International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18 (2026) pp. 1-16

Article

Research on the Infiltration and Inheritance Strategy of Chinese Traditional Music Culture in Higher Vocational Vocal Music Teaching

Shifang Yang [1,*]

[1] School of Preschool Education, Chongqing Youth Vocational & Technical College, Chongqing, 400700, China
* Correspondence author: 13677645698@163.com

Abstract: The practical application of traditional music culture in higher vocational vocal teaching can not only stimulate students' interest in vocal learning and give them a more comprehensive understanding of vocal knowledge, but also help students improve their own vocal perception ability and comprehensive vocal literacy. In this paper, oriented to the field of traditional music culture, we design and validate the named entity recognition algorithm based on the BERT-BILSTM-CRF model and the entity relationship extraction model based on the BERT-BiLSTM-Att to complete the visualization of the creation of the knowledge map of Chinese traditional music culture. On this basis, the online-offline hybrid vocal music teaching mode for higher vocational education that integrates the knowledge graph is constructed, and teaching experiments are conducted to investigate the effect on the inheritance of traditional Chinese music culture. The experiment shows that after the intervention of using the vocal music teaching method incorporating traditional music culture knowledge mapping, the students' mastery of traditional music culture knowledge, traditional music culture literacy, and vocal singing ability are significantly improved ($P<0.05$), and are better than the traditional teaching method, i.e., the method helps students to improve their own vocal music perception ability and comprehensive vocal music literacy. Based on this study, it is expected to help improve the teaching level of vocals in higher vocational education and the inheritance and development of Chinese traditional music culture.

Keywords: named entity recognition; entity relationship extraction; knowledge graph; traditional music culture; higher vocational vocal music teaching

1. Introduction

With the development of society, the diversification of modern music culture promotes the development of higher vocational vocal music teaching, optimizes the teaching methods of vocal music in colleges and universities, and makes music education present a good development trend [1]. However, the current vocal music teaching mode exists in which teachers' teaching focuses on vocal music technology and traditional music culture penetrates less, leading to the fact that most of the students are more inclined to become artisans of art and deviate from the path towards artists [2-3]. The problem is attributed to the traditional teaching of vocal music, students have less understanding of the culture and essence of music art, and to become a musical artist requires students to have a strong musical perception [4]. The development of science and technology is ever-changing, even though the development of new technology can reflect the progress of modern society, but the ancient oriental culture is carrying the treasure of China's long history, in the teaching of vocal music, the two should complement each other, reflecting the balance of science and technology and culture [5]. The traditional techniques of Western music can improve students' vocal skills, but incorporating the humanistic spirit contained in traditional music culture is an important way to help students improve their vocal literacy. The famous thinker Lao Tzu believed that a person's thinking and spiritual realm can influence the use of human senses, and without the support of cultural connotation in vocal music teaching, it is difficult for students to move towards the long-term development of music and art [6]. Therefore, it is necessary to actively integrate traditional Chinese culture into vocal music teaching, continuously optimize vocal music teaching methods, and improve the quality of vocal music teaching.

Traditional music culture is the treasure of the Chinese nation, containing rich emotional expression, philosophical thinking and artistic value, and higher vocational education, as an important stage of vocational personnel training, bears the mission of inheriting and promoting culture [7]. Incorporating traditional music culture into vocal music teaching can effectively help students understand the music tradition of their own nation, enhance cultural self-confidence, and then promote the inheritance and innovation of national culture [8-9]. Introducing traditional music culture into the classroom enables students not only to understand the social background and cultural connotations behind the musical works, but also to further cognize the unique status and contribution of Chinese culture in the world art stage, thus inspiring them to love and cherish the national art [10]. However, judging from the current teaching of vocal music in Chinese higher vocational colleges and universities, there is a relative gap in the quality of vocal music teaching due to the low basic literacy of the students and the teaching level of the teachers is yet to be improved [11-12]. In addition, the teaching concept is also a key factor limiting teachers' effective teaching [13]. At this stage, the teaching concepts of many teachers are introduced from western countries, which cannot effectively meet the current development process in China, and the vocal music teaching carried out by teachers cannot effectively mobilize students' emotional cognitive ability, which makes the development of vocal music teaching in Chinese higher vocational colleges and universities seriously hindered [14-16].

In addition, in the curriculum of vocals in higher vocational colleges and universities, traditional music culture accounts for a relatively low percentage [17]. Most of the courses focus on the singing techniques of western vocal works, neglecting the analysis and teaching of traditional musical works. The singularity of this course structure makes it difficult for students to fully understand the essence of folk music and limits their musical expression [18-19]. Even in courses involving traditional music, the teaching content often stays at a shallow level of knowledge introduction, such as simple repertoire appreciation or basic singing exercises, lacking in-depth analysis of the connotation, background and techniques of traditional music works [20]. Some teachers do not pay enough attention to traditional music culture, and the teaching process lacks systematicity and coherence, which makes it difficult for students to form a comprehensive knowledge of traditional music culture [21]. In this context, analyzing the significance of the application of traditional music culture in higher vocational vocal music teaching, analyzing the problems existing in its specific application, and exploring the penetration and inheritance strategies of traditional music culture applied to higher vocational music teaching are of great practical significance.

In this study, after annotating the cleaned and organized textual data of Chinese traditional music culture, the BERT+BILSTM+CRF and BERT-BiLSTM-Att models are introduced to extract the entities and relationships in the unstructured data, and the effectiveness of the methods is verified by comparing the performance of the established dataset of relationships of entities in Chinese traditional music culture with multiple models. Then we use Neo4j to construct a huge knowledge graph of Chinese traditional music cultural heritage, which lays the foundation for the graphical visualization of data in the field of traditional music culture. Finally, we construct a mixed teaching mode of higher vocational vocal music integrating the knowledge graph, and use vocal music teaching as a practical carrier to vigorously inherit and promote Chinese traditional music culture.

2. Extraction of knowledge of traditional Chinese music culture

2.1. Acquisition and labeling of traditional music and cultural terms

The data of this study mainly comes from China Intangible Cultural Heritage Website. After automatic crawling and data cleaning of the relevant contents of the representative national intangible cultural heritage projects of traditional music category (including extended projects and new projects), a total of 155 projects' information is obtained, which contains the project name, the serial number of the project, the number of the project, the time of its publication, the category, the region to which it belongs, the type, and the region or unit of its declaration, Detailed description, related successor information, and related project information.

On the basis of the above data, this study firstly summarizes and categorizes the terms of traditional music, and obtains a total of 16,829 entities in 6 categories, including the name of the non-heritage, unique term, inheritor's name, place name, work name, and name of tools and instruments, as shown in

Table 1. Among them, non-heritage name, inheritor's name, place name and work name have corresponding field categories in the source data, while the unique terms and tool and instrument names have more obvious hints of context in the detailed introduction and related project information to help manual identification and access.

Table 1. The Summary of Traditional Music Terminology

Physical category	Category mark	Example	Quantity
Unleft name	ICH-TITLE	The terms of river, old mouth and hui family banquet	1511
Exclusive term	ICH-TERM	The cross arms of the pipa, the wine song, the six words	8856
Inheritors	ICH-INHERITOR	Chen zi jing, shen hao first, shen xin state	221
Place name	ICH-PLACE	Tibetan autonomous prefecture, Tibetan region and ando	2765
Title	ICH-WORKS	"The village of the hero"	1358
Tool name	ICH-INST	Pipa, flute, suona, tubule	2118

The corpus is constructed based on the detailed introduction of the project and related project information, etc., and based on which manual word segmentation and lexical annotation are completed, and manual annotation is carried out on the above six categories of entities in the corpus with reference to the terminology summary data. The manual annotation adopts multi-person cross-annotation and verification to ensure the accuracy and reliability of the annotated content. On the basis of terminology annotation, a serialized corpus is constructed for use in annotating subsequent entity recognition models. The size of the completed corpus of national intangible cultural heritage in the category of traditional music is about 500,000 words.

2.2. Named Entity Identification

In this paper, named entity recognition is performed by BERT-BILSTM-CRF model, which consists of three main parts, firstly, the BERT pre-trained language model, the unstructured textual data of Chinese traditional music culture is transformed into vector form and extracted the rich semantic features embedded in Chinese traditional music culture, then the contextual features are further extracted by the BILSTM model, and finally Finally, the constraints are added by CRF to reduce the generation of error sequences, and the final labeled sequences are output.

2.2.1. BERT pre-trained language models

The BERT model is a language pre-training model. The structure of the model is shown in Figure 1.

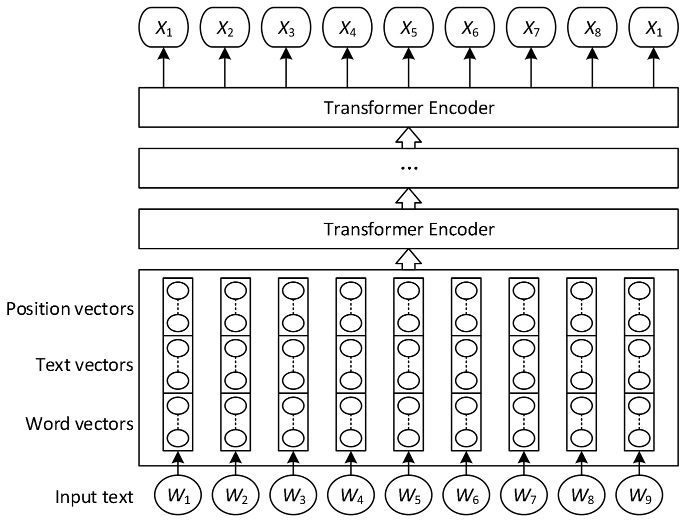

Figure 1. BERT Model

In this paper, after data screening and labeling of the original traditional music and culture data, the labeled text data are sliced and diced, and then vector representations are performed.The Transformer structure is a key part of BERT, which is a deep network based on the attention mechanism, and the weight sparsity matrix is adjusted by calculating the correlation degree between each word and the other words in the same sentence to get the feature vector of the word expression of the word. In this paper, we obtain the text sequence vectors with context-rich semantic features through the Encoder layer of the Transformer, and then output the vector, $\bar{x}$, as the Embedding layer of the Named Entity Recognition Model, which is inputted into the BILSTM model.

2.2.2. The BILSTM model

LSTM model is a special kind of recurrent neural network, the internal structure of hidden units in this network structure is very complex, through the introduction of memory units and gated memory units to save historical information, long-term state, and use gates to control the flow of information, effectively realizing the storage and updating of contextual information.

Each LSTM unit controls the information state through the three structures of forgetting gate, input gate and output gate, and the internal formula of the LSTM unit is:

$$f_t = Sigmoid\left(W_f \times \left[h_{t-1}, x_t\right] + b_f\right) \tag{1}$$

$$i_t = Sigmoid\left(W_i \times \left[h_{t-1}, x_t\right] + b_i\right) \tag{2}$$

$$o_t = Sigmoid\left(W_o \times \left[h_{t-1}, x_t\right] + b_o\right) \tag{3}$$

$$C_t = f_t * C_{t-1} + i_t * \tanh\left(W_c \times \left[h_{t-1}, x_t\right] + b_c\right) \tag{4}$$

$$h_t = o_t * \tanh\left(C_t\right) \tag{5}$$

As shown in the figure, there are three inputs to the LSTM, the current moment input x_t, the output value of the LSTM at the previous moment h_{t-1}, and the cell state at the previous moment C_{t-1}, and two outputs, the output value of the LSTM at the current moment h_t and the cell state at the current moment.The LSTM model realizes the selectivity of the information state through three gate structures Output. Where W and b denote the weights and bias terms, Eq. (1) is the oblivious gate state update formula, and $[h_{t-1}, x_t]$ denotes a longer vector formed by putting two vectors together.The sigmoid function serves to limit the output value of the gate to between 0 and 1. When the gate output is 0, any vector multiplied with it yields a 0 vector, which corresponds to nothing can be pass, and when the output is 1, any vector multiplied by it will not change anything, which is equivalent to nothing passing.

Eq. (1) determines how much of the cell state C_{t-1} from the previous moment is retained to the current moment C_t, and Eq. (2) is the state update formula for the input gate, which determines how much of the current network output x_t is saved to the state cell C_t. Eq. (3) is the formula for the state of the cell at the current moment, and Eqs. (4) and (5) are the formulas for the output gates, which determine how much of the control cell state C_t is output to the current output value h_t of the LSTM. Obviously, the hidden state h_t of the current LSTM cell depends on the previous hidden state h_{t-1}, but is not related to the next hidden state h_{t+1}, i.e., the information only flows forward in the unidirectional LSTM. This makes the LSTM model suffer from gradient vanishing or gradient explosion.

In this paper, we use the BILSTM model based on the BERT pre-trained language model to optimize the iterative problem of the recurrent neural network (RNN) model by the bidirectional semantic information, i.e., the potential semantic relationship, in the traditional music and culture data, to alleviate the phenomenon of gradient vanishing or gradient explosion, and to improve the long term memory capability of the sequence data.

2.2.3. CRF model

Conditional Random Fields (CRFs) are used as a conditional probability distribution model for named entity recognition. In the field of named entity recognition, its primary function is to pick out a labeling sequence with the highest probability as our labeling of the sentence among multiple possible labeling sequences. Although the BILSTM model is able to output the probability value of the label taking values, some of the labels output directly with the BILSTM model are not reasonable, the reason is that the correlation between labels and labels is not taken into account, for example, the head of the entity must not begin with I , the next label after the O label must not be I , the B-Dis label must be followed by I-Dis, and so on, therefore, in the BILSTM model is followed by a CRF layer to add a constraint mechanism, so that the output labels can be adjusted to make the result order of the labels more reasonable, thus improving the accuracy of the model. In the task of this paper, the main application is the linear chain conditional random field, the principle of which is shown in equation (6) as:

$$P(X|Y)=\frac{\exp(s(x,y))}{Z(x)} \tag{6}$$

where $Z(x)$ denotes the normalization factor and $Z(x)$ and $s(x,y)$ are calculated as:

$$Z(x)=\sum_{y'}\exp(s(x,y')) \tag{7}$$

$$s(x,y)=\sum_{i} Emit(x_i,y_i)+Trans(y_{i-1},y_i) \tag{8}$$

where $Emit(x_i,y_i)$ denotes the output probability of the LSTM, and $Trans(y_{i-1},y_i)$ denotes the corresponding transfer probability, which is also the value corresponding to the CRF transfer probability.

2.3. Relationship Extraction

In this section, a relational extraction model of Chinese traditional music and cultural knowledge based on BERT-BiLSTM-Att neural network model is proposed based on the framework of deep learning, which combines the advanced pre-trained language model and the attention mechanism.The BERT-BiLSTM-Att neural network model firstly applies the pre-trained Chinese BERT model from Google to the BERT-BiLSTM-Att for extracting the textual semantic features of traditional Chinese music and cultural knowledge, and then use the bidirectional recurrent neural network BiLSTM to model the sequential contextual semantic features, and finally use the attention mechanism to assign different weights to different words, and get the predicted relationship categories after the Softmax layer.The BERT-BiLSTM-Att neural network The modeling framework is shown in Fig. 2.

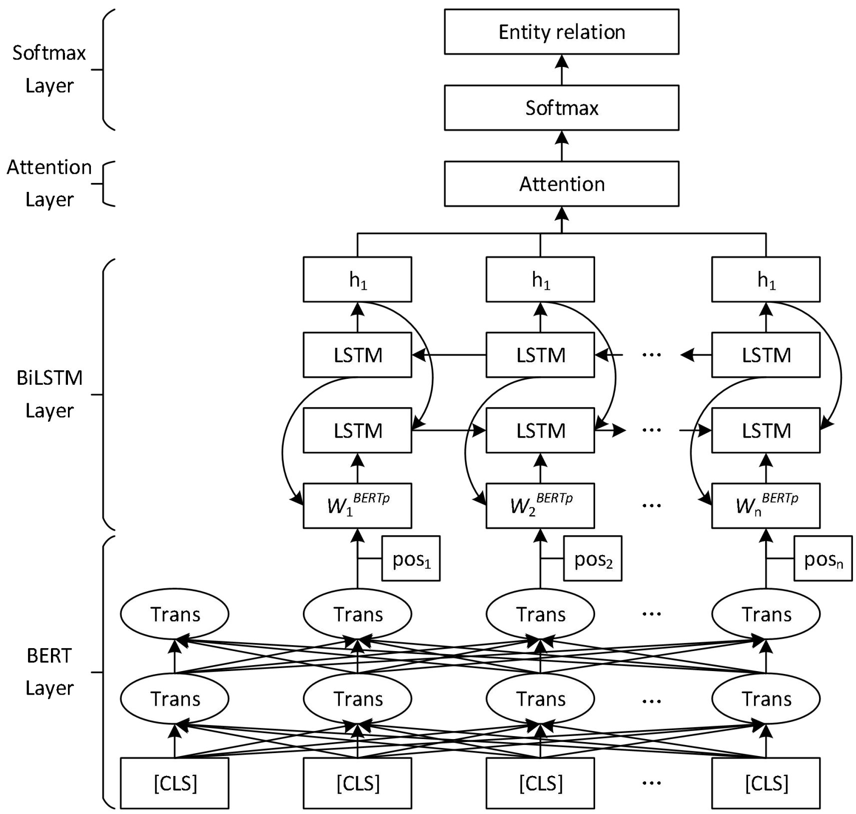

Figure 2. BERT-BiLSTM-Att neural network model framework

2.3.1. BERT-based semantic feature extraction

This layer utilizes BERT to extract semantic features of traditional Chinese music and cultural knowledge. Inputting the text into BERT can get the feature codes of each character, which can characterize very rich contextual semantic information, and at the same time, when modeling the information based on the bidirectional Transformer encoder, it takes into account both the past information and the future information, so as to appropriately transform and adjust the semantic information, and be able to dynamically characterize multiple meanings of a word.

The input sequence $W_1, W_2, \cdots, W_n$ of the BERT layer is in the form of word vectors containing word vectors, sentence vectors, and position vectors, and the input sequence is fed into the BERT pre-trained language model for the pre-training of word vectors, and then the feature coding of each word is obtained to imply a rich semantic feature of the sequence context. Usually, the closer the words are to the two entities in the sentence, the more they reflect the greater influence on the entity relationship, therefore, the positional distance of each word from the entity is taken into account, and these three vectors of the sentence are spelled as the features $W_1^{BERTp}, W_2^{BERTp}, \cdots, W_n^{BERTp}$.

2.3.2. Semantic modeling based on BiLSTM neural network

Based on the BERT semantic feature extraction although implies the rich semantic features of sequence context, and RNN recurrent neural network has a strong applicability in sequence modeling with high sensitivity to word order and positional features, therefore, this paper uses LSTM to model the sequence context semantic information.

The LSTM model is an optimization of the RNN, and the bidirectional LSTM network extracts features that take into account both the connection to the above and the connection to the below, and can learn features from both past and future sequences, and thus has an outstanding performance in relational classification applications. In the BiLSTM layer, the forward LSTM can model the text sequence from left to right, and its input at each moment is W_n^{BERTp}, which is combined with the coding vector h_{n-1}^R prior to that moment, and the output of the coding vector at that moment is:

$$h_n^R = LSTM\left(W_n^{BERTp}, h_{n-1}^R\right) \tag{9}$$

The backward LSTM, on the other hand, is modeled from right to left, and its input at each moment is W_n^{BERTp}, which, combined with the coding vector h_{n+1}^L after that moment, outputs the coding vector at that moment as:

$$h_n^L = LSTM\left(W_n^{BERTp}, h_{n+1}^L\right) \tag{10}$$

The output h_n^R of the forward LSTM and the output h_n^L of the backward LSTM are spliced together to obtain h_n, h_n which is the final feature encoding obtained by the Bidirectional Long Short-Term Memory Network (Bi-LSTM) modeling each word, denoted as $h_n = \left[h_n^R; h_n^L\right]$, and the final output of the hidden state encoding vector can be denoted as $H = \{h_1, h_2, \cdots, h_{n-1}, h_n\}$.

2.3.3. Attention layer

The most helpful information for relationship recognition may appear at different locations in the sentence, and the bidirectional LSTM captures local features of the sentence that do not make it easy to determine which words in the sentence are more important for relationship classification. Adding the Attention layer lies in using Attention to assign different weights to each word in response to the different effects of the words on the relationship categorization. The output representation of the Attention layer is calculated through equation (11):

$$h_\alpha = \sum_{t=1}^{T} \alpha_n \cdot h_n \tag{11}$$

where α_n denotes the weight under n moments, and α_n is calculated from the coding vector of the moment and the coding vector of the most recent moment, which is calculated as follows:

$$\alpha_n = \frac{\exp(h'_n)}{\sum_{n'=1}^{N} \exp(h'_{n'})} \tag{12}$$

$$h'_n = \tanh\left(W^n h_n + b^n\right) \tag{13}$$

h_n denotes the feature output at n moments, W^n is the weight matrix, and b^n is the bias term.

2.3.4. Softmax output layer

Input the output h_α of the attention layer into the fully-connected layer module, a network layer whose output vectors have dimensions equal to the number of relations, with each dimension corresponding to the predicted score of the i th relation. The vectors learned by the fully connected layer module are input to the Softmax layer to predict the probability of the category relations. Define y as the predicted result of the final solution for the relationship.

$$y\left(y|S\right) = soft\max\left(W^s h_\alpha + b^s\right) \tag{14}$$

$$y = \arg\max_y y\left(y|S\right) \tag{15}$$

where y is the relation category, W^s is the weight parameter learned by the classifier, and b^s is the classifier bias term.

3. Experimentation and Construction of Knowledge Maps of Traditional Chinese Music Culture

3.1. Experiment on identification of traditional music and cultural knowledge

3.1.1. Evaluation indicators

Assume that a total of N concepts, i.e., N categories of entities, are defined. For the i th category $\left(i = 1, 2, \cdots, N\right)$, in order to consider the recognition effect of the NER model on the entities of this category, the precision rate (Precision), the recall rate R_i (Recall), and the reconciled mean F_i (F1-measure) can be used as the evaluation metrics, and their Definitional equations are as follows:

$$P_i = \frac{Correct_i}{Predict_i} \times 100\% \tag{16}$$

$$R_i = \frac{Correct_i}{Label_i} \times 100\% \tag{17}$$

$$F_i = \frac{2 \times P_i \times R_i}{P_i + R_i} \times 100\% \tag{18}$$

where $Correct_i$ represents the number of entities of class i correctly recognized by the model, $Predict_i$ represents the number of entities of class i predicted by the model, and the accuracy of the model's recognition on entities of class i is measured by the precision rate; $Label_i$ represents the number of entities in the dataset labeled as entities of class i, and the proportion of entities recognized by the model on class i entities; F_i is the reconciled average of accuracy and recall, using this metric to measure the model's recognition effect on class i entities, and then using the weighted average to obtain the overall metric.

3.1.2. Experimental environment

The experiment is based on Ubuntu system version number 18.04 LTS, CPU is Intel(R) Core(TM) i7-10700K CPU @3.80GHZ, RAM is 2*8 dual-channel RAM, graphic card is NVIDIA GeForceRTX2080 SUPERr, and the software environment is Python3.8, PyTorch 1.12.

The parameter settings of the BERT-BiLSTM-CRF model in this paper are shown in Table 2. The hyper-parameter settings are batch size of 32, lr of 0.0005, Epoch iteration number of 100, the dropout mechanism is added to prevent overfitting, which is set to 0.5, and the maximum sequence length of each processing is 32, in which Adam is used for optimization in the BiLSTM module, and the BERT uses the version of $BERT_{BASE}$ using a 12-layer network with 768 dimensions of embedding, BiLSTM with 2 layers and 384 dimensions of hidden layers.

Table 2. Experimental Parameter Design

Hyperparameter	Parameter values
Batch_size	33
lr	0.0005
n_epochs	52
dropout	0.5
seq_length	33
optimizer	Adam

3.1.3. Experiments and analysis of results

In order to verify the effectiveness of the model in this paper, the named entity recognition methods based on HMM, CRF, BiLSTM, and BiLSTM-CRF are compared.In the experiments based on HMM and CRF, for the dataset in Table 1, the division method of 90% doing the training set and 10% doing the test set is taken, and in the experiments based on BiLSTM and based on BiLSTM-CRF, the division method of 80% do the division of training set, 10% do the validation set and 10% do the test set. The specific experimental results are shown in Table 3. From the table, it can be seen that the BERT-BiLSTM-CRF algorithm proposed in this paper has some improvement in each index.

Table 3. Comparison of experimental results

Model	P/%	R/%	F1/%
HMM	73.46	70.90	72.16
CRF	77.65	75.46	76.54
BiLSTM	83.14	81.58	82.35
BiLSTM-CRF	86.84	88.23	87.53
BERT-BiLSTM-CRF	97.52	95.47	96.48

The variation of F1 values in the first 50 epochs of the different models is shown in Figure 3. From the results, it can be seen that the experimental results of HMM alone, and CRF are relatively similar, with the HMM model reaching the highest F1 value of 71.15% in the 46th epoch, while the CRF model achieves an optimal solution of 75.48% F1 value in the 42nd round of training through the feature template. The results of BiLSTM alone are slightly better than the former, reaching 81.29% F1 value in the 48th round of training. And the model fused with BiLSTM-CRF, which obtains the historical information before and after the words on the basis of CRF, fuses the context, and can fuse complex semantic information, the result is improved by 5.15% compared to the model with CRF alone, and reaches 86.44% in the 49th round of training. And after using the pre-training model Bert with large-scale corpus, the recognition results are improved by another 10.12% compared to BiLSTM-CRF, and the F1 value reaches 96.48% in the 46th round of training. It is evident that the results of the Chinese traditional music and culture dataset can perform better with the addition of the pre-training model than the traditional word2vec static word vectors, which can better express the meaning of the word, embedded representation of the text, and the speed of convergence is significantly higher than the original model.

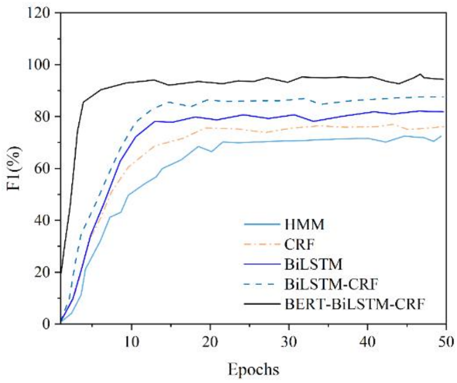

Figure 3. The experimental results vary with the number of iteration rounds

The F1 values of each model on each type of label are shown in Table 4.The classification results of each type of BERT-BiLSTM-CRF are shown in Table 5. By comparing the recall rate, accuracy rate and F1 value of each type of entities, we can see that the recognition effect of the two types of labels, i.e. unique terms of traditional music and culture and the names of traditional music and culture works, is excellent, with the F1 values of 98.13% and 97.29, respectively, which have certain regularity and simple structure, and thus the recognition effect is the most excellent. The rest of the related entities do not have clear boundaries, are ambiguous, have many nested entities, and are semantically complex, resulting in poor recognition results.

Table 4. The identification effect of various labels is compared

Model	**ICH-TITLE**	**ICH-TERM**	**ICH-INHERITOR**	**ICH-PLACE**	**ICH-WORKS**	**ICH-INST**
HMM	53.49	86.01	66.99	36.85	57.63	47.62
CRF	54.15	86.92	72.43	39.71	67.43	50.02
BiLSTM	55.79	89.74	80.61	41.31	77.83	52.22
BiLSTM-CRF	65.35	93.42	86.13	52.13	84.97	55.58
BERT-BiLSTM-CRF	83.13	98.13	91.04	89.38	97.29	95.94

Table 5. The model is compared to the identification effect of each label

Entity class	**P/%**	**R/%**	**F1/%**
ICH-TITLE	82.46	83.82	83.13
ICH-TERM	98.86	97.41	**98.13**
ICH-INHERITOR	97.42	85.44	91.04
ICH-PLACE	90.81	87.99	89.38
ICH-WORKS	97.91	96.67	**97.29**
ICH-INST	97.09	94.81	95.94

3.2. Experiments on the extraction of traditional music and cultural knowledge

The experimental results based on the BERT-BiLSTM-CRF model on this dataset are shown in Figure 4. As can be seen from the figure, the vertical axis is the percentage of each indicator and the horizontal axis is the number of iterations of the model. It can be found from this that the rate of change is faster in the first 15 iterations, after which the model gets a rapid convergence and reaches a relatively stable value after 37 times, after which there is less fluctuation, and in the 45th round, the F1 value reaches 92.01% to reach the extreme value.

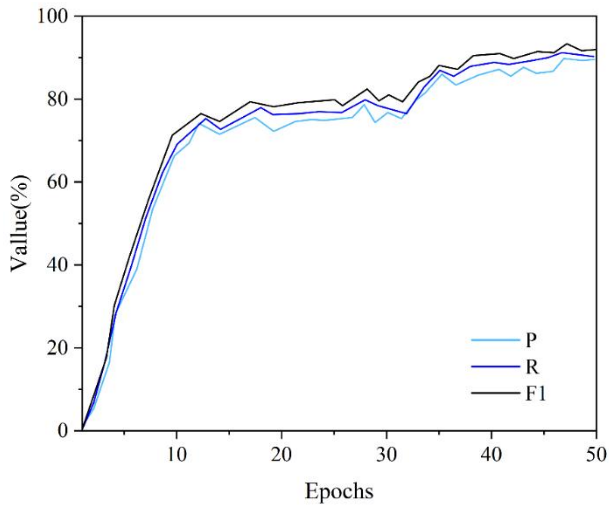

Figure 4. Results of the BERT-BiLSTM-CRF at model execution

The model proposed in this paper is compared with other models and the results of BERT-BiLSTM-Att model are compared in Table 6.The BERT-BiLSTM-Att relational extraction model achieves a precision rate of 93.03%, a recall rate of 91.12%, and an F1 value of 92.07%, which is a better performance than other models. The BERT-BiLSTM-Att model proposed in this paper adds Attention operation on the basis of BERT-BiLSTM, which improves the F1 value by 4.96%. The BERT-BiLSTM-Att model proposed in this paper has a better effect in the relational extraction of traditional music and cultural knowledge, and can effectively improve the training effect compared with the existing models, and can be It can be generalized and applied to more directions in the field of culture.

Table 6. The BERT-BiLSTM-Att model results are compared

Model	P%	R%	F1%
CNN-Att	63.72	60.85	62.25
SVM	75.01	70.01	72.42
Word2vec-BiLSTM	85.03	83.95	84.49
BERT-BiLSTM	88.02	86.23	87.11
BERT-BiLSTM-Att	93.03	91.12	92.07

3.3. Construction of Knowledge Maps of Chinese Traditional Music Culture

The application system of knowledge mapping of music-based traditional music culture describes the formalized features of music-based non-heritage in a more detailed way, and the information of music-based traditional music culture projects is displayed visually. While the official website of China Intangible Cultural Heritage and Baidu only display simple basic information such as heritage categories and declaration regions in the search interface, and most of the important information is hidden in the long text after the link jumps, compared with them, the knowledge framework of music-based NRM constructed in this paper is more detailed and complete as shown in Fig. 5, which can deeply reflect the external information characteristics of music-based NRM.

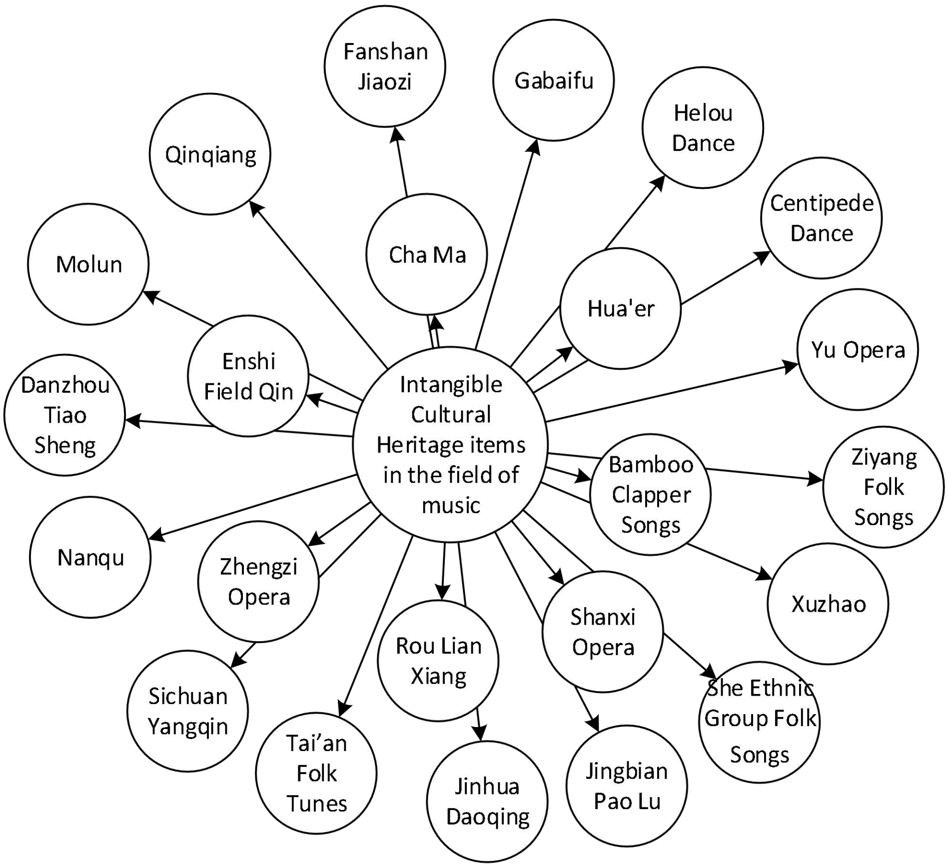

Figure 5. Partial visualization of music intangible cultural heritage items

4. The inheritance strategy of higher vocational vocal music teaching by integrating knowledge mapping

4.1. The construction of higher vocational vocal music teaching mode integrating knowledge mapping

4.1.1. Developing a systematic teaching and learning framework

Analyzed from the perspective of knowledge classification, the knowledge map of the vocal art instruction course can be divided into three layers: the basic layer, the application layer and the expansion layer. The basic layer should contain core concepts such as vocal principles, music theory knowledge and sight-singing and ear training, which are the foundation of the course; the application layer should focus on practical skills such as music analysis, singing processing and accompaniment, which serves as a bridge from theory to actual operation; the expansion layer should contain the historical background, genre and interdisciplinary connection of musical works, which broaden the students' academic horizons. In the process of integration, teachers should emphasize the vertical and horizontal correlation between knowledge points. At the vertical level, teachers need to clarify the role of the core concepts of the basic level in supporting the practical skills of the applied level, such as clarifying the specific influence of breath control on the emotional expression of musical works in the vocal skills; at the horizontal level, it is necessary to build a linkage mechanism between different knowledge modules, such as the music styles of a specific period of time are not only related to the historical background, but also have a direct impact on the selection of singing tones and the way of stage performance.

4.1.2. Implementation of online and offline co-teaching

The core of the blended teaching mode of college music course integrating knowledge mapping in the vocal art instruction course is to break through the single form of traditional teaching which is mainly based on offline classroom and supplemented by online resources, and to implement the synergistic teaching of "online independent learning + offline interactive teaching". The online independent learning link should take the knowledge map as the core carrier, assuming the dual functions of systematic knowledge transfer and open independent inquiry; the offline interactive teaching link should focus on practical training and in-depth interaction to promote the internalization of knowledge migration and innovative application. These two links have complementary functions and are connected back and forth, which together constitute a complete teaching closed loop of "input - internalization - output".

In the online independent learning session, teachers can design hierarchical learning resources and learning tasks with the help of knowledge mapping. Students can start music learning from different levels according to their own foundation. In the offline interactive teaching session, teachers can focus on the key contents, difficult parts and practical skills modules in the knowledge map, giving full play to the guiding role of teachers and expanding the depth and breadth of students' participation.

4.1.3. Optimization of hybrid teaching models

Teachers can strengthen the guidance and incentive mechanism for independent learning to help students improve their interest and learning efficiency. In view of some students' lack of proficiency in the use of knowledge mapping, teachers can use online platforms to publish a list of learning tasks based on knowledge mapping, and make clear the learning objectives and requirements of tasks at each stage. At the same time, teachers can establish a course point reward system to reward students who actively use knowledge mapping to carry out learning, share high-quality insights in online discussion forums, and excel in offline practical tasks, and incorporate them into the course evaluation system, so as to stimulate students' motivation and initiative in learning.

4.2. Analysis of Teaching Effectiveness

In order to enhance the homogeneity of the subjects in the experimental group and the control group, the subjects were free to participate in the experiment as students of the same major in the first semester of the first year of a higher vocational school, the subjects were randomly divided into the experimental group and the test group, 50 people in the experimental age 17-21, the average age of 18.43±0.91, and 50 people in the control group (32 boys, 19 girls, boys accounted for the control group number of 62.745%), aged 17-23, average age 18.45±1.12, in which the experimental group applied the vocal music teaching mode of this paper incorporating Chinese traditional music culture, and the control class still used the traditional teaching mode.

4.2.1. Data Analysis of Comparative Experiments

The results of the comparison between the pre-test and post-test of the control class and the experimental class are shown in Table 7. From the pre-test statistics, it was found that the mean total score of the control class was 22.04 and the mean total score of the experimental class was 21.99, and the comparison of the two samples revealed that the scores gained were not too high and there was not a big difference between the two samples. From the post-test data, it was found that the mean total score of the students in the control class was 24.22 and the mean total score of the students in the experimental class was 28.7, and the scores of the students in the two classes were improved, and the improvement of the experimental class was significantly higher than that of the control class. In terms of the three ability points, the experimental class showed more significant improvement in independent singing ability.

Table 7. The previous test and post-test descriptive statistics were measured

	Class	**df**	**Pre-test(M)**	**Post-test(M)**
Timeliness	Control class	49	3.00	3.40
	Experimental class	49	2.98	4.00
Strength	Control class	49	1.76	2.17
	Experimental class	49	1.78	2.48
Rhythm	Control class	49	3.4	3.9
	Experimental class	49	3.08	4.36
Timbre	Control class	49	2.16	2.08
	Experimental class	49	2.07	2.12
Interval	Control class	49	2.29	2.36
	Experimental class	49	2.29	2.88
Melody	Control class	49	2.36	2.67
	Experimental class	49	2.58	3.12
Traditional music knowledge	Control class	49	1.33	1.3

	Experimental class	49	1.22	1.94
Loud	Control class	49	1.83	1.94
	Experimental class	49	1.91	2.25
Low voice	Control class	49	0.94	1.02
	Experimental class	49	0.95	1.21
Affective expression	Control class	49	3.03	3.34
	Experimental class	49	3.16	3.64
Independent singing ability	Control class	49	10.33	11.63
	Experimental class	49	9.97	12.97
Traditional music quality	Control class	49	5.93	6.4
	Experimental class	49	6.09	7.88
Chorus ability	Control class	49	5.73	6.25
	Experimental class	49	5.92	5.95
Total score	Control class	49	22.06	24.22
	Experimental class	49	21.99	28.07

4.2.2. Results of pre- and post-test variability tests for control classes

The paper conducted a two-sample equal variance hypothesis test, i.e., P (T ≤ t) two-tailed test for the control class The results of the pre- and post-test variability test for the control class are shown in Table 8. Comparing the pre- and post-test means, the post-test scores were higher than the pre-test scores. According to the data analysis division <0.05 is a significant difference between the two samples, >0.05 difference is not significant, <0.01 is particularly significant. The table shows that the difference between the pre and post-test scores for pitch, intensity, rhythm, and melody is significant. Using traditional music teaching methods, students' independent singing ability, traditional music cultural literacy, and polyphonic choral ability all improved, of which independent singing ability improved especially significantly, traditional music cultural literacy, and polyphonic choral ability before and after the difference is not significant.

Table 8. Survey table for differences between comparison and comparison

	Pre-test(M)	**Post-test(M)**	**t Stet**	**P(T≤t)**
Timeliness	3.00	3.40	-2.22	0.05
Strength	1.76	2.17	-2.85	0.00
Rhythm	3.4	3.9	-2.43	0.05
Timbre	2.16	2.08	0.36	0.85
Interval	2.16	2.08	-0.39	0.72
Melody	2.36	2.67	-2.41	0.05
Traditional music knowledge	1.33	1.3	-0.15	0.91
Loud	1.83	1.94	-0.87	0.42
Low voice	0.94	1.02	-0.22	0.15
Affective expression	3.03	3.34	-1.62	0.05
Independent singing ability	10.33	11.63	-2.52	0.02
Traditional music quality	5.93	6.4	-1.23	0.025
Chorus ability	5.73	6.25	-1.58	0.15

4.2.3. Results of pre- and post-test variability tests for experimental classes

In this paper, the two-sample equal variance hypothesis test, i.e., P (T ≤ t) two-tailed test, was conducted on the experimental class, and the pre- and post-test difference tests of the experimental class are shown in Table 9. Comparing the pre- and post-test means, the post-test scores are all higher than the pre-test scores. Among them, only the timbre part of the pre- and post-test differences were not obvious (P>0.05), the emotional expression pre- and post-test differences were obvious (P<0.05), and the differences between pre- and post-tests for the rest of the parts were particularly obvious (P<0.01). After the intervention of the experimental class using the vocal pedagogy incorporating traditional music and cultural knowledge mapping, the students' independent singing ability, traditional music and cultural literacy, and polyphonic choral ability were significantly improved.

Table 9. Test of survey of differences before and after the experiment

	Pre-test(M)	**Post-test(M)**	**t Stet**	**P(T≤t)**
Timeliness	3.40	4.00	-4.68	0.00
Strength	2.17	2.48	-5.49	0.00
Rhythm	3.9	4.36	-5.58	0.00
Timbre	2.08	2.12	-3.74	0.58
Interval	2.08	2.88	-5.05	0.00

Melody	2.67	3.12	-3.78	0.00
Traditional music knowledge	1.3	1.94	-3.19	0.00
Loud	1.94	2.25	-3.69	0.00
Low voice	1.02	1.21	-4.74	0.02
Affective expression	3.34	3.64	-4.37	0.04
Independent singing ability	11.63	12.97	-4.04	0.00
Traditional music quality	6.4	7.88	-3.26	0.00
Chorus ability	6.25	7.95	-2.3	0.00

4.2.4. Results of the post-test test of variance for the experimental class in the control class

In this paper, a two-tailed P (T ≤ t) test was conducted on the posttest scores of the control and experimental classes using a two-sample equal variance hypothesis test. The posttest difference test for the control class and the experimental class, as shown in Table 10, found that the differences in the pitch, intensity, rhythm, high voice part, and low voice part were significant, i.e., the experimental class improved more significantly than the control class. The differences in the mastery of intervals, melody, and knowledge of traditional music culture were particularly significant, i.e., the experimental class improved more significantly than the control class. The differences in tone color and emotional expression were not significant, i.e., the two classes improved more significantly than the control class. In terms of the three ability points, the experimental class progressed more significantly than the control class ($P<0.05$), with independent singing ability and music perception ability being particularly significant ($P<0.01$). The above data shows that after 10 weeks of teaching experiment, the experimental class of vocal music teaching incorporating traditional music and cultural knowledge mapping significantly outperformed the control class of traditional music teaching in the development and improvement of the three ability points.

Table 10. Survey of differential survey of comparison class experiment

	Control class	**Experimental class**	**t Stet**	**P(T≤t)**
Timeliness	2.98	4.00	-2.71	0.01
Strength	1.78	2.48	-3.91	0.01
Rhythm	3.08	4.36	-2.64	0.01
Timbre	2.07	2.12	-0.12	0.75
Interval	2.29	2.88	-3.38	0.00
Melody	2.58	3.12	-2.48	0.02
Traditional music knowledge	1.22	1.94	-3.78	0.00
Loud	1.91	2.25	-2.95	0.00
Low voice	0.95	1.21	-2.79	0.05
Affective expression	3.16	3.64	-2.3	0.15
Independent singing ability	9.97	12.97	-3.08	0.00
Traditional music quality	6.09	7.88	-2.46	0.00
Chorus ability	5.92	7.95	-3.63	0.00

5. Conclusion

Since the 21st century, the inheritance and development of excellent traditional culture has become a very popular social topic. There are many ways to realize the inheritance and development of traditional culture, and the most direct and effective way is to integrate, inherit and develop traditional culture in teaching. This paper builds a knowledge map of traditional Chinese music culture and integrates it into the teaching of vocals in higher vocational education, constructing an online and offline integrated mixed vocals teaching mode in higher vocational education, and realizing the promotion of traditional Chinese music culture knowledge in teaching. The F1 values of the BERT-BILSTM-CRF model named entity recognition model and the BERT-BiLSTM-Att relationship extraction model used in the construction of the knowledge graph are 96.48, 92.07%, respectively, with good performance. It was found through the teaching experiments that after the intervention of vocal music teaching method integrating traditional music and cultural knowledge mapping, the students' independent singing ability, traditional music and cultural literacy, and polyphonic choral ability were significantly improved and better than the traditional teaching method. Therefore, the application of traditional music culture in higher vocational vocal music teaching should not only innovate the teaching concept and optimize the teaching method, but also pay attention to the students' subjective initiative in the vocal music classroom, so that the students can improve their music perception under the cultivation and guidance of traditional music culture, and subconsciously inherit and carry forward the traditional music culture.

About the Author
Shifang Yang, female, born in November 1984, postgraduate, lecturer. She graduated from the School of Music, Southwest University, majoring in Musicology. She is currently working at Chongqing Youth Vocational & Technical College, with her main research interests in vocal music performance and teaching research. She mainly teaches courses including Basic Vocal Music, Music Theory, and Solfège & Ear Training.

References

1. Li, Y. (2024). The Formation and Future Direction of Chinese Vocal Music in 100 Years of Development History. Pakistan Journal of Life & Social Sciences, 22(1).
2. Yang, Y. (2025). Innovative approaches and reformation strategies for vocal pedagogy in higher vocational music education. Journal of Humanities, Arts and Social Science, 9(4).
3. Yin, W. (2024). Innovations and practical exploration of vocal music teaching models in vocational colleges. Journal of Modern Educational Theory and Practice, 1(2).
4. Li, K. (2025). The path of vocal teaching in music education to cultivate students' aesthetic ability. International Journal of Educational Innovation and Science, 6(1), 31-39.
5. Wu, J. (2019, June). Systematic discussion and rational thinking of vocal music education under the current social background of China. In 2nd International Seminar on Education Research and Social Science (ISERSS 2019) (pp. 262-265). Atlantis Press.
6. Tiantong, G., & Sondhiratna, T. (2024). The Rapid Change of Singing Technique of Dong Folk Songs. Journal of Modern Learning Development, 9(7), 660-670.
7. Schippers, H. (2015). Applied ethnomusicology and intangible cultural heritage. The Oxford handbook of applied ethnomusicology, 134-156.
8. Lim, C. K. N., & Zhang, M. (2023). Chinese national music platformisation: A systematic review. Heliyon, 9(11).
9. Yunfan, Y., Charoensloong, T., & Yukolthonwong, S. (2024). The Influence of Western Music on Chinese Music and the Current Situation of Chinese Vocal Music in China. Journal of Modern Learning Development, 9(4), 586-594.
10. Julia, J., Iswara, P. D., & Supriyadi, T. (2019). Redesigning and implementing traditional musical instrument in integrated technology classroom. International Journal of Emerging Technologies in Learning (Online), 14(10), 75.
11. Ocariz, H. X. V., & Godall, P. (2018). Traditional music in the classroom: contributions for meaningful learning at school: Música tradicional en el aula: aportaciones para un aprendizaje significativo en la escuela. Revista Electronica de LEEME, (41), 16-34.
12. Murillo, R. E. (2017). The 21st Century Elementary Music Classroom and the Digital Music Curriculum: A Synergism of Technology and Traditional Pedagogy. Texas Music Education Research, 14, 27.
13. Lu, J. (2023, August). Research and Analysis of Vocal Music Teaching Practice in Universities Based on the Concept. In Proceedings of the 4th International Conference on Language, Art and Cultural Exchange (ICLACE 2023) (p. 283). Springer Nature.
14. Nian, L., & Wang, F. (2017). On the importance of emotional cultivation in vocal music teaching. International Journal of Technology, Management.
15. Sabina, V., & Senad, K. (2021). Cognitive-emotional music listening paradigm in professional music education. International journal of cognitive research in science, engineering and education, 9(1), 135-145.

16. Ovcharenko, N., Samoilenko, O., Moskva, O., & Chebotarenko, O. (2020). Innovative technologies in vocal training: Technological culture formation of future musical art teachers. Journal of History Culture and Art Research, 9(3), 115-126.
17. Özdek, A. (2015). The role of folk music as cultural heritage in the curriculum of vocational high schools of music: Pattern of Azerbaijan-Turkey. Anadolu Journal of Educational Sciences International, 5(3), 84-112.
18. Sun, L. (2021). A probe into the integration of traditional music culture in vocal music teaching in colleges and universities. Region-Educational Research and Reviews, 3(2), 65.
19. Kalyuzhnaya, V. P. (2025). “Folk Music Culture” as an Academic Discipline in the System of Secondary Vocational Education: Problems and Perspectives. Russian Musicology, (3), 148-156.
20. Zhang, P. (2020, November). The Significance and Ways of Constructing the Northwest Traditional Music Curriculum System in Lanzhou Vocational Colleges. In International Conference on Education Studies: Experience and Innovation (ICESEI 2020) (pp. 582-586). Atlantis Press.
21. Li, K. (2024). Innovative Strategies in Teaching Vocal Music for Higher Vocational Music Performance Majors. Art and Performance Letters, 5(2), 146-152.