[1] Wei Zhou

[1] Xueying Liu*

Application of LSTM Feature Extraction Model in the Construction of Opera Audio Database

Abstract: - Chinese traditional music plays a greater role in ideal society governed by rites and rituals. The musical culture of China is both ancient and modern. Throughout history, musicians have created several musical genres and used countless instruments. In order to appreciate Chinese traditional music, it is helpful to have a basic familiarity with three "keynotes": the classical traditional music of the elites; folk and ethnic music, and the most popular traditional Chinese instruments. In general, an opera audio database consists of audio recordings of opera performances. These recordings may be utilized for a variety of reasons, including research, education, and entertainment. To establish such a database, pertinent features must be extracted from the audio recordings to allow for effective storage, retrieval, and analysis of the data. Improvements in audio signal processing over the last several years have qualified for creating feature extraction methods that improve the quality and accuracy of the retrieved features. We first gather the audio recordings of opera performances using an appropriate recording device to build an opera audio database. Data preprocessing then eliminates noise and other undesired signals from the audio recordings to enhance the data quality. After that, to extract useful features from the audio recordings that may be utilized to define the sound's rates, the Long Short Term Memory (LSTM) technique is used. The components must then be organized and stored in a database for fast retrieval and analysis once extracted. A suitable database management system may be used to build the database. The database should be created to make it simple to search and filter the data using different criteria. The study results stated that building an opera audio database using novel feature extraction approaches is a challenging process that requires an understanding of opera music, database administration, and audio signal processing. The experimental results stated that the proposed model has provided an accuracy of 96%.

Keywords: Chinese traditional music, Opera audio database, audio recordings, feature extraction, Long Short Term Memory (LSTM).

1. Introduction

Chinese opera included songs and dances that were based on folk stories and were regarded as ancient forms of entertainment. It features rap, burlesque, folk music, dances, and other forms of expression that show uniqueness in commonality. It also includes performing arts, martial arts, acrobatics, music, dance, and literature. Opera is regarded among the three most ancient dramatic civilizations in history, with Greek comedy and tragedy, as well as Indian Sanskrit opera. People with dementia may have an emotional resonance when their long-term memories are stimulated, which heightens their interest in life and improves compliance. According to several kinds of research, traditional music therapy helps dementia patients' cognitive function [1]. The interaction of singing, dance, and music is opera's most distinctive element. Dancing, lyrics, music, literature, and drama are just a few of the many possible shapes that dance drama may take. It is also an extensive art, which is another one of its unique cultural qualities. People were deeply affected by certain mannerisms in opera, such as the exquisite arcs created by flinging one's arms. They get a strong feeling of color and distinct regional traits from opera clothes. Opera expresses characters via twisted and magnified visuals, which is immensely metaphorical [2].

A single sound produced in traditional Chinese music for a thousand years is the foundation. National musical instruments and unique performance designs are significant components of traditional Chinese music. The development of song, dance, and instrumental folklore, which provides the vibrant and unique foundation for the majority of Chinese musical compositions, is what gave birth to the national oddity of Chinese music. Traditional music is continually evolving and rising in popularity. Hence, when the younger generation enrolls in training, traditional Chinese culture is created and developed [3]. A musical aesthetic is an artistic feeling brought on by listening to music. It includes commonplace emotions from everyday life and unique emotion from musical art. The only automated categorization and calculation of common sentiment is the exclusive subject of traditional music emotion research [4]. Figure 1 shows the flowchart of the proposed method.

Opera is a modern art form that combines vocal music and drama. Yet, its production and performance are fundamentally based on the classic verbal music form known as opera performing art. This is unquestionably a

[1]Minjiang University，Tsai Chi-Kun Acadamy of Music,China,4350108 mjliuxy@163.com

demanding standard for vocal music performers. It is an effective method for raising vocalists' overall performance quality [5]. Many Chinese immigrants, including those from the Philippines, introduced China's music culture, particularly the diverse traditional music skills in Fujian and Guangdong, into the four regions of South Asia as recently as the middle of the 19th century. The South Asian Chinese developed a strong spiritual life through these musical skills [6]. An opera database collects information about operas, including composers, librettists, performers, productions, recordings, and other relevant details. An opera database can be a valuable resource for opera enthusiasts, scholars, and performers, as it can provide a comprehensive and organized repository of information about the vast and complex world of opera [7]. In this paper, we proposed LSTM feature extraction method for constructing an opera audio database.

The rest of the paper is as follows: Section 2 summarizes related works, Section 3 gives the proposed method, Section 4 provides the results and discussion, and Section 5 summarizes the conclusion.

2. Importance of opera audio

The research put out a systematic framework for Chinese Heritage and culture computing, composed of three conceptual levels: a Chinese CH preservation and development plan, a computing methodology, and a computable cultural ecology. Three parts make up the computing process: data collection and analysis, digital modeling and database creation, and data use and promotion. The lessons show how to use computer techniques for various stages of Chinese cultural heritage maintenance and growth, from digital inheritance and conservation to display and marketing [8]. An efficient semi-supervised learning (SSL) approach that utilizes the mean teacher method is presented in the study. To incorporate perturbations for various convolutional layers, a novel independent component (IC) module that combines batch normalization and drop block operations is specifically suggested [9]. The accuracy of speech emotion detection is discussed in the study, along with the effects of the classification strategy and the best feature and data augmentation combinations. In terms of classification, the suggested framework, a 1D convolutional neural network (1D CNN), and conventional machine learning methods [10].

The study [11] suggested a voice emotion recognition communication system. The system uses a sound data enhancing approach to pre-process audio for speech emotion identification and transforms the sound using MFCC into a spectrogram (Mel Frequency Cepstral Coefficient). The five emotions—peace, happiness, sad, anger, and fear—are then applied using CNN's (Convolutional Neural Network). The article suggested the creation of a multimedia database of Chinese musical instruments. The database contains, for each musical instrument, written descriptions, photographs, audio clips of playing methods, music clips, video clips of the making of the component and the recording process, and acoustic analysis resources. The motivation for the database's creation and its selection criteria are explained in depth [12]. The paper presented a state-of-the-art audio alignment method for complete recordings of a Mozart opera based on online Dynamic Time-Warping (ODTW). Additionally, the paper analyzed the tracker's most severe errors to identify three familiar sources of problems specific to the opera scenario. The paper proposes a mixture of a DTW-based music tracking system with specialized audio event detectors that affect the DTW algorithm in a leading fashion. [13]. The authors of [14] suggested a data augmentation approach that uses a vocoder-based speech synthesizer to transform genuine speech into a singing voice. The model-based style transfer is capable of producing singing voices of a high quality, which makes it possible to convert vast corpora of actual speech into singing voices that may be used in constructing an End to End (E2E) lyrics transcribing system. The study suggests developing multimodal virtual 3D models to promote and foster new chances to exploit and transmit cultural heritage for the sake of the Cultural and Creative Industries (CCI). The work provides the construction of multisensory environments associated with three major European theatres that are seen as research articles [15]. The paper proposed clothing artistic modeling based on a clustering approach for large amounts of data. The method being offered creates a database for the clustering of big data by first designing and building the set of attributes of the major data function sequential training set and then simultaneously establishing a second-order cone optimization method to rectify the big data. In doing so, the method is able to provide a database for the cluster analysis of big data [16].

3. Proposed method

In this section, we describe the LSTM feature extraction model in constructing an opera audio database. An LSTM (Long Short-Term Memory) model can be used for feature extraction in constructing an opera audio database. The basic idea behind using an LSTM model is that it can effectively capture long-term dependencies in the audio signals, which is particularly important in the context of opera music, where different musical themes and motifs may be repeated over more extended periods. Once the LSTM model is trained, it can extract features from the pre-processed audio data. These features may include time-series data such as tempo, rhythm, and melody. Figure 1 depicts the flow of the proposed methodology.

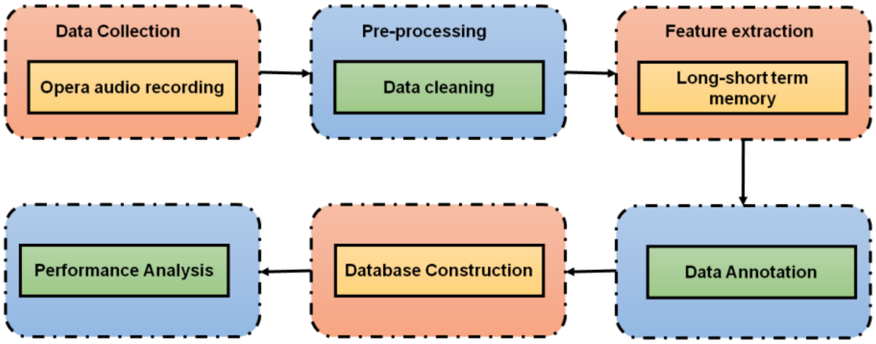

Figure 1: Flowchart of Proposed method

3.1 Dataset

Data from a specified dataset, labeled with several Cantonese opera singing genres, was utilized for the investigations in this research. The variety and popularity of opera singing genres affected the kind of data labels employed. The ten types of singing genres established by Hong Xiangnv, Luo Jiabao, Ma Shizeng, and others are combined in this work and labeled He, Bai, Chen, LPC, Xue, Luo, Hong, Ma, Deng, and Gui, respectively. For each opera singing type, audio was gathered from various settings, eras, and occasions. The data was utilized to minimize the effects of the audio quality overtime on the categorization effect and included as many typical opera vocal performances as feasible. As a full opera performance often necessitates the participation of several performers, constant singing portions from each kind of singing must be gathered as reliable experimental data. To ensure the proportion of the data, several opera pieces of varied durations but with the same overall duration were chosen from a comprehensive collection of opera recordings. The experimental data included 1000 cantos from 305 operas in WAV format, every having duration of 30 seconds and some examples frequency of 44,100 Hz [17]. Table 1 shows the information on the dataset.

Table 1: Dataset description

Number	Class	Labels	Operas Number	Number of Files (.Wav)
1	Bai Jurong	Bai	28	100
2	Chen Xiaofeng	Chen	32	100
3	Luo Pinchao	Lpc	29	100
4	Xue Juexian	Xue	30	100
5	Gui Mingyang	Gui	29	100
6	Luo Jiabao	Luo	36	100
7	Hong Xiannv	Hong	34	100
8	Ma Shizeng	Ma	35	100
9	Deng Zhiju	Deng	30	100
10	He Feifan	He	34	100

3.2 Data preprocessing

Data cleaning for opera audio involves steps to ensure the data is accurate, complete, and reliable. The first step in constructing an opera audio database is to gather high-quality audio recordings of operatic works, arias, and singers. It is essential to ensure that the recordings are of sufficient quality to accurately capture the nuances and subtleties of the music and the singers' performances. Opera audio recordings may contain background noise, such as coughing, rustling, or applause. It is essential to remove this noise to ensure the database is clean and accurate. This can be done using noise reduction techniques like filtering and spectral subtraction. It is common for different recordings of the same opera or aria to exist. Duplicates can be removed to ensure that the database is not bloated and that the analysis is focused on unique recordings.

3.3 Opera audio

An opera is a musical theater that combines singing, acting, and instrumental music. As such, opera audio refers to the sound produced during an opera performance, including the singing voices, orchestra, and any sound effects or recorded music that may be used. In opera, the singers are typically trained to project their voices over the orchestra and other instruments, so the vocals are often quite powerful and expressive. The orchestra, meanwhile, plays a crucial role in setting the mood and tone of the piece and can range from a small chamber ensemble to a full symphony orchestra.

In terms of audio quality, the sound of an opera performance can vary widely depending on the venue, the quality of the recording equipment, and the expertise of the sound engineers involved. However, with modern technology, it is possible to capture the full range and depth of opera performance and reproduce it in high-fidelity audio recordings. Pitch is defined explicitly by the vocal body's vibration frequency, which also indicates the pitch level. Also, different pitches inherently transmit various emotions. In particular, high-frequency tones may convey feelings like enthusiasm, haste, and brightness. Low-frequency tones, on the other hand, represent negative feelings like procrastination, despair, and decadence. In addition, different instruments have different pitch ranges.

3.4 Feature extraction using Long Short-Term Memory (LSTM)

The generated audio samples may need to be post-processed to remove noise or unwanted artifacts, and to adjust the volume and other parameters to make them sound more like authentic opera recordings. The temporal properties of the input vector are considered by LSTM, a kind of temporally recurrent neural network (RNN), to retrieve temporal contextual data. When compared to RNN in terms of structure, LSTM, a section of RNN, adds storage blocks within every neuron, namely input gates, output gates, and forgetting gates. To extract features from an LSTM model, you can use the output of one of the hidden layers as input to another machine learning model, such as a feedforward neural network or a support vector machine. The hidden layers of an LSTM model contain learned representations of the input sequence that capture relevant patterns and correlations. To construct an opera audio database using LSTM, the first step is to collect the audio data of the opera performances. Audio data can be collected from various sources, such as commercial recordings, live performances, and radio broadcasts. Once the audio data is collected, it must be preprocessed and segmented into smaller chunks or frames of fixed duration. Next, the segmented audio data can be used to train an LSTM model to recognize the different components of the opera, such as arias, recitatives, and choruses. The LSTM model can be prepared using a supervised learning approach. The input to the model is the segmented audio data, and the output is a label that indicates the type of the component (e.g., aria, recitative, or chorus). After training the LSTM model, it can automatically label the audio data in the opera audio database. The labeled data can be used for various applications such as music information retrieval, automatic playlist generation, and content-based recommendation.

It is important to note that constructing an opera audio database using LSTM requires many labeled data and computational resources. The quality of the labeled data and the design of the LSTM model can also have a significant impact on the performance of the system. Alternatively, you can use the LSTM model as a feature extractor by removing the output layer and using the output of the last hidden layer as input to another model. This approach is known as transfer learning, where you use the learned representations from a pre-trained model to improve the performance of another related task. When mappings input and output while taking contextual information into account, LSTM not only inherited the benefits of RNN but also addresses several issues like "long-range reliance" and "gradient disappearance" while generating lengthy segments. These issues are common challenges that RNNs have when trying to recall information over extended periods while calculating implicit contextual layers. The following equations from 1 to 2 illustrate the LSTM structure.

$$e_s = \sigma(U_e[g_{s-1}, v_s] + a_e) \quad (1)$$

$$j_s = \sigma(U_j[g_{s-1}, v_s] + a_j) \quad (2)$$

$$\widetilde{D}_s = tanh(U_{\widetilde{D}}[g_{s-1}, v_s] + a_{\widetilde{D}}) \quad (3)$$

$$D_s = e_s * D_{s-1} + j_s * \widetilde{D}_s \quad (4)$$

$$p_s = \sigma(U_p[g_{s-1}, v_s] + a_p) \quad (5)$$

$$g_s = p_s * tanh(D_s) \quad (6)$$

When v_s serves as the input sequence, combined with the hidden layer ht1's state above, it creates the forget gate ft through the activation function. U_e,U_j, $U_{\widetilde{D}}$, U_p and a_e , a_j , $a_{\widetilde{D}}$, and a_p stand for the weight parameters and biases, respectively. Moreover, $g_{s-1,}$ and v_s are used to compute input gate it and output gate ot. To decide whether it should discard the message, the forget gate ft and the pre-unit state D_{s-1} are combined.

3.5 Data Annotation

Data annotation for an opera audio database would typically involve identifying and labeling various aspects of the audio files, such as the composer, the name of the opera, the performers, the genre, and other relevant metadata. Spoken roles in opera are comparatively uncommon, and the proportion varies depending on the opera. Opera is a unique genre of music with a wide range of musical expressions. As a result, we have finished annotating the data for the operas in this area. We observed that the non-silence opera segment is composed of three distinct sections: pure music, song; and spoken. Only musical instruments may be used to create "pure music." Typically, these passages are included at the start and finish of each opera. Combining musical instruments with vocals creates a "song." For the majority of operas, this section has the longest running time. The opera's "Speech" section, or purely vocal section, comes last. Spoken roles in opera are comparatively uncommon, and the proportion varies depending on the opera.

3.6 Database Construction

Constructing an opera database involves organizing and compiling information about operas, including composers, librettists, performances, recordings, and other relevant details. Determine what types of opera audio you want to include in the database. For example, you might include full operas, individual arias, recitatives, and other pieces. Once you have determined what to include, collect audio from various sources such as professional opera recordings, live performances, and amateur recordings. You might also want to consider collaborating with opera companies and performers to obtain high-quality audio recordings. To ensure high-quality audio, you might need to clean and process the audio. This could include removing noise, adjusting levels, and applying equalization. Add metadata to the audio files, such as the title, composer, performer, date, and other relevant information. This will make it easier for users to search and sort through the database.

4. Result and discussion

The results of using the LSTM feature extraction model to build the opera database's audio are presented in this section. The parameters include equalisation, compression, spatialization, accuracy, and precision. The existing used techniques include Convolutional Neural Network based Long Short Term Memory CNN-LSTM [18], Bidirectional based Long Short Term Memory (Bi-LSTM) [19], and Optical Music Recognition (OMR) [20].

4.1 Accuracy

The accuracy of the Opera music is a complex matter that involves the interpretation of the music, the vocal or instrumental technique, and the ability to convey the emotional content of the music. Opera music, by its nature, is highly expressive and requires a high level of technical proficiency from the performers. The accuracy of the performance is pitch accuracy, rhythm accuracy, and dynamic accuracy. Figure 2 depicts the proposed and existing of the accuracy. Table 2 depicts the comparison of accuracy.

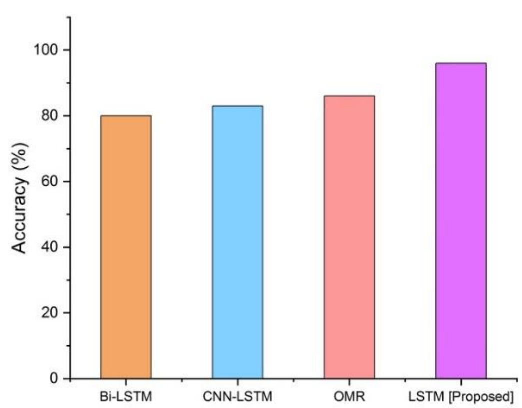

Figure 2: Accuracy of the proposed method and existing

Table 2: Comparison of accuracy

Techniques	Accuracy (%)
Bi-LSTM	80
CNN-LSTM	83
OMR	86
LSTM [Proposed]	96

4.2 Precision

The precision of a musical performance is influenced by many factors, including the skill and technique of the performer, the quality of the instrument or equipment being used, and the level of rehearsal and preparation that has gone into the performance. In the context of opera music, precision is especially important because of the complexity and intricacy of the music, as well as the fact that opera performances typically involve many performers, including singers, instrumentalists, and conductors, who must all work together seamlessly to create a cohesive and powerful musical experience. Achieving precision in an opera performance requires a high degree of skill, discipline, and coordination among all of the performers involved. Figure 3 shows the proposed and existing of precision. Table 3 shows the result of precision.

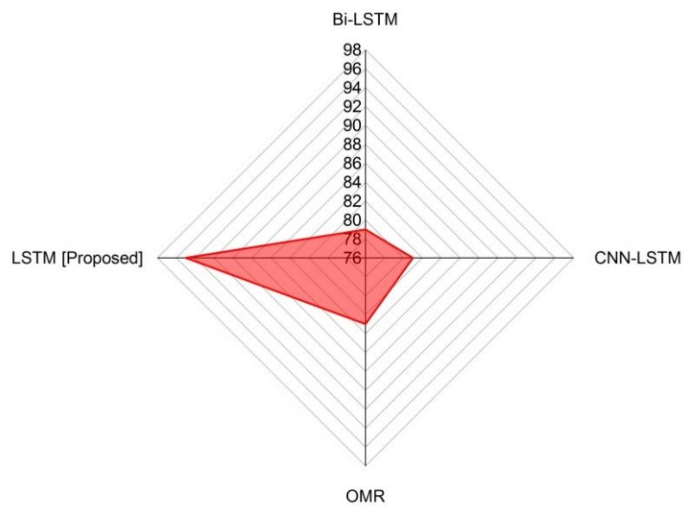

Figure 3: Precision of proposed and existing method

Table 3: Result of precision

Techniques	Precision (%)
Bi-LSTM	79
CNN-LSTM	81
OMR	83
LSTM [Proposed]	95

4.3 Compression

In music, compression refers to the process of reducing the dynamic range of a recording or live performance. Dynamic range refers to the difference between the loudest and quietest parts of a piece of music. Compression can help to make the quieter parts of a recording or performance more audible, while also preventing the louder parts from becoming too overpowering. Compression is often used in the recording and mixing process to create a more consistent and balanced sound. It can be applied to individual tracks or to the overall mix, and is often used to bring out specific elements of a recording, such as vocals or drums. Figure 4 denotes the proposed and existing method of compression. Table 4 denotes the comparison of compression.

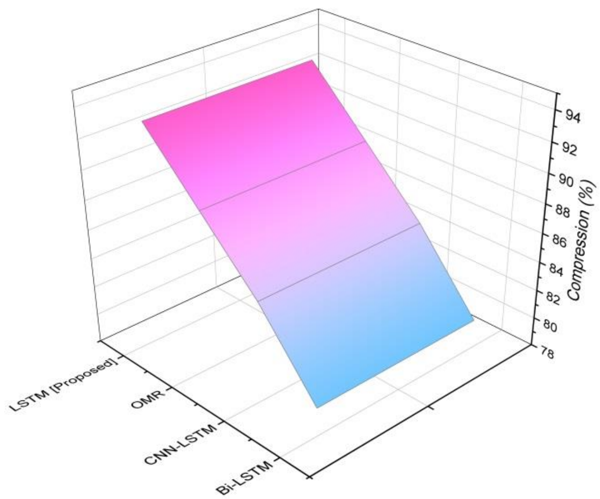

Figure 4: Compression of proposed and existing method

Table 4: comparison of compression

Techniques	Compression (%)
Bi-LSTM	80
CNN-LSTM	85
OMR	89
LSTM [Proposed]	93

4.4 Spatialization

Spatialization in audio refers to the process of creating a sense of space or dimensionality within a recorded or live sound. This can include creating the impression of sound sources moving through space, as well as creating a sense of depth and dimensionality within a recording or mix. Spatialization can be achieved through a variety of techniques, including the use of stereo or surround sound, reverb and other time-based effects, panning and level adjustments, and the use of specialized spatialization tools such as binaural or ambisonic processing. Figure 5 illustrates the proposed and existing of spatialization. Table 5 result of spatialization.

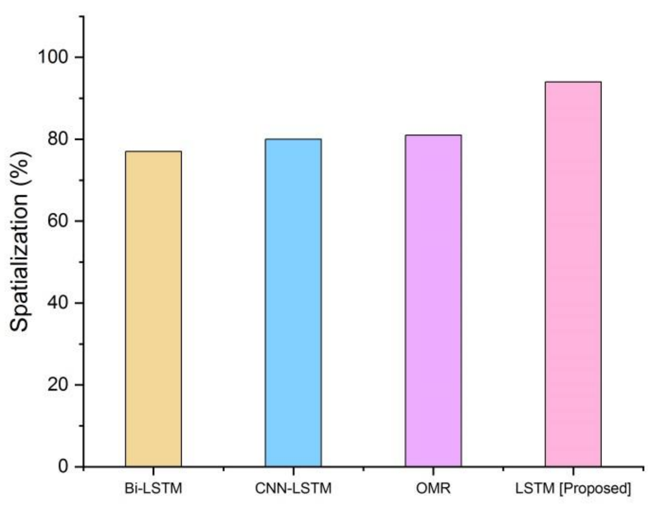

Figure 5: Spatialization of proposed and existing method

Table 5: Result of spatialization

Techniques	Spatialization (%)
Bi-LSTM	77
CNN-LSTM	80
OMR	81
LSTM [Proposed]	94

4.5 Equalization

Equalization can be an important tool in achieving a balanced and natural sound in opera music recordings and live performances. When applying EQ to opera music, it's important to consider the unique tonal characteristics of the human voice and the orchestral instruments typically used in opera productions.Skilled audio engineers and producers know how to use EQ effectively to enhance the overall sound of an opera recording or performance, while preserving the natural and dynamic qualities of the music. Figure 6 demonstrates the proposed and existing of equalization. Table 6 demonstrates the comparison equalization.

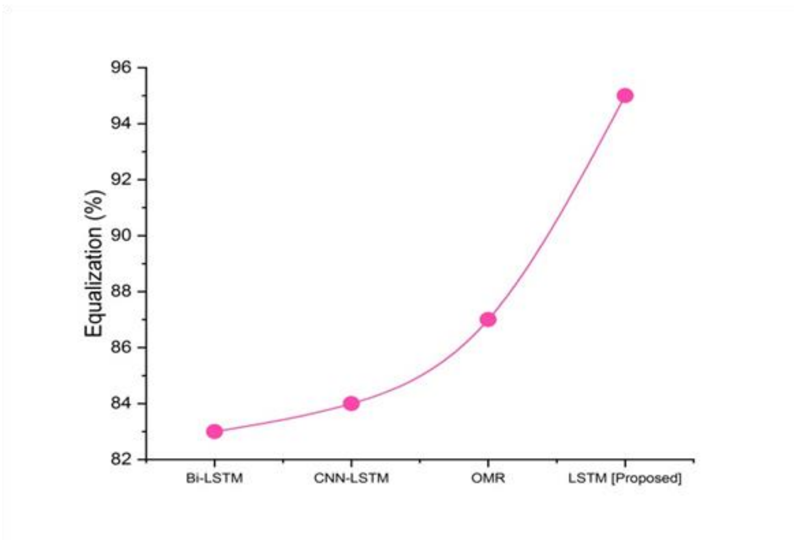

Figure 6: Equalization of proposed and existing method

Table 6: Comparison of equalization

Techniques	Equalization (%)
Bi-LSTM	83
CNN-LSTM	84
OMR	87
LSTM [Proposed]	95

5. Conclusion

Constructing an opera audio database using LSTM is a promising approach to developing a system that can accurately recognize and classify opera audio. LSTM is a type of neural network that is particularly well-suited for processing sequences of data, making it a good choice for analyzing opera audio, which is often composed of long, complex sequences of notes and sounds. By training an LSTM model on a large database of opera audio recordings, it is possible to develop a system that can accurately recognize different operatic works, arias, and singers. When compared the proposed method is more efficient than the existing methods. This system could be useful for a variety of applications, such as creating personalized playlists for opera fans, facilitating music education and research, and improving the accessibility of opera to a wider audience. However, constructing such a database requires a significant amount of resources, including high-quality audio recordings, metadata, and annotations. It also requires expertise in both opera and machine learning. Therefore, constructing an opera audio database using LSTM is a complex and challenging task, but it has the potential to yield significant benefits for the field of music and opera.

Acknowledgement：

The Ministry of Education's annual humanities and social science project：Research on the Inheritance Ecology of Min Opera from the Perspective of Art Anthropology(No.21YJC760118).

Reference

[1] A Chen, X., Li, D., Xu, H. and Hu, Z., 2020. Effect of traditional opera on older adults with dementia. Geriatric Nursing, 41(2), pp.118-123.
[2] Wang, Y. and Hu, X., 2020. Wuju opera cultural creative products and research on visual image under VR technology. IEEE Access, 8, pp.161862-161871.
[3] Lu, D., 2022. Inheritance and promotion of Chinese traditional music culture in college piano education. Heritage Science, 10(1), p.75.
[4] Xie, L. and Gao, Y., 2022. A database for aesthetic classification of Chinese traditional music. Cognitive Computation and Systems, 4(2), pp.197-204.
[5] Yao, L., Research on the inher
[6] itance of Chinese traditional vocal art in contemporary opera. Bai, X., 2022. Transnational Chinese Music in the Philippines: Chinese Traditional Music in Binondo Manila. Frontiers in Art Research, 4(5).
[7] Pavlica, L., The Operas Arrangements for Chamber Ensembles in the Moravian and Lower Austrian Monasteries.
[8] Li, M., Wang, Y. and Xu, Y.Q., 2022. Computing for Chinese cultural heritage. Visual Informatics, 6(1), pp.1-13.
[9] Zheng, X., Song, Y., Yan, J., Dai, L.R., McLoughlin, I. and Liu, L., 2020. An Effective Perturbation Based Semi-Supervised Learning Method for Sound Event Detection. In INTERSPEECH (pp. 841-845).
[10] Alnuaim, A.A., Zakariah, M., Alhadlaq, A., Shashidhar, C., Hatamleh, W.A., Tarazi, H., Shukla, P.K. and Ratna, R., 2022. Human-computer interaction with detection of speaker emotions using convolution neural networks. Computational Intelligence and Neuroscience, 2022.
[11] Lee, M.C., Chiang, S.Y., Yeh, S.C. and Wen, T.F., 2020. Study on emotion recognition and companion Chatbot using deep neural network. Multimedia Tools and Applications, 79, pp.19629-19657.
[12] Liang, X., Li, Z., Liu, J., Li, W., Zhu, J. and Han, B., 2019. Constructing a multimedia Chinese musical instrument database. In Proceedings of the 6th Conference on Sound and Music Technology (CSMT) Revised Selected Papers (pp. 53-60). Springer Singapore.
[13] Brazier, C. and Widmer, G., 2020. Towards reliable real-time opera tracking: combining alignment with audio event detectors to increase robustness. arXiv preprint arXiv:2006.11033.
[14] Basak, S., Agarwal, S., Ganapathy, S. and Takahashi, N., 2021, June. End-to-end lyrics recognition with voice to singing style transfer. In ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 266-270). IEEE.
[15] Bertocci, S., Lumini, A. and Cioli, F., 2022, July. Digital survey and 3D modeling to support the auralization and virtualization processes of three European theater halls: Berlin Konzerthaus, Lviv Opera House, and Teatro del Maggio Musicale in Florence. A methodological framework. In Proceedings of the 2nd Symposium: The Acoustics of Ancient Theatres, Verona, Italy (pp. 6-8).
[16] Luo, W., 2021. Analysis of artistic modeling of opera stage clothing based on big data clustering algorithm. Security and Communication Networks, 2021, pp.1-9.
[17] Chen, Q., Zhao, W., Wang, Q. and Zhao, Y., 2022. The sustainable development of intangible cultural heritage with AI: Cantonese opera singing genre classification based on CoGCNet model in China. Sustainability, 14(5), p.2923.
[18] Jin, C., Song, Z., Xu, J. and Gao, H., 2022. Attention-based Bi-DLSTM for sentiment analysis of Beijing opera lyrics. Wireless Communications and Mobile Computing, 2022.
[19] Wu, Y., Li, S., Yu, C., Lu, H., Weng, C., Zhang, L. and Yu, D., 2020. Peking opera synthesis via duration informed attention network. arXiv preprint arXiv:2008.03029.
[20] Chen, G.F. and Sheu, J.S., 2014. An optical music recognition system for traditional Chinese Kunqu Opera scores written in Gong-Che Notation. EURASIP Journal on Audio, Speech, and Music Processing, 2014, pp.1-12.