020.

[11] K. Kozioł, R. Stanisławski, and G. Bialic, "Fractional-order SIR epidemic model for transmission prediction of COVID-19 disease," *Applied Sciences*, vol. 10, no. 23, p. 8316, 2020.

[12] I. N. Yulita, M. I. Fanany, and A. M. Arymurthy, "Gesture recognition using latent-dynamic based conditional random fields and scalar features," *Journal of Physics: Conference Series*, vol. 812, article 012113, 2017.

[13] A. Allam and M. Krauthammer, "PySeqLab: an open source Python package for sequence labeling and segmentation," *Bioinformatics*, vol. 33, no. 21, pp. 3497–3499, 2017.

[14] F. Zhuang, Z. Qi, K. Duan et al., "A comprehensive survey on transfer learning," *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43–76, 2021.

[15] A. Khamparia, D. Gupta, V. H. C. de Albuquerque, A. K. Sangaiah, and R. H. Jhaveri, "Internet of health things-driven deep learning system for detection and classification of cervical cells using transfer learning," *The Journal of Supercomputing*, vol. 76, no. 11, pp. 8590–8608, 2020.

[16] M. Zhang, G. Geng, S. Zeng, and H. Jia, "Knowledge graph completion for the Chinese text of cultural relics based on bidirectional encoder representations from transformers with entity-type information," *Entropy*, vol. 22, no. 10, p. 1168, 2020.

[17] N. Kang, B. Singh, Z. Afzal, E. M. van Mulligen, and J. A. Kors, "Using rule-based natural language processing to improve disease normalization in biomedical text," *Journal of the American Medical Informatics Association*, vol. 20, no. 5, pp. 876–881, 2013.

[18] T. Kang, A. Perotte, Y. Tang, C. Ta, and C. Weng, "UMLS-based data augmentation for natural language processing of clinical research literature," *Journal of the American Medical Informatics Association*, vol. 28, no. 4, pp. 812–823, 2021.

[19] C. Mi, S. Zhu, and R. Nie, "Improving loanword identification in low-resource language with data augmentation and multiple feature fusion," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 9975078, 9 pages, 2021.

[20] J. Wei and K. Zou, "EDA: easy data augmentation techniques for boosting performance on text classification tasks," in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pp. 6382–6388, Hong Kong, China, 2019.

[21] X. Xing, R. Xie, and W. Zhong, "Model-based sparse coding beyond Gaussian independent model," *Computational Statistics & Data Analysis*, vol. 166, article 107336, 2022.

[22] M.-F. Tsai and H.-J. Tseng, "Enhancing the identification accuracy of deep learning object detection using natural language processing," *The Journal of Supercomputing*, vol. 77, no. 7, pp. 6676–6691, 2021.

[23] T. H. V. Phan and P. Do, "BERT+vnKG: using deep learning and knowledge graph to improve Vietnamese question answering system," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 7, 2020.

[24] A. H. A. Rahnama, M. Toloo, and N. J. Zaidenberg, "An LP-based hyperparameter optimization model for language modeling," *The Journal of Supercomputing*, vol. 74, no. 5, pp. 2151–2160, 2018.

[25] Y. Song, S. Tian, and L. Yu, "A method for identifying local drug names in Xinjiang based on BERT-BiLSTM-CRF," *Automatic Control and Computer Sciences*, vol. 54, no. 3, pp. 179–190, 2020.

[26] G. Yu, Y. Yang, X. Wang et al., "Adversarial active learning for the identification of medical concepts and annotation inconsistency," *Journal of Biomedical Informatics*, vol. 108, no. 2, article 103481, 2020.

[27] Y. Chen, C. Zhou, T. Li et al., "Named entity recognition from Chinese adverse drug event reports with lexical feature based BiLSTM-CRF and tri-training," *Journal of Biomedical Informatics*, vol. 96, article 103252, 2019.

[28] C. Lee, "LSTM-CRF models for named entity recognition," *IEICE Transactions on Information and Systems*, vol. 100, no. 4, pp. 882–887, 2017.

[29] J. Younes, E. Souissi, H. Achour, and A. Ferchichi, "Language resources for Maghrebi Arabic dialects' NLP: a survey," *Language Resources and Evaluation*, vol. 54, no. 4, pp. 1079–1142, 2020.

Research Article

Multisensor-Based Heavy Machine Faulty Identification Using Sparse Autoencoder-Based Feature Fusion and Deep Belief Network-Based Ensemble Learning

Yiqing Zhou [iD],[1] Jian Wang,[1] and Zeru Wang[2]

[1]*Computer Integrated Manufacturing System (CIMS) Research Center, College of Electronics and Information Engineering, Tongji University, Shanghai 201804, China*
[2]*CAD Research Center, College of Electronics and Information Engineering, Tongji University, Shanghai 201804, China*

Correspondence should be addressed to Yiqing Zhou; 1710334@tongji.edu.cn

Faulty identification plays a vital role in the area of prognostic and the health management (PHM) of the industrial equipment which offers great support to the maintenance strategy decision. Owing to the complexity of the machine internal component-system structure, the precise prediction of the heavy machine is hard to be obtained, thus full of uncertainty. Moreover, even for a single component, the feature representation of the acquired conditional monitoring signal can be different due to the different deployment of the sensor location and environmental inference, causing difficulty in feature selection and uncertainty in faulty identification. In order to improve the model identification reliability, a novel hybrid machine faulty identification approach based on sparse autoencoder- (SAE-) and deep belief network- (DBN-) based ensemble learning is proposed in this paper. First, six kinds of statistical features are extracted and normalized from multiple sensors monitoring the same target component. Second, the six extracted features are fused by the two-stage SAE proposed in this paper from the sensor dimension and feature dimension, respectively. The composite feature fused in the feature dimension is regarded as the comprehensive representation of the corresponding component. Finally, the fused features containing comprehensive representation of different components are utilized to predict the machine health condition by the ensemble of multiple deep belief classifiers. The effectiveness of the proposed method is validated by the two case studies of wind turbine gearbox and industrial port crane. The experimental result shows that the proposed ensemble learning approach outperforms other traditional deep learning approaches in terms of the prediction accuracy and the prediction stability when dealing with multisensor feature fusion and the precise faulty identification of the industrial heavy machine.

1. Introduction

With the development of the modern industrial manufacturing scale, the accurate faulty identification of the heavy industrial machine has been becoming increasingly important. The contemporary heavy machine has the characteristic of having highly complex internal component-system structure. Moreover, even for a certain component, multiple sensors are mounted on different locations to acquire complementary information. How to effectively utilize these multisensor information and raise the reliability of the machine faulty identification result remains a great challenge.

The conventional faulty identification approach is usually based on the historical conditional monitoring data which can be very useful for making the appropriate maintenance strategy to avoid catastrophic failure and save excessive maintenance costs. The traditional intelligent faulty identification procedure usually consists of three steps: feature extraction; feature fusion; and faulty identification.

1.1. Feature Extraction. During the feature extraction period, suitable statistical features are extracted and processed from the obtained sensor monitoring signal on the data space based on the expertise knowledge of signal processing to construct a suitable health indicator (HI) which can well

represent the machine health status and provide useful information for faulty identification. The construction approach of the HI can be categorized into two categories: single sensor-based approach and multisensor-based approach. The single sensor-based approach is totally based on the understanding of the extracted single source monitoring signal by selecting the appropriate statistical and signal processing analysis method. Gebraeel et al. [1] extracted the average amplitude and harmonic wave of the vibration signal as the health indicator to represent the bearing condition status. Malhi et al. [2] extract the root mean squared error and the peak value of the vibration signal to construct the health indicator of bearing status by using continuous wavelet transform. Kilundu et al. [3] construct the bearing HI indicator by using the singular spectrum analysis of the vibration signal. The single signal-based faulty identification approach can capture the faulty symptom of the target component with specific physical meaning; however, it can only capture the faulty symptom from a single data dimension with lower confidence level.

In order to raise the reliability of the faulty identification result, the multisignal-based approach is proposed which reflects the potential faulty symptom of the target component from multiple data dimensions [4–6]. In the multisignal-based approach, signals including forces, vibrations, temperatures, and acoustics are fused for the faulty identification task [7]. Compared to the single signal-based approach, the multisignal-based faulty identification approach can make the identification result more reliable [8–9]. Hao et al. [10] proposed a multisensor-based approach for the degradation identification of the mechanical component by evaluating the composite index which is combined with multiple sensor signals collected under multiple operational conditions. A motor faulty identification model based on sensor data fusion by using support vector machine and the short-term Fourier transform is proposed by Banerjee and Das [11].

1.2. Feature Fusion. The feature fusion is usually performed on the feature space where different kinds of statistical features extracted during the feature extraction phase are integrated into a composite feature by using technologies such as principal component analysis (PCA), kernel-based methods, and manifold learning. The feature fusion procedure is aimed at achieving considerable information compression and facilitating more effective feature. Xu et al. [12] fuse the time-domain and frequency-domain features extracted from the flame oscillation signal by using PCA. Wang et al. [13] propose the feature fusion and feature selection method from forces and vibration signals by using kernel PCA. Sun et al. [14] proposes a feature fusion method for handwritten recognition by using locally linear embedding.

1.3. Faulty Identification. During the faulty identification phase, the composite feature obtained from the feature fusion step is sent to a classifier for faulty identification [15–16]. Currently, various kinds of machine learning classifiers have been widely explored in the faulty classification stage; however, it is extremely hard for shallow-structured classifier to achieve precise faulty prediction for industrial
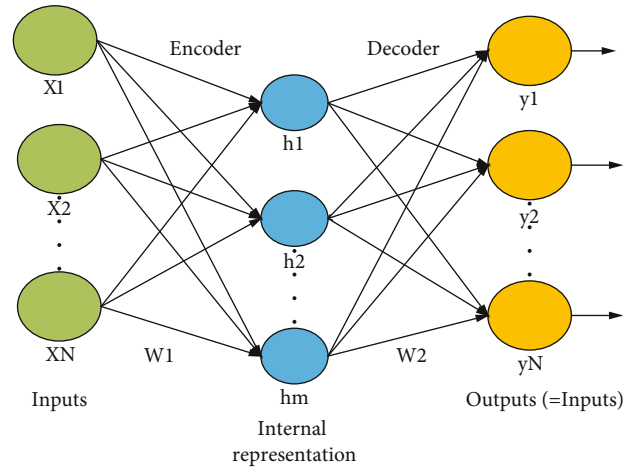
FIGURE 1: The detailed structure of autoencoder.
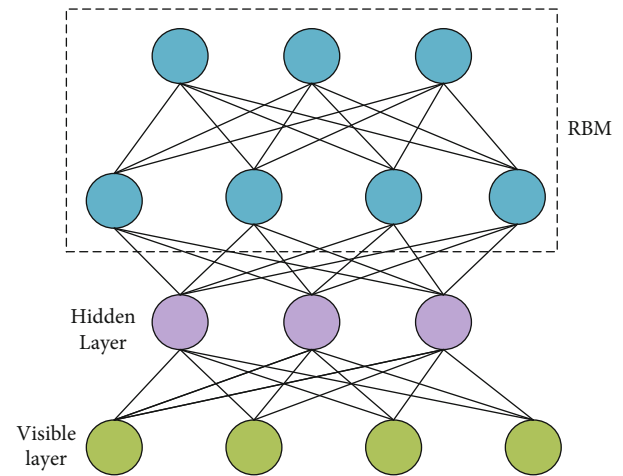
FIGURE 2: The structure of deep belief network.

heavy machine with complex internal component-system structure [17–19].

Recently, with the development of the deep learning technology, it has been widely used in feature extraction: feature fusion and faulty identification phase due to its powerful automatic feature learning and input-output mapping ability with its multi-layer structure. The deep learning-based faulty identification approach can release the dependence on the expertise knowledge of faulty identification, and it has been proved to be effective in improving the identification result [20–23]. However, there are three issues needed to be further considered.

(1) During the feature extraction period, most existing methods directly extract statistical features from multiple sensor signals. However, these sensors are mounted on different locations to acquire a variety of condition monitoring signals with different characteristics due to sensors' arrangement and environment inference. Therefore, the features extracted from these sensors are disordered and correlated
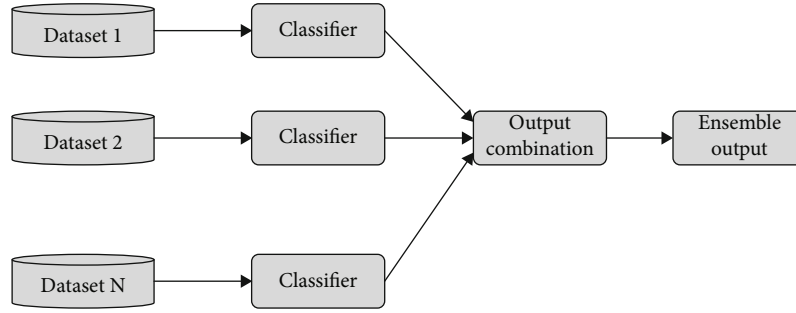
FIGURE 3: The scheme of homogeneous approach.

with different sources, increasing the difficulty of representative feature extraction. How to effectively merge these multisensor features to learn an over complete representation of these multisource information remains a challenge

(2) During the feature fusion period, conventional feature fusion methods such as PCA, NPE (neighborhood preserving embedding), and LLE (locally linear embedding) fail to learn high-level features due to the nonstationary characteristics of fault vibration signals. How to jointly utilize different kinds of extracted features to construct a more robust feature remains a challenge

(3) During the faulty identification period, most existing methods only use a single classifier as the final decision maker to output the final prediction result. Since the machine faulty reason can be related to multiple components due to its highly complex internal component-system structure, the prediction result of a single classifier can be full of uncertainly. Since different components can reflect machine faulty conditions from different aspects, how to construct a hierarchical faulty prediction approach reflecting the complex component-system relationship network remains a challenge

Dealing with the three issues listed above, this paper proposes a multisensor-based hierarchical machine faulty identification approach by using a two-stage sparse autoencoder-based feature fusion and the deep belief network-based ensemble learning. The innovation point of the proposed paper is listed as follows: Considering the first issue listed above, six kinds of statistical features are extracted and normalized from sensors monitoring the same component, and these features are sent to six different three-layer SAEs for feature extraction. The middle layers of six SAEs are extracted as the complete representation of the multisensor input signals. Considering the second issue listed above, the six extracted middle-layer features are concatenated and sent to a specially designed three-layer SAE for feature fusion to obtain a more robust feature. The six extracted middle-layer features are merged into a six-node compressed feature which is regarded as the composite health indicator of the target component. Considering the

third issue listed above, the composite health indicators obtained during the feature fusion period which reflects the comprehensive health statuses of different components are sent to multiple DBN classifiers to independently classify the machine faulty status, and the outputs of these classifiers are aggregated by using the Bayesian weighting strategy which represent the affiliated degree between the target component and the certain machine conditions. The rest of this paper is organized as follows: Section 2 briefly reviews the related research literature including the sparse autoencoder, the deep belief network, and the ensemble learning; Section 3 presents the framework of the proposed hybrid faulty prognostic approach; Section 4 describes the case study and the competition with other deep learning models; and finally, Section 5 summarizes the main contribution and future work of the proposed paper.

2. Methodology

2.1. Autoencoder and Sparse Autoencoder. The autoencoder is a kind of unsupervised feature learning approach which can learn the high-level representation of the raw input signal. The output layer of the autoencoder has the same dimension as the input layer which minimizes the reconstruction error between the input and the output so that the high-level features contained in the hidden layer can be obtained. The learning process of the autoencoder consists of two procedures: encoding and decoding. The encoding process acts as a feature extractor realizing the transformation from the raw input into the high-level feature, and the decoding process acts as a feature detector which reconstruct the input in the output layer based on the obtained high-level feature. The detailed structure of the autoencoder can be illustrated in Figure 1.

Assuming the n-dimensional raw input vector is $\{X1, X2, \cdots XN\}$, $Xi \in R^n$. During the encoding process, the raw input vector was transformed into the high-level feature represented in the following equations:

$$h_j = f\left(\sum_{i=1}^{N} w_{ij}^1 \times X_i + b_j^1\right), \tag{1}$$

$$f(z) = 1/(1 + \exp{(-z)}), \tag{2}$$
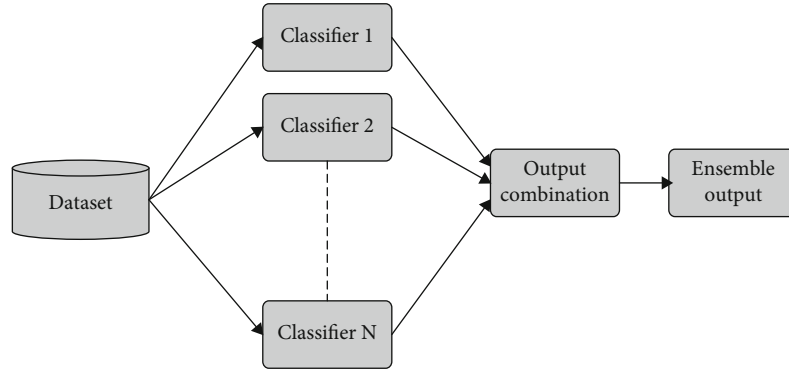
FIGURE 4: The scheme of heterogeneous approach.
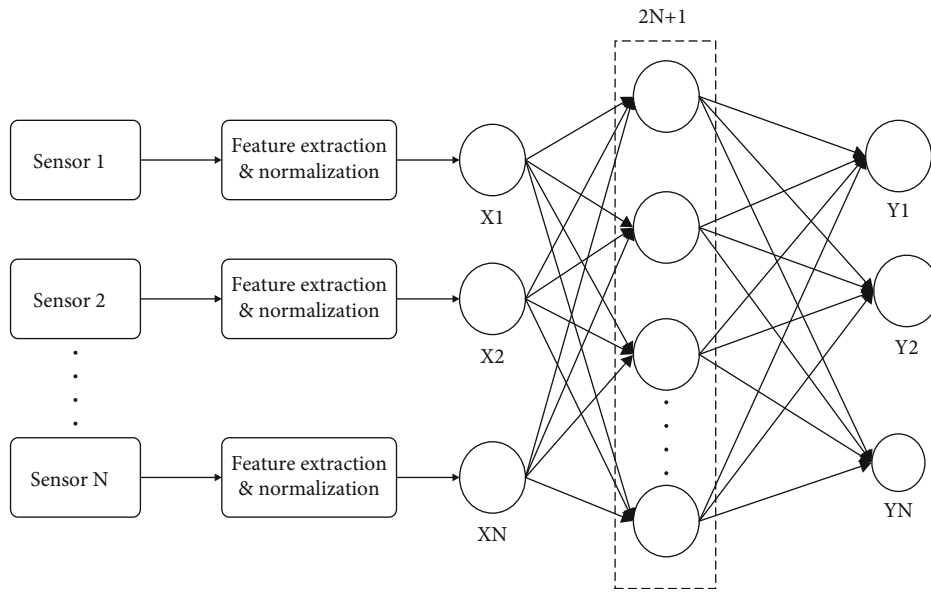
FIGURE 5: SAE fusion used for comprehensive feature representation.

where w_{ij}^1 denotes the weight between the ith dimension in the input layer and the jth dimension in the hidden layer. The parameter b_j^1 denotes the bias of the jth node in the hidden layer. The function $f(*)$ denotes the activation function which transforms the raw input vector into the high-level feature.

During the decoding process, the decoder transforms the hidden layer h_j into the output layer y, and the transform function can be illustrated in the following equation:

$$y_{k,b}(x) = f\left(\sum_{j=1}^{m} w_{jk}^2 \times h_j + b_k^2\right), \qquad (3)$$

where w_{jk}^2 denotes the jth dimension in the hidden layer and the kth dimension in the output layer. The parameter b_k^2 denotes the bias of the kth node in the output layer. In order to obtain the compressed high-level information of the raw
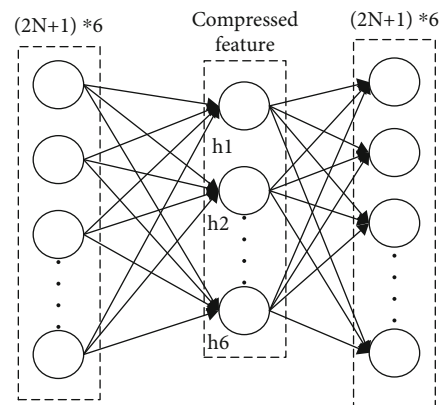
FIGURE 6: SAE fusion used for compressed feature representation.

input signal, the "sparse" restriction is applied on the hidden layer. The "sparse" restriction is aimed at making the statuses of the majority nodes at an inactive stage whose output
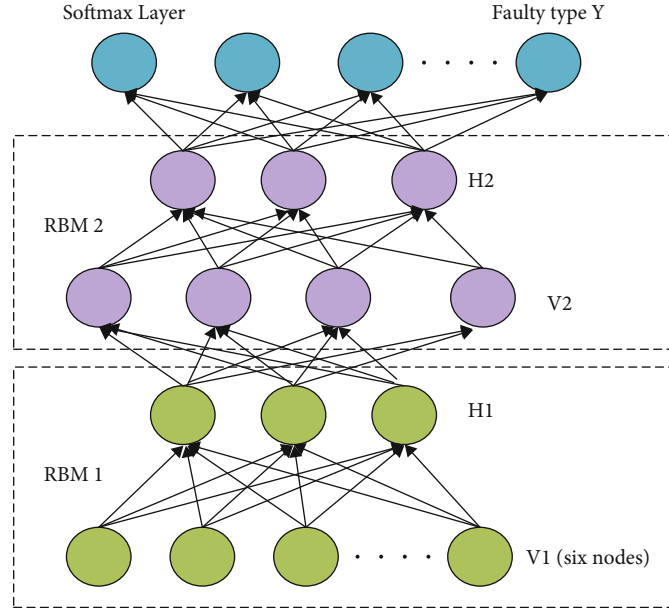
FIGURE 7: The detailed structure of the proposed DBNs.

is close to zero. The "sparse" restriction can be illustrated as shown in the following:

$$\widetilde{p}_j = \frac{1}{n} \times \sum_{i=1}^{n} \left[\alpha_j(x_i) \right]. \tag{4}$$

Here, the parameter $\widetilde{p}_j$ denotes the average activation value of the jth node of the hidden layer, and n denotes the number of the training samples. $\alpha_j(x_i)$ denotes the activation value of the jth hidden node of the ith sample.

During the training stage, the activated output of the hidden layer is restricted within the predefined sparse value as is illustrated in the following equation:

$$\sum_{j=1}^{n} KL\left(p \middle\| \widetilde{p}_j\right) = \sum_{j=1}^{n} p \times \log\left(\frac{p}{\widetilde{p}_j}\right) + (1-p) \times \log \frac{1-p}{1-\widetilde{p}_j}. \tag{5}$$

Here, the parameter p denotes the "sparse" parameter. The KL (Kullback-Leibler) divergence is used for the similarity evaluation between the "sparse" parameter p and the actual average activation $\tilde{p}$. Once the $\widetilde{p}_j = p$, the KL divergence equals zero, and the loss function of SAE can be represented as shown in the following equations:

$$J_{Sparse}(w, b) = J(w, b) + \beta \times \sum_{j=1}^{S_2} KL\left(p \middle\| \widetilde{p}_j\right), \tag{6}$$

$$
\begin{aligned}
J(w, b) = &\left[\frac{1}{n} \times \sum_{i=1}^{n} J\left(w, b\,;x^{(i)}, y^{(i)}\right) \right] \\
&+ \frac{\lambda}{2} \times \sum_{l=1}^{n_{l-1}} \sum_{i=1}^{S_l} \sum_{j=i}^{S_{l+1}} \left(w_{ji}^{(l)}\right)^2.
\end{aligned} \tag{7}
$$

Equation (6) denotes the loss function of the autoencoder, and Equation (7) denotes the "sparse" restriction. The sparse autoencoder is aimed at making the output as equivalent as possible to the input, obtaining the high-quality internal representation of the input vector. In this paper, the sparse autoencoder is used for the construction of the health indicator of individual component, containing the information of different sensor locations and feature characteristics.

2.2. Deep Belief Network. The deep belief network is a kind of unsupervised greedy layer-by-layer training algorithm which was proposed by Hinton in 2006 for solving the problem of structural optimization existed in the deep learning algorithm. The DBN is constructed by stacking multiple layers of restricted Boltzmann machines which is illustrated in Figure 2.

Each RBM consists of two layers with visible units and hidden units, respectively, and there is no connection between the same layers which only exists between the visible layer and the hidden layer. The training process of the DBN consists of two stages: pretraining the stacked RBM one by one in a greedy way and fine-tuning the whole network for obtaining an ideal performance. Since the DBN is composed of multiple stacked RBMs, it can be trained through contrastive divergence [24] in an unsupervised manner.

In recent years, the DBNs have been successfully used in faulty identification areas [25–27]. In this paper, multiple DBN classifiers are used independently for machine faulty classification from the aspect of different components.

2.3. Ensemble Learning. The ensemble learning is based on the notion that "two heads are better than one head." The performance of aggregating multiple classifiers has been
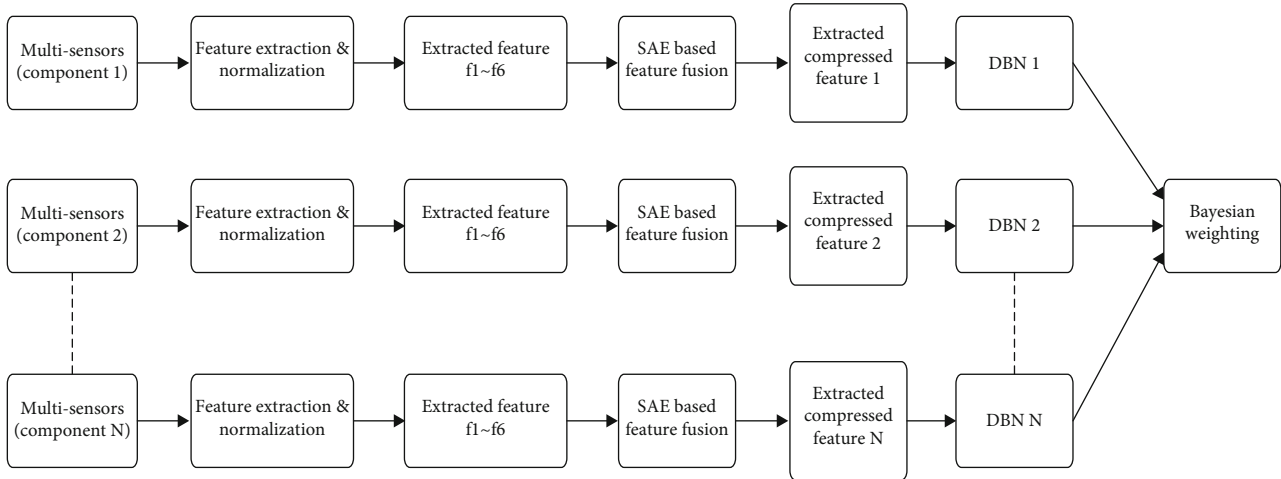
FIGURE 8: The overall framework of the proposed approach.

Input: Given the condition monitoring data collected from multiple sensors
Output: The machine faulty identification result
Step 1: Categorizing the multisensor data according to component
    The conditional monitoring data collected from the sensors monitoring the same
component with different sensing locations are categorized into the same group
Step 2: Feature extraction and feature fusion
    2.1 The categorized sensor data in step 1 is segmented into two categories, one for training and the other for testing
    2.2 The six time-domain features ($f1\sim f6$) illustrated in Equations (8)–(13) are extracted and normalized by using Equation (14)
    2.3 The six normalized features are sent to six SAEs, respectively, with the structure of Figure 5 for unsupervised training by using
Equations (1)–(7) and obtain the complete expression of the six normalized features, respectively
    2.4 The complete representation of the six normalized feature in Section 2.3 are concatenated and sent to the SAE with the struc-
ture of Figure 6 for unsupervised training by using Equations (1)–(7) and obtain the six-node compressed feature
    2.5 The six-node compressed features are extracted and regarded as the comprehensive representation of a certain target
component
Step 3: Machine faulty identification based on the proposed multi-DBN-based ensemble learning
    3.1 The constructed compressed comprehensive features with respect to multiple components are sent to multiple individual DBN
classifiers for predicting the machine health status by adding the Softmax prediction layer from the component perspective and the
individual classifiers are trained by using Equations (15)–(19)
    3.2 The Softmax outputs of the multiple DBN classifiers are aggregated based on the dynamically updated correlation weight
obtained by using the Bayesian weighting strategy as shown in Equations (20)–(22)
    3.3 Output the heavy machine condition result based on the maximum weighted output class
    3.4 Calculate the mean value of the correlation weight of each individual classifier during the training process
Step 4: Evaluate the proposed methodology
    Validate the proposed ensemble learning methodology in the testing datasets by using the mean value of the correlation weight
obtained during the training process

ALGORITHM 1: General procedure of the proposed methodology.

proved to be better than a single classifier in many fields [28–31]. By using the appropriate combination strategy, the ensemble learning can fully take advantage of each individual classifier so as to improve generalization [32]. The ensemble learning can be usually categorized into two types: sequential ensemble learning and parallel ensemble learning. In sequential ensemble learning, different algorithms are sequentially combined in the way that the first algorithm is used to generate a model, and then the second algorithm is used to correct the first model and so on which is also called boosting. In parallel ensemble learning, different algorithms are used for model training independently which is also

called bagging. There are two kinds of parallel ensemble learning namely homogeneous ensemble learning and heterogeneous ensemble learning [33].

2.3.1. Homogeneous Ensemble Learning. In the homogeneous ensemble learning scheme, the same types of classifier are used but with different training datasets. These datasets are collected from multiple data sources, and the data federation of multisources is achieved which is illustrated in Figure 3. This type of ensemble is also named as data variation ensemble, and it is mainly used for the management of industrial big data under large-scale scenarios.
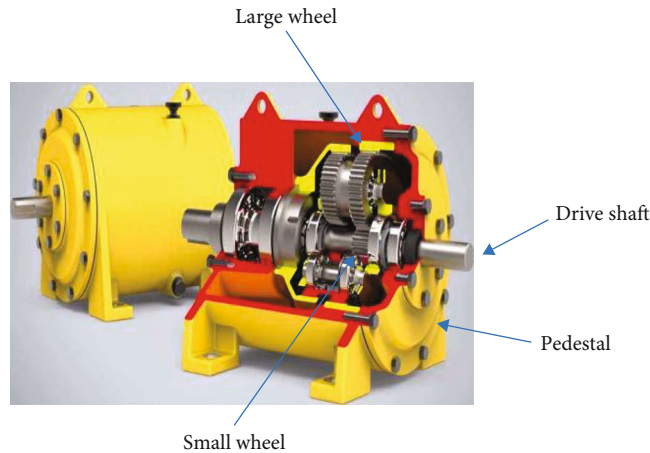
FIGURE 9: The component-system structure of the turbine gearbox.

TABLE 1: Description of datasets of Case Study I.

Datasets	Speed range (rpm)	Sample number in each condition	Condition status
		1080/1080	Healthy
		180/180	Rotor unbalance
Training/testing	800-1600/800-1600	180/180	Rotor misalignment
		180/180	Rotor friction
		180/180	Bearing loosing

2.3.2. Heterogeneous Ensemble Learning. In the heterogeneous ensemble learning scheme, a number of different types of classifiers but over the same training datasets are applied as can be seen in Figure 4, and the model diversity is achieved in this scheme. This type of ensemble is also called function variation ensemble which is mainly used for the generalization enhancement of model outputs.

In this paper, the homogeneous ensemble learning of several DBN classifiers are used for the construction of the hierarchical data processing framework of the component-system network.

3. Proposed Methodology

3.1. Data Preprocessing and Feature Fusion with SAE

3.1.1. Data Segmentation. The conditional monitoring data are obtained from multiple sensors monitoring the same target component but with the different installed locations. The obtained datasets were segmented into two categories: one for training and the other for testing.

3.1.2. Feature Extraction and Normalization. During the feature extraction and normalization period, the six time-domain features including the impulse $X_{Impulse}$, kurtosis $X_{Kurtosis}$, skewness $X_{Skewness}$, shape factor $X_{Shape\,factor}$, clearance factor $X_{Clearance\,Factor}$, and crest factor X_{Crest} are used

in this paper. Formulas for the six used features are presented as follows:

$$X_{Impulse} = \frac{\max\,(x(i))}{1/N\sum_{i=1}^{N}|x(i)|}, \tag{8}$$

$$X_{Kurtosis} = \frac{\sum_{i=1}^{N}(x(i) - \bar{x})^4}{N \times \left(\sqrt{\left(\sum_{i=1}^{n}(x(i))^2\right)/n}\right)^4}, \tag{9}$$

$$X_{Skewness} = \frac{1}{N} \times \frac{\sum_{i=1}^{N}(x(i) - \bar{x})^3}{\left(\sqrt{1/N \times \sum_{i=1}^{N}(x(i) - \bar{x})^2}\right)^3}, \tag{10}$$

$$X_{Shape\,factor} = \frac{\sqrt{1/N \times \sum_{i=1}^{N}(x(i))^2}}{1/N\sum_{i=1}^{N}|x(i)|}, \tag{11}$$

$$X_{Clearance\,Factor} = \frac{\max\,|x(i)|}{\left(1/N \times \sum_{i=1}^{N}|x(i)|^{1/2}\right)^2}, \tag{12}$$

$$X_{Crest} = \frac{\max\,(x(i))}{\sqrt{\left(\sum_{i=1}^{N}(x(i))^2\right)/N}}. \tag{13}$$

where $x(i)$ denotes a series of sensor signal and the parameter "N" refers to window length. The six extracted

Figure 10: Continued.

Kurtosis — Vertical sensor

Kurtosis — Horizontal sensor

Kurtosis — Lateral sensor

(c)

Skewness — Vertical sensor

Skewness — Horizontal sensor

Skewness — Lateral sensor

(d)
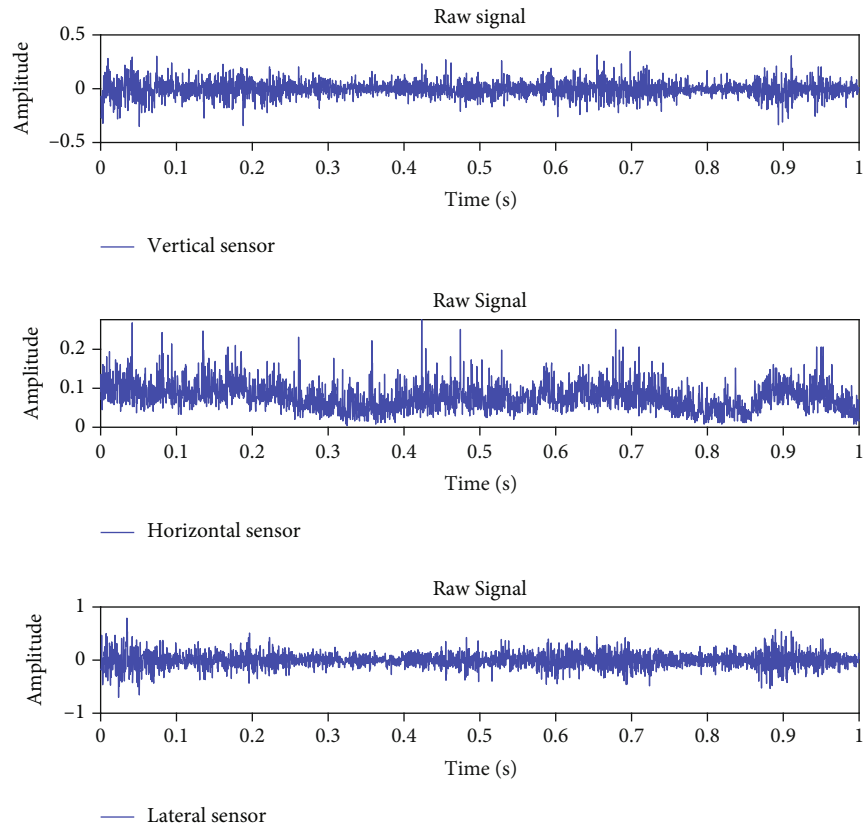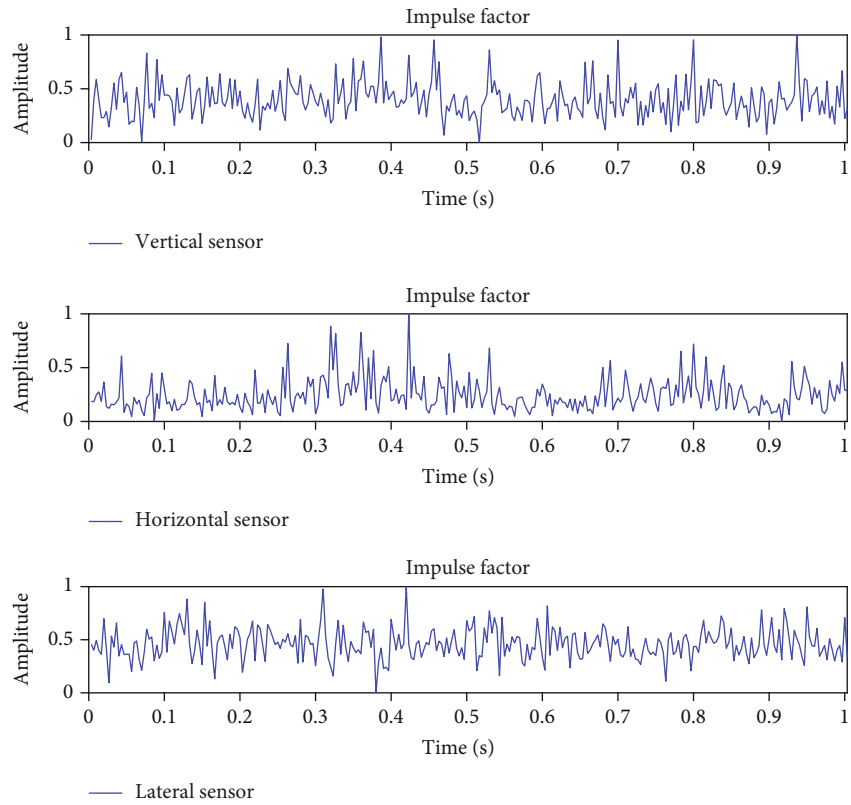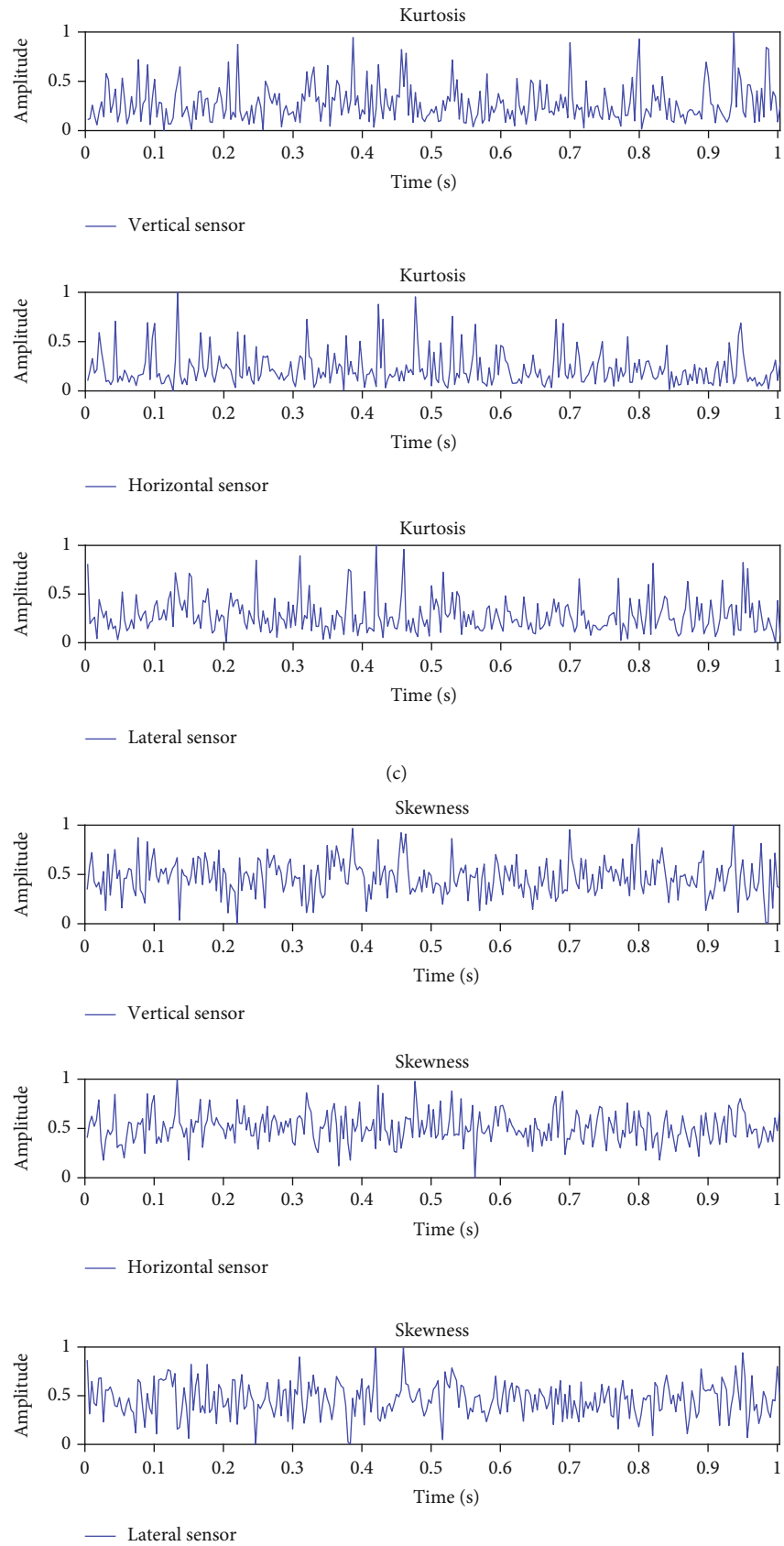
Figure 10: Continued.
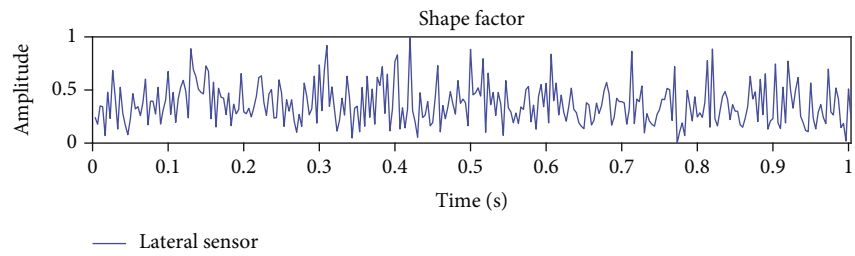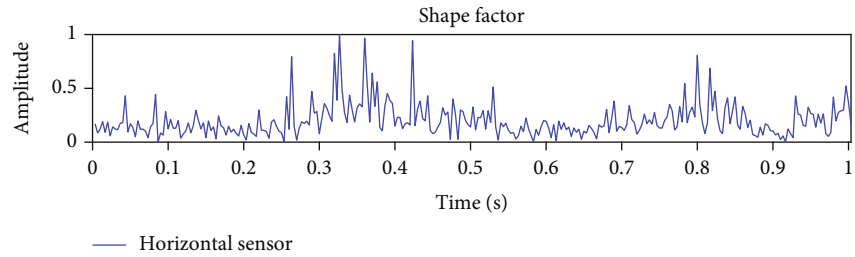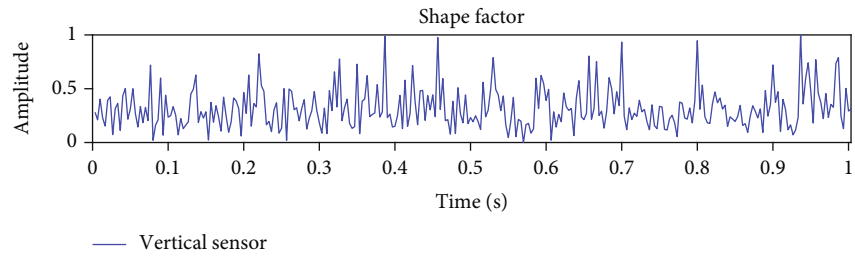
(e)
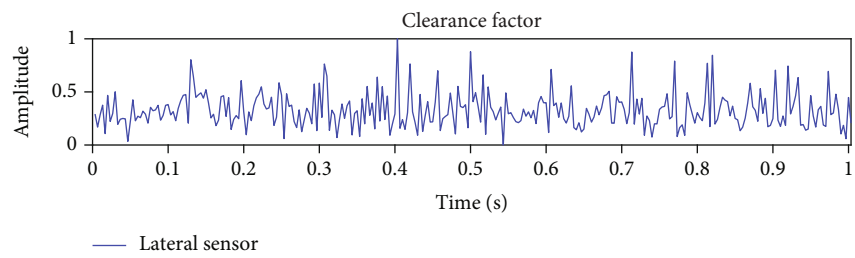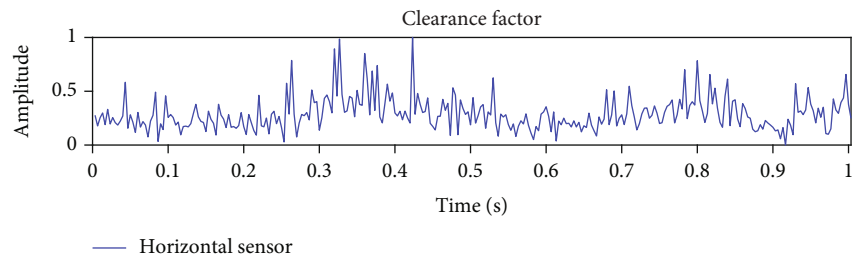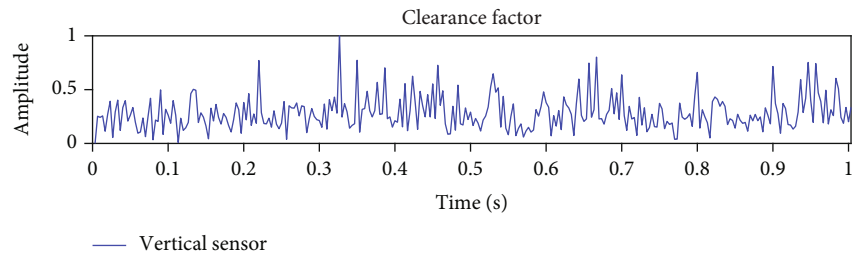
(f)

Figure 10: Continued.

(g)
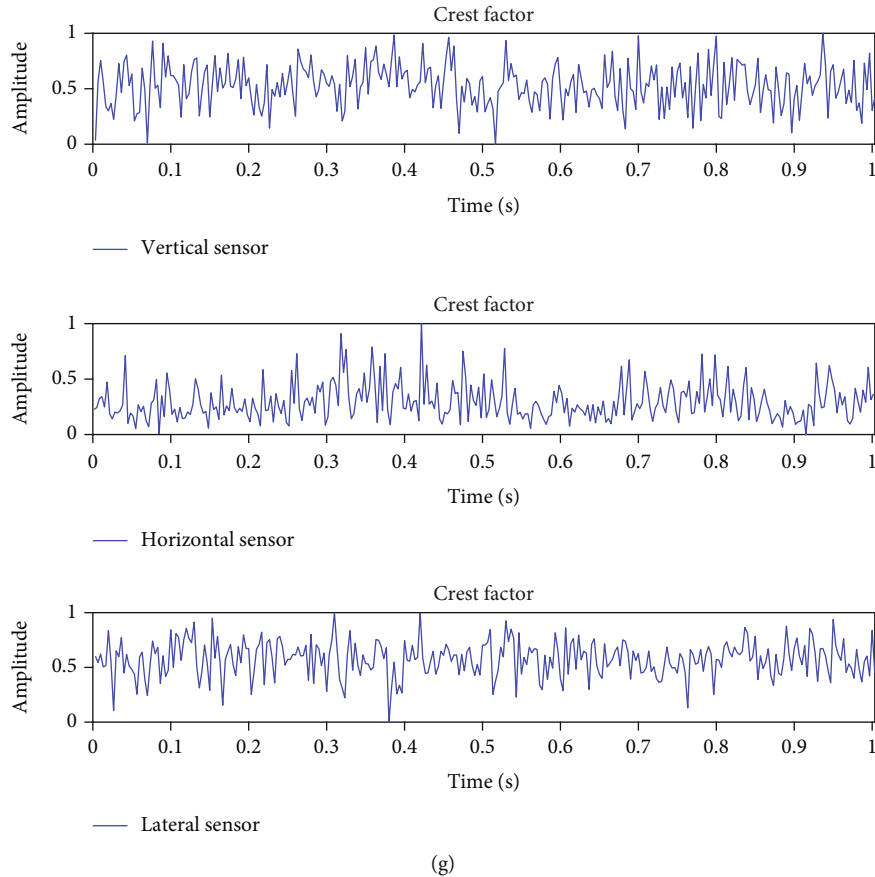
FIGURE 10: Six extracted normalized feature of turbine system of (a) raw signal, (b) impulse factor, (c) kurtosis, (d) skewness, (e) shape factor, (f) clearance factor, and (g) crest factor.

features are normalized by using Equation (14) where $x_{i,j}(t)$ denotes the extracted feature of the jth sensor within the ith signal segmentation. $\max(x_{i,j})$ and $\min(x_{i,j})$ represent the maximum feature value and the minimum feature value within the ith signal segmentation of the jth sensor.

$$\widehat{x}_{i,j}(t) = \frac{x_{i,j}(t) - \min(x_{i,j})}{\max(x_{i,j}) - \min(x_{i,j})}. \tag{14}$$

3.1.3. Feature Fusion and Feature Construction Based on SAE.

To construct the comprehensive health indicator of the target component, the proposed two-stage SAE-based feature fusion method is executed in this paper. First, six kinds of normalized features extracted from a group of sensors monitoring the same component are sent to six three-layer SAEs, respectively, with the structure as shown in Figure 5. Assuming that there are N sensors monitoring the same target component, the node number of the hidden layer is set as $2N + 1$ so that the layer can be forced to learn an over complete representation of each kind of feature obtained from multisensor signals [34]. Since there are six kinds of features extracted from the sensor signal

which is mentioned in Section 3.1.2, the six comprehensive features are concatenated and sent to the second SAE as shown in Figure 6. The node number of the hidden layer of the second SAE is set to 6 according to the parameter expertise adjustment in literature [34] so that the network is forced to learn a highly compressed representation of the six extracted features. The six-node compressed feature is extracted and regarded as the composite health indicator of the target component.

3.2. Ensemble Learning Based on DBNs

3.2.1. Individual Faulty Classification Based on Single DBN.

The extracted six-node compressed features are fed to multiple DBN models which are used as the subclassifier of the ensemble learning network. In this paper, the DBN is constructed by stacking two RBM layers and one softmax layer for faulty classification with the structure as shown in Figure 7, where "V" denotes visible vector and "h" denotes the hidden vector. The parameter "W" denotes the connection weight between the visible layer and the hidden layer.

The input dimension of the proposed DBN is set to six, and the output dimension of the Softmax layer is equivalent to the faulty type of the corresponding faulty

(a)

(b)

(c)

(d)
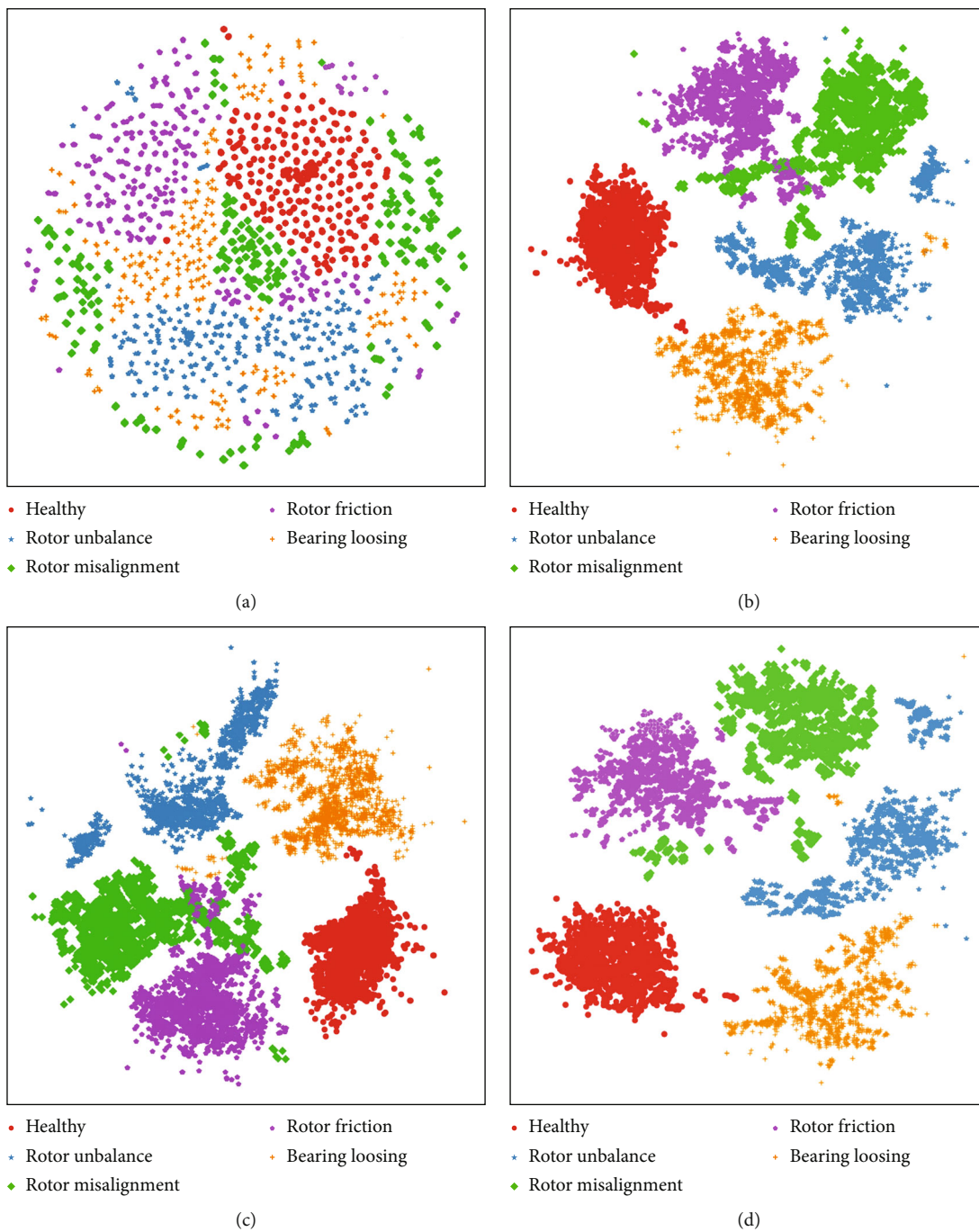
FIGURE 11: Continued.

(e)

(f)

(g)

(h)
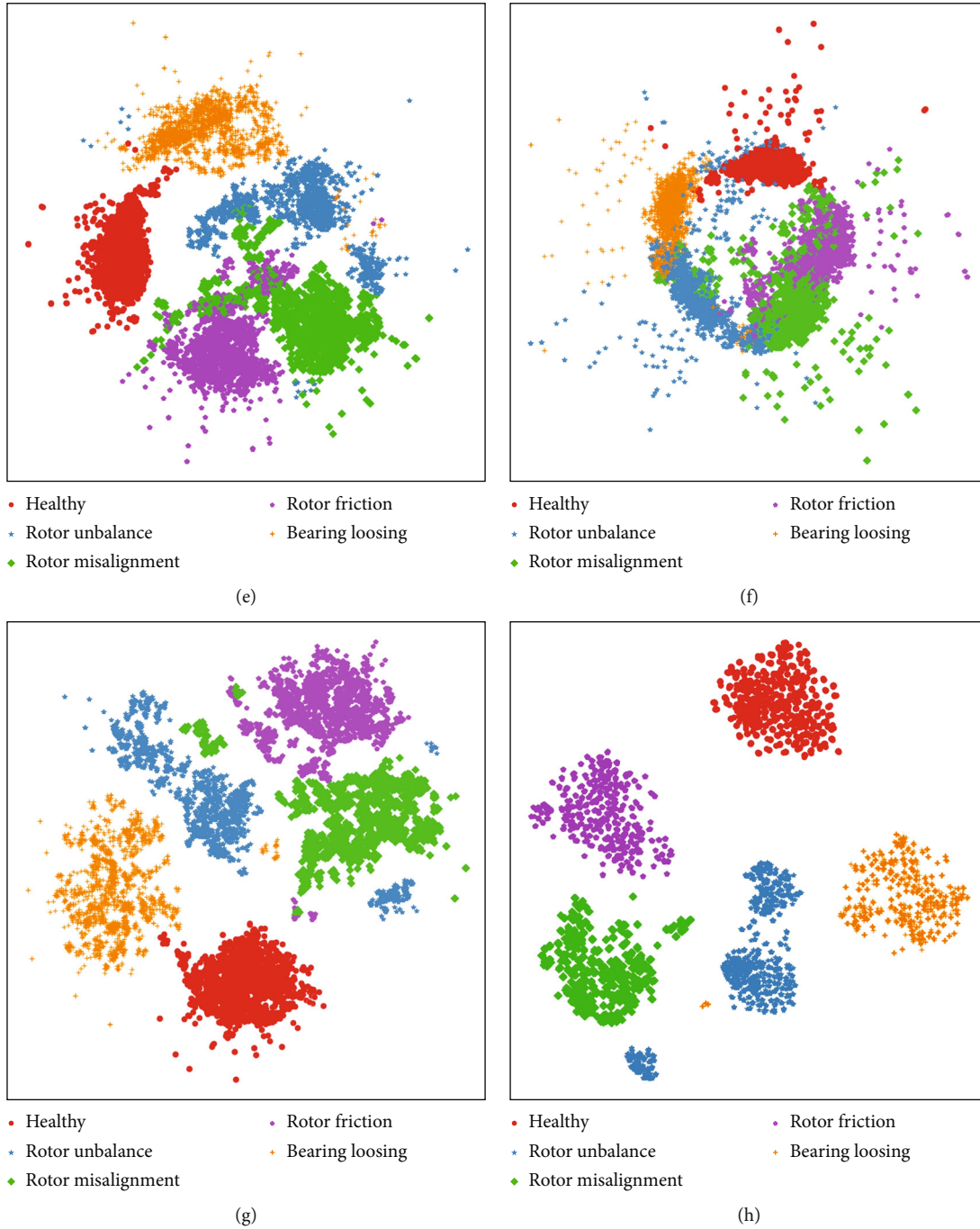
FIGURE 11: Feature visualization by t-SNE. (a) Raw sensor data. (b) Impulse+SAE extraction. (c) Kurtosis+SAE extraction. (d)Skewness+ SAE extraction. (e) Shape factor+SAE extraction. (f) Clearance factor+SAE extraction. (g) Crest factor+SAE extraction. (h) Six-node compressed feature.

classification task. The contrastive divergence proposed by Hinton et al. [24] is used for the fast training of the DBN model. Firstly, the conditional probability of the hidden units contained in RBM1 is obtained by using Equation (15). The Gibbs sampling is then employed to determine the state of the hidden units. Finally, the state of the visible units can be obtained through Equation (16), which can be regarded as the reconstruction of the former hidden layer. The param-

eter $w_{i,j}$ denotes the connection weight between the ith visible node and the jth hidden node. The parameters α and β denote the biases of the visible and hidden layer, respectively.

$$p(h_i = 1 | v; \theta) = 1 / \left[1 + \exp \left[-\alpha_i - \sum_{j=1}^{m} V_j \times w_{i,j} \right] \right], \quad (15)$$

TABLE 2: The configuration of the SAE used for feature extraction and feature fusion.

SAE for feature extraction		SAE for feature fusion	
Number of hidden layer	1	Number of hidden layer	1
Node per layer	3-7-3	Node per layer	42-6-42
Sparsity penalty	0.02	Sparsity penalty	0.02
Number of epochs	60	Number of epochs	60
Batch size	100	Batch size	100
Weight decay	0.0001	Weight decay	0.0001

$$p\left(v_j = 1 | h ; \theta\right) = 1 / \left[1 + \exp \left[-\beta_j - \sum_{i=1}^{h \text{ dimension}} h_i \times w_{i,j} \right] \right]. \quad (16)$$

The gradient variation can be obtained through maximizing the log-likelihood by performing stochastic gradient descent. The updating regulations of the parameter θ can be achieved by using the following equations:

$$\Delta W_{ij} = \sigma \times \left(\langle v_j \times h_i \rangle_{p(h|v)} - \langle v_j \times h_i \rangle_{\text{rec}} \right), \quad (17)$$

$$\Delta \beta_j = \sigma \times \left(\langle v_j \rangle_{p(h|v)} - \langle v_j \rangle_{\text{rec}} \right), \quad (18)$$

$$\Delta \alpha_i = \sigma \times \left(\langle h_i \rangle_{p(h|v)} - \langle h_i \rangle_{\text{rec}} \right). \quad (19)$$

Here, σ denotes the learning rate, $\langle * \rangle_{p(h|v)}$ denotes the expectation value of the conditional distribution $p(h|v)$, and $\langle * \rangle_{\text{rec}}$ denotes the distribution expectation of the reconstructed model. The stacked RBMs are first pretrained in an unsupervised way. After the DBN pretraining process is finished, the DBN model is fine-tuned starting from the last layer by using the labeled data. The softmax layer of representing the probability of the different faulty types are added for a classification problem, and the entire model parameters are optimized by using the backpropagation algorithm.

3.2.2. Ensemble Learning Based on Bayesian Weighting. Since the individual DBN classifier only represents the machine health status from the component perspective, in order to construct the correlation network between individual components and different machine faulty conditions, the ensemble learning approach based on Bayesian weighting strategy is applied for integrating the outputs result of multiple DBN classifiers. Assuming that there are N individual DBN classifiers corresponding to the health statuses of N individual components of the equipment, the machine health status can be represented as follows.

$$P_{\text{machine}}(S = S_i) = \sum_{j=1}^{N} \text{Corr}\left(\text{component}_j\right) \times P_{\text{component}_j}\left(S = S_i | \text{DBN}_j\right). \quad (20)$$

Here, the $\text{Corr}(\text{component}_j)$ denotes the correlation weight between the machine and the target component

TABLE 3: The configuration of individual DBN classifier.

Number of RBM	2
Node per layer	6-10-10-10-5
Learning rate	0.1
Number of epochs	60
Batch size	100
Weight decay	0.0002

when identifying the state S_i, and it is initialized with the weight value of $1/N$, representing that all the components have the same correlation relationship with the machine at the beginning;

$P_{\text{component}_j}(S = S_i | \text{DBN}_i)$ denotes the output of the softmax probability of the state S_i by the jth DBN which corresponds to the jth component. The specific mathematical expression is illustrated as shown in Equation (21), where the parameter K denotes the number of the faulty types of the classification task and the Z_{S_i} denotes the raw output of the state S_i

$$P_{\text{component}_j}\left(S = S_i | \text{DBN}_j\right) = \frac{e^{z_{S_i}}}{\sum_{z=1}^{k} e^z}. \quad (21)$$

If the true state of the of the machine state is exactly S_i, the postier correlation weight can be updated as shown in Equation (22); otherwise, the weight will not be updated. It should be noted that the updated weight assignment used here is only corresponding to the classification type of the state S_i, and the different machine faulty states have different component-system weight assignment.

$$\text{Corr}\left(\text{component}_j\right)_{\text{postier}} = \frac{P_{\text{component}_j}\left(S = S_i | \text{DBN}_j\right)}{\sum_{j=1}^{N} P_{\text{component}_j}\left(S = S_i | \text{DBN}_j\right)}. \quad (22)$$

3.3. The Combination of the SAE and the DBN-Based Ensemble Learning. The overall framework of the proposed approach is illustrated in Figure 8. First, the sensors monitoring the same component are categorized, and six classical time-domain features are extracted and normalized from these sensor data. Second, the SAE-based feature fusion is used for obtaining the compressed representation of the six normalized features which is also regarded as the
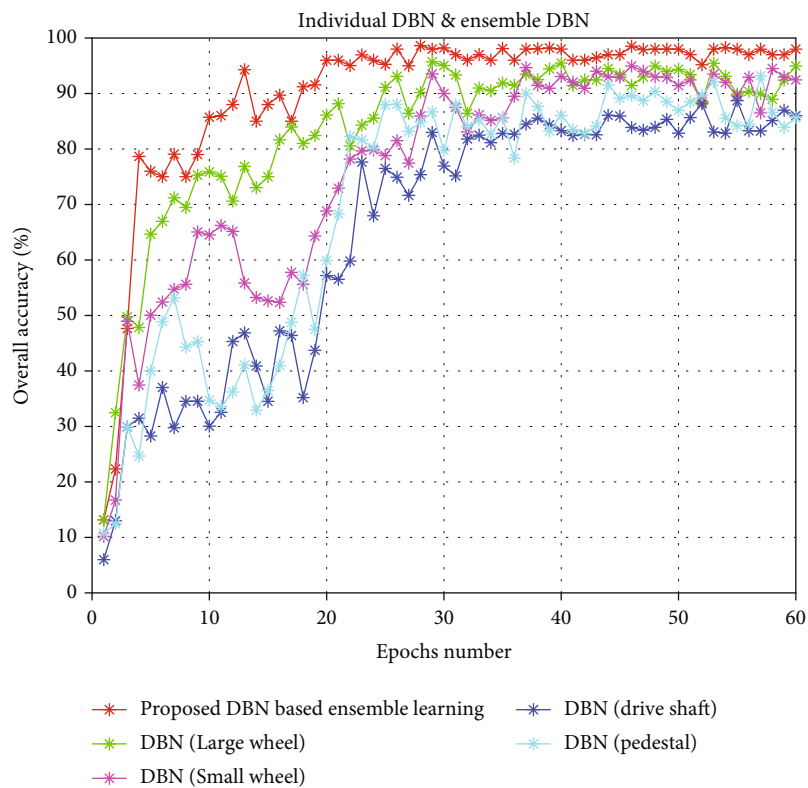
(a)
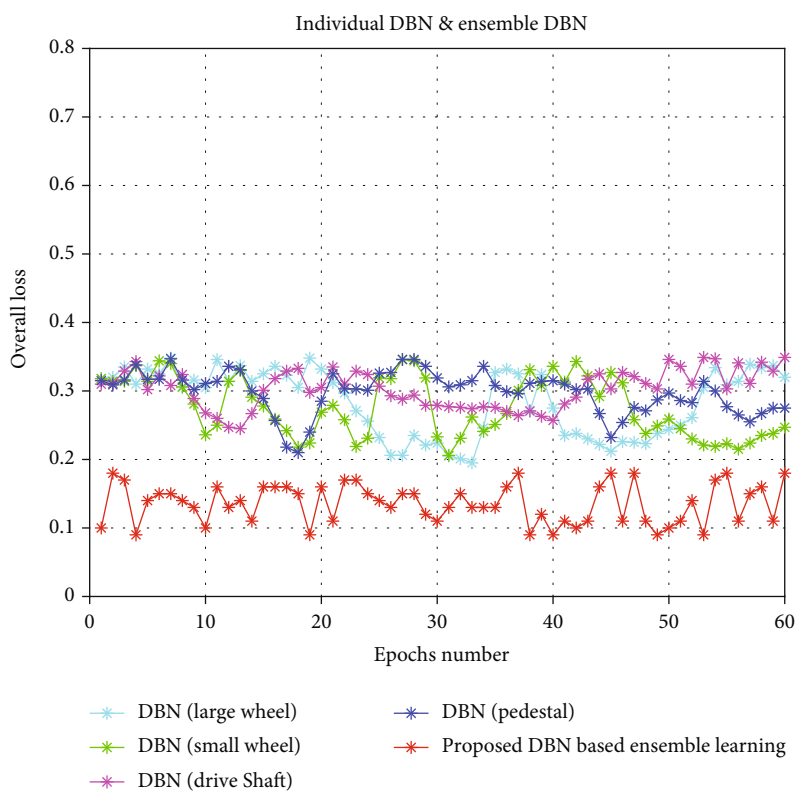
(b)

Figure 12: The performance of single individual classifier and the proposed DBN-based ensemble learning. (a) Testing accuracy curves. (b) Testing loss curves.

TABLE 4: The average accuracy of each condition of individual classifier and the proposed ensemble learning.

Classifier	Healthy	Rotor unbalance	Rotor misalignment	Rotor friction	Bearing loosing	Total accuracy	Variance
Proposed ensemble learning	97.33%	95.45%	96.42%	96.57%	95.40%	96.78%	0.66
DBN (large wheel)	93.01%	83.29%	84.17%	87.64%	92.87%	90.6%	15.98
DBN (small wheel)	91.23%	86.11%	88.92%	84.09%	92.02%	89.8%	11.29
DBN (drive shaft)	84.24%	82.19%	81.94%	80.35%	87.38%	83.73%	7.32
DBN (pedestal)	87.22%	84.50%	82.99%	85.68%	86.44%	86.29%	2.77

TABLE 5: The comparison of different prediction model.

Input	Classifier	Total accuracy	Variance
Six-node compressed feature	Proposed DBN-based ensemble learning (Bayesian weighting)	96.78%	0.66
	DBN-based ensemble learning (winner takes all)	95.32%	0.87
Six-node compressed feature (concatenated)	Single CNN	91.33%	2.06
	Single DBN	90.87%	2.35
	Single BPNN	91.64%	2.51
	Single SVM	91.03%	1.93

TABLE 6: The comparison of different model input.

Classifier	Input	Total accuracy
Proposed DBN-based ensemble learning (Bayesian weighting)	Six-node compressed feature	96.78%
	Impulse+SAE extraction	84.31%
	Kurtosis+SAE extraction	86.33%
	Skewness+SAE extraction	82.46%
	Shape factor+SAE extraction	83.92%
	Clearance factor+SAE extraction	85.37%
	Crest factor+SAE extraction	86.87%
	Raw sensor data without SAE extraction	55.24%

comprehensive representation of the target component. Third, the extracted compressed features representing the statuses of different components are sent to multiple individual DBN classifiers for the machine faulty identification from the component perspective. Finally, the multiple outputs of DBNs are aggregated through the Bayesian weighting strategy, and the correlation degree between the individual component and the whole machine system is constructed. The correlation weight is constantly updated based on the posterior true label during the training process. During the testing process, the ensemble weighting strategy applied on different DBN classifiers is defined based on the mean value of the correlation weight obtained during the training process. The general procedure of the proposed methodology can be illustrated in Algorithm 1.

4. Methodology Evaluation

4.1. Case Study I: Equipment Faulty Identification for Wind Turbine Gearbox

4.1.1. Data Description and Experimental Set up. The performance of the proposed approach is evaluated on the

FIGURE 13: The component-system structure of the crane trolley.

TABLE 7: The sensor deployment of crane trolley.

Subsystem	Monitoring sensor number
Pinion bearing	PB-01; PB-02; and PB-03
Drum shaft	DS-01; DS-02
Retarder input shaft	RIS-01; RIS-02
Retarder output shaft	ROS-01; ROS-02; and ROS-03

TABLE 8: Description of datasets of Case Study II.

Datasets	Speed range (rpm)	Sample number in each condition	Condition status
Training/testing	75-100/75-100	750/750	Healthy
		150/150	Wheel biting
		150/150	Wire rope over winding
		150/150	Brake failure

multisensor datasets of the wind turbine gearbox provided by the enterprise of Shanghai electric group. As shown in Figure 9, the wind turbine gearbox mainly consists of four main components namely large wheel, small wheel, drive shaft, and pedestal. Each component is monitored by three vibration sensors from the vertical, horizontal, and lateral directions. The sampling frequency is set as 36 kHz, and the sampling time is set as 1 s under the load speed ranging from 800 rpm to 1600 rpm. The time window is set as 10, where each sample contains 10 sampling points. Therefore, there are 3600 samples in total, in which half are for training and half are for testing. There are five kinds of gearbox conditions namely healthy, rotor unbalance, rotor misalignment, rotor friction, and bearing loosing. The specific details of the datasets of Case Study I are illustrated as shown in Table 1.

4.1.2. Feature Extraction and Normalization. In this paper, six kinds of features including the impulse factor, kurtosis, skewness, shape factor, clearance factor, and crest factor are extracted and normalized from each sample collected by sensors monitoring the same subsystem. Take the component of "driving shaft" as an example; the raw signal and the six kinds of normalized features are extracted from three vibration sensors from vertical, horizontal, and lateral directions, respectively, as shown in Figures 10(a)–10(g). It can be found that in Figure 10 that the amplitude of the raw signals and all the six extracted and normalized features of impulse, kurtosis, skewness, shape factor, clearance factor, and crest factor varies obviously due to the different monitoring locations of sensors which may increase not only the difficulty of feature selection but also the uncertainty in faulty identification. Therefore, it is necessary to find an effective method to merge the information collected from the three vibration sensors to obtain the comprehensive representative status of the target component, and in this, paper the proposed two-stage sparse autoencoder is introduced owning to its powerful feature compression and feature reconstruction ability.

4.1.3. The Feature Fusion. In order to demonstrate the robustness of the fused feature and to visualize the feature representation ability, the raw sensor data without SAE extraction, the six statistical features being further extracted by SAE and the six-node compressed composite feature fused by the proposed two-stage SAE model are represented on a 2-D feature map as shown in Figures 11(a)–11(g) by using the t-SNE (t-distributed stochastic neighbor embedding) technology. As shown in Figure 11(a), the t-SNE fails to separate the five turbine gearbox conditions with raw sen-

sor data of the multisensor. Most of the samples are mixed with each other, and it can only be distinguishable between the healthy and unhealthy state which may greatly influence the accuracy of the faulty identification model. In Figures 11(b)–11(g), most of the samples are correctly classified by the six extracted and normalized features fused by one stage SAE with less samples being mistakenly classified, where only a few marginal samples of rotor unbalance, rotor friction, and rotor misalignment are mistakenly mixed with each other, whereas in Figure 11(h), all the five conditions are perfectly separated clearly compared with the raw sensor data illustrated in Figure 11(a) and the six extracted and normalized features fused by one stage SAE illustrated from Figures 11(b)–11(g). Therefore, the proposed six-node compressed feature can greatly improve the data clustering ability because the compressed feature contains the compressed information of the six features rather than a single feature, which can be more effective when dealing with the nonstationary characteristics of faulty signals.

The configuration of the SAEs used for feature extraction and feature fusion is illustrated as shown in Table 2.

4.1.4. DBN-Based Ensemble Learning. The six-node compressed feature indicating the comprehensive representation of the four subsystems are extracted and used as the input of five individual DBN classifiers. The configuration of a single individual DBN classifier is illustrated as shown in Table 3. The DBN used as the individual classifier has five layers in total, where the node number of the input layer node is set as 6, which has the same dimension of the six-node compressed feature. The node number of the output layer is set as 5 according to the type number of the turbine gearbox health conditions. The overall testing accuracy curves and overall loss curves of each individual DBN and the proposed ensemble learning networks are illustrated as shown in Figures 12(a) and 12(b). It can be found that the proposed ensemble learning approach outperforms other five individual DBN classifiers in terms of the overall testing accuracy and the overall testing loss.

The experiment is repeated 10 times, and the average accuracy in predicting each kind of conditional status and the total accuracy is illustrated as shown in Table 4. It can be found that the proposed DBN-based ensemble learning approach achieves the highest average prediction accuracy in terms of predicting each kind of health condition and the total accuracy. The "variance" is used as the metric to evaluate stability of the faulty prediction model. From the variance evaluation of each classifier, it can be found that the performance of the proposed ensemble learning approach remains stable in predicting each kind of
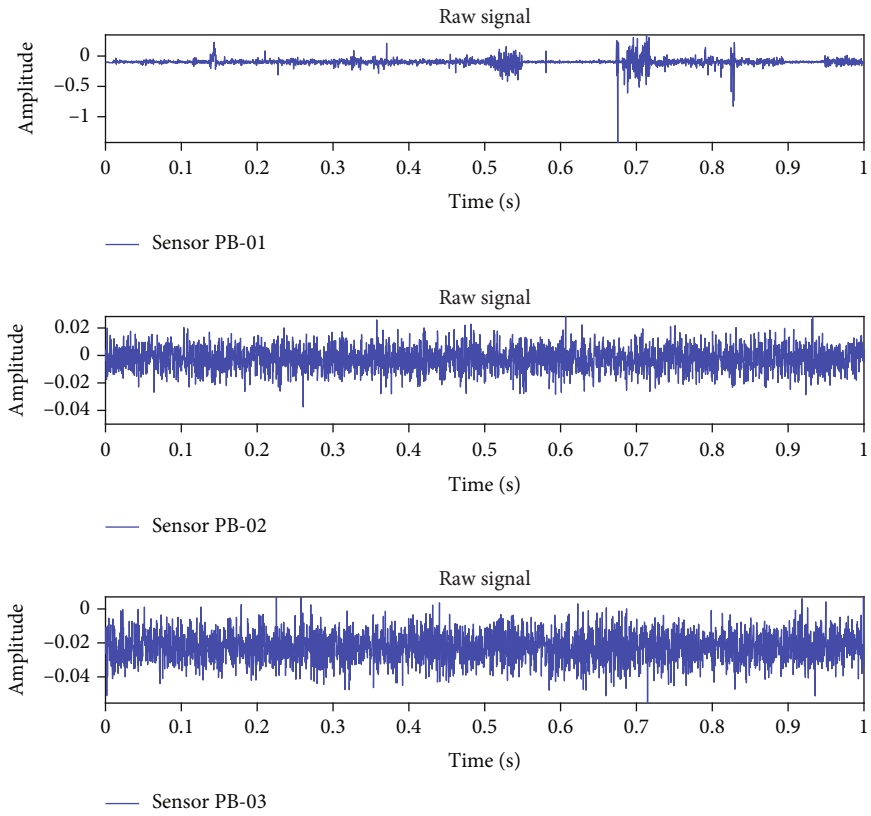
(a)

(b)

Figure 14: Continued.

(c)

(d)

FIGURE 14: Continued.

Sensor PB-01

Sensor PB-02

Sensor PB-03

(e)

Sensor PB-01

Sensor PB-02

Sensor PB-03
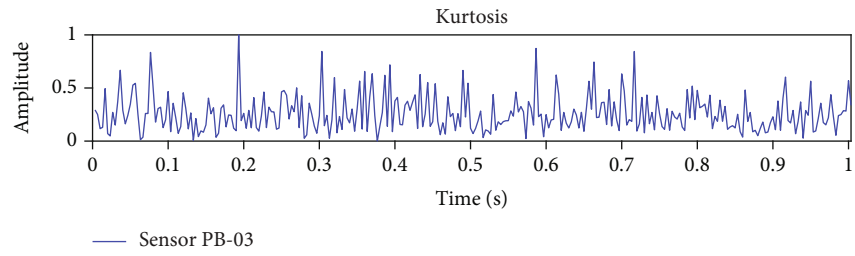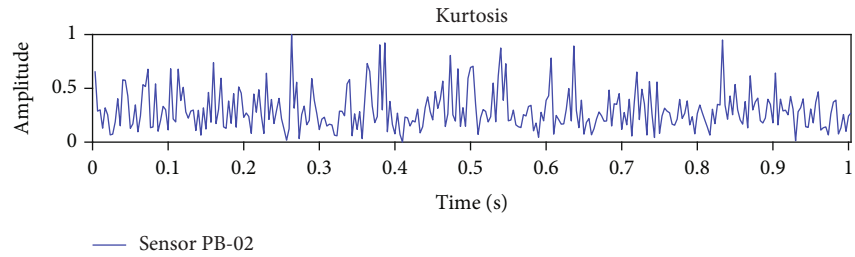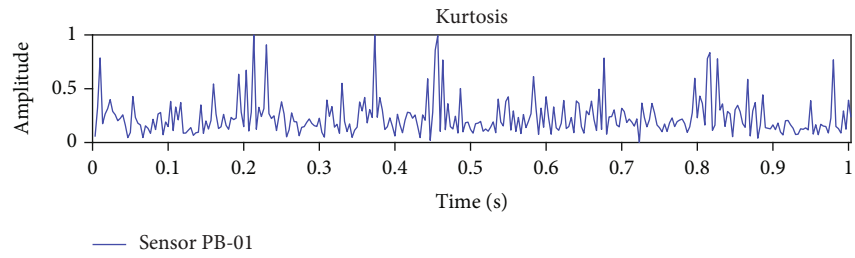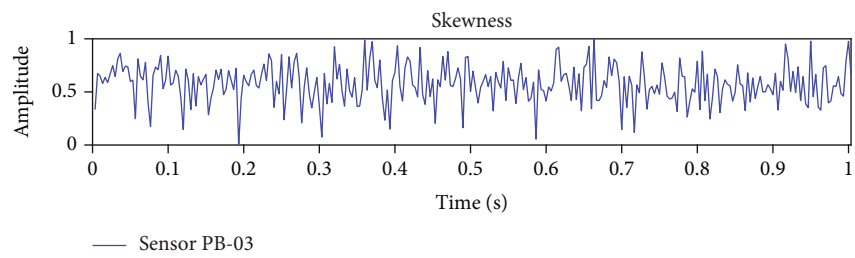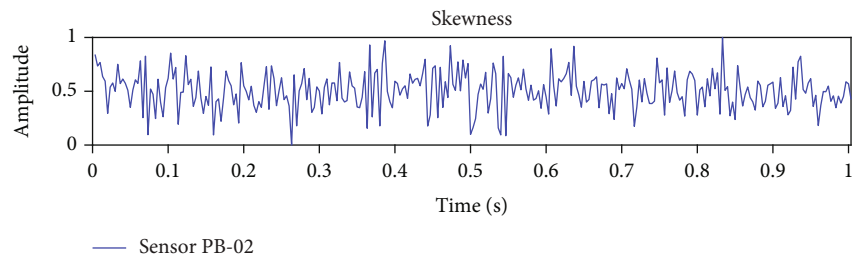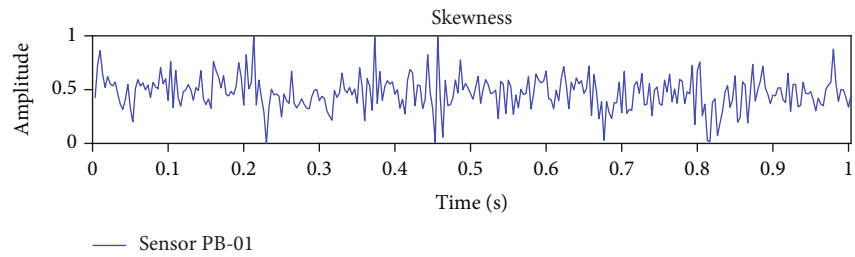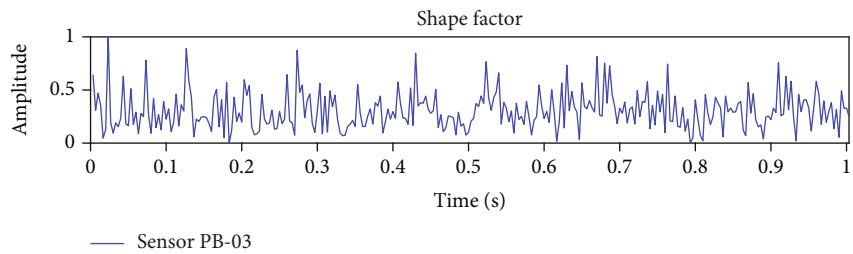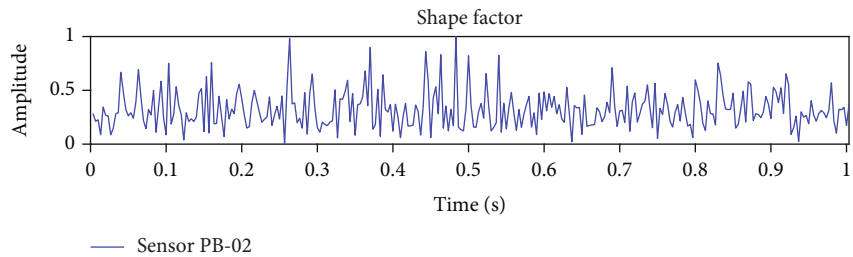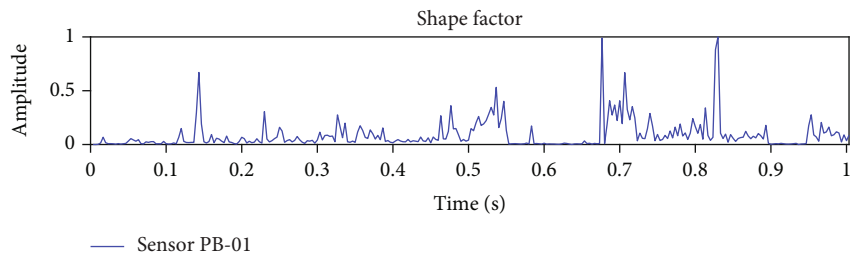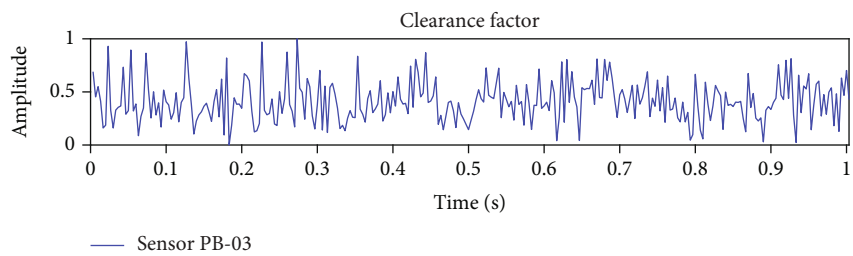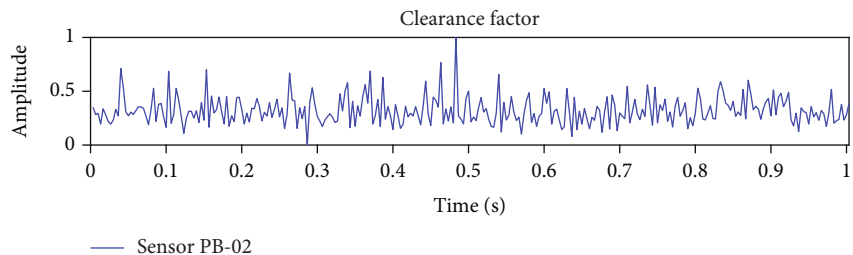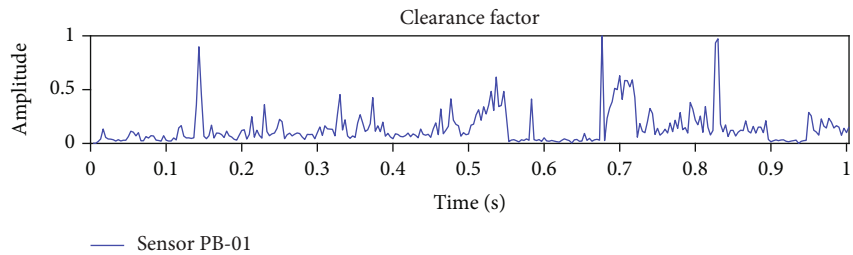
(f)

Figure 14: Continued.

(g)
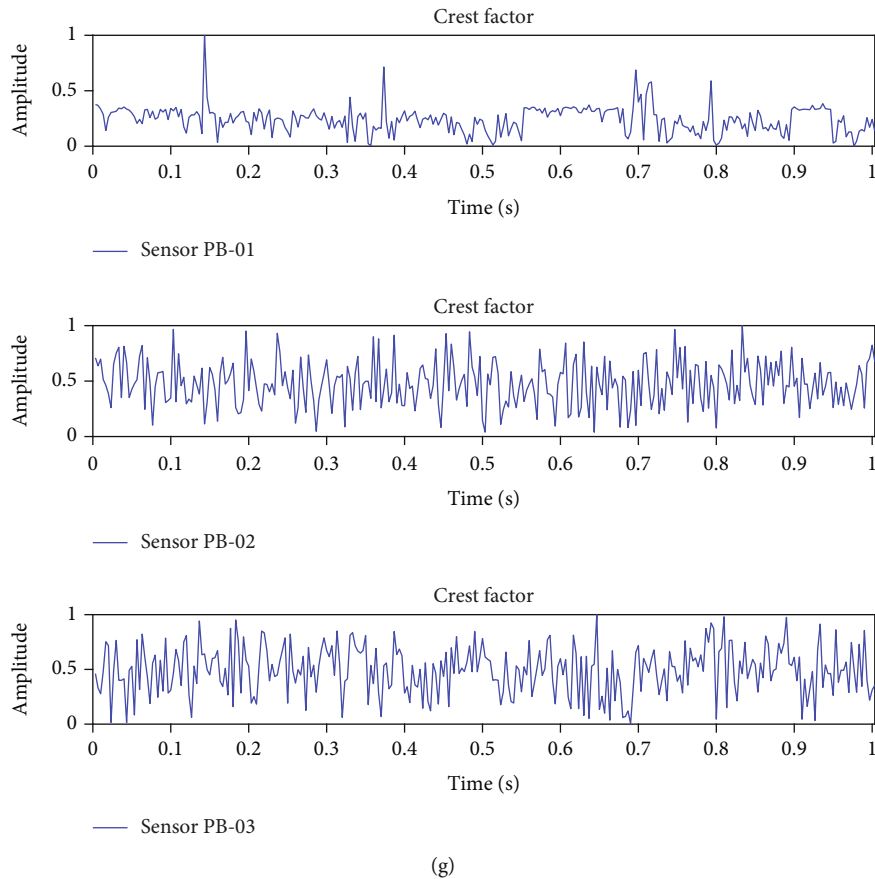
FIGURE 14: Six extracted normalized features of pinion bearing of (a) raw signal, (b) impulse factor, (c) kurtosis, (d) skewness, (e) shape factor, (f) clearance factor, and (g) crest factor.

conditional status while the performance of individual DBN classifier varies heavily in terms of the prediction of certain faulty type. The reason should be that different components have different sensitive degree to different machine faulty conditions, indicating the necessity of constructing a component-system affiliation network.

4.1.5. Comparison with Other Methods. In order to evaluate the effectiveness of the proposed two-stage SAE-based feature fusion and the DBN-based ensemble learning, the comparison experiments are carried out 10 times from the following two aspects, respectively.

(1) Proving the Effectiveness of the Proposed DBN-Based Ensemble Learning Classifier. In order to evaluate the proposed DBN-based ensemble learning classifier, we fix the model input as the six-node composite feature. The classifiers including the proposed DBN-based ensemble learning classifier based on Bayesian weighting, DBN-based ensemble learning classifier on winner takes all, single CNN, single DBN, single BPNN, and single SVM classifiers are used for comparison. It should be noted that when using single classifier, the six-node compressed features of different components are concatenated as a categorical input. The simulation result is illustrated as shown in Table 5. It can be found that the proposed DBN-based ensemble learning approach

based on Bayesian weighting strategy achieves the highest total accuracy of 96.78% while ensemble learning based on "winner takes all" strategy achieves the total accuracy of 95.32%. The total accuracy obtained by the rest four single classifiers are beyond 92%. Therefore, it can be concluded that the ensemble learning classifier can be very effective when dealing with the faulty prediction task of heavy machine with complex component-system structure.

From the stability evaluation of the different classifiers, it can be found that the proposed ensemble learning approach with the Bayesian weighting strategy achieves the lowest variance value of 0.66, indicating the model accuracy stability in predicting each kind of machine conditions. Moreover, it can be found that the ensemble learning approach with the "winner takes all" being used for comparison outperforms the other four single classifiers in terms of the prediction stability with the variance value of 0.87. The reason should be that although the ensemble learning model based on "winner takes all" is totally based on the principal of majority voting which cannot reflect the component-system affiliation relationship of the heavy equipment, it can reflect the component-system structure of the heavy machine which cannot be achieved by a single classifier.

(2) Proving the Effectiveness of the Proposed Two-Stage SAE-Based Feature Fusion. In order to evaluate the proposed two-

TABLE 9: The configuration of the SAE used for feature fusion.

Subsystem	Node per layer	Sparsity penalty	Batch sizes	Number of hidden layer	Weight decay
Pinion bearing	3-7-3	0.02	100	1	0.0001
Drum shaft	2-5-2	0.02	100	1	0.0001
Retarder input shaft	2-5-2	0.02	100	1	0.0001
Retarder output shaft	3-7-3	0.02	100	1	0.0001

TABLE 10: The configuration of the SAE used for feature compression.

Subsystem	Node per layer	Sparsity penalty	Batch sizes	Number of hidden layer	Weight decay
Pinion bearing	42-6-42	0.02	100	1	0.0001
Drum shaft	30-6-30	0.02	100	1	0.0001
Retarder input shaft	30-6-30	0.02	100	1	0.0001
Retarder output shaft	42-6-42	0.02	100	1	0.0001

stage SAE-based feature fusion, we fix the classifier as the proposed DBN-based ensemble learning classifier based on the Bayesian weighting strategy. The raw sensor data, the six statistical features with SAE extraction, and the six-node compressed features are used as the input for comparison. The simulation result is illustrated as shown in Table 6. It can be found that the proposed ensemble learning approach can achieve the highest total accuracy of 96.78% with the model input of the six-node compressed feature. The six statistical features with SAE extraction can achieve the accuracy ranging from 82% to 87% which can be inferior to the six-node compressed feature. The raw sensor data achieve the lowest total accuracy of 55.24%. The conclusion is consistent with the visualization effect of t-SNE illustrated in Figure 11.

4.2. Case Study II: Faulty Prognostic for Industrial Port Crane Trolley

4.2.1. Data Description and Experimental Set up. In Case Study II, the performance of the proposed approach is evaluated on the multisensor datasets of the crane trolley provided by the enterprise of the CSIC (China Shipping Industry Company). As shown in Figure 13, the crane trolley mainly consists of four subsystems namely the pinion bearing, drum shaft, retarder input shaft, and retarder output shaft. There are 10 vibration sensors with the sensor series number from PB-01 to ROS-03 installed inside the crane trolley, and the sensor deployment is illustrated as shown in Table 7. The sampling frequency in Case Study II is set to 24 KHz, and the sampling time is set as 1 s. The time window is set to 10, and there are 2400 samples in total containing 10 sampling points in each.

The specific details of the datasets of Case Study II are illustrated as shown in Table 8. There are four kinds of crane trolley conditions namely wheel biting, wire rope over winding, brake failure, and healthy.

4.2.2. Feature Extraction and Normalization. Same as the Case Study I, six kinds of features are extracted and normalized from sensors monitoring the same subsystem.

TABLE 11: The configuration of individual DBN classifier used in Case Study II.

Number of RBM	2
Node per layer	6-10-10-10-4
Learning rate	0.1
Number of epochs	60
Batch size	100
Weight decay	0.0002

Take the pinion bearing as an example, six kinds of normalized features are extracted from the "sensor PB-01," "sensor PB-02," and "sensor PB-03," respectively, as shown in Figures 14(a)–14(g).

4.2.3. The Feature Fusion and Feature Construction. Same as the Case Study I, the six extracted features from the sensor signal are fused by SAE to obtain the compressed feature representing the comprehensive health st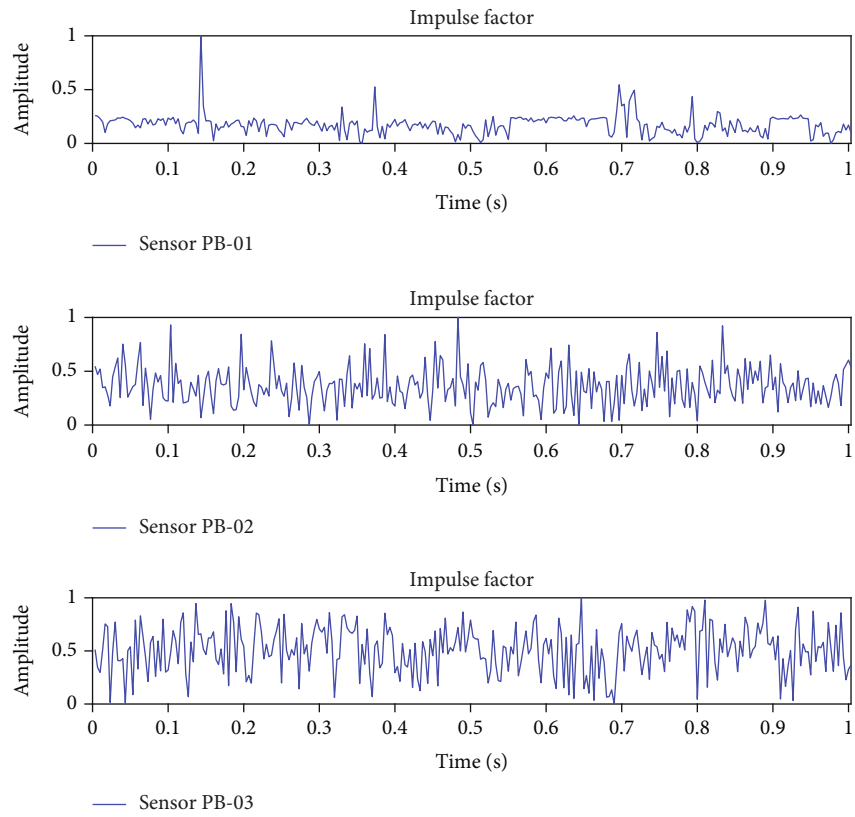atus of the target component. The configuration of the SAE of the four subsystems used for feature extraction and feature fusion is illustrated as shown in Tables 9 and 10.

4.2.4. DBN-Based Ensemble Learning. Same as the Case Study I, the six-node compressed feature indicating the comprehensive representation of the four subsystems are extracted and used as the input of four individual DBN classifiers. The configuration of a single individual DBN classifier used in Case Study II is illustrated as shown in Table 11. Since there are four kinds of crane trolley condition, the unit number of the output layer is set as 4. The overall accuracy curves and overall loss curves of each individual DBN and the proposed ensemble learning networks are illustrated as shown in Figures 15(a) and 15(b). It can be also found that in Case Study II, the proposed ensemble learning approach outperforms other four individual DBN classifiers in terms of the overall testing accuracy and the overall testing loss.
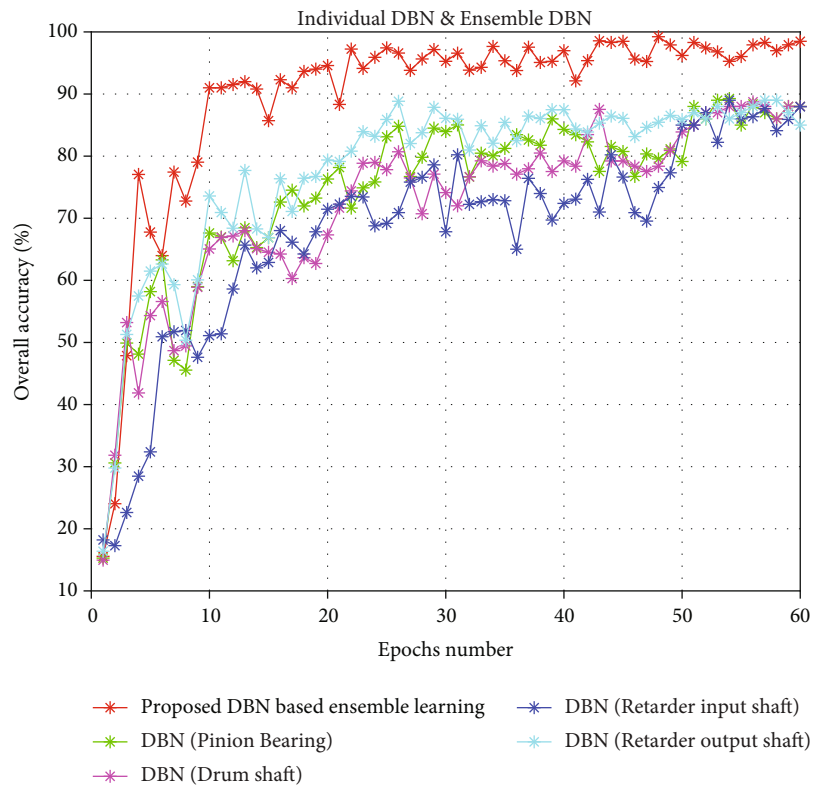
(a)
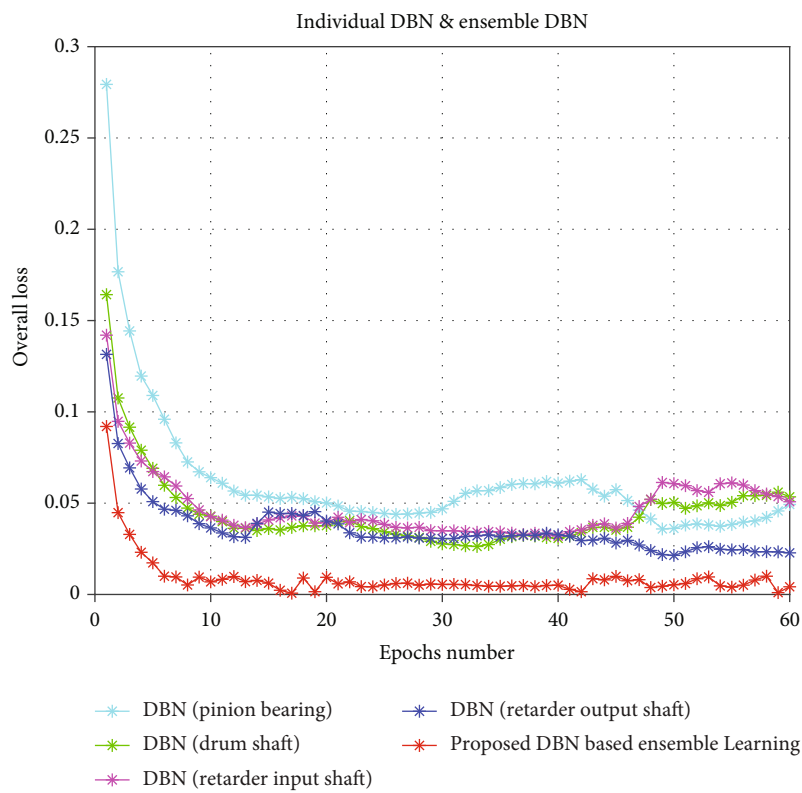
(b)

FIGURE 15: The performance of single individual classifier and the proposed DBN-based ensemble learning. (a) Testing accuracy curves. (b) Testing loss curves.

TABLE 12: The average accuracy of each condition of individual classifier and the proposed ensemble learning in Case Study II.

Classifier	Healthy	Wheel biting	Wire rope over winding	Brake failure	Total accuracy	Variance
Proposed ensemble learning	96.88%	96.54%	96.87%	98.35%	97.02%	0.65
DBN (pinion bearing)	85.53%	81.77%	88.59%	74.15%	84,02%	38.84
DBN (drum shaft)	82.87%	88.96%	85.58%	79.17%	83.51%	17.20
DBN(retarder input shaft)	82.11%	73.01%	77.84%	85.23%	80.83%	28.17
DBN (retarder output shaft)	84.85%	91.40%	84.79%	89.20%	86.21%	10.81

TABLE 13: The comparison of different prediction models.

Input	Classifier	Total accuracy	Variance
Six-node compressed feature	Proposed DBN-based ensemble learning (Bayesian weighting)	97.02%	0.65
	DBN-based ensemble learning (winner takes all)	93.06%	1.35
Six-node compressed feature (concatenated)	Single CNN	88.76%	8.74
	Single DBN	86.31%	7.81
	Single BPNN	86.97%	5.93
	Single SVM	87.79%	8.85

TABLE 14: The comparison of different model input.

Classifier	Input	Total accuracy
Proposed DBN-based ensemble learning (Bayesian weighting)	Six-node compressed feature	97.02%
	Impulse+SAE extraction	83.51%
	Kurtosis+SAE extraction	82.20%
	Skewness+SAE extraction	85.67%
	Shape factor+SAE extraction	83.91%
	Clearance factor+SAE extraction	81.49%
	Crest factor+SAE extraction	85.43%
	Raw sensor data without SAE extraction	62.33%

The experiment is repeated 10 times, and the average accuracy in predicting each kind of conditional status and the total accuracy is illustrated as shown in Table 12. It can be found that the proposed ensemble learning approach outperforms the other four individual DBN classifiers in terms of the average prediction accuracy in each condition. Moreover, the proposed ensemble learning approach outperforms the other four DBN classifiers in terms of the stability in predicting the four different kinds of crane trolley conditions based on the variance evaluation in Case Study II.

4.2.5. Comparison with Other Methods. Same as the Case Study I, the comparison experiments are carried out 10 times and evaluated from the input aspect and prediction aspect, respectively.

(1) Proving the Effectiveness of the Proposed DBN-Based Ensemble Learning Classifier. In comparison experiment of prediction aspect of Case Study II, we fix the model input as the six-node composite feature as shown in Table 13. It can be also found that the proposed DBN-based ensemble learning approach based on Bayesian weighting outperforms all the other four single classifiers and the ensemble learning

approach based on (winner takes all) with the highest total accuracy of 97.02%.

Moreover, it can be found that both ensemble learning models outperforms the other four single models in terms of the stability evaluated with lower variance which is consistent with the conclusion of Case Study I.

(2) Proving the Effectiveness of the Proposed Two-Stage SAE-Based Feature Fusion. In comparison experiment of the input aspect of Case Study II, we fix the classifier as the proposed DBN-based ensemble learning classifier based on the Bayesian weighting strategy as shown in Table 14. It can be also found in Case Study II that the proposed ensemble learning approach can achieve the highest total accuracy of 97.02% with the model input of the six-node compressed feature. The six fused features of impulse, kurtosis, skewness, shape factor, clearance factor, and the crest factor can achieve the inferior average accuracy ranging from 81% to 86% The raw sensor data achieve the lowest total accuracy of 62.33%.

The comparison result of Case Study II is consistent with the result of Case Study I, indicating the effectiveness of the proposed faulty identification approach also being applicable in the practical use of industrial crane trolley of Case Study II.

5. Conclusion and Future Work

5.1. Main Contribution of the Proposed Paper. In this paper, a novel hybrid faulty prognostic approach based on sparse autoencoder- and deep belief network-based ensemble learning is proposed. The main contribution is summarized as follows:

(1) Introducing the SAE-based feature extraction method so that the features extracted from different sensors are merged into one stream, releasing the influence caused by different sensor deployment

(2) Proposing the SAE-based feature fusion method so that the six kinds of extracted features are compressed to construct a composite feature which is regarded as the health indicator of a target component. The constructed composite feature has better robust ability when dealing with the nonstationary faulty vibration signal

(3) Proposing the DBN-based ensemble learning so that the machine internal component-system relationship can be well represented. The outputs of these DBN classifiers are aggregated by using the Bayesian weighting strategy which represent the affiliated degree between the health statuses of the target component and faulty probability of the certain type of machine faulty status

The proposed hybrid faulty prognostic approach is evaluated on two case studies of wind turbine gearbox and the industrial port crane trolley. Both case studies demonstrate that the proposed hybrid faulty identification approach outperforms other traditional faulty identification methods in terms of the prediction accuracy and prediction stability when dealing with the industrial heavy machine, indicating the necessity of applying the two-stage SAE- and the DBN-based ensemble learning.

5.2. Future Work. Although the proposed hybrid faulty identification method has been well evaluated in the proposed two case studies, there are still some issues needed to be considered. Firstly, the computation complexity of the proposed Bayesian weighting strategy used in the ensemble learning process should be taken into account. In the future, some more suitable weighting strategies should be designed which can not only release the computation burden but also reflect the component-system relationship. Moreover, the proposed hybrid faulty identification approach should be expected to be evaluated in other heavy machines such as vehicle system and aircraft engine system which also have a highly complex internal component-system structure.

Data Availability

The raw/processed sensor data of the two case studies are offered by the enterprises of "Shanghai Electric Group" and the "China Shipping Industrial Company," respectively. It cannot be shared due to the corporate confidentiality.

Conflicts of Interest

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

Acknowledgments

References

[1] N. Gebraeel, M. Lawley, R. Liu, and V. Parmeshwaran, "Residual life predictions from vibration-based degradation signals: a neural network approach," *IEEE Transactions on Industrial Electronics*, vol. 51, no. 3, pp. 694–700, 2004.

[2] A. Malhi, R. Yan, and R. X. Gao, "Prognosis of defect propagation based on recurrent neural networks," *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 3, pp. 703–711, 2011.

[3] B. Kilundu, X. Chiementin, and P. Dehombreux, "Singular spectrum analysis for bearing defect detection," *Journal of Vibration and Acoustics*, vol. 133, no. 5, article 051007, 2011.

[4] P. K. Varshney, "Multisensor data fusion," *Electronics & Communication Engineering Journal*, vol. 9, no. 6, pp. 245–253, 1997.

[5] T. Wang and H. Sun, "Design method of data acquisition in intelligent sensor based on web data mining clustering technology," in *International Conference on Education*, Shenyang, China, 2015.

[6] C. Axenie and J. Conradt, "Cortically inspired sensor fusion network for mobile robot egomotion estimation," *Robotics and Autonomous Systems*, vol. 71, pp. 69–82, 2015.

[7] K. Peng, R. Jiao, J. Dong, and Y. Pi, "A deep belief network based health indicator construction and remaining useful life prediction using improved particle filter," *Neurocomputing*, vol. 361, no. Oct.7, pp. 19–28, 2019.

[8] J. Wu, Y. Su, Y. Cheng, X. Shao, C. Deng, and C. Liu, "Multisensor information fusion for remaining useful life prediction of machining tools by adaptive network based fuzzy inference system," *Applied Soft Computing*, vol. 68, pp. 13–23, 2018.

[9] J. Gokulachandran and K. Mohandas, "Comparative study of two soft computing techniques for the prediction of remaining useful life of cutting tools," *Journal of Intelligent Manufacturing*, vol. 26, no. 2, pp. 1–14, 2013.

[10] H. Yan, K. Liu, X. Zhang, and J. Shi, "Multiple sensor data fusion for degradation modeling and prognostics under multiple operational conditions," *IEEE Transactions on Reliability*, vol. 65, no. 3, pp. 1416–1426, 2016.

[11] T. P. Banerjee and S. Das, "Multi-sensor data fusion using support vector machine for motor fault detection," *Information Sciences*, vol. 217, no. Complete, pp. 96–107, 2012.

[12] L. Xu, C. Tan, X. Li, Y. Cheng, and X. Li, "Fuel-type identification using joint probability density arbiter and soft-computing techniques," *IEEE Transactions on Instrumentation & Measurement*, vol. 61, no. 2, pp. 286–296, 2012.

[13] J. Wang, J. Xie, R. Zhao, L. Zhang, and L. Duan, "Multisensory fusion based virtual tool wear sensing for ubiquitous manufacturing," *Robotics and Computer-Integrated Manufacturing*, vol. 45, pp. 47–58, 2017.

[14] B. Y. Sun, X. M. Zhang, J. Li, and X. M. Mao, "Feature fusion using locally linear embedding for classification," *IEEE Transactions on Neural Networks*, vol. 21, no. 1, pp. 163–168, 2010.

[15] J. Ni, C. Zhang, and S. X. Yang, "An adaptive approach based on KPCA and SVM for real-time fault diagnosis of HVCBs," *IEEE Transactions on Power Delivery*, vol. 26, no. 3, pp. 1960–1971, 2011.

[16] N. Saravanan, V. N. S. K. Siddabattuni, and K. I. Ramachandran, "Fault diagnosis of spur bevel gear box using artificial neural network (ANN), and proximal support vector machine (PSVM)," *Applied Soft Computing*, vol. 10, no. 1, pp. 344–360, 2010.

[17] A. Malhi and R. X. Gao, "PCA-based feature selection scheme for machine defect classification," *IEEE Transactions on Instrumentation and Measurement*, vol. 53, no. 6, pp. 1517–1525, 2004.

[18] M. Aminian and F. Aminian, "Neural-network based analog-circuit fault diagnosis using wavelet transform as preprocessor," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 47, no. 2, pp. 151–156, 2000.

[19] C. Li, R. V. Sanchez, G. Zurita, M. Cerrada, D. Cabrera, and R. E. Vásquez, "Gearbox fault diagnosis based on deep random forest fusion of acoustic and vibratory signals," *Mechanical Systems & Signal Processing*, vol. 76-77, pp. 283–293, 2016.

[20] T. Junbo, L. Weining, A. Juneng, and W. Xueqian, "Fault diagnosis method study in roller bearing based on wavelet transform and stacked auto-encoder," in *The 27th Chinese Control and Decision Conference (2015 CCDC)*, Qingdao, China, 2015.

[21] 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.

[22] H. Shao, H. Jiang, H. Zhang, and T. Liang, "Electric locomotive bearing fault diagnosis using a novel convolutional deep belief network," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 3, pp. 2727–2736, 2018.

[23] L. Wen, X. Li, L. Gao, and Y. Zhang, "A new convolutional neural network-based data-driven fault diagnosis method," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 7, pp. 5990–5998, 2017.

[24] G. E. Hinton, S. Osindero, and Y. W. Teh, "A fast learning algorithm for deep belief nets," *Neural Computation*, vol. 18, no. 7, pp. 1527–1554, 2006.

[25] V. T. Tran, F. Althobiani, and A. Ball, "An approach to fault diagnosis of reciprocating compressor valves using Teager-Kaiser energy operator and deep belief networks," *Expert Systems with Applications*, vol. 41, no. 9, pp. 4113–4122, 2014.

[26] P. Tamilselvan and P. Wang, "Failure diagnosis using deep belief learning based health state classification," *Reliability Engineering & System Safety*, vol. 115, no. jul., pp. 124–135, 2013.

[27] Z. Chen, X. Zeng, W. Li, and G. Liao, "Machine fault classification using deep belief network," in *2016 IEEE international instrumentation and measurement technology conference proceedings*, Taipei, Taiwan, 2016.

[28] G. Brown, J. Wyatt, R. Harris, and Y. Xin, "Diversity creation methods: a survey and categorisation," *Information Fusion*, vol. 6, no. 1, pp. 5–20, 2005.

[29] F. Fernandez-Navarro, P. A. Gutierrez, C. Hervas-Martinez, and X. Yao, "Negative correlation ensemble learning for ordinal regression," *IEEE Transactions on Neural Networks & Learning Systems*, vol. 24, no. 11, pp. 1836–1849, 2013.

[30] J. Yang, X. Zeng, S. Zhong, and S. Wu, "Effective neural network ensemble approach for improving generalization performance," *Transactions on Neural Networks & Learning Systems*, vol. 24, no. 6, pp. 878–887, 2013.

[31] R. Coop, A. Mishtal, and I. Arel, "Ensemble learning in fixed expansion layer networks for mitigating catastrophic forgetting," *IEEE Transactions on Neural Networks & Learning Systems*, vol. 24, no. 10, pp. 1623–1634, 2013.

[32] C. Zhang, P. Lim, A. K. Qin, and K. C. Tan, "Multi-objective deep belief networks ensemble for remaining useful life estimation in prognostics," *IEEE transactions on neural networks and learning systems*, vol. 28, no. 10, pp. 2306–2318, 2017.

[33] V. Bolón-Canedo and A. Alonso-Betanzos, "Ensembles for feature selection: a review and future trends," *Information Fusion*, vol. 52, pp. 1–12, 2019.

[34] A. Ng, "Sparse auto-encoder," *CS294A Lecture notes*, vol. 72, pp. 1–19, 2011.

Research Article

Financial Risk Assessment of Enterprise Management Accounting Based on Association Rule Algorithm under the Background of Big Data

Yibin Lin,[1] Huabo Yue,[2] Haojie Liao [ID],[1,2,3] Dong Li,[3] and Ling Chen[4]

[1]*Chakrabongse Bhuvanath International Institute for Interdisciplinary Studies, Raja Mangala University of Technology, Tawan-Ok 10600, Thailand*
[2]*College of Graduate Studies, Master of Management Program in Management (International Program), Walailak University, 222 Thaiburi, Thasala, Nakhon Si Thammarat 80160, Thailand*
[3]*Accounting and Audit School, Guangxi University of Finance and Economics, Nanning, Guangxi Province 530003, China*
[4]*College of Graduate Studies, Bansomdejchaopraya Rajabhat University, Bangkok 10600, Thailand*

Correspondence should be addressed to Haojie Liao; 2017110007@gxufe.edu.cn

Under the background of the rapid development of data informatization, the evaluation and analysis of enterprise management accounting risk are the key link of enterprise sustainable development. Based on the theoretical basis of enterprise management accounting risk analysis, this paper selects 11 representative corporate financial risk indicators from financial risk indicators through correlation analysis and then selects 8 key indicators that affect corporate financial risk through association rules, to assess and analyze the company's financial risks. Finally, this paper takes 20 ST companies in China as the research object to carry out empirical research. According to the assessment and analysis of financial risks in the company's management accounting, it proposes measures from four modules to strengthen the company's ability to resist risks.

1. Introduction

At present, in an environment of shrinking global demand and economic decline, most companies have also entered a new normal of low profit margins and low growth rates. Enterprises not only face the difficulties of increasing operational risks but also face severe challenges such as increased financial pressure and rising comprehensive costs. In the process of operation, if an enterprise can accurately discover the potential risk factors around it, the sustainable development of the enterprise can be directly and strongly guaranteed. Therefore, the selection of corporate financial risk indicators and the construction of the overall model and other related theories have also played a promoting role and positive impact on the sustainable and healthy development of the company.

Data mining is to find the relationship and laws between information and data from massive, unstructured, and fuzzy actual data and provide strong support for the decision-making of enterprise management. Introducing data mining technology into financial analysis strategy has become the mainstream development direction nowadays. With the help of risk analysis, enterprises will be able to understand financial risks more comprehensively and scientifically and take relevant defensive measures with a more active attitude. Under the background of market economic downturn, Chinese enterprises are under increasing competitive pressure. The advent of the digital age has ushered in a new approach that emphasizes the presentation of regularities in large-scale data analysis compared to traditional what-if analysis methods. Information data is also an important resource in management accounting. Therefore, this method has been fully applied in the management accounting financial risk evaluation of enterprises. The specific performance is, based on data mining technology, linking the dynamic nature of enterprise development, focusing on timeliness

and practicality, and establishing an enterprise financial risk analysis model. Ensure that potential risk factors existing in the operation of the enterprise can be discovered in time and be actively resolved and dealt with.

Under the background of the rapid development of big data, this paper uses association rule algorithm to mine the key indicators affecting the financial risk of enterprise management accounting. Then, this paper evaluates the financial risk of enterprise management accounting through key indicators, finds the root cause of enterprise financial risk, and puts forward relevant suggestions.

2. Literature Review

Management accounting is a branch of accounting with the goal of improving enterprise management and economic benefits. Scholars have carried out research on the development trend and method innovation of management accounting [1, 2]. Haka and Heitger expounded the opportunities and challenges faced by management accounting and proposed research on the role of management accounting based on an international management accounting framework [3]. Meng et al. conducted a research summary of 327 literatures in China and found that the research methods of management accounting are more diversified, which is mainly reflected in the significant increase in the proportion of database-based empirical research and survey research [4]. Lixia and Shishui sorted out the relevant research on management accounting tools and the typical problems existing in practical application and proposed to strengthen the research on the preliminary budget, production cost standard, business process, and internal assessment mechanism of management accounting [5]. In management accounting, the traditional cost calculation method can no longer meet the needs of managers. Ponisciakova et al. proposed an activity-based costing method to replace the traditional cost calculation method, so that the cost can be collected and calculated more accurately [6]. Yang et al. discussed the research hotspots of management accounting in China and proposed that more attention should be paid to the case studies of management accounting models in the future [7].

In the era of big data, data mining and analysis have improved the use value of data and information in the accounting industry and promoted the development of the accounting industry [8]. Minxia analyzed the impact of big data technology on management accounting from two aspects of opportunities and challenges and proposed strategies such as strengthening data storage technology and building analysis platforms to promote the development of management accounting in the era of big data [9]. Mihăilă proposed that enterprises need to develop management accounting business and develop a system for recording and mining data information according to the needs of enterprise management [10]. In the context of big data, Hua and Wang put forward a management accounting information application framework based on enterprise information fusion, which provides a reference for the application and development of management accounting [11].

Through empirical tests, Yigitbasioglu found that information system flexibility and management accounting adaptability have a significant positive correlation and deepened the connection between management accounting and information systems by exploring the role of factors that promote management accounting [12]. In order to make better use of business analysis data, Appelbaum et al. discussed the impact of business analysis on management accounting from the perspective of enterprise systems and proposed a management accounting data analysis framework [13].

In order to make better use of business analysis data, Al-Maolegi and Arkok discussed the impact of business analysis on management accounting from the perspective of enterprise systems and proposed a management accounting data analysis framework [14, 15]. Huayang and Xingxing used the association rule algorithm to mine and analyze the sales accounting data of enterprises and formulate corresponding sales plans according to the analysis results to provide managers with more decision-making information [16]. Wamba et al. proposed an explanatory framework to evaluate the business value of big data through case studies and other methods and determined the implementation value of big data strategy in management accounting [17].

To sum up, scholars at home and abroad have done a lot of research on big data and enterprise management accounting, but there are relatively few researches that combine big data and enterprise management accounting to really apply big data methods to management accounting applications. Therefore, based on the association rule algorithm, this paper evaluates and analyzes the financial risk of enterprise management accounting through specific financial indicators.

3. Risk Assessment and Analysis of Enterprise Management Accounting Based on Association Rules

Data mining can solve some technical obstacles in the implementation of management accounting by acquiring and utilizing the information with massive data, complex structure, and implicit knowledge [18, 19]. And association rules are a main form of data mining, looking for the correlation of different items in the same event [20]. According to the mining results of association rules, the key financial indicators are screened out to evaluate and analyze the financial risks of enterprise management accounting and put forward relevant suggestions. The specific process is shown in Figure 1.

3.1. The Selection of Financial Risk Evaluation Index of Enterprise Management Accounting. Most Chinese enterprises mainly focus on financial indicators in the management process. The traditional method of selecting financial indicators is to judge based on personal experience or cognition, which accounts for a large proportion of subjectivity. This paper selects all the financial indicators of the enterprise, uses the method of data mining to analyze the correlation between the financial indicators, and selects the most representative financial indicators. The financial risk assessment index system of enterprise management accounting is
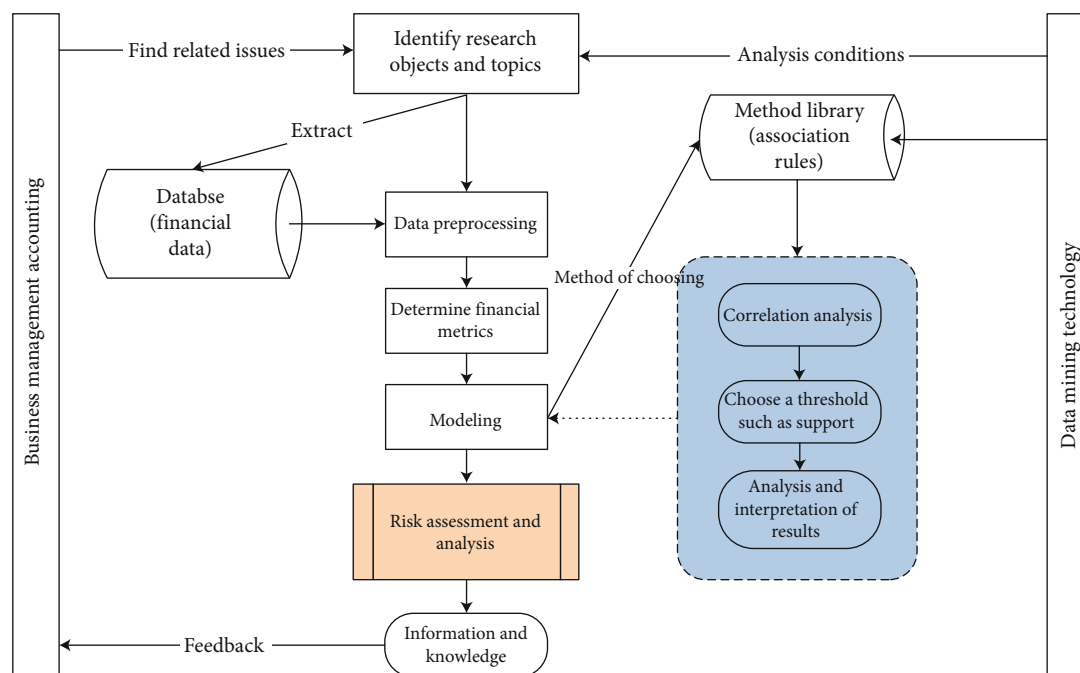
FIGURE 1: Enterprise management accounting risk analysis process based on association rules.

mainly divided into four aspects: enterprise profitability, enterprise solvency, enterprise operation ability, and enterprise growth ability. The specific indicators and meanings are shown in Table 1.

3.1.1. Corporate Profitability Indicators. Profitability is one of the most important core indicators in a company's financial statements, which reflects the company's ability to create profits. Among the profitability indicators, net profit margin and gross profit margin reflect the level of income created by the company's sales revenue within a certain period, and the level of profit margin reflects the strength of the company's capabilities. Basic earnings per share reflect the income created by each shareholder's equity, and the increase in total share capital will cause relative dilution of basic earnings per share. Total assets are composed of liabilities and shareholders' equity, and the return on total assets reflects the level of income generated by the funds invested by creditors and all shareholders. The return on total assets index is used to compare the relative levels of the same industry, and it reflects the level of the company's use of shareholders' equity and liabilities to create income. The level of return on total assets reflects the strength of the company's profitability.

3.1.2. Corporate Solvency Index. Corporate solvency indicators are divided into long-term solvency and short-term solvency. Insufficient solvency of the enterprise will affect the normal operation of the enterprise, causing the difficulty and cost of raising funds for the enterprise to increase and the decline of social reputation.

Short-term solvency indicators are generally measured by three factors: quick ratio, current ratio, and cash ratio. The company's quick ratio is an indicator to assess and measure the short-term solvency of a company. The stronger

the ability of an enterprise's quick assets to be realized in a short period of time, the stronger the short-term solvency of the enterprise. The current ratio reflects the relative level of the company's current assets and current liabilities. If the current ratio is too high or too low, it will have a greater impact on the company. The cash ratio reflects the ability of a company to cover short-term debt with its cash flow. The more cash on the company's accounts, the more conducive it is to repay the company's debt, but it also shows that the company's ability to use cash is insufficient.

The long-term solvency index is generally measured by the asset-liability ratio and the interest coverage ratio. The gearing ratio, also known as financial leverage, reflects the relative level of corporate liabilities and free assets. A reasonable level of assets and liabilities is conducive to the good operation of an enterprise. An excessively high asset-liability ratio will lead to higher risks for the enterprise, and a too low asset-liability ratio is not conducive to the expansion and rapid development of the enterprise. The interest coverage ratio measures whether the profit created by the enterprise can cover the level of interest expenses. The level of the interest coverage ratio indicates the strength of the long-term solvency of the enterprise.

3.1.3. Enterprise Operation Capability Index. The operational capability of an enterprise reflects the level of operational efficiency in the entire business process, including production, sales, and payment collection. The operational capability of the enterprise also reflects the operational efficiency and capability of the operator of the enterprise from the side. The higher the total asset turnover rate, the higher the market recognition of the products produced by the company, and the company's own assets can be quickly converted into sales revenue, bringing benefits to the company.

TABLE 1: Enterprise management accounting financial risk assessment indicators.

Financial indicator		Meaning
Corporate profitability	Gross profit margin	The ratio of current net profit to current sales revenue
	Net interest rate	The ratio of current sales revenue less cost of sales to current sales revenue
	ROE	The ratio of net profit for the period to the average value of net assets
	Basic earnings per share	The ratio of net profit for the current period to the total number of shares
	Return on total assets	The ratio of net profit for the period to the average value of total assets
Corporate solvency	Quick ratio	Total current assets less inventory to current liabilities ratio
	Current ratio	Ratio between total current assets and total current liabilities
	Cash ratio	Ratio of all cash plus marketable securities to current liabilities
	Assets and liabilities	Ratio of total liabilities to total assets
	Interest coverage ratio	The ratio of the current operating profit to the interest expense of the enterprise
Corporate operation capability	Total asset turnover	The ratio of net sales revenue to the average of total assets
	Accounts receivable turnover	The ratio between the net credit income for the period and the average balance of accounts receivable for the period
	Inventory turnover	The ratio between the cost of sales for the current period and the average inventory balance for the current period
Corporate growth ability	Total asset growth rate	The ratio of the increase in total assets for the current period to the total assets at the beginning of the period
	Operating income growth rate	Operating income of the current period minus the ratio of operating income of the previous period to the operating income of the previous period
	Net profit growth rate	The ratio between the net profit of the current period minus the net profit of the previous period and the net profit of the previous period

The level of the total asset turnover ratio reflects the strength of the enterprise's operating ability. A higher accounts receivable turnover ratio not only means that the company can quickly recover the sales payment but also reflects the customer's recognition of the company's products and the overall competitiveness of the company. The higher the inventory turnover rate, the less the inventory of the enterprise, the less the amount of funds occupied by the inventory, and the reduction of the production cost of the enterprise occupied by the inventory.

3.1.4. Enterprise Growth Capability Indicator.
Enterprise growth capability measures the development potential of an enterprise through its own production and operation activities. The rapid growth of enterprises is often positively related to the three factors of total assets, operating income, and market share of the enterprise. Total assets equal the sum of owners' equity and company liabilities. The increase in the growth rate of total assets indicates that the development trend of the enterprise is improving, but a reasonable asset-liability ratio is an indispensable factor in evaluating the growth rate of total assets. The higher the growth rate of operating income, the faster the growth rate of the enterprise. The increase in operating income is mainly composed of two factors: the increase in products sold by the enterprise and the increase in labor services provided. The level of net profit growth rate reflects the strength of the company's ability to generate income. The upgrade of the company's products leads to price increases and the increase in sales

cost, and the expansion of the company's scale will affect the growth of the company's net profit.

3.2. Correlation Analysis of Management Accounting Indicators.
Use the traditional method of analyzing the correlation coefficient of variables to analyze the correlation of the selected financial indicators. The calculation formula of the correlation coefficient is as follows:

$$r = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)}{\sqrt{\left(\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 \sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2\right)}}. \tag{1}$$

Equation (1) represents the correlation of two variables. Among them, X and Y represent two different variables, r represents the correlation coefficient of the two variables X and Y, and when r takes different values, the relationship between the two variables is different. When $r = 1$, the relationship between the two variables is completely positive; when $0 < r < 1$, the relationship between the two variables is linearly positive; when $r = 0$, the relationship between the two variables is completely uncorrelated; when $-1 < r < 0$, the relationship between the two variables is linearly negatively correlated; when $r = -1$, the relationship between the two variables is completely negatively correlated.

3.3. Definition of Risk Level of Enterprise Financial Indicators.
In the association rule algorithm, the mined data is generally discrete, while the corporate financial data is

TABLE 2: Classification of risk levels of corporate financial indicators.

Financial indicator	Risk level				
	1	2	3	4	5
Gross profit margin	>0.6	[0.2,0.6]	[−0.2,0.2]	[−0.6,−0.2]	<−0.6
Net interest rate	>0.4	[0.1,0.4]	[−0.1,0.1]	[−0.4,−0.1]	<−0.4
ROE	>0.4	[0.1,0.4]	[−0.1,0.1]	[−0.4,−0.1]	<−0.4
EPS	>1	[0.3,1]	[−0.3,0.3]	[−1,−0.3]	<−1
Current ratio	>3	[2, 3]	[1, 2]	[0.5,1]	[0,0.5]
Cash ratio	>3	[2, 3]	[1, 2]	[0.5,1]	[0,0.5]
Asset-liability ratio	[0,0.2]	[0.2,0.4]	[0.4,0.6]	[0.6,0.8]	[0.8,1]
Accounts receivable turnover	>100	[50,100]	[20, 50]	[5, 20]	[0, 5]
Inventory turnover	>100	[50,100]	[20, 50]	[5, 20]	[0, 5]
Total asset growth rate	>1	[0.3,1]	[0.1,0.3]	[0,0.1]	<0
Net profit growth rate	>0.5	[0.1,0.5]	[−0.1,0.1]	[−0.5,−0.1]	<−0.5

generally continuous. Therefore, when using association rules, it is necessary to convert continuous data into discrete data by defining financial data indicators, so that it presents discrete distribution characteristics. In addition, there are large differences in the value of financial indicators. This paper divides the risk level of enterprise financial indicators based on five levels, as shown in Table 2.

3.4. Financial Risk Assessment Model of Enterprise Management Accounting Based on Association Rules. Through association rules, the relationship between the financial risk indicators of management accounting is excavated, and the financial risk indicators of management accounting of enterprises are analyzed and evaluated, to find the reasons for the risk of management accounting of enterprises.

3.4.1. Establish the Concept Hierarchy Tree of Enterprise Management Accounting Financial Risk. Constructing a risk concept hierarchy tree is the key to establishing an enterprise management accounting financial risk assessment model, and the relationship between indicators at different levels can be more clearly expressed through the hierarchy tree. Through the analysis of low-level financial indicators, it is extended to high-level concepts, and finally, the level of enterprise management accounting risk is found.

The financial risk index system of enterprise management accounting is mainly divided into four modules: profitability, solvency, operation, and growth. There are two levels in each module, mainly the specific indicators of related modules. Therefore, the entire financial risk concept hierarchy tree is divided into four levels, as shown in Figure 2. The first layer is the root node to manage enterprise financial risks. The second layer is the four major modules of the enterprise's profitability, solvency, operating ability, and growth ability. The third layer is the comprehensive index of enterprise financial risk evaluation. The fourth layer is the specific financial indicators under the four modules, including gross profit rate, net profit rate, and account receivable turnover rate.

3.4.2. Data Mining Strategies with Different Support Thresholds. When performing data mining on each level, the minimum support threshold for financial indicators needs to be established. Generally, the lower the level of financial indicators, the smaller the corresponding minimum support threshold. Taking the profitability indicator as an example, as shown in Figure 3, the minimum support threshold for the second level is 20%, the minimum support threshold for the third level is 10%, and the minimum support threshold for the fourth level is 5%.

3.4.3. Result Output and Interpretation. By gradually adjusting the support thresholds of different levels, using the association rule algorithm to mine the data of enterprise management accounting risk indicators, analyzing, and comparing the performance of financial indicators under different support thresholds, the final key indicators for evaluating financial risks are obtained.

4. Empirical Research on Financial Risk Analysis of Enterprises

4.1. Selection of Sample Companies. China's ST (special treatment) company refers to a listed company whose net profit has been negative for two consecutive years or the stock exchange considers that the financial situation is abnormal. The management accounting risk of ST company is relatively large, which has a certain representative significance. Therefore, this paper selects 56 ST companies in China's listed companies from 2016 to 2021, removes some abnormal data and incomplete data in ST companies, and finally selects 20 companies as sample companies. The enterprise data in this article comes from the CHOICE financial terminal software data.

4.2. Selection of Financial Risk Indicators in Enterprise Management Accounting. According to the actual situation of the enterprises selected in this paper and the commonly used indicator system, a total of 16 financial indicators have
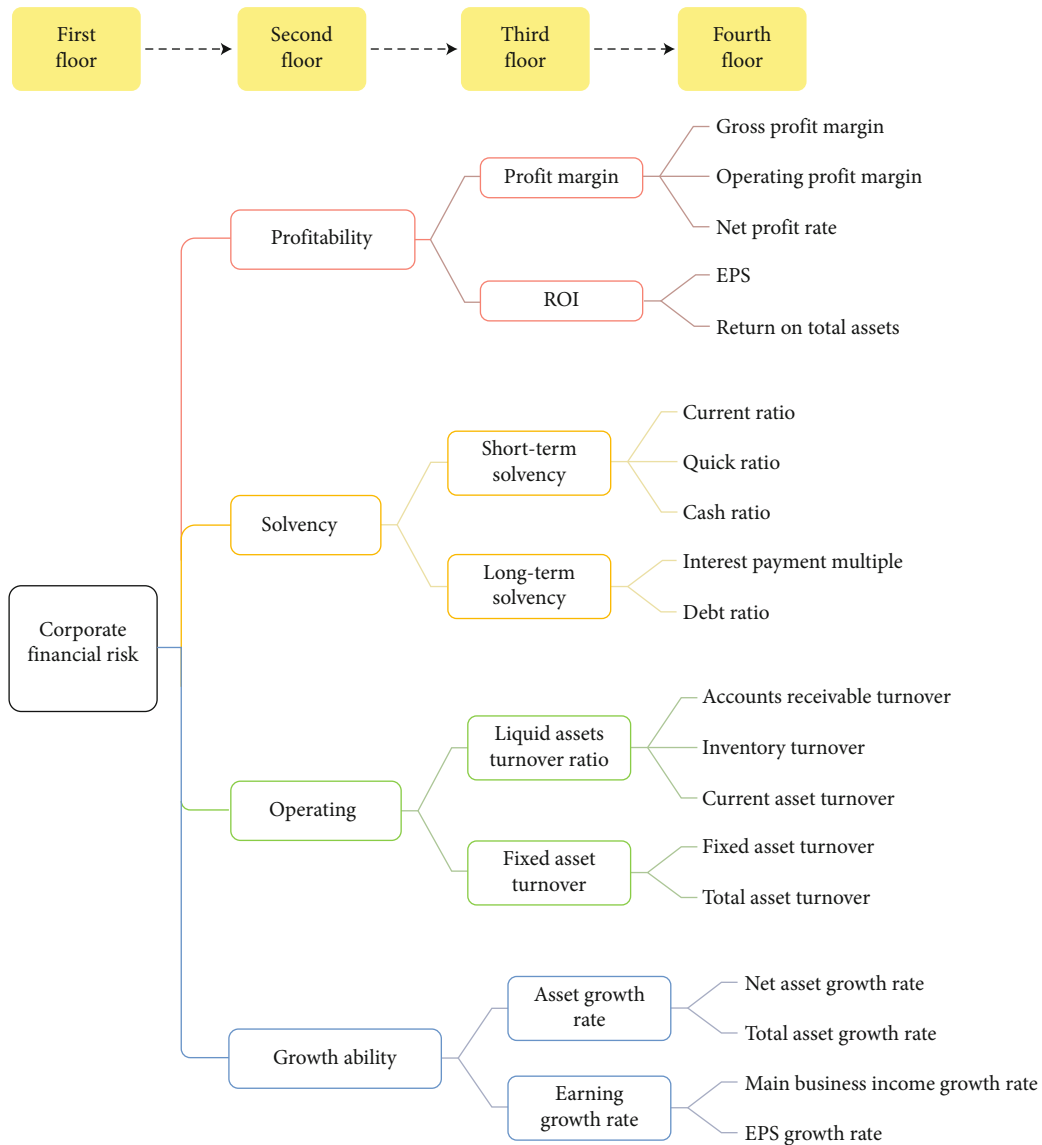
FIGURE 2: Concept hierarchy tree of enterprise financial risk.
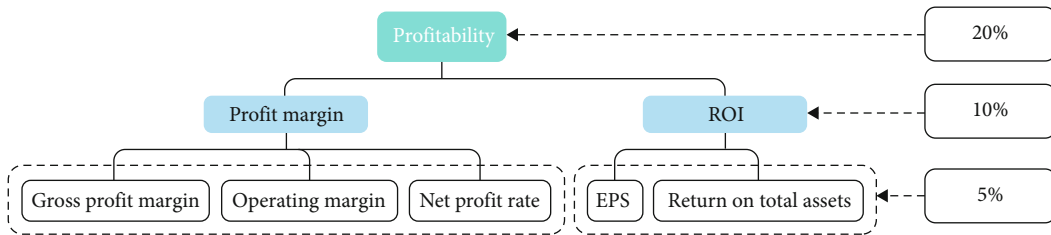
FIGURE 3: Profitability data mining strategy.

been collected, namely: (1) gross profit margin, (2) net profit margin, (3) return on assets, (4) basic earnings per share, (5) return on total assets, (6) quick ratio, (7) current ratio, (8) cash ratio, (9) asset-liability ratio, (10) Interest coverage ratio, (11) total asset turnover ratio, (12) accounts receivable turnover ratio, (13) inventory turnover rate, (14) growth rate of total assets, (15) growth rate of net profit, and (16) growth rate of operating income. These enterprise management accounting financial risk indicators have a certain correlation. Running SAS software through the correlation coefficient calculation module, the correlation coefficient between the indicators is obtained as shown in Table 3.

TABLE 3: Correlation of financial indicators.

X Y	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	1															
2	0.0912	1														
3	0.0754	0.112	1													
4	0.2413	0.2354	0.401	1												
5	0.1465	0.7512	0.805	0.7621	1											
6	0.1685	-0.063	0.0412	0.0846	0.06	1										
7	0.1896	-0.066	0.0385	0.0753	0.0587	0.948	1									
8	0.0621	0.0534	0.0432	0.1046	0.0824	0.3142	0.2764	1								
9	-0.342	0.2054	-0.231	-0.076	-0.129	-0.491	-0.509	-0.306	1							
10	0.343	-0.865	-0.045	-0.096	-0.12	0.1714	0.1611	-0.07	-0.031	1						
11	-0.514	0.1325	0.003	0.063	-0.014	-0.245	-0.246	-0.179	0.314	-0.182	1					
12	0.4321	-0.329	-0.013	-0.168	-0.087	0.0814	0.0943	0.037	-0.286	0.2468	-0.524	1				
13	0.5846	-0.682	0.0249	-0.121	-0.092	0.213	0.2198	-0.089	-0.364	0.8672	-0.423	0.4102	1			
14	0.0528	0.1954	0.191	0.4625	0.512	-0.012	-0.007	-0.041	0.1419	-0.12	-0.052	-0.129	-0.071	1		
15	-0.861	0.7658	0.5684	0.4895	0.547	-0.114	-0.142	0.021	0.213	-0.814	-0.005	-0.352	-0.698	0.5148	1	
16	-0.614	0.1103	0.0514	0.2147	0.171	-0.143	-0.153	-0.124	0.191	-0.068	0.314	-0.18	-0.104	0.1537	0.231	1

According to the correlation coefficient obtained by SAS software, the financial risk indicators with very high positive and negative correlation and the financial risk indicators with large correlation coefficient in the same module are screened and excluded. Finally, 11 financial indicators are obtained, namely: gross profit margin, net profit margin, return on equity, basic earnings per share, current ratio, cash ratio, asset-liability ratio, accounts receivable turnover ratio, inventory turnover ratio, total asset growth rate, and net profit growth rate.

4.3. Management Accounting Risk Assessment Based on Association Rules

4.3.1. Reconstruction of Financial Indicator Data Based on the Above Discretization Principles. After the correlation analysis of the sample data, the index values before the reconstruction of 11 financial indicators are obtained, and then the data is discretely reconstructed. The data of some indicators are listed in Table 4.

After the financial index values are reconstructed based on the above discretization principles, the continuous financial index data is converted into discrete financial index data as shown in Table 5.

4.3.2. Result Output and Interpretation. Run SAS software to perform association rule mining on the selected financial indicators through Apriori algorithm. Select different support thresholds and confidence values, and when the frequent itemsets of graded data under the corresponding thresholds are obtained, set the previous frequent itemsets as A, and recalculate the frequent itemsets under the new thresholds. Rescreening and recombining these data finally obtain frequent itemsets, looking for the number and rules of association rules between financial indicators.

The dataset is the financial indicator data of China ST company from 2016 to 2021, as shown in Table 6. When the support threshold is 0.6 and the confidence threshold is 0.3, 86 association rules are found. When the support threshold is 0.7 and the confidence threshold is 0.3, 34 association rules are found. When the support threshold is 0.8 and the confidence threshold is 0.7, 4 rule associations are found.

By mining and analyzing the association rules algorithm of the financial indicators of China ST Company and the relationship between the number of associations under different support thresholds and confidence thresholds, it is found that some key financial indicators always appear frequently when the enterprise management accounting crisis occurs. They are as follows: return on equity, basic earnings per share, current ratio, cash ratio, accounts receivable turnover ratio, inventory turnover ratio, total asset growth rate, net profit growth rate, and other eight financial indicators.

4.3.3. Analysis of financial risk in enterprise management accounting. Based on the data mining analysis of the association rules algorithm, the key indicators affecting the management accounting risk of the enterprise are obtained, namely: return on equity, basic earnings per share, current ratio, cash ratio, accounts receivable turnover ratio, inventory turnover ratio, total assets growth rate, and net profit growth rate. These key indicators fully reflect the degree of corporate management accounting risks from certain angles and provide a reference basis for investors to judge corporate profitability, debt repayment, operation, growth, and other aspects. According to the value of these key financial indicators, the risk level of financial indicators is judged, and the profitability, debt repayment, operation, and growth of enterprise management accounting risk are further obtained.

TABLE 4: Data related to corporate financial indicators.

| Securities name | Financial indicator | | | | |
	Gross profit margin (%)	Roe (%)	Net interest rate (%)	Current ratio	Cash ratio
ST Ankong	11.8128	-65.4141	-13.5296	0.7021	0.0503
ST Antai	9.3034	14.8200	5.9253	0.8800	0.2719
ST Aoma	22.5387	-22.7944	-1.1148	1.3781	0.8882
ST Baling	17.9890	-63.4871	-43.6758	0.9167	0.0501
ST Bolong	29.1277	-14.1991	-10.2818	2.1525	0.1654
ST Beiwen	45.7802	-36.4305	-16.8846	1.0729	0.0200
ST Busen	21.0667	-57.5625	-30.7033	1.4024	0.3107
ST Dayou	24.0102	-15.1652	-6.1860	0.7627	0.3828
ST Dazhou	31.9129	-55.1888	-11.0538	0.3197	0.0707
ST Deihao	9.1568	-27.9654	-12.7340	0.8109	0.2234
ST Dongyang	2.1689	-18.9127	-8.5264	1.2101	0.1904
ST Fangke	17.3389	-70.1883	-9.7050	0.5785	0.1297
ST Furen	43.1257	-26.4360	-11.7530	1.0971	0.0088
ST Gaosheng	20.9308	8.6339	4.3154	2.5588	0.7544
ST Guanfu	3.7516	3.8607	1.2019	0.5642	0.0693
ST Guangyi	15.4920	-9.5168	-5.5269	1.8558	0.5211
ST Guiyi	35.9591	-181.2189	-10.2387	0.4686	0.0196
ST Guoyi	-9.2097	0.9112	0.3145	0.7442	0.4949
ST Haitou	14.4275	6.2299	6.3737	1.1255	0.5584
ST Haiyue	4.0741	2.2894	1.7331	4.3110	2.1075

TABLE 5: Discrete risk levels of corporate financial indicators.

| Securities name | Risk level | | | | |
	Gross profit margin	Roe	Net interest rate	Current ratio	Cash ratio
ST Ankong	3	5	4	4	5
ST Antai	3	2	3	4	5
ST Aoma	2	4	3	3	4
ST Baling	3	5	5	4	4
ST Bolong	2	4	4	2	5
ST Beiwen	2	4	4	3	5
ST Busen	2	5	4	3	5
ST Dayou	2	4	3	4	5
ST Dazhou	2	5	4	5	5
ST Deihao	3	4	4	4	5
ST Dongyang	3	4	3	3	5
ST Fangke	3	5	3	4	5
ST Furen	2	4	4	3	5
ST Gaosheng	2	3	3	2	4
ST Guanfu	3	3	3	4	5
ST Guangyi	3	3	3	3	4
ST Guiyi	2	5	4	5	5
ST Guoyi	3	3	3	4	5
ST Haitou	3	3	3	3	4
ST Haiyue	3	3	3	1	2

TABLE 6: Number of association rules for corporate financial indicators.

Number of association rules		Support threshold		
		0.6	0.7	0.8
Confidence threshold	0.3	86	34	
	0.7			4

5. Conclusions

In the process of enterprise financial risk assessment, use association rules to screen financial risk indicators, find out key indicators that have a greater impact on enterprise financial risks, and improve the enterprise's defense ability against management accounting financial risks. It can be seen from the hierarchical tree of financial risk evaluation indicators established in this paper that enterprises need to systematically prevent financial risks. In the selection of specific financial indicators, enterprises should focus on the following aspects when preventing potential financial risks.

One is in the aspect of enterprise operation risk management. The focus should be on receivables turnover speed and inventory turnover. The turnover of accounts receivable is positively correlated with the flow rate of enterprise assets, which is directly related to the flow of enterprise assets.

The second is in the management of corporate profitability. The focus should be on earnings per share and return on equity, both of which are the focus of external investors. They are directly related to the company's profit rate of return, and they are positively correlated with each other. The higher the profit rate, the higher the earnings per share and the return on equity.

The third is in the evaluation of enterprise growth ability. We should focus on the growth of net profit and the growth rate of total assets, mainly because the growth rate of net profit is closely related to the business performance of the enterprise and is significantly related. Operational performance will primarily affect the evaluation and judgment of the company's future growth potential. The higher the operating efficiency of the enterprise, the greater the growth potential.

Fourth, in the evaluation of corporate solvency, the current ratio and cash ratio should be mainly concerned, both of which can realize a comprehensive judgment on the long-term and short-term solvency of the enterprise.

6. Discussion

In the process of this research, this paper still has some limitations. In the future, it is necessary to conduct in-depth exploration of data mining technology to improve the efficiency of data mining. And continue to explore data mining methods related to enterprise management accounting risk analysis. In addition, it will extend from enterprise management accounting risk analysis to research on enterprise management accounting risk early warning model.

Data Availability

The labeled dataset used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The authors declare no competing interests.

References

[1] Z. Xianhua and X. Guilin, "A review of the current situation and hotspots of management accounting research in my country under the background of "Big Wisdom Shifting to Cloud": visual analysis based on cite space," *Journal of Chongqing University of Arts and Sciences (Social Science Edition)*, vol. 40, no. 5, pp. 85–97, 2021.

[2] L. Feiran, "Management accounting in the digital era——a review of the symposium of the 2021 Chinese accounting society management accounting professional committee," *Finance and Accounting*, vol. 15, pp. 83–85, 2021.

[3] S. F. Haka and D. L. Heitger, "International managerial accounting research: a contracting framework and opportunities," *The International Journal of Accounting*, vol. 39, no. 1, pp. 21–69, 2004.

[4] M. Yan, L. Sun Jian, C. Lu, and J. Liu, "Review and prospect of Chinese management accounting research," *Accounting Research*, vol. 9, pp. 3–12, 2014.

[5] L. Lixia and W. Shishui, "Innovation and application of enterprise management accounting tools in the era of big data—-based on the perspective of cost control," *Friends of Accounting*, vol. 10, pp. 14–19, 2019.

[6] O. Ponisciakova, M. Gogolova, and K. Ivankova, "Calculations in managerial accounting," *Procedia Economics and Finance*, vol. 26, pp. 431–437, 2015.

[7] Y. Xia and G. Ye, "Discussion on the research hotspots of management accounting in my country under the background of big data: quantitative research on graphs based on CiteSpace," *Journal of Finance and Accounting*, vol. 28, pp. 33–38, 2017.

[8] G. Richins, A. Stapleton, T. C. Stratopoulos, and C. Wong, "Big data analytics: opportunity or threat for the accounting profession?," *Journal of Information Systems*, vol. 31, no. 3, pp. 63–79, 2017.

[9] L. Minxia, "The impact of big data technology on management accounting and its countermeasures," *Enterprise Economics*, vol. 1, pp. 103–108, 2018.

[10] M. Mihăilă, "Managerial accounting and decision making, in energy industry," *Procedia-Social and Behavioral Sciences*, vol. 109, pp. 1199–1202, 2014.

[11] H. Shuanglian and W. Ying, "Application of management accounting information under the background of big data," *Finance and Accounting Monthly*, vol. 3, pp. 86–95, 2019.

[12] O. Yigitbasioglu, "Firms' information system characteristics and management accounting adaptability," *International Journal of Accounting and Information Management*, vol. 24, no. 1, pp. 20–37, 2016.

[13] D. Appelbaum, A. Kogan, M. Vasarhelyi, and Z. Yan, "Impact of business analytics and enterprise systems on managerial accounting," *International Journal of Accounting Information Systems*, vol. 25, pp. 29–44, 2017.

[14] M. Al-Maolegi and B. Arkok, "An improved Apriori algorithm for association rules," 2014, http://arxiv.org/abs/1403.3948.

[15] L. Ke, G. Wei, J. Yuyan, and D. Pingping, "Optimization and implementation of parallel FP-growth algorithm based on spark," *Computer Applications and Software*, vol. 9, no. article 9, 2017.

[16] M. Huayang and L. Xingxing, "Research on big data analysis of sales accounting——algorithm based on R language association rules," *Journal of Finance and Accounting*, vol. 16, pp. 116–118, 2018.

[17] S. F. Wamba, S. Akter, A. Edwards, G. Chopin, and D. Gnanzou, "How 'big data' can make big impact: findings from a systematic review and a longitudinal case study," *International Journal of Production Economics*, vol. 165, pp. 234–246, 2015.

[18] S. Zhang, C. Zhang, and Q. Yang, "Data preparation for data mining," *Applied Artificial Intelligence*, vol. 17, no. 5-6, pp. 375–381, 2003.

[19] X. Wu, X. Zhu, G. Q. Wu, and W. Ding, "Data mining with big data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 1, pp. 97–107, 2013.

[20] C. Zhang and S. Zhang, *Association Rule Mining: Models and Algorithms*, Springer, Berlin Heidelberg, 2003.

Research Article

Analysis of Teaching Effect of English Classroom Mind Map Based on a Logistic Regression Model

Yan Deng and **Rui Zhu**

School of Foreign Languages, Jiangxi University of Science and Technology, Ganzhou 341000, China

Correspondence should be addressed to Yan Deng; yandeng@jxust.edu.cn

In order to improve the effect of English classroom teaching, this paper combines the pulse neural network to identify the teaching process of English classroom and conducts real-time supervision of classroom dynamics in combination with the actual situation. Moreover, this paper constructs an intelligent system, uses the recognition method based on the spiking neural network image to collect the real-time dynamics of classroom teaching, and combines the logistic regression model to evaluate the teaching effect. After constructing the system framework, the system process analysis is carried out. This paper combines experimental research to verify the model of this paper. The experimental research results show that the analysis model of teaching effect of mind map in English classroom based on the logistic regression model proposed in this paper has good results, which also verifies that the mind map teaching can play an important role in English classroom teaching.

1. Introduction

The teaching goal is the final result expected to be achieved by the teaching activity. Generally speaking, the teaching goal is a teaching result caused by a subjective wish. After the teaching activity is completed, it is a specific description of the behavioral state that the student needs to achieve [1]. In the process of English teaching, the reasonable degree of teaching objectives directly determines the teaching effect. The teaching goal of high school English courses is to develop the ability of autonomous learning and cooperative learning, form effective English learning strategies, and cultivate students' comprehensive language ability [2].

Students learn and master the basic knowledge of five aspects including pronunciation, vocabulary, grammar, functions, and topics. Knowledge is one of the important components of English ability, and it is the basis for effective use of language skills. Therefore, the effective use of mind maps in teaching can realize the organic connection between a series of related knowledge such as words, sentences, and grammar. Through the use of mind maps as a way of think-

ing for thinking and learning, it helps students to combine the actual practice of life. Moreover, in the process of drawing homework for students' real-life mind maps, it helps students to achieve effective mastery of basic theoretical knowledge [3].

Effectively master the four skills of listening, speaking, reading, and writing, and have the comprehensive ability to use these four skills. The high school English course also emphasizes that on the basis of developing students' comprehensive language ability, it focuses on improving students' ability to acquire information, deal with problems, analyze problems, and solve problems in English. Students should use a large number of special or comprehensive language practices. A comprehensive language ability lays the language foundation for actual communication. Therefore, from the perspective of learning strategies, listening, speaking, reading, and writing are not only learning content but also learning methods. Mind mapping is a graphical tool that is very convenient for language input and output, and there are nodes as the connection between knowledge. In actual teaching, applying mind maps to the four aspects of vocabulary, grammar, reading, and writing

can realize the overall improvement of students' practical use of English, information acquisition, problem analysis, and problem solving.

Let students gain a sense of satisfaction and accomplishment and have the awareness of solving practical problems. From the essence of education, learning is not only for acquiring knowledge but also for achieving awareness or emotional satisfaction. In the teaching process, teachers should guide students to transform interest into learning motivation, build self-confidence, and shape healthy personality. The purpose of high school English classroom teaching is not only to enable students to master the basic comprehensive abilities of listening, speaking, reading, and writing but also to enable students to learn from analogy to solve practical problems. English teaching can also be used as an alternative patriotism teaching, which can enhance the spirit of patriotism and the sense of national mission on the basis of further broadening the international horizon. In the process of teaching, there are often foreign cultural and historical articles. In reading teaching, using mind maps to analyze and compare the culture and history of other countries not only is a good way to learn knowledge but also can effectively improve students' thinking. The ability to solve the actual problems of the country can effectively cultivate the innovative spirit of students.

This article combines the logistic regression model to evaluate the teaching effect of English classroom mind map, which provides a theoretical reference for the subsequent reform of English classroom teaching.

2. Related Work

Literature [4], through the study of neurophysiological science and the classification and comparison of the connections between all things, found the radioactive phenomenon between them and analyzed the writing habits of human beings and those who have learning disabilities in humans. After training and observation, he finally created a mind map, a tool to help human thinking. Literature [5] pointed out: "This research result can be applied to information processing, that is, to summarize and synthesize research information, and can also be used to assist people in learning and research work such as thinking about complex problems and in-depth analysis of characterization information. Specifically speaking, Mind map users can comprehensively use information representation tools including pictures, graphics and words, etc., to clearly and systematically display the logical hierarchical relationship and thinking sequence relationship between concepts, and realize and improve the visualization of concept representation."

By reading the relevant literature of foreign researchers, I learned that the use of mind mapping tools in the education field will benefit both educators and learners. The most important thing is the creativity of the two in educational learning activities. Get a great improvement. With this ability of independent creation, the teaching effect of the school will be greatly improved [6]. I understand that the research scope of mind mapping by foreign scholars is getting wider

and wider and that mind mapping is applied and researched in more and more countries. Their research goals are not only limited to the basic concepts and connotations of mind maps but also more and more practical researches on the application of mind maps to subject teaching .

In these documents related to the study of mind maps, some authors mainly discussed the basic concepts and connotations of mind maps [7]. Literature [8] proposed that mind mapping is a tool that combines text and graphics. It was also pointed out that the information presented on the mind map is all related and must be drawn in a certain order. There are also many documents mainly for research on the meaning of mind mapping in foreign language classroom teaching. Researchers proposed that the use of mind mapping as a tool in foreign language classrooms to assist teaching has many benefits. In such a class, students are more efficient in language learning. Students no longer use read-only texts; they can also use mind maps to freely imagine, put forward their own unique insights, and deepen their understanding of language and text [9]. At the same time, communication between students can help them to cooperate fully. The use of this tool in foreign language classrooms is also beneficial to teachers. It can help teachers sort out their lesson preparation ideas, improve the quality and speed of teacher preparation, and help teachers formulate reasonable teaching plans. With the help of mind maps, it can also reasonably and effectively evaluate the level of knowledge learned by students [10]. There are also some documents that mainly focus on the issue of mind mapping to improve students' learning styles. Researchers believe that mind mapping, a teaching aid tool, can help students fully realize their dominant position in learning [11]. Students' learning in the classroom can no longer be passive acceptance, but active and autonomous learning. With the support of this tool, students will also actively learn to communicate with others and exchange views and opinions with each other. Their learning method is no longer single, but in multiple ways [12]. According to the research theme, when relevant experts conduct research on mind mapping, they mainly conduct research on English vocabulary, writing, reading, and oral expression [13]. In these aspects of research, we found that most of the research on mind mapping is to improve the memory of English vocabulary; secondly, there are also many researches on the use of mind mapping as a tool for English reading teaching [14]. There are also researches involved in spoken English and translation. These research themes affirm the feasibility and significance of mind mapping in English teaching [15]. These studies are mainly based on individual topics as research clues and clearly put forward that they are only studying the application of mind mapping in a certain professional aspect of English teaching and do not regard mind mapping as a complete teaching model, from the perspective of education management of the whole school. Set out to study its construction, implementation, evaluation, and management [16]. From the perspective of the English learning objects of the application of mind maps, mind maps can be used in various stages of learning such as elementary school, middle school, and university [17]. However, mind maps are mostly used in foreign
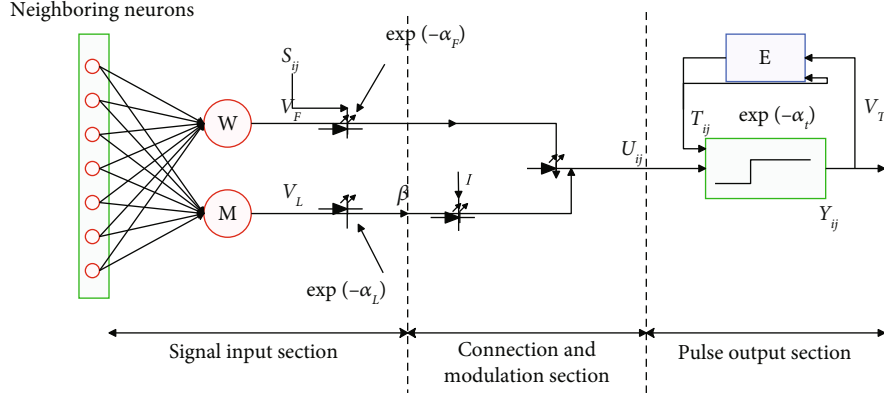
FIGURE 1: Basic structure diagram of PCNN model, quoted from the literature.

language teaching in colleges. There are few researches on the application of mind maps in English teaching to high school students and the least research on primary school students. The research in elementary school is mainly for vocabulary teaching [18]. Mind maps can be more widely used in English teaching, which is convenient for cultivating students' good language thinking from the beginning of the classroom, helping students to establish scientific thinking habits, instead of just using mind maps to help students memorize a few more English vocabulary [19].

3. Recognition of English Classroom Teaching Features Based on a Pulse Coupled Neural Network

This paper combines pulse-coupled neural networks to identify English classroom teaching features, which is convenient to provide data support for the subsequent teaching analysis of English classroom mind maps.

A pulse coupled neural network is a third-generation artificial neural network that is different from traditional neural networks. It is a single-layer two-dimensional locally connected feedback network proposed by the optimization and improvement of the Eckhorn model and the Rybak model. The basic model of PCNN consists of three parts: receiving domain, modulation domain, and pulse generation domain. Its structure diagram is shown in Figure 1.

The discrete equation corresponding to the pulse coupled neural network model is shown in

$$F_{ij}(n) = e^{-\alpha_F} F_{ij}(n-1) + V_F \sum_{k,l} M_{ijkl} Y_{kl}(n-1) + S_{ij}, \quad (1)$$

$$L_{ij}(n) = e^{-\alpha_L} L_{ij}(n-1) + V_L \sum_{k,l} W_{ijkl} Y_{kl}(n-1), \quad (2)$$

$$U_{ij}(n) = F_{ij}(n)\big(1 + \beta L_{ij}(n)\big), \quad (3)$$

$$E_{ij}(n) = e^{-\alpha_E} E_{ij}(n-1) + V_E Y_{ij}(n-1), \quad (4)$$

$$Y_{ij}(n) = \begin{cases} 1, & U_{ij}(n) > E_{ij}(n), \\ \\ 0, & \text{oaE\^u.} \end{cases} \quad (5)$$

It can be seen from the model structure diagram that the pulse coupled neural network model is divided into three parts. Among them, formulas (1) and (2) represent the signal receiving part, $F_{ij}(n)$ is the external stimulus, and $L_{ij}(n)$ is the link input from the peripheral neurons. Formula (3) represents the connection modulation part. Formulas (4) and (5) represent the pulse output part, and the generated modulation signal $U_{ij}(n)$ is compared with the threshold $E_{ij}(n)$ to obtain the pulse signal $Y_{ij}(n)$. In addition, there are four types of parameters in formulas (1)–(5):

(1) The time attenuation parameters are as follows: α_F, α_L, α_E. These three parameters control the attenuation speed of the input zone F, the connection input zone L, and the threshold $E_{ij}(n)$, respectively. α_E is particularly important, because its size directly affects the running time of PCNN. The larger the α_E, the slower the decrease of $E_{ij}(n)$, the longer the running time of the corresponding model, and vice versa

(2) The magnification factor are as follows: V_F, V_L, V_E. Among them, V_F and V_L control the size of external stimuli, while V_E affects the firing cycle of neurons

(3) The connection coefficient is β. β is a parameter in $U_{ij}(n)$. The larger the value, the greater the coupling degree of the neuron to the surrounding information, and it will also affect the ignition cycle of the central neuron

(4) The weight matrix are W and M. They, respectively, represent the connection strength between the input domain F and the connection input domain L and its surrounding neurons. The size is generally related to the distance between the neurons

When processing the image with the pulse-coupled neural network model, the pixels of the image and the neurons

of the PCNN form a two-dimensional neural network in one-to-one correspondence. First, in the signal input part, the receiving domain F receives the external stimulus S_{ij} (that is, each pixel of the image) and the feedback $F(n-1)$ from the previous iteration, which is connected to the input domain L to receive the stimulation of surrounding neurons. The signals F and L obtained in the receiving domain enter the connection modulation part together, and the internal activity item (that is, modulated signal) $U_{ij}(n)$ is obtained after modulation. After entering the pulse output part, the modulation signal $U_{ij}(n)$ will be compared with the dynamic threshold $E_{ij}(N)$ generated in the pulse generation domain. If $U_{ij}(n) > E_{ij}(n)$, the pulse generator will start and the neuron ignition output $Y_{ij} = 1$; otherwise, the neuron nonignition output $Y_{ij} = 0$. After performing the same operation on each pixel, the final result Y is obtained, and the above operation is repeated until the conditions are met. According to whether neurons are interconnected, it can be divided into two working modes: uncoupled and connected and coupled.

3.1. No Coupling Connection. Uncoupled connection means that the neurons between PCNNs have no connection relationship. At this time, the connection coefficient is $\beta = 0$, and the amplification coefficient is $V_F = 0$ in the feedback input domain. This ensures that each neuron of PCNN runs independently during work. At this time, the corresponding mathematical model formulas (1)–(5) are transformed into

$$F_{ij}(n) = e^{-\alpha_F} F_{ij}(n-1) + S_{ij}, \tag{6}$$

$$U_{ij}(n) = F_{ij}(n), \tag{7}$$

$$E_{ij}(n) = e^{-\alpha_E} E_{ij}(n-1) + V_E Y_{ij}(n-1), \tag{8}$$

$$Y_{ij}(n) = \begin{cases} 1, & U_{ij} > E_{ij}(n-1), \\ & \vdots? \\ 0, & \text{òaEû.} \end{cases} \tag{9}$$

In formulas (6)–(9), the feedback input F and the dynamic threshold E are both zero in the initial state of the PCNN. Generally, V_E takes a larger value, the time n is discrete, and the neuron ij ignites, and we can get

$$U_{ij}(0) = F_{ij}(n) = S_{ij}, \tag{10}$$

$$Y_{ij}(0) = 1, S_{ij} > 0. \tag{11}$$

Substituting (11) into (8), we can get

$$E_{ij}(0) = V_E. \tag{12}$$

The neuron ij ignites at $n = 0$ and the dynamic threshold $E_{ij}(0) = V_E$, $V_E \gg >S_{ij}$ at this time. Therefore, the neuron will not ignite immediately after the zero time. It can be
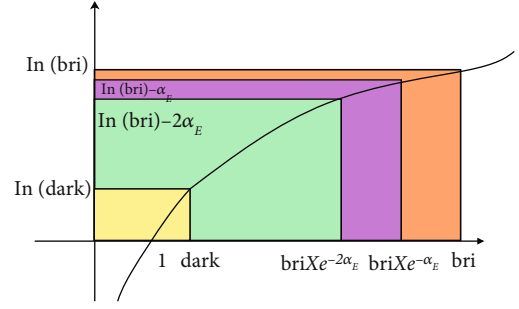
Figure 2: The relationship between PCNN and logarithmic transformation.

deduced that the neuron state at $n = 1$ is as follows:

$$\begin{aligned} F_{ij}(1) &= e^{-\alpha_F} S_{ij} + S_{ij}, \\ U_{ij}(1) &= F_{ij}(1), \\ Yij(1) &= 0, \\ E_{ij}(1) &= e^{-\alpha_E} V_E. \end{aligned} \tag{13}$$

From the state value of the neuron at $n = 0$ and $n = 1$, the general formula can be derived:

$$\begin{aligned} U_{ij}(1) &= S_{ij}(1 + e^{-\alpha_F} + \cdots + e^{-n\alpha_F}) = V_E e^{-(n-1)\alpha_E}, \\ E_{ij}(n) &= e^{-n\alpha_E} V_E, \\ Y_{ij}(n) &= 0. \end{aligned} \tag{14}$$

From this, the first and second ignition timings can be calculated:

$$\begin{aligned} n_1 &= 1 + \frac{1}{\alpha_E} \ln \frac{V_E}{c S_{ij}}, \\ n_2 &= 1 + \frac{1}{\alpha_E} \ln \frac{V_E}{c S_{ij}} + \frac{1}{\alpha_E} \ln \frac{c S_{ij} + V_E}{c' S_{ij}}. \end{aligned} \tag{15}$$

It is deduced that the neuron ignition at time m meets the conditions:

$$n(m) = 1 + n_1 + m n_2 = 1 + \frac{1}{\alpha_E} \ln \frac{V_E}{c S_{ij}} + m \frac{1}{\alpha_E} \ln \frac{c S_{ij} + V_E}{c' S_{ij}}. \tag{16}$$

From formula (16), it can be concluded that the neuron firing cycle of PCNN is

$$T_{ij} = n(m) - n(m-1) = \frac{1}{\alpha_E} \ln \frac{\left((1 - e^{-(n_1+1)\alpha_F})/(1 - e^{-\alpha_F})\right)S_{ij} + V_E}{\left((1 - e^{-(n_1+1)\alpha_F})/(1 - e^{-\alpha_F})\right)S_{ij}}. \tag{17}$$

It can be seen from formula (17) that when there is no coupling and the surrounding neurons have no influence, the ignition cycle has a very strong relationship with the

(a) Two-dimensional wavelet transform   (b) AWT (2,1) transformation
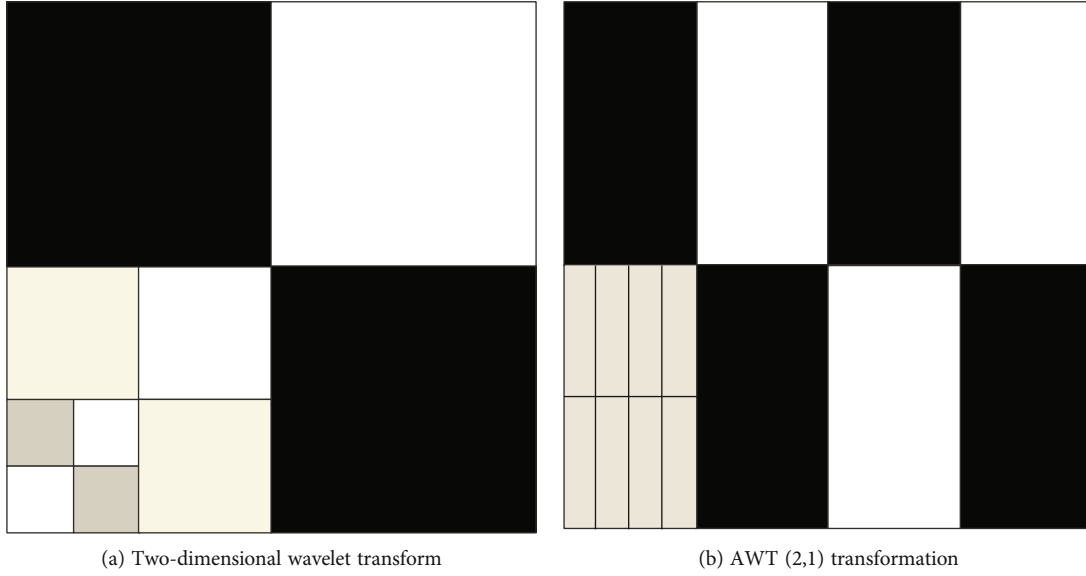
FIGURE 3: Comparison of two-dimensional wavelet transform and AWT transform.

external stimulus S_{ij}. The larger the S_{ij}, the shorter the ignition period, and the smaller the S_{ij}, the longer the ignition period.

3.2. Coupling Connection.

The coupling connection of PCNN is the key to transfer information between neurons, as shown in formulas (1)–(5). It uses the link input L to modulate the feedback input F. We assume that there are two interconnected neurons ij and kl, where $S_{ij} > S_{kl}$, and other parameters are the same. If two neurons fire at the same time at $t = 0$, the firing cycle of the neuron mentioned above; that is, from formula (17), the second point of neuron ij will precede neuron kl after the second point. At this time, as a coupling neuron of neuron kl, neuron ij transmits ignition information to neuron U_{kl} through modulation signal U_{kl}, that is,

$$U_{kl} = S_{kl}(1 + \beta L_{kl}). \tag{18}$$

If $U_{kl} > S_{ij}$, it can be calculated that neuron kl will pre-ignite. It is obtained that the coupled neuron kl of neuron ij will ignite with the ignition of neuron ij when it satisfies

$$S_{kl} \in \left[\frac{S_{ij}}{1 + \beta L_{kl}}, S_{ij} \right]. \tag{19}$$

In image processing, there are generally 9 neuron couplings around a neuron, and the neurons meeting the capture conditions will be fired in advance. For β and L, the larger the value, the wider the capture range.

In order to make digital image enhancement more convenient and effective, the pulse-coupled neural network model is simplified, and its discrete mathematical description is

$$F_{ij}(n) = S_{ij}(n), \tag{20}$$

$$L_{ij}(n) = e^{-\alpha_L} L_{ij}(n-1) + V_L \sum_{k,l} W_{ijkl} Y_{kl}(n-1), \tag{21}$$

$$U_{ij}(n) = F_{ij}(n)\left(1 + \beta L_{ij}(n)\right), \tag{22}$$

$$E_{ij}(n) = e^{-\alpha_E} E_{ij}(n-1) + V_E Y_{ij}(n-1), \tag{23}$$

$$Y_{ij}(n) = \begin{cases} 1, & U_{ij}(n) > E_{ij}(n), \\ & \raisebox{-3pt}{\textbf{?}} \\ 0, & \text{oaEû.} \end{cases} \tag{24}$$

The image pixels are stretched through logarithmic transformation to achieve the effect of image enhancement. The output of the pulse coupled neural network also has a logarithmic transformation relationship, as shown in Figure 2.

Among them, the abscissa is the gray value of the image pixel, and the ordinate is the corresponding logarithmic result. Bri is the largest pixel value, and its corresponding dark is the smallest pixel value. Assuming that the initial value of E_{ij} is $E_{ij}(0) =$ bri, the neuron ij is in a state of no ignition at the beginning, and the output is 0. After the first iteration of PCNN, the threshold $E_{ij}(1) =$ bri $\times e^{-\alpha_E}$ at time 1 can be obtained from formula (23). At this time, all neurons with gray values are between bri $\times e^{-\alpha_E}$ and bri ignite. At this time, the firing excitation of neurons is set to ln (bri). Next, perform the second iteration, and we can get $E_{ij}(2) =$ \$bri\$ $\times e^{-2\alpha_E}$. The ignition gray value range of the neuron becomes between bri $\times e^{-2\alpha_E}$ and bri $\times e^{-\alpha_E}$, and the gray excitation value of this segment of neuron ignition is denoted as ln (\$bri\$) $- \alpha_E$. It keeps iterating until the smallest gray value dark of the entire image is also ignited. In this way, we can define the gray-scale excitation value of neuron ignition at different moments as the image that has been
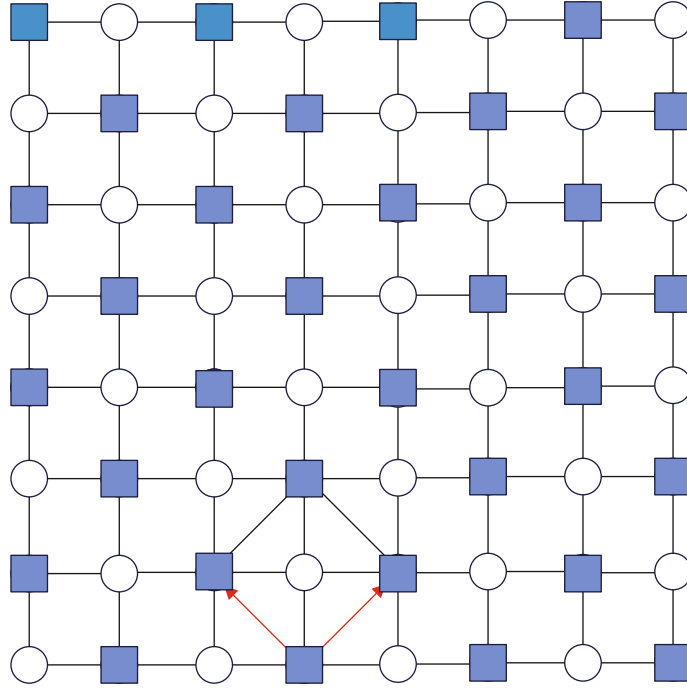
FIGURE 4: Schematic diagram of integer lattice decomposition.
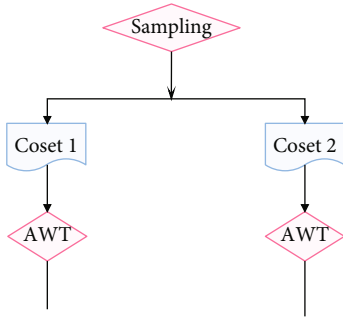
FIGURE 5: Directionlet transformation structure diagram.

enhanced, as in

$$\text{enh}S_{ij} = \ln\,(\text{bri}) - \alpha_E(k-1). \tag{25}$$

In the formula, $\text{enh}S_{ij}$ is the gray value of the enhanced image, and k is the ignition moment of neuron ij. Formula (25) is logarithmic, and each ignition moment of neuron ij has a one-to-one correspondence with the enhanced image value. In fact, we have assumed that the input of each neuron is not affected by the previous output of surrounding neurons, which means that formula (25) is the logarithmic transformation value of the grayscale image. However, due to the influence of $\alpha_E(k-1)$, the result is different from the true logarithmic relationship.

The traditional wavelet transform is isotropic; that is, the same number of transformations is performed on the horizontal and vertical directions at the same scale. However, when dealing with two-dimensional images, the ideal results

cannot be obtained. Anisotropic wavelet (AWT) refers to the transformation of horizontal and vertical directions at different times under the same scale. For example, in an AWT, n_1 one-dimensional wavelet transforms are performed in the horizontal direction, and n_2 one-dimensional wavelet transforms are performed in the vertical direction. If $n_1 = n_2$, it is the original isotropic; if $n_1 \neq n_2$, it is anisotropic wavelet transform. The extensibility of v is determined by its opposite sex rate $\rho = n_1/n_2$.

Figure 3(a) is a general two-dimensional wavelet transform; the number of transformations in the horizontal and vertical directions is the same. However, the AWT (2,1) transform in Figure 3(b) performs two one-dimensional wavelet transforms in the horizontal direction and only one one-dimensional wavelet transform in the vertical direction.

The so-called integer lattice refers to a set of points composed of two sets of linearly independent vectors d_1 and d_2 to form an integer lattice Λ, and d_1 and d_2 must be integers, $\Lambda \in Z^2$. The formula is expressed as

$$\Lambda = \{c_1 d_1 + c_2 d_2 | c_1, c_2 \in Z\}. \tag{26}$$

However, Λ can be represented by a non-unique generator matrix M_Λ, and it can be set up as

$$M_\Lambda = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}, a_1, a_2, b_1, b_2 \in Z, c = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}. \tag{27}$$

Then, Λ can be expressed as $\Lambda = cM_\Lambda$. This means that

Λ can be determined by M_Λ and the entire plane Z^2 can be divided into $|\det M_\Lambda|$ (the absolute value of the M_Λ determinant) cosets about the lattice Λ. Each coset corresponds to a displacement vector $S_k = (S_{k1}, S_{k2}), k = 0, 1, 2 \cdots, |\det (M_\Lambda)| - 1$. When performing the Directionlet transformation, the transformation direction is the d_1 direction with a slope of $r_1 = b_1/a_1$, and the queue direction is the d_2 direction with a slope of $r_2 = b_2/a_2$. Various combinations of d_1 and d_2 directions are different, so that different sampling matrices are obtained. Therefore, image information of anisotropic images with different directions is obtained.

For the Directionlet transformation of the image, the first choice is to use the sampling matrix M_Λ to sample the image to obtain $|\det (M_\Lambda)|$ cosets. Then, these cosets undergo an anisotropic wavelet transform in the transform direction and the queue direction, that is, AWT. In this way, the Directionlet sparse representation of the image can be obtained, and the wavelet transform will not produce dense wavelet coefficients due to its same direction and limited direction. This is of great help to our subsequent processing work.

Next, this paper uses Figure 4 to advance one to explain how cosets are decomposed.

As shown in Figure 4, the sampling matrix is selected as

$$M_\Lambda = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \tag{28}$$

and the slope of the transformation direction is $r_1 = 1$, and the slope of the queue direction is $r_2 = -1$, which means that the transformation is performed along the $\pm 45°$ direction. The whole grid space is divided into two cosets of black point and white point, the displacement vector of black point is $S_0 = [0, 0]$, and the displacement vector of white point is $S_1 = [0, 1]$. The structure diagram of the entire Directionlet transformation is shown in Figure 5.

The time attenuation constant is α_L, α_E, the amplification factor is V_L, V_E, the connection factor is β, and the weight matrix is W. The meanings and functions of these parameters have been introduced in the section of this paper, so I will not discuss them here.

In the pulse-coupled neural network image enhancement in this paper, because all the gray values are required to be fired only once, the final enhanced image is determined by its ignition time k. Therefore, the threshold amplification factor V_E must be large enough so that after the neuron ij is ignited, it will not reignite for a long time (at least until the neuron with the smallest external stimulation, that is, the pixel with the smallest gray value has also ignited).

Correspondingly, the time attenuation coefficient α_E of the threshold value needs to be small enough to ensure that the attenuation speed of the threshold value $E_{ij}(n)$ is slow enough, so that different gray levels of the image have different output results at different ignition moments. If the attenuation coefficient is very small, it will cause a lot of redundancy in the system, so we can set

$$\alpha_E \le \ln (\text{bri}|(\text{bri} - 1)). \tag{29}$$

β controls the relationship between input and external stimuli. The larger the value, the more the neuron is affected by nearby neurons and the easier it is to be captured. Therefore, the more complex the image of a certain area, the less the influence of the central neuron should be, and the simpler the area image, the greater the influence of the surrounding neurons and the neurons should be captured in advance. This will also avoid the problems of overcapture in areas with small differences and undercapture in areas with large differences due to the same connection coefficient β. Therefore, the standard deviation will be used as the β value in this paper, that is,

$$\beta = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_{ij} - \mu)^2}. \tag{30}$$

The standard deviation is the quantity that reflects the degree of spatial dispersion within a certain group. Here, it represents the difference between a neuron and its neighboring neurons. As shown in formula (30), N represents the total number of neurons in a certain neighborhood of the image, x_{ij} represents the gray value of each neuron in this neighborhood, and μ is the average gray value of the neurons in the neighborhood. The larger the value of β, the greater the difference between neurons in the neighborhood. Therefore, a larger connection coefficient is required to make it captured. The smaller the value of β, the neighborhood can be captured without the need for such a large connection coefficient.

The weight matrix W represents the degree of connection between the neuron and surrounding neurons. Generally, we set its size as the reciprocal of the square of the Euclidean distance of the neuron, that is,

$$W = \begin{bmatrix} \frac{1}{2} & 1 & \frac{1}{2} \\ 1 & 0 & 1 \\ \frac{1}{2} & 1 & \frac{1}{2} \end{bmatrix}. \tag{31}$$

This article will use image contrast (C, Contrast), peak signal-to-noise ratio (PSNR, Peak Signal-to-Noise-Rate) and information entropy H as the evaluation criteria to judge the enhancement effect. The formula for image contrast is

$$C = \sqrt{\frac{\sum_{(x,y)} (I(x,y) - v)^2}{x \times y}}. \tag{32}$$

Among them, $I(x, y)$ represents the gray value of a certain pixel of the image, μ is the average gray value of the image, and x and y are the pixel values of the image. The peak signal-to-noise is shown in

$$\text{PSNR} = 10 \times \log \left(\frac{255^2}{\text{MSE}}\right). \tag{33}$$

MSE is the mean square error, that is, the average of the

TABLE 1: Image enhancement evaluation indicators of different algorithms.

| | | Algorithm | | |
Index	Lena original image	Logarithmic enhancement	Wavelet transform enhancement	Algorithm enhancements in this paper
Contrast C	274.60	326.86	544.62	624.53
Peak signal-to-noise ratio PSNR	40.11	40.06	40.03	40.153
Information entropy H	6.638	6.479	7.175	7.348
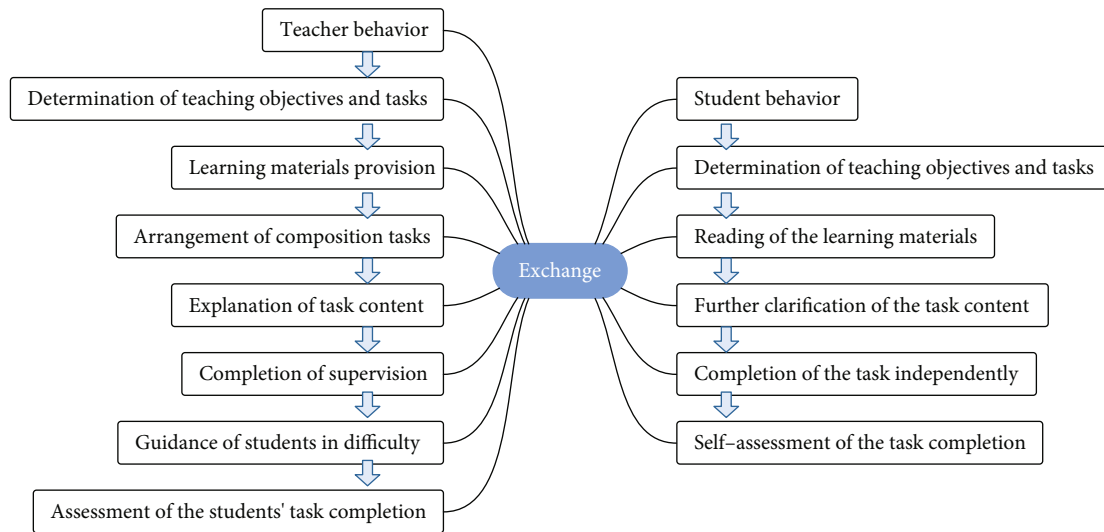
FIGURE 6: The design of the teaching process of personal mind map.
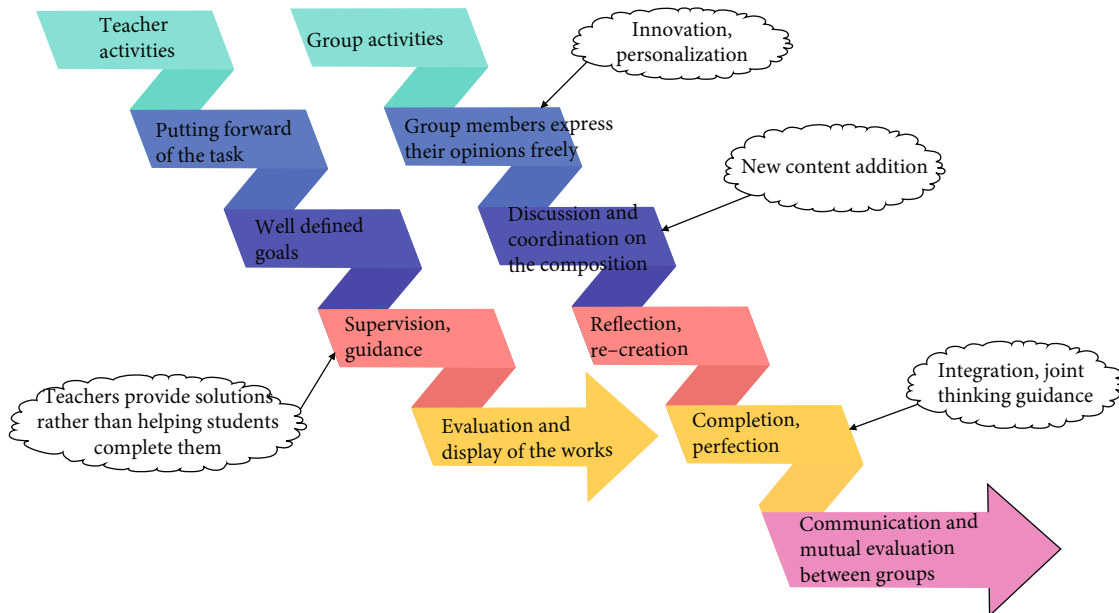
FIGURE 7: The strategy of drawing collective mind maps.
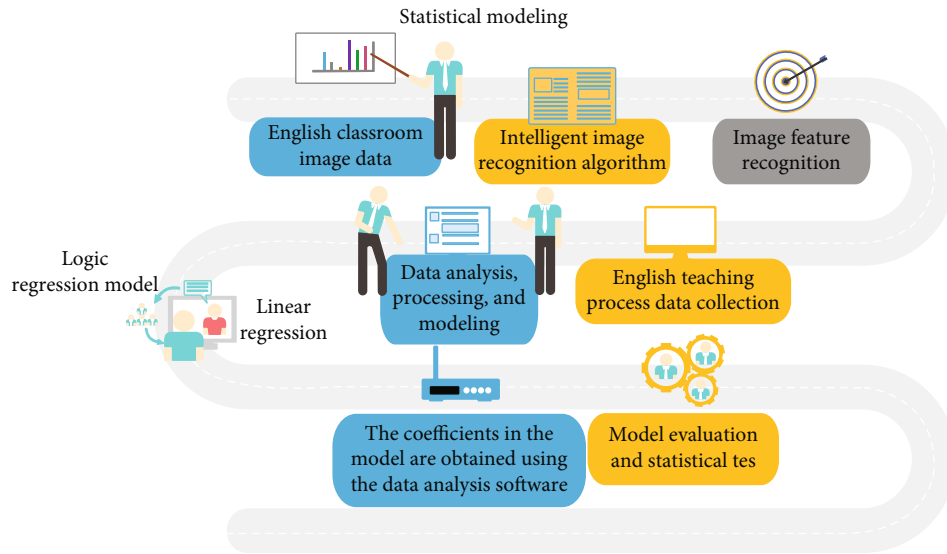
FIGURE 8: An analysis model of the English classroom mind mapping teaching effect based on the logistic regression model.

TABLE 2: The effect of feature recognition in English classroom teaching.

Number	Teaching feature recognition	Number	Teaching feature recognition	Number	Teaching feature recognition
1	94.51	20	94.06	39	93.09
2	91.61	21	92.31	40	93.20
3	88.55	22	95.78	41	93.15
4	91.95	23	87.95	42	88.21
5	88.07	24	95.62	43	93.87
6	93.84	25	90.12	44	91.84
7	92.28	26	92.49	45	92.29
8	87.31	27	88.48	46	88.48
9	95.51	28	93.80	47	95.00
10	93.27	29	90.90	48	89.54
11	92.72	30	91.55	49	89.41
12	88.39	31	91.20	50	95.35
13	92.86	32	92.04	51	88.78
14	91.96	33	94.51	52	95.99
15	88.92	34	89.89	53	87.49
16	93.73	35	88.76	54	94.64
17	92.73	36	88.70	55	90.53
18	90.40	37	89.73	56	89.12
19	88.66	38	89.27		

squared errors of each data. The formula for the information entropy of the image is

$$H_1 = -\sum_{i=0}^{255} P_i \ln P_i. \tag{34}$$

P_i is the probability of a certain gray value of the image. Table 1 shows the data results of these three evaluation indicators in the enhanced images of different algorithms.

It can be seen from the table that although logarithmic enhancement increases the contrast, the information entropy decreases. Although the wavelet transform method has achieved certain results, the peak signal-to-noise ratio drops slightly. The algorithm in this paper has achieved the greatest enhancement in contrast and information entropy, and there is no drop in the peak signal-to-noise ratio. Therefore, it can be judged that the algorithm in this paper is better than the other two algorithms.

Teaching feature recognition
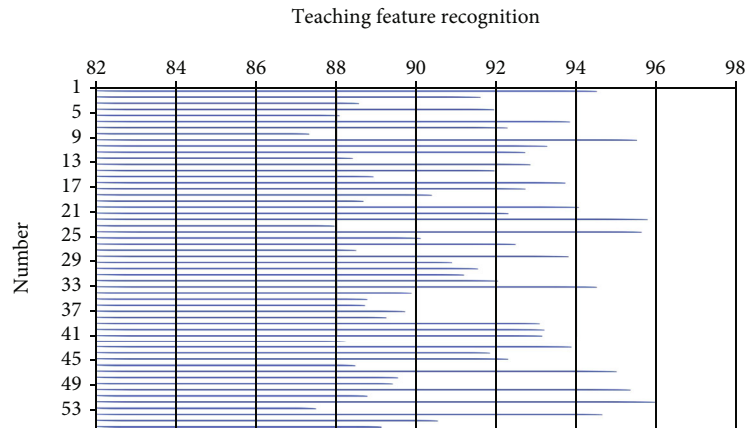
FIGURE 9: Statistical diagram of the effect of real-time feature extraction in English class.
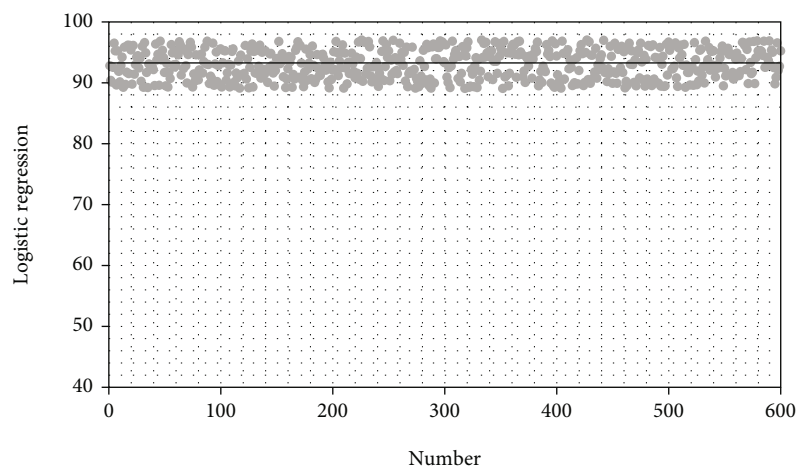
FIGURE 10: The effect evaluation of the analysis model of teaching effect of mind map in English classroom based on the logistic regression model.

4. Analysis of the Teaching Effect of Mind Map in English Class Based on a Logistic Regression Model

In the teaching process, different teaching media and teaching methods are used because of the different topics selected, and different mind map drawing processes are formed. There are two main strategies for mind map construction: personal mind map drawing strategy and collective mind map drawing strategy. The two drawing strategies, the teacher's activity and the student's activity process, are, respectively, introduced in the form of a flowchart below. The personal mind map is drawn as shown in Figure 6.

In this model, there are two parts: teacher activities and group activities, as shown in Figure 7.

This article uses the third part of the image recognition method based on the pulse neural network to identify the classroom teaching process, collect the real-time dynamics of classroom teaching, and evaluate the teaching effect with the logistic regression model. The analysis model of the English classroom mind mapping teaching effect based on the logistic regression model is shown in Figure 8.

After constructing the above intelligent model, this paper evaluates the feature recognition effect of intelligent English classroom teaching in this paper through the simulation platform, and the results are shown in Table 2 and Figure 9.

The above research shows that the English classroom feature recognition based on spiking neural network proposed in this paper has a good effect. On this basis, this paper evaluates the analysis model of the English classroom mind mapping teaching effect based on the logistic regression model. The results obtained through logistic regression are shown in Figure 10.

From the above research, the analysis model of teaching effect of mind map in English classroom based on the logistic regression model proposed in this paper has good results, which also verifies that the mind map teaching can play an important role in English classroom teaching.

5. Conclusion

As an effective thinking technology and cognitive tool, the mind map can not only guide and extend the thought

process but also record the thought process and carry out a clearer visual expression. At the same time, when establishing an individual's self-knowledge network, it is also conducive to the exchange and transmission of personal knowledge and information and affects the individual's cognition. From the previous discussion, we know that English learning is not only a language activity but also a thinking activity. Moreover, it is a procedural activity in which learners use their background knowledge, language skills, and thinking skills to communicate through words or oral language. This article combines the logistic regression model to evaluate the teaching effect of mind map in English classroom. The research results show that the analysis model of teaching effect of mind map in English classroom based on the logistic regression model proposed in this paper has good results, which also verifies that the mind map teaching can play an important role in English classroom teaching.

Data Availability

The labeled dataset used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The authors declare no competing interests.

Acknowledgments

References

[1] J. C. Richards, "Teaching English through English: proficiency, pedagogy and performance," *RELC Journal*, vol. 48, no. 1, pp. 7–30, 2017.

[2] L. Susanty, Z. Hartati, R. Sholihin, A. Syahid, and F. Y. Liriwati, "Why English teaching truth on digital trends as an effort for effective learning and evaluation: opportunities and challenges: analysis of teaching English," *Linguistics and Culture Review*, vol. 5, no. S1, pp. 303–316, 2021.

[3] A. S. Fatimah, S. Santiana, and Y. Saputra, "Digital comic: an innovation of using Toondoo as media technology for teaching English short story," *English Review: Journal of English Education*, vol. 7, no. 2, pp. 101–108, 2019.

[4] B. Ayçiçek and T. Yanpar Yelken, "The effect of flipped classroom model on students' classroom engagement in teaching English," *International Journal of Instruction*, vol. 11, no. 2, pp. 385–398, 2018.

[5] N. Guzachchova, "Zoom technology as an effective tool for distance learning in teaching English to medical students," *Бюллетень науки и практики*, vol. 6, no. 5, pp. 457–460, 2020.

[6] M. S. Hadi, "The use of song in teaching English for junior high school student," *English Language in Focus (ELIF)*, vol. 1, no. 2, pp. 107–112, 2019.

[7] A. Mahboob, "Beyond global Englishes: teaching English as a dynamic language," *RELC Journal*, vol. 49, no. 1, pp. 36–57, 2018.

[8] H. Sundari, "Classroom interaction in teaching English as foreign language at lower secondary schools in Indonesia," *Advances in Language and Literary Studies*, vol. 8, no. 6, pp. 147–154, 2017.

[9] A. Gupta, "Principles and practices of teaching English language learners," *International Education Studies*, vol. 12, no. 7, pp. 49–57, 2019.

[10] B. S. Abdelshaheed, "Using flipped learning model in teaching English language among female English majors in Majmaah University," *English Language Teaching*, vol. 10, no. 11, pp. 96–110, 2017.

[11] T. A. Ashraf, "Teaching English as a foreign language in Saudi Arabia: struggles and strategies," *International Journal of English Language Education*, vol. 6, no. 1, pp. 133–154, 2017.

[12] O. Tarnopolsky, "Principled pragmatism, or well-grounded eclecticism: a new paradigm in teaching English as a foreign language at Ukrainian tertiary schools?," *Advanced Education*, vol. 5, no. 10, pp. 5–11, 2018.

[13] N. I. Sayakhan and D. H. Bradley, "A nursery rhymes as a vehicle for teaching English as a foreign language," *Journal of University of Raparin*, vol. 6, no. 1, pp. 44–55, 2019.

[14] D. A. W. Nurhayati, "Students' perspective on innovative teaching model using Edmodo in teaching English phonology: " a virtual class development"," *Dinamika Ilmu*, vol. 19, no. 1, pp. 13–35, 2019.

[15] A. B. Rinekso and A. B. Muslim, "Synchronous online discussion: teaching English in higher education amidst the covid-19 pandemic," *JEES (Journal of English Educators Society)*, vol. 5, no. 2, pp. 155–162, 2020.

[16] A. S. N. Agung, "Current challenges in teaching English in least-developed region in Indonesia," *SOSHUM: Jurnal Sosial Dan Humaniora*, vol. 9, no. 3, pp. 266–271, 2019.

[17] M. A. Saydaliyeva, E. B. Atamirzayeva, and F. X. Dadaboyeva, "Modern methods of teaching English in Namangan state university," *International Journal on Integrated Education*, vol. 3, no. 1, pp. 8-9, 2020.

[18] L. B. Kelly, "Preservice teachers' developing conceptions of teaching English learners," *TESOL Quarterly*, vol. 52, no. 1, pp. 110–136, 2018.

[19] A. Coşkun, "The application of lesson study in teaching English as a foreign language," *İnönü Üniversitesi Eğitim Fakültesi Dergisi*, vol. 18, no. 1, pp. 151–162, 2017.

Research Article

Infrared and Visible Image Fusion in a Multilevel Low-Rank Decomposition Framework Based on Guided Filtering and Feature Extraction∗

Chao Fang [ID],[1] **Xin Feng** [ID],[1,2] **Haifeng Gong** [ID],[2] **and Xicheng Lou** [ID][1]

[1]*School of Mechanical Engineering, Key Laboratory of Manufacturing Equipment Mechanism Design and Control of Chongqing, Chongqing Technology and Business University, Chongqing 400067, China*
[2]*Engineering Research Centre for Waste Oil Recovery Technology and Equipment of Ministry of Education, Chongqing Technology and Business University, Chongqing 400067, China*

Correspondence should be addressed to Xin Feng; 149495263@qq.com

A novel infrared and visible image fusion method in a multilevel low-rank decomposition framework based on guided filtering and feature extraction is proposed to address the lack of edge information and blurred details in fused images. Based on multilevel low-rank decomposition, the fusion strategy of base part and detail contents has been improved. Firstly, the source infrared and visible images are decomposed to the base part coefficients and n-level detail content coefficients by multilevel low-rank decomposition. Secondly, the base part coefficients are learned by the VGG-19 network to get the weight map, and then, the improved weight map is obtained by guided filtering, and the coefficients of the base part are fused to acquire the fused base part coefficients. The n-level detail content coefficients are fused using the rule of dynamic level measurement with maximum value and then reconstructed to obtain the final fused detail content coefficients. Finally, the fused base part and detailed content information are superimposed to get the final fusion result. The results show that the fusion algorithm can effectively preserve the edge and detail features of the source image. Compared with other state-of-the-art fusion methods, the proposed method performs better in objective assessment and visual quality. The average value of evaluation metrics *EN* and *MI* have been improved by 0.5337 and 1.0673 on the six pair images.

1. Introduction

Image fusion is an enhancement technology. Image fusion is aimed at combining different images to generate a steady and informative image, which can facilitate subsequent processing and help in decision making. Recently, many fusion methods have been proposed to fuse the features in infrared and visible images into a single image [1]. The visible images usually have high spatial resolution and large detail contrast but are easily affected by harsh environments and climatic conditions. Infrared images depict the temperature or radiation of an object, which is not easily affected by the environment and climatic conditions. However, infrared images contain a few shortcomings, such as inconspicuous texture details and poor resolution. So we can make full use of different modalities to convey complementary information. It applies in a lot of applications, such as surveillance [2], object detection, and target recognition [3–5]. The methods of multiscale transforms [6, 7] and representation learning [8] are generally used in image fusion field.

The traditional multiscale transformation method decomposes the source images into base parts and detail content at distinct dimensions. The base part mainly represents the contours and edges of the source image, and the detail content contains more detailed texture information. The base part and detail content are fused according to predefined rules in the transform domain. Then, the final fused image is obtained through inverse multiscale transformation [9]. There are some typical algorithms, such as the discrete wavelet transform (DWT) [10], contourlet transform [11,

12], shearlet transforms [13], and multilevel decomposition latent low-rank representation(MDLatLRR) [14]. These decomposition methods can be consistent with human visual characteristics but are easy to introduce artifacts. Hence, many other approaches have attracted great attention, such as sparse representation and low-rank representation.

In the sparse domain, the sparse representation(SR) [15] and dictionary learning [16] are widely used in image fusion. For instance, Li et al. [17] proposed a novel multimodal fusion method via three-layer decomposition and SR. Also, there are many methods that combine SR and other approaches for image fusion, such as low-rank representation (LRR) [18]. Zhu et al. [19] proposed a novel multimodality image fusion method based on image decomposition and sparse representation in which the texture components can be preserved well by a sparse representation based method. In [20], Liu et al. proposed a fusion method based on convolutional sparse representation (CSR) in which the detail of the source image can be retained well by multilayer features that can learn more about it in [21]. Besides, the joint sparse representation (JSR) [22] and cosparse representation [23] are also used in sparse domain. Although SR-based methods can improve image fusion performance, these methods are too time-consuming in dictionary learning operations [24]. These issues have prompted a growing study in deep learning to replace dictionary learning in SR.

In deep learning-based fusion methods, deep features of the source images can be extracted to reconstruct the fused images. For example, the VGG-19 [25], ResNet-50 [26], and DenseFuse [27] network architecture are commonly used in deep learning-based methods. Ma et al. [28, 29] proposed a multimodal image fusion based on adversarial networks, which improves the performance of image fusion to a large extent. Although the deep learning-based methods have performed well in image fusion, these methods still have some drawbacks, such as the deeper the network; the choice of parameters can be more complex.

To preserve many of the edge and detail features of the source image, we proposed a multilevel low-rank decomposition framework based on guided filtering and feature extraction algorithm for infrared and visible image fusion. This solution uses the MDLatLRR method to decompose the original images to extract the detail content coefficients. The fused detail content coefficients can be obtained by dynamic level measurement with maximum value. After superimposing these detail content coefficients, the edge and structure information of the original images will be well retained and improve the display of the object. Then, the VGG-19 network is used to extract the significant area, structure, and object characteristics of the base part coefficients, and the weight maps can be produced according to the base part's activity level. In order to better preserve the edge information of the base part, the improved weight map is obtained by guided filtering. The improved weight map and the base part coefficients then make the Hadamard product to acquire the fused base part coefficients. Finally, the fused base part and detailed content information are superimposed to get the final fusion result. After the above fusion scheme process, the experimental results show that the proposed method significantly outperforms the comparison methods in image information retention. The significant contributions of this paper are summarized as follows:

(1) We introduce MDLatLRR to decompose the source images and determine the optimal number of decomposition layers for infrared and visible image fusion

(2) Base part fusion: to obtain more features information, we use the VGG-19 network and guide filtering to fuse the base part. Firstly, the base part coefficients are learned by the VGG network to get the weight map. The weight map obtained in this way can well adapt the base part coefficients of the source image with a block distribution of pixel information. And then, the improved weight map is obtained by guided filtering, which can effectively preserve edge information and reduce noise in the weight map. Finally, the fused base part coefficients are acquired by multiplying the improved weight map and the base part coefficients

(3) Detailed content fusion: it is well known that the larger the detail content coefficient is, the more information it contains. The n-level detail content coefficients are fused using dynamic level measurement with maximum value and then reconstructed to obtain the final fused detail content coefficients, which can preserve more sufficient detail content information from the source images

(4) We first conducted ablative experiments on the number of decomposition layers of MDLattLRR and the number of layers of VGG-19 and finally selected a five-layer VGG-19 network to sufficiently extract features

The remainder of this thesis is as follows. Section 2 introduces multilevel decomposition latent low-rank representation to decompose the source images. Section 3 presents the fusion method of this paper. The base part is fused by the VGG-19 network and guide filtering. The detail content is fused by dynamic level measurement with maximum value. Section 4 presents the structure of the proposed image fusion algorithm. The experimental results are discussed and presented in Section 5. Section 6 summarizes this paper.

2. Multilevel Decomposition Latent Low-Rank Representation

In this section, the method of MDLatLRR is introduced. Liu et al. [30] proposed the method of low-rank representation (LRR) which can extract features from the input data. LRR is a method to explore the structure of data multisubspace by finding the lowest rank representation among the data.

However, this method can not work well when the input data is inadequate and damaged. In order to obtain good performance, the latent low-rank representation (LatLRR)

of theory [31] is proposed. The method utilizes more data to acquire the dictionary. In addition, the salient features can be extracted from the source data [31] by using the method of LatLRR. More specially, the single-level decomposition LatLRR (DLatLRR) problem is formulated as

$$V_d = S \cdot Q(I), \tag{1}$$

$$I_d = R(V_d), \tag{2}$$

$$I_b = I - I_d, \tag{3}$$

where I is the source image. $Q(\cdot)$ represents the two-stage operator, which composes of reshuffling and the sliding window technique. S denotes the projection matrix which is obtained by LatLRR. V_d means the decomposed result of the source image. $R(\cdot)$ is the operator which reconstructs the detail image based on detail content. I_d and I_b, respectively, signify detail content and the base part from the source image.

Due to DLatLRR, a multilevel latent low-rank representation (MDLatLRR) [14] is formed which is able to extract saliency features from the source image. The method of MDLatLRR is formulated as

$$V_d^i = S \cdot Q\left(I_b^{i-1}\right), \tag{4}$$

$$I_d^i = R\left(V_d^i\right), \tag{5}$$

$$I_b^i = I_b^{i-1} - I_d^i, I_b^0 = I, i = [1, 2, \cdots, r], \tag{6}$$

where i and r represent the present and the highest decomposition layers, respectively. V_d^i means the ith-level decomposition result of the source image. I_d^i and I_b^i, respectively, signify the ith-level detail content the base part the source image. I_b^0 indicates the source image. In the end, a base part and r detail contents are obtained in different decomposition levels.

The framework of MDLatLRR is described in Figure 1. The source image I is decomposed base part I_b^1 and detail content I_d^1 by DLatLRR. In order to obtain more feature information from the base part, the I_b^1 is further decomposed I_b^2 and I_d^2 by DLatLRR. If the decomposition layer is r, it will get r detail contents and a base part. As a result, the fused image can show more information from the source image. Nevertheless, with the decomposition layer increasing, the artifacts will introduce more. An important problem was how to select a suitable decomposition layer. The detailed description is in Section 5.1. Next, we will introduce the fusion method of the base part and the detail content, respectively.

3. Fusion Method

The source images are decomposed base parts and detail contents using the method of MDLatLRR. The base part contains edge information and basic contour information. Simonyan and Zisserman [25] employed the VGG network for the first time to extract features from images of different levels and obtained excellent results. As the level of decomposition increases, the amount of information contained in the base layer becomes less and less. Using the VGG network to extract the base part, more helpful information will be identified and integrated. The generated weight map will contain more useful information. Then, in order to contain more edge information, guided filtering [32] is used to smooth the weight map. Finally, the fused base part can be acquired by multiplying the refined weight and the source of images. In contrast to the base part, the detailed content contains more structural and textural information. The fused detail content is obtained by using the rule of taking the maximum for dynamic measurement [33].

3.1. Fusion of Base Parts.
VGG-19 is a convolutional neural network with 19 layers, including 16 convolutional layers and three fully connected layers [34]. The structure of VGGNet is straightforward, using the same size convolutional kernel size (3×3) and maximum pooling size (2×2) for the whole network. The performance can be improved by continuously deepening the network structure. The structure diagram is shown in Figure 2. For the fused base part to contain more information, a five-layer VGG-19 network is used to extract the base part to form the feature map.

For the base part I_{b1}^i and I_{b2}^i, $\{\phi_1^m\}_{m=5}^{512}$ and $\{\phi_2^m\}_{m=5}^{512}$ indicate the deep features extracted from base part by the fifth convolutional level of VGG-19. As shown in Figure 2, the 5th convolutional layer is conv3-512, so there are 512 deep features in the each base part. In addition, the pooling process will resize the the feature map, which is $1/2^5$ times of the original size. Impacted by [18], the l_1-norm of $\{\phi_k^m\}_{m=5}^{512}(x, y)\}$ can transform into the activity level survey of the original detail content, where $k \in \{1, 2\}$. Hence, the activity level map C_k^i is shown in

$$C_k^i(x, y) = \left\| \{\phi_k^m\}_{m=5}^{512}(x, y) \right\|_1. \tag{7}$$

The soft-max operator is used to obtain the initial weight maps $\widehat{W}_k$, which is shown in

$$\widehat{W}_k(x, y) = \frac{C_k(x, y)}{\sum_{n=1}^{j} C_n(x, y)}, \tag{8}$$

where j is the amount of weight map, which is set to $j = 2$. W_k denotes the value of the weight map.

Using the upsampling operator, the final weight map is obtained that is consistent with the size of the detail content. The final weight is shown in

$$W_k(x, y) = \widehat{W}_k(x + p, y + q), p, q \in \{0, 1, 2, \cdots, 15\}. \tag{9}$$

In order to retain more edge information in the base part, guided filtering is used to smooth the final weight map W_k. The detailed calculation procedure for the guided filtering is described in [32]. First, W_k is processed to obtain the binary image, which is calculated by

$$P_k^n = \begin{cases} 1, \text{ if } W_k^n = \max\left(W_1^n, W_2^n\right), \\ 0, \text{ others}, \end{cases} \tag{10}$$
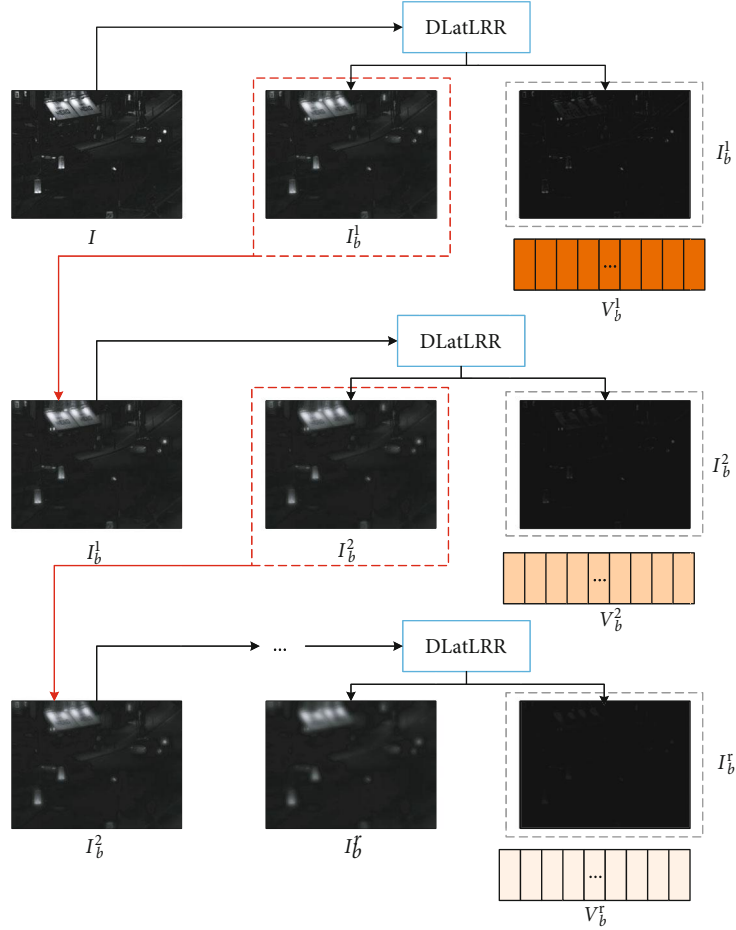
FIGURE 1: MDLatLRR decomposition diagram.

where P_k^n denotes the value of the nth pixel of the kth image in the binary image, and W_k^n means the value of the nth pixel of the kth image in the weight image.

Then, using the source image I_1 and I_2 as guided image, the guided filtering is applied to P_1 and P_2, as shown in

$$\overline{W}_1 = G_{r,\varepsilon}(P_1, I_1), \qquad (11)$$

$$\overline{W}_2 = G_{r,\varepsilon}(P_2, I_2), \qquad (12)$$

where $\overline{W}_1$ and $\overline{W}_2$ denotes the refined weight map, which is smoothed by the guided filtering. G is the guided filtering function. r and ε represent the parameters of guided filtering. If it is too smooth, it will cause the image of the edge and feature to be inconspicuous. The values of r and ε parameters are set in the experimental Section 5.1.

The fused base part I_b is calculated by

$$I_{bf} = \overline{W}_1 \cdot I_{b1}^r + \overline{W}_2 \cdot I_{b2}^r. \qquad (13)$$

3.2. Fusion of Detail Content. In general, the greater the level of coefficient activity, the more information is contained in the image. To make the fused image include rich information, we use a fusion method called the dynamic level mea-

surement with maximum value to fuse the detail content. The variance of each image patch over 3×3 or 5×5 windows is calculated as a measure of activity. The activity measure is associated with the pixel in the center of that window. The active measurements at the corresponding position are either taken as the maximum or the average, which is closed to each other. Since the activity measure in [35] corresponds to the cascading of a linear high-pass filter with a nonlinear high-pass filter, it has no clear physical meaning. In our implementation, we use the maximum absolute value within the window as the activity measure associated with the center pixel. Consistency verification can be understood as a switch to image B in the transform domain if the central pixel value is from image A and most of the surrounding pixel values are from image B. The fusion strategy diagram of the detailed content matrices is shown in Figure 3.

Firstly, the energy E_1 and E_2 are calculated for the corresponding local regions of the infrared and visible detail content, as shown in

$$E_k(x, y) = \sum_{m=-(M_1-1)/2}^{(M_1-1)/2} \sum_{n=-(N_1-1)/2}^{(N_1-1)/2} \left| V_{dk}^{i,j}(x+m, y+n) \right|^2, \quad (14)$$
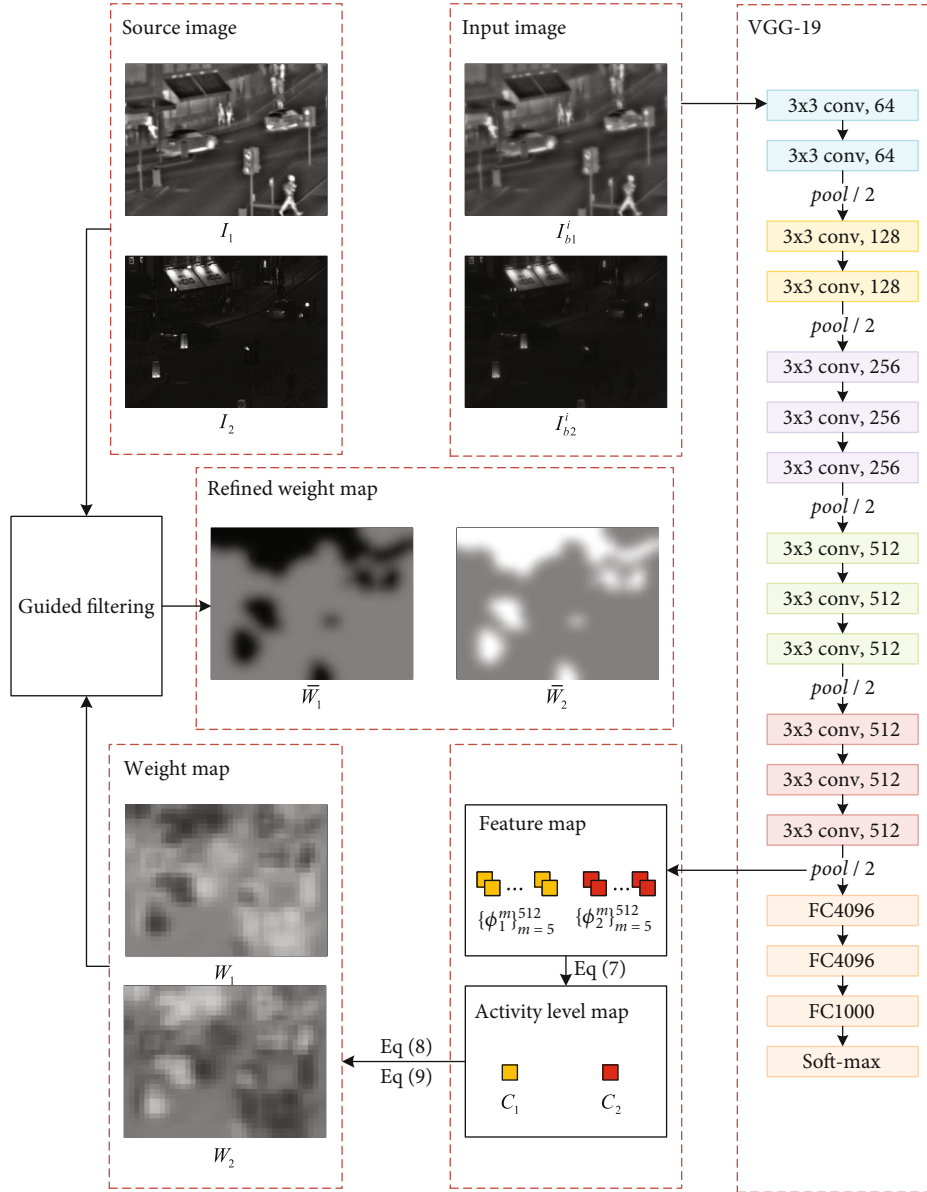
FIGURE 2: The procedure of base part fusion.

where $E_k(x, y)$ denotes the magnitude of the local energy, $k \in \{1, 2\}$. $m \times n$ defines the size of the local area, which is set to $m = n = 3$.

Then, the local area matching degree S_k is calculated by

$$S(x, y) = \frac{2 \sum_{m=-(M_1-1)/2}^{(M_1-1)/2} \sum_{n=-(N_1-1)/2}^{(N_1-1)/2} \left| V_{d1}^{i,j}(x+m, y+n) V_{d2}^{i,j}(x+m, y+n) \right|^2}{E_1(x, y) + E_2(x, y)}. \tag{15}$$

When the two images are strongly correlated, the weighted average is used. Conversely, the coefficient with higher local energy is used. The fused detail content vector

$V_{dk}^{i,j}$ is acquired by

$$V_{dk}^{i,j}(x, y) = \begin{cases} w^{\max}(x, y) V_{d1}^{i,j}(x, y) + w^{\min}(x, y) V_{d2}^{i,j}(x, y) & E_1(x, y) \geq E_2(x, y), \\ w^{\min}(x, y) V_{d1}^{i,j}(x, y) + w^{\max}(x, y) V_{d2}^{i,j}(x, y) & E_1(x, y) E_2(x, y), \end{cases} \tag{16}$$

$$w^{\min}(x, y) = \frac{1}{2} - \frac{1}{2} \left[\frac{1 - S(x, y)}{1 - \alpha} \right], w^{\max}(x, y) = 1 - w^{\min}(x, y), \tag{17}$$

where α is the matching threshold, which is set to $0.5 \sim 1$. $w^{\min}$ and $w^{\max}$ are the weighting factors.

The strategy is used to all detail content vector $V_{dk}^{1:r}$. The detailed content fusion procedure is shown in Figure 3.
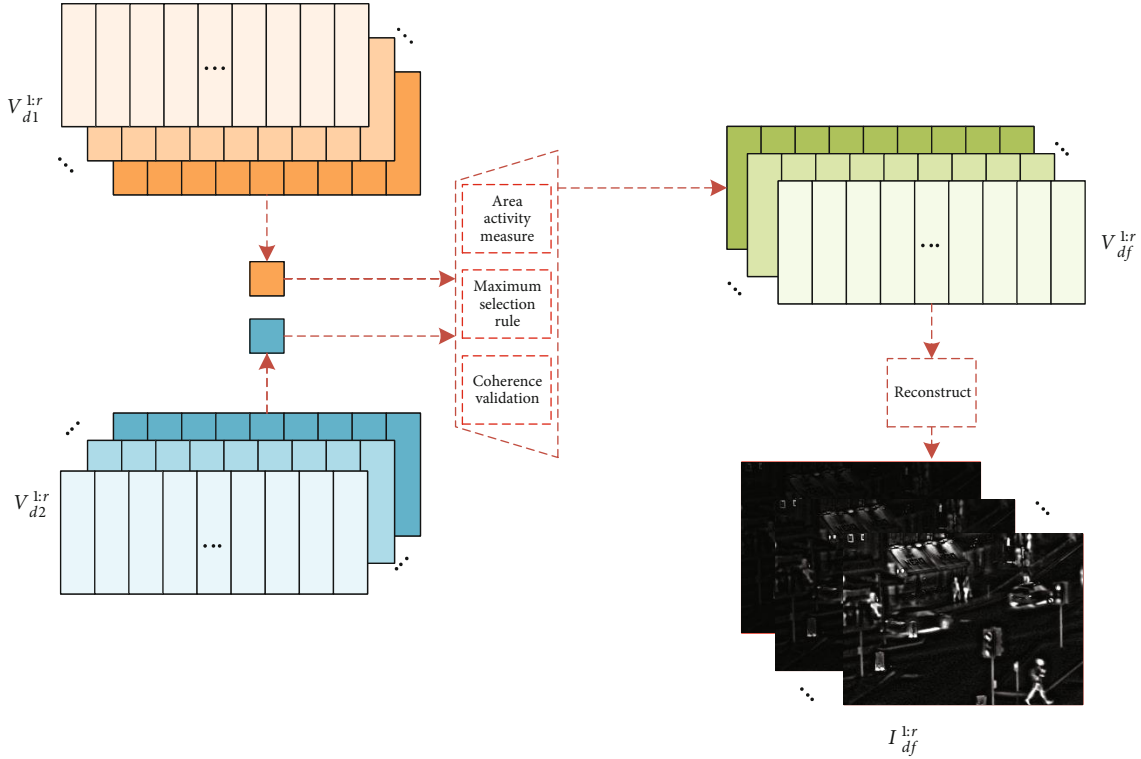
FIGURE 3: The procedure of detail content fusion.

Every detail content I_{df}^{i} is obtained by

$$I_{df}^{i} = R\left(V_{df}^{i}\right) \quad i = [1, 2, \cdots r], \tag{18}$$

where $R(\cdot)$ denotes the refactor operator, which is mainly used to reorganize vectors into image blocks.

3.3. Reconstruction. The fused base part I_{bf} and detail content I_{df}^{i} is superposed to reconstruct the fused image I_f, as shown in

$$I_f = I_{bf} + \sum_{i=1}^{r} I_{df}^{i}. \tag{19}$$

4. Structure of Fusion Algorithm

We develop a novel infrared and visible image fusion method called a multilevel low-rank decomposition framework based on guided filtering and feature extraction. The source images are denoted as I_1 andf I_2, which are preregistered. The proposed algorithm framework in this paper is shown in Figure 4.

The general steps of the proposed algorithm in this paper are shown in Algorithm 1.

5. Experiments

The aim of experiment is to give a supporting evidence for the proposed method. The experiment in this paper is com-posed of experimental settings, ablation experiment, subjective evaluation, and objective evaluation.

5.1. Experimental Settings. In our experiment, our infrared and visible images were collected from [36], which contains a lot of registered infrared and visible images from a different scene. We randomly selected six pairs of images to compare the fusion results is shown in Figure 5. From left to right, these pairs, respectively, named *Men in front of house*, *Bench*, *Bunker*, *Man_in_doorway*, *Soldier_in_trench_1*, and *Lake*.

The parameter setting for GF. According to [37], the value of r and ε is set to 45 and 0.3. The stride of the sliding window is set to 1, which can decompose the source images into patches. The window size is set to 16×16. The number of decomposition layers of the MDLatLRR and the number of network layers of the VGG-19 network to extract the base part of the feature map will be obtained from the subsequent ablation experiments.

Six classical infrared and visible image fusion methods are applied to conduct the same experiment for comparison, containing a generative adversarial network for image fusion(FusionGAN) [28], the joint-sparse representation model (JSR) [38], the JSR model with the method of saliency detection(JSR_SD) [39], multilevel decomposition method of MDLatLRR [14], two approaches based on deep learning VGG-19 [40], and ResNet50 [41].

In order to get a quantitative comparison at different methods, four quality metrics are used for the fused images. These are as follows: entropy [42] measures the amount of information contained in the fused image based on
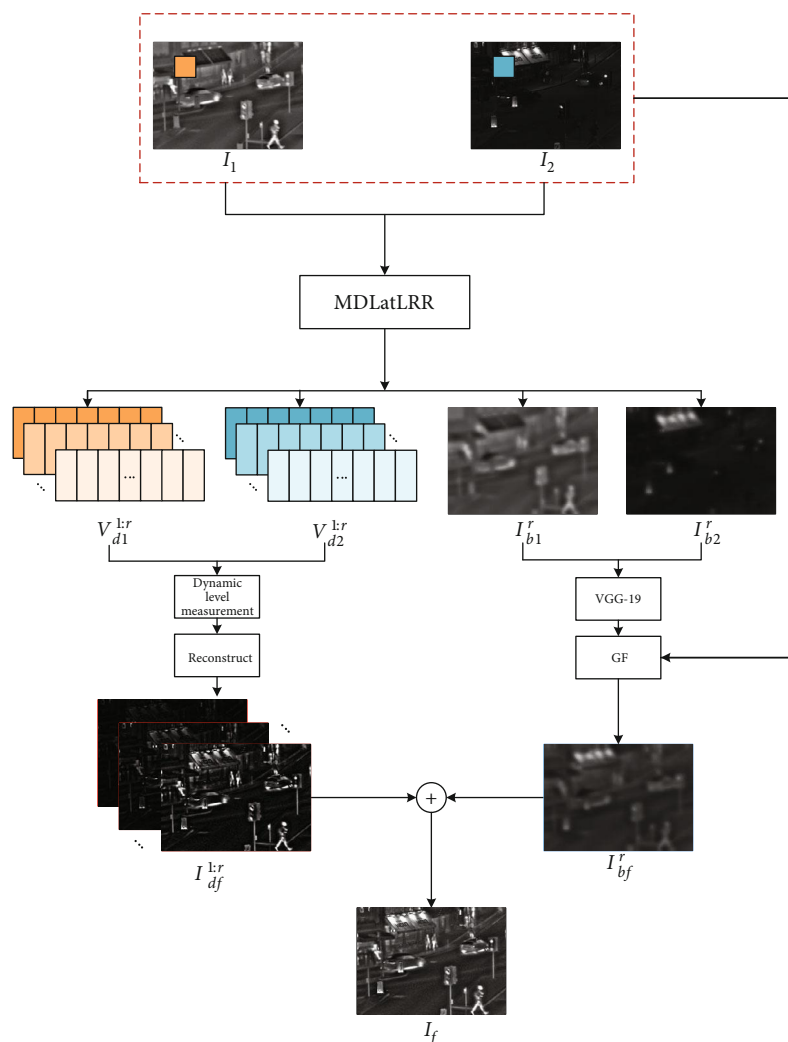
FIGURE 4: The proposed algorithm framework.

information theory; mutual information [43] represents a measure of the amount of information transferred from the source image to the fused image; Q_{abf} [44] indicates the quality of edge information acquired from the source images; MS-SSIM [45] only counts the structural information based on the refined structural similarity. The larger these metrics are, the better result of fusion quality will be.

All the fusion algorithms experiments are prosecuted in MATLAB R2020a on 3.95 GHz AMD(R) Ryzen(R) 5 3500X 6-Core Processor with 16 GB RAM and Win 10 64-bit operating system. The graphics card is GeForce RTX 2070 SUPER 8G.

5.2. Ablation Experiment

5.2.1. Ablation Experiments for Decomposing Layers. To select the best decomposition layer for the proposed method, five pairs of images in Figure 5 are implemented in the proposed algorithm in different decomposition layers. To test the decomposition layers of MDLatLRR, the layer is set from 1 to 4. The decomposition level of fused results for five pairs

of the source image is shown in Figure 6. With the increase of MDLatLRR decomposition level, the fused image luminance and contour are improved. However, it introduces the artifact around the object and makes some detailed information degradation. To obtain better fusion quality, the fewer artifacts, the better.

The experimental results are shown in Figure 7, which is obtained by the above quality metric. As can be seen, it is not the case that the greater the number of layers, the greater the value of the quality evaluation index. When the number of decomposition layers is at the first level, the value of EN and MI is more prominent than other layers. It suggests that the first layer can make the fused image contain more information from the source image. In addition, there are several images value of Q_{abf} that is best at the second layer. That indicates that more edge information is preserved in the fused image with the increasing decomposition layers. As for MS-SSIM, the first two layers show better values. It shows that the structure of the fused image is similar to the source image. However, when the decomposition layer is more than two layers, the fusion performance will decline.

Input:
The source of image I_1 and I_2.
Output:
Fused image I_f.
/∗ Part 1: multilevel DLatLRR decomposition. ∗/.
1: **for** each $K \in [I_1, I_2]$**do**
2:   **for** each $i \in [1, r]$**do**
3: Run DLatLRR decomposition on K to obtain $\{I_{b1}^i, V_{d1}^{1:i}\}, \{I_{b2}^i, V_{d2}^{1:i}\}$
4:   **end for**
5: **end for**
   /∗ Part 2: fusion of base parts. ∗/
6: **for** each $k \in \{1, 2\}$**do**
7:   Input image I_{bk}^i is extracted by the 5th layer of VGG-19 network to acquire $\{\phi_k^m\}_{m=5}^{512}$;
8:   Transform the l_1-norm of $\{\phi_k^m\}_{m=5}^{512}$ into the activity level map C_k by the Equation (7);
9:   Calculate the final weight map W_k via Equations (8) and (9);
10:   Use the guided filtering to smooth the final weight to obtain the refined weight $\bar{W}_k$ via Equations (11) and (12).
11: **end for**
12: Calculate the fused base parts I_{bf} via Equation (13).
   /∗ Fusion of detail content. ∗/
13: **for** each $i \in \{1, 2, \cdots, r\}$**do**
14:   Apply the dynamic activity level with the maximum value on $\{V_{d1}^{1:i}, V_{d2}^{1:i}\}$ to obtain the fused vector $\{V_{df}^{1:i}, V_{d2}^{1:i}\}$ as Equation (16);
15:   Reconstruct the vector $\{V_{df}^{1:i}, V_{d2}^{1:i}\}$ to I_{df}^i via Equation (18).
16: end for
   /∗ Reconstruction ∗/.
17: Superpose the fused base part I_{bf} and detail content I_{df}^i to reconstruct the fused image I_f, as shown in Equation (19).

ALGORITHM 1: Framework of the proposed algorithm for image fusion
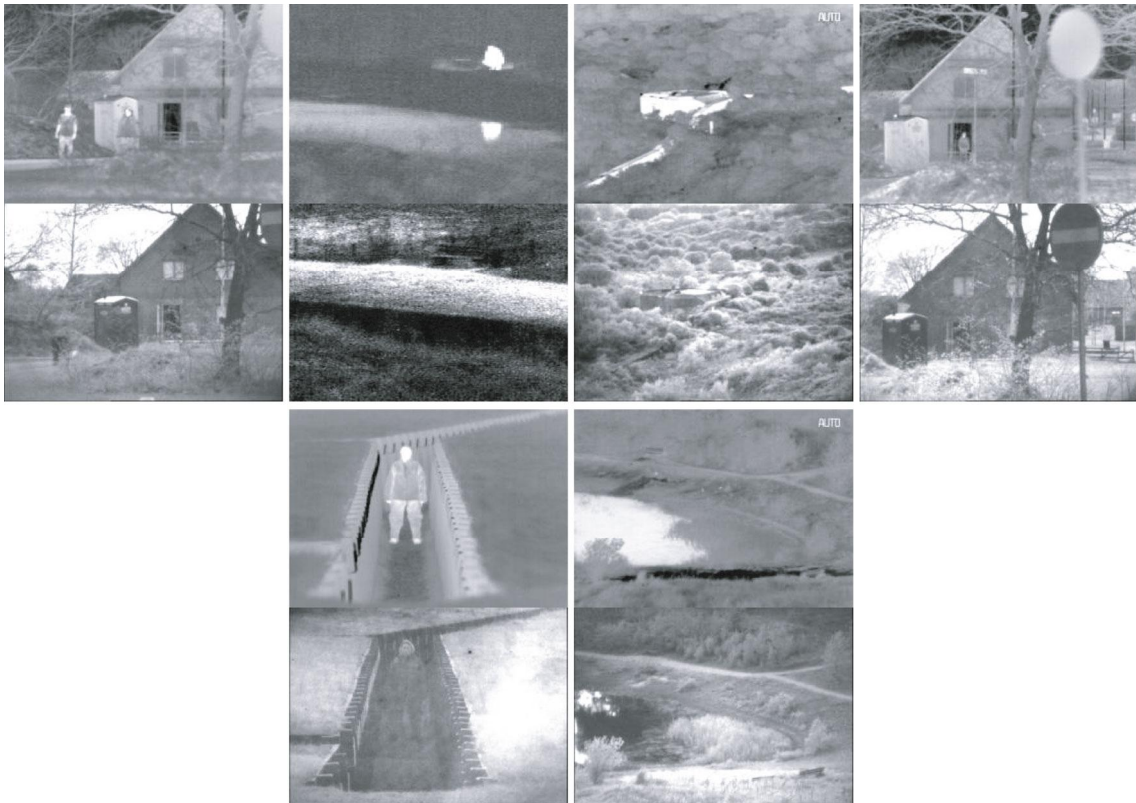
FIGURE 5: Six pairs of multimodal infrared and visible images. The top is infrared images, and the bottom is visible images.

1 layer

(a)

2 layers

(b)

3 layers

(c)

4 layers

(d)

1 layer

(e)

2 layers

(f)

3 layers

(g)

4 layers

(h)

1 layer

(i)

2 layers

(j)

3 layers

(k)

4 layers

(l)

1 layer

(m)

2 layers

(n)

Figure 6: Continued.

3 layers

(o)

4 layers

(p)

1 layer

(q)

2 layers

(r)

3 layers

(s)

4 layers

(t)
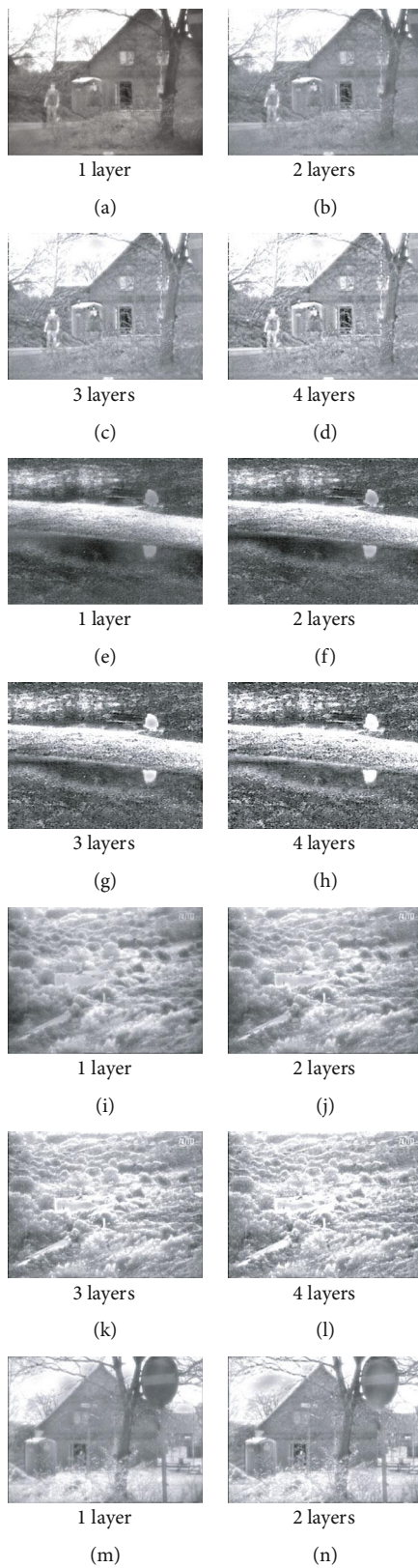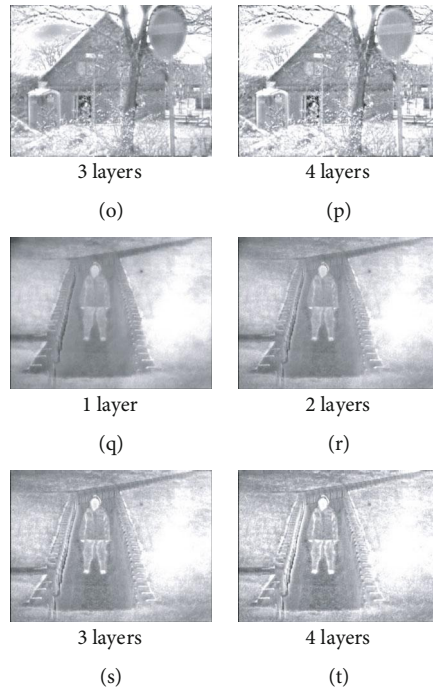
Figure 6: The each level fusion result is decomposed by MDLatLRR.
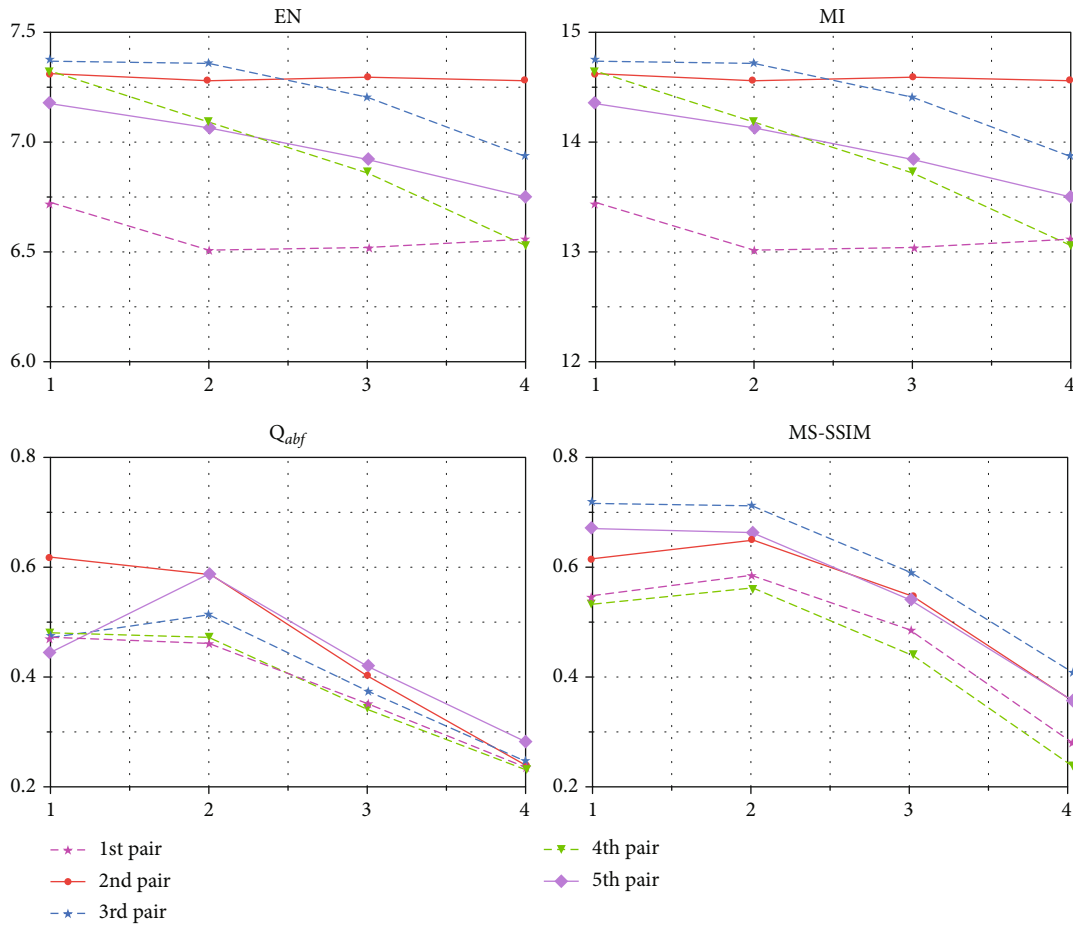
Figure 7: The decomposition layer is set from 1 to 4. The values of four evaluation metrics are acquired by MDLatLRR with different layers.

1 layer

(a)

2 layers

(b)

3 layers

(c)

4 layers

(d)

5 layers

(e)

1 layer

(f)

2 layers

(g)

3 layers

(h)

4 layers

(i)

5 layers

(j)

1 layer

(k)

2 layers

(l)

3 layers

(m)

4 layers

(n)
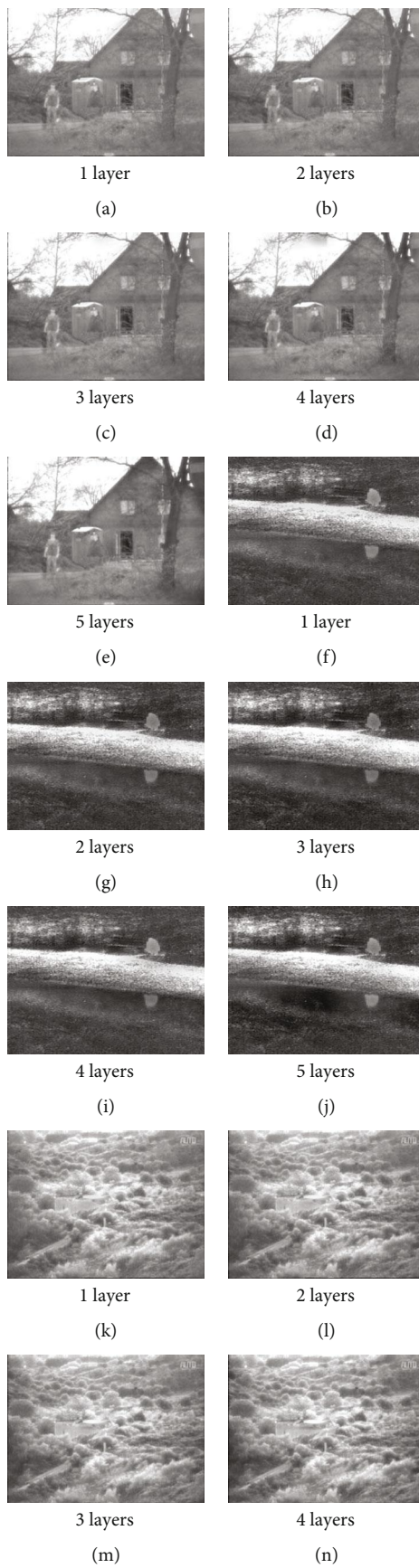
FIGURE 8: Continued.
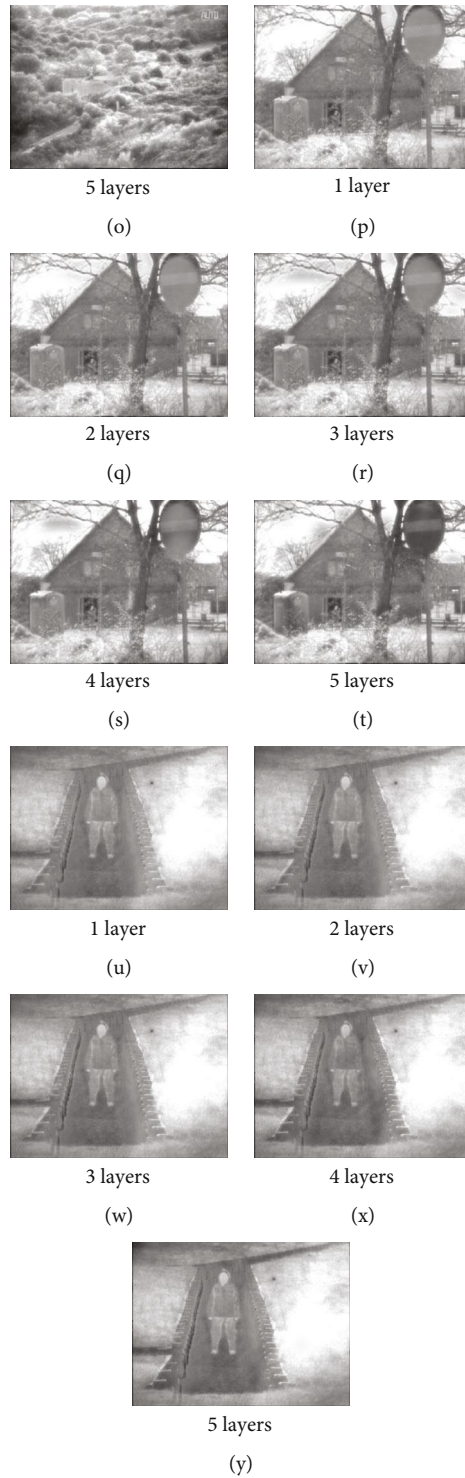
Figure 8: The fusion results for different VGG-19 network layers.

It is because that the detailed content obtains more luminance and contour information from the base part. This information can not be fused well by the detail content fusion method. By the way, the larger the value of the above evaluation index, the better the effect of fused images. On the basis of the above analysis, the decomposition of MDLatLRR is set one in our proposed algorithm.

5.2.2. Ablation Experiments for VGG-19 Network Layers. In order to select an appropriate number of layers for the VGG-19 network of the proposed method, five pairs of images in Figure 5 are used for ablation experiments of VGG-19 network layers. The layer is set from 1 to 5, which represents *relu_1_1*, *relu_2_1*, *relu_3_1*, *relu_4_1*, and *relu_5_1*, respectively. The fusion results for different
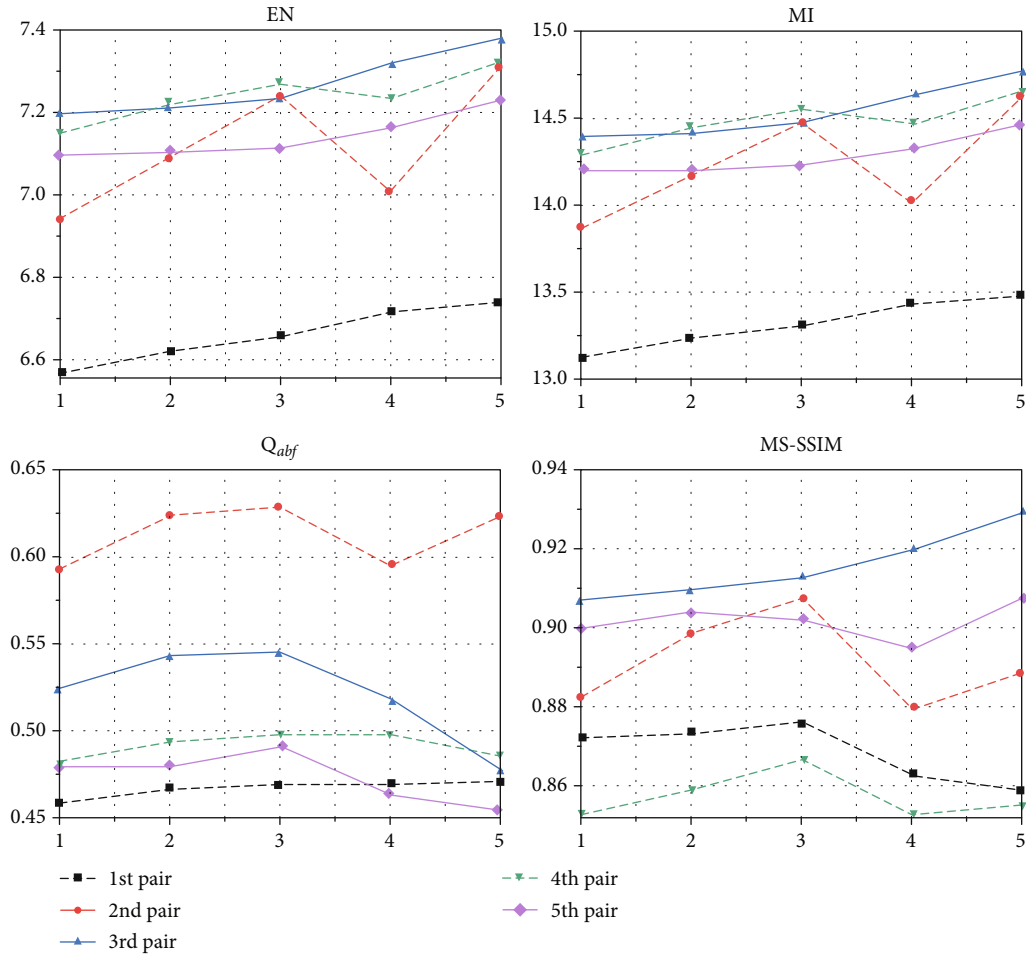
FIGURE 9: The values of four evaluation metrics are acquired by VGG-19 with different network layers.

VGG-19 network layers are shown in Figure 8. As can be seen from the Figure 8, compared with the others layers of VGG-19 network, the fifth layers of VGG-19 network extract more detail features and salient target information from the source image. For example, Figure 8(t) of the extraction of traffic sign in the fifth layer is better than the others layers, so we choose 5-layer VGG-19 network to extract features.

The experimental results of different network layers are shown in Figure 9. As can be seen, with the number of network layers increases, the values of evaluation metrics EN and MI become larger. It represents that the five-layer VGG-19 network can extract more feature information from the source images. The evaluation metric Q_{abf} is basically the best when the source images are extracted using a three-layer VGG-19 network. It indicates that the three-layer VGG-19 network can extract edge information well. As for the evaluation of MS-SSIM, the MS-SSIM values of the third and fifth pairs are the best when using the five-layer VGG-19 network. The MS-SSIM values of the first, second, and fourth pairs are the best when using the three-layer VGG-19 network. To sum up the above, we select a 5-layer VGG-19 network to extract features.

5.3. Subjective Evaluation. Figure 10 shows the subjective fusion results of the first pair of images. Figures 10(a) and 10(b) are the original images. The object of man in the red box and the grass in the green box obtained by JSR and JSR_SD are fuzzy, and the fused images obtained by JSR and JSR_SD have significantly more the visible components than the infrared ones. The fused images obtained by FusionGAN have more the infrared components than the visible ones. In addition, the fused images obtained by MDLatLRR, VGG-19, and ResNet-ZCA are less artifact but the detailed texture information in the visible image is not well preserved. As shown from Figure 10(i), the object of man in the red box and the grass in the green box are the clearest compared with other methods. The proposed method adds more detailed texture information to make the same as a visible image while containing the infrared image of thermal radiation information. It has excellent visibility. Figure 11 shows the subjective fusion results of the second pair of images. Figures 11(a) and 11 (b) are the original images. It can be seen from Figure 11(i) that the pixel consistency of the object edge structure is the best for the fused images. The objects of red and green boxes obtain more texture information. Figure 12 shows the subjective fusion results
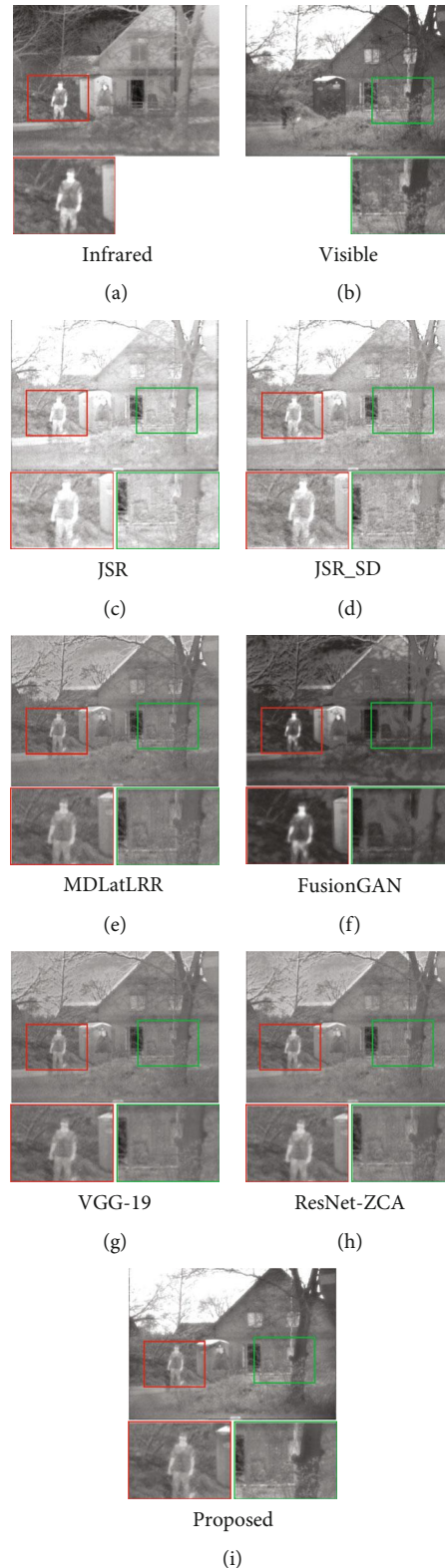
FIGURE 10: Comparison of subjective fusion results using different methods in the first pair images.

of the third pair images. Figures 12(a) and 12 (b) are the original images. In Figure 12(i) of the proposed method, the building in the red box contrasts with its surroundings, and the chromatic aberration is consistent with the visible image. The grass in the green box has more texture information. Figure 13 shows the subjective fusion results of the fourth pair of images. Figures 13(a) and 13(b) are the original images. In Figure 13(i), the brand

Infrared   Visible

(a)   (b)

JSR   JSR_SD

(c)   (d)

MDLatLRR   FusionGAN

(e)   (f)

VGG-19   ResNet-ZCA

(g)   (h)

Proposed

(i)
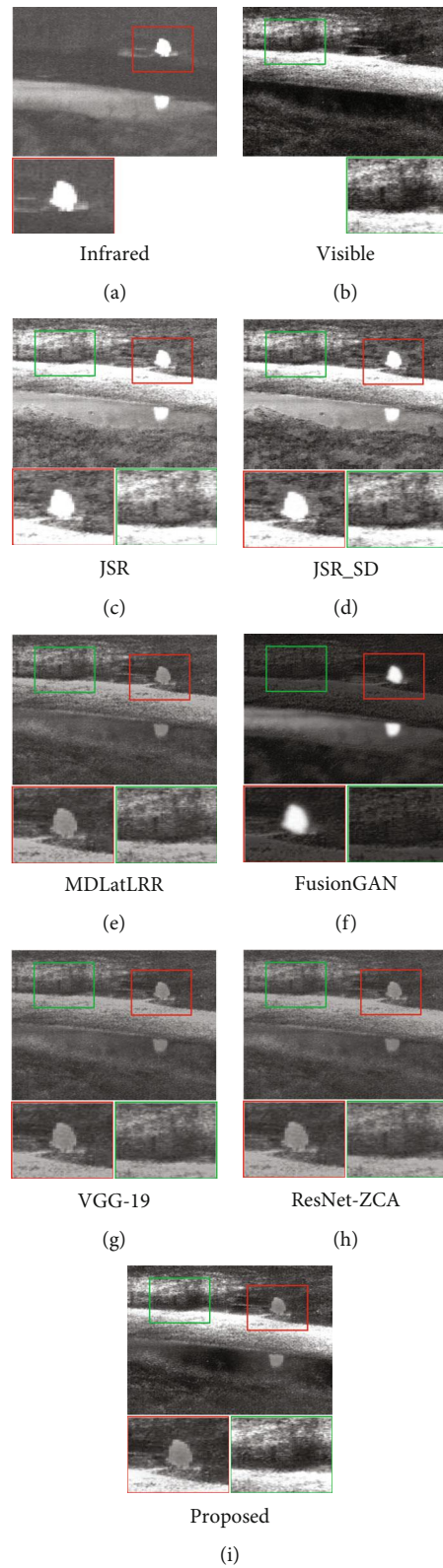
FIGURE 11: Comparison of subjective fusion results using different methods in the second pair images.

in the green box is the most recognizable compared with the results of other methods. The object in the red box contains more edge feature information. Besides, target object information is lost in some images, such as Figures 13(c), 13(d), and 13(f). The proposed method performs well and has good visibility. Figure 14 shows the subjective fusion results of the fifth pair images. Figures 14(a) and 14(b) are the original images.

Infrared

(a)

Visible

(b)

JSR

(c)

JSR_SD

(d)

MDLatLRR

(e)

FusionGAN

(f)

VGG-19

(g)

ResNet-ZCA

(h)

Proposed

(i)
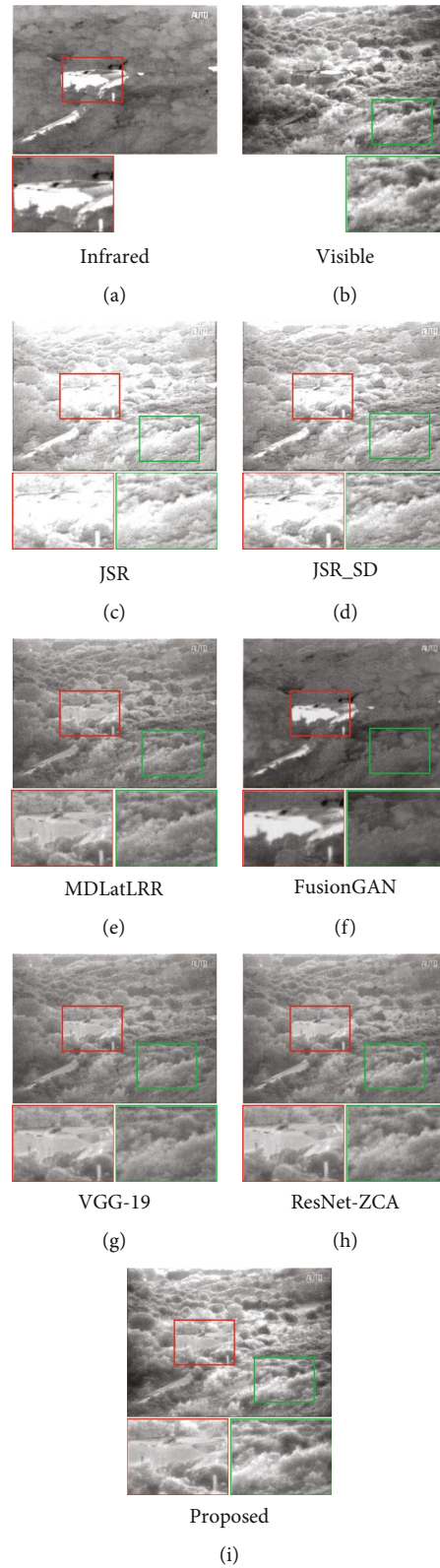
FIGURE 12: Comparison of subjective fusion results using different methods in the third pair images.

Figure 15 shows the subjective fusion results of the sixth pair of images. Figures 15(a) and 15(b) are the original images. From the target of red boxes and the detail features of green boxes in Figures 15(c)–15(i), the proposed method contains more details information from the source images compared with other methods. The proposed method adds more detail texture information to make the same as a visible image while containing the

Infrared   Visible

(a)   (b)

JSR   JSR_SD

(c)   (d)

MDLatLRR   FusionGAN

(e)   (f)

VGG-19   ResNet-ZCA

(g)   (h)

Proposed

(i)
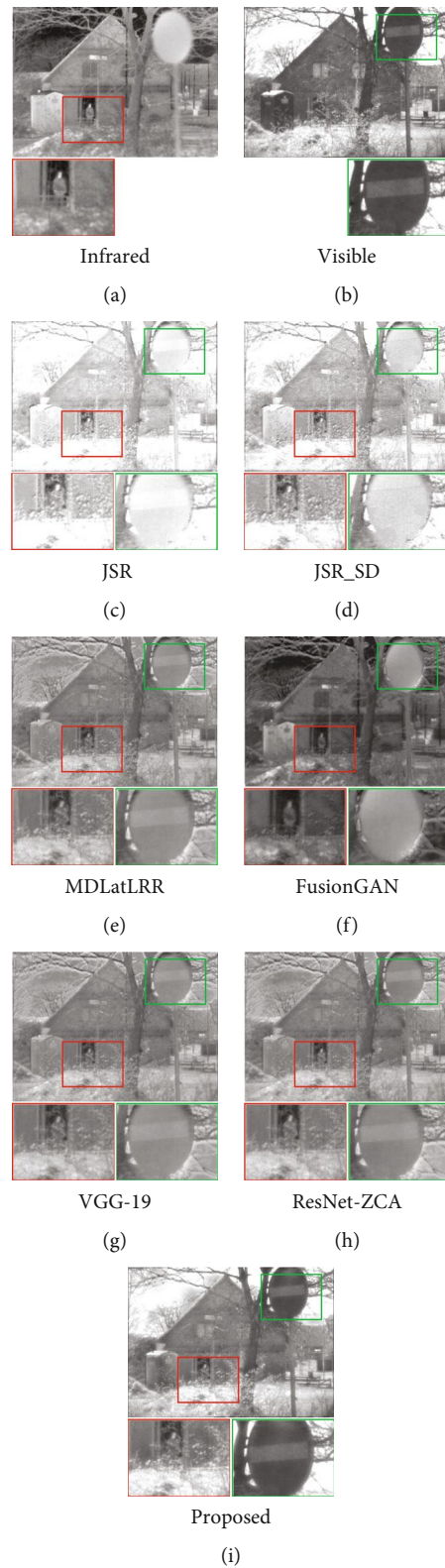
FIGURE 13: Comparison of subjective fusion results using different methods in the fourth pair images.

infrared image of thermal radiation information. Compared with other methods, the object in the red box and green box are more texture information, and the contrast between light and dark details is sharp. The structure is the most consistent with the original images. In addition, we randomly chose 20 pairs of images from [36] to verify the performance of the proposed method in Figure 16.

Infrared

(a)

Visible

(b)

JSR

(c)

JSR_SD

(d)

MDLatLRR

(e)

FusionGAN

(f)

VGG-19

(g)

ResNet-ZCA

(h)

Proposed

(i)
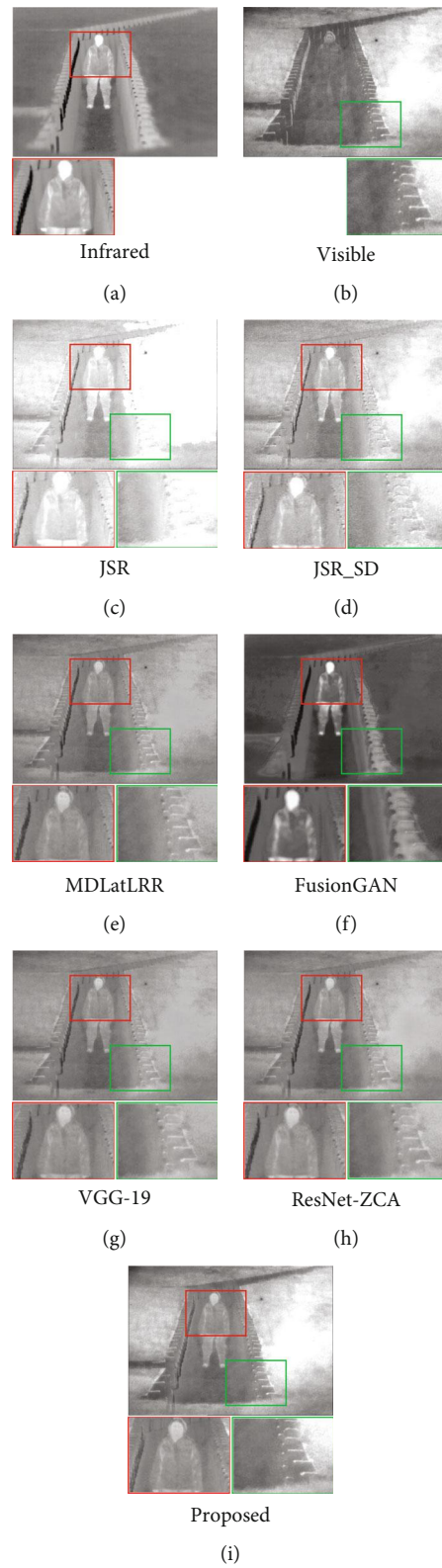
FIGURE 14: Comparison of subjective fusion results using different methods in the fifth pair images.

Infrared

(a)

Visible

(b)

JSR

(c)

JSR_SD

(d)

MDLatLRR

(e)

FusionGAN

(f)

VGG-19

(g)

ResNet-ZCA

(h)

Proposed

(i)
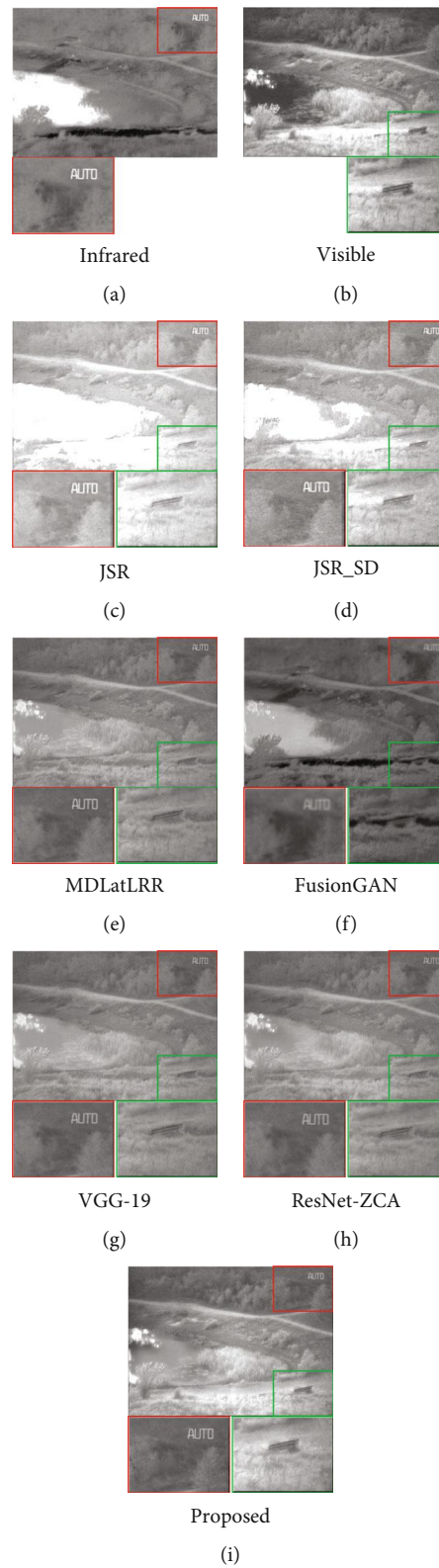
FIGURE 15: Comparison of subjective fusion results using different methods in the sixth pair images.

1st pair

(a)

2nd pair

(b)

3rd pair

(c)

4th pair

(d)

5th pair

(e)

6th pair

(f)

7th pair

(g)

8th pair

(h)

9th pair

(i)

10th pair

(j)

11th pair

(k)

12th pair

(l)

13th pair

(m)

14th pair

(n)
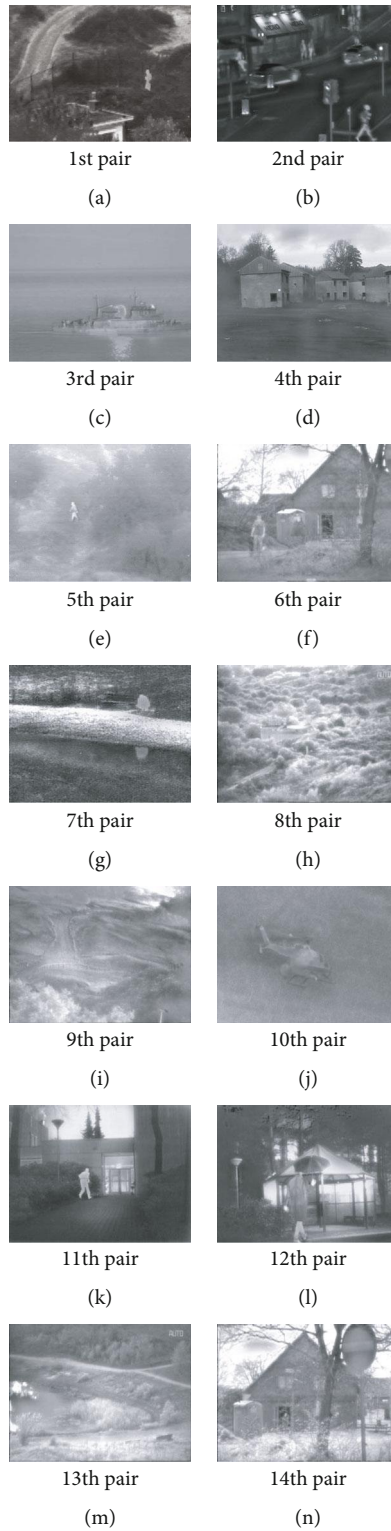
FIGURE 16: Continued.
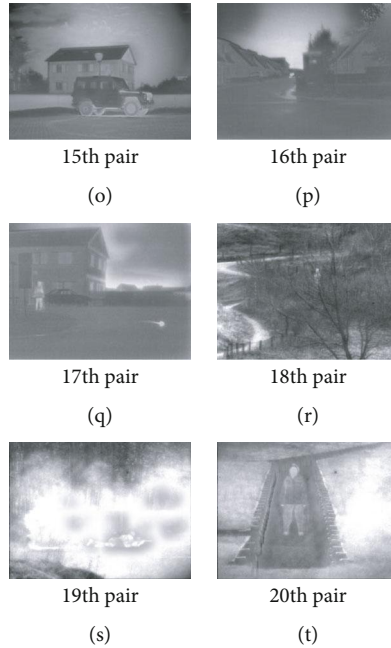
FIGURE 16: The fusion results using the proposed method on the twenty pairs of images.

5.4. Objective Evaluation. For exhibiting the attractive characteristic of our proposed method, four evaluation metrics are applied to compare the fusion property of six popular fusion methods and our proposed algorithm. In the tables, the best values are shown in italics.

In Table 1, the evaluation metrics EN, MI, and Q_{abf} are the best. It indicates that the proposed method contains more detailed information from the original images and edge information. In addition, the MS-SSIM is not the best, but the gap between the proposed method of MS-SSIM and the best value by MDLatLRR is tiny. As mentioned in Section 5.1, EN and MI measure the amount of information from the source image in the fused image. But EN is susceptible to noise. As shown in Figure 12(d), the object of the JSR_SD fusion image is distorted and has apparent artifacts. That is why EN and MI perform undeniable advantages in the third pair images. MS-SSIM counts the structural information based on the refined structural similarity, and the artifacts and the distortion of image structure will lower this metric. That will result in poor visibility. The proposed method has obvious advantages in the MS-SSIM index, which contains little noise and distortion of the structure. It is crucial for infrared and visible images. In Tables 2–5, the proposed method mostly performs the best in EN, MI, and MS-SSIM index. It shows that our algorithm makes the fused image contain more information from the source image and the structure of the fused image is similar to the source image. In addition, for objective evaluation, we provide Table 6 which contains the average values of all test images on different metrics. The evaluation metrics obtained by the proposed method are the best except the values of t/s. It indicates that our proposed method contains more feature information from the source image.

TABLE 1: Objective fusion results on the first pair images when using different algorithms.

Fusion algorithm	EN	MI	Q_{abf}	MS-SSIM	t/s
Proposed method	6.7289	13.4578	0.4735	0.8586	18.1600
FusionGAN	6.4955	12.9910	0.2303	0.7476	3.3096
JSR	6.1779	12.3558	0.2953	0.8146	2.1414
JSR_SD	6.4545	12.9090	0.2866	0.7746	130.4739
MDLatLRR	6.4837	12.9675	0.4261	0.8907	74.0113
VGG-19	6.4450	12.8901	0.3526	0.8700	7.2790
ResNet-ZCA	6.5132	13.0264	0.3640	0.8758	4.1634

TABLE 2: Objective fusion results on the third pair images when using different algorithms.

Fusion algorithm	EN	MI	Q_{abf}	MS-SSIM	t/s
Proposed method	7.3758	14.7516	0.4807	0.9295	17.6002
FusionGAN	6.4505	12.9010	0.1658	0.4494	3.1953
JSR	6.5060	13.0121	0.3221	0.8279	2.1392
JSR_SD	7.0878	14.1756	0.3023	0.8210	129.8800
MDLatLRR	6.7717	13.5433	0.4630	0.8631	73.3203
VGG-19	6.7090	13.4181	0.3185	0.8070	7.4550
ResNet-ZCA	6.7676	13.5352	0.3522	0.8280	4.1687

And the structure of the fused image is similar to the source image, better than the others compared methods. Among all the compared methods, the proposed method is in the middle level in terms of time consumption in Tables 1–6. Based on the above analysis, our fusion algorithm is effective.

Table 3: Objective fusion results on the fourth pair images when using different algorithms.

Fusion algorithm	EN	MI	Q_{abf}	MS-SSIM	t/s
Proposed method	7.3181	14.6363	0.4857	0.8551	17.5412
FusionGAN	6.8485	13.6971	0.2294	0.6862	3.2032
JSR	5.5820	11.1640	0.2739	0.7258	2.1193
JSR_SD	6.4993	12.9987	0.2803	0.7439	129.8348
MDLatLRR	6.8080	13.6160	0.4395	0.8770	73.2556
VGG-19	6.7667	13.5333	0.3614	0.8516	8.1230
ResNet-ZCA	6.7676	13.5352	0.3522	0.8280	4.2020

Table 4: Objective fusion results on the fifth pair images when using different algorithms.

Fusion algorithm	EN	MI	Q_{abf}	MS-SSIM	t/s
Proposed method	7.1776	14.3552	0.4494	0.9082	17.5217
FusionGAN	6.3209	12.6418	0.2147	0.7218	3.2608
JSR	5.8003	11.6006	0.2944	0.8217	2.1742
JSR_SD	6.9258	13.8516	0.2944	0.8249	129.8357
MDLatLRR	6.5841	13.1682	0.5273	0.8954	73.3878
VGG-19	6.5430	13.0859	0.3988	0.8693	7.4692
ResNet-ZCA	6.6948	13.3896	0.4062	0.8813	4.2912

Table 5: Objective fusion results on the sixth pair images when using different algorithms.

Fusion algorithm	EN	MI	Q_{abf}	MS-SSIM	t/s
Proposed method	7.1558	14.3115	0.4775	0.8748	28.5617
FusionGAN	6.5194	13.0387	0.2329	0.7299	3.2456
JSR	6.1612	12.3224	0.2862	0.7932	2.2352
JSR_SD	6.9259	13.8518	0.2660	0.7287	143.9548
MDLatLRR	6.5695	13.1389	0.4522	0.8823	47.2279
VGG-19	6.5451	13.0901	0.3526	0.8565	11.6805
ResNet-ZCA	6.5782	13.1565	0.3584	0.8609	5.1654

Table 6: Objective fusion results of average value on the six pairs images when using different algorithms.

Fusion algorithm	EN	MI	Q_{abf}	MS-SSIM	t/s
Proposed method	7.1786	14.3572	0.4981	0.8856	17.0943
FusionGAN	6.5198	13.0395	0.2155	0.6521	2.8039
JSR	6.2254	12.4509	0.3211	0.7993	1.8635
JSR_SD	6.8701	13.7401	0.3098	0.7706	114.3775
MDLatLRR	6.6449	13.2899	0.4766	0.8763	58.9798
VGG-19	6.5957	13.1914	0.3658	0.8435	7.3742
ResNet-ZCA	6.6489	13.2978	0.3684	0.8500	3.8398

Table 7: Objective fusion results on the second pair images when using different algorithms.

Fusion algorithm	EN	MI	Q_{abf}	MS-SSIM	t/s
Proposed method	7.3155	14.6310	0.6220	0.8875	3.1809
FusionGAN	6.4838	12.9677	0.2200	0.5777	0.6088
JSR	7.1251	14.2502	0.4546	0.8122	0.3714
JSR_SD	7.3270	14.6541	0.4292	0.7303	22.2860
MDLatLRR	6.6527	13.3053	0.5512	0.8494	12.6757
VGG-19	6.5654	13.1307	0.4109	0.8066	2.2384
ResNet-ZCA	6.5488	13.0976	0.3711	0.7990	1.0481

6. Conclusion

This paper proposes a multilevel low-rank decomposition method based on guided filtering and feature extraction for infrared and visible image fusion. The VGG-19 network and guide filtering are used in the base layer fusion to obtain the weight map. Then, the final base layer is acquired by multiplying the initial base layer and the weight map. As for detail content fusion, we are using the dynamic activity level with maximum value to obtain the final detail content. The results exhibit that our proposed method has an attractive performance in retaining the object detail features information and edge feature information compared with other fusion methods in both subjective and objective. The proposed method can be applied in target detection and recognition in daily computer vision. In addition, there are some drawbacks to our proposed algorithm. With decomposition increasing, more luminance and contour information are introduced, aggravating the fused performances. The artifacts are more bright to interfere with the targets. In the latter work, we will be committed to reducing artifacts' effect and enhancing the fusion performance with the number of decomposition layers increasing.

Data Availability

The figures data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

References

[1] S. Li, X. Kang, L. Fang, J. Hu, and H. Yin, "Pixel-level image fusion: a survey of the state of the art," *information Fusion*, vol. 33, pp. 100–112, 2017.

[2] V. Shrinidhi, P. Yadav, and N. Venkateswaran, "IR and visible video fusion for surveillance," in *2018 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, pp. 1–6, Chennai, India, March 2018.

[3] M. X. Jiang, C. Deng, J. S. Shan, Y. Y. Wang, Y. J. Jia, and X. Sun, "Hierarchical multi-modal fusion RCN with attention model for RGB-D tracking," *Information Fusion*, vol. 50, pp. 1–8, 2019.

[4] C. Li, X. Liang, Y. Lu, N. Zhao, and J. Tang, "RGB-T object tracking: benchmark and baseline," *Pattern Recognition*, vol. 96, article 106977, 2019.

[5] Y. Zhu, B. Zhu, H. H. Liu, and K. Qin, "A model-based approach for measurement noise estimation and compensation in feedback control systems," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, pp. 8112–8127, 2020.

[6] J. Chen, X. Li, L. Luo, X. Mei, and J. Ma, "Infrared and visible image fusion based on target-enhanced multiscale transform decomposition," *Information Sciences*, vol. 508, pp. 64–78, 2020.

[7] A. Vishwakarma and M. K. Bhuyan, "Image fusion using adjustable nonsubsampled shearlet transform," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 9, pp. 3367–3378, 2019.

[8] R. A. Borsoi, T. Imbiriba, and J. C. M. Bermudez, "Super-resolution for hyperspectral and multispectral image fusion accounting for seasonal spectral variability," *IEEE Transactions on Image Processing*, vol. 29, pp. 116–127, 2020.

[9] G. Piella, "A general framework for multiresolution image fusion: from pixels to regions," *Information fusion*, vol. 4, no. 4, pp. 259–280, 2003.

[10] Z. Wang, X. Li, H. Duan, X. Zhang, and H. Wang, "Multifocus image fusion using convolutional neural networks in the discrete wavelet transform domain," *Multimedia Tools and Applications*, vol. 78, no. 24, pp. 34483–34512, 2019.

[11] K. Seethalakshmi and S. Valli, "A fuzzy approach to recognize face using contourlet transform," *International Journal of Fuzzy Systems*, vol. 21, no. 7, pp. 2204–2211, 2019.

[12] H. Wei, Z. Zhu, L. Chang et al., "A novel precise decomposition method for infrared and visible image fusion," in *2019 Chinese Control Conference*, pp. 3341–3345, Guangzhou, China, July 2019.

[13] W. Ahmad, S. Vagharshakyan, M. Sjöström, A. Gotchev, R. Bregovic, and R. Olsson, "Shearlet transform-based light eld compression under low bitrates," *IEEE Transactions on Image Processing*, vol. 29, pp. 4269–4280, 2020.

[14] H. Li, X.-J. Wu, and J. Kittler, "MDLatLRR: a novel decomposition method for infrared and visible image fusion," *IEEE Transactions on Image Processing*, vol. 29, pp. 4733–4746, 2020.

[15] S. Maqsood and U. Javed, "Multi-modal medical image fusion based on two-scale image decomposition and sparse representation," *Biomedical Signal Processing and Control*, vol. 57, article 101810, 2020.

[16] Q. Hu, S. Hu, and F. Zhang, "Multi-modality medical image fusion based on separable dictionary learning and Gabor filtering," *Signal Processing: Image Communication*, vol. 83, article 115758, 2020.

[17] X. Li, F. Zhou, and H. Tan, "Joint image fusion and denoising via three-layer decomposition and sparse representation," *Knowledge-Based Systems*, vol. 224, article 107087, 2021.

[18] H. Li and X.-J. Wu, "Multi-focus image fusion using dictionary learning and low-rank representation," in *International Conference on Image and Graphics*, pp. 675–686, Springer, Cham, 2017.

[19] Z. Zhu, H. Yin, Y. Chai, Y. Li, and G. Qi, "A novel multimodality image fusion method based on image decomposition and sparse representation," *Information Sciences*, vol. 432, pp. 516–529, 2018.

[20] Y. Liu, X. Chen, R. K. Ward, and Z. J. Wang, "Image fusion with convolutional sparse representation," *IEEE Signal Processing Letters*, vol. 23, no. 12, pp. 1882–1886, 2016.

[21] Y. Liu, X. Chen, H. Peng, and Z. Wang, "Multi-focus image fusion with a deep convolutional neural network," *Information Fusion*, vol. 36, pp. 191–207, 2017.

[22] X. Ma, S. Hu, S. Liu, J. Fang, and S. Xu, "Multi-focus image fusion based on joint sparse representation and optimum theory," *Signal Processing: Image Communication*, vol. 78, pp. 125–134, 2019.

[23] R. Gao, S. A. Vorobyov, and H. Zhao, "Image fusion with cosparse analysis operator," *IEEE Signal Processing Letters*, vol. 24, no. 7, pp. 943–947, 2017.

[24] Y. Bin, Y. Chao, and H. Guoyu, "Efficient image fusion with approximate sparse representation," *Multiresolution and Information Processing*, vol. 14, no. 4, article 1650024, 2016.

[25] K. Simonyan and A. Zisserman, "Very deep convolutional networks for largescale image recognition," http://arxiv.org/abs/1409.1556.

[26] 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.

[27] H. Li and X.-J. Wu, "Densefuse: a fusion approach to infrared and visible images," *IEEE Transactions on Image Processing*, vol. 28, no. 5, pp. 2614–2623, 2018.

[28] J. Ma, W. Yu, P. Liang, C. Li, and J. Jiang, "Fusiongan: a generative adversarial network for infrared and visible image fusion," *Information Fusion*, vol. 48, pp. 11–26, 2019.

[29] J. Ma, P. Liang, W. Yu et al., "Infrared and visible image fusion via detail preserving adversarial learning," *Information Fusion*, vol. 54, pp. 85–98, 2020.

[30] G. Liu, Z. Lin, and Y. Yu, "Robust subspace segmentation by low-rank representation," *Icml*, vol. 1, article 8, 2010.

[31] G. Liu and S. Yan, "Latent low-rank representation for subspace segmentation and feature extraction," in *2011 international conference on computer vision*, pp. 1615–1622, Barcelona, Spain, Nov 2011.

[32] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 6, pp. 1397–1409, 2012.

[33] H. Li, B. Manjunath, and S. K. Mitra, "Multisensor image fusion using the wavelet transform," *Graphical Models and Image Processing*, vol. 57, no. 3, pp. 235–245, 1995.

[34] X.-C. Lou and X. Feng, "Multimodal medical image fusion based on multiple latent low-rank representation," *Computational and Mathematical Methods in Medicine*, vol. 2021, 16 pages, 2021.

[35] P. J. Burt and R. J. Kolczynski, "*Enhanced image capture through fusion*," in *1993 (4th) international Conference on Computer Vision*, pp. 173–182, Berlin, Germany, May 1993.

[36] A. Toet, "TNO image fusion dataset," *Data in Brief*, vol. 15, article 249, 2017.

[37] S. Li, X. Kang, and J. Hu, "Image fusion with guided filtering," *IEEE Transactions on Image Processing*, vol. 22, no. 7, pp. 2864–2875, 2013.

[38] Q. Zhang, Y. Fu, H. Li, and J. Zou, "Dictionary learning method for joint sparse representation-based image fusion," *Optical Engineering*, vol. 52, no. 5, article 057006, 2013.

[39] C. Liu, Y. Qi, and W. Ding, "Infrared and visible image fusion method based on saliency detection in sparse domain," *Infrared Physics & Technology*, vol. 83, pp. 94–102, 2017.

[40] H. Li, X.-J. Wu, and J. Kittler, "Infrared and visible image fusion using a deep learning framework," in *2018 24th international conference on pattern recognition (ICPR)*, pp. 2705–2710, Beijing, China, August 2018.

[41] H. Li, X.-J. Wu, and T. S. Durrani, "Infrared and visible image fusion with ResNet and zero-phase component analysis," *Infrared Physics & Technology*, vol. 102, article 103039, 2019.

[42] J. W. Roberts, J. A. Van Aardt, and F. B. Ahmed, "Assessment of image fusion procedures using entropy, image quality, and multispectral classification," *Journal of Applied Remote Sensing*, vol. 2, no. 1, article 023522, 2008.

[43] M. Hossny, S. Nahavandi, and D. Creighton, "Comments on 'information measure for performance of image fusion'," *Electronics Letters*, vol. 44, no. 18, pp. 1066-1067, 2008.

[44] C. Xydeas and V. Petrovic, "Objective image fusion performance measure," *Electronics Letters*, vol. 36, no. 4, pp. 308-309, 2000.

[45] K. Ma, K. Zeng, and Z. Wang, "Perceptual quality assessment for multi-exposure image fusion," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 3345–3356, 2015.

Research Article

Application of Improved Algorithm Based on Four-Dimensional ResNet in Rural Tourism Passenger Flow Prediction

Xi Chen [1,2,3] **and Donglai Cong** [1,2,3]

¹College of Geography and Tourism, Harbin University, Heilongjiang, Harbin 150086, China
²Heilongjiang Province Key Laboratory of Cold Region Wetland Ecology and Environment Research, Heilongjiang, Harbin 150086, China
³Harbin Institution of Wetland Research, Heilongjiang, Harbin 150086, China

Correspondence should be addressed to Donglai Cong; congdonglai@126.com

In rural tourism, precise visitor flow forecasts may aid management in making better decisions. It aids in the reduction of visitor crowds and trash. It also has the potential to improve visitor security. As a result, it is critical to continue to encourage tourism's long-term growth. However, regional tourist flow of rural tourism has the characteristics of high volatility, complex nonlinearity, and susceptibility to seasonal influences. Moreover, a single neural network model cannot learn both temporal and spatial correlation. Therefore, by examining the variables impacting regional tourist flow and integrating residual networks with fully linked networks, this research offers an enhanced Quad-ResNet model for forecasting regional tourist flow of rural tourism. To be specific, this model learns spatial correlation through deep convolution; combines four residual networks to learn temporal proximity, similarity, periodicity and tendency; and uses one fully connected network to learn seasonal effects. Furthermore, this study compares the Quad-ResNet model with LSTM, CNN, and ST-ResNet models on the same dataset for regional tourist flow prediction experiments. The findings reveal that the Quad-ResNet model has less error and is substantially simpler to train and predict than the LSTM model, making it more suited for predicting regional visitor flows in rural tourism. For relevant stakeholders, the generated model may serve as a useful decision-making tool.

1. Introduction

Nowadays, citizens are suffering from more and more pressure from work as well as lives. Under this background, rural tourism has turned into a new kind of tourism form and has become an important component of tourist products. Rural tourism is not only based on agricultural tourism activities but also includes multifaceted tourism activities such as walking, mountain climbing, and riding a horse, adventure, sport and health tourism, hunting and fishing, educational travels, cultural and traditional trips, and some regional folk tourism activities. It provides stressed people a good chance to get in touch with nature and relax themselves. What is more important is that it can be one of the economic driving forces for rural areas and contributes to regional development significantly [1]. Also, rural tourism can bring jobs and support retail development [2]. In other words, studying tourism development is one of the effective ways to promote rural development [3]. Developing rural tourism is significant to the new rural economy construction. Thus, people attach more and more importance to rural tourism space as a suitable tourism carrier [4].

The development of rural tourism has a brilliant future in China. According to the latest estimates of the National Tourism Administration, the total annual tourist flow of rural tourism tourists in China has reached 300 million with revenue of more than 40 billion yuan, accounting for nearly one-third of the total tourism. It is reported that there are about 400 rural tourism demonstration sites in China, covering 31 provinces, autonomous regions, and municipalities in the mainland. In the three golden weeks (the Spring Festival, the May Day, and the National Day) every year, the

proportion of urban residents choosing rural tourism accounts for about 70%, which forms about 60 million person-times of rural tourism market. From 2009 to 2013, China's rural tourism revenue increased at an average annual rate of 43%. In October 2018, the National Development and Reform Commission and 13 other departments jointly issued the Action Plan for Promoting the Development and Upgrading of Rural Tourism (2018-2020), proposing to "encourage and guide social capital to participate in the development of rural tourism" and increase supporting policy supports for rural tourism development. Previously, the No. 1 document of the CPC Central Committee in 2018 clearly stated requirements on "implementing high-quality projects for leisure agriculture and rural tourism." This good news indicates that China's rural tourism will usher in a new round of investment and consumption upsurge. There is no doubt that rural tourism has become a new growth point of tourism [5].

Rural tourism-related attractions and enterprises are usually small in scale, and many new sites are open and operating in a short time. Managers often lack business management experience in the tourism industry. This tends to tourist attraction saturation, tourist congestion, and other problems, which then affect tourists' travel experience and the sustainable development of rural tourism. Forecasting the tourist flow is an important aspect in tourism economy in the rural area [6]. Accurate prediction of the tourist flow of rural tourism can provide decision-making references for managers. It helps to avoid the tourists gathering and material waste. In addition, it can also improve the safety of tourists. Thus, it is of great significance to further promote the sustainable development of tourism [7].

With the development of computational techniques, specifically in machine learning and deep learning, there are more and more algorithms that can be applied for tourist flow prediction. Traditional methods usually consist of linear regression, grey forecasting method, ARIMA, time series, Markov forecasting model, etc. [8–18]. Although these methods have been applied in various studies and get certain achievements, they failed in accurately predicting regional tourist flow with strong fluctuation and complex nonlinear. They as well cannot represent the spatial correlation. Advancements in deep learning provide a new solution for prediction problems. Algorithms like convolutional neural network, long and short memory neural network, and residual network made it possible for predicting tourist flow in consideration of elements of time and space [19–27].

Regional tourist flow of rural tourism is susceptible to seasonal effects, with stronger short-term correlation and strong volatility. To address this problem, this paper presents an improved Quad-ResNet model for tourist flow forecasting of rural tourism. A rural tourism site named Jiayuguan in Gansu province has been selected for the case study. By comparing the prediction results with other models like LSTM, CNN, and ST-ResNet, it can be found that the developed model has better accuracy. Findings from the study contribute to the knowledge body of tourism management. It can as well serve as a decision tool for managers of rural tourism.

2. Tourist Flow

2.1. Overview. This research focuses on the use of pedestrian positioning data and weather and holiday data to study the changing pattern of regional tourist flow of rural tourism and predict it in the next five minutes. The longitude and latitude of the tourist area are divided into the $M \times N$ grid map, where a grid map represents a region; then, the regional tourist flow of the mth row and nth column grid at time t can be expressed as follows:

$$f_t^{x,y} = |\{k_i \in (x,y) \bigwedge k_i \in K|, \qquad (1)$$

where K is the set of location points of pedestrians in the tourist scene at moment t, k_i is a location point in the set K, $k_i \in (x,y)$ denotes that k_i is in the $X \times Y$ grid map, and $|\cdot|$ refers to the base of the set.

2.2. Influence Factor Analysis

2.2.1. Spatial Effect. According to the first law of geography, everything is related to each other, and the correlation between neighboring things is likely to be stronger, with the closer things are to each other the more closely related they are likely to be. Therefore, the tourist flow to neighboring areas of a tourist attraction can be influenced by the movement of people, which means that there is spatial proximity. The change in tourist flow in two similar areas shows some correlation, expressed as spatial similarity.

2.2.2. Seasonal Effect. The term "seasonality" is often used in studies of rural tourism's regional visitor flow. This illustrates the unpredictably unequal distribution of flow across time as a result of weather and scheduling choices, resulting in a unique low season in the tourist business. For example, in winter when the weather is cold, regional tourist flow may show a downward trend. For example, in winter when the weather is cold, regional tourist flow may show a downward trend. In the case study of regional tourist flow of rural tourism, the factors to be considered are weather and official holidays.

2.2.3. Time Effect. The tourist flow to an attraction at a certain moment in time is influenced by the previous moments and can be specifically classified as similarity, proximity, tendency, and periodicity depending on the length of the time interval.

The difference in regional tourist flow of rural tourism between 12:00 and 11:00, 10:00 and 9:00 indicates that the majority of the areas have a small difference in tourist flow, reflecting the similarity in time.

Figure 1 shows a graph of the total tourist flow in all zones at five-minute intervals, from which it can be seen that the flow is low and decreasing from 0:00 to 5:00, increasing from 5:00 to 11:00, with a peak for the day around 12:00 and then oscillating until 21:00, after which the tourist flow starts to decrease gradually. For each hour of the day, the tourist flow is influenced by the neighboring moments, and it also influences the tourist flow in the following moments, reflecting the proximity of the area.
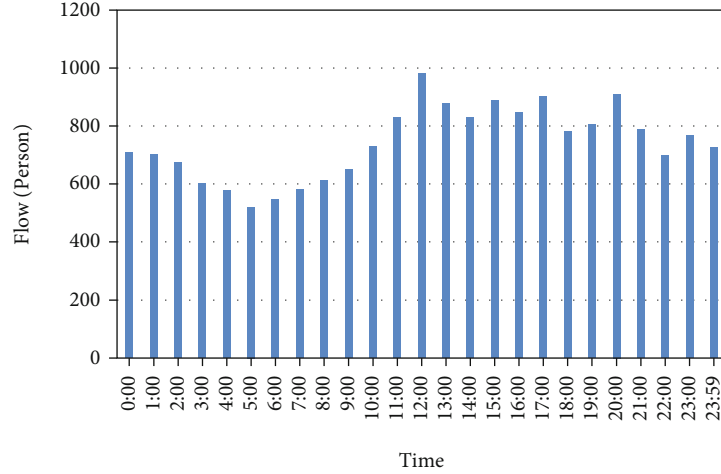
FIGURE 1: Tourist flow.

Figure 2 shows the length of time for 6 months, with an interval of 15 days, for each 9:00 am passenger flow, which increases at the same moment as the temperature warms, reflecting the tendency of regional tourist flow of rural tourism.

In Figure 3, it can be seen that the daily flow trends are broadly similar and that the tourist flow from 2019.9.3 to 2019.9.5 at the same time of day is similar, suggesting the periodicity of the regional tourist flow of rural tourism. Also, the overall trend shows a relatively stable, which suggest the periodicity as well.

2.3. Quad-ResNet Model. The structure of the Quad-ResNet model is shown in Figure 4. The model is divided into five parts, with four residual networks to simulate temporal proximity, similarity, periodicity, and tendency, and a two-layer fully connected network to simulate seasonal effects. The residual network consists of one convolutional layer, L residual cells, and one convolutional layer, which is structured to simulate spatial proximity and similarity. The outputs of the four residual networks, $o_p^{[L+2]}$, $o_q^{[L+2]}$, $o_r^{[L+2]}$, and $o_s^{[L+2]}$ are fused into o_f by a parametric matrix, and o_f is fused with the output o_e of the fully connected network. Finally, the fused output is mapped to [1] by the tanh function.

The structures of the four residual networks are the same. Take the part of proximity as an example, the residual cells based on BN (batch normalization) can be calculated as follows [28]:

$$b^{[l+1]} = d^{[l+1]} * g\left(b^{[l]}\right), \qquad (2)$$

$$b^{[l+2]} = d^{[l+2]} * g\left(b^{[l+1]}\right) + b^{[l]}, \qquad (3)$$

where $*$ refers to the convolution operations, g is the activation function ReLU, and $d^{[l+1]}$ and $d^{[l+2]}$ are learnable parameters.

The length of the input sequence for the proximity part is expressed as l_p, the time interval for proximity is $p = 1$, and the input sequence is $x_{t-l_p * p}$, $x_{t-(l_p-1)*p}$, ..., x_{t-p}. Then,

the final output of the proximity part is denoted as $o_p^{[L+2]}$.

The length of the input sequence for the similarity part is expressed as l_q, the time interval for proximity is $q = 14$, and the input sequence is $x_{t-l_q * q}$, $x_{t-(l_q-1)*q}$, ..., x_{t-q}. Then, the final output of the proximity part is denoted as $o_q^{[L+2]}$.

The length of the input sequence for the periodicity part is expressed as l_r, the time interval for proximity is $r = 312$, and the input sequence is $x_{t-l_r * r}$, $x_{t-(l_r-1)*r}$, ..., x_{t-r}. Then, the final output of the proximity part is denoted as $o_r^{[L+2]}$.

The length of the input sequence for the tendency part is expressed as l_s, the time interval for proximity is $s = 2,018$, and the input sequence is $x_{t-l_s * s}$, $x_{t-(l_s-1)*s}$, ..., x_{t-s}. Then, the final output of the proximity part is denoted as $o_s^{[L+2]}$.

The proximity, similarity, periodicity, and tendency components are fused by a parametric matrix. The output o_f after fusion is shown in the following equation:

$$o_f = d_p * o_p^{[L+2]} + d_q * o_q^{[L+2]} + d_r * o_r^{[L+2]} + d_s * o_s^{[L+2]}, \qquad (4)$$

where $*$ refers to Hadamard product; d_p, d_q, d_r, and d_s are learnable parameters, which are used to adjust the influence level of proximity, similarity, periodicity, and tendency.

After combining o_f and o_e, the predicted value $\hat{x}_t$ can be obtained through tanh function:

$$\hat{x}_t = \tanh\left(o_f + o_e\right).. \qquad (5)$$

The loss can be calculated through MSE:

$$L(\theta) = (x_t - \hat{x}_t)^2.. \qquad (6)$$

3. Experimental Design and Analysis

3.1. Model Tuning Experiments

3.1.1. Dataset. The raw data of tourist flow was sourced from the Tencent Location Big Data website, which crawled the
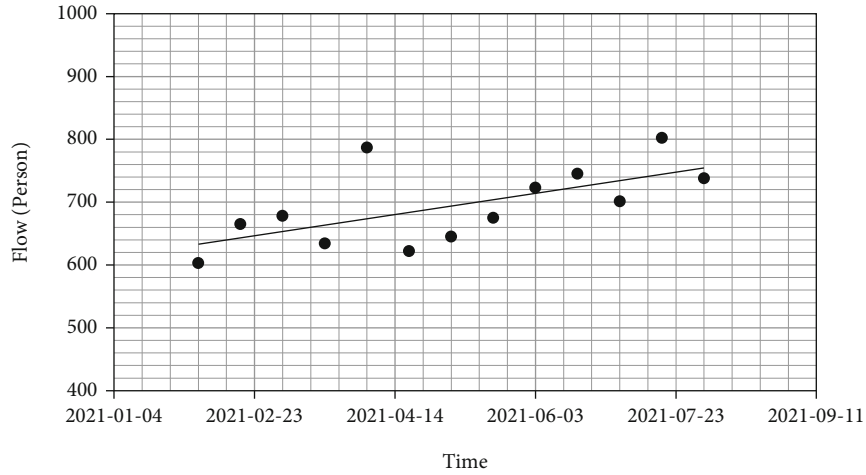
Figure 2: Regional tourist flow of rural tourism from 2021-02-03 to 2021-08-02.
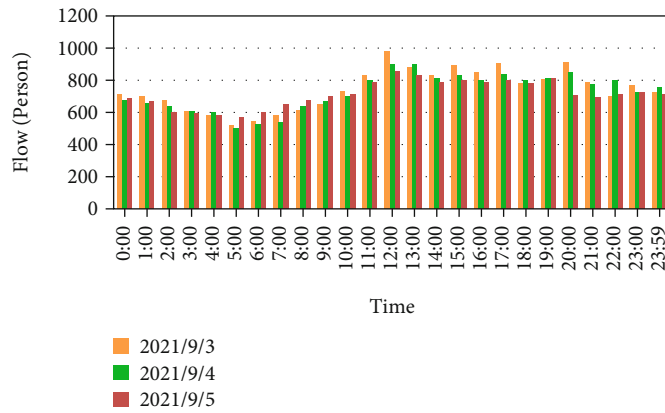

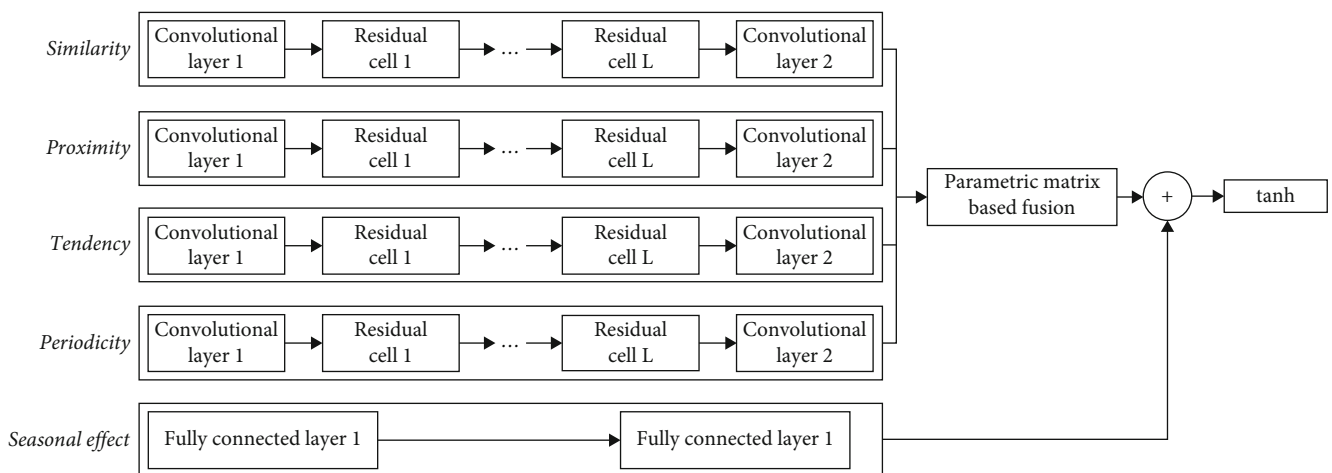Figure 3: Regional tourist flow of rural tourism in successive 3 days.

Figure 4: Structure of Quad-ResNet model.

APP location data of Jiayuguan in Gansu Province, from February 3, 2021 to August 2, 2021, with a time interval of 8 minutes, which was converted into the data of regional tourist flow of rural tourism, containing a total of 47,828 moments.

The historical weather data of Jiayuguan was used as the weather data of this attraction, the weather data contains 13 periods with these four attributes: weather conditions, temperature, wind, and wind direction. The holiday data is sourced from the open API, with weekdays marked as 0, weekend rest days are identified as 1, and legal holidays are identified as 2.

3.1.2. Parameter Setting. The Quad-ResNet model is trained using the Adam optimization algorithm with batch size set to 22, learning rate set to 0.0002, and loss function set to MSE. Interval is in units of 2 weeks. Since the time interval for the data set is 6 minutes, the proximity interval p is 1, the similarity interval q is 14, the periodic interval r is 312, and the trend interval s is 2,018.

3.1.3. Model Assess Standard. The RMSE (root mean squared error) is used to evaluate the model and is calculated as shown below. A smaller RMSE means that the model has less error and higher accuracy.

$$\text{RMSE} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(x_i - \widehat{x_i})^2}.. \qquad (7)$$

3.1.4. Model Selection. The parameter selection experiments mainly focus on the selection of proximity sequence length, periodic sequence length, similarity sequence length, tendency sequence length, and the number of residual cells.

Taking the proximity sequence length selection experiment as an example, the experimental results are shown in Figure 5. When $l_p = 3$, the RMSE is the smallest, indicating that the accuracy of the model is the highest at this time. When $l_p = 7$, the RMSE is larger than before, indicating that a longer proximity sequence not only fails to improve the accuracy of the model but also may lead to a higher accuracy of the model. This indicates that the longer proximity sequences do not improve the accuracy of the model but may lead to a decrease in the accuracy of the model. Therefore, the length of the proximity sequence $l_p = 3$ was chosen as the model with the lowest error and the highest accuracy. According to Figures 6, the length of the periodic series $l_r = 3$, the length of the similarity series $l_q = 5$, the length of the tendency series $l_s = 2$, and the number of residual cells is 3.

3.2. Comparative Experiment

3.2.1. Experiment Design. In order to verify the validity of the models, the models ST-ResNet, LSTM, CNN, and Quad-ResNet are selected to compare the performance of the models, where the deep learning models ST-ResNet, LSTM, and CNN set the same parameter values as in Quad-ResNet. These three models are all widely applied deep learning models for prediction.

A residual neural network (ResNet) is a kind of artificial neural network (ANN) that is based on known pyramidal cell constructions in the cerebral cortex. Residual neural networks do this by the use of skip connections, or shortcuts, to bypass certain layers. Typical ResNet models are built with double- or triple-layer skips with nonlinearities (ReLU) in between and batch normalization. To learn the skip weights, an extra weight matrix may be utilized; these models are referred to as HighwayNets. DenseNets are models with several parallel skips. A nonresidual network is referred to as a plain network in the setting of residual neural networks. Both vanishing gradients and degradation (accuracy saturation) may be mitigated by adding skip connections to a sufficiently deep model, but there are two primary reasons for doing so. Upstream layer muted during training and previously skipped level amplified during training. Using the simplest example, just the weights for the link between the neighboring layer and the upstream layer are changed. A single nonlinear layer or all of the intermediate layers must be linear for this to operate. A weight matrix should be learnt for the link that was missed if it is not already (a HighwayNet should be used). As a result of skipping, the network's early training phases use fewer layers [29]. There are fewer layers to propagate through, decreasing the effect of gradients disappearing. In time, the network will progressively return to the levels that were bypassed. The closer it gets to the manifold at the conclusion of training, the more quickly it learns. It is easier for a neural network with no leftover pieces to explore the whole range of features. Perturbations that lead it to depart the manifold demand more training data in order to recover.

The long short-term memory (LSTM) architecture is a kind of recurrent neural network (RNN) utilized in the area of deep learning. In contrast to conventional feedforward neural networks, LSTMs have feedback connections. It is capable of processing not only individual data points (such as photos) but also complete data sequences (such as speech or video). For instance, LSTM may be used to do tasks such as unsegmented, linked handwriting recognition, voice recognition, and anomaly detection in network traffic, or IDSs (intrusion detection systems). A typical LSTM unit is made up of four components: a cell, an input gate, an output gate, and a forget gate. The cell retains data across arbitrary time periods, and the three gates control the inflow and outflow of information. The LSTM networks are highly suited for categorizing, analyzing, and forecasting time series data, since there may be unpredictable delays between significant occurrences in a time series. LSTMs were designed to address the vanishing gradient issue that may occur when regular RNNs are trained. In various situations, LSTMs outperform RNNs, hidden Markov models, and other sequence learning algorithms due to their relative insensitivity to gap length.

The name "convolutional neural network" indicates that the network employs a mathematical operation called convolution. Convolutional networks are a specialized type of neural networks that use convolution in place of general matrix multiplication in at least one of their layers. Convolutional neural networks (CNNs or ConvNets) are a kind of artificial neural network used most often in deep learning
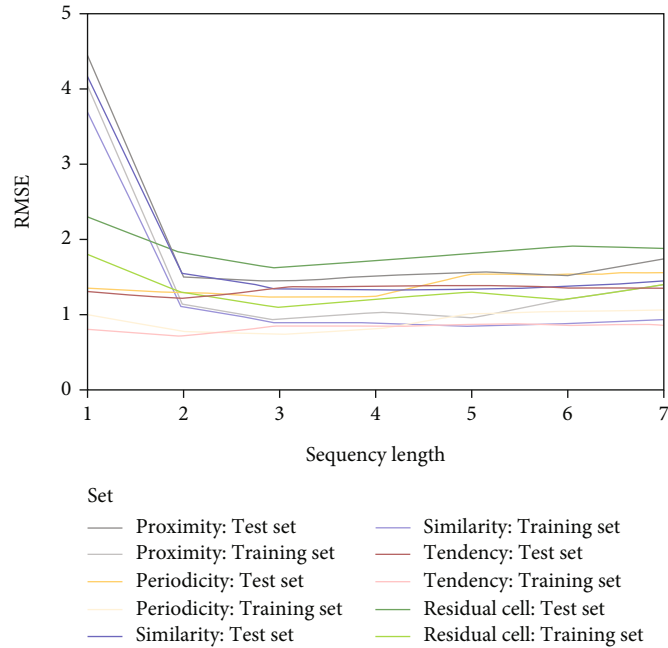
FIGURE 5: Parameter selection.
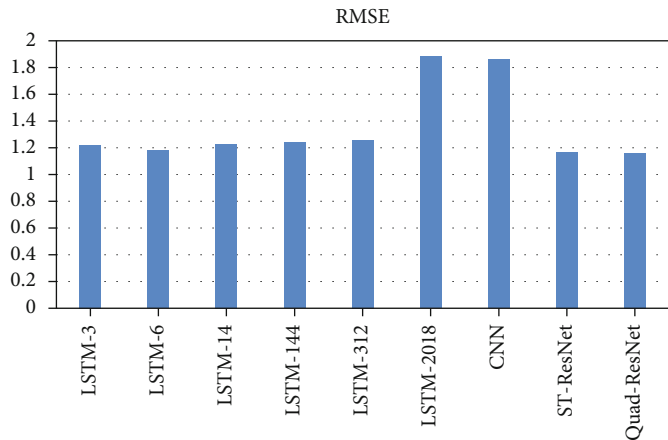
FIGURE 6: Experiment result.

for image analysis. They are also referred to as shift invariant or space invariant artificial neural networks (SIANN), due to the shared-weight design of the convolution kernels or filters that slide along input features and produce translation-equivariant outputs known as feature maps. Contrary to popular belief, the majority of convolutional neural networks are equivariant to translation, not invariant. They are used in image and video recognition, recommender systems, image classification, image segmentation, medical image analysis, natural language processing, brain-computer interfaces, and financial time series analysis. CNNs are specialized multilayer perceptrons. Multilayer perceptrons are often used to refer to completely linked networks, in which each neuron in one layer is coupled to every neuron in the next layer. Due to their "complete connectedness," these networks are prone to overfitting data. Regularization techniques that are often used to avoid overfitting include punishing parameters during training (such as weight decay) or cutting connectivity (skipped connections, dropout, etc.). CNNs tackle regularization differently: they use the hierarchical structure of data to construct patterns of increasing complexity by embossing smaller and simpler patterns in their filters. Thus, on a connectivity and complexity scale, CNNs are at the bottom end. Convolutional networks were motivated by biological processes since their connection pattern mirrors that of the animal visual brain. Individual cortical neurons react to stimuli only within a small section of the visual field referred to as the receptive field. Different neurons' receptive areas partly overlap to encompass the whole visual field. In comparison to other image classification methods, CNNs need

comparatively less preprocessing. This implies that the network learns to optimize the filters (or kernels) via automatic learning, while these filters are hand-engineered in conventional techniques. This independence from past knowledge and human interference is a significant benefit in feature extraction.

3.2.2. Experiment Result Analysis. The results of each model can be seen in Figure 6. It can be seen that RMSE of Quad-ResNet is the smallest, meaning that the accuracy is the highest. The RMSEs of LSTM-3, LSTM-6, and LSTM-14 are larger than that of Quad-ResNet, meaning that LSTM model can capture short time relevance effectively and get great predicted results. However, the spatial dependence can cause an important influence on the predicted results. Thus, it cannot just consider time dependence; otherwise, you cannot improve accuracy any further.

The RMSEs of the LSTM-3, LSTM-6, LSTM-14, LSTM-144, and LSTM-312 models are relatively similar and significantly smaller than those of the LSTM-2018 model, while the RMSEs of the models decrease and then increase as the lookback increases, suggesting that the LSTM models may have difficulty capturing very long-term temporal correlations.

The RMSE of the CNN model is significantly larger than that of the Quad-ResNet, LSTM-3, LSTM-6, LSTM-14, LSTM-144, and LSTM-312 models, and only slightly smaller than that of the LSTM-2018 model, indicating that the shallow CNN may have difficulty in capturing sufficient spatial correlation, resulting in poor prediction accuracy of the model.

The RMSE of the ST-ResNet model is larger than that of the Quad-ResNet model, which is due to the fact that the ST-ResNet model only considers temporal proximity, periodicity, and trend, but not temporal similarity, resulting in no further improvement in the accuracy of the model. The impact of short-term temporal correlation on the accuracy of regional tourist flow of rural tourism prediction is significantly greater than that of long-term temporal correlation. The Quad-ResNet model considers the temporal similarity, and therefore, the experimental results are better than those of ST-ResNet. The experimental results are better than those of the ST-ResNet model.

4. Conclusion

People nowadays are under increasing pressure due to the demands of both their professional and personal life. A new kind of tourism is emerging, and rural tourism is becoming a significant part of the overall product mix. China's rural tourist industry has a bright future. However, managers in the tourist sector often lack business management expertise. This may lead to overcrowding, congestion, and other issues that negatively impact visitors' travel experiences and rural tourism's long-term viability. For rural tourism businesses, forecasting regional tourist flow is an essential part of their business plan and strategy. This paper proposes a method for predicting regional tourist flow of rural tourism using spatiotemporal residual networks. This approach can predict the regional tourist flow of rural tour-

ism based on pedestrian location data, weather and holiday data and thus find the hotspots of tourist attractions. In this research, three models are selected to compare the performance of the Quad-ResNet model, and the Quad-ResNet model is proved to be more suitable for predicting regional tourist flow of rural tourism. However, there are still some aspects that need to be improved, as the regional tourist flow of rural tourism in this study was obtained from pedestrian data. The factors considered in this paper fail to represent all the influences and do not consider unexpected events or unknown causes. Therefore, if other influences can be explored and analyzed, the prediction model can be enhanced and its accuracy can be further improved.

Data Availability

The labeled dataset used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no competing interests.

Acknowledgments

References

[1] J. J. Villanueva-Álvaro, J. Mondéjar-Jiménez, and F. J. Sáez-Martínez, "Rural tourism: development, management and sustainability in rural establishments," *Sustainability*, vol. 9, no. 5, p. 818, 2017.

[2] S. Wilson, D. R. Fesenmaier, J. Fesenmaier, and J. C. Van Es, "Factors for success in rural tourism development," *Journal of Travel Research*, vol. 40, no. 2, pp. 132–138, 2001.

[3] M. Cawley and D. A. Gillmor, "Integrated rural tourism:: concepts and practice," *Annals of Tourism Research*, vol. 35, no. 2, pp. 316–337, 2008.

[4] J. M. Yachin and D. Ioannides, "Making do in rural tourism: the resourcing behaviour of tourism micro-firms," *Journal of Sustainable Tourism*, vol. 28, no. 7, pp. 1003–1021, 2020.

[5] J. M. G. Martínez, J. M. M. Martín, J. A. S. Fernández, and H. M. Guerrero, "An analysis of the stability of rural tourism as a desired condition for sustainable tourism," *Journal of Business Research*, vol. 100, pp. 165–174, 2019.

[6] Y. Yao, Y. Cao, X. Ding et al., "A paired neural network model for tourist arrival forecasting," *Expert Systems with Applications*, vol. 114, pp. 588–614, 2018.

[7] B. Q. Wu, J. F. Wu, X. T. Shi, T. G. Zhang, C. C. Deng, and S. S. Wu, "Visiting probability model: a new method for tourist volume forecasting," *Asia Pacific Journal of Tourism Research*, vol. 24, no. 12, pp. 1155–1168, 2019.

[8] V. Bianco, O. Manca, and S. Nardini, "Electricity consumption forecasting in Italy using linear regression models," *Energy*, vol. 34, no. 9, pp. 1413–1421, 2009.

[9] S. V. Kumar and L. Vanajakshi, "Short-term traffic flow prediction using seasonal ARIMA model with limited input data,"

European Transport Research Review, vol. 7, no. 3, pp. 1–9, 2015.

[10] F. Serin, Y. Alisan, and A. Kece, "Hybrid time series forecasting methods for travel time prediction," *Physica A: Statistical Mechanics and its Applications*, vol. 579, p. 126134, 2021.

[11] I. Svetunkov and J. E. Boylan, "State-space ARIMA for supply-chain forecasting," *International Journal of Production Research*, vol. 58, no. 3, pp. 818–827, 2020.

[12] Y. C. Hu, "Constructing grey prediction models using grey relational analysis and neural networks for magnesium material demand forecasting," *Applied Soft Computing*, vol. 93, p. 106398, 2020.

[13] Y. C. Hu, P. Jiang, Y. J. Chiu, and Y. W. Ken, "Incorporating grey relational analysis into grey prediction models to forecast the demand for magnesium materials," *Cybernetics and Systems*, vol. 52, no. 6, pp. 522–532, 2021.

[14] H. F. Xu, B. Liu, and Z. Fang, "New grey prediction model and its application in forecasting land subsidence in coal mine," *Natural Hazards*, vol. 71, no. 2, pp. 1181–1194, 2014.

[15] P. H. Xu, L. K. Yin, Z. T. Yue, and T. Zhou, "On predictability of time series," *Physica A: Statistical Mechanics and its Applications*, vol. 523, pp. 345–351, 2019.

[16] W. Wang, J. Chen, J. Wang, J. Chen, and Z. Gong, "Geography-aware inductive matrix completion for personalized Point-of-Interest recommendation in smart cities," *IEEE Internet of Things Journal*, vol. 7, no. 5, pp. 4361–4370, 2019.

[17] F. Yang, X. Y. Tang, Y. X. Gan, X. D. Zhang, J. C. Li, and X. Han, "Forecast of freight volume in Xi'an based on gray GM (1, 1) model and Markov forecasting model," *Journal of Mathematics*, vol. 2021, Article ID 6686786, 6 pages, 2021.

[18] Y. Yin and P. J. Shang, "Forecasting traffic time series with multivariate predicting method," *Applied Mathematics and Computation*, vol. 291, pp. 266–278, 2016.

[19] S. Zhao, M. Hu, Z. Cai, Z. Zhang, T. Zhou, and F. Liu, "Enhancing Chinese character representation with lattice-aligned attention," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–10, 2021.

[20] Y. Dai, T. Jin, Y. Song, S. Sun, and C. Wu, "Convolutional neural network with spatial-variant convolution kernel," *Remote Sensing*, vol. 12, no. 17, p. 2811, 2020.

[21] H. Alaeddine and M. Jihene, "Deep residual network in network," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 6659083, 9 pages, 2021.

[22] Q. B. Guo, X. J. Wu, J. Kittler, and Z. Q. Feng, "Self-grouping convolutional neural networks," *Neural Networks*, vol. 132, pp. 491–505, 2020.

[23] S. Zhao, M. Hu, Z. Cai, and F. Liu, "Dynamic modeling cross-modal interactions in two-phase prediction for entity-relation extraction," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–10, 2021.

[24] Y. M. Qian, S. H. Chen, J. C. Li et al., "A decision-making model using machine learning for improving dispatching efficiency in Chengdu Shuangliu airport," *Complexity*, vol. 2020, 16 pages, 2020.

[25] A. Quddus, A. S. Zandi, L. Prest, and F. J. E. Comeau, "Using long short term memory and convolutional neural networks for driver drowsiness detection," *Accident Analysis & Prevention*, vol. 156, p. 106107, 2021.

[26] F. Rousseau, L. Drumetz, and R. Fablet, "Residual networks as flows of diffeomorphisms," *Journal of Mathematical Imaging and Vision*, vol. 62, no. 3, pp. 365–375, 2020.

[27] M. Mirkhan and M. R. Meybodi, "Restricted convolutional neural networks," *Neural Processing Letters*, vol. 50, no. 2, pp. 1705–1733, 2019.

[28] S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," in *International Conference on International Conference on Machine Learning*, pp. 448–456, Lille, France, 2015.

[29] K. Y. Zhang, Y. J. Wang, and J. Wang, "Study on the method of setting the spatial weight matrix," *Regional Economic Review*, vol. 1, pp. 19–25, 2017.

Hindawi

Research Article

An Informatization Model of Scientific Computing for Mining Association Rules Used in Teaching Management Evaluation

Chen Chen (ID)

Graduate School of Nanjing Forestry University, Nanjing 210037, China

Correspondence should be addressed to Chen Chen; hx@njfu.edu.cn

The management information system is a kind of information-based, people-oriented, and socialized service as its purpose. It adopts advanced technology to realize the management of teachers, students, and other related personnel while effectively improving the quality of teaching. This article first analyzes the overview of data mining technology. Secondly, it analyzes the related algorithms of association rules. Finally, this article introduces the demand analysis of the university teaching quality evaluation system based on the scientific calculation of association rules. According to the demand analysis results, the system module and the corresponding submodule function modules are determined, and finally, the system architecture model is designed. Experimental results show that this model is useful for providing the effectiveness of information management.

1. Introduction

With the innovation of decision-making science and the development of modern management theory, scientific calculation based on association rule mining is gradually introduced into the application of management evaluation in universities. This makes the working methods of university administrators gradually change from traditional empirical management methods to modern management theories and methods. This change has made university administrators more creative. The traditional empirical management method only one-sidedly emphasized the particularity of education, while ignoring the commonality between public management and education management. It focuses its work on management methods based on experience. The scientific management method allows managers to combine their own experience with scientific methods to improve the management of universities.

2. Overview of Data Mining

2.1. Data Mining Definition.
At present, data mining technology is widely used in various fields. This technology can effec-

tively extract the data needed by enterprises from the market, and it can form a data set of these data. The data set can be structured or unstructured, and the corresponding results can be obtained by processing and analyzing these data sets. Of course, when universities use this technology to mine data in the market, they will not only mine related data for one field or industry. They generally involve many different regions, and the fusion of these regional data is the intersection of these different regions. The relationship between data mining and other disciplines is shown in Figure 1 [1].

2.2. Functions of Data Mining.
The task of data mining is to find hidden and meaningful information and knowledge from the database. In terms of its functions, it is generally divided into the following categories:

2.2.1. Concept Description.
The purpose of concept description is to briefly and comprehensively describe the data in the database. Statistics as we know is actually one of the simplest conceptual descriptions. They use data volume, mean value, square difference, or online analytical processing technology (OLAP) to perform multidimensional analysis and
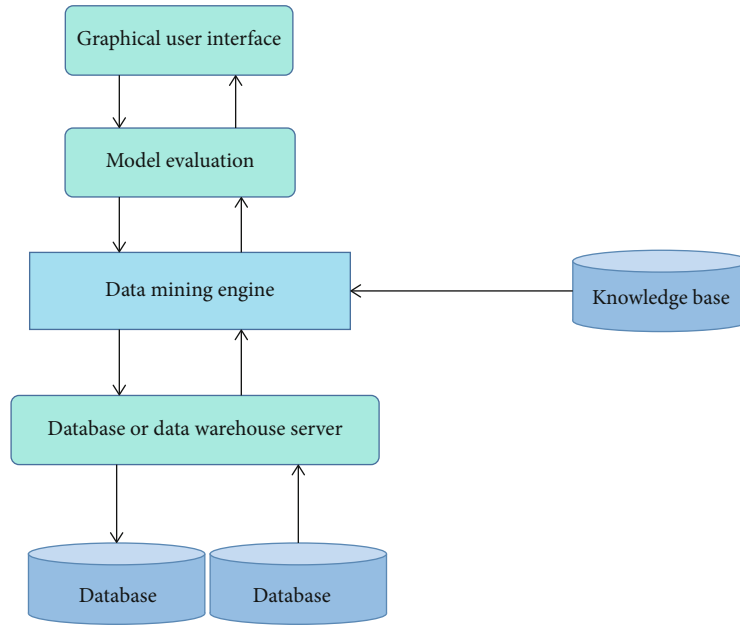
FIGURE 1: Typical data mining system.

query operations on data, and use histograms, line graphs, and other intuitive ways to display [2, 3].

2.2.2. Correlation Analysis. The core meaning of correlation analysis is to compare and correlate data obtained in different fields and then obtain a high-quality result by looking for the connection between the data. With the continuous update and expansion of the database, the content of its management and analysis has become more and more extensive, and the fields involved have become more and more extensive, which makes the database attract the attention of people from all walks of life. Association analysis finds eye-catching associations and connections with a large number of business transaction databases, so that users can make business predictions and decision-making. The main methods of association analysis are the classic Apriori algorithm and FP tree growth algorithm.

2.2.3. Classification and Prediction. The main purpose of classification is to extract important data, which can be used to predict location objects to obtain conceptual models, and relevant personnel will classify these conceptual models obtained from data information. When the predicted value of a data value is missing or unexpected, it is usually called a prediction. There are many applications for classification and prediction.

2.2.4. Cluster Analysis. The object of cluster analysis is data. The main idea is to analyze the original data. The original data refers to the data without descriptive information and the data without any classification mode. After that, the data is divided into different types of structures. Contrary to classification, clustering is to classify data objects without prior knowledge, which is convenient for stratifying survey information and grouping events with similar information together [4].

2.2.5. Outlier Analysis. Some data objects in the database may not follow the behavior or pattern of the compiled data. These data objects are called outliers. In order to reduce the impact of these anomalies, many mining algorithms isolate them. However, these individual sections are not always useful and sometimes contain important information. For example, when fraud is detected, anomalies usually indicate that fraud is more likely and more likely to stimulate our interest [5].

3. The Process of Data Mining

Data mining refers to the processing and analysis of database instances and the establishment of data patterns or functions to describe the characteristics and relationships between data. The process of data mining is shown in Figure 2.

4. Methods of Data Mining

4.1. Neural Network Algorithm. The research of artificial neurons comes from the theory of brain neurons. It is recognized that the complex nervous system is composed of a large number of neurons, thus forming a complex organization structure. If the human body feels an external stimulus, it will immediately respond accordingly. For example, if a person's hand is accidentally punctured by a needle, he will shrink his hand unconsciously. The working process of the neural network is quite complicated and transparent to the user. The user can only see the input and output and cannot see the process of transferring information between neurons. A neural network method is a data mining algorithm that simulates the human nervous system [6].

4.2. Genetic Algorithm. A genetic algorithm is an algorithm based on biological genetic variation theory to model the biological evolution process. The idea of the algorithm is to first
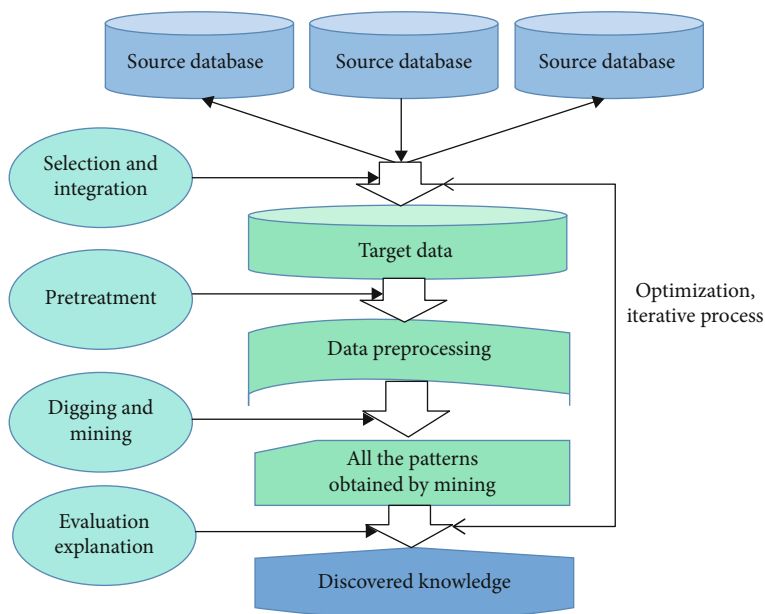
FIGURE 2: Data mining process.
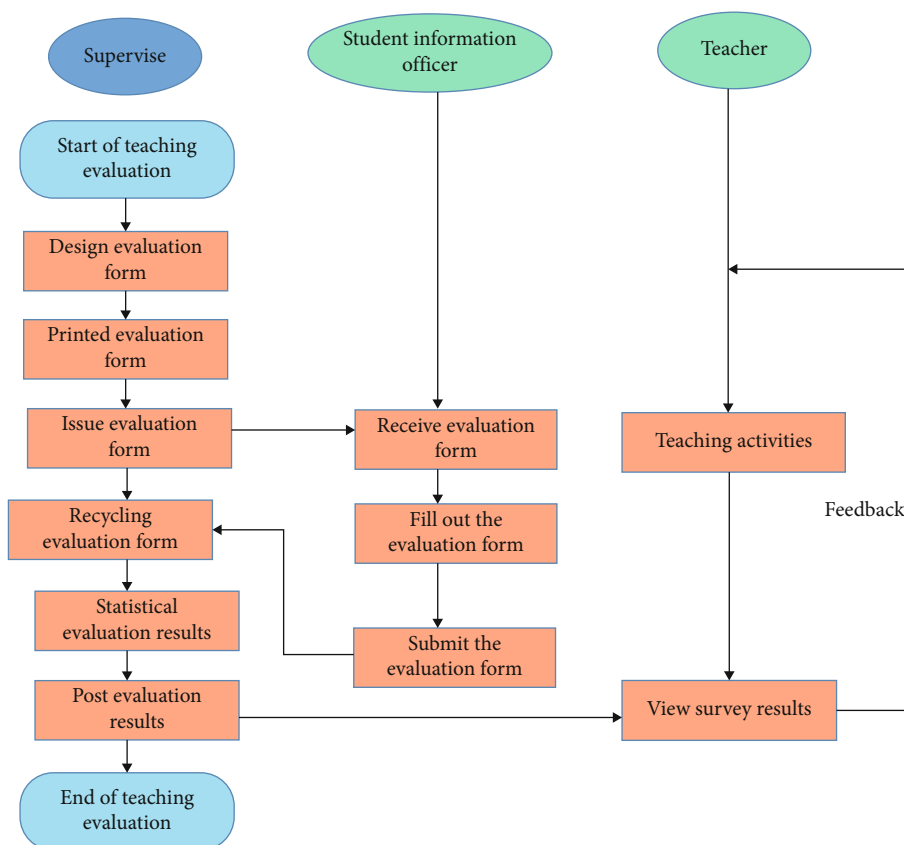
FIGURE 3: Flow chart of teaching quality evaluation activities.

identify the research object, initialize the population and encode the object, then randomly generate multiple intersections and mutation points with a small probability to generate new individuals, and add new individuals to the original pop-ulation. When the new ontology breeds the next generation, look for a population with high adaptability. The adaptability can be calculated by an adaptability function, which must be determined according to specific research questions [7].
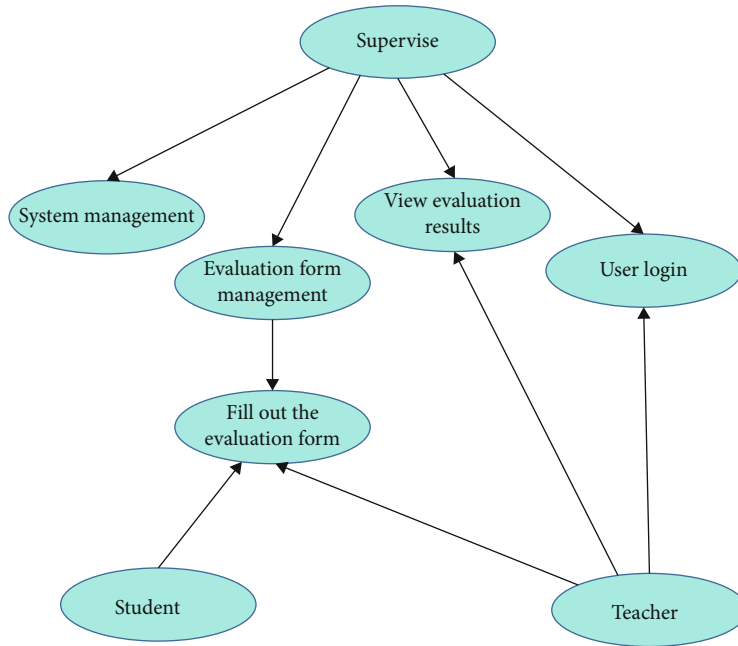
FIGURE 4: Use case diagram at the top of the system.
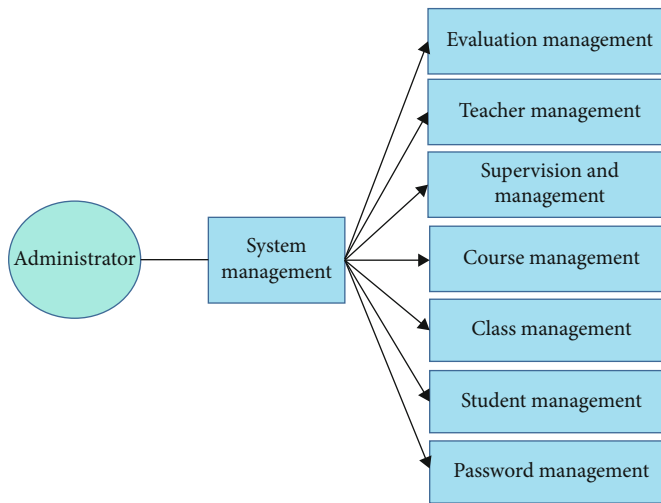
FIGURE 5: System management use case diagram.

4.3. Decision Tree Method. In essence, a decision tree classifies data through some rules and completes the screening of information through such a process. It can better block data and has strong applicability. This article uses the decision tree method to evaluate the management of colleges and universities; the following is a detailed elaboration.

It can be known from Shannon's information theory that the amount of information $I(a_i)$ of an event a_i with a probability of $p(a_i)$ can be calculated as shown in

$$I(a_i) = p(a_i) \log_2 \frac{1}{p(a_i)}. \tag{1}$$

If $a_1, a_2, a_3, \cdots, a_n$, for n mutually incompatible events, there is only one occurrence between them; then there is the following average amount of information as shown in

$$I(a_1, a_2, \cdots, a_n) = \sum_{i=1}^{n} I(a_i) = \sum_{i=1}^{n} p(a_i) \log_2 \frac{1}{p(a_i)}. \tag{2}$$

In the above formula, the base of the logarithm is any value and can be changed. But different values correspond to different units, usually 2 can be taken, and the following
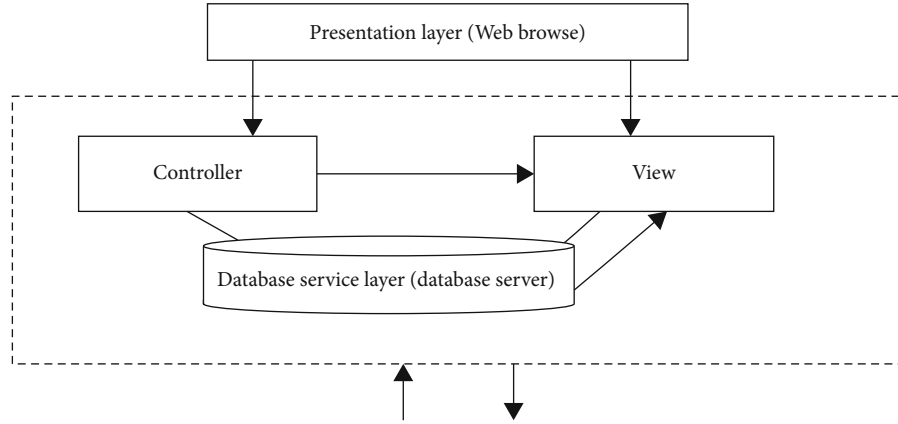
FIGURE 6: Web three-tier model based on B/S approach.

conditions are specified; when $p(a) = 0$, formula (3) is obtained:

$$I(a_i) = p(a_i) \log_2 \frac{1}{p(a_i)} = 0. \tag{3}$$

The decision tree is now defined in more detail. Assuming that the set sample is S, $|S|$ is the number of selected training samples; $|S|$ can be divided into $C_1, C_2, C_3, \cdots, C_n$ different categories. Among them $|C_1|, |C_2|, |C_3|, \cdots, |C_n|$ represents the number of elements of the class [8]. In this way, the probability of belonging to C_i in the set s is shown in

$$p(S_i) = \frac{|C_i|}{|S|}, \tag{4}$$

$$\text{Entropy}(S, A) = \sum \left(\frac{|Sv|}{|S|} \right) - \text{Entropy}(Sv). \tag{5}$$

The information gain of attribute A on set S is represented by $\text{Gain}(S, A)$, as shown in

$$\text{Gain}(S, A) = \text{Entropy}(S) - \text{Entropy}(S, A). \tag{6}$$

If there is no single evaluation object, an evaluation matrix M needs to be established, and these variables need to be normalized. The matrix is represented by the following equation:

$$M = x_{11} \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ x_{p1} & x_{p1} & \cdots & x_k \end{bmatrix}. \tag{7}$$

The optimal value and the worst value are determined by

$$F^+ = \left(\left(\max_i x_{ij} | j \in J^+ \right) \text{ or } \left(\min_i x_{ij} | j \in J^- \right) \right)^T, \tag{8}$$

$$F^- = \left(\left(\min_i x_{ij} | j \in J^+ \right) \text{ or } \left(\max_i x_{ij} | j \in J^- \right) \right)^T. \tag{9}$$

$d + (i)$ is the difference between the evaluation value and the best point, and $d - (i)$ is the difference between the evaluation value and the worst point; then formulas (10) and (11) are obtained:

$$d^+(i) = \sqrt{\sum_{j=1}^{m} \left(x_{ij} - x_j^+ \right)}, \quad 1 \le i \le n, \tag{10}$$

$$d^-(i) = \sqrt{\sum_{j=1}^{m} \left(x_{ij} - x_j^- \right)}, \quad 1 \le i \le n. \tag{11}$$

Through the calculation formula, the deviation degree is obtained as shown in

$$c_i = \frac{d^-(i)}{d^+(i) + d^-(i)}, \quad 1 \le i \le n. \tag{12}$$

In the process of evaluation, it is necessary to check the reasonableness, so that the evaluation value can be corrected. This paper uses the consistency test method, D represents the comparison matrix between the evaluation value and the test value, d_{ik} represents the evaluation value, and d_{kj} represents the test value and is calculated with

$$d_{ik} d_{kj} = d_{ij}, \tag{13}$$

$$CI = \frac{\lambda_{\max} - n}{n - 1}. \tag{14}$$

The classification using the decision tree method is divided into two stages: the purpose of the first stage is to obtain the corresponding knowledge and results from the obtained data. The main content is to use the data in the training set to build a decision tree and create a decision tree model. The second step is to use the generated decision tree model to classify unknown data samples. The data samples are sorted from the root node of the decision tree, branching

down at a time, until a certain leaf point is reached. At this time, the class represented by the leaf point is the object class.

The main content of the decision tree method is to build a high-precision, small-scale decision tree. The construction of the decision tree can be completed in two stages. The first step is to build a decision tree: this step is the process of building a decision tree by obtaining data from the sample set. Generally speaking, the practical sample data set has a certain general level, which is suitable for the data set of historical and practical needs. The second stage is the pruning of the decision tree: this stage is mainly carried out by modifying the decision tree that survived the previous stage. It checks the preliminary rules that appeared during the tree formation process and prunes branches that affect the accuracy. This can reduce the impact of noise data on classification accuracy.

The decision tree method can classify data to complete data screening. This classification process is more important. But this method has drawbacks when processing data sets. When dealing with the binary classification problem of the tree in the model solution, it is necessary to calculate the Gini index of all possible values of each possible dimension and then lay a layer of binary tree at the depth or leaf point of the tree. The binary solution tree created at this time can better classify the data. However, when the data scale is large, creating a solution tree containing all values in each dimension of all templates requires time and computing resources. Researchers proposed a solution tree acceleration algorithm model based on eigenvalue interval division, which can be used to classify large-scale data sets.

This article classifies decision tree methods into data mining methods. Decision tree methods are not only a data classification method but also a data mining method. The decision tree method can be applied to the teaching system, because the decision number method can remove useless information from a large amount of data and extract useful information. After this is applied to the teaching system, the system can delete useful learning websites for students and filter out vulgar websites.

The decision tree method is based on the "tree" as the method model, and the construction of the decision tree can be divided into two stages. The first step is to create a decision tree: the process of creating a decision tree can be described from a set of training instructions. Generally speaking, the training sample data set is a historical and comprehensive data set used to analyze and process data according to actual needs. The second step is to prune the decision tree: pruning the decision tree is the process of analyzing and changing the decision tree created in the previous step. It is mainly used to verify the initial rules created in the process of constructing the decision tree and the data in the new sample set (called the test data set), and it will remove the branches that affect the accuracy of the prediction.

4.4. Bayesian Method. A Bayesian network is a probabilistic reasoning method, which can effectively infer and process incomplete and uncertain data. It can also effectively deal with incomplete and noisy data sets and solve the inconsis-

tency and freedom of data. Bayesian classification is a statistical classification method that can predict the probability of class members. It uses the values in sampling features to calculate the probability of a sample falling into a specific category, and then, it assigns the sample to the most probable category. Bayesian classification shows high accuracy and speed when applied to large databases [9].

4.5. Method Based on Rough Set. As a convenient calculation method, the original column does not require additional information. For example, the probability distribution in statistics and the membership degree of fuzzy columns can simplify data and obtain knowledge based on the information provided by the data. It can overcome the shortcomings of traditional methods of processing uncertain information, and it can be combined with it to obtain more accurate results, which can improve the ability to deal with uncertain and incomplete information. The original set method first separates the value of the attributes in the information system through estimation and then divides each attribute into equivalence classes and then uses the equivalence relations in the set to subtract the attributes in the information system and finally obtains the minimum decision-making relation. The current relational database management system and the new data warehouse management system have laid a solid foundation for original data mining.

5. Association Rule Mining Algorithm and Its Application in Teaching Management Evaluation

5.1. Association Rule Mining-Related Algorithms

5.1.1. Apriori algorithm. The Apriori algorithm is an algorithm for mining association rules. It uses the idea of two-stage mining, and it is executed based on multiple scans of the transaction database. In the first traversal, calculate the frequency set of data items with all elements of 1, and iteratively perform the following steps: connect the results obtained from the k-th traversal, return the length $k + 1$, and set $k + 1$ frequent items to be retained. At the end of each traversal, if it is found that no frequency set has been created in this traversal, stop. Therefore, the Apriori algorithm must pass a set of transactions $c + 1$ times, where c is the number of elements in the longest frequency set [10]. The algorithm idea of Apriori algorithm is simple and easy to implement. It uses a recursive statistical algorithm to generate a continuous element set, which greatly reduces the size of the continuous element set and this can obtain good performance. But the Apriori algorithm also has obvious shortcomings. First of all, some transaction items in the Apriori algorithm can be judged after the first scan that there is no need to scan again, but the results are scanned multiple times, which greatly reduces the efficiency of the algorithm. During the scanning operation, the database must be reanalyzed for each iteration. When calculating the support of the candidate set, a complete scan is always performed from the beginning of the transaction database, which will cause multiple repeated scans. The traditional Apriori algorithm

generates a large number of candidate sets when solving frequent 2-item sets, which increases the computational complexity of the algorithm. In addition, the Apriori algorithm creates a large number of continuous sets. Due to the large amount of data, more candidate name sets can be created. When there are 1000 frequent 1-itemsets I, the number of candidate 2-itemsets will exceed 1 million, which is an exponential growth. These have high requirements on the running time and running space of the machine, which makes the execution efficiency of the algorithm very low [11].

However, the Apriori algorithm still has some shortcomings that need to be improved. Details are as follows:

(1) Some transaction items in the Apriori algorithm can be judged after the first scan that there is no need to scan again, but the result is scanned multiple times. This greatly reduces the efficiency of the algorithm. In response to this problem, the researchers proposed an improved method, specifically: when the Apriori algorithm generates k-item frequent itemsets, remove some infrequent itemsets, so as not to combine them into candidates again to remove some special transaction records. The counting problem is no longer considered when generating $(k + 1)$-item frequent itemsets

(2) The traditional Apriori algorithm always performs a complete scan from the beginning of the transaction database when calculating the support of the candidate set, which will cause multiple repeated scans. Therefore, researchers have proposed a dynamic pattern counting DIC algorithm, which sets markers to divide the database. It inserts a new candidate set generated each time at any starting point, and when scanning a candidate set later, it stops scanning until it encounters a position where the candidate set already exists. Therefore, the algorithm scans the database less times than the Apriori algorithm

(3) The traditional Apriori algorithm generates many candidate sets when solving frequent 2-itemsets, which increases the computational complexity of the algorithm. Therefore, researchers have proposed a DHP algorithm for compressing candidate sets. The DHP algorithm solves frequent itemsets by decomposing the transaction database to construct a hash table. This algorithm drastically reduces the number of candidate sets with the increase of iterations, thus effectively reducing the computational complexity of the algorithm

In this paper, the Apriori algorithm is analyzed in detail. The algorithm can be widely used, and it also has many advantages. The Aprioi algorithm uses a layer-by-layer iterative search method. The algorithm is simple and clear. It does not require complicated theoretical derivation, and it is easy to implement. This algorithm is suitable for association rule mining in transactional databases, and it is also suitable for rare data sets. According to previous research, this algorithm is only suitable for association rule mining of rare data sets, that is, data sets less than a frequency and length.

The Aprioi algorithm is mainly composed of two processes: connection and pruning. In connection, assume that the items in the item set are sorted alphabetically. l1 and l2 are any two item sets in Lk-1; the condition that l1 and l2 can be connected is $(l_1[1] = l_2[1]) \Lambda (l_1[2] = l_2[2]) \Lambda \cdots \Lambda (l_1[k - 2] = l_2[k - 2]) \Lambda (l_1[k - 1] = l_2[k - 1])$, and the result itemset produced by connecting l1 [1] and l2 [2] is $l_1[1] l_2[2] \cdots\cdots l_1[k - 1] l_2[k - 1]$. Pruning is mainly to scan the database to determine the support count of each candidate set in C.

5.1.2. FP Growth Algorithm.
The FP growth algorithm is an effective algorithm for retrieving frequent itemsets in large databases. This algorithm does not generate a tedious process of generating candidate sets when exploring frequent element sets but uses a divide-and-conquer method. First, it compresses the database that provides frequent elements into a frequent model tree (FP tree), while it retains information related to the element set. Then, it divides the compressed database into a set of condition databases, each condition database is associated with a frequent item, and each condition database is retrieved separately.

The FP growth algorithm compresses the database and provides a set of repeated elements in the repeated model tree. Through two database checks, the frequency of the database is compressed into a complex pattern tree (CE tree) in a descending order according to its mean value, and frequent patterns are obtained by mining the frequent pattern tree. However, there are still performance bottlenecks in the algorithm when constructing the complex model tree. The characteristic of the FP growth algorithm is that it only needs to scan the database twice without creating candidate algorithms. Compared with that of the Apriori algorithm, the FP growth rate performance of this algorithm has been greatly improved, but the FP growth algorithm also has insufficient problems, and more researchers are needed to improve it.

However, the FP growth algorithm still has some shortcomings that need to be improved. Details are as follows:

(1) From the content of this article, we know that the FP growth algorithm compresses all records in the transaction database into a tree (FP tree). The more data content, the larger the content using the FP growth algorithm, so for large databases, it is unrealistic to construct a memory-based FP tree. Therefore, researchers have proposed a method of parallel processing of the database, by segmenting the database data, each segmentation point is individually mined, and finally, the results are merged

(2) According to the content of the article, the FP growth algorithm has a performance bottleneck in the process of mining frequent patterns. Therefore, the researchers proposed an improved FP growth algorithm and its distributed parallel implementation. They improved the FP growth algorithm and pruned the complete pattern tree based on the

frequent closed pattern itemset strategy to reduce the space search and improve the efficiency of algorithm mining

(3) According to the content of this paragraph, the FP growth algorithm is a classic association rule mining algorithm, but its existence query efficiency is low. When mining frequent FP tree, it needs to traverse FP tree repeatedly. Aiming at the above shortcomings, a method using two-dimensional tables combined with hash table technology is proposed to improve the traditional FP growth algorithm. The improved algorithm shortens the running time, and the smaller the support, the more obvious this advantage

(4) According to the content of the article, the FP growth algorithm needs to scan the database twice to compress the database of frequent itemsets into a frequent pattern tree. If it is a small database, the amount of calculation is not large, but for a large database, the time required to scan twice is very long, and the FP growth algorithm also requires a lot of memory. Therefore, researchers have proposed an improved FP tree frequent itemset mining algorithm. Through the test of database data and a comparative analysis with the traditional FP growth algorithm, it can be seen that the improved FP growth algorithm can save a lot of memory usage

The FP growth algorithm is an effective algorithm for retrieving frequent itemsets in large databases. In this process, the algorithm uses a divide-and-conquer method. Then, the specific steps of the divide and conquer method are as follows: first, scan the entire database to find frequent 1-itemsets, and then sort them according to the sort order of the number of endorsements. Use frequent 1-itemsets to create an item header and association based on frequent 1-itemsets in the FP tree in the second step. Point the table 1-itemsets and sort and reset the transaction of the entire database. Secondly, FP tree is composed of "Null" as the root point, so the entire database is divided into several conditional databases (projection database type), each database is associated with a frequent item set, and each conditional database is mined.

5.1.3. Partition Algorithm. The algorithm only needs to traverse the set transaction twice. The algorithm first divides the transaction set into multiple blocks so that each block can be loaded into memory. In the first traversal, all blocks are loaded into the memory one by one, and Apriori is used to find the collection of data item frequency sets within the range of each block. Because the search is limited to one block, these data item frequency sets are not the last frequency set of data items. These sets are combined to obtain a superset of the frequency set of a set of data items. In the second traversal, for each data item set obtained in the first traversal, the actual support is calculated. If the support is less than the support specified by the user, the data item

set is deleted. At the end of the second traversal, a collection of data item frequency sets is obtained.

5.1.4. Carma Algorithm. Some existing association rules require users to enter the minimum trust and minimum support before running the mining algorithm. The algorithm needs to run several times to determine whether the minimum confidence and minimum support are too high or too low. The Carma algorithm provides feedback to the user during the operation. Users can adjust the minimum support at any time based on the feedback information. If the user is satisfied with the output, the algorithm can be stopped at any time [12].

6. Application of Association Rule Mining Algorithm in Teaching Evaluation

In terms of teaching evaluation, teaching evaluation is an important part of education. It is a guarantee for guiding education, cultivating high-quality talents, helping society to make full use of educational achievements, and promoting the healthy development of education. Traditional teaching evaluation is mostly implemented with reference to relevant evaluation index systems and questionnaires, and it pays more attention to the results of evaluation. It is used as a basis for teacher promotion and evaluation. Applying association rules to teaching evaluation data will dig out some useful data, which will guide the teaching process and provide support for managers' decision-making.

6.1. Application of Association Rule Mining Algorithm in Student Courses. As the school continues to expand the scale of enrollment, the scale of teaching has also continued to expand, the expansion of majors has accelerated, and the amount of data has also increased geometrically. It is difficult for teaching administrators and class teachers to find out the relationship between previous courses and subsequent courses directly based on the distribution of student test scores and then make decisions such as making and modifying teaching plans based on this.

The correlation analysis of data mining is used here, mainly using association rules to analyze the rationality of the curriculum. For example, according to the analysis of association rules, the A course and the B course need to be set up in the C course. Whether students are unable to understand and accept a certain course, whether the market needs this course, and whether there are special requirements for a certain course, make reasonable arrangements that are conducive to improving student performance and employment. It is also possible to set whether the A course and B course have a contextual arrangement according to the mining mode of association rules, adjust the teaching plan, and promote student learning. It can also be judged based on this that the A course should not be arranged with the B course. This can avoid the negative effects of arranging two courses together and making students unable to grasp. In this way, the correlation between courses contained in a large amount of data can be effectively discovered, and the correlation between courses and courses can be discovered

through data mining of the database in the student performance management system.

6.2. Research on the Evaluation of University Informatization Teaching Management

6.2.1. Informatization Teaching Management Design. Informatization education design is based on learning theory, education theory, and communication theory to improve the process and method of students' academic performance. It can fully and appropriately use a variety of strategies and methods to develop modern information technology and information education design. It can organize all links and elements in the learning process scientifically and reasonably. Informatization education planning is known as the core of college teachers' informatization learning ability, which is also the theoretical basis and practical guide for carrying out educational activities under informatization conditions. At the same time, university teachers need to understand a certain educational process.

6.2.2. Implementation of Informatization Teaching Management. Under the influence of educational ideology in the information age, the specific implementation of information-based learning activities is an important part of college teachers' teaching. The implementation of information-based learning is a broad, professional, and flexible activity. Therefore, in addition to traditional teaching capabilities, it also puts forward requirements for college teachers' educational technology, teaching equipment, and professional experimental equipment. Relevant personnel should use the ability of technology to guide college students to use technology to practice, so that they can effectively manage the teaching and learning process, as well as the ability to express language, the ability to organize the classroom, and the ability to write on the blackboard. In the implementation of information-based learning, it is especially necessary to emphasize that in an information-based environment, it is necessary to cultivate the innovative potential of college students, the development and understanding of collaborative learning, and the ability of university teachers to explore learning.

6.2.3. Evaluation of Informatization Teaching Management. Information-based learning evaluation is one of the most powerful tools for college teachers' self-improvement and professional development. This is one of the most effective ways to understand the effectiveness of teaching and the degree of knowledge gained by students. It is necessary to comprehensively evaluate all aspects of the learning goals set by teachers and students to diagnose and correct problems in the learning process in time. School teachers should make full use of the results of the feedback. Evaluating the informatization of teaching is an important part of teachers' informatization teaching ability in the information age. Prior to this, the concept of educational evaluation has undergone major changes. Due to the limitations of traditional teaching evaluation methods, university teachers cannot do this. There is little or no attention to the educational evaluation process. Therefore, it is necessary to advocate multiple learning evaluations and individual learning evaluations.

7. Design of Teaching Management Evaluation System Based on Association Rule Mining Scientific Calculation

7.1. Analysis of Big Data Platform Architecture. The range of data stored is very wide. In data mining, the choice of fuzzy question will affect the actual value of data mining. Therefore, when designing the system, the functional requirements and various functional modules of the system should be clarified. In the operation process of the management information system, the orderly management and automatic processing of the computer can be used to reduce the waste of manpower and material resources. In this way, paperless office can be realized and a complicated work can be simplified. This can also liberate management personnel, allowing them to concentrate on other related tasks, thereby improving work efficiency and quality. The relevant personnel must fully guarantee the accuracy, timeliness, and consistency of the data, so as to ensure the accuracy, completeness, and dynamics of the various information resources provided by the system in real time. This enables colleges and universities to obtain quality assurance in terms of management administration, logistics education, and teaching [13].

The overall management level directly depends on the level of data collection and processing to a certain extent, but the complexity of college data management and the clear and reasonable division of responsibilities are closely linked. Nowadays, accurate and efficient data management methods are indispensable. Improving management transparency and cooperation between different departments requires the same or different departments in different jurisdictions. Employees interact online in real time to achieve common goals or perform their own tasks. Management requirements are the core of this system, so the goal of this system is to achieve key management functions with relatively complete functions.

7.2. Feasibility Analysis

7.2.1. Timing Feasibility. In terms of timing and feasibility, the preparation work for the development of the system is sufficient. Through the investigation of many domestic universities, we have a more comprehensive understanding of the advantages and disadvantages of the university management information system used by the universities. The system has been designed in detail. If the system is successfully developed, it can make college information management more standardized and rigorous, which will bring more efficient business processing and smarter data analysis and processing to colleges and universities, and make college data analysis and processing more efficient [14].

At present, foreign management information systems have been developed relatively well. Many enterprises and scientific research institutions have already used data mining

and other related technologies to analyze their internal data through specialized data analysis companies in order to obtain useful information. In contrast, the domestic management information system is still in the development stage, and it is in its infancy in terms of data correlation analysis and useful rule mining. Under such circumstances, China's university informatization management should fully adapt to the needs of development, and relevant personnel should accelerate the informatization process in order to be further improved.

7.2.2. Economic Feasibility. Evaluating the economic benefits of the developed project is the main purpose of the economic feasibility analysis. Based on the existing university management information system, the system has made many improvements and related improvements, such as handling data consistency between different departments and associations. For the development of the rule analysis module, the development funds of higher education can be obtained in an economic form. Once the system is developed and implemented, it can improve the quality and efficiency of employment and assist in the information management of colleges and universities. Therefore, it is economically feasible to develop the management information system of the University of Mining and Technology in accordance with federal regulations.

7.3. System Function. The teaching management and teaching operation involved in the teaching quality evaluation system include experts, teachers, students, and teaching management personnel. The activity flow chart of teaching quality evaluation is shown in Figure 3.

When analyzing system requirements, it is necessary to consider functional requirements and non-functional requirements. The so-called functional requirements are the description of the services that the system is expected to provide, how to respond to inputs, and the behavior of the system under specific circumstances. Nonfunctional requirements are types of requirements that are not directly related to the specific functions of the system [15]. The system management use case in Figure 4 also includes evaluation management, teacher management, administrator management, course management, class management, password management, and other functions, which are described in detail as shown in Figure 5.

7.4. System Architecture Design Model. Application of the three-tier Web model based on the BS approach: the three-tier BS architecture is shown in Figure 6.

8. Conclusions

This paper designs an information-based teaching management system that integrates association rule mining algorithms. The association rule mining algorithm can achieve the purpose of mining required information from massive data through certain association rules. The information mined by this method can be used for colleges and universities to implement informatization teaching. Management provides a scientific and effective basis. This article uses this method to excavate the required information from the database for the implementation of teaching management to realize information-based teaching management. Future research will continue to test the operating performance and other functional performance of this system, verify the practical applicability of this system, and expand the scope of application of this system.

Data Availability

The 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.

Acknowledgments

References

[1] Y. Suo, "Application of data mining technology in optimization of mobile communication network," *International Core Journal of Engineering*, vol. 7, no. 7, pp. 122–201, 2021.

[2] Z. Zhenyi, J. Zhou, G. G. Singh, A. Roobaea, M. Mehedi, and R. Saeed, "An improved association rule mining algorithm for large data," *Journal of Intelligent Systems*, vol. 30, no. 1, pp. 20–43, 2021.

[3] C. Yanan, "Research on data mining of intelligent education evaluation from the perspective of the knowledge graph," *Journal of Physics: Conference Series*, vol. 1955, no. 1, pp. 233–289, 2021.

[4] L. Jianguo and Z. Sheng, "Application research of data mining technology in personal privacy protection and material data analysis," *Integrated Ferroelectrics*, vol. 216, no. 1, pp. 13–58, 2021.

[5] H. Junzhong, "A study on the integration of computer network information security prevention and web data mining technology," *Journal of Physics: Conference Series*, vol. 1915, no. 3, pp. 344–402, 2021.

[6] R. J. Hoon, K. S. Il, L. K. Min, K. M. Kyum, and L. Y. Mook, "Optimization of position and number of hotspot detectors using artificial neural network and genetic algorithm to estimate material levels inside a silo," *Sensors (Basel, Switzerland)*, vol. 21, no. 13, pp. 21–34, 2021.

[7] L. A. Araújo, M. D. Rafael Menali Oliveira Jr., C. S. J. e. Silva, and L. R. Gomide, "Appropriate search techniques to estimate Weibull function parameters in a Pinus spp. plantation," *Journal of Forestry Research*, vol. 32, no. 6, pp. 2423–2435, 2021.

[8] Z. Guo, S. Yu, F. Huang, X. Fan, and J. Huang, "Landslide susceptibility zonation method based on C5.0 decision tree and K-means cluster algorithms to improve the efficiency of risk management," *Geoscience Frontiers*, vol. 12, no. 6, pp. 249–267, 2021.

[9] K. V. Dunaevskaya, L. V. Kiselev, and V. B. Kostousov, "Study of a method for calculating the current accuracy in map-aided navigation problem," *Gyroscopy and Navigation*, vol. 12, no. 1, pp. 50–60, 2021.

[10] R. Makara and W. L. Seung, "Evaluation of asphalt overlay pretreatments against reflective crack using association rule mining," *Journal of Transportation Engineering, Part B: Pavements*, vol. 147, no. 3, pp. 34–54, 2021.

[11] S. HaiYan, Z. Huan, and X. ZhiHao, "Research on personalized recommendation system based on association rules," *Journal of Physics: Conference Series*, vol. 1961, no. 1, pp. 110–204, 2021.

[12] Z. Yang, "Research on the application of university teaching management evaluation system based on Apriori algorithm," *Journal of Physics: Conference Series*, vol. 1883, no. 1, pp. 233–287, 2021.

[13] H. Liu, L. Hongcheng, and Z. Yang, "Research on the construction of teaching quality evaluation system," *Journal of Physics: Conference Series*, vol. 1673, no. 1, p. 012055, 2020.

[14] G.-h. Lei, "Research on the evaluation and monitoring system of college physical education based on artificial intelligence," *Journal of rResiduals Science & Technology*, vol. 13, no. 8, pp. 210–233, 2016.

[15] L. Hong, W. Li, and L. Dawei, "Inquiry learning evaluation system under the environment of cyber learning space," *Humanities and Social Sciences*, vol. 4, no. 2, pp. 42–274, 2016.

Research Article

Infrared and Visible Image Fusion Based on Iterative Control of Anisotropic Diffusion and Regional Gradient Structure

Jingyu Ji [ID],[1] **Yuhua Zhang** [ID],[1] **Zhilong Lin** [ID],[1] **Yongke Li** [ID],[1] **Changlong Wang** [ID],[1] **Yongjiang Hu** [ID],[1] **and Jiangyi Yao** [ID][2]

[1]*Department of UAV, Army Engineering University, Shijiazhuang, Hebei, China*
[2]*Equipment Simulation Training Center, Army Engineering University, Shijiazhuang, Hebei, China*

Correspondence should be addressed to Yongjiang Hu; jjy565827034@126.com

To improve the fusion performance of infrared and visible images and effectively retain the edge structure information of the image, a fusion algorithm based on iterative control of anisotropic diffusion and regional gradient structure is proposed. First, the iterative control operator is introduced into the anisotropic diffusion model to effectively control the number of iterations. Then, the image is decomposed into a structure layer containing detail information and a base layer containing residual energy information. According to the characteristics of different layers, different fusion schemes are utilized. The structure layer is fused by combining the regional structure operator and the structure tensor matrix, and the base layer is fused through the Visual Saliency Map. Finally, the fusion image is obtained by reconstructing the structure layer and the energy layer. Experimental results show that the proposed algorithm can not only effectively deal with the fusion of infrared and visible images but also has high efficiency in calculation.

1. Introduction

In recent years, UAVs have played an increasingly important role in many fields due to their high flexibility, low cost, and easy operation, which are often used for battlefield reconnaissance, battle situation assessment, target recognition, and tracking in the military. Now, image sensors in UAVS can acquire multiple types of images such as multispectral images, visible images, and infrared images [1]. However, due to the limitation of environmental conditions such as light, imaging with only one sensor will be affected by certain factors and cannot meet the requirements of practical applications. The combination of multiple imaging sensors can overcome the shortcomings of a single sensor and obtain more reliable and comprehensive information. The imaging sensors commonly used in UAVs are infrared sensors and visible sensors. The infrared sensors use the principle of thermal radiation to obtain images with larger infrared targets, but the targets are not clear and the edges are blurred [2]. The visible sensors use the principle of light reflection to obtain clear images with

clear details, but under low-visibility conditions, the images have limitations. Research has found that the effective combination of infrared images and visible images can result in a more comprehensive and accurate scene or target, which provides strong support for subsequent task processing [3].

The more widely used methods in the field of infrared and visible image fusion can be roughly classified into MST-based methods [4], sparse representation-based methods [5], spatial domain-based methods [6], and deep learning-based methods [7]. At present, the most researched and applied methods are MST-based methods, including wavelet transform [8], Laplacian pyramid transform [9], nonsubsampled shear wave transform [10], and nonsubsampled contourlet transform [11]. These methods decompose the source images in multiple scales, then fuse them separately according to certain fusion rules, and finally get the fusion result through inverse transformation, which can extract the salient information in the images and get better performance. For example, nonsubsampled contourlet transform is utilized by Huang et al. [11] to decompose the source images to obtain precise decomposition. However,

due to the lack of spatial consistency in the traditional MST methods, structural or brightness distortion may appear in the result.

In addition, image fusion methods with edge preserving filtering [12] are also receiving attention. Edge-preserving filtering can effectively reduce the halo artifacts around the edges in the fusion results while retaining the edge information of the image contour and has a good visual performance. Popular methods are mean filtering [13], bilateral filtering [14], joint bilateral filtering [15], and guided filtering [16]. These methods complete decomposition according to the spatial structure of the images to achieve spatial consistency, so as to achieve the purpose of smoothing the texture and preserving edge detail information. For example, Zhu et al. [16] proposed a novel fast single-image dehazing algorithm by using guided filtering to decomposition the images, and it obtained good performance. The edge-preserving fusion algorithms maintain spatial consistency and effectively improve the phenomenon of fusion image distortion or artifacts, but there are certain limitations: (1) it will introduce detail "halos" at the edges; (2) when the input images and the guide images are inconsistent, the filtering will be insensitive or even fail; and (3) it is difficult to meet the requirements of fusion performance, time efficiency, and noise robustness simultaneously.

Inspired by the previous research, this article focuses on reducing "halos" at the edges to retain the edge structure information and obtaining better decomposition performance in both noise-free and noise-perturbed images. In this paper, a new infrared and visible image fusion method based on iterative control of anisotropic diffusion and regional gradient structure operator is proposed. Anisotropic diffusion is utilized to deconstruct the source image into a structure layer and a base layer. Then, the structure layer is processed by using the gradient-based structure tensor matrix and the regional structure operator. Due to the weak detail and high energy of the base layer, the Visual Saliency Map (VSM) is utilized to fuse the base layer. By reconstructing the two prefusion components, the final fusion image can be obtained.

The main contributions of the proposed method can be summarized as follows:

(1) A novel method of infrared and visible image fusion is proposed. The anisotropic diffusion model with a control iteration operator is proposed to adaptively control the number of iterations, so the image is decomposed adaptively into a structure layer with rich edges and detail information and a base layer with pure energy information. Especially, the computational efficiency is greatly improved

(2) The regional structure operator is proposed into the structure tensor matrix, which can effectively extract information such as image details, contrast, and structure. It can also greatly improve the detection ability of weak structures and obtain structure images with good prefusion performance

(3) Since anisotropic diffusion can effectively deal with noise, the proposed method also has a good perfor-

mance on noisy image fusion. In addition, the algorithm is widely used and it is also suitable for other types of image fusion

The paper is organized as follows. Section 2 briefly reviews the anisotropic diffusion and structure tensor theory and introduces new operators. Section 3 describes the proposed infrared and visible image fusion algorithm in detail. Section 4 introduces related experiments and compares with several current advanced algorithms. Finally, the conclusion is discussed in Section 5.

2. Related Theories

2.1. Anisotropic Diffusion Based on Iterative Control. Anisotropic diffusion [17] can be utilized to smooth the image and maintain the image details and edge information. Compared with other filtering methods, it is more suitable for image decomposition processing. The anisotropic diffusion equation is expressed as

$$I_t = \text{div}\left(v(x, y, t)\nabla I\right) = v(x, y, t)\Delta I + \nabla v \cdot \nabla I, \tag{1}$$

where $v(x, y, t)$ is the flux function or diffusion rate of diffusion, ∇ is the Laplacian operator, Δ is the gradient operator, and t is the time or scale or iteration. Equation (1) can be regarded as a discrete square matrix, and the four nearest neighbor discretizations of Laplacian can be used:

$$I_{i,j}^{t+1} = I_{i,j}^t + \mu\left[v_N \cdot D_N I_{i,j}^t + v_S \cdot D_S I_{i,j}^t + v_W \cdot D_W I_{i,j}^t + v_E \cdot D_E I_{i,j}^t\right], \tag{2}$$

where $I_{i,j}^{t+1}$ is the coarser resolution image at $t + 1$ scale, which is influenced by $I_{i,j}^t$. μ is a constant with $0 \le \mu \le 1/4$. D_N, D_S, D_W, and D_E are the nearest difference values in the four directions of North, South, West, and East, respectively, which can be defined by

$$
\begin{aligned}
D_N I_{i,j} &\equiv I_{i-1,j} - I_{i,j}, \\
D_S I_{i,j} &\equiv I_{i+1,j} - I_{i,j}, \\
D_W I_{i,j} &\equiv I_{i,j-1} - I_{i,j}, \\
D_E I_{i,j} &\equiv I_{i,j+1} - I_{i,j},
\end{aligned}
\tag{3}
$$

and v_N, v_S, v_W, and v_E are the conduction coefficients or flux functions in the four directions of North, South, West, and East.

$$
\begin{aligned}
v_{N_{i,j}}^t &= g\left(\left\|(\nabla I)_{i+1/2,j}^t\right\|\right) = g\left(\left|D_N I_{i,j}^t\right|\right), \\
v_{S_{i,j}}^t &= g\left(\left\|(\nabla I)_{i-1/2,j}^t\right\|\right) = g\left(\left|D_S I_{i,j}^t\right|\right), \\
v_{W_{i,j}}^t &= g\left(\left\|(\nabla I)_{i,j-1/2}^t\right\|\right) = g\left(\left|D_W I_{i,j}^t\right|\right), \\
v_{E_{i,j}}^t &= g\left(\left\|(\nabla I)_{i,j+1/2}^t\right\|\right) = g\left(\left|D_E I_{i,j}^t\right|\right),
\end{aligned}
\tag{4}
$$

where $g(|\bullet|)$ is a monotonically decreasing function with $g(0) = 1$ and $g(\bullet)$ is the "edge stop" function or the differential coefficient, which has a very important influence on the noise suppression and edge retention ability of anisotropic diffusion. The image format in this paper is obtained by image processing technology.

$$g(\nabla I) = e^{-(\|\nabla I\|/k)^2},$$
$$g(\nabla I) = \frac{1}{1 + (\|\nabla I\|/k)^2}. \tag{5}$$

The scale space weighed by these two functions is different. The first function is for the abrupt areas with large gradients, namely, the edge and detail areas. The second function is for flat areas with small gradients. Both functions consist of a free parameter k.

The anisotropic diffusion is a differential iterative process, in which the number of iterations is a key issue. If it is overiterated, it will lead to oversmoothing; but if the number of iterations is not enough, the detail components cannot be separated effectively. Moreover, the number of iterations for noisy images and the number of iterations for noise-free images are also uncertain. Therefore, an iterative control operator θ is introduced to control k, thereby adaptively controlling the number of iterations and reasonably separating structural information such as gradients and details. And it can also be improved in computational efficiency.

$$k = \begin{cases} (\theta + 1)K_0, & \text{if } \max(D_N, D_S, D_W, D_E) > 0, \\ (1 - \theta)K_0, & \text{others,} \end{cases} \tag{6}$$

where K_0 is the empirical value for controlling the diffusion strength, which usually is set by 30. It can be seen from Equation (6) that the value of k is related to the edge strength of the region boundary, and the value of k is updated through positive and negative excitation by θ to obtain the optimal number of iterations. Get the most effective and accurate separation results.

The anisotropic diffusion of the image I is simply represented by aniso(I). After the image is diffused through anisotropy, since the iterative control operator can precisely control the number of iterations, almost all the vibration and repetitive context can be effectively preserved in the structure layer, while the energy information and weak edges are preserved in the base layer. Figure 1 shows the base layer and structure layer images obtained after anisotropic diffusion decomposition. It can be clearly seen that the images are basically consistent with the theoretical analysis.

2.2. Gradient-Based Structure Tensor Matrix. Gradient is the rate of change, which is reflected by the difference between a central pixel and surrounding pixels. It can be used to accurately reflect the texture details, contour features, and structural components in the image. The structure tensor is an effective method to analyse the gradient problem, and it has been applied to a variety of image processing tasks.

Input   Base   Structure
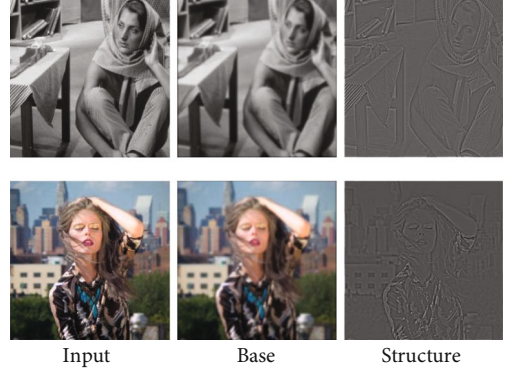
FIGURE 1: The results of anisotropic diffusion decomposition.

The gradient operator [18] is described as follows. For a local window $\Theta(x, y)$ of any $\varepsilon \longrightarrow 0^+$ in the direction β, the square of the change of the image $I(x, y)$ at the point (x, y) is

$$(dI)^2 = \|I(x + \varepsilon \cos \beta, y + \varepsilon \sin \beta) - I(x, y)\|_2^2 \tag{7}$$
$$\approx \sum_{\Theta(x,y)} \left(\frac{\partial I}{\partial x} \varepsilon \cos \beta + \frac{\partial I}{\partial y} \varepsilon \sin \beta \right).$$

In any direction β at the point (x, y), the change rate $C(\beta)$ of the local features of the image $I(x, y)$ is

$$C(\beta) = \sum_{\Theta(x,y)} \left(I_x \cos \beta + I_y \sin \beta \right)^2$$
$$= (\cos \beta, \sin \beta) \begin{bmatrix} \sum_{\Theta(x,y)} \left(\frac{\partial I}{\partial x} \right)^2 & \sum_{\Theta(x,y)} \left(\frac{\partial I}{\partial x} \frac{\partial I}{\partial y} \right) \\ \sum_{\Theta(x,y)} \left(\frac{\partial I}{\partial y} \frac{\partial I}{\partial x} \right) & \sum_{\Theta(x,y)} \left(\frac{\partial I}{\partial y} \right)^2 \end{bmatrix} (\cos \beta, \sin \beta)^T. \tag{8}$$

To make better analysis of gradient features and effectively realize image processing, the structure tensor matrix S is introduced. And $C(\beta)$ can be expressed as

$$C(\beta) = (\cos \beta, \sin \beta) S (\cos \beta, \sin \beta)^T, \tag{9}$$

where

$$S = \begin{bmatrix} M & N \\ N & O \end{bmatrix},$$
$$M = \sum_{\Theta(x,y)} \left(\frac{\partial I}{\partial x} \right)^2,$$
$$N = \sum_{\Theta(x,y)} \left(\frac{\partial I}{\partial x} \frac{\partial I}{\partial y} \right),$$
$$O = \sum_{\Theta(x,y)} \left(\frac{\partial I}{\partial y} \right)^2. \tag{10}$$

FIGURE 2: Schematic diagram of the proposed algorithm.

FIGURE 3: Six pairs of source images.

The two extreme values of the structure tensor S can be expressed as

$$\lambda_{1,2} = \frac{1}{2}\left((M+O) \mp \sqrt{(M-O)^2 + (2N)^2}\right). \tag{11}$$

The structural characteristics of the local area of the image are related to the extreme value of the matrix. Generally, if the two extreme values are relatively small, it indicates that the region does not have gradient 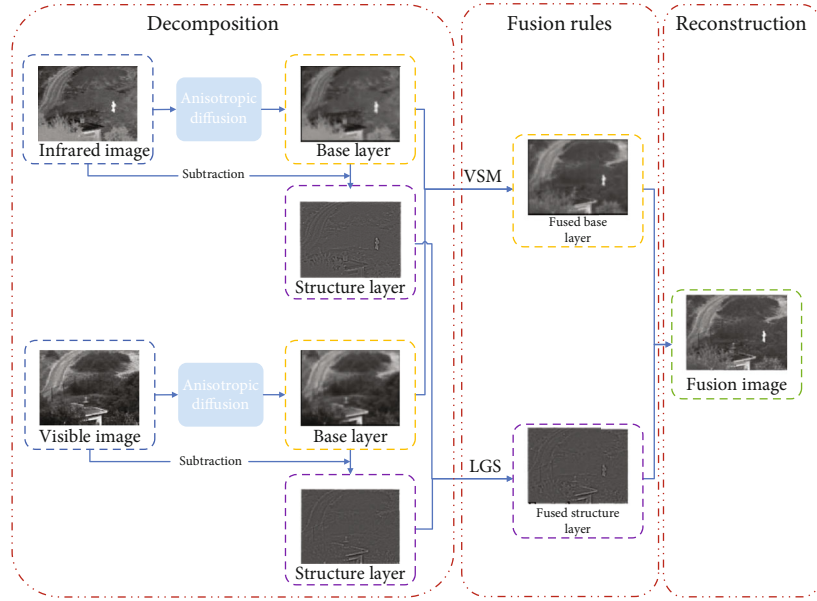characteristics; that is, the region is located in the isotropic part. Otherwise, it means that the local area of the area has obvious changes and contains certain structural information, because in the image area saliency measurement, a wide range of structure types are involved. Finally, the structural saliency operator SSO is defined according to [19] as

$$\text{SSO} = \sqrt{(\lambda_1 + \lambda_2)^2 + 0.5(\lambda_1 - \lambda_2)^2}. \tag{12}$$

3. Fusion Framework

Based on the above theories, a new image fusion framework is constructed, as shown in Figure 2. Different from the tra-

ditional decomposition scheme, in order to make better use of the useful information in the original image, first, the iterative control anisotropic diffusion is utilized to decompose the source image into base and structure components. At this time, most of the gradients and edges can be effectively preserved in the structure layer, and the base layer contains the remaining energy information. Then, according to the characteristics of each layer, different fusion rules are introduced to acquire the prefusion of each layer. Among them, for the fusion of the structure layer, the prefusion is effectively realized through the regional gradient structure; for the base layer, the prefusion is performed through the VSM. Finally, the fusion result is obtained by reconstructing the two prefusion layers.

3.1. Anisotropic Decomposition.
Let the source images $\{I_n(x,y)\}_{n=1}^N$ be all coregistered. The base layer is obtained through the anisotropic diffusion model in the previous section with smooth edges:

$$I_n^B(x,y) = \text{aniso}(I_n(x,y)), \tag{13}$$

where $I_n^B(x,y)$ is the nth base layer and $\text{aniso}(I_n(x,y))$ represents the anisotropic diffusion process on the nth source
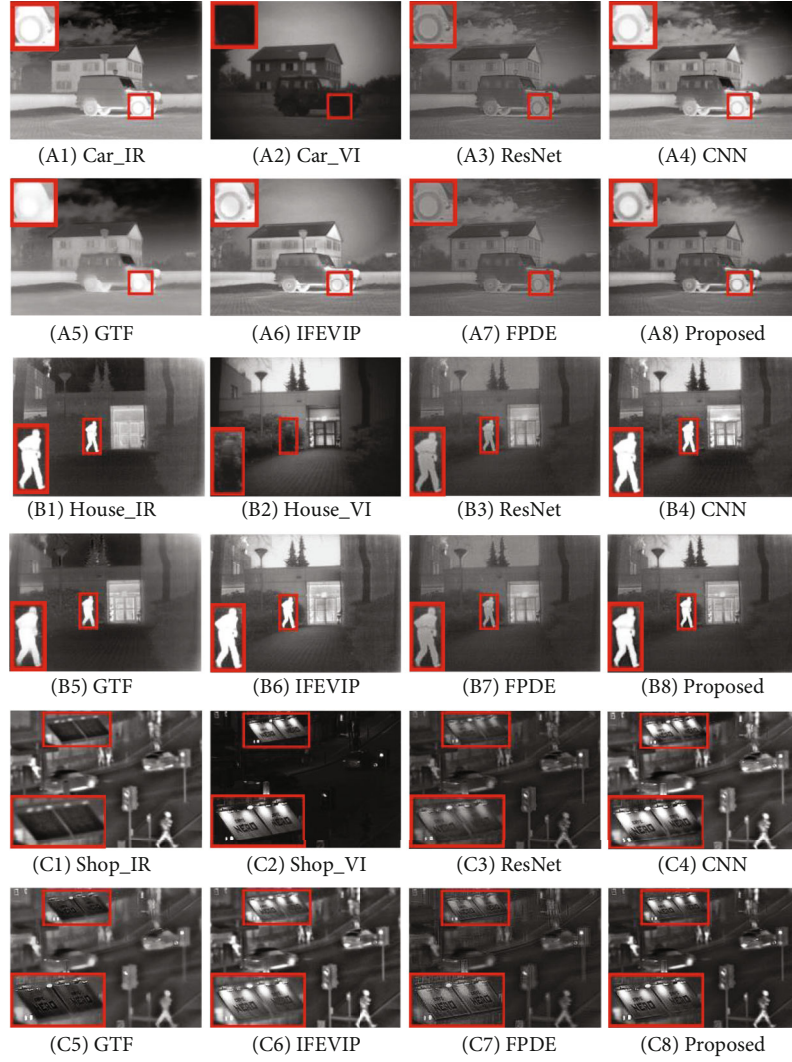
FIGURE 4: Fused results of infrared/visible images.

image. The structure layer is obtained by subtracting the base layer from the source image.

$$I_n^S(x, y) = I_n(x, y) - I_n^B(x, y). \tag{14}$$

After anisotropic decomposition, a structure layer with rich outline and texture details and a base layer with intensity information can be obtained.

3.2. Fusion of Structure Layers. Since the structure saliency operator (SSO) in the previous section can effectively detect the gradient structure information of the images, SSO can be used to prefuse the structure layers. However, due to the lack of intensity variables, SSO cannot accurately detect the weak feature information in the images. In order to improve the structure detection ability, the regional structure operator (RSO) is introduced to improve the performance of SSO. RSO is the regional structural component with (x, y) as the center position; then, the regional gradient structure (RGS) can be expressed as

$$RGS_I(x, y) = RS_I(x, y) \cdot SS_I(x, y), \tag{15}$$

where $SS_I(x, y)$ is the salient image produced by SSO, and $RS_I(x, y)$ represents the regional structure feature at position (x, y), which can be expressed as

$$RS_I(x, y) = \sum_{-N}^{N} \sum_{-N}^{N} I(x - N, y - N), \tag{16}$$

where N controls the size of the region and influences the efficiency and effect of fusion. Through comparing the RGS of the input image, the structure saliency map $M_1(x, y)$ of the image $I_1^S(x, y)$ is calculated:

$$M_1(x, y) = \begin{cases} 1, & \text{if } RGS_1(x, y) > RGS_2(x, y), \\ 0, & \text{otherwise.} \end{cases} \tag{17}$$

$M_2(x, y)$ can be obtained in the same way. In addition, $M_n(x, y)$ can be future refined by

(A1) Snow_IR	(A2) Snow_VI	(A3) ResNet	(A4) CNN
(A5) GTF	(A6) IFEVIP	(A7) FPDE	(A8) Proposed
(B1) Tree_IR	(B2) Tree_VI	(B3) ResNet	(B4) CNN
(B5) GTF	(B6) IFEVIP	(B7) FPDE	(B8) Proposed
(C1) Walking night_IR	(C2) Walking night_VI	(C3) ResNet	(C4) CNN
(C5) GTF	(C6) IFEVIP	(C7) FPDE	(C8) Proposed
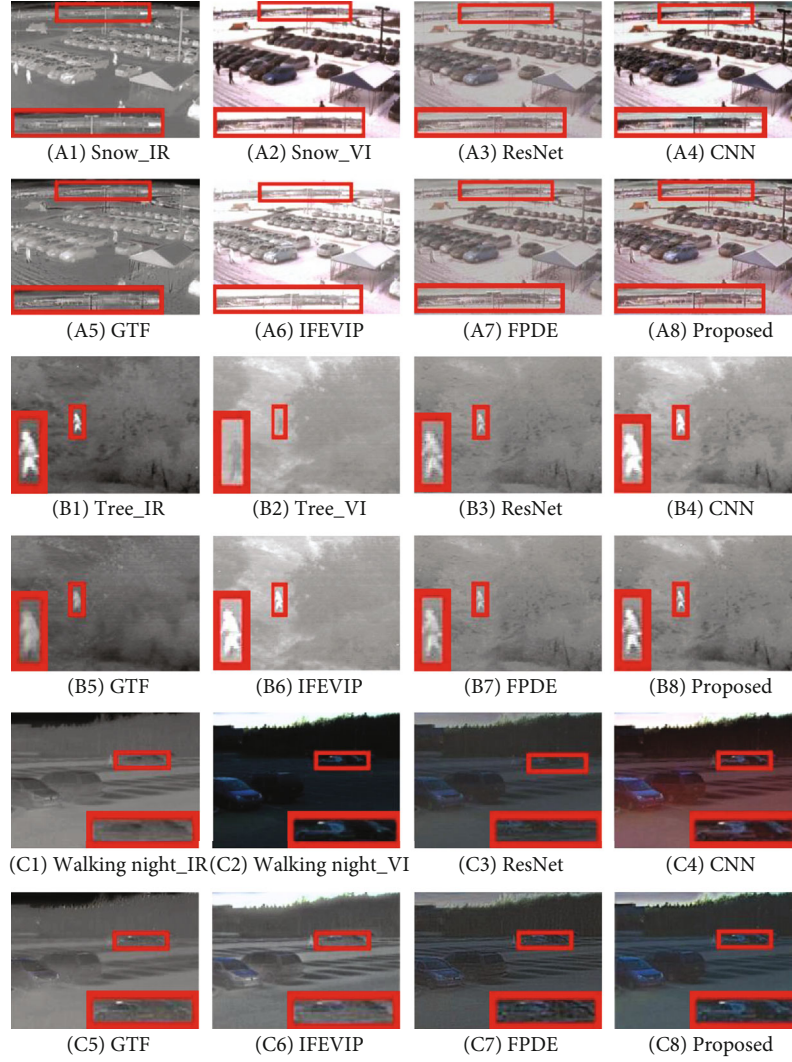
FIGURE 5: Fused results of infrared/visible images.

$$\tilde{M}_I(x,y) = \begin{cases} 1, & \text{if } \sum_{(a,b)\in\Theta} M_I(x+a,y+b) > 0.5 \times \text{size}(\Theta), \\ 0, & \text{otherwise,} \end{cases} \tag{18}$$

where Θ is a central local area in (x,y) whose size is $T \times T$. Therefore, the prefusion structure layer image $F_I^S(x,y)$ can be expressed by

$$F_I^S(x,y) = \tilde{M}_1(x,y)\bullet I_1^S(x,y) + \tilde{M}_2(x,y)I_2^S(x,y). \tag{19}$$

3.3. Fusion of Base Layers. Since the base layers contain less details, the weighted average technology based on VSM [20] is used to fuse the base layer F_I^B.

First, VSM is constructed; let I_P represent the intensity value of a pixel p in the image I. The saliency value $V(p)$ of pixel p is defined as

$$V(p) = \sum_{j=0}^{L-1} M_j |I_P - I_j|, \tag{20}$$

where j represents the pixel intensity, M_j represents the number of pixels whose intensity is equal to j, and L represents the number of gray levels (in this case, 256). If two pixels have the same intensity value, their saliency values are equal. Then, normalize $V(p)$ to [0,1].

Let V_1 and V_2 denote the VSM of different source images, and $I_1^B(x,y)$ and $I_2^B(x,y)$ denote the base layer images of different source images, and the final prefusion base layer image is obtained by weighted average

$$F_I^B(x,y) = \left(0.5 + \frac{V_1 - V_2}{2}\right) I_1^B(x,y) + \left(0.5 - \frac{V_1 - V_2}{2}\right) I_2^B(x,y). \tag{21}$$

After obtaining these two prefusion components, the final fusion image F_I is

TABLE 1: Quantitative index of image fusion results.

Source images	Index	ResNet	CNN	GTF	TIF	FPDE	Proposed
Car	EN	6.798	6.519	7.119	7.090	6.569	**7.564**
	Q_{ABF}	0.313	0.349	0.246	0.398	0.299	**0.401**
	Q_{CB}	0.401	0.395	0.337	0.367	0.395	**0.439**
	MI	1.735	1.082	1.243	1.638	1.432	**2.866**
	SSIM	1.543	1.372	1.310	1.330	1.435	**1.988**
	PSNR	58.253	58.430	58.127	57.280	58.349	**59.124**
House	EN	6.592	6.618	6.952	6.960	6.870	**6.998**
	Q_{ABF}	0.375	0.345	0.253	0.402	0.387	**0.418**
	Q_{CB}	0.463	0.471	0.338	0.468	0.477	**0.490**
	MI	1.301	1.251	1.080	1.417	1.222	**1.456**
	SSIM	1.534	1.436	1.374	1.418	1.305	**1.991**
	PSNR	58.143	**59.462**	59.066	58.225	58.356	59.184
Shop	EN	6.275	6.062	6.627	6.266	6.403	**6.703**
	Q_{ABF}	0.461	0.205	0.481	0.589	0.328	**0.598**
	Q_{CB}	0.478	0.451	0.434	0.481	0.445	**0.497**
	MI	1.723	0.812	1.746	1.682	1.308	**1.896**
	SSIM	1.334	1.031	1.221	1.325	1.299	**1.993**
	PSNR	59.012	59.885	59.466	59.308	59.345	**59.936**
Snow	EN	7.012	6.818	5.919	6.911	6.918	**7.723**
	Q_{ABF}	0.543	0.551	0.498	0.577	0.565	**0.579**
	Q_{CB}	0.554	0.483	0.478	0.517	0.612	**0.661**
	MI	2.045	1.914	1.598	2.368	1.999	**2.657**
	SSIM	1.367	1.238	1.164	1.222	1.029	**1.989**
	PSNR	55.034	56.630	56.122	55.237	56.856	**57.083**
Tree	EN	6.744	6.696	6.696	6.229	6.801	**6.962**
	Q_{ABF}	0.301	0.311	0.348	0.341	0.333	**0.379**
	Q_{CB}	0.423	0.424	0.417	0.430	0.415	**0.469**
	MI	1.354	1.109	1.452	1.090	1.027	**1.635**
	SSIM	1.578	1.512	1.469	1.483	1.496	**1.995**
	PSNR	58.013	57.641	57.242	57.325	57.934	**58.823**
Walking Night	EN	6.348	5.734	6.059	6.281	6.786	**7.011**
	Q_{ABF}	0.405	0.346	0.309	0.471	0.418	**0.474**
	Q_{CB}	0.313	0.353	0.291	0.344	0.350	**0.386**
	MI	2.019	2.130	2.189	2.139	2.005	**2.335**
	SSIM	1.346	1.153	1.232	1.208	1.206	**1.989**
	PSNR	55.999	56.047	55.843	54.825	55.470	**56.943**

$$F_I = F_I^S(x, y) + F_I^B(x, y). \tag{22}$$

4. Experimental Analysis and Results

In order to verify the effectiveness and reliability of the algorithm in this paper, multiple pairs of images are utilized for experimental verification, and the results are analysed through subjective vision and objective quantitative evaluation. After setting the algorithm parameters, the experimental results are displayed and discussed.

4.1. Experimental Setting. As shown in Figure 3, six pairs of source images are employed in the experiment, which can be obtained from the public websitehttp://imagefusion.org/. All the experiments are implemented using MATLAB 2018a on a notebook PC. And five recent methods are compared in the same experimental environment for verification, such as image fusion with ResNet and zero-phase component analysis (ResNet) proposed by Li et al. [21], image fusion with the convolutional neural network (CNN) proposed by Liu et al. [22], gradient transfer and total variation minimization-based image fusion method (GTF) proposed by Ma et al. [23], image
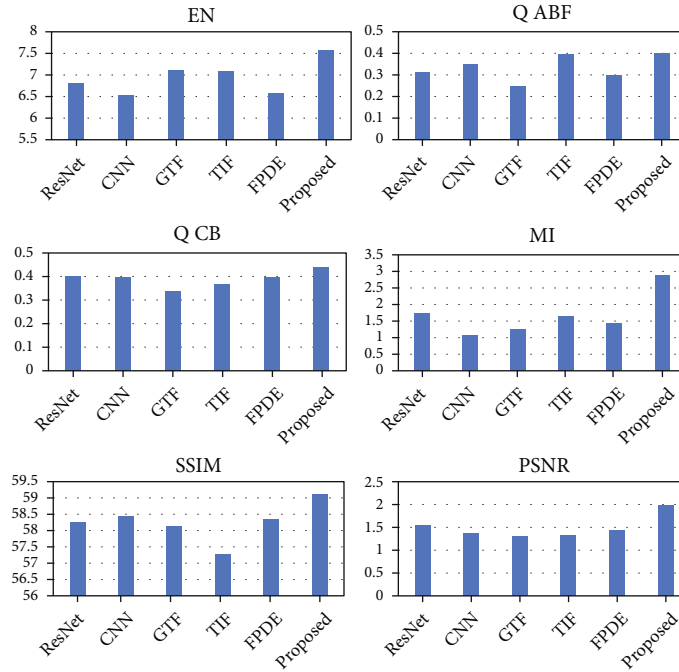
FIGURE 6: The bar chart comparison of EN, $Q_{AB/F}$, Q_{CB}, MI, SSIM, and PSNR values of various fusion methods for the car example.
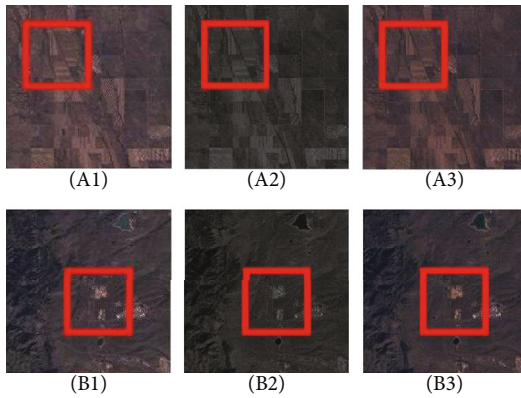

FIGURE 7: Fused results of multispectral/panchromatic images.

fusion through infrared feature extraction and visual information preservation (IFEVIP) proposed by Zhang et al. [24], and multisensor image fusion based on fourth-order partial differential equations (FPDE) proposed by Bavirisetti et al. [25]. In addition, the fusion performance is quantitatively evaluated by six indicators, including entropy (EN) [26], edge information retention ($Q_{AB/F}$) [27], Chen-Blum's index (Q_{CB}) [28], mutual information (MI) [29], structural similarity (SSIM) [30], and peak signal-to-noise ratio (PSNR) [31].

4.2. Image Fusion and Evaluation. Figures 4 and 5 are six pairs of infrared and visible image fusion examples. Figures 4(a1), 4(b1), and 4(c1) and Figures 5(a1), 5(b1), and 5(c1) are infrared images, and Figures 5(a2), 5(b2), and 5(c2) and Figures 4(a1), 4(b1), and 4(c1) are infrared images. Figures 4(a2), 5(b2), and 5(c2) are visible images; Figures 4(a3)–4(a8), 4(b3)–4(b8), and 4(c3)–4(c8) and Figures 5(a3)–5(a8), 5(b3)–5(b8), and 5(c3)–

5(c8) are the fusion results obtained by different methods. The content in the red box in the figure is the part to be emphasized.

4.2.1. Subjective Evaluation. It can be seen from Figures 4 and 5 that the fusion images obtained by the ResNet and GTF methods have lower contrast than the results obtained by the proposed method. Although the structure is better preserved, the details are relatively weakened and lost. The IFEVIP method maintains a good contrast, but the visual effect is too enhanced, especially in the partially enlarged areas, resulting in obvious error in the result. The FPDE method has the phenomenon of blurred internal features. The CNN method has obtained a relatively good fusion result, but its image is somewhat unnatural, and the colour of the result in Figure 5(c4) contains errors. Therefore, the proposed method can effectively separate the component information of different images, preserve the useful information of the source images into the fusion images, and obtain the best visual performance in the aspect of edge and detail preservation.

4.2.2. Objective Evaluation. Except for subjective evaluation, the fusion results are quantitatively evaluated, and the results are shown in Table 1, in which the best results are labelled in bold. According to the data in the table, it can be seen that the objective evaluation of the proposed method is significantly higher than other methods. In all quantitative evaluations, only a few places are not optimal, but they do not affect the advantages of the method in this paper. In addition, Figure 6 shows the bar chart comparison of EN, $Q_{AB/F}$, Q_{CB}, MI, SSIM, and PSNR values of various fusion methods for the car example.

In summary, for infrared and visible fusion, the method in this paper has a good performance both subjectively and objectively.
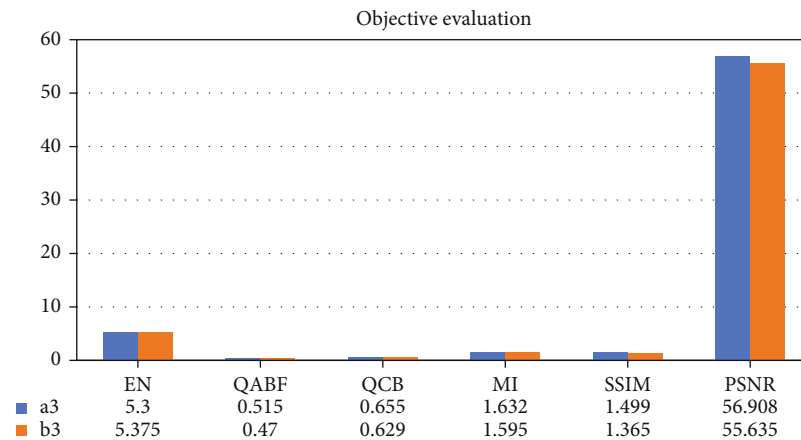
FIGURE 8: The bar chart comparison of EN, $Q_{AB/F}$, Q_{CB}, MI, SSIM, and PSNR values of the proposed method for the multispectral/panchromatic images.

TABLE 2: Computational efficiency of different methods.

Method	ResNet	CNN	GTF	IFEVIP	FPDE	Proposed
Time (s)	23.07	23.03	2.91	1.34	5.78	0.78

4.3. Extended Experiment. After experimental verification, the proposed fusion algorithm is equally effective for remote sensing images. To illustrate the effectiveness of this algorithm, two different sets of panchromatic and multispectral satellite remote sensing images are shown in Figure 7.

Figures 7(a1) and 7(b1) are multispectral images with high spectral resolution and low spatial resolution. Figures 7(a2) and 7(b2) are panchromatic images with high spatial resolution and low spectral resolution. The corresponding fusion results are shown in Figures 7(a3) and 7(b3). As can be seen from the content in the red box in Figure 7, the fusion results have both high spatial resolution and high spectral resolution, and the fused images have a strong ability to express structure and details. The objective evaluation results are shown in Figure 8. It can be seen from the visual and objective results that this algorithm can effectively retain high-spatial and hyperspectral information and can improve the accuracy of subsequent processing of remote sensing images.

4.4. Computational Efficiency. The methods tested in this paper are all carried out in the same experimental environment. The average implementation time of six pairs of images is compared as shown in Table 2. It can be seen that the calculation efficiency of the proposed algorithm has a considerable advantage over the comparison algorithms.

5. Conclusions

In this paper, an infrared and visible image fusion algorithm based on iterative control of anisotropic diffusion and regional gradient structure is proposed. The algorithm makes full use of the advantages of anisotropic diffusion and improves the decomposition efficiency and effect through iterative control operators. The regional gradient structure operator is intro-

duced to fully extract the detailed information in the structure layer to obtain a better fusion performance. Many experimental results show that this algorithm is significantly better than existing methods in terms of subjective and objective evaluation. In addition, higher calculation efficiency and stronger antinoise performance can be obtained, and the algorithm can be effectively applied to other types of image fusion situations.

Data Availability

The data used to support the findings of this paper are available from http://imagefusion.org/.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

References

[1] Y. Liu, L. Dong, Y. Chen, and W. Xu, "An efficient method for infrared and visual images fusion based on visual attention technique," *Remote Sensing*, vol. 12, no. 5, p. 781, 2020.

[2] D. Paquin, D. Levy, and L. Xing, "Hybrid multiscale landmark and deformable image registration," *Mathematical Biosciences & Engineering*, vol. 4, no. 4, pp. 711–737, 2017.

[3] X. Zhang, A. Zhang, and X. Meng, "Automatic fusion of hyperspectral images and laser scans using feature points," *Journal of Sensors*, vol. 2015, Article ID 415361, 9 pages, 2015.

[4] X. Yan, H. Qin, J. Li, H. Zhou, J. G. Zong, and Q. Zeng, "Infrared and visible image fusion using multiscale directional nonlocal means filter," *Applied Optics*, vol. 54, no. 13, p. 4299, 2015.

[5] Z. Zhu, H. Yin, Y. Chai, Y. Li, and G. Qi, "A novel multi-modality image fusion method based on image decomposition and sparse representation," *Information Sciences*, vol. 432, pp. 516–529, 2018.

[6] Y. Li, F. Li, B. Bai, and Q. Shen, "Image fusion via nonlocal sparse K-SVD dictionary learning," *Applied Optics*, vol. 55, no. 7, pp. 1814–1823, 2016.

[7] F. Salvetti, V. Mazzia, A. Khaliq, and M. Chiaberge, "Multi-image super resolution of remotely sensed images using residual attention deep neural networks," *Remote Sensing*, vol. 12, no. 14, p. 2207, 2020.

[8] A. Yehia, M. Safy, and A. S. Amein, "Fusion of high-resolution SAR and optical imageries based on a wavelet transform and IHS integrated algorithm," *International Journal of Engineering Research in Africa*, vol. 52, pp. 62–72, 2021.

[9] T. Rabie, M. Baziyad, and I. Kamel, "Enhanced high capacity image steganography using discrete wavelet transform and the Laplacian pyramid," *Multimedia Tools and Applications*, vol. 77, no. 18, pp. 23673–23698, 2018.

[10] M. Zhao and Y. Peng, "A multi-module medical image fusion method based on non-subsampled shear wave transformation and convolutional neural network," *Sensing and Imaging*, vol. 22, no. 1, 2021.

[11] X. Huang, G. Qi, H. Wei, Y. Chai, and J. Sim, "A novel infrared and visible image information fusion method based on phase congruency and image entropy," *Entropy*, vol. 21, no. 12, p. 1135, 2019.

[12] C. Q. Kang, Y. Zheng, and Y. X. Liu, "Multi-focus image fusion algorithm based on edge-preserving filtering," *Laser & Infrared*, vol. 43, no. 7, pp. 814–817, 2013.

[13] J. Duan, C. Long, and C. Chen, "Multifocus image fusion using superpixel segmentation and superpixel-based mean filtering," *Applied Optics*, vol. 55, no. 36, pp. 10352–10362, 2016.

[14] C. Bo, "Infrared and visible images fusion based on gradient bilateral filtering," in *2016 3rd International Conference on Systems and Informatics (ICSAI)*, Shanghai, China, 2016IEEE.

[15] T. Luo, Z. Liu, Z. Pan, and M. Zhang, "A virtual-real occlusion method based on GPU acceleration for MR," in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, Osaka, Japan, 2019IEEE.

[16] Z. Zhu, H. Wei, G. Hu, Y. Li, G. Qi, and N. Mazur, "A novel fast single image dehazing algorithm based on artificial multi-exposure image fusion," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–23, 2021.

[17] P. Perona and J. Malik, "Scale space and edge detection using anisotropic diffusion," *IEEE Computer Society Workshop on Computer Vision*, vol. 12, no. 7, pp. 629–639, 1990.

[18] L. Jin, H. Liu, X. Xu, and E. Song, "Improved direction estimation for Di Zenzo's multichannel image gradient operator," *Pattern Recognition*, vol. 45, no. 12, pp. 4300–4311, 2012.

[19] Z. Zhou, S. Li, and B. Wang, "Multi-scale weighted gradient-based fusion for multi-focus images," *Information Fusion*, vol. 20, pp. 60–72, 2014.

[20] M. M. Cheng, N. J. Mitra, X. Huang, P. H. 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, 2014.

[21] H. Li, X. J. Wu, and T. S. Durrani, "Infrared and visible image fusion with ResNet and zero-phase component analysis," *Infrared Physics & Technology*, vol. 102, article 103039, 2018.

[22] Y. Liu, X. Chen, J. Cheng, H. Peng, and Z. Wang, "Infrared and visible image fusion with convolutional neural networks," *International Journal of Wavelets, Multiresolution and Information Processing*, vol. 16, no. 3, p. 1850018, 2018.

[23] J. Ma, C. Chen, C. Li, and J. Huang, "Infrared and visible image fusion via gradient transfer and total variation minimization," *Information Fusion*, vol. 31, pp. 100–109, 2016.

[24] Y. Zhang, L. Zhang, X. Bai, and L. Zhang, "Infrared and visual image fusion through infrared feature extraction and visual information preservation," *Infrared Physics & Technology*, vol. 83, pp. 227–237, 2017.

[25] D. P. Bavirisetti, G. Xiao, and G. Liu, "Multi-sensor image fusion based on fourth order partial differential equations," in *2017 20th International Conference on Information Fusion (Fusion)*, Xi'an, China, 2017IEEE.

[26] Y. Chibani, "Additive integration of SAR features into multispectral SPOT images by means of the a trous wavelet decomposition," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 60, no. 5, pp. 306–314, 2006.

[27] C. S. Xydeas and V. Pv, "Objective image fusion performance measure," *Military Technical Courier*, vol. 56, no. 4, pp. 181–193, 2000.

[28] C. Yin and R. S. Blum, "A new automated quality assessment algorithm for night vision image fusion," in *2007 41st Annual Conference on Information Sciences and Systems*, Baltimore, MD, USA, 2007.

[29] G. Qu, D. Zhang, and P. Yan, "Information measure for performance of image fusion," *Electronics Letters*, vol. 38, no. 7, pp. 313–315, 2002.

[30] W. Zhou, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[31] P. Jagalingam and A. V. Hegde, "A review of quality metrics for fused image," *Aquatic Procedia*, vol. 4, pp. 133–142, 2015.

Research Article

The Improved Nonparametric Regression Model for the IoT Link Load Balancing Control Algorithm

Qingping Xue [ID]

Henan University of Animal Husbandry and Economy, Zhengzhou 450044, China

Correspondence should be addressed to Qingping Xue; 81017@hnuahe.edu.cn

In order to improve the load balance control effect of the Internet of Things link, this paper combines the nonparametric regression model to improve the load balancing algorithm of the Internet of Things link. Moreover, this paper proposes a load balancing research strategy based on data plane data flow, which is aimed at improving the load balancing problem of data flow in the link. The data center network uses a multilayer fat tree topology to store information in the flow table corresponding to the switch for data flow processing and forwarding operations. In addition, this paper constructs a load balancing model for intelligent Internet of Things link and verifies the model in this paper through experimental research. The research results show that the load balancing control algorithm for Internet of Things link based on the nonparametric regression model proposed in this paper can effectively improve the internal scheduling of the Internet of Things link system and promote the load balancing effect.

1. Introduction

With the in-depth study of WSN by scientific researchers, they have found that it has a wide range of applications but also has shortcomings, such as sensor node energy, communication distance, computing, and storage capabilities that are very limited. Since each sensor node is powered by a battery with limited energy, the complicated deployment environment brings a lot of inconvenience to the battery replacement work [1]. In addition, with the exhaustion of sensor node energy and the addition of new sensor nodes, the topology of the network will change. In the traditional network hierarchy, only two adjacent layers can communicate with each other, and it is difficult to share information with each other between two nonadjacent layers. Therefore, there are great limitations, which can no longer be well adapted to the development needs of WSN [2]. How to design a good performance energy optimization algorithm to meet the high requirements of WSN on the network life cycle has become the research focus and difficulty of WSN in recent years [3].

High-rate wireless sensor network (HRWSN) integrates information perception, processing, and transmission and is widely used in military, industrial and agricultural control, biomedicine, environmental monitoring, disaster relief, and other fields. It is a key technical link from sensor networks to the Internet of Things and pervasive computing. In order to meet high-quality video surveillance, diversified information collection, complex task processing, high-precision positioning, and other applications, it is urgent to introduce information-rich images, audio, and video into high-rate sensor networks. High-rate sensor networks have the characteristics of rich sensing information, complex processing tasks, node movement, dynamic changes in topology, and severe energy limitation.

This paper combines the nonparametric regression model to improve the Internet of Things link load balancing algorithm, builds an intelligent model, and verifies the effectiveness of this model through research.

2. Related Work

Firstly, the research and practical application of WSN related theories are reviewed. Researchers have actively carried out research and practical application of WSN related theories and have put forward energy optimization schemes

for sensor nodes one after another. Literature [4] proposed the LEACH (Low Energy Adaptive Clustering Hierarchy) algorithm, a typical representative of clustering routing algorithms. Through periodic random selection of cluster heads, each node has an equal probability of being a cluster head, and the network energy consumption is balanced to the entire network; [5] uses simulated annealing algorithm to divide the clusters, and the remaining energy of the nodes is considered when selecting the cluster heads; the number of times the nodes serve as cluster heads is considered in the literature [6]; the pLEACH algorithm proposed in the literature [7] divides the network, and the node with the highest remaining energy is selected as the cluster head in each subarea; literature [8] balances energy consumption by adjusting the data transmission rate; literature [9] improves the selection method of cluster heads and determines the remaining energy of nodes and neighbor nodes. The number and location information are considered; literature [10] uses fuzzy theory to calculate the probability of a node as a cluster head to improve the threshold formula; the LEACH-EC algorithm proposed in literature [11] optimizes the selection of cluster heads through the remaining energy and connectivity of the node In recent years, some new intelligent algorithms have also been proposed. Literature [12] uses particle swarm optimization (PSO) to optimize the selection of cluster heads; literature [13] uses the ant colony algorithm to optimize data transmission. Path. Literature [14] based on the cell membrane optimization algorithm to complete the clustering of nodes through the three factors of concentration, energy, and distance; in addition, some researchers have proposed new energy supply methods, but they are limited by the cost of WSN, and it is difficult to realize the consideration of volume. In literature [15], wireless charging is used for the sensor node to deal with the energy hole problem, and the literature [16] uses a smart solar energy collection system to provide stable power supply to the sensor node.

The data exchange process involves multiple sets of data fusion and data processing, which puts forward certain requirements for load balancing. The related work is analyzed below. The traditional data center network is generally a tree topology with two or three layers of switching structure. Although it can accommodate tens of thousands of servers, with the rapid expansion of the data center scale and the continuous influx of various new services, there is a tree-based data center. The structure shows the shortcomings of low utilization of network resources and poor scalability; so, the data center network topology has begun to transform to a new type of network structure. The structure of literature [17] has become a new and popular network topology with its simple structure, easy deployment, and the advantages of non-blocking transmission and half-bandwidth. According to its characteristics, the fat-tree structure generally adopts ECMP (equal cost multipath-based) or its improved technology to realize traffic forwarding. ECMP, as its name implies, uses multiple equivalent paths to forward traffic. It uses the diversity of the underlying path. When a packet arrives, it selects a path by modulo the number of paths obtained by the hash value of the

header of the packet. In this way, the flow will be forwarded along the same path. Since the distribution probability of the value after the header of each flow is equalized, each path has an equal utilization rate, and to a certain extent, the flow can be balanced to achieve the effect of load balancing. However, ECMP also has its own characteristics and limitations. One of the most critical issues is that although most of the data center network traffic is small, a small amount of continuous large flow (we will define it as a stable flow later) accounts for the vast majority. For most bandwidth [18], once two or more long-lived stable flows choose the same egress for forwarding, congestion will occur, resulting in unbalanced distribution of link traffic, and seriously affecting network transmission performance. The test results in literature [19]. It means that the transmission performance can be reduced by 50% in this case. If it is just to balance the link traffic, you can also use packet-based polling and forwarding. This method forwards each message to each outlet in turn according to the FIFO principle, almost ensuring that the number of packets forwarded by each path is the same, but even if it is, the improved balance polling algorithm [20] also cannot avoid a problem; that is, packets arrive out of order. From the perspective of TCP flow, the out-of-sequence arrival of packets is regarded as link congestion, which reduces the sending window and significantly degrades TCP performance. In contrast, ECMP is more suitable for the fat-tree structure. How to implement a dynamic flow routing and scheduling to solve the congestion problem caused by stable flow aggregation in ECMP has become the focus of this paper. For traditional networks, data center network load balancing is performed on each network node, and it is impossible to obtain all node attributes and link information in real time during network operation. It is very difficult to achieve the abovementioned dynamic load balancing effect. But fortunately, with the emergence of SDN technology, the controller can communicate with switches that support the OpenFlow protocol through the OpenFlow protocol [21] to obtain real-time statistical information that can truly and effectively reflect the network status. Based on these statistical information, we can perform calculations more accurately and efficiently from a global perspective, which also makes it possible to implement the abovementioned dynamic load balancing algorithm.

3. Nonparametric Regression Model and Local Polynomial Regression Estimation

For n observations $(x_1, Y_1), \cdots, (x_n, Y_n)$, the relationship between the response variable Y and the covariate x is defined by the following equation:

$$Y_i = r(x_i) + \varepsilon_i, \quad \mathrm{E}(\varepsilon_i) = 0, \quad i = 1, 2, \cdots, n. \quad (1)$$

Among them, r is the regression function, the variable x is also called feature, and the estimate of $r(x)$ is denoted by $\hat{r}_n(x)$.

Under the conditions of Nadaraya-Watson kernel estimation, we consider choosing an estimator $a \equiv \hat{r}_n(x)$ to

minimize the sum of squares $\sum_{i=1}^{n} (Y_i - a)^2$. We define the weight function $w_i(x) = K((x_i - x)/h)$ and choose $a \equiv \widehat{r}_n(x)$ to minimize the following weighted sum of squares:

$$\sum_{i=1}^{n} w_i(x)(Y_i - a)^2. \tag{2}$$

The solution is

$$\widehat{r}_n(x) \equiv \frac{\sum_{i=1}^{n} w_i(x) Y_i}{\sum_{i=1}^{n} w_i(x)}. \tag{3}$$

It happens to be a nuclear regression estimate.

We use a local polynomial of p order to improve the estimation. For a value u in a neighborhood of x, the polynomial is

$$P_x(u\,;a) = a_0 + a_1(u - x) + \frac{a_2}{2!}(u - x)^2 + \cdots + \frac{a_p}{p!}(u - x)^p. \tag{4}$$

In a neighborhood of x, we use the following polynomial to approximate a smooth regression function $r(u)$:

$$r(u) \approx P_x(u\,;a). \tag{5}$$

We choose $\widehat{a} = (\widehat{a}_0, \cdots, \widehat{a}_p)^T$ that minimizes the following local weighted sum of squares to estimate $a = (a_0, \cdots, a_p)^T$:

$$\sum_{i=1}^{n} w_i(x)[Y_i - P_x(X_i\,;a)]^2. \tag{6}$$

It is estimated that $\widehat{\mathbf{a}}$ depends on x. $\widehat{\mathbf{a}}(x) = (\widehat{a}_0(x), \cdots, \widehat{a}_p(x))^T$, and the estimate of $r(\cdot)$ is $\widehat{r}_n(u) = P_x(u\,; \widehat{a})$. We set the following:

$$X_x = \begin{bmatrix} 1 & x_1 - x & \cdots & \dfrac{(x_1 - x)^p}{p!} \\ 1 & x_2 - x & \cdots & \dfrac{(x_2 - x)^p}{p!} \\ \vdots & \vdots & \vdots & \\ 1 & x_n - x & \cdots & \dfrac{(x_n - x)^p}{p!} \end{bmatrix}. \tag{7}$$

W_x is the diagonal matrix of $n \times n$, where $w_{ij} = w_i(x)$. Formula (6) is written as

$$(Y - X_x a)^T W_x (Y - X_x a). \tag{8}$$

By minimizing formula (8), the weighted least squares estimate is obtained:

$$\widehat{a}(x) = (X_x^T W_x X_x)^{-1} X_x^T W_x Y. \tag{9}$$

$\widehat{r}_n(x) = \widehat{a}_0(x)$ is the inner product of the first row of $(X_x^T W_x X_x)^{-1} X_x^T W_x$ and Y.

The local polynomial regression is estimated as

$$\widehat{r}_n(x) = \sum_{i=1}^{n} l_i(x) Y_i. \tag{10}$$

Among them,

$$\mathbf{L}(x)^T = (l_1(x), \cdots, l_n(x)), \quad L(x) = e_1^T (X_x^T W_x X_x)^{-1} X_x^T W_x. \tag{11}$$

The mean and variance of this estimate are

$$\mathrm{E}(\widehat{r}_n(x)) = \sum_{i=1}^{n} l_i(x) r(x_i),$$

$$\mathrm{Var}(\widehat{r}_n(x)) = \sigma^2 \sum_{i=1}^{n} l_i(x)^2 = \sigma^2 \|L(x)\|^2. \tag{12}$$

The risk (mean square error) is

$$R(h) = \mathrm{E}\left(\frac{1}{n} \sum_{i=1}^{n} [\widehat{r}_n(x_i) - r(x_i)]^2\right). \tag{13}$$

We need to choose the smoothing parameter h. Ideally, we want to choose the $R(h)$ that minimizes v. However, $R(h)$ depends on the unknown function $r(x)$. Instead, the estimated $\widehat{R}(h)$ of $R(h)$ will be minimized. We use the missing crossvalidation score defined below to estimate risk.

The missing crossvalidation score is defined as

$$\mathrm{CV} = \widehat{R}(h) = \frac{1}{n} \sum_{i=1}^{n} \left[Y_i - \widehat{r}_{(-i)}(x_i) \right]^2. \tag{14}$$

Among them, $\widehat{r}_{(-i)}$ is the estimate obtained from the i-th data point (x_i, Y_i).

$\widehat{r}_{(-i)}$ is defined as

$$\widehat{r}_{(-i)}(x) = \sum_{j=1}^{n} Y_j \ell_{j,(-i)}(x). \tag{15}$$

Among them,

$$\ell_{j,(-i)}(x) = \begin{cases} 0, & j = i, \\ \dfrac{\ell_j(x)}{\sum_{k \neq i} \ell_k(x)}, & j \neq i. \end{cases} \tag{16}$$

In other words, the weight of x_i is 0, and the other weights are reregularized so that their sum is 1.

The intuitive meaning of crossvalidation is due to

$$
\begin{aligned}
\mathrm{E}\left(Y_i - \hat{r}_{(-i)}(x_i)\right)^2 &= \mathrm{E}\left(Y_i - r(x_i) + r(x_i) - \hat{r}_{(-i)}(x_i)\right)^2 \\
&= \sigma^2 + \mathrm{E}\left(r(x_i) - \hat{r}_{(-i)}(x_i)\right)^2 \approx \sigma^2 + \mathrm{E}(r(x_i) - \hat{r}_n(x_i))^2.
\end{aligned}
\tag{17}
$$

Predictive risk is equal to $R(r, \hat{r}_n) + \sigma^2$[21], and then

$$
\mathrm{E}\left(\hat{R}\right) \approx R + \sigma^2.
\tag{18}
$$

$\mathrm{E}\left(\hat{R}\right)$ equals predictive risk. In this way, the crossvalidation score is an almost unbiased estimate of risk.

For linear smoothers, there is a simple formula for calculating $\hat{R}(h)$.

$\hat{r}_n$ is a linear smoother; then, the missing crossvalidation score $\hat{R}(h)$ can be written as

$$
\hat{R}(h) = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{Y_i - \hat{r}_n(x_i)}{1 - L_{ii}}\right].
\tag{19}
$$

Among them, $L_{ii} = \ell_i(x_i)$ is the i-th diagonal element of the smooth matrix L.

The smoothing parameter h can be selected by minimizing the crossvalidation score $\hat{R}(h)$.

If we do not minimize the crossvalidation score, we can use another approximation method called generalized crossvalidation. Here, we replace each L_{ii} with its average $n^{-1}\sum_{i=1}^{n}L_{ii} = V/n$, and $V = \mathrm{tr}(L)$ is the effective degree of freedom. This will minimize the following equation:

$$
\mathrm{GCV}(h) = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{Y_i - \hat{r}_n(x_i)}{1 - v/n}\right]^2.
\tag{20}
$$

Generally, the bandwidth that minimizes the generalized crossvalidation score is close to the bandwidth that minimizes the crossvalidation score.

By using approximate $(1 - x)^{-1} \approx 1 + 2x$, we can obtain

$$
\mathrm{GCV}(h) \approx \frac{1}{n}\sum_{i=1}^{n}\left[Y_i - \hat{r}_n(x_i)\right]^2 + \frac{2v\hat{\sigma}^2}{n} \equiv C_p.
\tag{21}
$$

Among them, $\hat{\sigma}^2 = 1/n\sum_{i=1}^{n}[Y_i - \hat{r}_n(x_i)]^2$. Generally, for different choices of $\Theta(n, h)$, many common bandwidth selection criteria can be written as

$$
\begin{aligned}
B(h) &= \Theta(n, h) \times \frac{1}{n}\sum_{i=1}^{n}\left[Y_i - \hat{r}_n(x_i)\right]^2, \\
L\left(\hat{h}\right) &= \frac{1}{n}\sum_{i=1}^{n}\left[\hat{r}_n(x_i) - r(x_i)\right]^2.
\end{aligned}
\tag{22}
$$

Then, all $\hat{h}, \hat{h}_0$ and h_0 tend to 0 at the rate of $n^{-1/5}$. More-over, for some normal numbers $C_1, C_2, \sigma_1, \sigma_2$, there are

$$
\begin{aligned}
n^{3/10}\left(\hat{h} - \hat{h}_0\right) &\longrightarrow N\left(0, \sigma_1^2\right) & n\left[L\left(\hat{h}\right) - L\left(\hat{h}_0\right)\right] &\longrightarrow C_1\chi_1^2, \\
n^{3/10}\left(h_0 - \hat{h}_0\right) &\longrightarrow N\left(0, \sigma_2^2\right) & n\left[L(h_0) - L\left(\hat{h}_0\right)\right] &\longrightarrow C_2\chi_1^2.
\end{aligned}
\tag{23}
$$

In this way, the relative convergence rate of $\hat{h}$ is

$$
\frac{\hat{h} - \hat{h}_0}{\hat{h}_0} = Op\left(\frac{n^{3/10}}{n^{1/5}}\right) = Op\left(n^{-1/10}\right).
\tag{24}
$$

This slow convergence rate indicates that it is difficult to estimate the bandwidth. The convergence rate is essentially a question of bandwidth selection, because the following formula also holds:

$$
\frac{\hat{h} - h_0}{h_0} = Op\left(\frac{n^{3/10}}{n^{1/5}}\right) = Op\left(n^{-1/10}\right).
\tag{25}
$$

4. Load Balancing Control Algorithm for Internet of Things Link Based on the Nonparametric Regression Model

The new SDN architecture puts forward the idea of completely separating the controller and the switch. The control plane realizes flexible control of data packets on the forwarding plane through custom programming, which provides a good platform for the innovative development of new networks and applications in the future. It connects the interaction between business applications, network services, and network devices through open interfaces, adopts a centralized control method, has more flexible software programmability, and can realize different network business logic. Among them, the network control layer provides an abstract network global view for the business application layer through the northbound interface API, so that the application layer can directly control the behavior of the network. It can also provide various application services for the network, mainly including network policies, security policies, QoS policies, and cloud services. The control layer and the underlying SDN switch realize data grouping and forwarding processing through the "control-forward" communication interface OpenFlow protocol. The biggest difference between an SDN network and a traditional computer network is that it can customize and flexibly manage the functions of network switching equipment through software programming. Among them, SDN has the three basic characteristics of separation of control and forwarding, open programmable interfaces, and logically centralized control, which brings great convenience to both network managers and application developers. The new SDN architecture is shown in Figure 1, which consists of four parts: an application plane, a control plane, a data plane, and a side management plane. Different interface protocols are used for communication
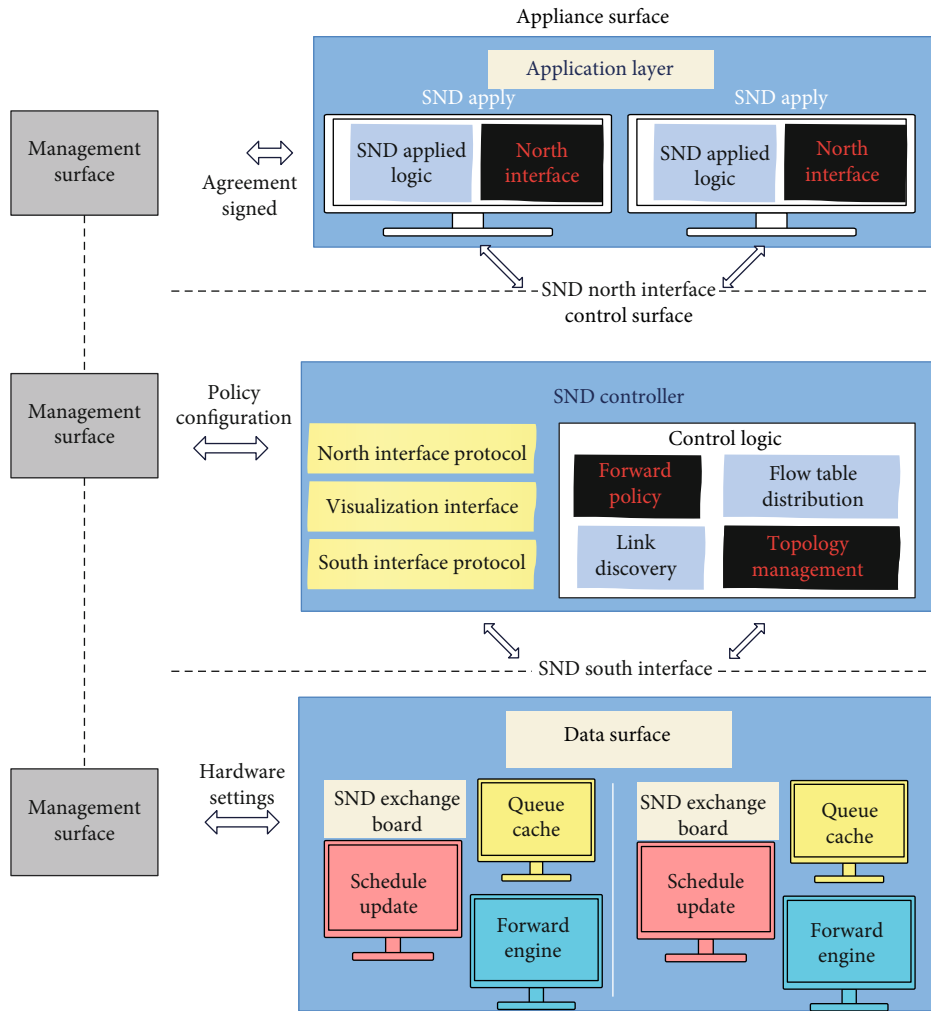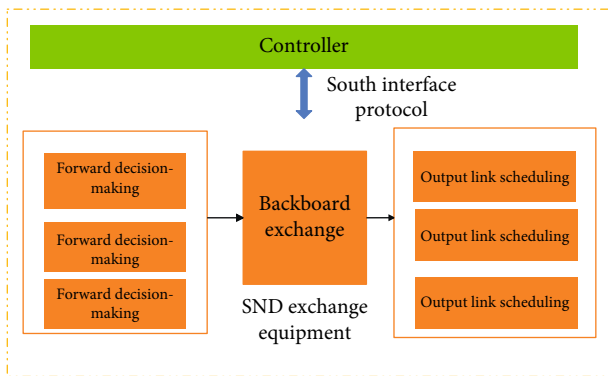
FIGURE 1: SDN architecture.

FIGURE 2: SDN switching equipment architecture.

and interconnection between the various planes, so that the entire architecture has better perception and control capabilities.

Different from the traditional network switching equipment, the SDN switching equipment architecture is shown in Figure 2. The controller is connected to the inherent switching equipment of the SDN data plane through the southbound interface protocol to implement data forwarding. This decoupling architecture design reduces the complexity of network device design and improves the scalability of hardware switching devices.

The OpenFlow protocol is the most commonly used southbound interface protocol for connecting the data plane and the control plane in the SDN architecture. It is a universal, open, and vendor-independent interface protocol. In the SDN network architecture, the OpenFlow protocol is a communication interface used for information exchange between the controller and the switch. The schematic diagram of the OpenFlow protocol architecture is shown in Figure 3. It allows the controller in the control plane to access and manage the switching equipment of the data plane by issuing flow table rules. That is to say, when a newly arrived data packet reaches the OpenFlow switch, it first searches for a match in the flow table. If there is a matching item, it will directly perform the forwarding process. If there is no matching item, the OpenFlow switch will upload the data packet to the controller through the interactive
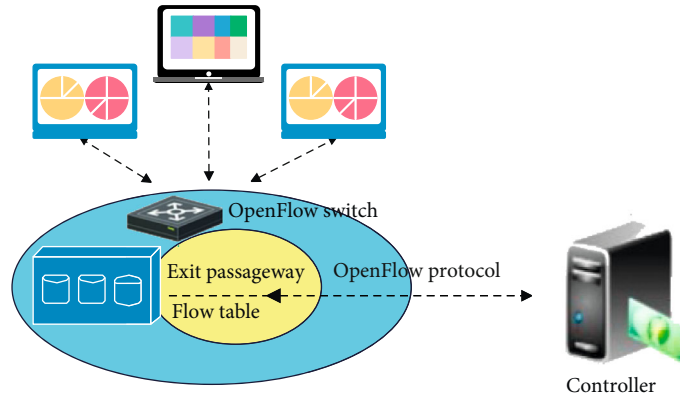
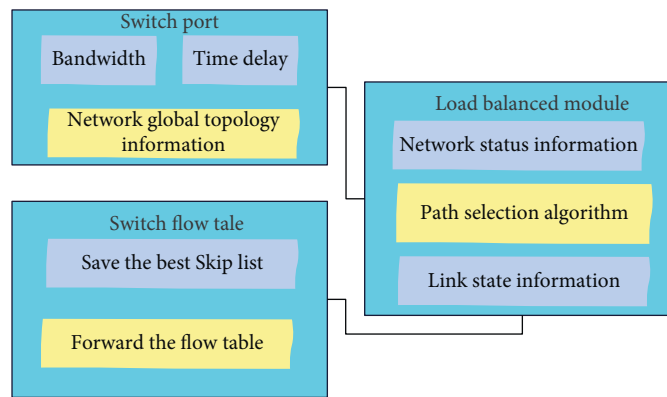Figure 3: Schematic diagram of OpenFlow protocol architecture.

Figure 4: Overall architecture diagram of load balancing of data plane data flow.

communication interface OpenFlow protocol. Then, the controller uses routing calculation, forwarding decision, and other processing and then sends the forwarding mode of the data flow in the data packet to the OpenFlow switch through the OpenFlow protocol. Moreover, the forwarding mode is stored in the flow table of the switch for subsequent search and matching operations of the flow table.

According to the characteristics of data center network traffic that is relatively concentrated and burst, this paper proposes a load balancing research strategy based on data plane data flow, which is mainly aimed at the load balancing problem of data flow in the link. Data center networks usually adopt a multilayer fat tree topology. A top-level switch at the sending end can access the state information of the neighboring receiving-end switch or the bottom-level receiving-end switch through the next-level switch and save the information in the corresponding flow table of the switch for operations such as data flow processing and forwarding. The overall architecture diagram of the data flow processing of the switch data plane is shown in Figure 4, which includes a switch port module, a load balancing module, and a switch flow table module.

Under the premise of satisfying network connectivity and coverage, unnecessary communication links are eliminated through power control or backbone network node selection. The topology control algorithm is to simplify the densely deployed topology shown in Figure 5(a) to a simple

topology through power control or the selection of backbone network nodes, respectively, as shown in Figure 5(b) or Figure 5(c), so as to achieve the purpose of improving node energy utilization rate and prolonging the network lifetime. In Figure 5, circles represent nodes, black nodes are selected backbone network nodes, and dotted lines represent communication links between nodes.

The architecture diagram of the load balancing application is shown in Figure 6. Mininet implements the data center topology, Iperf simulates virtual machine communication, and the proxy&iperfcontroller script is used to collect the traffic matrix and maintain the iperf process. The database is used to record network throughput and link load information. The web is used to display network status information and algorithm control operations. The main logic of the load balancing application is implemented on the floodlight controller.

As can be seen in Figure 6, the load balancing application is mainly composed of four components: link load collection components, traffic collection components, load balancing algorithm components, and routing configuration components. The entire load balancing application has two asynchronous execution processes, namely, the flow analysis process and the link load analysis process.

Each user in the rate sensor network can include multiple types of services. The system bandwidth is B, the number of users in the active period is K, the number of subcarriers

(a) Dense deployment

(b) After the power control algorithm is optimized

(c) After the backbone network scheduling algorithm is optimized
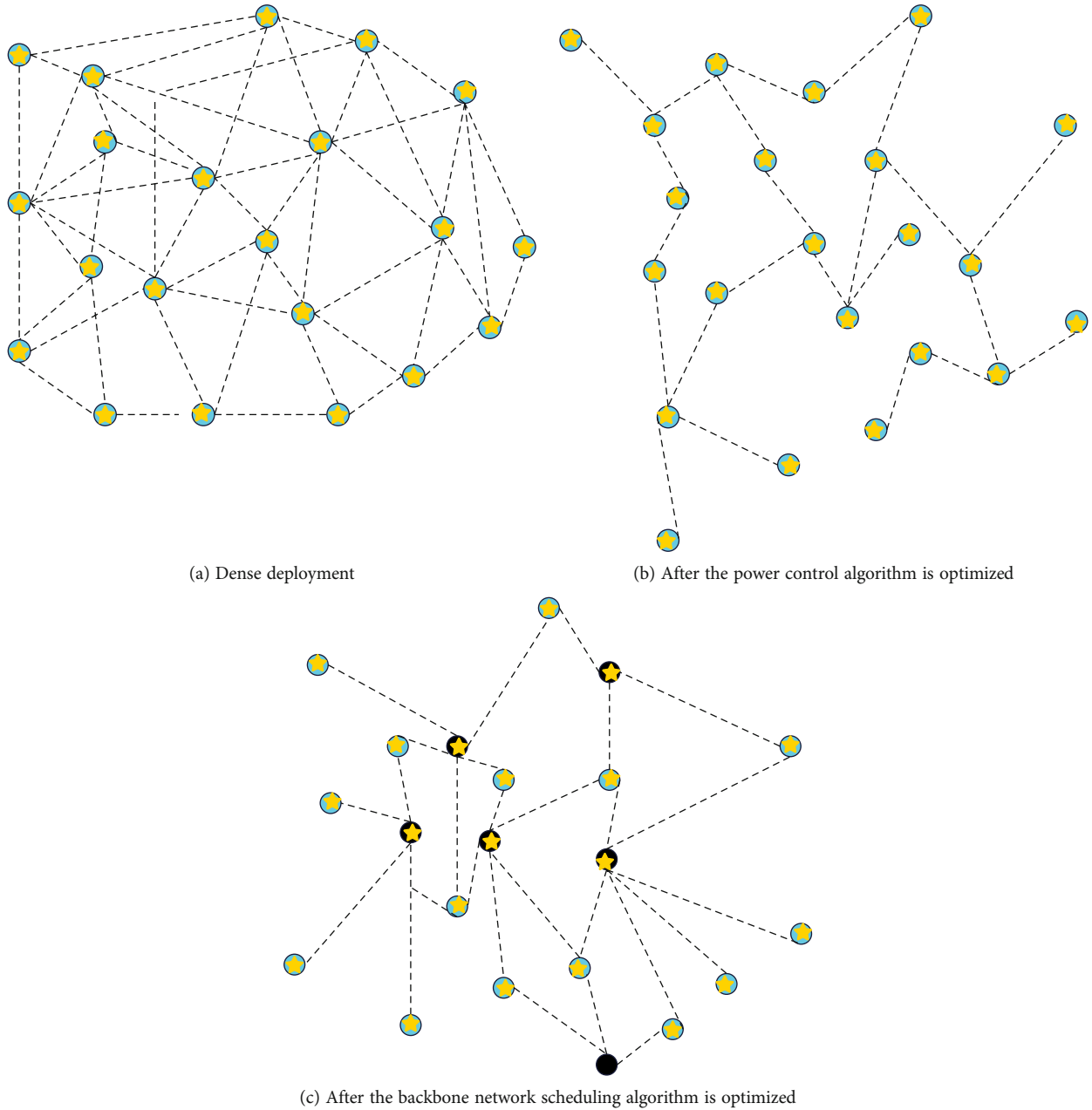
FIGURE 5: Topology control.

is N, and the bandwidth of each subcarrier is B/N. It is assumed that the system can obtain complete instantaneous state information of the channel and report the information to the base station in time through the error-free feedback channel. Assuming that N subcarriers are equally divided into n subchannels, the subchannels composed of subcarriers are almost flat fading channels. In order to improve spectrum utilization, adaptive modulation and coding (AMC) technology can be used on each subchannel. The system divides the system time in units of data frames, and each data frame time can be subdivided into S time slots. Assuming that the channel gain of each subchannel remains unchanged within each frame time, at the beginning of each frame, The base station estimates the state information CSI

of each subchannel in the link according to the information fed back by the user. According to the two-dimensional allocatable resources (subchannels, time slots) of the system, the shared channel is divided into multiple time-frequency units as the smallest unit of resource allocation; that is, a certain subchannel is occupied in frequency, and a time slot is occupied in time. $(1 < n < N)$ and $(1 < s < S)$, respectively, represent the subchannel sequence number and the sequence number within the data frame, so that there are a total of $N \times S$ block time-frequency units in a frame, and each time-frequency unit can only be one user occupied.

At the MAC layer, the base station allocates an independent data queue with limited capacity for each user, and the service order of the queue is first-come, first-served. Based
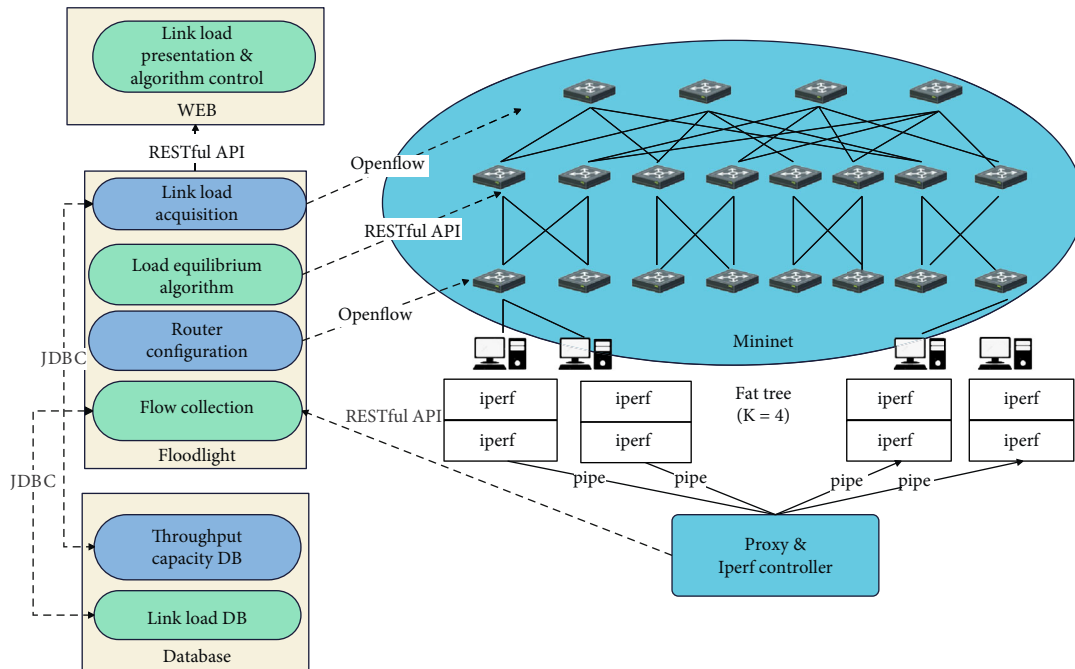
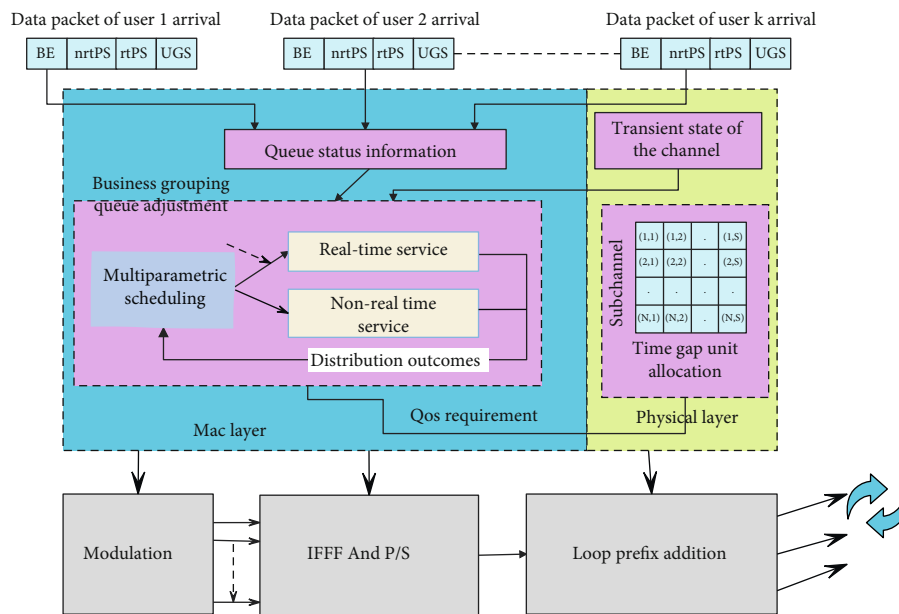FIGURE 6: Load balancing application architecture diagram.

FIGURE 7: Multiservice system scheduling model.

on the cross-layer multi-ervice scheduling model diagram shown in Figure 7, the scheduling algorithm divides users into emergency users and nonemergency users according to different parameters of the MAC layer and the physical layer. Then, it adaptively sorts the packets according to the relationship between the user's information and the maximum packet delay that can be tolerated by the user's service. The groups are divided into emergency user compensation group, emergency user noncompensation group, nonemer-gency user compensation group, and nonemergency user noncompensation group. The physical layer sorts the carrier resources and then allocates subchannels and corresponding time slots to each packet according to the sorting results provided by the scheduling algorithm, the channel state information of each user, the number of packets in the user buffer, and the QoS requirements of the service. Finally, the number of packets sent by each user is fed back to the scheduling model.
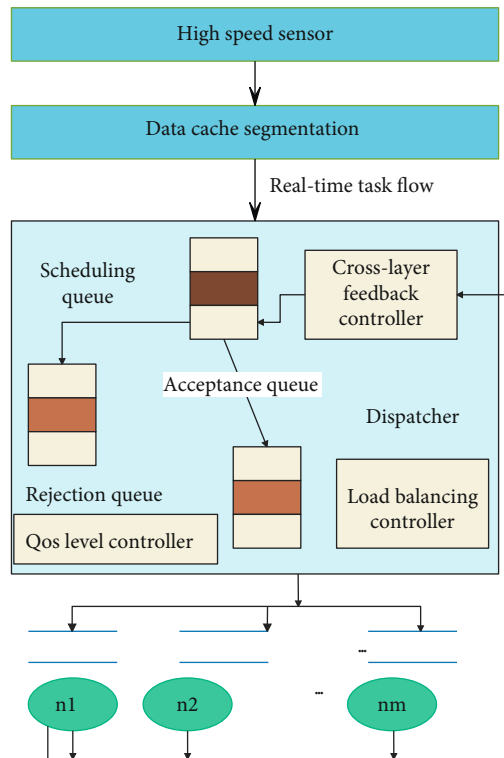
Figure 8: Multidimensional crosslayer scheduling model.

Under normal circumstances, high-speed sensor network nodes not only have the ability to sense and collect data but also have strong storage and processing capabilities. The real-time communication signal collected by the high-speed sensor node is first divided into multiple data blocks (tasks), and these tasks can be distributed to each high-speed sensor node for parallel processing. Nodes have certain storage and processing capabilities. Due to the real-time requirements of the task, the high-rate sensor network system not only needs the correctness of the task processing result but also ensures that the task can be completed within the specified time, and the final completion time of the task depends on the completion time of the last subtask. When the task is divided into multiple subtasks, all tasks of these complex applications need to be reasonably scheduled and allocated to the processing units of heterogeneous high-speed sensor nodes through a certain scheduling strategy, and task scheduling with the minimum completion time of the entire task is pursued. The quality of the scheduling algorithm directly affects the throughput rate of the task system, the success rate of scheduling, load balancing, and the QoS of the task. Scheduling algorithms for real-time tasks can be divided into two categories: static priority scheduling algorithm and dynamic priority scheduling algorithm.

The predetermined strategy of static priority scheduling is to schedule before the system starts running. Strict static scheduling cannot reschedule tasks while the system is running. The static scheduling algorithm has the advantages of simple implementation, low scheduling overhead, and good predictability when the system is overloaded. Generally, a static scheduling algorithm is used to schedule periodic tasks. The dynamic priority scheduling algorithm dynamically allocates the priority of tasks, and its purpose is to have greater flexibility in resource allocation and scheduling. The dynamic scheduling algorithm is usually used to schedule aperiodic (aperiodic) tasks. Because the emergency hospitals used by high-rate sensor networks have a high degree of uncertainty, the time of task arrival, the size of the data, and other parameters are uncertain, and the data is burst; so, dynamic scheduling algorithms should be adopted. Scheduling algorithms for real-time tasks are divided into preemptive and nonpreemptive scheduling, that is, whether the task will be interrupted by other tasks during execution. In this paper, a nonpreemptive dynamic scheduling algorithm is used to meet the requirements, which can greatly reduce the switching cost between execution tasks, and is especially suitable for emergency hospitals with large data volume and strict delay requirements.

In a heterogeneous high-speed sensor network system, the same task can be allocated to different nodes for processing, different nodes have different processing capabilities, and the same task has different processing times on different nodes. The scheduling model can be divided into a distributed scheduling model and a centralized scheduling model. In the distributed scheduling model, the arrival of local tasks is independent and can be scheduled in parallel with others. In the centralized scheduling model, all tasks are centrally processed in the scheduling processing unit. In this paper, a centralized scheduling model is adopted. Compared with distributed scheduling, the centralized scheduling model has two notable features. (1) The centralized scheduling model is simpler and easier than the distributed scheduling model. (2) By using backup scheduling, it is convenient to realize fault-tolerant scheduling. In order to achieve the scheduling goal, this paper proposes a scheduling model based on crosslayer information on the basis of the traditional centralized scheduling model, as shown in Figure 8.

The schematic diagram of the communication between the sensor network and the IPv6 host is shown in Figure 9. IP-based wireless sensor networks are easier to achieve data interaction than other heterogeneous networks that use specific protocols. The Internet gateway enables end-to-end communication between sensing nodes and Internet hosts.

Figure 10 shows the network topology structure diagram generated by the BA model in a circular monitoring area with a radius of 500 m, and 200 nodes are randomly deployed. The initial network consists of three interconnected node 3. At each subsequent time step, a new node and two links are added to the network. In Figure 10, circles represent nodes, black nodes are selected backbone network nodes, and dotted lines represent communication links between nodes.

On the basis of the above model, the effect of the IoT sensor parameter regression processing is performed on
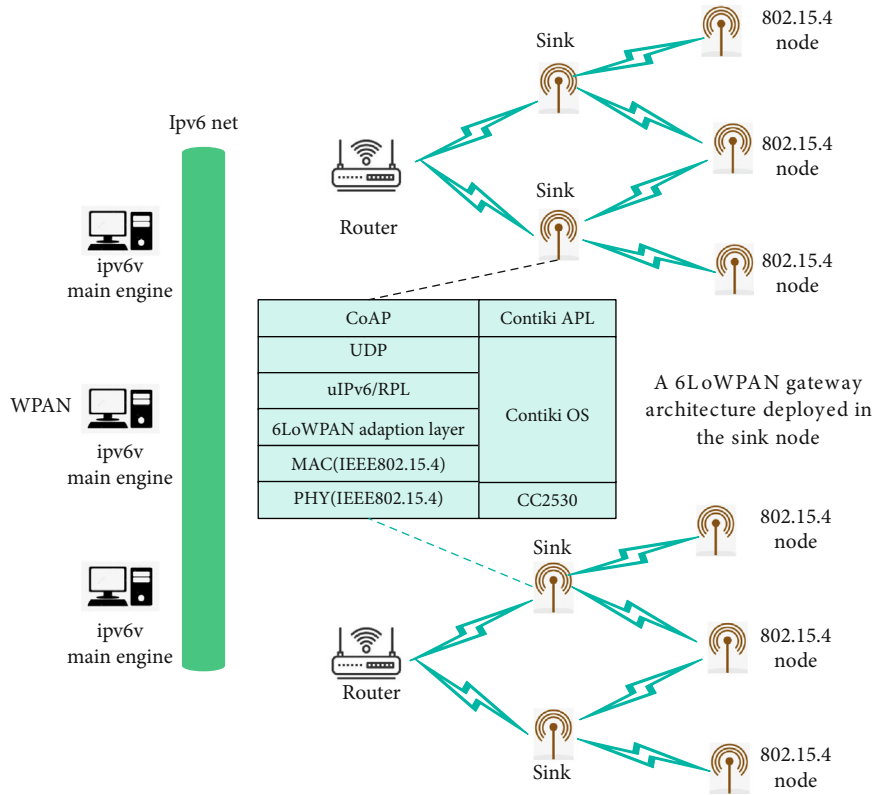
FIGURE 9: Schematic diagram of the communication between the sensing node and the host.
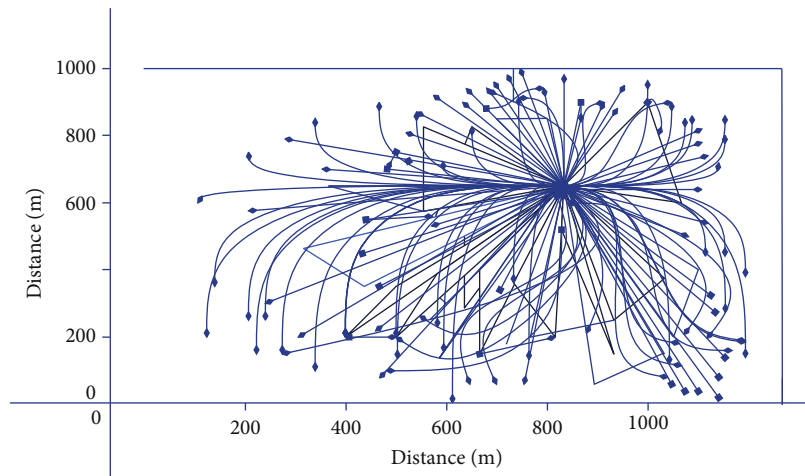

FIGURE 10: Topological structure diagram.

the model in this paper, and the results shown in Table 1 are obtained.

From the effect of IoT sensor parameter regression processing, the effect evaluation of the IoT sensor parameter regression method proposed in this paper is above 85; so, it has a good effect.

From the above research, it can be seen that the nonparametric regression model proposed in this paper can play a certain role in the sensor parameter regression of the Inter-net of Things. The load balancing control effect of this model is verified on this infrastructure, and the results shown in Table 2 below are obtained.

From the load balancing effect, the evaluation of the load balancing control algorithm based on the nonparametric regression model of the Internet of Things link proposed in this paper is above 70; so, it has a good effect.

It can be seen from the above research that the IoT link load balancing control algorithm based on the

TABLE 1: Parameter regression processing effect of IoT sensors.

No.	Regression processing	No.	Regression processing	No.	Regression processing
1	89.43	18	91.64	35	92.95
2	88.94	19	90.41	36	94.91
3	90.57	20	87.55	37	94.25
4	88.25	21	94.54	38	85.99
5	91.28	22	87.49	39	90.23
6	86.36	23	86.15	40	89.30
7	91.22	24	85.81	41	95.88
8	88.62	25	85.19	42	85.78
9	92.72	26	89.15	43	85.16
10	92.48	27	94.11	44	88.09
11	91.90	28	92.19	45	92.55
12	89.48	29	90.50	46	95.43
13	89.28	30	89.68	47	85.64
14	87.38	31	87.41	48	87.28
15	85.65	32	85.24	49	86.58
16	95.01	33	92.62	50	94.45
17	93.56	34	94.32	51	90.50

TABLE 2: Evaluation of load balancing effect.

No.	Load balancing	No.	Load balancing	No.	Load balancing
1	71.82	18	82.38	35	79.52
2	68.98	19	88.59	36	79.19
3	88.17	20	83.01	37	82.58
4	89.62	21	86.47	38	83.30
5	88.69	22	89.10	39	66.47
6	71.73	23	88.69	40	66.30
7	83.45	24	77.16	41	83.59
8	79.97	25	89.37	42	80.13
9	78.78	26	85.60	43	75.10
10	88.82	27	67.69	44	68.21
11	85.77	28	80.33	45	85.59
12	80.09	29	70.69	46	73.88
13	85.27	30	72.45	47	78.26
14	68.70	31	68.75	48	69.68
15	76.42	32	88.52	49	66.93
16	81.77	33	77.57	50	78.96
17	69.17	34	75.55	51	75.55

nonparametric regression model proposed in this paper can effectively improve the internal scheduling of the IoT network link system and promote the load balancing effect.

5. Conclusion

The high-rate sensor network handles multiple mixed services including data, voice, video, web browsing, file transfer, and multimedia applications. In different applications, data, voice, and video information have different requirements for transmission quality parameters such as bandwidth and delay. The amount of data in high-rate sensor networks is relatively large, various types of services have different requirements for the network QoS mechanism, and the resource constraints of sensor nodes are relatively more severe. Therefore, ensuring the QoS requirements of multiple services while meeting the energy constraints of high-rate sensor networks has put forward an urgent need for the optimization of the overall performance of high-rate sensor networks. How to coordinate the characteristic parameters of each sublayer in the high-rate sensor network to improve the overall network performance is a new challenge to the hierarchical protocol design of traditional wireless sensor networks. The effect evaluation of the IoT sensor parameter regression method proposed in this paper is above 85, the evaluation of the load balancing control algorithm based on the nonparametric regression model of the Internet of Things link proposed in this paper is above 70, and this paper combines the nonparametric regression model to improve the Internet of Things link load balancing algorithm, builds an intelligent model, and verifies the effectiveness of this model through research.

Data Availability

The labeled dataset used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares no competing interests.

Acknowledgments

References

[1] J. Pan, Q. Xie, H. Chiang et al., ""From the nature for the nature": an eco-friendly antifouling coating consisting of poly (lactic acid)-based polyurethane and natural antifoulant," *ACS Sustainable Chemistry & Engineering*, vol. 8, no. 3, pp. 1671–1678, 2020.

[2] B. Behroozpour, P. A. M. Sandborn, M. C. Wu, and B. E. Boser, "Lidar system architectures and circuits," *IEEE Communications Magazine*, vol. 55, no. 10, pp. 135–142, 2017.

[3] J. Barowski, M. Zimmermanns, and I. Rolfes, "Millimeter-wave characterization of dielectric materials using calibrated FMCW transceivers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 66, no. 8, pp. 3683–3689, 2018.

[4] Y. Jiang, S. Karpf, and B. Jalali, "Time-stretch LiDAR as a spectrally scanned time-of-flight ranging camera," *Nature Photonics*, vol. 14, no. 1, pp. 14–18, 2020.

[5] L. J. Xu, X. Lin, Q. He, M. Worku, and B. Ma, "Highly efficient eco-friendly X-ray scintillators based on an organic manganese halide," *Nature Communications*, vol. 11, no. 1, pp. 1–7, 2020.

[6] Q. Y. Cheng, X. L. Zhao, Y. X. Weng, Y. D. Li, and J. B. Zeng, "Fully sustainable, nanoparticle-free, fluorine-free, and robust

superhydrophobic cotton fabric fabricated via an eco-friendly method for efficient oil/water separation," *ACS Sustainable Chemistry & Engineering*, vol. 7, no. 18, pp. 15696–15705, 2019.

[7] N. Maring, P. Farrera, K. Kutluer, M. Mazzera, G. Heinze, and H. de Riedmatten, "Photonic quantum state transfer between a cold atomic gas and a crystal," *Nature*, vol. 551, no. 7681, pp. 485–488, 2017.

[8] H. Mohapatra and A. K. Rath, "Detection and avoidance of water loss through municipality taps in India by using smart taps and ICT," *IET Wireless Sensor Systems*, vol. 9, no. 6, pp. 447–457, 2019.

[9] A. Seri, G. Corrielli, D. Lago-Rivera et al., "Laser-written inte-grated platform for quantum storage of heralded single pho-tons," *Optica*, vol. 5, no. 8, pp. 934–941, 2018.

[10] Z. Meng, J. Li, C. Yin et al., "Dual-band dechirping LFMCW radar receiver with high image rejection using microwave pho-tonic I/Q mixer," *Optics Express*, vol. 25, no. 18, pp. 22055–22065, 2017.

[11] A. Al-Halafi and B. Shihada, "UHD video transmission over bidirectional underwater wireless optical communication," *IEEE Photonics Journal*, vol. 10, no. 2, pp. 1–14, 2018.

[12] H. C. Kumawat and A. B. Raj, "Extraction of Doppler signature of micro-to-macro rotations/motions using continuous wave radar-assisted measurement system," *IET Science, Measure-ment & Technology*, vol. 14, no. 7, pp. 772–785, 2020.

[13] Y. Wang, L. Ma, Y. Xu, and W. Xiang, "Computationally effi-cient energy optimization for cloud radio access networks with CSI uncertainty," *IEEE Transactions on Communications*, vol. 65, no. 12, pp. 5499–5513, 2017.

[14] F. Zhang, Q. Guo, and S. Pan, "Photonics-based real-time ultra-high-range-resolution radar with broadband signal gen-eration and processing," *Scientific Reports*, vol. 7, no. 1, pp. 1–8, 2017.

[15] T. Zhong and P. Goldner, "Emerging rare-earth doped mate-rial platforms for quantum nanophotonics," *Nano*, vol. 8, no. 11, pp. 2003–2015, 2019.

[16] Z. Sabouri, A. Akbari, H. A. Hosseini, A. Hashemzadeh, and M. Darroudi, "Eco-friendly biosynthesis of nickel oxide nano-particles mediated by okra plant extract and investigation of their photocatalytic, magnetic, cytotoxicity, and antibacterial properties," *Journal of Cluster Science*, vol. 30, no. 6, pp. 1425–1434, 2019.

[17] T. Zhong, J. M. Kindem, J. G. Bartholomew et al., "Nanopho-tonic rare-earth quantum memory with optically controlled retrieval," *Science*, vol. 357, no. 6358, pp. 1392–1395, 2017.

[18] K. Soga and L. Luo, "Distributed fiber optics sensors for civil engineering infrastructure sensing," *Journal of Structural Integrity and Maintenance*, vol. 3, no. 1, pp. 1–21, 2018.

[19] Y. J. Ma, A. Tadros, J. Du, and E. Y. Chang, "Quantitative two-dimensional ultrashort echo time magnetization transfer (2D UTE-MT) imaging of cortical bone," *Magnetic Resonance in Medicine*, vol. 79, no. 4, pp. 1941–1949, 2018.

[20] S. Sharaf and M. E. El-Naggar, "Eco-friendly technology for preparation, characterization and promotion of honey bee propolis extract loaded cellulose acetate nanofibers in medical domains," *Cellulose*, vol. 25, no. 9, pp. 5195–5204, 2018.

[21] C. Zhang, P. Xiao, F. Ni et al., "Converting pomelo peel into eco-friendly and low-consumption photothermic biomass sponge toward multifunctial solar-to-heat conversion," *ACS Sustainable Chemistry & Engineering*, vol. 8, no. 13, pp. 5328–5337, 2020.

Research Article

Intelligent Monitoring of Basketball Teaching Action Optimization considering Heterogeneous Grouping Algorithm

Bingyang Wang ᵔ

Kunsan National University, Kunsan City, Jeollabuk-do 54150, Republic of Korea

Correspondence should be addressed to Bingyang Wang; wang881120@kunsan.ac.kr

With the continuous development of social economy, sports have become an important part of people's daily life, especially ball sports are becoming more and more popular. However, in physical education, the standardization and rationality of teaching actions are extremely important. In view of these needs and limitations, a heterogeneous grouping algorithm is introduced in this paper, taking basketball teaching as the entry point, the situation of heterogeneous grouping teaching and learning are analyzed under different teaching organization through combing the business flow of basketball teaching, and the modified factor algorithm is used to conduct a process evaluation of basketball teaching. The simulation experiment results show that the heterogeneous grouping algorithm is effective, can support the intelligent monitoring of basketball teaching action optimization, and promote the improvement of basketball teaching effect, and students can get more room and level for improvement.

1. Introduction

With the continuous development of society and economy, the physical condition and learning situation of students have increasingly attracted the attention of society and families [1]. More families pay more attention to academic performance and neglect physical exercise, which often leads to lack of physical fitness [2]. Traditional physical education is often passive teaching, that is, singular and passive teaching. This method often fails to conduct effective analysis and grouping according to students' interests and specialties, which inhibits students' self-development [3, 4]. In view of these limitations, the scholars within the industry have made in-depth research and practical exploration. Typically, for instance, mixed teaching is performed for the students with different skills and different knowledge. Students learn from each other in the group to increase their enthusiasm for autonomous learning, thereby achieving the improvement of students' self-exercise ability, giving full play to their subjectivity. To a certain extent, it has changed the teaching from the single teaching in the past to the active and personalized teaching [5]. On the premise of fully considering the

physical fitness of students, it is extremely effective to effectively formulate targeted basketball teaching courses. Especially in today's continuous development of sports, it is necessary to promote more interaction among students, teachers and students, to promote students' enthusiasm for physical exercise, and to give full play to students' creativity and innovation.

Different from the existing models abroad, the current teaching model in our country mostly uses random grouping and free combination for classified competitions. Such grouping often leads to a large difference in overall skill level, which inhibits the collaborative learning exploration. [6]. It is worth noting that students are the future of the country, and their health determines the long-term development of the nation and the country's economic revitalization. Because of this, the physical fitness of students should receive more attention. The guidance of students' physical health is mainly based on guiding and heuristic strategies. It is not only an attempt of practical education, but also an important direction for participation. Compared with foreign practice, China has relatively few interventions for students, and the frequency is not enough [7, 8].

Therefore, the heterogeneous grouping method is gradually being valued and favored by the industry and scholars. Its essence is a collective teaching, that is, the students are artificially grouped according to different characteristics and technical levels. Such grouping is mainly to distinguish students in the same group. Meanwhile, the relative gap between such groupings is small, and students can help each other and grow up according to their own characteristics. This small group teaching method is grouped effectively based on the resources, allowing students to experience improvement and progress in the collective atmosphere, through mutual help and promotion of excellence, to ensure that the members of the whole group progress together.

In view of these needs and deficiencies, basketball teaching is sorted out based on the heterogeneous grouping algorithm; the intelligent monitoring of basketball teaching actions is used to count the situation of heterogeneous grouping teaching and learning, to use the modified factor algorithm to carry out evaluation of basketball teaching process, and aims to further improve the effectiveness and teaching results of basketball teaching.

2. Objects and Methods of Research

2.1. Research Object.
Set up the corresponding research objects. From the perspective of basketball teaching, two classes are randomly selected for experimental comparison. Each class has a certain basketball foundation, of which one class is an experimental class and the other is a control class.

2.2. Research Method.
The corresponding research methods are used to realize the level analysis, teaching comparison, teaching feedback, and effect evaluation analysis of the students in the experimental class and the control class. In order to eliminate external influence factors and errors as much as possible, while realizing the comparison of the two classes, other external conditions are the same.

2.2.1. Experimental Purpose. In order to effectively verify the effectiveness of the heterogeneous grouping algorithm, explore the teaching effect of the experiment.

2.2.2. Pre-Experiment Diagnostic Test. For the students in the experimental classes and the control class, an effective level assessment is carried out, and T test of the data is used to judge the validity of the heterogeneous grouping.

2.2.3. Heterogeneous Grouping Method. According to the students' physical fitness and technical level, three levels are equally divided. Meanwhile, the specific students are divided into 6 groups, each with 5 people, relatively equal in strength, and there are students with high, medium, and low basketball skills. In the meantime, reasonable teaching cooperation and teaching competitions are carried out.

2.2.4. Check the Content and Standards of the Experimental Effect Index Evaluation. In the teaching process of the heterogeneous grouping algorithm, cooperative interaction is realized, and cooperative teaching is carried out in the way

of intragroup competition, which further clarifies the concept of students' collective thinking, enhances students' cohesion, experiences cooperation and unity of ball games, and improves the enthusiasm and effect of basketball physical exercise.

During the process of experimental teaching, each group must carry out an effective implementation assessment, and its specific assessment indicators are based on teaching competition of each technology.

(1) The Teacher's Evaluation of the Student's Learning Process. During each stage, the teacher scores according to whether the corresponding learning goal is achieved, the progress of the subject, the spirit of unity, etc., and effectively scores according to the threshold range of [1, 4].

(1) First Conduct Self-Assessment and Peer Evaluation. During the course of a fixed competition, students' self-assessment and peer evaluation are carried out. As shown in Table 1, the average value of students' peer scoring is calculated.

(2) Modified Factor Method. According to the comprehensive evaluation of the relevant teachers and students on the learning process, peer evaluation/self-assessment is set as the specific modification factor of the group evaluation, so that the individual's final score can truly and effectively feedback the specific situation of the individual in the group learning process.

(2) Summative Evaluation of Student Academic Performance. In real physical education, in view of the standardized technical actions, timely feedback of effects, and improvement of technical level that students should have in the process of physical education, three specific assessment contents are selected for effective assessment and effective evaluation of learning effects in this paper, which accounts 60% in the basketball technical assessment.

(3) The Special Technical Evaluation Scores and the Proportions of Each Part of the Experimental Class and the Control Class. The total evaluation score of the experimental class is set as 100, the process evaluation accounts for 40%, and the end of study evaluation accounts for 60%;

The total score of the control class is set as 100, the evaluation score of the end of learning accounts for 60%, and the evaluation score of the teaching competition accounts for 40%.

2.2.5. On the Basis of Calculating the Relevant Scores, the Mean Value of Each Item Is Calculated, the Statistical Software Is Used for Effective Statistical Analysis, and Effective Heterogeneity Test Is Performed for Basketball Skills and Internal Motivation [9, 10]. First set up the corresponding intelligent sensor; the data transmission delay of the terminal request location and the server base station location can both be set as d, the clock deviation of the two locations is set to Δt, the server base station receives

TABLE 1: Contents of self-assessment and peer evaluation of students for basketball as in ordinary colleges and universities a special item.

Evaluation index and score	4 points	3 points	2 points	1 point
Technical levels	Very skilled	Relatively skilled	Have a certain level of technology	Poor technology
Physical quality level	Easy to adapt to the game	General adaptation to the game	Reluctantly adapt to the game	Cannot adapt to the game
Progress	Very obvious	Relatively obvious	Some progress	Less progress
Participation	Very proactive	Relatively proactive	Average attitude	Passive, lack of confidence

the corresponding data at T2 and T3, respectively, and reply the response data to the requesting terminal. After the data requesting terminal receives the corresponding data, it can be quantitatively calculated according to formula (1).

The deviation calculated by each requesting terminal can be corrected according to the corresponding atomic clock to realize the time synchronization between the data request and the data server [11–14].

$$\begin{cases} T_2 = T_1 + d + \Delta t \\ T_4 = T_3 + d - \Delta t \end{cases} \Rightarrow \begin{cases} d = \dfrac{(T_2 - T_1) + (T_4 - T_3)}{2} \\ \Delta t = \dfrac{(T_2 - T_1) - (T_4 - T_3)}{2} \end{cases}. \tag{1}$$

On this basis, the data request terminal applies for the corresponding data, which includes the number of terminal nodes, the period of the time slot, the start and end time of the time slot, and other corresponding metadata information. These parameters are all analyzed by the data request terminal for corresponding initial settings.

After receiving the corresponding data from the data server, the data requesting terminal determines the specific time of the data requesting terminal according to the information of the data, so that the corresponding configuration can be completed according to the static time slot; the specific calculation is shown in formula (2), when setting. The data requesting terminal needs data to be sent, and data transmission is performed.

$$t_i = t_0 + (i - 1) \times 2l + k \times T, k \in \{0, 1, \cdots\}. \tag{2}$$

It should be noted that when the data requesting terminal retransmits the data in the time slot of the server's continued transmission, first, the retransmission of part 1 is continued, and the time at this time can be calculated by formula (3); when the retransmission part 1 does not receive the response data from the data server, it enters the retransmission part 2. At this time, the retransmitted time slot can be calculated by formula (4). When the corresponding data is not received in the part 2 of the retransmission, it is similar to 1 and 2; it enters retransmission part 3 for continued transmission, and the retransmission time slot at this time can be calculated by formula (5). However, if the response data from the data server is still not received, the terminal

requesting node stops the continued data transmission, enters the planting state, and waits for the administrator's next cycle.

$$t_1 = t_0 + (n + r_1) \times 2l, r_1 \in \left\{0, 1, \cdots, \frac{n}{5}\right\}, \tag{3}$$

$$t_2 = t_0 + \left(\frac{6n}{5 + r_2}\right) \times 2l, r_2 \in \left\{0, 1, \cdots, \frac{n}{25}\right\}, \tag{4}$$

$$t_3 = t_0 + \left(\frac{31n}{25 + r_3}\right) \times 2l, r_3 \in \left\{0, 1, \cdots, \frac{n}{125}\right\}. \tag{5}$$

When a data terminal node encounters emergency data transmission, it first needs to be authenticated by the data server, and the node can communicate with the terminal through the data server to allocate shadow time slots. The specific calculation is shown in formula (6):

When data is being transmitted, the data requesting terminal performs data transmission in the shadow time slot of the data server.

$$t_j = t_0 + (2j - 1) \times l + t \times k, k \in \{1, 2, \cdots\}. \tag{6}$$

This article selects the corresponding student groups for research, one of which is used as the experimental class, and the other is used as the corresponding verification class, as shown in Figure 1.

The research method is shown in Figure 2. It can be seen from the results that the corresponding effective rate can be obtained in multiple measurements before and after the intervention. The average effective rate of the data used for quantitative analysis is higher than 80%, which meets the statistical requirements.

In the actual investigation and research, the corresponding teachers were selected for effective communication to ensure the progress and effect of the experiment.

On this basis, a lot of sports theory experts and sports teachers are also interviewed, and the corresponding consultation and analysis are conducted from the current physical fitness and exercise situation.
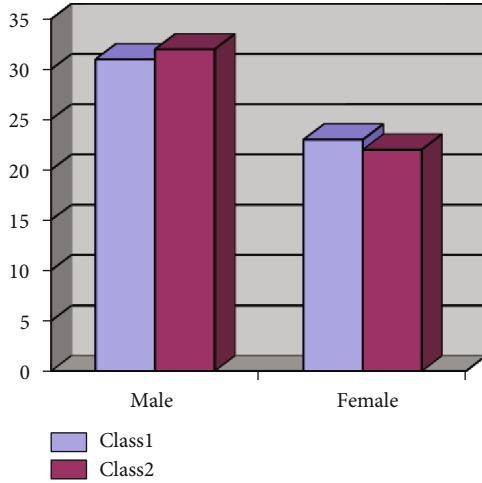
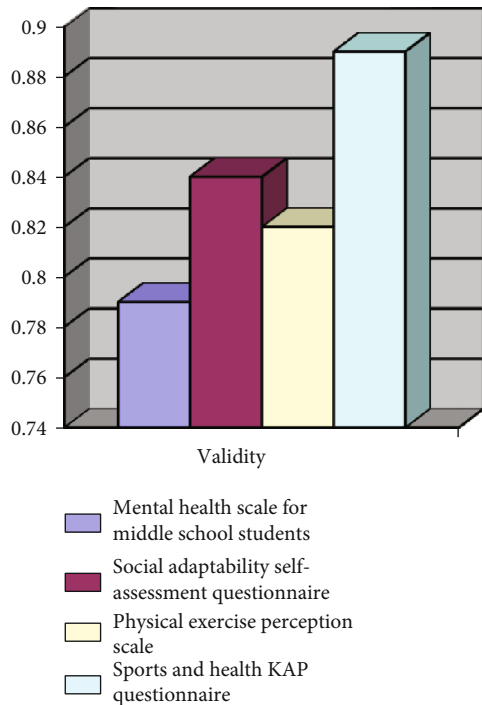FIGURE 1: Statistics of the number of subjects.

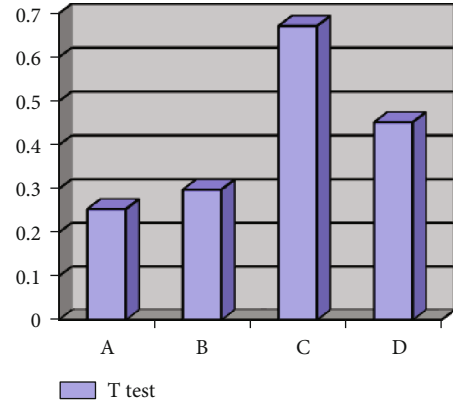FIGURE 2: List of basic information of each scale (questionnaire).

FIGURE 3: Examination of the physical health indicators of the students in the experimental class and the control class before the experiment.

average method, the weighted average method, and the maximum method [15–17]; no matter which method is used, the image feature can be changed effectively and even cause the loss of data. It can be solved by using the gray conversion formula as

$$\text{Gray} = 0.299R(x, y) + 0.587G(x, y) + 0.114B(x, y), \quad (7)$$

$$R(x, y) = G(x, y) = B(x, y) = \text{Gray}. \quad (8)$$

Secondly, the grayscaled image is effectively denoised. During the process of video processing, possible noise comes from image processing. Generally, image denoising processing shall be performed for the edge detection, image segmentation, and effective feature extraction and other processes. In this paper, a filter is used for denoising processing.

Finally, the image is binarized. Image binarization is to use the effective motion knowledge and background of the image to separate and to provide a basis for image detection. The corresponding threshold is used for effective selection; the overall and partial binarization analysis of the image can be analyzed.

3.2. Stadium Sideline Detection. Because the position of the shooting camera is fixed during basketball teaching, the sideline of the video does not change. However, during the actual teaching process, because of the influence of external factors such as different colors and billboards, the difficulty in detection of sidelines in basketball courts has been increased. Generally, the industry uses corresponding operators for extraction, such as Sobel operator and Roberts operator, but these operators usually have limitations such as complex recognition process and long operation time.

In this paper, Canny's operator is used for effective court sideline detection. Its main advantage is that it can use detection, positioning, response, and other aspects for effective analysis:

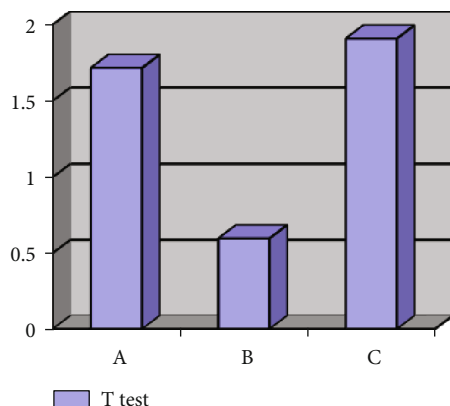(1) The process of basketball teaching videos can effectively realize as many marks as possible

3. Analysis and Realization of the Key Technology of System

In view of the research status of basketball teaching, the heterogeneous grouping algorithm proposed in this paper effectively realizes the optimal intelligent monitoring of basketball teaching actions, namely, according to image processing, basketball court sideline monitoring, sports goal monitoring, and action tracking, and the effective analysis of basketball teaching action analysis is realized.

3.1. Preprocessing of Video Images. First, the grayscale of basketball teaching action videos shall be realized. The traditional grayscale methods usually mainly include the

Figure 4: Before the experiment, technical assessment performance test of the experimental class and the control class students.

Figure 5: After the experiment, the increase rate and test of the physical health indicators of the experimental class and the control class.

Figure 6: Test of the improvement range of students' technical performance in the experimental class and the control class.

(2) The process edge detection of basketball instructional video should be as close to the edge as possible

(3) The edge of basketball instructional video can only be marked once and cannot be affected by external image noise
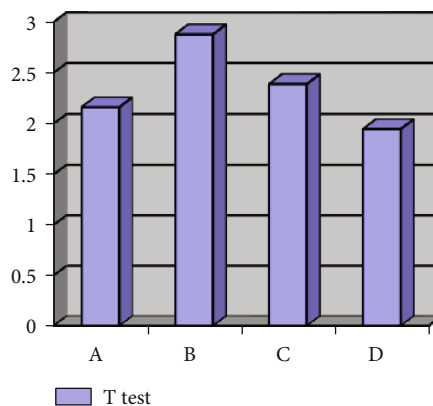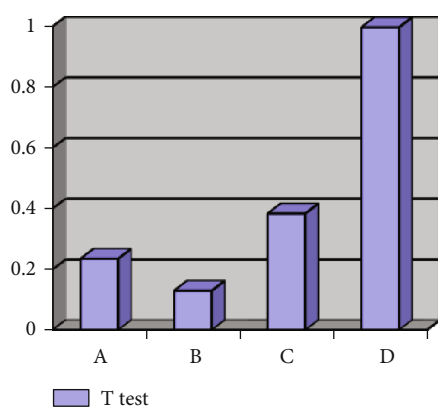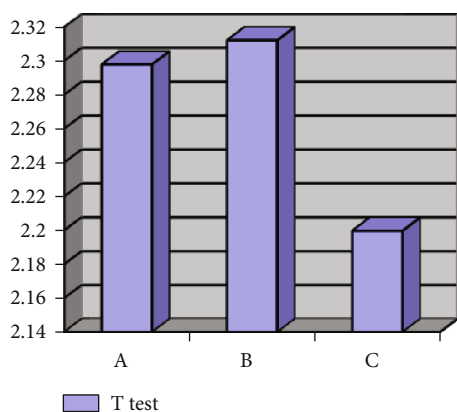

Figure 7: Statistics and inspection of some technical indicators in the 10-min teaching competition between the experimental class and the control class.

The main algorithm steps include the following:

(1) Denoise first. It cannot get better analysis results that the edge detection algorithm is applied to unprocessed images. Therefore, before processing, it is necessary to perform Gaussian smoothing of image noise to eliminate the adverse effects of response

(2) Secondly, look for the corresponding brightness gradient in the image and use 4 directions to detect the edges, which are vertical, horizontal, and diagonal directions, to identify the maximum value and edge direction of each point in the image

(3) The edge of the image is tracked, and the threshold value is used to determine the brightness gradient value, which effectively solves the problem of large brightness gradient

3.3. Moving Target Detection. During the actual impact sampling process, moving target detection is used to realize the segmentation of the moving foreground, and the feature analysis of the two-dimensional image is carried out with the help of grayscale, texture, and edge. It is a method to detect moving objects by comparing the current frame in the image sequence with the background reference model. Its performance depends on the background modeling technology used [18–21].

3.4. Moving Target Tracking Algorithm. Moving target tracking is to establish a corresponding matching problem based on the target shape, color, texture, and other related characteristics between consecutive images. The purpose is to obtain the coordinate position of the moving target in each frame of image and then associate these data to obtain the target.

4. Results and Analysis

4.1. Experimental Results and Analysis. Effectively test and compare the indicators and basketball skills of the
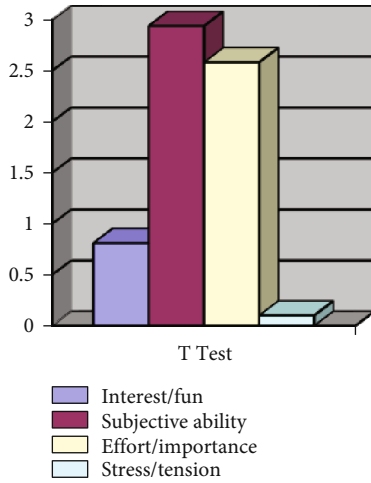
Figure 8: Test results of various dimensions of internal motivation of students in the experimental class and the control class.
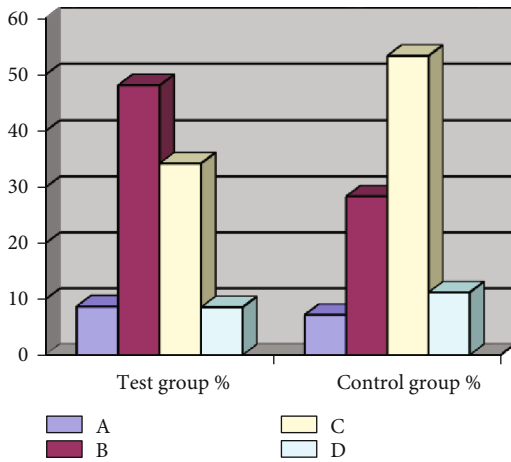

Figure 9: Comparison of the physical fitness classification of middle school students after the intervention and during the back test.
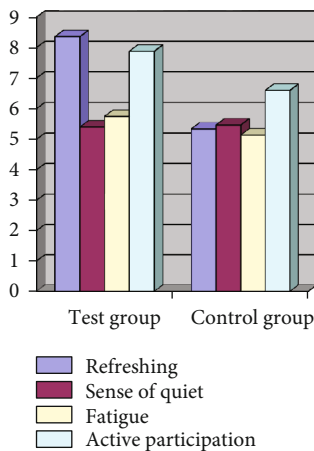

Figure 10: Comparison of changes in feelings of middle school students' physical exercise after intervention and during back test.
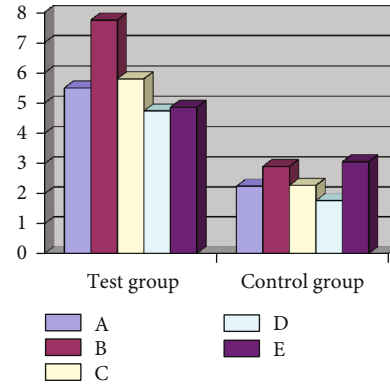

Figure 11: Comparison of awareness rate of KAP questionnaire after intervention.

experimental class and the control class. The test results are effectively tested. The specific results are shown in Figures 3 and 4. Among them, standing long jump is represented by A, vital capacity index is represented by B, grip strength and body mass index is represented by C, solid ball is represented by D, the three basketball skills are represented by a pass A with both hands on the chest, a figure eight dribble is represented by B, and a shot with a single hand on the shoulder is represented by C.

4.2. Comparative Analysis of Experimental Results

4.2.1. Analysis and Test of the Results of the Physical Health Indicators of the Two Groups of Students After the Experiment. The two forms of teaching organization can improve students' physical function level, and there is no significant difference. A is standing long jump, B is vital capacity index, C is grip strength and body mass index, and D is a solid ball (Figure 5).

4.2.2. Process Evaluation Analysis of Experimental Class Students. The function of evaluation is mainly to reflect students' progress in learning in a timely manner and to encourage students to actively reflect and summarize the learning process.

4.2.3. The Summative Evaluation Analysis of the Students' Academic Performance in the Experimental Class and the Control Class

(1) Comparison of Technical Performance between Experimental Class and Control Class. At the end of the semester, the experimental classes and the control class were both evaluated for technology, and three techniques were used for analysis. The experimental results showed that the analysis between groups has a clear sense of autonomous learning. Through mutual help between the groups, an effective improvement of the technical level has been achieved. The competition among groups makes students have common goals and become more autonomous and motivated (Figure 6).
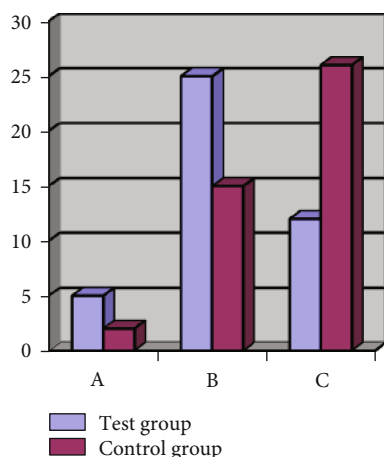
FIGURE 12: Comparison of data of action skills assessment in heterogeneous group teaching experiments.

(2) Comparison of Competition Ability. According to the 10-minute teaching and competition situation of the experimental class and the control class, violations, fouls, assist, steals, and other indicators representing the students' competition ability are counted. A is an aggression foul, B is a walking violation, C is an assist, and D is a steal (Figure 7).

(3) Comparison of Various Dimensions of Internal Motivation of the Students in the Experimental Class and the Control Class after the Experiment. By comparing the experimental classes and the control class, students have significant differences in terms of relevance, subjective initiative, and importance (Figure 8).

The results show that heterogeneous grouping teaching can give full play to the characteristics of students and realize mutual learning, mutual communication, and mutual influence. Because students are heterogeneous, students with a certain gap will continuously make progress and study harder to improve their technical level. Meanwhile, they can also recognize their own shortcomings and work harder to ensure their leading position in the class.

As shown in Figure 9, after the intervention of sports and nutrition, the excellent rate of physical fitness in the test group was significantly higher, increased by 8%, the good rate increased by about 20%, and the unqualified rate decreased. From the results, the comprehensive combination of sports and nutrition has a significant positive effect on the improvement of students' physical fitness, which can ensure that students develop good exercise habits and improve their physical fitness.

The results after the comprehensive intervention are shown in Figure 10. The students in the test group are more active and motivated, especially when compared with the nonintervention group, the results are more obvious. Meanwhile, the sense of quietness and fatigue have changed, indicating that the students are actively participating in physical exercises, but they still have a certain degree of fatigue, which is worth noting and paying attention to.

4.2.4. The Influence of Sports and Healthy Exercise Prescriptions on Students' Health Knowledge, Beliefs, and Behaviors. From the results in Figure 11, it can be seen that after the intervention, the students' scores on the theoretical knowledge of physical education increased by about 50%, and their participation in healthy behaviors even increased by about 53%. The effect was very obvious. The simulation experiment shows that the intelligent sensor is effective and can effectively monitor.

4.2.5. The Impact of Sports and Health Exercise Prescriptions on Students' Social Adaptability and Mental health. Physical education activities are group activities that accomplish teaching tasks through the interactive behavior of teaching, learning, and training between teachers and students and between students.

4.3. Analysis of Motion Skill. The analysis and assessment of movement skills is the most direct factor for students to learn basketball. Students can express the effect of the whole set of movements through specific situations [22–26]. If each student's movement skills are not up to standard, it may affect the overall performance and movement of the team. Therefore, for the assessment of movement skills, the most basic assessment analysis is realized. The specific relevant assessment content is designated by the basketball teacher and completed by the students independently. The specific assessment data is shown in Figure 12:

Independent sample $T = 2.285$. According to the above table, after 16 weeks of study, the students in the test group and the control group have significantly widened the gap in basketball skills. The students in the test group who adopted the heterogeneous grouping teaching model have made more obvious progress.

5. Conclusions

With the continuous development of social economy, ball games in colleges and universities are becoming more and more popular. Most of these ball games require teamwork to complete. However, in the actual physical education process, effective teaching is difficult to achieve due to the quality and skills of teachers and the guidance of students, and it is easy to cause students' skills to fail to improve. In view of these limitations, this article combs the business logic of basketball teaching based on the heterogeneous grouping algorithm, analyzes the teaching motives of the two groups under different teaching organizations, fully guides students to exert their subjective initiative, enhances students' learning enthusiasm, and improves students' level of basketball skills. Finally, the modified factor algorithm is used to evaluate the basketball teaching process students' evaluation and promote the improvement of basketball teaching. The simulation experiment results prove that the heterogeneous grouping algorithm is effective and can support the intelligent monitoring of basketball teaching action optimization. Students can better grasp the specific teaching form in the organizational form of cooperative teaching and promote the learning, communication, and communication between

students and can more objectively evaluate students' learning effects.

Data Availability

The 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.

Acknowledgments

References

[1] X. Wang, "An optimization method of basketball teaching and training system design based on motion capture technology," *Revista de la Facultad de Ingenieria*, vol. 32, no. 12, pp. 498–504, 2017.

[2] H. Xu and Z. Yuan, "Effective path research on cultivating university students' sportsmanship with university basketball teaching as an example," *Kuram ve Uygulamada Egitim Bilimleri*, vol. 18, no. 2, pp. 1–9, 2018.

[3] A. Dania and S. Harvey, "Teaching basketball to sampling-year athletes: a game-centered and situated learning perspective," *Journal of Physical Education and Sport*, vol. 20, no. 3, pp. 109–117, 2020.

[4] L. Xia, "Iterative learning control: an optimization paradigm [bookshelf]," *IEEE Control Systems*, vol. 37, no. 2, pp. 185-186, 2017.

[5] S. Dutta, S. H. Jacobson, and J. J. Sauppe, "Identifying NCAA tournament upsets using balance optimization subset selection," *Journal of Quantitative Analysis in Sports*, vol. 13, no. 2, pp. 10–18, 2017.

[6] Z. Yang and M. Engineering, "Distributed virtual environment basketball equipment embedded systems' research and development," *Mathematical Problems in Engineering*, vol. 4, no. 2, Article ID 5584125, 12 pages, 2021.

[7] L. Yin and R. He, "Target state recognition of basketball players based on video image detection and FPGA," *Microprocessors and Microsystems*, vol. 80, no. 48, pp. 103–114, 2020.

[8] A. D. Heishman, M. A. Curtis, E. N. Saliba, R. J. Hornett, S. K. Malin, and A. L. Weltman, "Comparing performance during morning vs. afternoon training sessions in intercollegiate basketball players," *Journal of Strength & Conditioning Research*, vol. 31, no. 6, pp. 1557–1562, 2017.

[9] N. Yamada and K. Miyatsuji, "Study on teaching method of the shot of basketball: relationship between success rate of shots and throwing ability," *Bulletin of Social Welfare Kobe Shinwa Womens University*, vol. 12, no. 5, pp. 77–82, 2015.

[10] S. Harvey, M. L. Smith, Y. Song, D. Robertson, R. Brown, and L. R. Smith, "Gender and school-level differences in students' moderate and vigorous physical activity levels when taught basketball through the tactical games model," *Journal of Teaching in Physical Education*, vol. 35, no. 4, pp. 349–357, 2016.

[11] F. Di Rienzo, P. Joassy, T. Kanthack et al., "Effects of action observation and action observation combined with motor imagery on maximal isometric strength," *Neuroscience*, vol. 418, no. 4, pp. 82–95, 2019.

[12] R. Nakamichi, "A study on teaching material of ball games in school physical education: from the viewpoint of "traveling" of basketball," *Journal of Hokkaido University of Education Education*, vol. 65, no. 4, pp. 291–301, 2015.

[13] R. Ji, "Research on basketball shooting action based on image feature extraction and machine learning," *IEEE Access*, vol. 8, no. 5, pp. 138743–138751, 2020.

[14] W. Lin, "On the optimization of sentence imitation in primary school English teaching from the perspective of strong memes," *English Language Teaching*, vol. 10, no. 5, pp. 11–18, 2017.

[15] Y. Goto, O. Hayashi, and T. Saeki, "Fundamental study on the developing teaching material of the basketball game: from the view point of change in a game aspect depended on number of players and court size," *Macromolecules*, vol. 36, no. 1, pp. 153–161, 2002.

[16] M. Mcmahon, C. J. Poster, and I. I. Session, "Poster Session II, July 14th 2010 – Abstracts: mobile multimedia learning and basketball coaching," *Procedia Engineering*, vol. 2, no. 2, pp. 3459–3459, 2010.

[17] N. Tang and P. Li, "Study on cooperative learning teaching mode in university tennis teaching," *Lecture notes in electrical engineering*, vol. 204, no. 2, pp. 545–552, 2013.

[18] Y. Wu and F. Zhang, "Research on the influence of sports and nutrition matching on improving students' physique based on intelligent sensor," *Computational Intelligence and Neuroscience*, vol. 2021, no. 2, Article ID 3556131, 7 pages, 2021.

[19] M. N. Martín-Lara and A. Ronda, "Implementation of modeling tools for teaching biorefinery (focused on bioethanol production) in biochemical engineering courses: dynamic modeling of batch, semi-batch, and continuous well-stirred bioreactors," *Energies*, vol. 13, no. 21, p. 5772, 2020.

[20] G. N. de Souza, D. F. de Deus, V. Tadaiesky, I. M. de Araújo, D. C. Monteiro, and Á. L. de Santana, "Optimizing tasks generation for children in the early stages of literacy teaching: a study using bio-inspired metaheuristics," *Soft Computing*, vol. 22, no. 20, pp. 6811–6824, 2018.

[21] R. G. Li and H. N. Wu, "Secure communication on fractional-order chaotic systems via adaptive sliding mode control with teaching–learning–feedback-based optimization," *Nonlinear Dynamics*, vol. 95, no. 2, pp. 1221–1243, 2019.

[22] H. Guo, N. Ajmeri, and M. P. Singh, "Teaching crowdsourcing: an experience report," *IEEE Internet Computing*, vol. 22, no. 6, pp. 44–52, 2018.

[23] F. N. Leite, E. S. Hoji, and H. A. Junior, "Collaborative teaching and learning strategies for communication networks," *the International Journal of Engineering Education*, vol. 34, no. 2, pp. 527–536, 2018.

[24] A. P. See, P. Khandelwal, N. Patel, and M. A. Aziz-Sultan, "Teaching NeuroImages: dynamic vertebral artery insufficiency," *Neurology*, vol. 87, no. 20, pp. e245–e246, 2016.

[25] B. Ray, C. Alan, M. Pherson et al., "Re-examining the effects of verbal instructional type on early stage motor learning," *Human Movement Science*, vol. 44, no. 4, pp. 168–181, 2015.

[26] J. P. Parsons, "Exercise-induced bronchoconstriction," *Otolaryngologic Clinics of North America*, vol. 47, no. 1, pp. 119–126, 2014.

Research Article

Economic Crisis Early Warning of Real Estate Companies Based on PSO-Optimized SVM

Xinhui Li (iD)

Department of Applied Economics, Cheongju University, Cheongju 25803, Republic of Korea

Correspondence should be addressed to Xinhui Li; lxinhui91@cju.ac.kr

For the possible economic crisis of real estate companies, by analyzing the shortcomings of support vector machine (SVM) model, after optimization with particle swarm optimization (PSO), the PSO-SVM model was established by changing economic condition parameters and using data to warn the real estate economic crisis. Then, this model is used to warn the economic capacity of four real estate companies in Beijing. The results show that this model can further predict and analyze the solvency index, operating ability index, development ability index, profitability index, cash flow index, and comprehensive financial index of real estate companies and come to the conclusion that profitability index is the most important, and higher profitability can resist the harm brought by economic crisis. Finally, it analyzes the economic crisis that real estate companies may encounter and puts forward specific suggestions.

1. Introduction

In the national economic and social development, the real estate industry, as the tertiary industry, occupies a very important position. It has become one of the pillar industries of the national economy [1, 2]. However, the real estate industry has the characteristics of large value, rigidity, high added value and can appreciate with the development of urbanization and modernization, which makes the real estate industry very attractive for investment. According to statistics, the real estate industry accounts for about 25% of the foreign investment attracted since China's reform and 40% of the projects that attract investment through the real estate industry [3]. As a high value-added industry, the real estate industry can not only provide a lot of accumulation for the country but also speed up the pace of urban construction, promoting the development of urban economy [4].

However, real estate companies still have some economic problems [5–8], such as (1) demand is determined by supply, and urbanization planning is divorced from the process of marketization. Under the condition of market economy, the site selection, scale, scope and timing of the transfer of

urban land use right should be determined by the demand of the market rather than the supply of the government. Therefore, the correct approach of the government on this issue is to only determine the supply according to the principle of demand, only through planning and coordination can the supply operation of market maintain a benign state. (2) One sided pursuit of high-level industrial structure, and industrial structure adjustment is divorced from market-oriented orientation. In the market economy, the change of industrial structure is the result of the law of average profit. The market orientation has both blind and reasonable aspects. The rational thinking of the government is to respect its rationality and adjust its blindness. The problem now is, or regardless of the business direction of investors, issue orders, and set targets, and lease to whoever pays a high price. Or do not respect the market orientation of investors, forcibly change the investment direction and unilaterally pursue high-level industries with high technology and high consumption. (3) The immature real estate market leads to blind high land price and low land price. Land itself is not a commodity. In a market economy, its price is not determined by social necessary labor, but the capitalization of land rent, that is, it is determined by the rate

of return on investment in land and is affected by the relationship between supply and demand. However, the "buyer's market" of China's real estate market has not really taken shape. To determine the lease price of land largely depends on "patting the head." There are two prominent problems, one is the blind high price behavior, the other is the administrative and policy needs, competing to lower the land price, which brings the blindness of land investment. Therefore, it is urgent to establish a model that can predict the economic crisis of real estate companies.

The support vector machine (SVM) was first introduced in the kernel function; SVM has the advantages of universality, robustness, effectiveness, and simple calculation point, which can better solve the classification problem in the case of small samples, nonlinearity, and high dimension, and has strong generalization ability [9]. However, whether the parameters of support vector machine are appropriate or not will have a great impact on the training and convergence of data [10–14]. Particle swarm optimization (PSO) has not only global optimization ability but also efficient convergence and strong local optimization ability [15–18]. Therefore, the PSO algorithm is used to optimize the SVM parameters and make SVM conduct the problem and then applied to the prediction of real estate economic crisis, which has attracted the interest of researchers at home and abroad.

Based on the PSO-SVM model, Quandang et al. [19] analyze the measures taken by different real estate companies to deal with economic crises. Through the study of these preventive measures, a mechanism model for predicting the economic crises of real estate companies is established. Zhaojun et al. [20] adopt the single-variable SVM early warning model, selects the financial indicators of different real estate companies as the method of economic crisis early warning, and indirectly reflects the financial crisis of some real estate companies through the financial indicators selected by the research. However, due to the limitations of the early warning model, the research still has some shortcomings, which can be regarded as descriptive research statistics. Jiang [21] takes into account the shortcomings of only using the financial status of real estate companies as early warning indicators, and the indicators of different real estate companies' operating capabilities, solvency, and development capabilities are different. The SVM model optimized by PSO is adopted to provide different real estate companies. By giving certain weights to different financial indicators of different real estate companies, the discrimination value of weighted average value is obtained. This model can be a good early warning of the economic crisis of real estate companies. Wang et al. [22] optimize the SVM model and use the index ratio as the evaluation method of the model. Ten bankrupt companies and 10 companies with good financial performance are selected for research, and the economic crisis early warning model of real estate companies is effectively verified. Zhou et al. [23] use the PSO-SVM model to analyze the bankrupt companies in the United States. The results of the study show that the country's macroeconomic policies have a reference value to a certain extent on whether the company has a financial crisis. Its impact is greater, which is consistent with the economic crisis encountered by actual real estate companies, and has a good early warning and judgment.

By combining PSO with SVM, this paper analyzes the solvency index, operating ability index, development ability index, profitability index, cash flow index, and comprehensive financial index among different real estate companies and then uses PSO to find the best penalty factor C and kernel function parameters σ and establish the best SVM model, so as to carry out relevant early warning and analysis of the economic crisis of real estate companies.

2. Establishment of SVM Model

The main idea of SVM is to make the samples to the feature space through a nonlinear mapping function K and carry out linear regression operation in the high-dimensional space.

SVM model functions are as follows:

$$f(x, \omega) = \sum_{j=i}^{N} \omega_j K(x, x_i) + b, \tag{1}$$

where $\omega_j = (\omega_0, \omega_1, \cdots, \omega_N)^T$ is the weight vector, $x_i = (x_0, x_1, \cdots, x_N)^T$ is the training sample set, b is the offset, and $K(x, x_i)$ is the kernel function satisfying Mercer condition.

The regression problem of equation (1) can be transformed into the problem:

$$\min_{\omega, b, \xi, \xi^*} \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^{l} \left(\xi_i + \xi_i^*\right). \tag{2}$$

The constraint conditions are equations (3)–(5):

$$f(x_{i,}\omega) - y_i \leq +\varepsilon, i = 1, 2, \cdots, 1, \tag{3}$$

$$y_i - f(x_{i,}\omega) \leq \xi_i + \varepsilon, i = 1, 2 \cdots, 1, \tag{4}$$

$$\xi_i \xi_i^{*\cdot} \geq 0, i = 1, 2, \cdots, l, \tag{5}$$

where C is the penalty factor, ε is the coefficient of insensitive loss function, and ξ_i, ξ_i^* is the relaxation factor. The above problems could be conducted through using the Lagrange method. Finally, the financial index ratio model can be obtained:

$$a = \sum_{i+1}^{l} a_i K(x, x_i) + b, \tag{6}$$

where a is the ratio of financial indicator, α_i is the support vector coefficient of the training sample, x_i is the support vector, and x is the experiment thing. In this paper, the scheme kernel function K is adopted, and the radial basis function (RBF) function is selected:

$$K(x, x_i) = \exp\left(\frac{\|x - x_i\|^2}{2\sigma^2}\right). \tag{7}$$

During regression, the kernel function parameters need to be confirmed σ.

The ideal function model of financial index ratio should be

$$V_a = V_0 + S_F\, a, \tag{8}$$

where S_F represents the scale factor of financial index ratio and V_0 is the zero status of financial indicator ratio. S_F and V_0 are constants independent of temperature, but due to the influence of temperature, S_F and V_0 have certain errors with the nominal value, and the difference between the output financial data value and the financial index ratio and the ideal linear relationship is also very large, so it is necessary to compensate the financial index ratio, which includes solvency index, operating ability index, development ability index, profitability index, cash flow index, and comprehensive financial index. During compensation, the financial index ratio after compensation can be calculated by outputting two signals V_T and V_a of the financial index ratio to form the test sample data and substituting them into equation (7).

3. SVM Parameter Optimization Method Based on PSO

In the SVM regression model, penalty factor C and kernel function parameters σ have a great impact on the performance of the system. The parameter selection process of SVM is equivalent to an optimization process. It may be the optimal solution for every point. Therefore, the optimal solution can be obtained by finding the point with the smallest generalization error. When the PSO is used to optimize the parameter selection of SVM, it is effectively improve the parameter optimization ability [24, 25].

3.1. PSO Principle and Improvement. The PSO algorithm is an intelligent optimization algorithm derived from the flight foraging behavior of birds. All points in the search space of the algorithm are the complete set of problem solutions. Each particle represents a potential optimal solution. Each particle has its own position and speed, as well as a fitness value.

In d-dimensional space, the position of the i particle is $x_i = (x_{i1}, x_{i2}, \cdots, x_{id})$ and the velocity $v_i = (v_{i1}, v_{i2}, \cdots, v_{id})$. In each iteration, the particle updates its position by tracking two optimal solutions, one is the individual extremum Pbest, $p_i = (p_{i1}, p_{i2}, \cdots, p_{id})$, and the other is the global optimal solution Gbest, $p_g = (p_{g1}, p_{g2}, \cdots, p_{gd})$. The particle updates its velocity and position by

$$v_i^{k+1} = \omega v_i^k + c_1 r_1 \left(p_i^{k-}\right) + c_2 r_2 \left(p_g^{k-}x_i^k\right), \tag{9}$$

$$x_i^{k+1} = x_i^k + \beta v_i^{k+1}, \tag{10}$$

where ω is inertia weight, c_1 and c_2 are learning factors, the values are both 2, and r_1 and r_2 are random numbers between $[0, 1]$. β is the constraint factor, the values is usually

1. ω is the most important parameter. Choosing an appropriate weight is very important to the performance of the algorithm. The common method is to use the weight linear decreasing method ω from the maximum to the minimum value, and this method enhances the search ability of the algorithm, but the actual search iteration process is a complex nonlinear change process. The linear decline method often can not well reflect this feature, which is easy to lead to insufficient improvement in the accuracy and speed of the algorithm.

Therefore, the calculation method is as follows:
When $f \le f_a$,

$$\omega = \omega_{\min} - \frac{(\omega_{\max} - \omega_{\min})(f - f_{\min})}{f_a - f_{\min}}. \tag{11}$$

When $f > f_a$,

$$\omega = \omega_{\max}, \tag{12}$$

where $\omega_{\min}$ and $\omega_{\max}$ is the minimum value and maximum value, respectively. f, $f_{\min}$, and f_a are the current particle fitness value, minimum fitness value, and average fitness value, respectively. The fitness function in this paper is the mean square error of particles.

3.2. PSO Optimization SVM. According to the previous description, PSO optimizes SVM as follows:

(1) Initialize PSO algorithm parameters. Including SVM penalty factor C and kernel function parameters σ, population size, initial position and velocity of particles, individual extremum p_i, and global extremum p_g

(2) Update the inertia weight ω according to equations (11) and (12).

(3) Update the velocity and position of the particles by the equations (9) and (10).

(4) Taking the mean square error of regression as the function, the individual optimal Pbest and global optimal Gbest are recorded

(5) If the conditions are met, the result will be output. Otherwise, skip to step 2

3.3. PSO-Optimized SVM Classification Model

3.3.1. Initialization

(a) When the particle swarm optimization algorithm processes large sample data, the calculation iteration speed is very slow, so first take any size sample from the experimental sample as the training sample in the whole iterative optimization process and obtain the best regression parameters after many experiments

(b) Initialize particle m to 20 × 2, where 20 is the population size, and then find the optimal penalty factor C and kernel function parameters σ

(c) Initializes the speed of particles

(d) Using the k-fold cross-validation method, the normalized subsamples and the initialized particle swarm are substituted into the svmtrain function to calculate the initial fitness value

(e) Initialize the global extremum and individual extremum, where the global extremum is the optimal value of the initial fitness value of all particles, and the individual extremum is the initial fitness value of each particle

3.3.2. Iterative Optimization

(a) By updating the velocity and position of the particles, the parameters are obtained. If the boundary value is exceeded, the boundary value is taken to generate new particles

(b) The fitness value of the new particle is calculated and compared with that of the previous generation. If it is better than the fitness value of the individual optimal particle of the previous generation, it will be regarded as the optimal solution of the current particle

(c) If the particle is better than the fitness of the current population optimal solution, it will be regarded as a new population optimal solution

3.3.3. End Iteration. Check whether the iteration reaches the maximum number of iterations. In the iterative process, if it is found that after a certain number of iterations, the optimal moderate value no longer changes; it can be determined that the maximum number of iterations has been reached at this time. If it has, end the optimization. Return the global optimal C sum at the same time σ.

4) Will get C and σ substitute SVM to establish a classification model to test and analyze the economic crisis early warning model of real estate

4. Results and Discussions

This article analyzes the financial status of four real estate companies in Beijing (China Overseas Land, replaced by A, China Resources Land, replaced by B, Beijing Urban Construction, replaced by C, and China Communications Land, replaced by D) to analyze the solvency index, operating ability index, development ability index, profitability index, cash flow index, and comprehensive financial index are used as the economic capabilities between companies, and then, the PSO-optimized SVM model is used for early warning analysis, in order to provide the experiment support for the economic crisis of real estate companies.

4.1. Solvency Index. Solvency index refers to an important management index of enterprise financial management,

which refers to the ability of an enterprise to repay due debts (including principal and interest). From the perspective of creditors of real estate enterprises, what they are most concerned about is the solvency of real estate enterprises, because the basic purpose of creditors' lending funds is to obtain interest income. If the real estate enterprise fails to pay the principal and interest on time, it is bound to suffer certain losses for creditors. Therefore, if real estate enterprises want to have a better financing environment, they must match a strong solvency. Only in this way, real estate enterprises can collect more funds for investment, so as to bring more profits to investors.

Short-term solvency refers to the degree to which an enterprise's current assets guarantee the timely and full repayment of current liabilities. It is an important indicator to measure an enterprise's current financial ability (especially the liquidity of current assets). The measurement indicators of an enterprise's short-term solvency mainly include current ratio, quick ratio, and cash current liability ratio. The liquidity of current assets determines the strength of short-term solvency, that is, the speed at which assets are converted into cash. The stronger the liquidity of current assets of real estate enterprises, the corresponding short-term solvency is also strong. However, long-term solvency refers to the ability of an enterprise to repay long-term liabilities. The measurement indicators of long-term solvency of enterprises mainly include asset liability ratio, property right ratio, contingent liability ratio, interest earned multiple, and interest bearing liability ratio.

Figure 1 shows the solvency indicators of different companies. It can be seen that B company has the best short-term solvency, followed by A company, and D company has the worst short-term solvency. The main reason may be that the funds of B company are maximized, which will indirectly increase the profitability of the enterprise, and the short-term current liabilities of the real estate company can be well guaranteed, so as to meet the short-term debt capacity of the enterprise. On the contrary, it will occupy too much capital of real estate enterprises, which will increase the opportunity cost of capital. From the perspective of long-term solvency, A company is the best and that of D company is still the worst. The possible reason is that the interest protection multiple between companies is different. If the profit before interest and tax is greater than the interest expense, the real estate enterprise may have the possibility of debt. If the interest coverage ratio is at a low level, it means that real estate enterprises will encounter losses, debt insecurity, and other problems. The higher the interest coverage ratio, the higher the degree of protection provided by the profits of real estate companies for paying debt interest, and the stronger their long-term solvency.

4.2. Operating Capacity Index. Operating capacity is measured by asset management ratio. Common indicators that can reflect the operation and management efficiency of real estate companies include total asset turnover ratio, accounts receivable turnover ratio, inventory turnover ratio, and fixed asset turnover ratio.
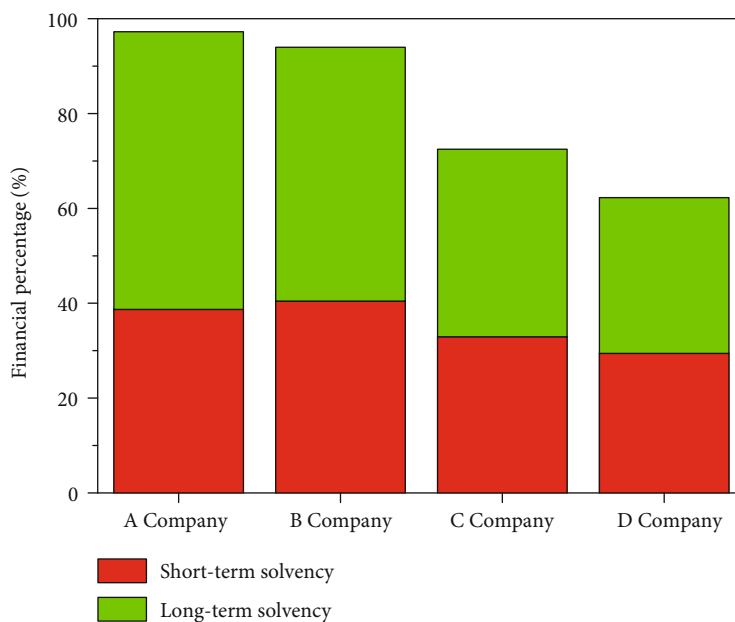
FIGURE 1: Solvency index of different companies.

By calculating the total asset turnover rate, we can compare the total asset operation capacity of real estate enterprises. The indicator reflecting the operating capacity of total assets is the total asset turnover rate. The total asset turnover can be calculated by subtracting the sales cost from the current sales revenue of the real estate enterprise and then comparing with the average amount of assets in the previous period. The inventory turnover rate is obtained by calculating the ratio of main business cost to the average of opening inventory amount and closing inventory amount. This indicator can be used to assess the performance of relevant business processes of real estate enterprises. The turnover rate of fixed assets is usually used to reflect the operating capacity of fixed assets. The so-called fixed asset turnover rate is the average value of the net operating income of the food industry compared with the fixed assets at the beginning of the previous period and the fixed assets at the beginning of the current period. This indicator can describe the current turnover of fixed assets of real estate enterprises.

Figure 2 shows the operating capacity index between different companies. It can be seen that the turnover rates of D company are high, followed by B company and A company, and the turnover rate of C company is the worst. The above results show that D company has high operation and management efficiency and good asset turnover capacity, which can ensure the solvency of the enterprise to a certain extent, avoid some bad debt reserves, and then improve the operating profit of the enterprise. Therefore, if real estate enterprises want to maintain a high turnover rate of fixed assets, they must make reasonable investment and pay attention to the portfolio structure. The high turnover rate of fixed assets means that the operation capacity of real estate enterprises is strong.

4.3. Development Capacity Index.

The development process of real estate enterprises can show the production and operation of some enterprises. In the process of enterprise development, it will constantly tap its own potential, which is the basic situation of development ability. At present, a common index of development capacity includes sales revenue growth ratio, net profit growth ratio, capital accumulation ratio, and total asset growth ratio.

The growth rate of sales revenue can be used to explain the change of sales revenue in a certain period. Through this index, we can judge the market operation and change trend of business performance. The net profit growth rate is the ratio of the current net profit minus the previous net profit to the previous net profit. Generally, the capital accumulation rate can be calculated by comparing the change amount of owner's equity in the current period with the amount of owner's equity at the beginning of the previous period. This index is generally used to evaluate the current capital accumulation of enterprises. The growth rate of total assets is mainly the result of calculating the change of assets compared with the total assets at the beginning of the previous period. It can generally be used to measure the change of total assets of real estate enterprises in the current period.

Figure 3 shows the development capacity trend of different real estate companies. The ordinate indicates the level of development ability. The larger the value is, the stronger the development ability of the company is. The abscissa indicates the value changes of different growth rates. The larger the value range, the better the future financial improvement ability of the company. It can be seen that A company and B company have the best financial improvement ability in the future and lack the current development ability, while C company and D company have the strongest current development ability and limited future financial
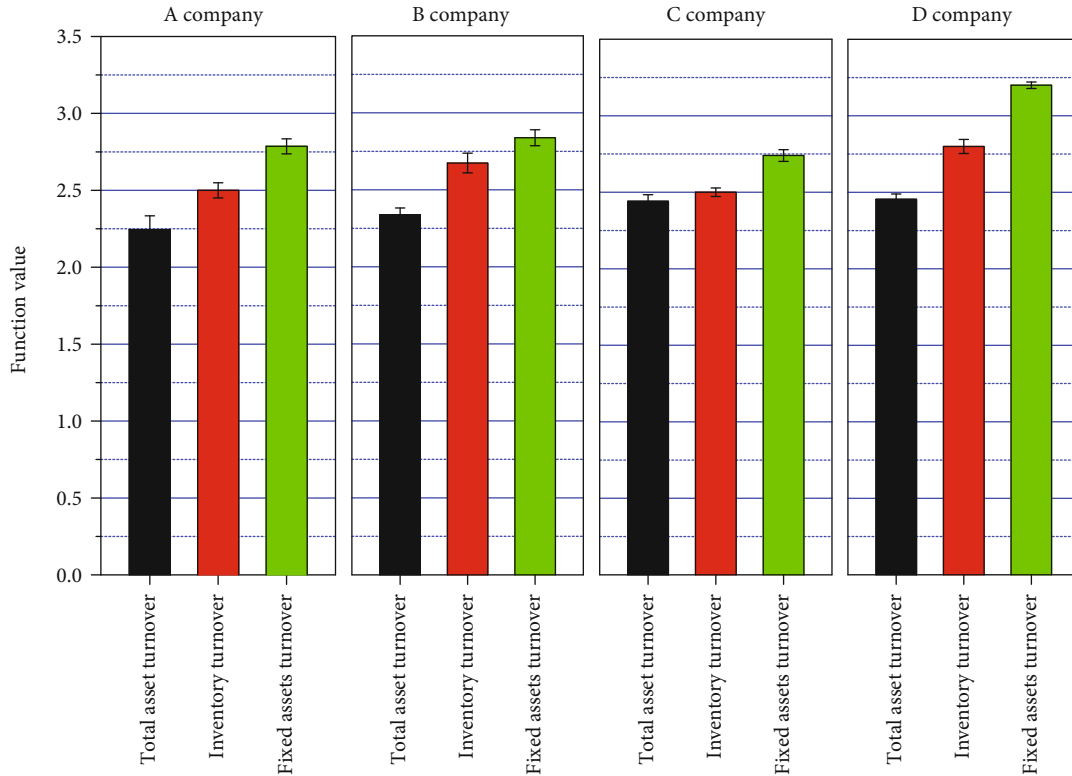
FIGURE 2: Operating capacity indicators of different companies.
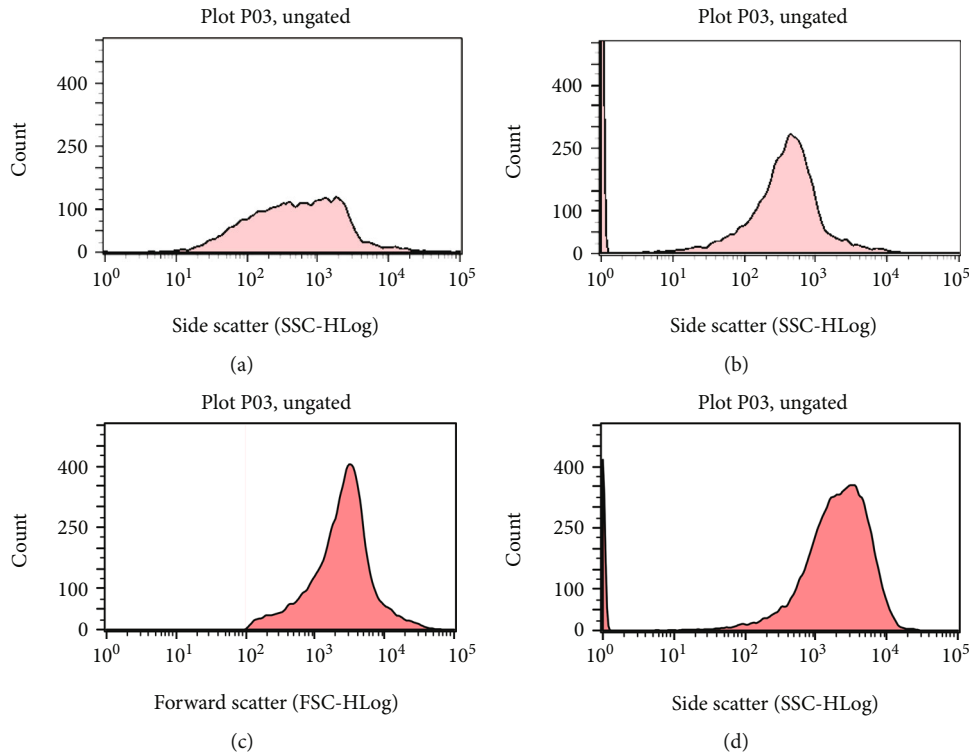

(a)

(b)

(c)

(d)

FIGURE 3: Development capacity of different companies: (a) A company, (b) B company, (c) C company, and (d) D company.

improvement ability. Considering the comprehensive factors, the development capacity index of B company is the best.

4.4. Profitability Index. Profitability is a very valuable indicator for both internal managers and external information users. Generally speaking, profitability means the ability of
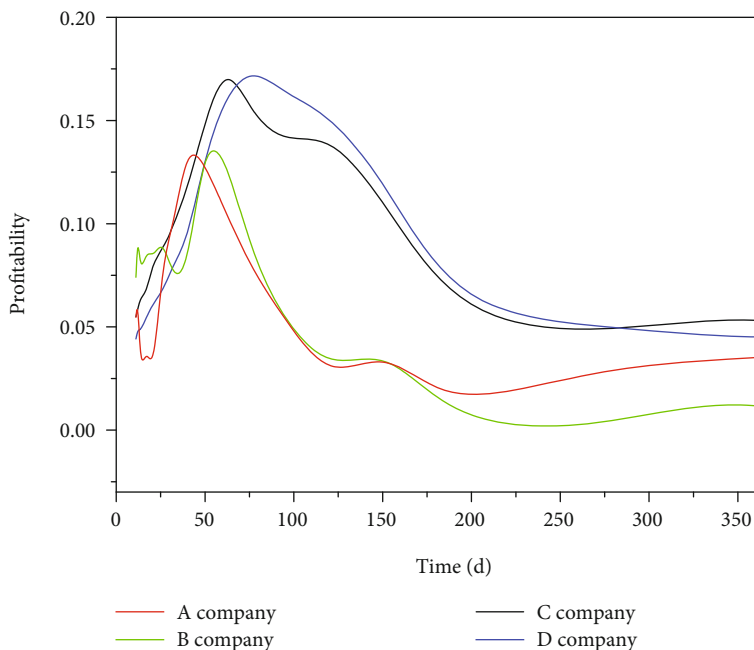
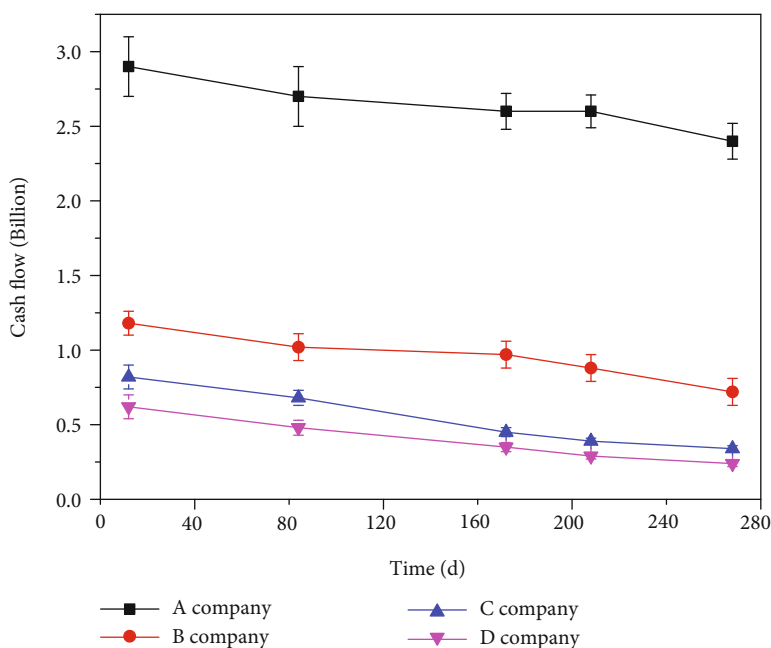FIGURE 4: Profitability of different companies.

FIGURE 5: Cash flow of different companies.

real estate enterprises to earn profits. Usually, the profit index is connected with the current operating income, operating profit and asset scale of the enterprise.

Figure 4 shows the profitability of different companies. It can be seen that the profitability of A company and B company is poor, and the profitability of A company is lower than that of B company in the early stage and slightly higher than that of B company in the later stage. The similar profitability of C company and D company indicates that the

two companies have good profitability, which is also consistent with the actual profits of the company.

4.5. Cash Flow Index. Cash flow analysis generally includes cash flow structure analysis, liquidity analysis, cash acquisition ability analysis, financial elasticity analysis, and income quality analysis. Real estate enterprises pay more attention to cash flow; if the cash flow of real estate enterprises breaks, the impact on real estate enterprises will be a heavy blow.
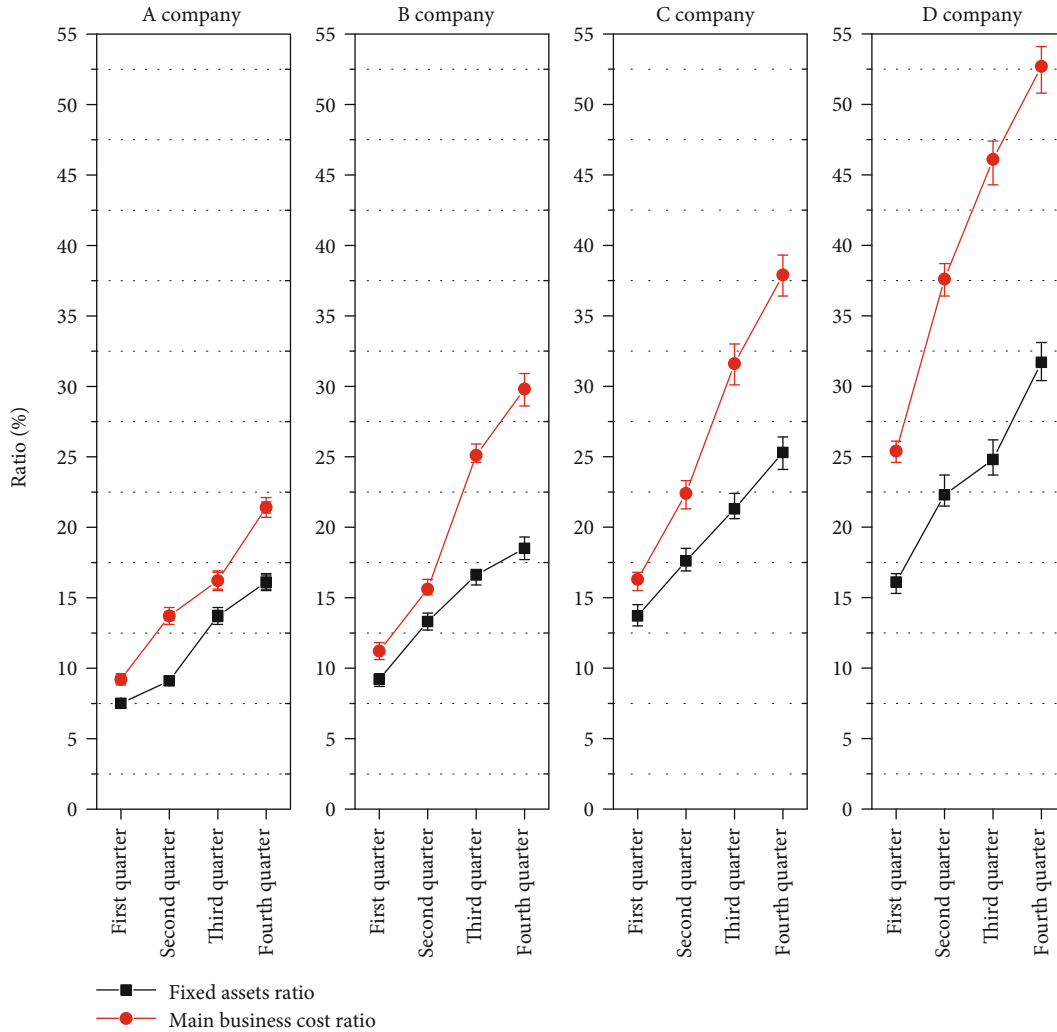
FIGURE 6: Comprehensive financial capacity of different companies.

Enterprises will be unable to pay workers' wages, leading to workers' strike, interruption of cooperation with other companies, and in serious cases, enterprise bankruptcy. Considering that the liquidity of assets is inversely proportional to profitability, it is impossible for real estate enterprises to retain too much cash. They should mainly be careful with the ability of enterprises to realize profits. Through the cash flow index, the matching relationship between the cash flow generated by the enterprise in the current period and the total liabilities can be calculated. It is generally believed that the higher the index, the better, indicating that the financial situation of the enterprise in the current period is better.

Figure 5 shows the cash flow of different real estate companies. It can be seen that A company has good financial ability, with the highest cash flow and the lowest cash flow of D company. However, combined with the profit index, although A company has more cash, its profit is the lowest, which shows that A company should strengthen its financial management ability and make rational planning and use of cash in order to deal with the possible financial crisis.

4.6. Comprehensive Financial Index. In addition to the above five indicators, the comprehensive financial indicators mainly refer to the ratio of fixed assets and the cost ratio of main business. The larger the ratio of fixed assets, it indicates that there is no idle waste of fixed assets in real estate enterprises. This indicator is mainly used to measure the proportion of the main business cost of the real estate enterprise in the main business income. If this indicator is too large, it indicates that the real estate enterprise should take the main business cost seriously, control the cost of the real estate enterprise, and do a good job in its own cost management.

Figure 6 shows the comprehensive financial capacity values of different real estate companies. It can be seen that the main business cost rates of the four companies are higher than the fixed assets ratio in each quarter, and the fixed assets ratio and main business cost ratio of D company are the highest, indicating that D company has good fixed assets and good ability to resist the economic crisis; A company's ability to resist the economic crisis is poor, so it needs to prepare for the possible economic crisis in advance.
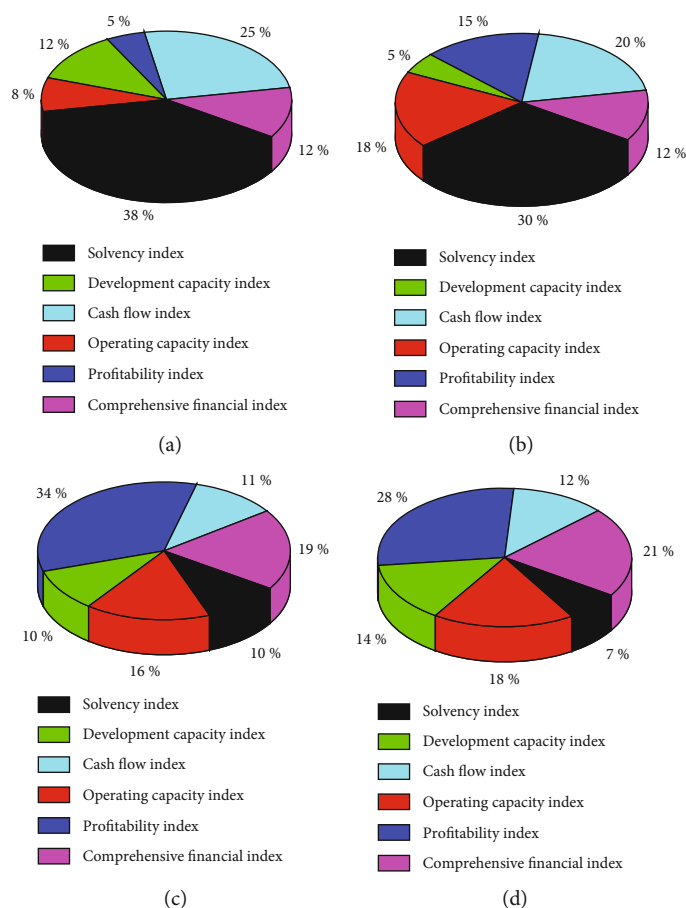
FIGURE 7: Economic capacity of different companies: (a) A company, (b) B company, (c) C company, and (d) D company.

Figure 7 shows the financial and economic capability analysis of four real estate companies. It can be seen from it that the profitability index of C company is the highest and that of company A is the worst. Therefore, the management and output of profitability should be strengthened. Combined with the analysis of the above four companies, the economic capacity of real estate companies based on PSO optimized SVM has the following suggestions:

(1) Real estate enterprises should be careful with the profitability and realize the growth of net profit for development projects. Choose appropriate projects for investment, do not invest blindly, otherwise, it will bring a serious burden to the enterprise. Because only by finding the growth point of profit, the enterprise will not fall into financial crisis

(2) Reasonable use of financial leverage, real estate enterprises will inevitably have to carry out financing. When financing, we must consider the impact of financial leverage on enterprise operation. Too high financing cost will aggravate the operation burden of enterprises and even bring experience risk

(3) Cash plays a small role in the enterprises. Enterprises must be careful with the same sales money to prevent the rupture of the enterprise's capital chain.

Of course, we cannot keep too much idle funds, which is bound to cause a certain waste. For repayment projects such as loans that are about to expire, sufficient funds should be prepared in time. So as not to affect the enterprise credit

(4) To strengthen cost control, real estate enterprises attach the significance to the impact of costs on enterprise operation. Excessive operating costs will increase the burden on enterprises and reduce their benefits. Therefore, real estate enterprises should strive to reduce daily operating costs and development costs, but the premise is to ensure the quality of goods and services

(5) Real estate enterprises should strengthen market development. At present, there are many real estate enterprises, and the products developed tend to be homogeneous. How to better base themselves on the local or regional market will be a very important problem. Only when sales increase, the corresponding sales revenue will go up

5. Conclusion and Outlook

Aiming at the possible economic crisis of real estate companies, this paper establishes an optimal PSO-SVM model

based on the existing SVM model and then uses PSO for optimization analysis. Then, the model is applied to four real estate companies in Beijing to analyze the economic ability of the company with solvency index, operating ability index, development ability index, profitability index, cash flow index, and comprehensive financial index. The results show that the profitability index is the most important for real estate companies. Only companies with good profitability can be able to deal with the possible economic crisis. In addition, the PSO-SVM model can effectively warn the economic crisis of real estate companies and provide a certain application basis and theoretical support for the large-scale application of the model.

Data Availability

The labeled dataset used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares no competing interests.

References

[1] S. Hao and H. Huang, "Prediction algorithm for the ship traffic based on PSO-SVM," *Microcomputer and Its Applications*, vol. 34, no. 5, pp. 1–3, 2015.

[2] T. Liu, X. U. Dacheng, and H. Zhao, "Research of temperature compensation method for MEMS accelerometer based on SVM optimized by PSO," *Modern Electronics Technique*, vol. 41, no. 10, pp. 58–66, 2018.

[3] C. Xu, L. Li, J. Li, and C. Wen, "Surface defects detection and identification of lithium battery pole piece based on multi-feature fusion and PSO-SVM," *Access*, vol. 9, pp. 85232–85239, 2021.

[4] Y. Song, X. Xie, Y. Wang, S. Yang, W. Ma, and P. Wang, "Energy consumption prediction method based on LSSVM-PSO model for autonomous underwater gliders," *Ocean Engineering*, vol. 230, p. 108982, 2021.

[5] K. Hu, M. Jiang, H. Zhang, S. Cao, and Z. Guo, "Design of fault diagnosis algorithm for electric fan based on LSSVM and Kd-tree," *Applied Intelligence*, vol. 51, no. 2, pp. 804–818, 2021.

[6] Y. Wang, M. Zhang, X. Tang, F. Peng, and R. Yan, "A kMap optimized VMD-SVM model for milling chatter detection with an industrial robot," *Journal of Intelligent Manufacturing*, vol. 20, no. 3, pp. 134–143, 2021.

[7] X. Wang, X. Zhang, and M. M. Shahzad, "A novel structural damage identification scheme based on deep learning framework," *Structure*, vol. 29, pp. 1537–1549, 2021.

[8] B. Wei, B. Liu, D. Yuan, Y. Mao, and S. Yao, "Spatiotemporal hybrid model for concrete arch dam deformation monitoring considering chaotic effect of residual series," *Engineering Structures*, vol. 228, no. 1, p. 111488, 2021.

[9] J. Li, J. Lu, C. Chen, J. Ma, and X. Liao, "Tool wear state prediction based on feature-based transfer learning," *The International Journal of Advanced Manufacturing Technology*, vol. 113, no. 11-12, pp. 3283–3301, 2021.

[10] L. Huang, K. Wu, W. Huang et al., "Detection of fusarium head blight in wheat ears using continuous wavelet analysis and PSO-SVM," *Agriculture*, vol. 11, no. 10, p. 998, 2021.

[11] S. Raj and K. C. Ray, "ECG signal analysis using DCT-based DOST and PSO optimized SVM," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 3, pp. 470–478, 2017.

[12] M. Duan, "Short-time prediction of traffic flow based on PSO optimized SVM," in *2018 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)*, pp. 41–45, Xiamen, China, 2018.

[13] P. J. García-Nieto, E. García-Gonzalo, J. A. Vilán Vilán, and A. Segade Robleda, "A new predictive model based on the PSO-optimized support vector machine approach for predicting the milling tool wear from milling runs experimental data," *International Journal of Advanced Manufacturing Technology*, vol. 86, pp. 1–12, 2015.

[14] Y. Li and Z. Li, "Forecasting of coal demand in China based on support vector machine optimized by the improved gravitational search algorithm," *Energies*, vol. 12, no. 12, pp. 2249–2267, 2019.

[15] J. J. Cheng, "Research on data mining technology of ship traffic based on cloud platform of internet of things," *Ship Science and Technology*, vol. 6, no. 3, pp. 82–89, 2018.

[16] K. X. Zhang and T. Universtiy, "Design of ship traffic management system based on internet of things architecture," *Ship Science and Technology*, vol. 12, no. 1, pp. 111–120, 2018.

[17] F. Yang, H. E. Zhengwei, H. E. Fan, and S. O. Fan, "Research on waterway depth prediction method based on depth network and ship traffic flow," *Journal of Wuhan University of Technology(Transportation Science & Engineering)*, vol. 5, pp. 54–62, 2019.

[18] Y. Liang and H. Zhang, "Ship track prediction based on AIS data and PSO optimized LSTM network," *International Core Journal of Engineering*, vol. 6, no. 5, pp. 23–33, 2020.

[19] M. A. Quandang, F. Jiang, Q. Fan, and R. Zhu, "Application of PSO-unbiased grey Markov model in ship traffic flow prediction," *Navigation of China*, vol. 5, pp. 34–42, 2019.

[20] G. U. Zhaojun, L. I. Bing, and T. Liu, "Network traffic prediction based on PSO-Elman model," *Modern Electronics Technique*, vol. 10, no. 2, pp. 4–9, 2019.

[21] A. N. Jiang, "Forecasting nonlinear time series of surrounding rock deformations of underground cavern based on PSO-SVM," *Rock and Soil Mechanics*, vol. 28, no. 6, pp. 1176–1180, 2007.

[22] B. Wang, Y. Sun, X. Ji, Y. Huang, L. Ji, and L. Huang, "Soft-sensor modeling for lysine fermentation processes based on PSO-SVM inversion," *CIESC Journal*, vol. 63, no. 9, pp. 3000–3007, 2012.

[23] W. Zhou, M. Chen, Z. Yang, and X. Song, "Real estate risk measurement and early warning based on PSO-SVM," *Socio-Economic Planning Sciences*, vol. 77, p. 101001, 2021.

[24] L. Fang and Z. Q. Zhang, "QoS prediction of web service based on PSO-SVM," *Journal of information and computational science*, vol. 12, no. 11, pp. 4415–4423, 2015.

[25] S. Chen, Y. Zheng, Z. Jiang, and L. Fangwei, "Network security situation prediction method based on PSO-SVM," *Application Research of Computers*, vol. 23, pp. 25–31, 2015.