

Retraction

Retracted: Research on the Model of Music Sight-Singing Guidance System Based on Artificial Intelligence

Complexity

Received 23 January 2024; Accepted 23 January 2024; Published 24 January 2024

Copyright © 2024 Complexity. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1].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

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

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

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

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

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

References

- [1] T. Zhe, "Research on the Model of Music Sight-Singing Guidance System Based on Artificial Intelligence," *Complexity*, vol. 2021, Article ID 3332216, 11 pages, 2021.

Research Article

Research on the Model of Music Sight-Singing Guidance System Based on Artificial Intelligence

Tong Zhe 

Northwest Normal University School of Music, Lanzhou 730072, Gansu, China

Correspondence should be addressed to Tong Zhe; zhetong@nwnu.edu.cn

Received 7 April 2021; Revised 27 April 2021; Accepted 29 April 2021; Published 11 May 2021

Academic Editor: Zhihan Lv

Copyright © 2021 Tong Zhe. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Intelligent Guided Learning System (ITS) is a computer system that uses computers to imitate the experience and methods of teaching experts to assist in teaching; ITS provides learners with personalized learning resources and adaptive teaching methods, thus reducing students' dependence on teachers and realizing independent learning. After years of research and development, and with the help of artificial intelligence technology, ITS has basically formed a stable structure and unified implementation specifications and has given birth to many excellent products in disciplines such as basic computer, language, medicine, and mathematics. Based on the above background, this paper takes the system structure of universal ITS as the basis, combines the characteristics of music sight-singing subject in teaching contents and teaching methods, researches the intelligent guidance system model of music sight-singing, and completes the design of the overall system architecture. The specific strategies and implementation methods of the teaching methods, resource recommendation, and ability assessment components of the teaching model are studied. The research includes the definition of the difficulty characteristics of the score, the design of the score recommendation algorithm, the design of the sight-singing scoring algorithm, and the experimental analysis of each algorithm. It is proposed that the difficulty feature-based score recommendation algorithm is the core component of resource recommendation in the teacher model, and the sight-singing scoring algorithm is the basis for ability assessment and learner model update.

1. Introduction

Personalized teaching is the key to improve teaching and cultivate innovative talents, but in the traditional teaching environment, one-to-one personalized teaching for each student is difficult to achieve due to the limited teaching resources and teachers' strength [1]. With the rise of Internet teaching, the limitations of traditional teaching methods in time and space are solved, and the web-based teaching platform builds virtualized teaching scenarios and teaching resources so that teaching behaviors are no longer limited to classrooms and classrooms. Students can learn freely through online learning platforms with unlimited time and location, and the platforms provide various services for learners like private teachers to provide one-on-one guidance and help [2]. However, simply providing learners with learning resources and learning environment cannot truly realize personalized teaching but also requires the learning

platform to actively interact with learners, understand their individual characteristics, and provide targeted teaching services for different learners by imitating the teaching style of teachers. The intelligent tutoring system uses artificial intelligence technology to simulate and learn the teaching style of human teachers, provide personalized learning paths and learning suggestions for different learners according to their needs, and understand learners' learning habits and preferences through verbal communication, behavioral analysis, and simulation during the learning process to dynamically guide learners to complete the mastery of knowledge [3].

In recent years, artificial intelligence technology has made significant progress in deep learning, intelligent control, data mining, and artificial neural networks, which has greatly promoted the development of intelligent guidance systems, formed a standardized structural framework, and given birth to all many excellent products in disciplines

such as basic computer, language, medicine, and mathematics [4]. The discipline of music sight-reading is a course that aims to develop students' ability to read, sing, memorize, and recognize music, as well as perceive, analyze, express, and imagine music, based on proficiency in the basic principles of music and is a basic skill that all musicians engaged in music must master has traditionally been valued by people engaged in professional music education [5]. With the continuous development of the music education system, the teaching concept and teaching methods of sight-singing and ear training have been innovated; however, the teaching form has been using classroom teaching and training for many years, and this boring teaching form is not only inefficient but also greatly affects the students' interest in learning and lack of learning initiative [6]. Although more and more multimedia tools and teaching software are being used in the classroom for sight-singing teaching, these tools (software) can only assist in teaching, and the existing online courses and online learning platforms, which only provide learners with rich learning resources and different learning pathways, do not yet meet the needs of intelligent teaching [7]. There is a lack of research on the design and implementation methods of intelligent guidance systems, and only a few visual singing learning systems are still inadequate in terms of personalization and intelligence [8]. The main manifestations are the lack of targeting in the management and pushing of learning resources and the reliance on manual judgment in the assessment of learners' ability level (proficiency of knowledge acquisition), which cannot be fully automated and intelligent. The assessment method that only considers the differences between the template and the performance has a large error [9]. In addition, the same problem as with learning resources is that there is no standardized classification of the difficulty level of the score, and it relies on the subjective judgment of experts. The questions in the assessment of sight-singing teaching can only come from the expert database or textbooks, and the personal preference and professional level of the questions will affect the balance of the difficulty level of the test questions, which cannot guarantee the fairness, accuracy, and objectivity of the assessment [10]. Therefore, it is a challenge for the current sight-reading system to achieve automatic scoring and to provide questions based on the difference in difficulty of the score [11].

In order to get rid of the above dilemma and realize the intelligence of music sight-singing teaching, in addition to the research on the design of the intelligent guide system for music sight-singing, it is also necessary to propose effective algorithmic support for the solution of key problems such as personalized recommendation of music scores and singing scoring and to develop an effective guide strategy, as shown in Figure 1. This study takes the existing research results of intelligent guidance system as the theoretical basis, combines the disciplinary characteristics of music sight-singing, and deeply investigates the model and its design method in the intelligent guidance system of music sight-singing, accumulating valuable theoretical basis for the implementation of the intelligent guidance system; at the same time, the algorithm proposed in this paper provides a feasible solution

to the problems of personalized teaching resources pushing and ability assessment in music sight-singing. The algorithm proposed in this paper also provides a feasible solution to the problems of personalized teaching resources and ability assessment in music sight-singing, which has a positive significance to further improve personalized teaching of music sight-singing.

2. Related Work

The prototype of intelligent guidance system is a computer-assisted teaching system that appeared in the middle of the 20th century, and with the development of artificial intelligence technology and its in-depth application in the field of education, it developed into an intelligent teaching system with cognitive ability by the end of the 20th century [12]. The definition of an intelligent guidance system is still inconclusive, but it is generally accepted that an intelligent guidance system is a learning support system that synthesizes and analyzes the learning of all types of students in a course with the help of artificial intelligence technology, combining the teaching knowledge and teaching models of specific subjects, and proposes a suggested learning solution that is decision-making for learning. Various definitions of its component structure also exist, where expert module, teaching module, student module, and interactive interface are recognized as four essential components [13]. The grasp of the difficulty level of a score in music teaching not only helps students to choose the appropriate score for their level as learning material but also improves the balance and fairness in teaching assessment. Although there are no clear criteria for classifying scores or classifying them according to performers' subjective interpretations, there are classifications for score difficulty on score websites, music grade exams, and textbook layouts [14]. For example, the sheet music download website classifies scores into 10 difficulties based on approximate pitch, scale, harmony, and rhythm [15].

The British ABRSM, which provides a grading system for evaluating music performance and comprehension, is divided into 8 grades based on scales and harmonics, playing methods, and so on. In the case of the American NYSSMA, there are 6 grade standards based on scales and octaves, scales, playing speed, and so on, which are subjectively classified out in conjunction with the difficulty of the music. The syllabus of the Sight-Singing and Ear Grading System (SSMA) also classifies the teaching content from easy to difficult according to the complexity of the pitch and rhythm of the score [16]. The current research on difficulty level recognition of piano scores is mainly based on the use of existing textbooks or scores with clear difficulty level labels given by authoritative organizations, and the use of linear regression support vector machines and other methods to construct a model for difficulty recognition trying to use the model to predict the difficulty level of a large number of unlabeled scores in order to achieve automatic classification of scores by difficulty level [17]. Melodic complexity has been studied for a long time, and it has been used as the main indicator of investigation in studies such as the arousal

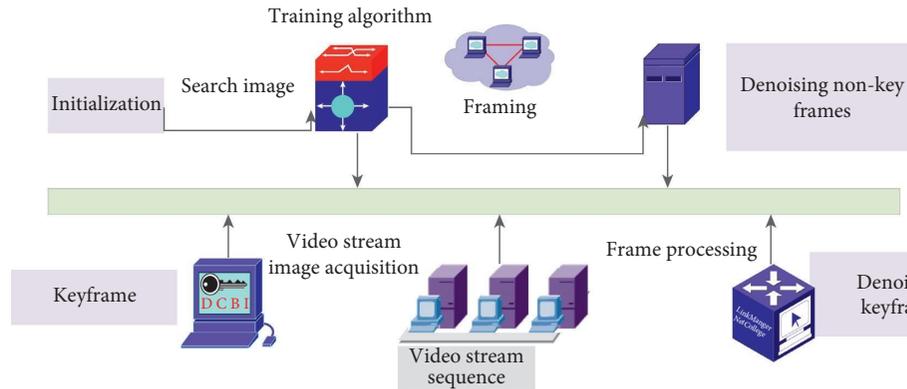


FIGURE 1: Structure of the general intelligent guided learning system model.

potential of artistic stimuli and the analysis of hedonic value (preference) music popularity factors [18]. These studies illustrate the potential of melodic complexity as a diagnostic tool for determining when melodic contingencies occur, or what aspects of melody (pitch combinations, hierarchical structure, sequential order, etc.) are more important for parsing together melodic structures (similarity ratings, melodic repertoire, etc.) [19–23].

Models of melodic complexity were first proposed by Eugene Narmour, who suggested that the implicit realization model performed well in the description of melodic complexity of original folk songs but lacked consideration of the musical and cultural context of the listeners. For this reason, Eerola et al. proposed an expectation-based model (EBM) to estimate complexity, and they found that tonality (modified by rhythmic position and pitch duration), the interval principle of range direction and range variation, rhythmic principle slicing, and rhythmic variability showed significant ability in predicting listeners' complexity judgments. Model construction in addition to feature definition has algorithm selection. Cilibrasi et al. questioned the validity of musical clustering algorithms and used standardized information distance as a generalized similarity metric to construct models. In a recent study, Eerola et al. used information theory and expectation violation as a model for constructing melodic complexity and validated it on seven datasets where overall static ratings of melodic complexity have been collected and performed well. The traditional steps of music scoring are first extracting the testers' singing music, extracting melodic features from the singing input signal, and then comparing the extracted melodic features with the template melody for similarity and measuring the music performance by their differences. Among the melodic features, there are three main types of melodic features: reference frequency, Mel to spectral coefficients, and pitch. For different features, various researchers have proposed different extraction methods, such as MFCC extraction using fast Fourier transform and pitch features using SHS algorithm. The dynamic time regularization (DTW) algorithm is a resource recommendation algorithm based on the difficulty of the score, which can improve the learning efficiency and assessment effectiveness by pushing the score with the appropriate difficulty level for the students during the study

or quiz. DTW algorithm is currently used more often on the difference comparison, using distance to represent the similarity. Different studies select different types of features or combine different types of features according to their scoring purposes and focus, and the differences in scoring effects all depend on the selected features.

3. Music Sight-Singing Domain Knowledge Model

Sight-singing is a highly technical and fundamental subject in the field of music teaching. The content of sight-singing teaching includes both the study of basic music theory and the application of theoretical knowledge skills, so it also has the characteristics of a comprehensive subject. The connotation of sight-singing teaching is to cultivate students' musical perception, understanding, appreciation, and creativity through their training in rhythm, pitch, tonal sense, and ability to read music, so that students can master the correct pitch, rhythm, and the ability to express musical emotions, so that they can more easily understand the image of music, perceive the style, type, and theme of music, and lay a solid foundation for future music learning. This means that the teaching of sight-singing is not only about improving students' ability to sing from sight but also about developing their musical abilities. Research has shown that the development of musical ability in sight-singing is achieved by improving students' sense of rhythm, melody, intensity, and harmony by.

- (1) Rhythm sense is the learner's ability to perceive and reproduce the rhythmic changes in a musical work. The development and training of rhythm sense enhances the learners' ability to regulate their breath and tempo during musical performance and is one of the basic tools for developing students' ability to perceive, understand, and create music.
- (2) Melody is the organic combination of musical elements such as key, rhythm, and beat in a musical work and contains a rich artistic image. Accurately perceiving and expressing the melody of a musical work is a key indicator of one's comprehensive musical ability. Only by accurately perceiving the

melody of a musical work can one have a systematic and comprehensive understanding of music as a whole and only then can use the best singing method to express the artistic image in music.

- (3) Strength is the main means of expressing emotion in music works, grasping the strength structure of music works in order to properly show the emotion of music; the cultivation of the sense of strength is the exercise of the ability to express the emotion of music, which can enhance the infectious power and attractiveness of students singing.
- (4) Harmony is the formation of a fixed acoustic combination of two or more tones according to specific rules and standards. Harmony reflects some of the characteristics of music in terms of style and genre and is important for the establishment of musical themes, and the cultivation of a sense of harmony is conducive to improving students' sense of hierarchy and three dimensionality of musical expression and enhancing the ability to coordinate and cooperate.

4. Sight-Singing Subject Knowledge Classification

The discipline of sight-singing has a long history and has become a basic discipline of music learning in China after nearly one hundred years of development, with a relatively complete teaching system and relatively stable teaching content. This paper summarizes and classifies the knowledge contained in the discipline of sight-singing by analyzing the teaching content of the existing textbooks on sight-singing. Because of the distribution of chapters in the textbook, the table only describes the content of each chapter in general, and the chapters involving the same knowledge points are combined. Although the above textbooks have their own focus and characteristics in terms of content layout, system, concept, and method, the specific contents involved are similar. Each section is divided into two parts: the fundamentals of music theory (key, beat, and tone) and sight-reading exercises (tonality, rhythm, and intervals). The foundation of music theory is the prerequisite of sight-reading, and only by accurately identifying the tones and keys in the score can we sing them out; the sight-singing practice is the key to improve the ability of sight-singing, and only through repeated practice and self-summary can we grasp the rhythm and intervals in the melody of music, as shown in Figure 2. In addition, the connotation of the teaching of sight-singing is the cultivation of students' musical ability, so the subject knowledge in addition to the fundamentals of music theory and sight-singing techniques should also include musical ability. In summary, the knowledge of the subject of sight-singing can be divided into the following three categories.

- (1) Music theory includes relatively simple basic theory (score reading, intervals, chords, rhythms, beats, etc.) and relatively advanced harmony, polyphony, tune, melody, orchestration, and so on. In

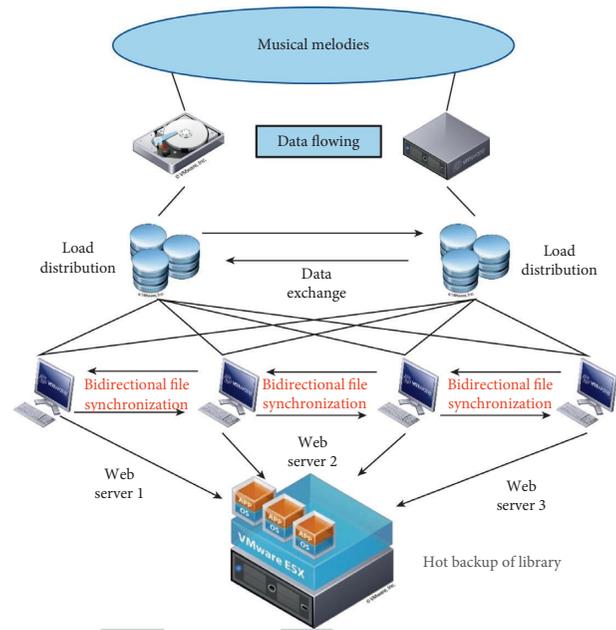


FIGURE 2: Rhythmic and interval changes in musical melodies.

this article, the fundamentals of music theory are divided into three types: basic concepts, musical notation, and musical elements, each of which includes several knowledge points. The basic concepts of the pentatonic score, such as clef, bar, chant, and tone row, are the main ones. Music notation is mainly the various notes and notations that make up a score, such as notes, mutation notation, and rests. The basic and compound elements are the musical elements in the score based on one or more symbols, such as intervals, beats, and tones.

- (2) For sight-singing techniques, the main purpose of sight-singing exercises is to master and strengthen sight-singing techniques. The reason why it is called "technology" is because the sight-singing practice is concerned with the student's ability to perform the basic elements of music (intervals, timbre, intensity, speed, and rhythm) and the ability to appreciate the organization of musical forms (melody, harmony, polyphony, tune, etc.), for example, the ability to sing simple-to-complex intervals and rhythmic patterns with pitch and rhythmic accuracy, the ability to accurately identify the tonality and modality of any score, and the ability to express basic chords and changing chords in their entirety with assistance.
- (3) The development of musical ability requires the exercise and improvement of students' sense of rhythm, melody, intensity, and harmony, that is, the training of students' ability to perceive and express the rhythmic, melodic, intensity, and harmonic elements of musical works. Students can improve their sense of melody through complex melodies, improve their sense of rhythm and harmony through special exercises (e.g., model singing, composition singing, and beat-playing), and train their ability in rhythm,

melody, intensity, and harmony in a comprehensive manner through whole song sight-singing.

5. Sight-Singing Knowledge Model

Sight-singing subject knowledge modeling is the basis for personalized instruction in the Sight-Singing Intelligent Guided Learning System. The knowledge model provides a structured description of domain knowledge, and this description reflects the relationships between knowledge and between knowledge and instructional resources. The use of the knowledge model representation then enables personalized instructional resource management and push. According to the relevant principles of pedagogy, there are three main relationships between the knowledge points of a subject: interdependence, parent-child relationship, and mutual independence (parallel relationship).

Dependency relationship is defined according to the process of knowledge acquisition: if knowledge a depends on knowledge b , then the order of learning is b first and then a . Parent-child relationship is the inclusion (hierarchical) relationship between knowledge, and the parent node knowledge and the child node knowledge in the knowledge structure tree are the parent-child relationship; there is no dependency and inclusion between knowledge points in parallel relationship, and they are independent of each other in knowledge structure and knowledge acquisition, and they are independent and do not affect each other in knowledge structure and knowledge acquisition. Dependency and parent-child relationship need to be explicitly reflected in the construction of knowledge model; that is, the structure of describing knowledge points should contain two attributes of dependent knowledge and parent knowledge. Learning resources are the support objects of knowledge point learning, and diverse data types can provide different ways of displaying and interacting with learning resources, so as to meet the learning choices preferred by different learners.

According to the classification of sight-singing subject knowledge, sight-singing teaching resources mainly include example resources, sight-singing skills learning resources, music ability training resources, and teaching videos. The example resources, sight-singing skills learning resources, and music ability training resources are mainly various types of music files, such as pentatonic images, music modeling audio, MIDI files; the teaching videos are the curriculum resources that explain the knowledge and training skills. The learner model is to describe and quantify learner information using complex data structures so that the computer can effectively identify the learning characteristics of learners. In an intelligent tutoring system, the individual characteristics of learners are the main decision basis for the system to provide personalized learning services to them. The algorithm analyzes the learner's characteristic data to assess the learner's learning status and ability level and then provides him/her with suggestions and guidance on learning resources and learning methods under the established rules based on the assessment results. The learner model standard defines the characteristics of learners from a general

perspective, but, in practice, it needs to be refined, modified, and customized in accordance with the design purpose of the system and the characteristics of learners in specific disciplines in order to better provide personalized services for learners and improve the learning effect. In this paper, the learner model is designed from four aspects: basic information, learning achievements, learning behaviors, and interest preferences, taking into account the characteristics of the learners of visual singing.

- (1) For basic information, sight-singing is a basic subject of music teaching, and there are courses related to it in China from the compulsory education level, to professional music colleges, and then to the music faculties of comprehensive universities. Due to the different stages of education and the time of learning music, the needs of the learners in the content and methods of teaching and learning vary greatly, so the basic information in addition to name, ID, gender, age, school, and major, but also need to include school age, music specialties, music field, and other content. Music expertise refers to a learner's previous musical achievements, especially the grades obtained in vocal exams, and so on. Music field refers to the music subdiscipline in which the learner mainly studies or works, such as dance, instrumental music, and singing. Basic learner information is defined as follows: basic information = (name, gender, age, school age, major, school, musical specialization, musical field, ID, password).
- (2) For learning achievements, sight-singing is a highly technical subject, and the content of the instruction includes both the learning of basic music theory and the emphasis on the application of theoretical knowledge skills. Therefore, learners' achievements are examined and evaluated at two levels: proficiency in knowledge and proficiency in sight-singing. The proficiency of knowledge is the learners' familiarity with the basic knowledge of music theory, such as the concept of beat, rhythm, chords, and other musical elements; the level of sight-singing ability mainly includes the accuracy of pitch, rhythm, singing speed, rhythm, melody, strength, and harmony. The specific content of learning achievement is as follows: knowledge proficiency = (beat, rhythm, rhythmic pattern, tonality, modulation, interval, chord) and sight-singing ability level = (pitch, rhythmic accuracy, singing speed accuracy, sense of rhythm, sense of melody, sense of intensity, sense of harmony).
- (3) Learning behavior mainly refers to the learning process of specific subject content, which in turn includes three major categories: basic knowledge learning, sight-singing skill practice, and music ability development. Therefore, learning behavior not only records the accumulated learning hours and times, the start time of this learning, and the end of this learning, but also records the learning hours and times of each learning content in one learning

process. The learning behaviors are described as follows: learning behavior = (total learning hours, total number of learning times, this start time, this end time, total basic learning hours, total skill practice hours, total ability development hours, total basic learning times, total skill practice times, total ability development times).

- (4) Interest preferences in the learning of sight-singing include the frequency of using various types of learning resource and, the frequency of learning different learning styles. The learning resources include images (pentatonic pictures), digital scores (MIDI), audio, video, and learning tools; the learning methods include resinging, enough singing, model singing, and whole song singing. The preference level is calculated by the cumulative number of times and hours of use. The definition is as follows: interest preference = (image, digital score, audio, video, learning tool, reprise, composition, model singing, whole song singing). In summary, the structure of the model designed in this paper for learners of sight-singing is shown in Figure 3.

6. Design and Verification of Key Algorithms

6.1. Learning Resource Recommendation Algorithm. To implement a recommendation algorithm based on the difficulty of a score, we first need to extract features that can indicate the difficulty level of a score. There is no unified standard for defining the difficulty features of a score. The SMPLevelGuidelines defined by SheetMusicPlus, a sheet music sharing website, uses the approximate number of pitch, scale, harmony, and rhythm occurrences in a score as the difficulty level indicator: the Sight-Singing and Ear Training. In the study of the calculation of the complexity of the score, eight complexity indicators are proposed based on the information entropy theory: average interval, pitch ambiguity, level distribution entropy, interval distribution entropy, note density, rhythmic variation, and harmonic distribution. In the study of the identification of the difficulty of piano scores, 25 difficulty-related features defined 25 features related to difficulty such as key velocity entropy, beat variation, and chords. After analyzing the existing score difficulty features, combined with the discipline characteristics of music sight-reading and the application scenario of score difficulty in this system, 17 score difficulty-related features are defined in this paper, and the detailed description and extraction method are shown in Figure 4.

In the process of learning and assessment, it is often necessary to select appropriate scores from the question bank according to the difficulty level; for example, students often want to practice intensively on scores with the same difficulty level as the learned scores. For this reason, we propose a resource recommendation algorithm based on the difficulty of the score, which can improve the learning efficiency and assessment effectiveness by pushing the score with the appropriate difficulty level for the students during the study or quiz. In the previous sections, the validity of the score difficulty feature is verified, and the

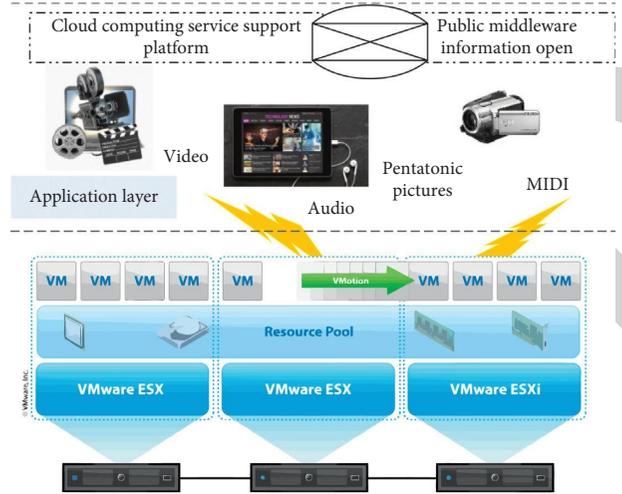


FIGURE 3: Learner model structure.

differentiation of the score difficulty by various similarity calculation methods on the feature space is also experimentally compared. Supported by the experimental results, this chapter proposes a collaborative filtering algorithm based on Euclidean distance to construct a resource recommendation system. The recommendation algorithm follows the following steps to recommend sheet music for learners to learn and evaluate in real time, and the specific process is shown in Figure 5.

The evaluation of sight-singing in this study uses midi score files as a standard reference, and MIDI describes the notes in the score in terms of pitch values and pitch length (duration of pitch), so it is necessary to extract the pitch characteristics and their duration in the sight-singing audio. The MIDI files use semitone values to represent the pitch (Pitch) of the notes, and the semitone, and the semitone and the fundamental frequency have a correspondence as expressed in the following equation:

$$P = 12 \log \frac{f_d}{q} + 69, \quad (1)$$

where 69 is the semitone value corresponding to the international standard tone A, f_d indicates the fundamental frequency, and 144 is the frequency difference between the two semitones. After extracting the fundamental period corresponding to each frame in the audio using the YIN algorithm, the fundamental frequency is obtained by taking the reciprocal of the fundamental period and then converting it into the number of semitones in MIDI using the above formula. Figure 6 shows the pitch sequence curve extracted from the audio, from which the ups and downs of the pitch over time can be seen. However, due to the accuracy of the algorithm and the influence of the noise data, there are many wild points in the pitch sequence, and for better matching and comparison with the template, the pitch sequence needs to be smoothed. The commonly used smoothing algorithms are linear smoothing and median smoothing; however,

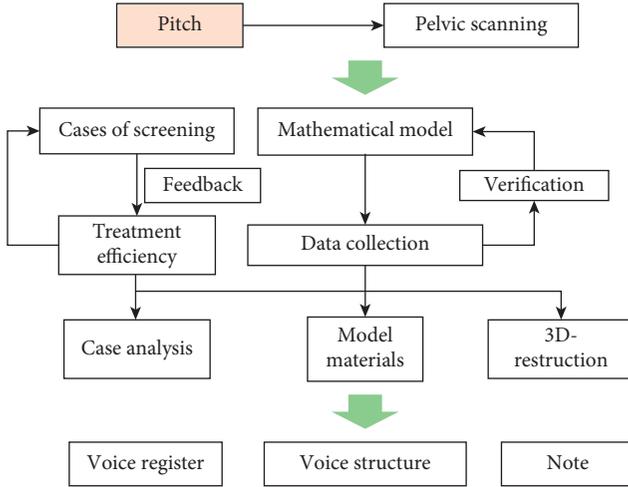


FIGURE 4: List of difficulty characteristics of the score.

both methods require multiple smoothing and will cause the normal pitch to be time-shifted, so this paper proposes a scenario-optimized smoothing method. The model examines the coordinated development of these three subsystems from a holistic perspective:

$$U = \sqrt{\frac{1}{q} \sum_{i=1}^q X_{ii}}, \quad (2)$$

$$E_i = \sum_{j=1}^q (x_j + x_{ij}),$$

where E_i denotes a subsystem i in the 3E system, E_1 , E_2 , and E_3 denote the development levels of the energy subsystem, the economic subsystem, and the environmental subsystem, respectively, x_{ij} denotes the j th evaluation index of subsystem i , and X is the corresponding weight of this evaluation index. From (1), it is easy to calculate and is essentially the geometric average of the development level of the three subsystems, and there is no specific requirement for the value of the comprehensive development level of the system to be positive, negative, or zero. It can reflect the functional degree of the development level of the whole 3E system, but it cannot reflect the submission degree of the subsystems matching each other's development level.

The distance-based coordination model assumes the premise that the combined development levels of the energy, economic, and environmental subsystems must all be positive and is calculated as

$$E_{ii} = \sum_{j=1}^q (x_j + x_{ij}) \times \delta_{ij}, \quad (3)$$

where E_{ii} denotes the degree of matching of development levels among subsystems; V denotes the functional degree of the overall development level of the 3E system, which is essentially a weighted average of the development levels of each subsystem.

When evaluating the degree of coordination between the energy subsystem and the economic subsystem,

$$Q_{ii} = \sum_{i=1}^q E_{ii} \times E_{ij}, \quad (4)$$

q ($q > 2$) is the coordination coefficient, which is generally taken as 2; E_1 and E_2 denote the weights of the subsystems, respectively. When measuring the coordination degree among the three systems, that is, the total coordination degree of the Q system, E_1 and E_2 in the equation are expanded according to the system covariance theory and the discretization principle, respectively:

$$Q_{ii} = \alpha E_1 + \beta E_2. \quad (5)$$

In the YIN algorithm-based fundamental extraction, the audio signal is divided into frames, and the pitch sequences we obtained are also in frames. Assuming that each singer is aiming to sing the score accurately, the wild points that exist should not last long, and they should be preceded and followed by a smooth pitch subsequence. In fact, by analyzing the duration of wild points in the pitch sequences of a large number of sight-singing audio, it was found that a large number of wild points appeared in the middle of two smooth signals and lasted for only 1-2 frames. Therefore, we can smooth the pitch sequences as follows.

- (1) Take the frames with equal and close pitch values in the pitch sequence as a subsequence and count the number of frames
- (2) Iterate through the frames of each pitch subsequence to find the wild point with frame number between 1 and 2 and whose preceding and following pitch frames are greater than 2
- (3) Set the pitch value corresponding to the found wild point to the average value of its preceding and following subsequence pitches
- (4) Median filtering is applied to the optimized pitch sequence

To compare the smoothing effect, Figure 7 shows the effect plots using median smoothing and optimization followed by smoothing, respectively. It can be seen that it is difficult to process the wild points of the above cases with effective number of smoothing and smoothing parameters by using only conventional smoothing methods.

The above DTW algorithm is designed for the similarity calculation of all time series, and in the sight-singing scoring of this paper, the matching of pitch sequences can be improved by combining the special characteristics of the data. The matching process of pitch sequences in sight-singing scoring has the following two characteristics compared with the general matching of speech signals.

- (1) Sight-singing is the singing of each note according to the melody in the score, and the duration of each note in the score is much longer than the duration of one frame in the fundamental detection, usually in a multiplicative relationship. This means that the pitch sequence extracted from the sung audio is composed of many segments with the same pitch value as a subsequence. This means that when using the DTW

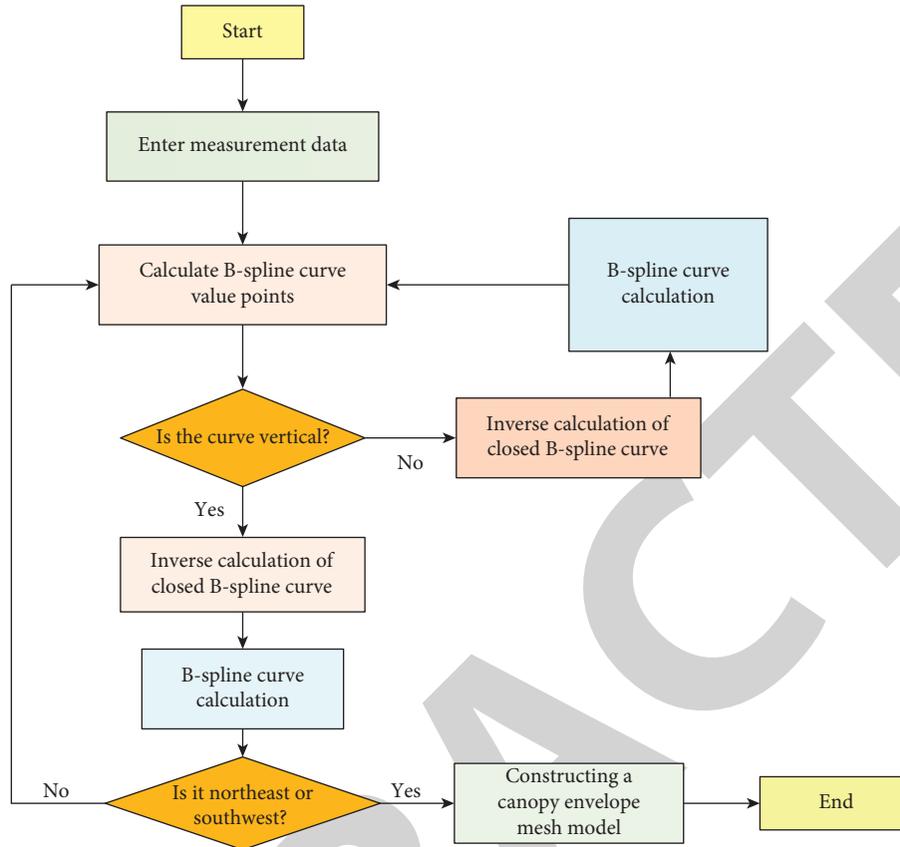


FIGURE 5: Flowchart of sheet music recommendation algorithm.

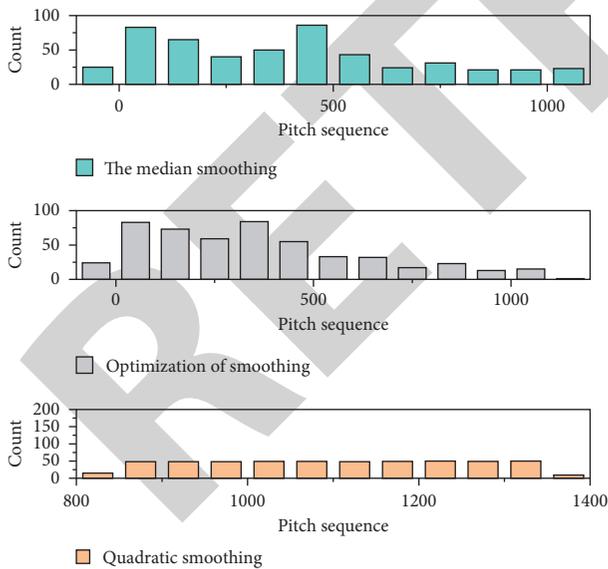


FIGURE 6: Pitch sequence diagram.

algorithm for pitch sequence similarity calculation, it is possible to match by subsequence, which can reduce the computation time of the algorithm exponentially.

- (2) There are three directions for the selection of the next path point in the original DTW, where horizontal

and vertical choices indicate that a point in sequence x can correspond to multiple consecutive points in y , and vice versa. In the sight-singing scoring, assuming that the pitch sequence of the sight-singing audio is X and the pitch sequence of the template is y , if multiple consecutive points in y are aligned with a point in x , it means that the singer missed some notes; in this case, these points in y are originally the position of the error, which can be ignored in the path selection of the DTW, and finally it can be corresponded to the neighboring points. In other words, the points in y can be skipped in the minimum cumulative distance path selection, and this improvement can also shorten the calculation time.

6.2. Analysis of Experimental Results. The time value of a note in a score is defined relative to the time value of a beat; for example, the time of a quarter note is a quarter of the time required to play a beat in the current score, and the time of a beat is determined by the tempo of the score. The tempo indicates how fast or slow the score is sung and is usually recorded as the number of beats per minute, the reciprocal of which is the time of a beat (in minutes); for example, 120 beats/minute means that a beat needs to be sung for 0.5 seconds. In the actual process of singing, the time of each beat is grasped by the singer himself; it is difficult to sing each note, precisely the standard time. Even if the same

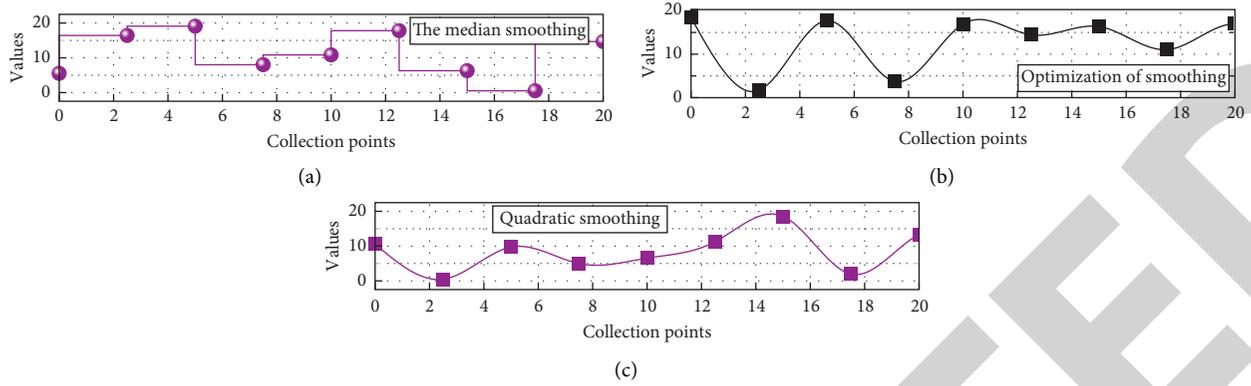


FIGURE 7: The effect of median smoothing and optimal smoothing.

person sings the same song several times, there are differences in individual notes and overall rhythm. Therefore, the pitch sequence cannot be matched to the template strictly according to time, and dynamic regularization operations such as time shifting and scaling of the pitch sequence are required. Considering this objective requirement, this paper uses the DTW algorithm from the previous section to implement template matching. Scoring the sight-singing requires calculating the similarity between the pitch sequence in the audio sequence and the template pitch sequence, and some abnormal data in the extracted pitch sequence are not caused by human factors, if the differences are calculated directly according to the temporal alignment as shown in Figure 8, where the blue represents the extracted pitch sequence and the brown is the pitch sequence of the template score, due to the different lengths of the two sequences. From the figure, we can see that the *a* section and the *b* section do not correspond to each other, and the comparison of the specific data shows that the *b* section is actually the singing part of the *a* section, which is caused by the singing error of the previous notes, so it is obviously unreasonable to calculate the difference between these two sequences directly. The use of DTW algorithm can avoid this problem very well, as shown in Figure 9; the *a* section and *b* section can be matched correctly now.

In order to select the appropriate similarity calculation method for the resource recommendation system, this experiment applies the kmi classification algorithm based on different distances in the feature space defined in the previous section to classify the 8noteS dataset and selects the best similarity calculation method by comparing the accuracy of the classification. The similarity calculation methods involved in the evaluation are Euclidean distance, Manhattan distance, cosine similarity, and Pearson coefficient. The experiment was validated by a 5-fold crossover method, and the experiment was repeated 5 times, and the average accuracy of the 5 times was used as the evaluation index, and finally the distance calculation method with the highest accuracy was selected and used for the resource recommendation system.

The experimental results are shown in Figure 10. The experimental results show that it was found that a large number of wild points appeared in the middle of two smooth

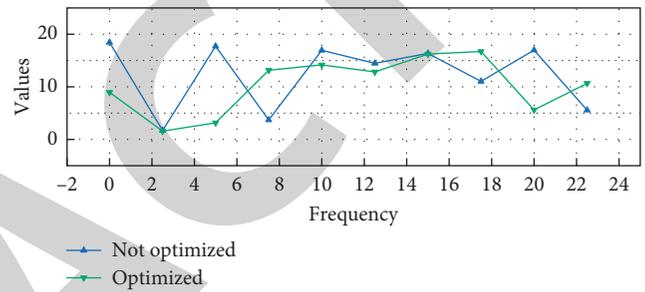


FIGURE 8: Pitch sequence diagram.

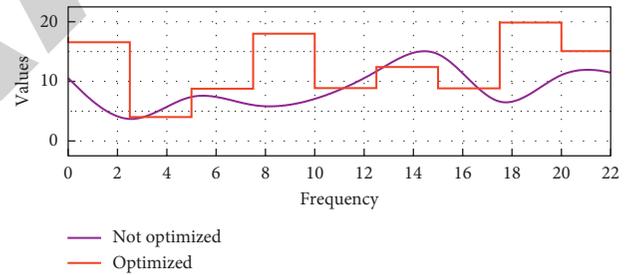


FIGURE 9: Pitch sequence DTW matching chart.

signals and lasted for only 1-2 frames. Therefore, the resource recommendation system in this paper uses the Euclidean distance to measure the similarity of difficulty between scores. In this paper, we evaluate the effectiveness of the feature set by its accuracy in the difficulty classification of the feature set by its accuracy in the difficulty classification of sheet music. If the feature set performs well in the difficulty classification, it means that these features can effectively represent the difficulty level of sheet music, and vice versa; it means that the feature set cannot effectively distinguish the difficulty difference between sheet music. For this purpose, we collected sheet music with difficulty level labels from the free music website 8noteS as the dataset for this experiment. 8noteS is currently the largest number of sheet music sharing websites with difficulty level labels that are available for free download.

The scores are divided into four difficulty levels, beginner, easy, intermediate, and advanced, and are available in both PDF and MIDI file formats. The advantages of MIDI

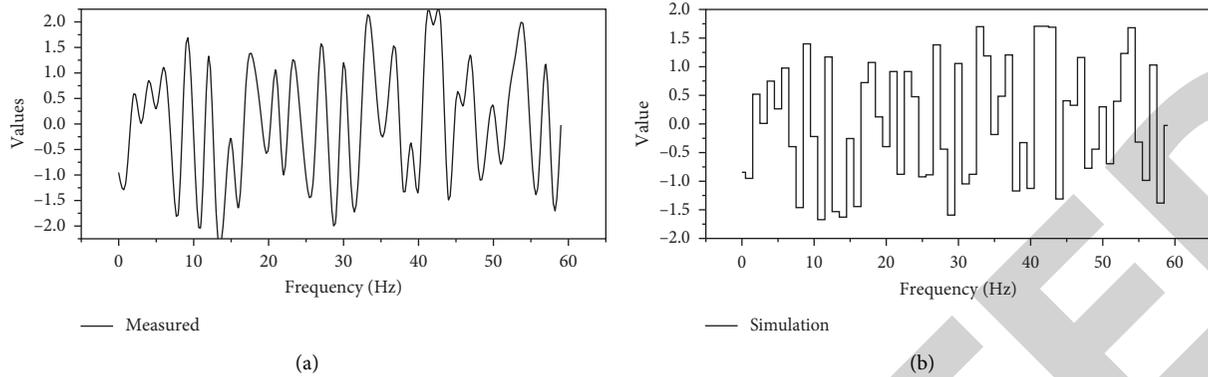


FIGURE 10: MIDI score quantization diagram.

format in score analysis have been discussed in the previous chapter, so this paper only collects scores in MIDI format from this website. Using web crawler tools, a total of 1780 scores were finally available for free download and in the correct format, of which 1395 were monophonic and monophonic, and the distribution of each difficulty level was beginner, 84; easy, 427; intermediate, 739; and advanced, 145. The unevenness of the sample categories in the classification algorithm can lead to overfitting of the model. In order to make the distribution of difficulty levels in the samples relatively even, the final 380 scores were taken as the final experimental data set, which is referred to as 8notes dataset in the following for the convenience of description. The data preprocessing mainly includes two steps: data validity check and feature standardization, which means to check the data format, eliminate the wrong files, and fix the errors; feature standardization means to normalize the extracted features and unify the feature outline. The starting and ending times of notes do not correspond to the starting and ending times of the beat, which will lead to the inability to correctly identify the notes during feature extraction. Therefore, before feature extraction, the score should be quantized first to fix these misaligned notes.

7. Conclusion

This paper provides a usable reference for the design and implementation of a personalized learning system for music sight-reading by studying the model and design of an intelligent guide system for music sight-reading; it also proposes a difficulty feature-based score recommendation algorithm and a pitch sequence matching-based sight-reading scoring method for solving two key problems in personalized recommendation of resources and assessment of sight-reading ability in an intelligent guide for music sight-reading. In this paper, a set of features describing the difficulty of a score is proposed, and its differentiation of score difficulty is better than those in the existing literature. Due to the technical nature of sight-singing teaching, the identification and differentiation of the difficulty of the score is the key to personalizing the sight-singing teaching resources. Only by accurately recommending teaching resources that match the difficulty of the score with the ability

of the learner can the learner use the resources effectively, master the knowledge quickly, and avoid the loss of interest in learning. Although there are many studies on intelligent guidance systems, most of them are based on general models, while intelligent learning is closely related to the characteristics of specific subjects, so the research on intelligent guidance systems can only be practical if they are combined with specific subjects. In this study, the models of intelligent tutoring systems are based on standard models and designed with the teaching characteristics of the subject of sight-singing, which are relevant and usable to a certain extent.

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Consent

Informed consent was obtained from all individual participants included in the study references.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

- [1] N. J. Foss and T. Saebi, "Fifteen years of research on business model innovation," *Journal of Management*, vol. 43, no. 1, pp. 200–227, 2017.
- [2] S. Welch, R. Johns, G. Nicholls et al., "Incorporation of a theme-based research model into a pharmacy department research program," *Journal of Pharmacy Practice and Research*, vol. 49, no. 3, pp. 262–269, 2019.
- [3] S. Ghosh, A. Watson, O. E. Gonzalez-Navarro et al., "Speed breeding in growth chambers and glasshouses for crop breeding and model plant research," *Nature Protocols*, vol. 13, no. 12, pp. 2944–2963, 2018.
- [4] P. d. Reffye, "Two decades of research with the GreenLab model in agronomy," *Annals of Botany*, vol. 127, no. 3, pp. 281–295, 2021.
- [5] H. Verbeek, S. M. G. Zwakhalen, J. M. G. A. Schols, G. I. J. M. Kempen, and J. P. H. Hamers, "The living lab in

- ageing and long-term care: a sustainable model for translational research improving quality of life, quality of care and quality of work,” *The Journal of Nutrition, Health & Aging*, vol. 24, no. 1, pp. 43–47, 2020.
- [6] C. Qing, X. Wu, X. Li, and W. Zhu, “Performance analysis of weather research and forecasting model for simulating near-surface optical turbulence over land,” *Optik*, vol. 188, pp. 225–232, 2019.
- [7] O. P. Lewandowska and M. A. Schmuckler, “Tonal and textural influences on musical sight-reading,” *Psychological Research-Psychologische Forschung*, vol. 84, no. 7, pp. 1920–1945, 2020.
- [8] A. Stenberg and I. Cross, “White spaces, music notation and the facilitation of sight-reading,” *Scientific Reports*, vol. 9, no. 1, pp. 1–12, 2019.
- [9] E. Huovinen, “Early attraction in temporally controlled sight reading of music,” *Journal of Eye Movement Research*, vol. 11, no. 2, pp. 1–30, 2018.
- [10] J. C. Smith, “Hidden in plain sight: a music therapist and music educator in A public school district,” *International Journal of Music Education*, vol. 36, no. 2, pp. 182–196, 2018.
- [11] H. Mitchell and B. Roger, “The moot audition: preparing music performers as expert listeners,” *Research Studies in Music Education*, vol. 39, no. 2, pp. 195–208, 2017.
- [12] P. Lelieveldt and H. B. B Evert, “City muscscapes: live music in the cities of groningen and utrecht, 2010–2016,” *Journal of Humanities and Social Sciences*, vol. 3, no. 1, pp. 42–54, 2018.
- [13] F. Delogu, “That does not sound right: sounds affect visual ERPs during a piano sight-reading task,” *Behavioural Brain Research*, vol. 367, pp. 1–9, 2019.
- [14] L. Herrero and N. Carriedo, “The role of cognitive flexibility and inhibition in complex dynamic tasks: the case of sight reading music,” in *Current Psychology*, pp. 1–13, Springer, Berlin, Germany, 2020.
- [15] T. Herbert, “Public military music and the promotion of patriotism in the British provinces, c. 1780–c. 1850,” *Nineteenth-Century Music Review*, vol. 17, no. 3, pp. 427–444, 2020.
- [16] A. J. Bovin, “The effects of frequent use of a web-based sight-reading software on eighth graders’ music notational literacy,” *Journal of Music, Technology and Education*, vol. 11, no. 2, pp. 131–147, 2018.
- [17] Ç. Özen, “The analysis and comparison of pre-service music teachers’ attitudes towards the piano lesson and their sight-reading skills in terms of certain variables,” *Journal of Education and Practice*, vol. 8, no. 29, pp. 162–167, 2017.
- [18] M. E. Düzbastılar, “The comparasion of music teacher candidates’ instrument test grades and instrument sight reading grades in talent exam with individual instrument exam grades during four-years education period,” *Journal of New Results in Science*, vol. 17, no. 1, pp. 315–323, 2020.
- [19] A. Adamyany, “Sight-reading as an important factor in the professional growth of future music teacher,” *Revista Vortex*, vol. 8, no. 2, pp. 1–16, 2020.
- [20] Y. Liu, J. Wang, J. Li et al., “Zero-bias deep learning for accurate identification of Internet of things (IoT) devices,” *IEEE Internet of Things Journal*, vol. 84, p. 2627, 2634.
- [21] S. Niu, Y. Liu, J. Wang et al., “A decade survey of transfer learning (2010–2020),” *IEEE Transactions on Artificial Intelligence*, vol. 1, no. 2, pp. 151–166, 2020.
- [22] J. Yang, J. Wen, B. Jiang et al., “Blockchain-based sharing and tamper-proof framework of big data networking,” *IEEE Network*, vol. 34, no. 4, pp. 62–67, 2020.
- [23] J. Yang, M. Xi, B. Jiang et al., “FADN: fully connected attitude detection network based on industrial video,” *IEEE Transactions on Industrial Informatics*, vol. 17, no. 3, pp. 2011–2020, 2021.