Feature Recognition Method based on Computer-Aided Technology and its Application in Music Teaching

Ning Fu[1,*] and Xia Peng[2]

[1]Normal College, Xinxiang Vocational and Technical College, Xinxiang 453000, China, funing@xxvtc.edu.cn
[2]Normal College, Xinxiang Vocational and Technical College, Xinxiang 453000, China, pxxppxxp2826@163.com

Corresponding author: Ning Fu, funing@xxvtc.edu.cn

Abstract. Due to the influence of factors such as strong music specialty, complex music theory knowledge and various changes, it is difficult to identify music features in the process of music teaching. Therefore, a music feature recognition system based on computer aided technology is proposed. The physical sensing layer of the system is equipped with sound sensors in different locations to collect the original music signals, and the digital signal processor is used to analyze and process the music signals. The network transmission layer will process the music signal and transmit it to the music signal database in the system application layer. The music feature analysis module in the application layer adopts dynamic time warping algorithm to obtain the maximum similarity between the test template and the reference template, realize the feature recognition of music signal, and identify music form and music emotion corresponding music feature content according to the recognition results. The experimental results show that the computer-aided music teaching system for music theory knowledge learning, works of music appreciation, music composition activity provides many functions, such as for teachers and students to provide a lot of learning resources, through the network technology to help music learners learn effectively and quickly, rich music knowledge, expand horizons, meet different users' personalized requirements.

Keywords: feature recognition, computer aided technology, music teaching, signal processing, data collection.
DOI: https://doi.org/10.14733/cadaps.2023.S4.123-132

1 INTRODUCTION

Music is one of the most common sound media. How we should adopt better information technology means to process music signal, this has become an inevitable trend, now has caused people's attention. How to quickly and accurately extract the information you need from a large amount of information has a very important application prospect. Zubair and Abu Mansor [1]

mentioned that music is different from other subjects in teaching to a certain extent because of its particularity of acting directly on people's emotional organs with sound. First, music works only through the auditory system, whereas other disciplines work mainly through the visual system. At this point. At least make music teaching in the choice and use of media and other subjects produce differences; Secondly, music does not follow the organization logic of language, it has its own unique way of information organization. Therefore, the processing and analysis methods of language information may not be suitable for music information analysis. Moreover, language works on man by means of meaning, while music works on man by means of feeling; Language can express meaning in individual words, music can express meaning in whole, and so on. In addition, compared with other subjects, cognition in music teaching is the way, means, not the goal, aesthetic and feeling is its goal. From this point, we can say that music is a higher level of cognition, a special mode of cognition of music formed on the basis of specific cognition of music. Therefore, used mainly for aim at specific cognitive teaching system design theory to solve the problem of music teaching is necessarily flawed, need to have your own music teaching objective law of teaching system design theory, the music education should reflect the aesthetic value, creative development value, performance and the social value idea, etc.

With the continuous development of modern Internet technology, networking and informatization have become the mainstream of the current era. As the mainstream of the Times, how to meet the development trend of modern Internet learning and the needs of modern mainstream groups through modern information means and computer aided technology has become the focus of current music teaching thinking. A computer-aided music teaching system is like a piano, a tape recorder, or a blackboard or slide with staff drawn on it, but it is a combination of all these and many others. Wu and Chen [2] mentioned that the educational revolution brought about by the change of educational tools is a long and complicated process. We need to know that it's still people who operate computers. Properly used, the computer aided music teaching system can greatly improve the efficiency and quality of education. Improper use, will only be in the advanced educational tool system on the old teaching content repeated, is only a low-level process of high automation.

Through the effective application of digital information of musical skill training elements in the teaching process, a digital music course auxiliary teaching system is constructed. In music classroom teaching, using the computer multimedia auxiliary teaching, will greatly enrich the teaching technique of expression and representation, so that students in both visual and auditory greatly expanded the space and time, is good for students to receive information from multiple channels, multiple perspectives, the expanded aspect of knowledge, enhance the thirst for knowledge, to cultivate the students' thinking ability, improve the efficiency of classroom teaching, promote the implementation of quality-oriented education.

2 MUSIC FEATURE RECOGNITION METHOD

In music recognition, the extraction of pitch and time feature of each independent note is one of the keys to music signal analysis. It can be said that in music, note pitch and time are the basis of each piece of music, and also the essential difference between a piece of music and other pieces of music; Change the pitch or timing of any note in a piece and the whole piece will be different from the original piece. Similarly, if we can choose a reasonable way to express the result of music pitch and time value, we can accurately judge the merits and demerits of the result of feature extraction module. Therefore, it is very important to extract the pitch and time feature of a sound.

2.1 Fundamental Nature of Music

Compared with text and pictures, music takes up more storage space. If the original music data is used for machine learning, both storage and computing efficiency will be too high, and there are a lot of redundant data in the original data, or even noise data, which will have adverse effects on the results. Through the feature engineering of music data, the data describing music can be

greatly compressed, and then the original music data can be mapped to several dimensions concerned by the system. In this way, redundant data is removed and the magnitude of data is greatly reduced, making music data computable. Through these features, the algorithm of machine learning can realize the classification, clustering, recommendation, analysis and other operations of music data. Through this method, the computer can 'understand' music, and the method can be widely used in the field of music data retrieval and management, so as to reduce a lot of manual management operations, at the same time, the music service field can provide more accurate, humanized service through this method.

Sound has four characteristics: pitch, length, intensity and pitch, which are related to the frequency and duration of the sound body vibration respectively. Physical quantities such as amplitude and spectral distribution correspond to the three basic elements of a piece of music: melody, rhythm and harmony. At the same time, basic music theory includes notes, length, pitch, volume, tone, name, fundamental frequency, beat, rhythm and speed. Sounding instruments can be roughly divided into two categories: pipe chords. Ning et al. [3] mentioned that the sound principle of a tubular instrument is the vibration of air in a pipe, while a string instrument is the vibration of the string itself. In order to analyze the structure of musical form, the method of searching for similar melody is adopted, and the efficiency and accuracy of searching are improved by the three-step recognition method of preliminary recognition, key recognition and supplementary recognition, and the characteristics of rhythm and harmony of musical form are also taken into account.

2.2 Feature Recognition of Music

To describe a specific piece of music accurately, there is a wide range of things involved, including at least rhythm, melody, harmony, timbre and so on. Music recognition method is mainly based on short-term identification methods of frequency domain analysis, the signal of time domain analysis is the analysis and extraction of music signal of time domain parameter, short-time autocorrelation function is mainly used to estimate music notes of base band signal, short-time energy and short-time zero crossing ratio used in the detection of music clips, and short time average magnitude identification for strength of notes. The speech frames of the music feature test template and reference template follow the search path of the DTW algorithm to expand the music feature matching. In real life, people's perception of music is mainly measured by timbre, tone and rhythm. Bakkouri and Afdel [4] mentioned that timbre refers to the feeling of human ear on the sound quality of music. The specific combination of overtones forms a specific timbre, while overtones refer to all frequencies except the lowest frequency in the musical signal; Pitch refers to the human ear's perception of the high or low tone of the music, which is proportional to the high or low frequency of the music signal; Rhythm refers to the regularity of music changes. The more beats per unit time, the stronger the sense of rhythm, and vice versa.

When multiple instruments are playing simultaneously, timbre recognition of multi-timbre sounds is extremely challenging. The results of sound analysis can be useful for automatic indexing and browsing of music data when people are searching for melodies played by specific instruments. When describing the sound produced by a given instrument, a large amount of relevant information can be lost if only a single parameter value is used. Voice tone depends not only on the sound spectrum content, but also depends on the time spectrum changes, after the audio section, model, respectively, in the process of training to train the audio section of each section of each audio as a single sample, it also increased the size of training data. In the prediction, the results of all audio slices of music are aggregated to realize the voting mechanism in the prediction.

Algorithm based on the new model of the study, based on the music score more sound combination of polyphony, told CNN model algorithm for adaptive adjustment, introducing the simulation of the human auditory attention to build a music recognition classification benchmark model, the model of the key is due to the differences in brain structure of the automatic filter, when they hear music Attention will focus on the rhythm of the music and the main voice parts, so

as to identify the music instrument is different, therefore, the model proposed in this paper the modeling method, the simulation of the human brain operation process, combining with the characteristics of the background music instruments and attention models are introduced to key component corresponding to different instrument band set corresponding feature weights. The flow chart of music feature extraction algorithm is shown in Figure 1.

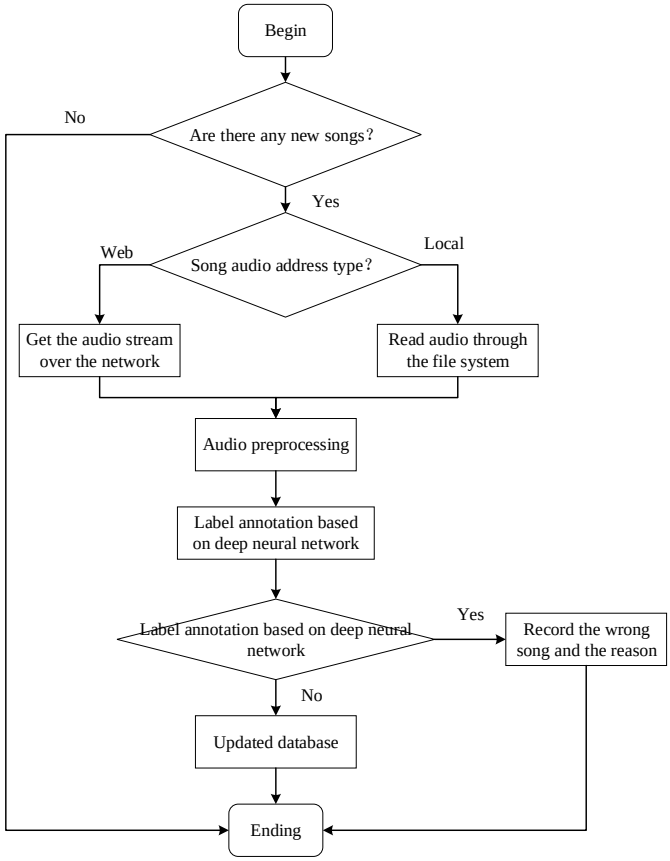

Figure 1: Flow chart of music feature extraction algorithm.

3 DESIGN OF MUSIC TEACHING SYSTEM

3.1 Concepts and Techniques Related to Music Teaching

The traditional teaching mode is classroom teaching, piano as the teaching tool, students under the guidance of teachers in a collective form of solfeggio, listening cognitive learning process. Wang et al. [5] mentioned that teachers cannot take into account the individual differences of students and carry out targeted training. More importantly, if you install special teaching software for the computer, it itself will constitute a very perfect auditory training system.

The learners choose targeted training program, and then launched an interactive teaching by computer unit, presents on multimedia learning content and interpretation of the relevant knowledge and skills, can immediately following the end of unit of learning into the evaluation phase, after completed and passed all the testing subject can selectively into the next unit.

Solfeggio training in this environment is exactly the same, except that the human-machine relationship is transformed from human-listening to computer to human-listening to computer. The whole teaching process of solfeggio and ear training does not need teacher intervention at all, and the learning time and training items are controlled by learners themselves, which completely eliminates the contradiction that classroom teaching cannot take care of the individual differences of learners, so as to truly realize the individuality of education [6].

Computer multimedia assisted teaching is to make use of the integration and interactive characteristics of multimedia, fully use color, sound, animation, pictures and other forms, intuitive, novel, active image and sound display teaching content. This way of information expression is very attractive, allowing students to acquire information through multiple sensory organs and absorb knowledge to the maximum extent. The unique interaction of computer network can realize the two-way transmission of teaching information, improve students' active participation in the teaching process, and provide a good environment for students' enthusiasm and creativity in learning. Through human-computer interaction, the ideal model of individual learning in collectivized teaching situation can be realized [7]. Therefore, it is suitable for the teaching of many courses and has a very wide application prospect in music classroom teaching. The music teaching auxiliary training process is shown in Figure 2.

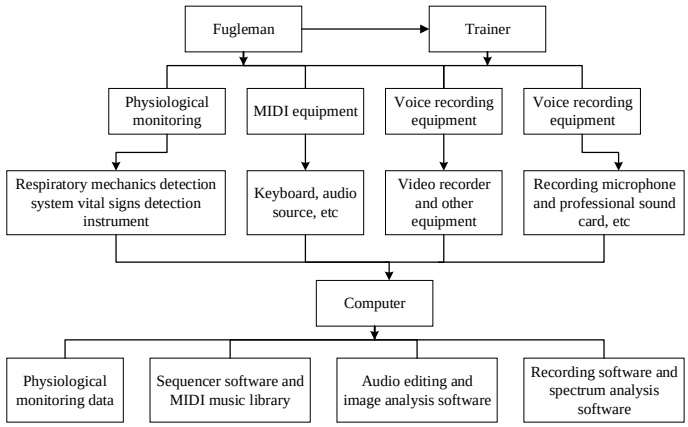

Figure 2: Music teaching auxiliary training process.

3.2 Concepts and Techniques Related to Music Teaching

The design of music teaching system is mainly to use scientific methods to analyze the specific teaching, teaching music appreciation of our homework procedures evaluation results, planning process. It mainly by using the modern education, computer multimedia technology, advanced concepts, scientific research and teaching music appreciation problem and the main purpose is to solve our problems in music teaching, methods and procedures, and the goal of treatment is music teaching, teaching methods, teaching process, teaching habits, interest and the teaching of multimedia editing, Gradually realize our teaching process, teaching effect, teaching resources, teaching effect and simple production.

Music appreciation course is a basic design based on music theory. Therefore, determine the teaching music teaching objectives and analysis stage of the content arrangement, excluding learning needs and learning content of learners, as well as analysis [8].

Music education realizes scientific, accurate, objective, just and fair evaluation. With the reform and development of music education in China, foreign advanced education concepts have

gradually attracted the attention of the education system, and China has gradually learned from the western educational ideas of music teaching, focusing on students' emotions, attitudes and values.

The establishment of a standard music course evaluation system is based on evaluation and implementation, which is to establish the values of the course, the basic concepts of the course as well as the objectives and conditions. As for the goal of music teaching and specific analysis content, mainly through the analysis of the results of teaching conditions and teaching to achieve. Music teaching method is based on the change of teaching conditions, teaching methods, teaching development achievements and teaching diversity.

Music teaching system design theory is mainly determined from the beginning of teaching objectives. Teaching conditions and teaching effect of teaching methods exist, we are mainly on the existing knowledge of music combined to solve the problems in teaching and put forward solutions. In order to reduce the complexity of teaching design, reasonable allocation for different types of music teaching, so as to develop good psychological quality, solve practical problems in the teaching process, improve teaching innovation ability evaluation, hands-on practice and ability, scientific spirit, positive mood [9].

Music teaching system is a complete, complementary existing teaching method, but also an inevitable trend of information and network development. The system is mainly used to assist the construction of music knowledge map, which can provide basic services for intelligent music search and recommendation. A large amount of knowledge in the music knowledge map can help the upstream system capture users' intentions more accurately and realize more personalized recommendation services. The construction of music knowledge graph needs the support of a large amount of music domain information [10]. In addition to the extraction of common music entities and relationships, music label information is also an important part of music knowledge graph. Music is mainly the carrier of audio. Compared with other music labels, extracting music labels directly from audio is more direct and closer to the original expression of music. Therefore, audio-based label labeling system is more valuable to the construction of music knowledge graph. The music assisted teaching main process is shown in Figure 3.

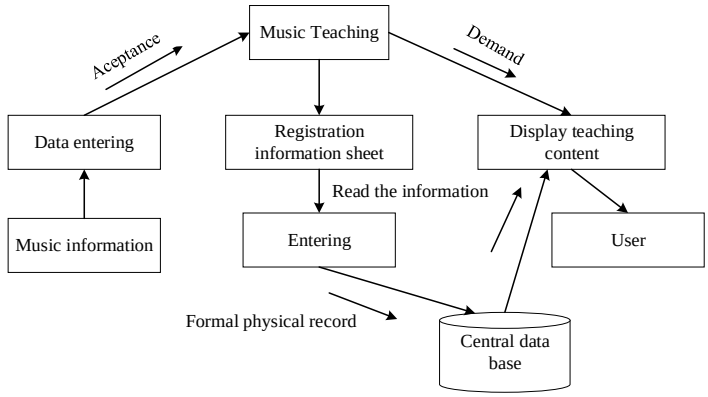

Figure 3: Music assisted teaching main process.

4 SYSTEM TEST AND RESULT ANALYSIS

This paper will not consider the balance of the data set, first of all, the instrument data is divided into wind instruments and stringed instruments, among which stringed instruments are divided into stringed instruments and plucked instruments. In the classification process, one-to-many or

one-to-one classification method can be adopted. One-to-many means that you treat one instrument as one type of data, the other five instruments as another type of data, and then you take one of the remaining five types of data, and the other four types of data as one type, and so on, until the data is completely separated. One-to-one means that the data is first divided into two categories based on certain characteristics, and then further subdivided in each category until it is separated. The clustering analysis of different model parameters is shown in the Figure 4. As can be seen from the figure, the distribution of labels of different categories is different, and the proportion of positive and negative samples of different labels is different. The proportion of negative samples will greatly exceed the proportion of positive samples. If all test samples are marked as negative in the model, high accuracy can be achieved, but this evaluation method is not suitable.

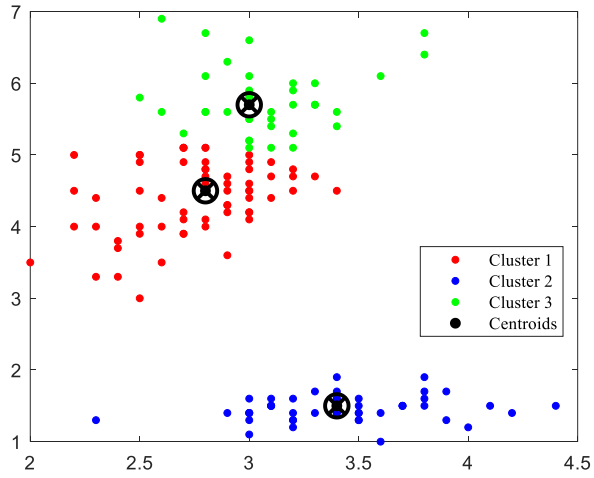

Figure 4: Cluster analysis of different model parameters.

In order to observe the characteristics of output of each layer, the output of each layer of test data is projected onto the 2-dimensional plane by linear dimensionality reduction method. The Figure 5 shows the increment of accuracy of different music data enhancement methods. It can be seen from the figure that Mendelssohn sample, Chopin sample and Handel sample are basically separated from other types, so the accuracy of these three types is relatively high. The support vector machine model is used to classify the output of each layer, and it is found that the accuracy of classification is the highest in the last layer, which is higher than the accuracy of the original data. It is proved that the proposed music enhancement method is effective for music classification task, and can provide more abundant music modes for the data set while enlarging the scale of the data set, which is beneficial to network training and improves the music classification performance of the model.

The Figure 6 shows the Schematic diagram of audio segmentation process. The section length selected in this paper is 5 seconds, and the overlap rate is 50%. The last part of the sample that is less than the section length is discarded. Using overlapping way audio segmentation mainly for the following reasons: first, the audio signal has a certain continuity, shard, without overlapping segmentation points on both sides of the audio signal is a physical partition may cause a certain loss of information, this information includes the music signal continuous information and turn the sound spectrum window.

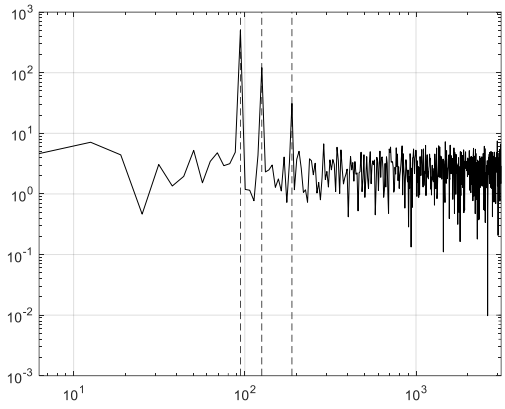

Figure 5: Incremental accuracy of different music data enhancement methods.

Secondly, the overlapping method can further increase the number of slices when the length of slices is constant. Finally, the overlapped parts appear in different positions in different samples, which can play an enhancement effect similar to the translation mode in image data enhancement. It shows that the attention mechanism plays a positive role in the network learning of music features, and the sequence features of each moment are given the corresponding attention weight, so that the sequence information can be integrated more effectively and the overall feature of music can be better represented.

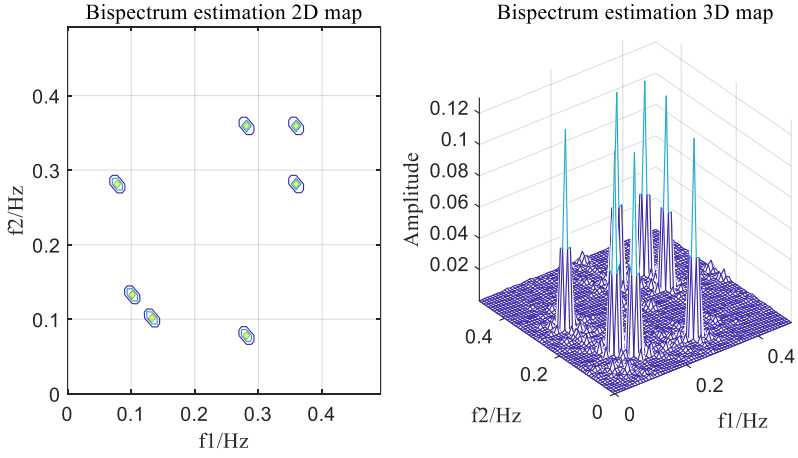

Figure 6: Schematic diagram of audio segmentation process.

The Figure 7 shows the accurate comparison of classification of overlapping rates of different audio segments. As can be seen from the figure, the accuracy of using the overlapping method is improved compared with that of not using the overlapping method in the three slicing periods. This is on the one hand because the overlapping method can retain the continuous information on both sides of the slicing point, and on the other hand because the overlapping method has the data enhancement effect of translating the audio signal. In addition, it can be seen from the results that blindly increasing the overlap rate cannot improve the classification performance of the model.

This is because the model oversamples some audio signals during training due to the large proportion of repeated data of different samples, which affects the final classification effect. At the same time, it also shows that the structure not only has good performance in music genre types, but also applies to a wider range of music labels, has good generalization ability, and has a stronger ability to extract music features than other convolutional structures. In addition, this paper finds that after further feature extraction by deep confidence network, the recognition accuracy of basically all music genres is improved, which further illustrates the superiority of deep confidence network in music genre recognition and classification.

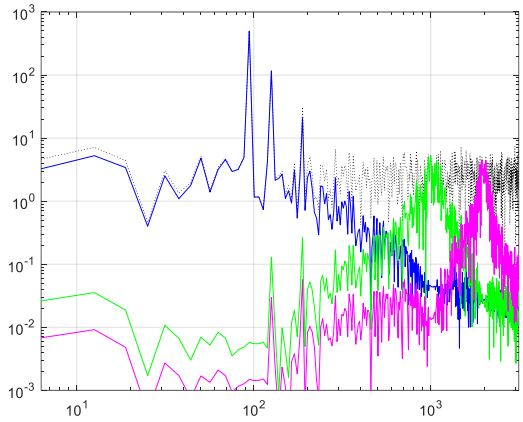

Figure 7: Accurate comparison of classification of overlapping rates of different audio segments.

5 CONCLUSION

With the rapid development of digital multimedia technology, the number of digital music on the Internet is increasing rapidly. How to manage these massive music resources is a thorny problem that each major music media platform faces. The development of music classification technology is of great significance to the research and application of efficient organization, retrieval and recommendation of music resources.

Music annotation is difficult, and limited annotation data is not conducive to the training of deep learning model. Therefore, this paper adopts a variety of music data enhancement methods to enhance music data combined with the characteristics of music signals.

According to the characteristics of sound spectrum, a music classification model based on convolutional neural network is proposed. The attention mechanism is used to assign different attention weights to the output of cyclic neural network at different times, so as to get a better representation of the overall characteristics of music.

Finally, this paper designs and implements an audio-based music label labeling system based on the proposed music classification method, which realizes the labeling of music genre, emotion and scene music labels, and provides data support for the construction of music domain knowledge.

This paper can also attempt to conduct in-depth research from the following aspects:

(1) Try to use multi-modal data as input. In this paper, only audio is used as the data input for classification. In addition to audio, lyrics, MV and other data also contain rich information that can be used for classification. Multi-modal music data can be used to master the classification characteristics of music more comprehensively.

(2) Further analysis of features extracted by convolution. Through deconvolution operation, the sound spectrum features extracted by each layer of convolution can be visualized. The results of visualization help us better understand the extraction process of sound spectrum features by convolution, which is conducive to the design of convolution structure.

(3) Try to learn the original audio signals of music directly. Although sound spectrum is adopted as a unified representation of music in this paper, the problem of manual selection of features is avoided, but in the process of converting sound spectrum, music information may be lost. It can be considered that the original audio signal of music can be used as the input of network to extract music features, which may help to improve the classification performance.

Ning Fu, https://orcid.org/0000-0002-7930-1329
Xia Peng, https://orcid.org/0000-0003-3515-6542

REFERENCES

[1] Zubair, A.-F.; Abu Mansor, M.-S.: Auto-recognition and part model complexity quantification of regular-freeform revolved surfaces through delta volume generations, Engineering with Computer, 36(2), 2020, 511-526. https://doi.org/10.1007/s00366-019-00710-7
[2] Wu, J.-R.; Chen, B.-D.: English vocabulary online teaching based on machine learning recognition and target visual detection, Journal of Intelligent & Fuzzy Systems, 39(2), 2020, 1745-1756. https://doi.org/10.3233/jifs-179948
[3] Ning, F.-W.; Shi, Y.; Cai, M.-.L.; Xu, W.-Q.: Various realization methods of machine-part classification based on deep learning, Journal of Intelligent Manufacturing, 31(8), 2020, 2019-2032. https://doi.org/10.1007/s10845-020-01550-9
[4] Bakkouri, I.; Afdel, K.: Computer-aided diagnosis (CAD) system based on multi-layer feature fusion network for skin lesion recognition in dermoscopy images, Multimedia Tools and Applications, 79(29), 2020, 20483-20518. https://doi.org/10.1007/s11042-019-07988-1
[5] Wang, Z.-Y.; Xin, J.-C.; Wang, Z.-Q.; Qian, W.: Computer-aided dementia diagnosis based on hierarchical extreme learning machine, Cognitive Computation, 13(1), 2020, 34-48. https://doi.org/10.1007/s12559-019-09708-1
[6] Prichard, S.: A mixed-methods investigation of preservice music teaching efficacy beliefs and commitment to music teaching, Journal of Research in Music Education, 65(2), 2017, 237-257. https://doi.org/10.1177/0022429417710387
[7] Shin, J.: Becoming a music teacher: Korean music student teachers? Perceptions of their student teaching experience, Music Education Research, 21(5), 2019, 549-559. https://doi.org/10.1080/14613808.2019.1673328
[8] Han, R.-C.; Leung, B.-W.: A survey on Weifang teachers' attitudes toward teaching Chinese folk music, International Journal of Music Education, 35(2), 2017, 202-215. https://doi.org/10.1177/0255761415620521
[9] Yancey, N.-R.: Music and teaching-learning in nursing: harmony and rhythm in co-constructing new knowings, Nursing Science Quarterly, 33(1), 2020, 21-26. https://doi.org/10.1177/0894318419881802
[10] Shuo, W.; Ming, M.: Exploring online intelligent teaching method with machine learning and SVM algorithm, Neural Computing & Applications, 34(4), 2022, 2583-2596. https://doi.org/10.1007/s00521-021-05846-6