IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 21 July 2023, accepted 19 September 2023, date of publication 9 October 2023, date of current version 18 October 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3323214

RESEARCH ARTICLE

A Hybrid Intelligence-Based Integrated Smart Evaluation Model for Vocal Music Teaching

RAN WU
School of Music, Shandong University of Arts, Jinan, Shandong 250012, China
e-mail: 13791127952@163.com

ABSTRACT The smart evaluation for teaching effect has received much attention, especially in field of vocal music. Currently, such evaluation mainly relies on expert rating, which costs much human labors. Fortunately, the machine learning and deep learning-based techniques have been applied to evaluation affairs in a number of areas. This work takes the vocal music teaching as the main object, and introduces several typical intelligent algorithms to construct a smart evaluation workflow. Thus in this paper, a hybrid intelligence-based integrated smart evaluation model for vocal music teaching is proposed. First of all, a comprehensive evaluation system is formulated from mechanism as the main feature space. Then, convolutional neural network, long short-term memory network and multi-layer perceptrons are employed to establish a novel integrated structure as the main evaluation model. To assess the proposed technical framework in this paper, a case study is conducted and some simulation experiments are carried out for this purpose. The experimental results show that the proposal can well realize automatic evaluation for vocal music teaching.

INDEX TERMS Hybrid intelligence, smart evaluation, neural networks, intelligent algorithms.

I. INTRODUCTION

Piano educators are the guides of piano scholars, and excellent piano educators can closely integrate people's aesthetic sentiment and artistic perception. This perceptual content can effectively promote the level of piano scholars and play a variety of roles in piano education. Piano education is called by many people the art of "cultivating sentiment and enhancing literary and artistic temperament" [1]. Through piano education, students' aesthetic ability can be improved and cultivated. The artistic aesthetic characteristics can improve the creative ability of piano learners. The ultimate goal of piano education is not to turn students into skilled piano playing machines, but to train students to be creators who can create beautiful repertoire. Music is a very perceptive and expressive existence, and artistic aesthetic characteristics are the basis of creation. Only by allowing students to have a profound perception and resonance of the artistic aesthetics in the piano can they find the core and essence of the piano performance. There is no upper limit to art, especially the piano with infinite possibilities. Students who learn piano need not only learn piano playing techniques, but also more importantly is to form one's own unique cognition of piano, which is the most important role of artistic aesthetic characteristics to piano education.

Foreign research on education level evaluation started earlier than China. At the beginning of the 20th century, the establishment of an education level evaluation system were carried out at many countries in Europe and the United States, but at that stage, only a few schools had a relatively formal education level evaluation process to fairly and effectively evaluate the education level of teachers. By the 1950s, formal evaluation standards for teachers' educational level were formulated, and the educational level evaluation system was scarcely used in colleges and universities until the 1970s. After that, the research on teacher education level evaluation has developed rapidly [2], [3]. Research in this field has made a qualitative leap At the end of the last century.

In the process of playing, children usually have active learning desires and interests. During this process, self-efficacy will gradually emerge, and the individual's behavior will be subtly integrated into the subject's social relationship.

The associate editor coordinating the review of this manuscript and approving it for publication was Debashis De.

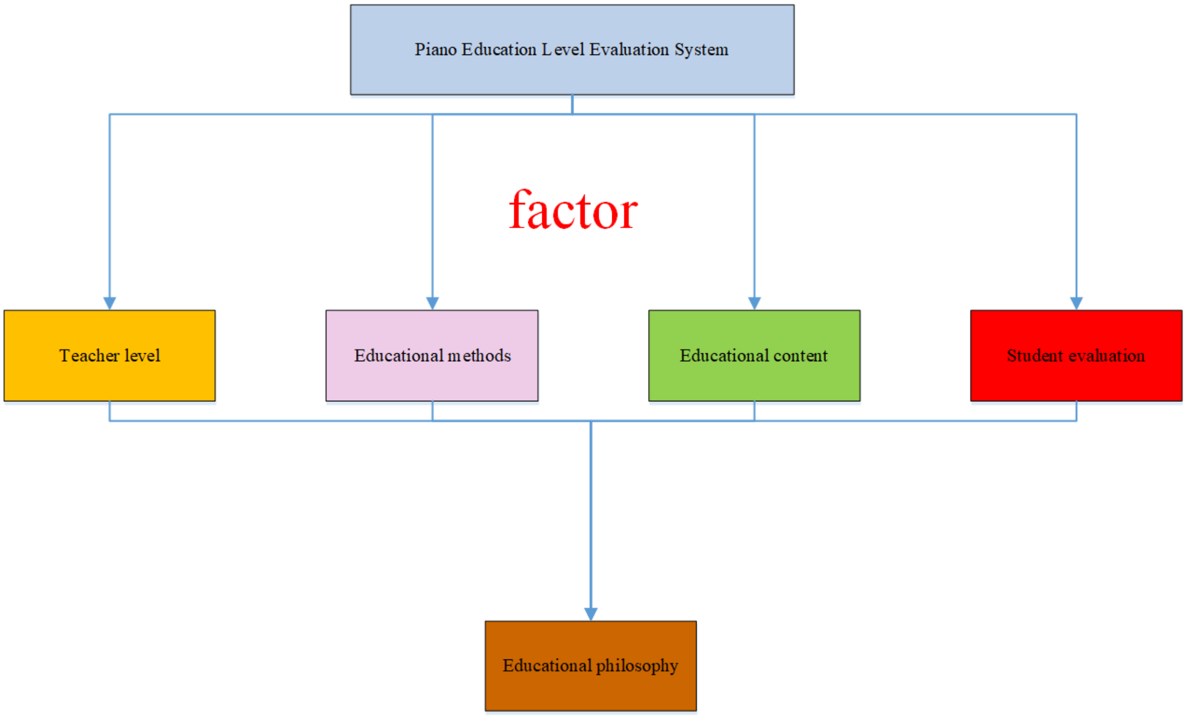

FIGURE 1. **The framework of the piano teacher's piano level evaluation system.**

For example, there are many traditional nursery rhymes that express emotion through word rhythm, phonetic rhythm, rhythm, and melody. In the enlightenment stage of children's piano learning, teachers can cultivate children's language, perception, thinking and other abilities, so that children use various sense organs of the body to cooperate with each other, cultivate their sense of rhythm, music and intonation, and develop their imagination and creativity. In the practice of piano teaching, teachers should abandon traditional ideas and teaching modes that pursue external value in traditional teaching, choose innovative teaching concepts, and improve the flexibility and interest of piano learning.

In order to make students no longer bored with piano learning, teachers need to abandon boring dead practice, pay attention to students' experience and feelings, and stimulate students' love for music. For the corresponding elderly people learning piano, although the number of elderly students participating in piano learning has increased year by year, the investment in piano teaching for the elderly has not increased accordingly. Problems such as insufficient piano classrooms, insufficient number of pianos, and low teacher-student ratio have not been improved. Neither hardware nor software can meet the actual needs of piano teaching for the elderly. From the current configuration of an ordinary university for the elderly, a district-level university for the elderly generally has one or two piano classrooms and one or two pianos, but it has to recruit dozens to hundreds of students.

Teachers have to take group lessons, using a piano to teach a dozen or even dozens of people at the same time. From the development process of Chinese piano, the development process of Chinese piano education, and the important role that piano education plays in contemporary early childhood education and elderly education in China, we can see that the level of piano education of piano educators is related to the enthusiasm of piano scholars to learn piano. The improvement of piano level is very important, and the demand for piano education in China is large, and the piano education level of piano educators varies, so it is indispensable to evaluate the piano education level of piano teachers.

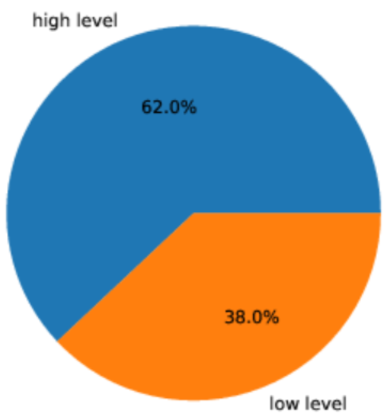

FIGURE 2. **The data distribution.**

Therefore, this paper establishes an evaluation system and an evaluation model for the evaluation of piano education level. This paper has three contributions:

- This paper established a piano education level evaluation system.
- This paper collect and establish a piano teacher's piano education level evaluation data set.

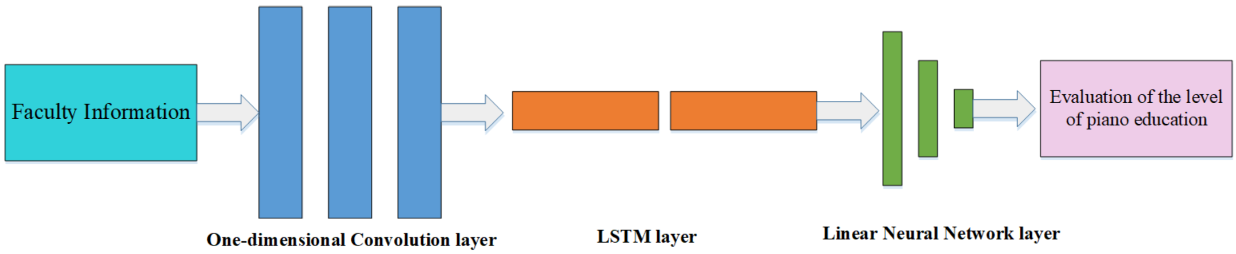

FIGURE 3. Model structure of the evaluation model of piano education level.

- This paper build a piano level evaluation model based on one-dimensional CNN, LSTM, and linear neural network.

II. RELATED WORK

In the field of music and art education, piano education plays a pivotal role in the cultivation of hearing ability, imagination ability, observation ability, memory ability, expressive ability, etc. [4], [5]. As the continuous development of my country's socialist economy, the cultivation of piano ability has been focused on by parents and educators [6], [7]. The piano education has been paid attention to from the ideological point of view, and new ideas have been continuously opened up, thereby improving the artistic quality of young people [8], [9]. The education level of piano educators plays a pivotal role in the whole process of piano education [10], [11]. The level of piano education can almost directly determine the quality of piano education [12], [13]. Therefore, a scientific and effective evaluation of the piano education level of piano educators is crucial and an indispensable process.

Deep learning, a major component in AI, has developed rapidly in recent years and is widely applied in various fields [14], [15]. Bao Wenxia et al. adopted a multi-channel convolutional neural network with joint loss to conduct image recognition research on wheat scab, segmented wheat images through U-net, extracted features using a multi-channel convolutional neural network, and learned with a joint loss function. And it achieved 100% accurate recognition of wheat scab disease [16]; Long et al. used deep learning network of transfer learning to carry out image recognition on four kinds of Camellia diseases, AlexNet to pre-train on Imagenet was used, and a brand new fully connected layer was designed, an average accuracy of 91.25% was obtained [17]. Abdullah-Al Nahid et al. used a multi-channel CNN to identify Chest Radiographs to diagnose pneumonia. The classification accuracy reached 97.92%, a very reliable detection method was provided [18].

Daoud et al. used deep learning to extract image features and manual extraction to process breast ultrasound images, classify breast tumors, and the average classification accuracy reached more than 95%, which can accurately detect breast cancer through breast ultrasound images [19]. Traditional system evaluation methods are mainly based on feature extraction. Features are usually extracted manually, and then the evaluation system is evaluated based on these extracted features combined with traditional machine learning methods. The performance of such methods depends on the quality of manually extracted features, which are usually extracted heuristically, making it difficult to capture factors comprehensively and effectively. Due to the excellent application of deep learning in the fields of text, speech, and images, prediction methods based on deep learning have gradually been paid attention to and proposed by researchers. With the powerful representation ability of deep learning, researchers have effectively modeled the evaluation system [20], [21].

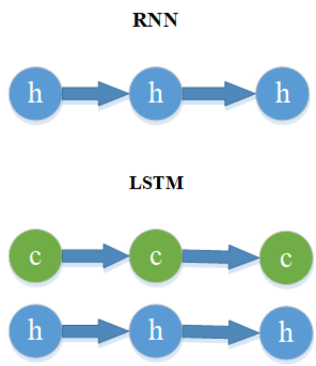

FIGURE 4. Schematic diagram of unit state.

III. EVALUATION SYSTEM OF PIANO EDUCATION LEVEL FOR TEACHERS

To scientifically and effectively evaluate the level of piano education, this paper establishes the evaluation factors of the level of piano education from four aspects: the teacher's personal level, the teaching content, the teaching method, and the teacher's educational philosophy. As an evaluation standard, a complete piano education level evaluation system has been established. The first is the personal level of teachers, the personal level of teachers which is divided into amateur level and professional level, the educational content of teaching piano is divided into rich and not rich, followed by the educational method of teaching piano, divided into one-to-one and one-to-many, and finally is the teacher's educational philosophy, divided into strict and loose. The students' extensive evaluation has two grades: high level and

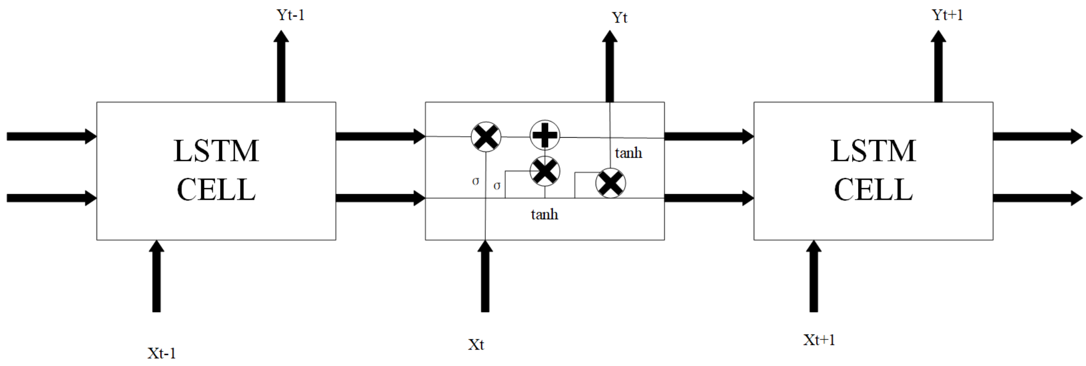

FIGURE 5. LSTM structure diagram.

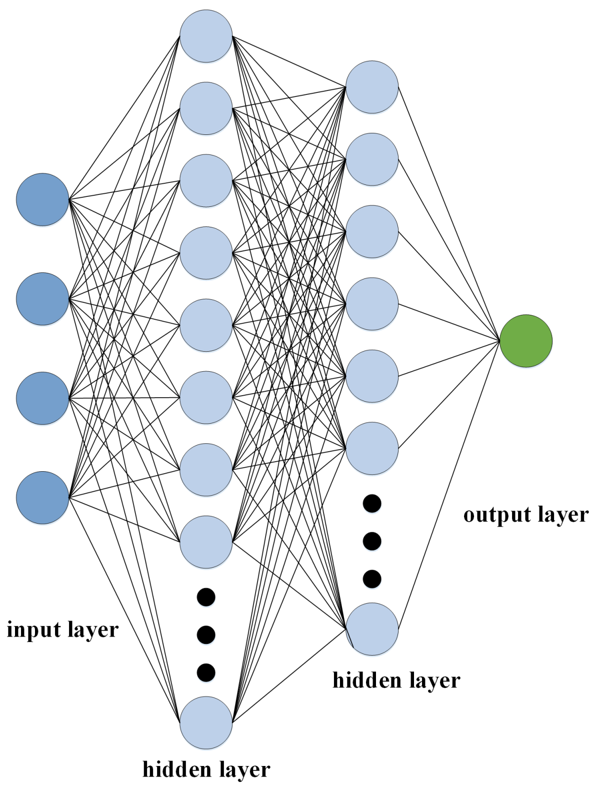

FIGURE 6. Schematic diagram of the linear neural network structure.

low level, which represent the teacher's piano education level. The relationship diagram of the whole evaluation system can be seen in Figure 1.

This paper collects information about the evaluation system and student evaluations of 1000 piano teachers in Jiangsu to establish a data set. Among them, 620 cases of teachers' piano education level are evaluated as high level, and 380 cases are evaluated as low level. The data distribution can be seen in Figure 2. And based on this information, the data is encoded so that the model can be used for training and prediction. The encoding method used is one-hot encoding. For a set of N kinds of category values, a status register with a length of N bits to encode each category value was used in one-hot encoding. Corresponding to an independent register bit, only one register bit is 1 and the rest are 0. Table 1 shows the values of teacher level, educational content, educational method, and educational philosophy after one-hot encoding.

IV. MODEL DESIGN

A. MODEL

The piano education level evaluation model designed in this paper is implemented based on one-dimensional convolution, long short-term memory network (LSTM) and linear neural network. The specific structure includes three-layer one-dimensional convolution, two-layer LSTM and three-layer linear neural network, as well as the composition of activation functions and the structure of the model. The schematic diagram of the model is shown in Figure 3.

CNN is a variant of multilayer perceptron (MLP) network, which was first used in 1980. Since its introduction, it has performed well in detection and classification problems and has been widely used. The convolutional layer is the core of the CNN. 1D-CNN is a feedforward neural network, which belongs to the classic deep neural network. The input is one-dimensional data, a one-dimensional structure is adopted in the convolution kernel, and the output of each convolution layer and pooling layer is also a one-dimensional feature vector.

The role of the activation function is to better solve complex problems. The more commonly used activation functions are the Sigmoid activation function, the Relu activation function, and the tanh activation function. In recent research, the Relu activation function is the most commonly used activation function, which can Yes, the neural network converges quickly, and its formula is shown as:

$$y = MAX(0, x) \tag{1}$$

LSTM is an improved recurrent neural network (RNN) that can overcome the problem that RNN cannot deal with long-distance dependencies, and is currently more widely used. The hidden layer of the original RNN has only one state (h), which is very sensitive to short-term inputs. Moreover, adding one more state, namely c (called the cell state), to save the long-term state, as shown in Figure 4.

A chained form of repeating neural network modules is included in all RNNs. In standard RNNs, this repeated

TABLE 1. Dataset encoding of piano education levels.

Factor	Category	Coding	Category	Coding
Teacher Level	Amateur Level	(0,1)	Professional Level	(1,0)
Educational content	Not rich	(0,1)	Rich	(1,0)
Educational methods	One-to-many	(0,1)	One-to-one	(1,0)
Educational Philosophy	Strict	(0,1)	Relaxed	(1,0)
Student evaluation	Low level	(0,1)	High level	(1,0)

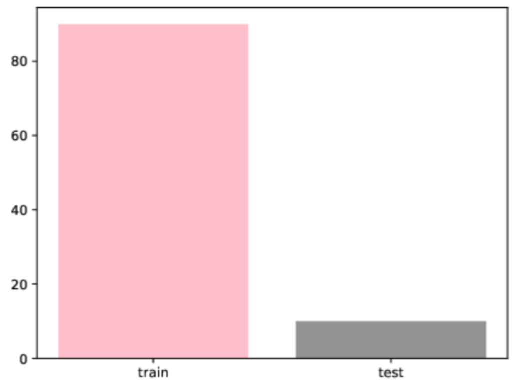

FIGURE 7. The ratio of training set and test set.

structural module has a quite simple structure, such as a tanh layer. LSTMs have the similar structure, while a different structure occurred in repeated modules. In addition to a single neural network layer, here are four, which is interacting in a specific way. The state of the cell throughout and the horizontal line that goes through the cell are the core to the LSTM. Cell states are similar with conveyor belts. Running directly on the entire chain, with only a few small linear interactions, it would be easy for the information to flow over it to remain the same. Gates can be implemented to selectively let information through, by using a sigmoid neural layer and a point-by-point multiplication operation. Each element of the output for the sigmoid layer (which is a vector) is a real number ranging between 0 and 1, representing the weight to let the useful information through. For example, 0 indicates "don't let any information through" and 1 indicates "let all information through". The input gate, forget gate and output gate are used to help LSTM realizes the protection and control of information. The structure of LSTM can be seen in Figure 5.

The main difference between a linear neural network and a perceptron is that the activation function of the perceptron can only output two possible values, while the output can take any value, and its activation function is a linear function. The linear neural network uses the Widrow-Hoff learning rule (Least Mean Square Rule), that is, the LMS (Least Mean Square) algorithm to improve the weights and biases of the network. In this paper, a four layer linear neural network is used, and the number of neurons in the linear neural network is small, so that the results can be well fitted even when the amount of data is small. Figure 6 shows the linear neural network designed in this paper.

B. LOSS FUNCTION

Since the collected piano teacher's piano education level dataset is an unbalanced dataset, if it is directly trained, the model will be unbalanced and the accuracy of the evaluation will not be high enough. The Taylor cross entropy loss function can decrease the influence of noise in the label on the model, increase the robustness of the model, and enhance the fitting ability of the model. Given a function g(x), assuming that the $n - th$ order is derivable, when $x = x_0$ you x_0, as shown in:

$$g(x) = \sum_{i=0}^{\infty} \frac{g^{(i)}(x_0)}{i!} (x - x_0)^i \tag{2}$$

According to $LCCE_{CC}(f(x), y) = -\log f_y(x)$, the following formulas can be defined:

$$g(f(x)) = -\log f_y(x) \tag{3}$$

$$g\left(f_y(x)\right) = \sum_{i=0}^{\infty} \frac{g^{(i)}\left(f_y(x_0)\right)}{i!} \left(f_y(x) - f_y(x_0)\right)^i \tag{4}$$

where $f_y\left(x_{\underline{n}}\right) = 1 and \forall i \geq 1$. Then, the following can be deduced:

$$g^{(i)}\left(f_y(x_0) = 1\right) = (-1)^i (i-1)! \tag{5}$$

The cross entropy loss at this time is as follows:

$$L_{CCE}(f(x), y) = g\left(f_y(x)\right) = \sum_{i=1}^{\infty} \frac{\left(1 - f_y(x)\right)^i}{i} \tag{6}$$

Obviously, n cannot tend to infinity, in reality n will tend to a finite term, then the Taylor cross entropy loss can be expressed as Equation 7, where t is the Taylor series.

$$L_{t-CE}(f(x), y) = \sum_{i=1} \frac{\left(1 - f_y(x)\right)^i}{i} \tag{7}$$

Label smoothing is a method to prevent over-fitting and can play a regularization role. The probability distribution of the original label can be expressed as:

$$P_i = \begin{cases} 1, & i = y \\ 0, & i \neq y \end{cases} \tag{8}$$

The probability distribution after label smoothing can be expressed as:

$$P_i = \begin{cases} 1 - \varepsilon, & i = y \\ \dfrac{\varepsilon}{K - 1}, & i \neq y \end{cases} \tag{9}$$

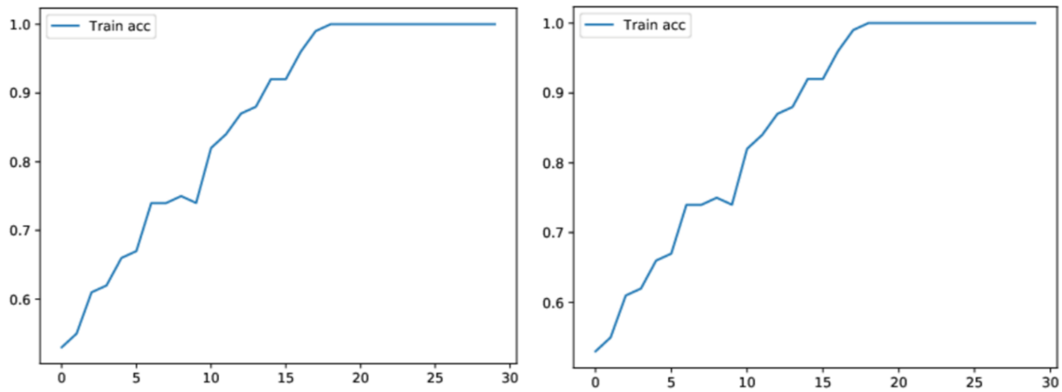

FIGURE 8. Training set accuracy curve and training set loss curve.

TABLE 2. Model comparison results.

Model	Accuracy
Random forest	80%
Support vector machines	80%
XGBoost	90%
This paper	100%

where K is the total number of categories, and ε is a custom smaller hyperparameter. The loss function at this time is expressed as:

$$LOSS_i = \begin{cases} (1-\varepsilon)^* L_{t-CE}(f(x), y) = \Delta_1 \\ \varepsilon^* L_{t-CE}(f(x), y) = \Delta_2 \end{cases} \tag{10}$$

where Δ_1 and Δ_2 respectively represent:

$$\Delta_1 = (1-\varepsilon)^* \sum_{i=1} \frac{\left(1 - f_y(x)\right)^i}{i}, \quad i = y \tag{11}$$

$$\Delta_2 = \varepsilon^* \sum_{i=1} t \frac{\left(1 - f_y(x)\right)^i}{i}, \quad i \neq y \tag{12}$$

Of the above two formulas, the i respectively satisfies the following conditions:

$$\Delta_1 : i = y \tag{13}$$

$$\Delta_2 = i \neq y \tag{14}$$

V. RESULTS AND ANALYSIS

A. EXPERIMENTAL ENVIRONMENT

The testing platform used in this study: the hardware condition is Intel Core i7-9700f processor, 16G memory; the software condition is WIN10 system, python3.7, and the deep learning framework used is pytorch.

B. EXPERIMENTAL RESULTS

1000 sets of data are divided into training set and test set, 900 sets of data are used as training set, and 100 sets of data are used as test set. The ratio of training set and test set can be seen in Figure 7. The RAdam optimizer is used as the optimizer, and the taylor cross entropy loss function is employed as the loss function. After training for 30 epochs, the test is performed. The evaluation index for evaluating the model performance is the accuracy rate. The formula for the accuracy rate is shown as:

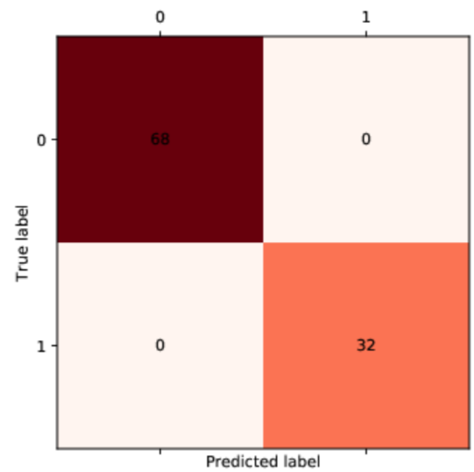

FIGURE 9. The test set confusion matrix.

$$ACC = \frac{(TP + TN)}{ALL} \tag{15}$$

After 30 epochs of training, the evaluation accuracy change curve and loss change curve of the piano education level evaluation model are in Figure 8.

According to the training results, it can be found that the accuracy of the training set of the piano level evaluation model proposed in this paper rises steadily, and can eventually reach 100% accuracy, while the loss also reaches the level of zero error after a steady decline, which shows that the piano education level evaluation model has a good fitting ability during the training process. Then the test set was used to test the piano education level evaluation model proposed in this paper, and the accuracy rate on the test set also reached 100%, which shows that the piano education level evaluation model proposed in this paper can effectively evaluate teachers level of piano education. The test set confusion matrix can be seen in Figure 9.

To further analyze the evaluation ability of the model, the performance of the proposed model is compared with three machine models of random forest, support vector machine and XGBoost. The comparison results can be seen in Table 2. From the table, it can be seen that among the three machine learning models, XGBoost has the highest classification accuracy, reaching 90%, while the classification accuracy of random forest and support vector machine is both 80%. The classification accuracy of the model is not as good as the proposed model. This is because the dataset is an unbalanced dataset, and there is a large data gap between the two categories, and it is difficult for traditional machine learning models to obtain a good classification accuracy on an unbalanced training set. The backbone network part of the proposed model is composed of convolutional neural network and LSTM, and is trained with Taylor cross entropy loss function and label smoothing which can reduce the impact of data imbalance, which makes the model in this paper resistant to training set noise. The data is imbalanced with the training set, resulting in better performance on the test set. This proves that the evaluation ability of the proposed model is better.

VI. CONCLUSION

Piano education is an important component of art education. The piano education level of piano teachers directly affects the quality of piano education. Therefore, scientific and effective evaluation of the piano education level of piano teachers is very important for piano education. Aiming at this key problem, this study proposes a piano education level evaluation model based on convolutional neural network. First, a piano education level evaluation system for piano teachers was built, and then the relevant data of 1000 piano teachers was collected according to the system, and a piano teacher level evaluation data set was established. This paper built the evaluation model of teachers' piano education level based on one-dimensional convolutional neural network, LSTM and linear neural network. And the evaluation model of piano education level is trained and tested by using the piano teacher level evaluation data set. The accuracy of the proposed model on the test set has reached 100%. Then, the proposed model is compared with the three machine learning models of random forest, support vector machine and XGBoost. The classification accuracy of the model is higher than these three machine learning models. The results show that the proposed method can evaluate the piano education level of piano teachers scientifically and effectively.

REFERENCES

[1] A. Nguyen, H. N. Ngo, Y. Hong, B. Dang, and B.-P.-T. Nguyen, "Ethical principles for artificial intelligence in education," *Educ. Inf. Technol.*, vol. 28, no. 4, pp. 4221–4241, Apr. 2023, doi: 10.1007/s10639-022-11316-w.

[2] H. K. Wachtel, "Student evaluation of college teaching effectiveness: A brief review," *Assessment Eval. Higher Educ.*, vol. 23, no. 2, pp. 191–212, Jan. 1998.

[3] M. Svinicki, "Reflective faculty evaluation: Enhancing teaching and determining faculty effectiveness," *Eval. Pract.*, vol. 17, no. 2, pp. 189–190, 1993.

[4] J. Li, "Analysis of piano curriculum education and cultivation of creative thinking ability," *Region-Educ. Res. Rev.*, vol. 2, no. 1, p. 6, Mar. 2020.

[5] S. W. Smoliar, J. A. Waterworth, and P. R. Kellock, "PianoFORTE: A system for piano education beyond notation literacy," in *Proc. 3rd ACM Int. Conf. Multimedia*, 1995, pp. 457–465.

[6] Y. S. Kaleli, "The effect of computer-assisted instruction on piano education: An experimental study with pre-service music teachers," *Int. J. Technol. Educ. Sci.*, vol. 4, no. 3, pp. 235–246, Jun. 2020.

[7] W. Zhu, "Study on the problems existing in Chinese piano education and improvement measures," *Educ. Sci., Theory Pract.*, vol. 18, no. 6, pp. 11–19, 2018.

[8] I. Yücetoker, Ç. E. Angi, and T. Kaynak, "Evaluation of asynchronous piano education and training in the COVID-19 era," *Educ. Res. Rev.*, vol. 16, no. 4, pp. 109–117, 2021.

[9] S. Akbulut, "'Turkish five': Their contributions to contemporary piano music and piano education in Turkeyy," *Online Submission*, vol. 8, no. 3, pp. 354–369, 2011.

[10] Z. Cui, F. Xue, X. Cai, Y. Cao, G.-G. Wang, and J. Chen, "Detection of malicious code variants based on deep learning," *IEEE Trans. Ind. Informat.*, vol. 14, no. 7, pp. 3187–3196, Jul. 2018.

[11] Y.-Z. Xiao, Z.-Q. Tian, J.-C. Yu, Y.-S. Zhang, S. Liu, S.-Y. Du, and X.-G. Lan, "A review of object detection based on deep learning," *Multimedia Tools Appl.*, vol. 79, no. 33, pp. 23729–23791, Sep. 2020.

[12] H. Liang, X. Sun, Y. Sun, and Y. Gao, "Text feature extraction based on deep learning: A review," *EURASIP J. Wireless Commun. Netw.*, vol. 2017, no. 1, pp. 1–12, Dec. 2017.

[13] N. C. Dang, M. N. Moreno-García, and F. De la Prieta, "Sentiment analysis based on deep learning: A comparative study," *Electron.*, vol. 9, no. 3, p. 483, 2020.

[14] V. Talreja, M. C. Valenti, and N. M. Nasrabadi, "Multibiometric secure system based on deep learning," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, Nov. 2017, pp. 298–302.

[15] R. A. Aral, S. R. Keskin, M. Kaya, and M. Haciomeroglu, "Classification of trashnet dataset based on deep learning models," in *Proc. IEEE Int. Conf. Big Data*, Dec. 2018, pp. 2058–2062.

[16] B. Wenxia, W. Dezhao, H. Gensheng, L. Dong, W. Nian, and Y. Xianjun, "Rice pest identification in natural scene based on lightweight residual network," *Trans. Chin. Soc. Agricult. Eng.*, vol. 37, no. 16, pp. 145–152, 2021.

[17] M. Long, C. Ouyang, H. Liu, and Q. Fu, "Image recognition of Camellia oleifera diseases based on convolutional neural network & transfer learning," *Trans. Chin. Soc. Agricult. Eng.*, vol. 34, no. 18, pp. 194–201, 2018.

[18] M. Masud, N. Sikder, A.-A. Nahid, A. K. Bairagi, and M. A. AlZain, "A machine learning approach to diagnosing lung and colon cancer using a deep learning-based classification framework," *Sensors*, vol. 21, no. 3, p. 748, Jan. 2021.

[19] M. I. Daoud, A. A. Atallah, F. Awwad, M. Al-Najjar, and R. Alazrai, "Automatic superpixel-based segmentation method for breast ultrasound images," *Exp. Syst. Appl.*, vol. 121, pp. 78–96, May 2019.

[20] X. W. Gao, R. Hui, and Z. Tian, "Classification of CT brain images based on deep learning networks," *Comput. Methods Programs Biomed.*, vol. 138, pp. 49–56, Jan. 2017.

[21] O. S. Lih, V. Jahmunah, T. R. San, E. J. Ciaccio, T. Yamakawa, M. Tanabe, M. Kobayashi, O. Faust, and U. R. Acharya, "Comprehensive electrocardiographic diagnosis based on deep learning," *Artif. Intell. Med.*, vol. 103, Mar. 2020, Art. no. 101789.

RAN WU received the bachelor's and master's degrees from the Shandong University of Arts, in 2001 and 2014, respectively. She is currently a Lecturer with the Shandong University of Arts. Her research interests include smart education and big data analysis.