

## Research Article

# Design and Application of the Piano Teaching System Integrating Videos and Images

**Lina Li** 

*Suzhou University Academy of Music, Anhui, Suzhou 234000, China*

Correspondence should be addressed to Lina Li; [lilinasj@ahszu.edu.cn](mailto:lilinasj@ahszu.edu.cn)

Received 18 May 2022; Revised 11 June 2022; Accepted 15 June 2022; Published 27 June 2022

Academic Editor: Lianhui Li

Copyright © 2022 Lina Li. 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.

The art of piano playing has been continuously entering into people's life. However, with the continuous improvement of science and technology and living standards, the traditional teaching mode can no longer meet the piano teaching mode. The teaching of piano is different from traditional subjects, such as Chinese and mathematics. It requires students to experience the artistic characteristics and the live atmosphere of the players brought by the piano. This study integrates video and image teaching methods with piano teaching. Videos and images can more intuitively show the live atmosphere brought by piano players and musical artistic features brought by the piano. At the same time, this study uses the convolutional neural network (CNN) method to study the relevant features of videos and images of piano teaching. These features are mainly the characteristics of piano music, the behavior of players, and the basic knowledge of a piano. The research results show that the clustering method can effectively classify the features of videos and images in piano teaching, and the maximum classification error is only 1.89%. The CNN method also has high performance in predicting the relevant features of piano teaching videos and images. *Accuracy*. The largest prediction error is only 2.23%, and the linear correlation coefficient also exceeds 0.95. This set of the piano teaching mode that combines videos and images is beneficial to both teachers and students.

## 1. Introduction

The traditional teaching method is to use the blackboard or PPT to impart knowledge. With the continuous progress of computer-aided teaching methods and hardware equipment, the teaching mode has also undergone great changes [1]. Computers can help people with some tedious tasks. It can also display teaching content in the form of pictures or flow charts. It more intuitively shows teaching knowledge to the students. This is different from the traditional teaching mode. The piano teaching mode is different from the traditional methods of mathematics and Chinese subjects. Mathematics and Chinese teaching modes are knowledge and methods that need to be taught in textbooks. However, a piano is a teaching mode that needs to be understood [2, 3]. Piano art not only embodies the meaning of music itself but also embodies the humanistic feelings of the connotation of music. Therefore, the piano teaching mode is not a simple superposition of knowledge, it also requires students to

experience the connotation of piano art. This requires the teaching mode of piano art to be improved. Traditional teaching methods can only allow students to learn the basics of piano, such as musical notation, musical notes, etc. It does not allow students to appreciate the humanities in it [4, 5]. The PPT method can allow students to learn piano knowledge in the form of pictures, but this method is far from enough. A piano is an artistic method that combines musical instruments and music, and it is also closely related to the emotional performance of the audience. The piano teaching classroom also requires students to combine the music scene and the form of the piano performance to carry out effective learning. The video method can record the ambient atmosphere, music atmosphere, and the content of the piano performance at the music scene, and it can effectively record the performance of the audience and the performer. Piano learners can learn more knowledge from piano-playing videos [6, 7]. The image can effectively record the key piano performance, it can capture the mood and

performance content of the piano player very well. However, the content of videos and images needs to be grasped and captured efficiently by the pianist, which requires big data technology to capture the key content of videos and images and relevant knowledge. This study intends to design a piano teaching mode by combining videos and images, which will better capture the content of the piano performance and the atmosphere of the scene. The research on the relationship between the content of videos and images and the content of the piano performance requires big data technology to analyze.

A convolutional neural network (CNN) can better extract the spatial features of data and it can also better map the relationship between input and output [8]. It has been widely used in many fields, such as image recognition and feature extraction. CNN can have higher efficiency and computing power than manual methods, so it can handle the relationship between large amounts of data very well. Video methods will contain many image features, and CNN can have better feature extraction capabilities in both videos and images [9, 10]. This research integrates videos and images to design a piano teaching mode, which mainly designs feature extraction. It also does not involve temporal features, so the CNN method was chosen in this study. The CNN method is also an intelligent algorithm that has developed rapidly in recent years, and many research objects have complex mapping relationships. However, there is relatively strong nonlinear correlation between these complex data. It is difficult for manual methods to deal with these problems with experience or professional knowledge. CNN methods can use these complex data to find correlations between data. At present, big data technology has produced algorithms related to spatial feature extraction or temporal feature extraction [11, 12]. There are relatively few temporal features involved in piano teaching videos and images recognition. Therefore, this study only selects the spatial features of piano teaching videos and images. CNN has relatively high requirements on computers and hardware devices. With the continuous advancement of science and technology, there will be huge amounts of data in every field. The larger the magnitude of the data, which requires deeper CNN to complete the feature extraction task [13, 14]. The birth of GPU technology makes it possible for CNN to develop to a deeper level. The computing power of GPU technology will be many orders of magnitude superior to the computing power of CPU. The piano video technology has more image features, and the GPU technology provides more technical support for the feature extraction of the video method.

This research will solve the shortcomings of the traditional piano teaching mode, it uses video and image methods to design a new teaching method for piano teaching. At the same time, CNN technology is used to extract spatial features existing in piano video and image teaching, and it can also be used to complete the extraction of nonlinear relationships. This set of an intelligent piano video and image teaching mode will not only effectively extract the piano player's behavior information and music and art content embodied by the piano but also improve students' learning

efficiency and interest. Compared with the teaching mode of the blackboard or PPT, the teaching mode of videos and images will attract more students' attention. When students' learning attention and learning interest are improved, piano learners will gain more piano teaching knowledge.

In this study, a new piano teaching mode was designed using the form of videos and images. At the same time, this study uses the CNN method to extract the relevant features of piano video and image teaching. This study will be introduced from five aspects. Section 1 introduces and analyzes the defects of the piano teaching mode and the research significance of the CNN method in piano teaching. The research status of the mode of piano teaching is introduced in Section 2. The piano teaching mode and the CNN method combining videos and images are introduced in Section 3. Section 4 analyzes the feasibility and accuracy of CNN for feature extraction of videos and images in piano teaching mode, which is the core part of this research. Section 5 summarizes the full text.

## 2. Related Work

The piano is an art form that can reflect the art of music and humanistic feelings, and it can also integrate the audience and the performer into an art form. There is a big difference between piano teaching and the teaching mode of traditional subjects. Many researchers have conducted a lot of research on the teaching mode of a piano. Cui [15] uses augmented virtual reality technology AR to study the relevant skills of piano teaching. This study mainly proposes an online piano teaching mode. It selected a number of college students in Heilongjiang to study the relevant situation of the piano, which mainly includes the learners' subjective attitudes and personal learning progress. The results of the study found that 89% of students had a significant improvement in the learning of musical terminology, which was mainly in reading sheet music and using music materials. This online piano teaching mode provides a new teaching mode and development direction for piano teaching. Pi [16] believed that piano education can alleviate teaching fatigue and improve the teachers' happiness index. This study discusses the relationship between work fatigue and piano education. It mainly uses qualitative and quantitative methods to study the relationship between piano education and the teachers' happiness index. The results of the study show that piano teaching can improve work motivation and efficiency. At the same time, it found that the happiness index of most teachers in piano teaching is at a low level, and this research will help to improve the happiness index of piano teachers. Liu [17] has found that the current piano teaching mode lacks comprehensiveness and scientificity, and it has been unable to use the development of the piano teaching mode. This study uses BP neural network technology to establish a mode of the music signal and the piano teaching performance score. It selects well-known piano works as the test set to verify the effect of the mode. The research results show that this method can effectively verify the working level of the piano. It can accurately

provide players with a certain scoring reference. This method can not only improve the musical level of piano performance but it can also help to improve the talent of a piano. Xue and Jia [18] found that information technology has brought about tremendous changes in the development of all areas of life. However, the teaching and learning of music and musical instruments is a huge challenge for teachers and students in remote areas, and this artistic discipline is extremely disadvantageous to teachers and students in remote areas. This study uses the multi-information classification (MSC) algorithm of artificial intelligence technology to study the piano teaching mode in remote areas. This algorithm will be spread through the wireless network, it can effectively spread the teaching knowledge of a piano. This teaching mode can efficiently classify the piano information data. The research results show that this piano teaching mode is beneficial to teachers and students in remote areas. Li [19] found that deep learning technology and artificial intelligence technology will provide more technical support for the improvement of modern piano teaching quality. It studies note detection methods for piano teaching using convolutional neural networks. The input to the CNN neural network is the music signal for piano teaching. The results of the study show that the intelligent piano teaching method can improve students' interest in learning piano. This method will improve children's interest and efficiency in learning piano. Huang and Ding [20] proposed a back-propagation neural network piano teaching evaluation system. This can solve the problem that traditional note recognition methods are easily affected by noise. It used the optimized BPNN algorithm to accurately measure the pitch of the note and the time value of the note. The research results show that this mode can effectively correct the pitch problem in the piano teaching process, which is 5.21% higher than the traditional method. The optimized BPNN algorithm can significantly improve the error correction accuracy of the player's note and pitch, which is beneficial to improve the teaching quality of piano teaching. Guo et al. [21] used wireless network technology to realize the intelligent piano teaching mode. It uses the regression fitting algorithm and the Relief F weight algorithm to extract the characteristics of the piano teaching process. The results of the study found that the use of intelligent algorithms can quantitatively analyze the relevant characteristics of the piano teaching process. This is conducive to the reform of the piano teaching mode. Through the above literature review, it can be found that artificial intelligence technology has been applied in the process of piano teaching, and it is mainly used to identify signal features, such as notes. This study mainly uses CNN to identify the features in the teaching process of piano videos and images. This is an innovative study. Most of the studies mainly use neural network methods to study the timbre and note characteristics of piano teaching. However, this study utilizes the CNN method to perform feature extraction on videos and images of piano teaching.

### 3. Piano Video and Image Teaching Program, Design, and Algorithm Introduction

*3.1. The Significance of the CNN Method for Piano Teaching.* In order to solve the shortcomings of the traditional piano teaching mode, this research introduces video and image methods into the piano teaching classroom, which realizes a new piano teaching scheme. The CNN method will assist in the recognition of video and image features. There are huge data features in the videos and images in the piano teaching plan, and these features have relatively large correlation with the musical features of the piano and the behavioral features of players. It is difficult to teach these characteristics only by relying on teachers' professional knowledge and teaching experience. CNN methods can efficiently and accurately identify and extract these features. Then, the CNN method can also realize the mapping of piano teaching features to video and image information. These relevant information can be transmitted to piano learners through computer-aided systems. If the CNN method is not used to extract piano-related features in the piano teaching classroom plan, this will limit the display of piano features. Some music or player characteristics cannot be visually displayed through video and image methods. In short, the CNN method is more important for the piano teaching system, especially it involves the relevant information of the videos and images of the piano.

*3.2. The Design and CNN of Piano Teaching Scheme Integrating Videos and Images.* The goal of this research is to integrate video and image methods into a piano teaching method, and then it utilizes a CNN method to identify the relevant features of piano videos and images. The identified features will be visually displayed to the students or teachers through the computer-aided system. This will improve the learning interest and learning efficiency of piano learners. Figure 1 shows the piano teaching scheme and workflow that integrates videos and images. First, the features of the piano videos and images are processed into data between 0 and 255. These data will then be normalized to be between 0 and 1. It needs to feed piano-related video and image data into the CNN algorithm in the form of an input layer. However, these data need to be classified by a classification algorithm before being input into CNN, which is beneficial to improve the accuracy of prediction. The classification algorithm will perform effective classification processing according to set classification criteria. This research needs to map the feature relationships between piano videos, images and piano music features, player behavior features, and piano connotations. After these three features are processed, it will be visually displayed to teachers and students through a computer-aided system. Although this piano teaching method seems to be more complicated. However, once this teaching scheme is trained, it only takes a few seconds to achieve the mapping of relevant features. In an actual piano teaching session, this only takes a few seconds. This is also an efficient and accurate

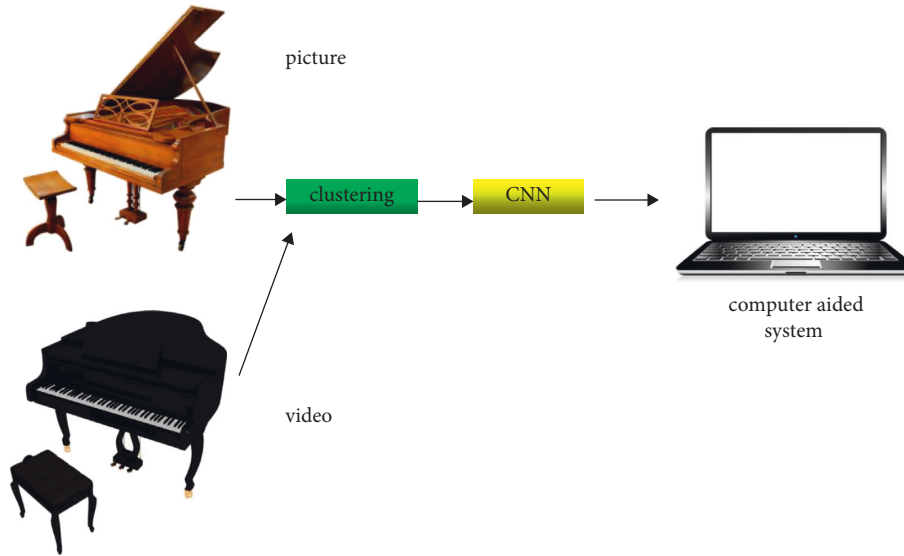


FIGURE 1: System design of piano teaching scheme integrating videos and images.

way from a time perspective. The computer-aided system will display the piano teaching features extracted by CNN to the students or teachers. Videos or images of piano teaching are also displayed to students or teachers through computer-aided systems.

The CNN method has great advantages in feature extraction and data mapping. It is also good at processing huge amounts of data and features. There is also a huge amount of data and associated features in the videos and images in piano teaching classes. These advantages of the CNN method are precisely to deal with the characteristics in piano teaching. Figure 2 shows the workflow of the CNN method. The CNN method has certain similarities with the fully connected neural network, which also utilizes the mechanism of forward propagation and back propagation. At the same time, its gradient descent is also carried out by means of derivation. However, it is more efficient than the fully connected neural network because it has a certain weight sharing mechanism. The weights of each layer are not connected to each other, it is selectively connected, which reduces the amount of parameter calculation. CNN can generally extract features with strong correlations. It will selectively filter features with weak correlation. CNN generally consists of multiple layers of convolutional layers, pooling layers, and activation functions. The process of feature filtering in CNN is carried out through the filters and strides of the convolutional layers. The output layer of CNN will perform error operation with the label data of piano teaching. The gradient descent method will calculate derivation based on the error between the predicted value and the actual value of the piano teaching. The step size used in this study is 1, and the number of filters is set to 64, which is a relatively common numerical range. Meanwhile, in order to fully exploit the features of piano videos and images, the learning rate is set to 0.0001.

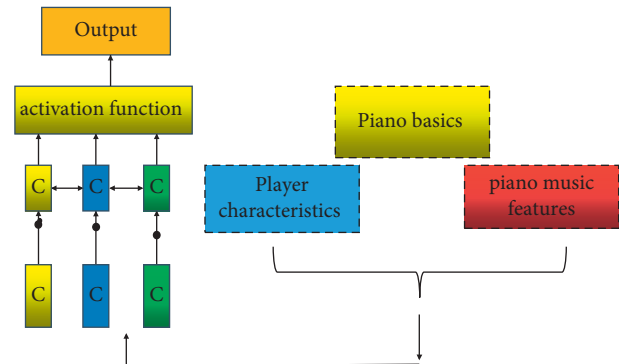


FIGURE 2: The workflow of the CNN method.

The difference between CNN and the fully connected neural network mainly reflects the existence of more hyperparameters, which are the source of feature selection and feature filtering. Different hyperparameter combinations will affect the accuracy and convergence of calculation results. Equation (1) shows the computational relationship that is satisfied between the hyperparameters.  $s$  represents the step size of feature selection.  $p$  represents the padding step size of the matrix.  $k$  represents the number of CNN filters.

$$w' = \frac{(w + 2p - k)}{s} + 1. \quad (1)$$

CNN is similar to other neural networks. It also has forward propagation mechanism and back propagation mechanism. These weights and biases are derived using automatic differentiation techniques. Equations (2) and (3) show the unfolded shape of weight derivative calculation at each layer. This is also the expanded form of loss function calculation.

$$E = \frac{1}{2} \sum_{k=1}^m [d_k - f(\text{net}w_k)]^2 \quad (2)$$

$$= \frac{1}{2} \sum_{k=1}^m \left[ d_k - f \left( \sum_{j=0}^n \omega_{jk} y_j \right) \right]^2,$$

$$E = \frac{1}{2} \sum_{k=1}^m [d_k - f(\text{net}w_k)]^2 \quad (3)$$

$$= \frac{1}{2} \sum_{k=1}^m \left[ d_k - f \left( \sum_{j=0}^n \omega_{jk} y_j \right) \right]^2$$

$$= \frac{1}{2} \sum_{k=1}^m \left[ d_k - f \left[ \left( \sum_{j=0}^n \omega_{jk} f \left( \sum_{i=0}^q u_{ij} \chi_i \right) \right) \right] \right]^2.$$

The derivation operation is a method to find optimal weights and optimal biases. Equations (4) and (5) show the calculation methods for the derivation of weights and biases. It can be seen that the gradient descent method is used here. Equation (6) shows the computation between each convolutional layer.

$$\Delta u_{ij} = -\eta \frac{\partial E \partial}{\partial u_{ij}}, \quad (4)$$

$$\Delta \omega_{ji} = -\eta \frac{\partial E}{\partial \omega_{ji}}, \quad (5)$$

$$\frac{\partial E}{\partial k_{ij}^{\zeta}} = \sum_{u,v} (\delta_j^{\zeta})_{uv} (p_i^{\zeta-1})_{uv}. \quad (6)$$

**3.3. Introduction to Piano Feature Data Classification Algorithm.** The video and image data in piano teaching differ greatly in magnitude and numerical size. If these data are fed into the CNN network layer together, this will lead to the problem of uneven weight distribution. The uneven distribution of weights leads to large errors in the results. Therefore, before feeding the video and image data of piano teaching into CNN, this data needs to be classified. The purpose of classification processing is to effectively classify different features. This allows the same type of features to have the same data distribution. Figure 3 shows the computational flow of the classification algorithm. It can be seen from Figure 3 that the classification algorithm can classify data with the same distribution characteristics together, and it can also separate different characteristics. In this study, the classification method of clustering was selected, which grouped the data of the same category into one category and processed different data separately. This classification method is processed according to the distance of data features.

When processing the data of piano teaching videos and images, the clustering method is mainly based on the distance of the data. Equation (7) shows the expression for the Euclidean distance, one of the commonly used distance

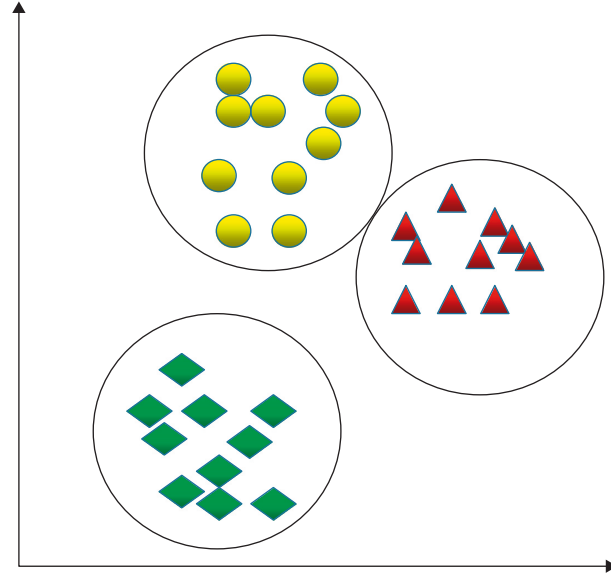


FIGURE 3: Clustering methods classification methods and principles.

measurements. Euclidean distance measures the distance between two points.

$$\text{dict}_{ed} = \sum_{k=1}^m (x_{ik} - x_{jk})^2. \quad (7)$$

Equation (8) shows the Chebyshev distance measurement method. It measures the difference in distance between the coordinate values of points in space. Equation (9) shows the Minkowski's distance measurement method. This is also a variant of the Euclidean and Chebyshev distances. There is a  $p$ -value here, and when  $p$  takes different values, it represents a different distance measurement method.

$$\text{dict}_{cd} = \lim_{t \rightarrow \infty} \left( \sum_{k=1}^m |x_{jk} - y_{ik}|^t \right)^{1/t}, \quad (8)$$

$$\text{dict}_{\min d} = \sqrt[p]{\sum_{k=1}^m |x_{ik} - y_{jk}|^p}. \quad (9)$$

Equation (10) shows the evaluation index of the external performance of classification. Among them,  $a$ ,  $b$ ,  $c$ ,  $d$  represent different features of piano instruction videos and images. Equation (11) shows the Rand statistic, where  $P$  is the precision and  $R$  is the recall.

$$R = \frac{a + d}{a + b + c + d} \quad (10)$$

$$F = \frac{(\beta^2 + 1)PR}{\beta^2 P + R}. \quad (11)$$

## 4. Result Analysis and Discussion

**4.1. Classification and Prediction Error Analysis of Piano Videos and Images.** The goal of this research is to use the

classification algorithm and the CNN algorithm to study the video and image features in the piano teaching process. In this study, video and image information of piano art courses in many colleges and universities in Hangzhou was selected as the research dataset. It will be divided into training set and test set. The test set is also a dataset derived from a college piano art course. This will ensure the reliability of algorithm verification.

In the piano teaching system integrating videos and images, the first step is to use the clustering algorithm to classify the related features of videos and images. The classification accuracy will affect the prediction accuracy of the CNN algorithm for videos and images in piano teaching. Figure 4 shows the classification errors of three features for videos and images of pianos. It can be seen from Figure 4 that the classification errors of the three related features of the piano are all within an acceptable range, and all the classification errors are within 2%. This is acceptable enough for teaching content with videos and images of the piano. The largest classification error is only 1.89%, and this part of the error mainly comes from the classification of the characteristics of piano players. The smallest classification error is only 1.23%, and this part of the error mainly comes from the classification of piano music features. This is mainly because there is a relatively large mutation in the characteristics of piano players. For the classification of the piano music feature method, this part of the error is only 1.52%. The classification errors of these three main features are all within 2%. This is a reliable error for both teachers and students of piano teaching.

After the three features of the piano videos and images are effectively classified, the CNN algorithm is required to predict the three features. This is a critical step for the piano teaching mode that combines videos and images. The video and image features of the piano predicted by CNN will be intuitively displayed to teachers and piano learners. The accuracy of the CNN algorithm in the piano teaching mode is also the key to the success of the fusion video and image piano teaching system. Figure 5 shows the prediction errors for three video and image features in the piano teaching mode. Overall, the CNN method has high feasibility and accuracy in predicting the characteristics of piano videos and images. This has high credibility for both teachers and students of piano lessons. All forecast errors are within 2.5%. The largest prediction error is only 2.23%, which mainly comes from the prediction of the characteristics of piano players. The characteristics of the pianist have a great relationship with the scene of the piano performance, and there is a great mutation in this part of characteristics. It is not just about the basics of the piano itself, so this part of the error is the biggest. Although this part of the error is the largest, it is also within a reasonable and acceptable range. The smallest error is also from the characteristics of piano music, and this part of the error is only 1.54%. The characteristics of piano music are closely related to the notes and spectrum of the piano. However, the mutation of this part of the characteristics is relatively small. This is because knowledge of the musical aspects of the piano is also less abrupt. The prediction error of CNN mainly comes from the

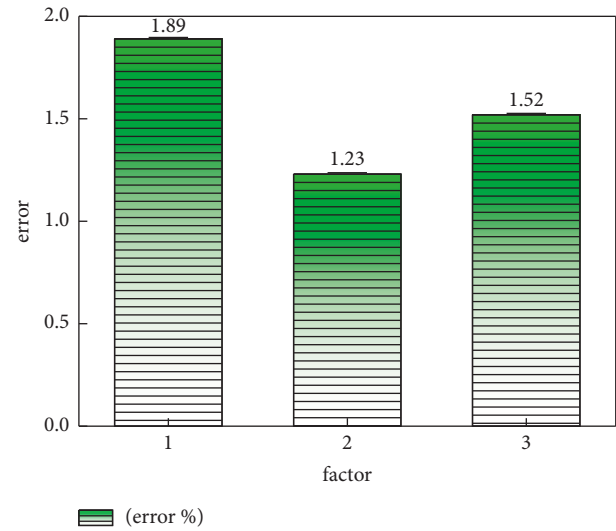


FIGURE 4: The classification errors of piano videos and images using clustering methods.

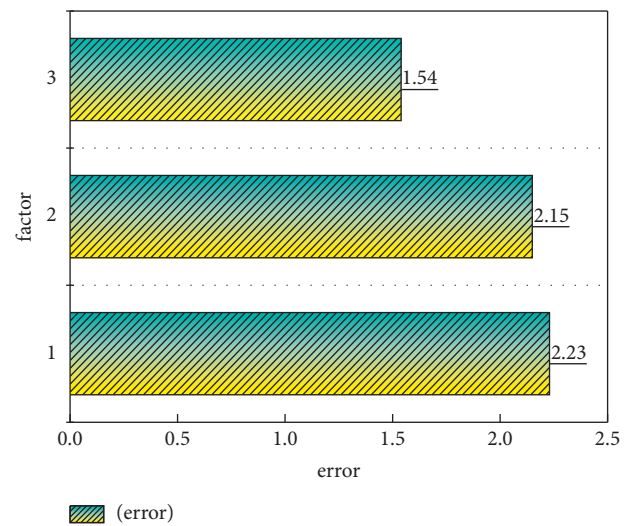


FIGURE 5: Three feature prediction errors of piano teaching mode using the CNN method.

error of the model and the error of the data. The reason for the relatively large error in this part is also that there is a certain cumulative error in the data when using the clustering method. In general, the CNN method can better complete the video and image prediction tasks in the piano teaching course.

*4.2. Analysis of Three Characteristics of Piano Teaching Videos and Images.* In this section, we discuss the piano music features, player behavior features, and basic piano features involved in videos and images in piano teaching. We selected 30 different sets of data to verify the accuracy and feasibility of CNN in predicting three features of piano teaching fusing videos and images. Figure 6 shows the distribution of predicted and actual values of musical features for piano teaching. In Figure 6, the green area represents the error of

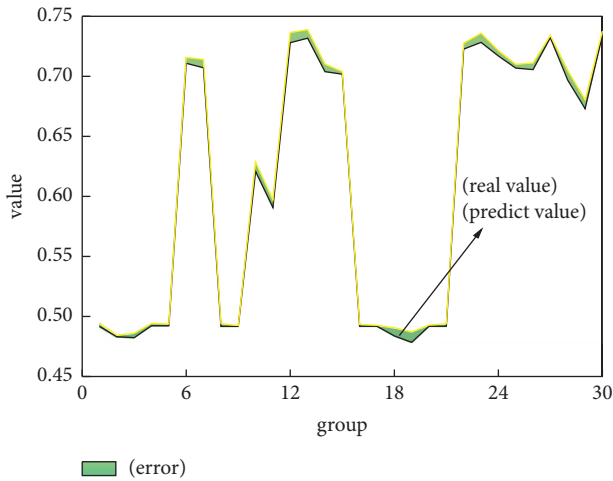


FIGURE 6: The predictive distribution of musical features in piano teaching.

the predicted value of music features in piano teaching. The black lines represent the predicted values of the musical features of piano teaching. The yellow lines represent actual piano music characteristic values. In general, CNN can better capture the peaks and trends between different sets of data. Although there are many peaks and troughs in different groups of piano music characteristics, CNN can still capture the characteristics of these data as well. The green area represents the error between the predicted value and the actual value. It can be seen that the distribution of prediction errors for the musical characteristics of the piano is relatively small, and these prediction errors are relatively small. This further illustrates the feasibility and accuracy of CNN in predicting the characteristics of piano teaching music.

The linear correlation coefficient can further demonstrate the accuracy of CNN in predicting the characteristics of piano teaching music. If the linear correlation coefficient is closer to 1, it means that CNN has good performance in predicting the musical features of piano teaching. If the linear correlation coefficient is closer to both sides of the function  $y = x$ , it means that the linear correlation coefficient is closer to 1. Figure 7 shows the linear correlation coefficient distribution of the musical features of the piano teaching mode fused with videos and images. In general, all data points are relatively close to the linear function  $y = x$ , which means that the linear correlation coefficient exceeds 0.95. It can also indicate that the predicted value of the musical characteristics of piano teaching is relatively close to the actual value. In other words, the CNN method has been shown to be accurate in predicting the musical features of piano teaching fusing videos and images. This is an algorithm and the teaching system that students or teachers can trust.

The player's behavioral characteristics are also an important indicator in piano teaching. The player's behavioral characteristics can only be shown in the form of videos. It is difficult for traditional teaching models to capture the behavioral characteristics of performers. This shows that the piano teaching mode integrating videos and images designed

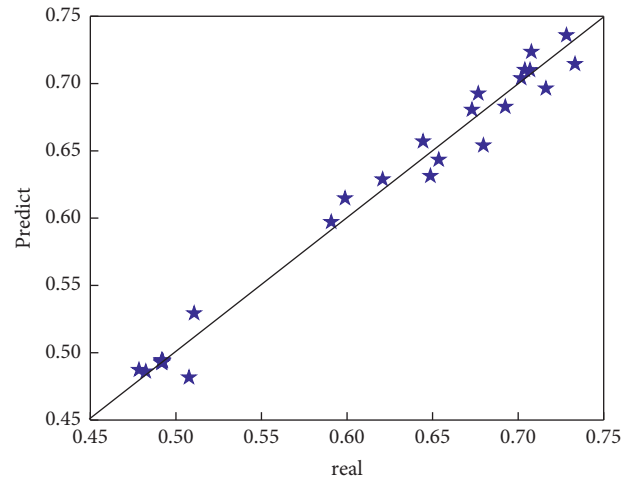


FIGURE 7: The correlation distribution of music feature prediction in piano teaching.

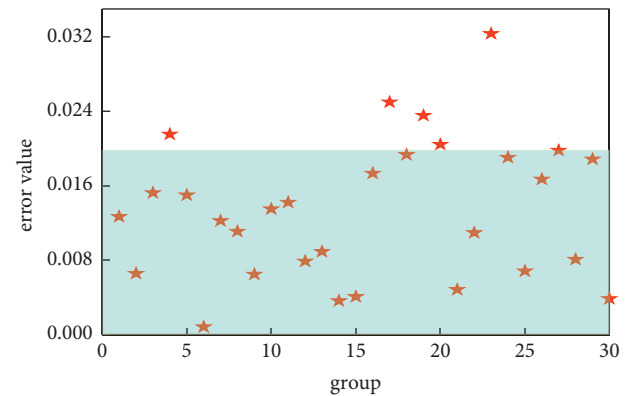


FIGURE 8: Scatter plot of player characteristics distribution in piano teaching.

in this study is innovative to a certain extent. Figure 8 shows the distribution of the prediction errors of the player's behavioral characteristics in the piano teaching mode. In Figure 8, the blue area represents data with prediction errors within 2%, a range where prediction effects can be distinguished. Overall, CNNs also have high reliability in predicting player behavior. Although the player's behavior information in the piano teaching mode has great volatility, it also has a great correlation with the piano performance scene. However, CNN has high reliability in predicting player behavior information for piano teaching. Most forecast errors are distributed within 2%. Only a small number of prediction errors exceed 3%. Basic knowledge of piano is also an important part of piano teaching. Figure 9 shows the distribution box plot of predicted and actual values for the basic knowledge of piano teaching. For the box plot, if the size of the box plot of the predicted value and the distribution of values are consistent with actual values, this indicates that the mode has a strong predictive ability. In general, the predicted value is basically the same as the actual value of the box, whether it is the size of the box or the distribution of the data. Basic knowledge of piano is

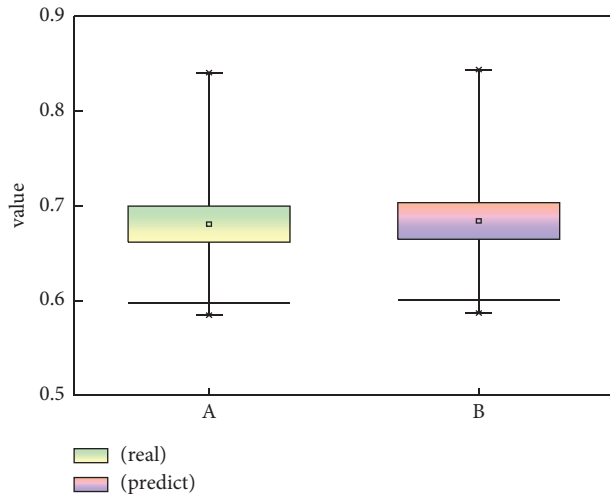


FIGURE 9: Predicted box distribution for the basics of piano teaching.

relatively easy to predict compared to the other two characteristics of piano teaching. CNN has basically reached a trustworthy level in predicting the basic knowledge of piano teaching.

## 5. Conclusion

With the advancement of technology and the improvement of living standards, the traditional teaching mode has limited the development of piano teaching. Piano teaching is a mode that is different from traditional subjects. It requires learners to experience the artistic information brought by piano. The player's on-the-spot performance will also have a certain impact on piano teaching.

This research introduces the teaching mode of videos and images into the teaching of piano, and it also uses CNN to predict the characteristics of piano music, player's behavior, and basic knowledge of piano in the piano teaching mode. Before using CNN prediction, this study also uses clustering methods to achieve the classification of piano video and image data. In terms of clustering, it shows certain feasibility in classifying relevant data of videos and images of piano teaching. The largest classification error is only 1.89%, and this part of the error comes from the relevant characteristic data of piano players. CNN also shows high accuracy in predicting the piano teaching features that fuse videos and images. The highest prediction error is only 2.23%, and this part of the prediction error also comes from the behavior characteristics of piano players. However, this part of the linear correlation coefficient exceeds 0.95. Overall, the design of this research is that the piano teaching mode that integrates videos and images will improve students' learning interest and efficiency, and CNN and clustering methods also show high accuracy in processing related features of piano videos and images.

## Data Availability

The dataset is available upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] A. Sepp, H. Ruismaki, and L. Hietanen, "Student teachers' and teacher educators' pedagogical reflections on piano courses in Finnish primary school teacher education," *Research Studies in Music Education*, vol. 4, no. 13, Article ID 221076997, 2022.
- [2] A. Van and F. Villiers, "Psychological attributes of primary school piano learners preparing for regional and national music competitions in South Africa," *Muziki-Journal of Music Research in Africa*, vol. 18, no. 2, pp. 120–135, 2022.
- [3] W. H. Luo and B. Ning, "Toward piano teaching evaluation based on neural network," *Scientific Programming*, vol. 2022, no. 5, Article ID 6328768, 2022.
- [4] C. Lee and B. Leung, "Instrumental teaching as 'the noblest and the most under-praised job': multiple case studies of three Hong Kong instrumental teachers," *Music Education Research*, vol. 24, no. 1, pp. 42–55, 2021.
- [5] Y. T. Niu, "Penetration of multimedia technology in piano teaching and performance based on complex network," *Mathematical Problems in Engineering*, vol. 11, no. 26, Article ID 8872227, 2021.
- [6] Y. K. Yang, "Piano performance and music automatic notation algorithm teaching system based on artificial intelligence," *Mobile Information Systems*, vol. 4, no. 1, Article ID 3552822, 2021.
- [7] Y. Zheng and B. Leung, "Perceptions of developing creativity in piano performance and pedagogy: an interview study from the Chinese perspective," *Research Studies in Music Education*, vol. 9, no. 25, Article ID 211033473, 2021.
- [8] H. Z. Li, "Piano education of children using musical instrument recognition and deep learning technologies under the educational psychology," *Frontiers in Psychology*, vol. 12, no. 10, Article ID 7051116, 2021.
- [9] J. F. Wang and F. F. Yang, "A fractional-order CNN hyperchaotic system for image encryption algorithm," *Physica Scripta*, vol. 96, no. 3, Article ID 035209, 2021.
- [10] J. F. Zhao, X. Mao, and L. J. Chen, "Learning deep features to recognise speech emotion using merged deep CNN," *IET Signal Processing*, vol. 12, no. 6, pp. 713–721, 2018.
- [11] T. Hsieh and J. Kiang, "Comparison of CNN algorithms on hyperspectral image classification in agricultural lands," *Sensors*, vol. 17, no. 6, Article ID 1724, 2020.
- [12] Z. L. Hu, Q. Zhao, and J. Wang, "The prediction model of cotton yarn intensity based on the CNN-BP neural network," *Wireless Personal Communications*, vol. 102, no. 2, pp. 1905–1916, 2018.
- [13] X. X. Jin, P. Peng, and Z. Huang, "Analysis of multi-level capital market linkage driven by artificial intelligence and deep learning methods," *Soft Computing*, vol. 24, no. 11, pp. 8011–8019, 2020.
- [14] G. N. Li, X. W. Zhao, C. Fan, X. Fang, F. Li, and Y. Wu, "Assessment of long short-term memory and its modifications for enhanced short-term building energy predictions," *Journal of Building Engineering*, vol. 43, no. 9, Article ID 103182, 2021.
- [15] K. X. Cui, "Artificial intelligence and creativity: piano teaching with augmented reality applications," *Interactive Learning Environments*, vol. 4, no. 13, Article ID 2059520, 2022.
- [16] J. Pi, "Effect of ergonomics-based piano teaching on teachers' physical and mental health and the improvement of sense of



- happiness,” *Journal of healthcare engineering*, vol. 2022, no. 4, Article ID 9174441, 2022.
- [17] X. Y. Liu, “Research on piano performance optimization based on big data and BP neural network technology,” *Computational Intelligence and Neuroscience*, vol. 2022, no. 4, Article ID 1268303, 2022.
- [18] X. M. Xue and Z. H. Jia, “The piano-assisted teaching system based on an artificial intelligent wireless network,” *Wireless Communications and Mobile Computing*, vol. 4, no. 13, Article ID 5287172, 2022.
- [19] W. Y. Li, “Analysis of piano performance characteristics by deep learning and artificial intelligence and its application in piano teaching,” *Frontiers in Psychology*, vol. 12, no. 2, Article ID 751406, 2022.
- [20] N. S. Huang and X. X. Ding, “Piano music teaching under the background of artificial intelligence,” *Wireless Communications and Mobile Computing*, vol. 2022, no. 13, Article ID 5816453, 2022.
- [21] R. Guo, J. N. Ding, and W. H. Zang, “Music online education reform and wireless network optimization using artificial intelligence piano teaching,” *Wireless Communications and Mobile Computing*, vol. 4, no. 14, Article ID 6456734, 2021.