Yanlong Niu[1*]

Automated Classification of Chinese Traditional Music Genres Using Multi-Modal Knowledge Graph Convolutional Networks

Journal of Electrical Systems

Abstract: - In the digital music era, effective classification of music genres is crucial for organizing vast music databases. This study addresses the challenges of manual annotation by proposing an automated approach for classifying Chinese traditional music genres. Leveraging a MIDI dataset, the music excerpts undergo preprocessing using the Adaptive Multi-Scale Improved Differential Filter (AMSIDF) to enhance signal quality. Subsequently, distinctive features are extracted using the Synchro-Transient-Extracting Transform (STET), enriching the data with essential musical characteristics. The preprocessed and feature-enriched data is then fed into a Multi-Modal Knowledge Graph Convolutional Network (MKGCN) for classification. The MKGCN model is adept at integrating multi-modal data sources, enabling the fusion of audio features with metadata or textual information. Through knowledge graph convolutional layers, the MKGCN model captures intricate relationships within the data, facilitating accurate classification of Chinese traditional music genres. Then MKGCN optimized with Black Winged Kite Algorithm (BWKA) for accurate classification of Chinese traditional music such as Dance music, Metal, Rural, Classical, and Folk genres. This study achieves significant advancements in the automated classification of Chinese traditional music genres. The proposed Chinese Traditional Music Classification using Multi-Modal Knowledge Graph Convolutional Network (CTMC-MKGCN-BWKA) approach is implemented in Python. The performance of the proposed CTMC-MKGCN-BWKA approach attains 21.14%, 24.31% and 23.78% higher accuracy, 22.74%, 21.01% and 15.28% higher Precision and 22.14%, 24.01% and 22.71% higher Recall compared with existing methods such as Chinese Traditional Music Classification using Depp learning (CTMC-DL), Chinese Traditional Music Classification using Deep Neural Network (CTMC-DNN), Chinese Traditional Music Classification using Deep Belief Network (CTMC-DBN).

Keywords:Chinese traditional music, Multi-Modal Knowledge Graph Convolutional Network, music excerpts , Black Winged Kite Algorithm , Musical Instrument , deep learning, education

1. INTRODUCTION

In the contemporary digital music landscape, the effective classification of music genres stands as a crucial endeavor for efficiently organizing vast music databases [1]. Due to consumers' varied tastes and the expansion of digital music services,

accurate genre classification is indispensable for audience management, collection curation, and personalized music recommendation systems [2-4]. However, there are a lot of practical, scalability, and accuracy issues with manual music genre annotation. As a result, the need for automated approaches to classify music genres has become increasingly pressing in recent years [5-8].

In addressing the aforementioned challenges, various solutions have been proposed to automate the classification of music genres. These solutions typically leverage advanced techniques like ML and DL to analyze audio features and extract meaningful patterns for genre classification [9-11]. However, many existing approaches still suffer from limitations such as reliance on handcrafted features, lack of robustness to variations in audio signals, and difficulty in capturing complex relationships between different musical elements. Consequently, there remains a need for more sophisticated and effective methodologies to overcome these drawbacks and achieve accurate genre classification in digital music repositories [12-15].

The goal of this study is to overcome the shortcomings of current approaches and make notable improvements in the accuracy of genre classification by proposing a unique strategy for the automated categorization of Chinese traditional music genres [16-18]. Leveraging a MIDI dataset and advanced preprocessing techniques such as the AMSIDF and the STET, the approach enriches the data with essential musical characteristics. The preprocessed and feature-enriched data is then fed into a MKGCN, optimized with the Black Winged Kite Algorithm (BWKA), to capture intricate relationships within the data and facilitate accurate classification of Chinese traditional music genres. Through this innovative approach, a substantial contribution is made to the advancement of automated music genre classification, paving the way for enhanced organization and exploration of Chinese traditional music repositories [19, 20].

[1] [1] School of Music and Dance, Zhengzhou University of Science and Technology, Zhengzhou, Henan, 450064, China
[*]Corresponding author e-mail: niuyanlong_0512@163.com

Copyright © JES 2024 on-line : journal.esrgroups.org

The following succinctly describes the primary contributions of this paper:

- Introduction of a novel approach for automated classification of Chinese traditional music genres, leveraging advanced preprocessing techniques and multi-modal knowledge graph convolutional networks.
- Optimization of the classification model with the Black Winged Kite Algorithm (BWKA) to enhance classification accuracy.
- Implementation and evaluation of the proposed approach in Python, with comparative analysis against existing techniques to demonstrate its effectiveness and superiority in genre classification accuracy.
- The novelty lies in the integration of advanced preprocessing techniques, MKGCN, and optimization algorithms to achieve accurate classification of Chinese traditional music genres, addressing the limits of current approch and contributing to the advancement of automated music genre classification methodologies.

II. RECENT RESEARCH WORK: A BRIEF REVIEW

Several works were have presented previously in literatures were depending on the Chinese traditional music classification using deep learning. Few of them were mentioned here.

He [21] conducted a study looking at how popular digital music resources are becoming in the age of digital music. Music genres play a crucial role in music description. Music labels are essential for discovering and categorizing digital music resources. However, manual annotation for classification in a vast music database consumes significant time and resources, failing to meet contemporary demands. The paper was divided into various local Musical Instrument Digital Interface (MIDI) music passages, and the main study discoveries and advances were playing style analysis, extracting passage features, and sequencing passage features. This process involved extracting note feature matrices, identifying topics and section divisions depend on these matrices, exploring and extracting effective features based on segment themes, and composing feature sequences. Due to the limitations of standard classification methods, classifiers struggle to grasp temporal and semantic music information effectively.

Xu [22] investigated the emphasis on the value of everyday ideological and political education as well as the teaching of courses on political theory and ideology in higher education. Higher criteria have been set for this type of education due to the ongoing progress of information technology, the individualized growth of pupils, and the reform and creativity of political and ideological education. While several colleges have experimented with novel teaching approaches, they frequently disregarded the assessment aspect and their own advancement. Instead of focusing solely on injecting fresh ideas into educational reform and ideological education, they failed to prioritize enhancing their own quality. Consequently, these limitations led to unsatisfactory learning outcomes. This study focused on defining deep learning and applying it to the ideological and political education of college students in order to address these problems. It examined the ways in which deep learning presents novel learning models, intelligent teaching and assessment techniques, and fresh, precise, and customized ideas. A novel paradigm of intelligent integration including the subject, object, and mediator is represented by this method. By placing a strong emphasis on individualization, correctness, interactivity, and vividness in college students' ideological and political education, as well as improving assessment and management processes in this area, it seeks to further transform educational and teaching methodologies. This study focused on defining DL and applying it to the ideological and political education of college students in order to address these problems. It examined the ways in which DL presents novel learning models, intelligent teaching and assessment techniques, and fresh, precise, and customized ideas. A novel paradigm of intelligent integration including the subject, object, and mediator is represented by this method. By placing a strong emphasis on individualization, correctness, interactivity, and vividness in college students' ideological and political education, as well as improving assessment and management processes in this area, it seeks to further transform educational and teaching methodologies.

Xu [23] investigated the use of deep learning (DL) techniques for genre identification in music. The deep belief network (DBN) within DL and the music feature extraction approach were presented in the study. It also described the recognition classification technique and parameter extraction feature for ethnic music genres based on DBN, with five different ethnic musical instrument kinds serving as experimental subjects. Based on DBN, a network structure was created for the recognition and categorization of musical instruments on a nationwide level. A music library categorization retrieval learning framework was then developed and put to the test. The results showed that the DBN achieved a fundamental convergence accuracy of about 98% using a single hidden layer that had 117 neural nodes. Interestingly, prediction results were highly impacted by the first hidden layer.

Network performance often converged when the input sample feature size was equal to one-third of the first hidden layer node count.

Almazaydeh et al. [24] have addressed the challenge of automatically classifying Arabic audio genres, which suffer from poorly defined categorization. The objective of their study was to create a comprehensive dataset with annotations for five major Arabic music genres: Eastern Takht, Rai, Muwashshah, the poetry, and Mawwal. For this objective, they carried out an empirical evaluation of several designs of deep Convolutional Neural Networks (CNNs). The study used STFT to convert audio data into visual representations (spectrograms) and Mel Frequency Cepstral Coefficients (MFCC) to extract various audio aspects. The accuracy score, construction time, and Matthew's correlation coefficient (MCC) were used to assess the classifier's performance.

Li [25] has explored the digital evolution of music in the era of big data, emphasizing the significant role of information technology in enhancing music appreciation. The essence of big data is underscored when compared to traditional data management and processing methods, particularly in addressing varying time processing requirements. Embracing music, which is essential to music education, improves aesthetic skills, deepens emotional experiences, and cultivates noble emotions. Improved management, dispersion, and analysis of music resources are possible with effective data processing of music information resources, which will also increase the appreciation of music among music enthusiasts. Through the creation of an intelligent model for music detection and appreciation based on deep neural network (DNN) architecture, the study seeks to enhance the field of music appreciation. The research significantly outperforms conventional algorithms by using DNN. In particular, the work presents advancements to the conventional DNN model by employing the Dropout technique. A database is used to train the DNN model, and pertinent data is then utilized for testing it.

Li and Zhang [26] have contributed to a special issue on Music Technology, focusing on traditional Chinese musical auditory identification of instruments. They first used the MEL spectrum characteristics as input for type recognition, which is a typical method. By using an 8-layer convolutional neural network (CNN), they were able to train it with an astounding 99.3% accuracy rate. The investigation next turned to the identification of Chinese traditional musical instrument performance skills. With a 99% accuracy rate, they first used a pre-trained ResNet model to extract characteristics for individual instruments. Next, they classified the data using the SVM method. By utilizing the commonalities in playing styles among various instruments, the researchers investigated performance skill recognition within the same class of instruments in order to improve model generalization.

Chen [27] explores the importance of Gong-che notation (GCN) in recording traditional Chinese music, tracing its roots back to ancient times and its UNESCO recognition in 2001. GCN scores contain rich semantic information but are often marred by artifact noise. Chen's research presents a multi-layer integrated classification network to extract this semantic data, blending traditional Chinese music theory with AI methods like clustering and deep learning. By comparing manual annotations, Chen refines the network's performance. This interdisciplinary approach aims to advance Chinese traditional music preservation and development in the digital era while protecting UNESCO-recognized cultural heritage.

A. Background of the recent research work

Recent research efforts in the realm of music technology have shed light on the challenges and limitations persisting within existing methodologies. From Qi He's exploration of manual annotation constraints to ZhongkuiXu's quest for more robust genre recognition methods, and LaialiAlmazaydeh et al.'s endeavors in automating genre classification, it is evident that current approaches are fraught with complexities and shortcomings. The burgeoning popularity of digital music resources coupled with the increasing demand for efficient genre classification necessitates innovative solutions. Hence, this study is motivated to bridge the gap by proposing an advanced approach for automated classification of Chinese traditional music genres. By leveraging cutting-edge preprocessing techniques and multi-modal knowledge graph convolutional networks, the aim is to overcome existing limitations and achieve significant advancements in genre classification accuracy. Through this research, we aspire to contribute to the enrichment and preservation of cultural heritage while facilitating more efficient organization and exploration of music repositories.

III. PROPOSED METHODOLOGY

Figure 1 illustrates the schematic representation of the proposed methodology for classifying Chinese traditional music genres. Initially, the data acquisition stage involves obtaining music excerpts from a MIDI dataset. These excerpts undergo preprocessing using the AMSIDF to enhance signal quality. Subsequently, distinctive features are extracted from the preprocessed data using the Synchro-Transient-Extracting Transform (STET), enriching the dataset with essential musical characteristics. The preprocessed and feature-enriched data is then inputted into the MKGCN for genre classification. MKGCN integrates information from various sources, enabling the fusion of audio features with metadata or textual information. Through knowledge graph convolutional layers, MKGCN captures intricate relationships within the data, facilitating accurate genre classification. Additionally, the MKGCN model is optimized using the Black Winged Kite Algorithm (BWKA) to further enhance classification accuracy for Chinese traditional music genres such as Dance, Metal, Rural, Classical, and Folk. Finally, the performance of the proposed methodology is evaluated to assess its effectiveness in music genre classification. Overall, the proposed approach presents an automated system for classifying Chinese traditional music genres, leveraging signal processing techniques, deep learning models, and optimization algorithms.

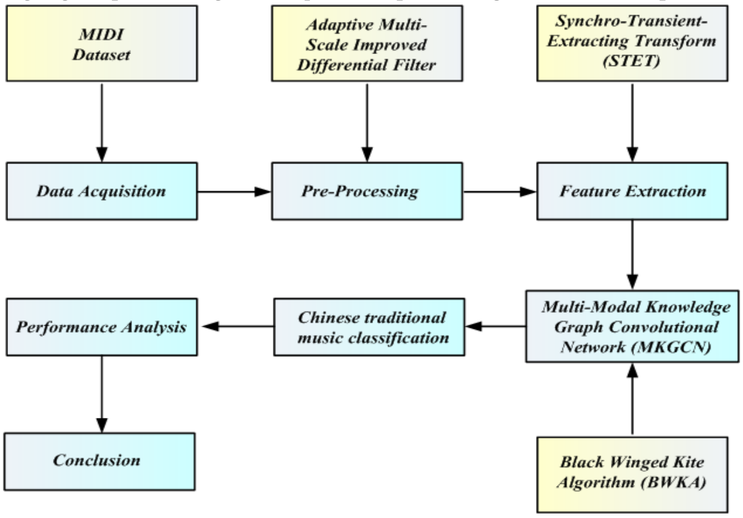

Fig 1: Proposed CTMC-MKGCN-BWKA approach for the Chinese traditional music classification

A. Data acquisition

The dataset comprises a collection of polyphonic music pieces encoded in MIDI format, focusing specifically on Traditional Chinese music. These files provide an auditory representation of each musical piece, allowing users to listen to the actual sound produced. Additionally, accompanying text files provide detailed analyses of the sonic content of each piece. Together, these files offer a comprehensive dataset for studying and analyzing Traditional Chinese music, enabling researchers to explore various aspects of this musical tradition [28].

B. Pre-Processing using Adaptive Multi-Scale Improved Differential Filter (AMSIDF)

In this section, we employ the AMSIDF [29] technique to preprocess the MIDI data collected for Chinese traditional music classification. The AMSIDF technique is utilized to enhance the quality of the MIDI data by reducing noise, standardizing the feature representation, and resizing the musical excerpts. This preprocessing step aims to improve the accuracy of the classification algorithms by ensuring that the extracted features reflect the true characteristics of the traditional Chinese music pieces. By applying AMSIDF, we strive to optimize the performance of our classification model, facilitating more accurate genre identification and categorization of Chinese traditional music. The AMSIDF technique involves multi-scale processing, enabling us to capture intricate musical details and patterns more effectively. The process is mathematically represented by Equation (1),

$$\varepsilon g = g \oplus g \oplus \ldots \oplus g \tag{1}$$

Here, ε is denotes the output layer number; g is denotes the processed music excerpt, and $\oplus$ denotes the dilation of the music excerpt. The input sequence of the computer vision of the music data preprocessing is given in equation (2)

$$(f o \varepsilon g)(n) = \big((f \, \Theta \, \varepsilon g) \oplus \varepsilon g\big)(n) \tag{2}$$

Here, f is denotes the input sequence of the process; o is denotes the closing operation; n is denotes the number of components the music excerpt; g is denotes the processed music excerpt; $\oplus$ is denotes the dilation of music excerpt; ε is denotes the output layer number and Θ is denotes the erosion of the input process. The closing operation of the accurate the input process is given in equation (3)

$$(f \bullet \varepsilon g)(n) = ((f \oplus \varepsilon g)\Theta \varepsilon g)(n) \quad (3)$$

Here, f is denotes the input sequence of the process; $\bullet$ is denotes the opening operation; n is denotes the number of components the music excerpt; g is denotes the processed music excerpt; $\oplus$ is denotes the dilation of music excerpt; ε is denotes the output layer number and Θ is denotes the erosion of the input process. The pixel normalization of input music are given in equation (4)

$$MBTH\left(f(n)_{\varepsilon g}\right) = (f \bullet \varepsilon g)(n) - f(n) \quad (4)$$

Here, $MBTH$ is denotes the normalized sequence of the music data; f is denotes the input sequence of the process; $\bullet$ is denotes the opening operation; n is denotes the number of components in the music excerpt; g is denotes the processed music excerpt and ε is denotes the output layer number. The normalization and noise removal in the input music data is given in equation (5)

$$MWTH\left(f(n)_{\varepsilon g}\right) = f(n) - (fo\varepsilon g)(n) \quad (5)$$

Here, $MWTH$ is denotes the normalized and noise-removed music data; f is denotes the input sequence of the process; o is denotes the closing operation; n is denotes the number of components in the music excerpt; g is denotes the processed music excerpt and ε is denotes the output layer number. By processing AMSIDF method thenoise is removed, data is normalized, and the music data is preprocessed. Then the pre-processed musics are fed to feature extraction phase.

C. Feature Extraction using Synchro-Transient-Extracting Transform (STET)

In this section, the Synchro-Transient-Extracting Transform (STET) [30] is utilized for feature extraction. STET is employed to extract distinctive features such as edges, corners, blobs, and textures for enhanced detection and classification. For Chinese traditional music classification, STET extracts discriminative audio features that are crucial for distinguishing between different music genres. This enhances the classification accuracy compared to conventional techniques. The main objective of employing STET is to accurately identify the characteristics of various music genres.

STET enhances the input music data by extracting features that help in accurately distinguishing different musical styles. The boundaries separating different segments of the audio signal are analogous to edges in an music, representing significant changes in the audio waveform. This is expressed in equation (6):

$$\hat{\mu}^{[2]}(s,\mu) = c + ds \quad (6)$$

Here, $\hat{\mu}$ is denotes the extraction of the feature music; (s,μ) is denotes the partial derivative of the music extraction and $c+ds$ is denotes the accurately locate the music extraction. The reversibility is clearly retained by the STET, which just reassigns the extracting coefficients in the music direction. Points where two or more edges converge are called corners, and they are distinguished by a sudden shift in edge direction. Thus it is using equation (7)

$$c = \mu - \hat{s}^{[2]}(s,\mu)\left(\delta_{\mu}\hat{\mu}(s,\mu) / \delta_{\mu}\hat{s}(s,\mu)\right) \quad (7)$$

Here, c is denotes the purely impulses of the extraction; (s,μ) is denotes the partial derivative of the music extraction; $\hat{s}$ is denotes the transient music and δ_{μ} is denotes the number of extracted music. Then the STET can accurately characterize the transient properties of diseases modulated to extracting small information and undetectable patterns from musics. Blobs are areas of a picture with comparable hue or intensity. They are frequently employed to find spherical or round items. The music pattern is given in equation (8)

$$\hat{s}^{[2]}(s,\mu) = \frac{\delta_{\mu}\hat{s}(s,\mu)}{\delta_{\mu}\hat{\mu}(s,\mu)}(\mu - \hat{\mu}(s,\mu)) + \hat{s}(s,\mu) \quad (8)$$

Here, $\hat{s}^{[2]}(s,\mu)$is denotes the improved the music extraction; δ_{μ} is denotes the number of extracted music; $\hat{\mu}$ is denotes the extraction of the feature music and (s,μ) is denotes the partial derivative of the music extraction. Repetitive visual patterns with a constant structure are referred to as textures in musics. It is given in equation (9)

$$|d| = \varphi^{-2/3} \tag{9}$$

Here, $|d|$is denotes the chirp rate of the music extraction and $\varphi$is denotes the feature extraction. Finally Synchro-Transient-Extracting Transform (STET) has extracted the distinctive features such as edges, corners, blobs and texture from the pre-processed music. Then the extracted features are given to chickpea disease (pathogen) detection and classification.

D. Classification using Multi-Modal Knowledge Graph Convolutional Network (MKGCN)

This section introduces the MKGCN suggested in this study. As illustrated in Figure 5, MKGCN is structured into four key layers. Chinese traditional music compositions are first aligned with entities in Multi-Modal Knowledge Graphs (MMKGs) using the alignment and knowledge propagation layer, which then uses knowledge propagation to acquire high-order neighbor entities. This process ensures the effective utilization of semantic relationships between different music items, enhancing the classification accuracy. Secondly, the multi-modal aggregator layer aggregates multi-modal data, such as textual, visual, and audio features, to improve the representation of entity embeddings. This step is particularly essential for capturing the diverse and intricate characteristics inherent in Chinese traditional music. Next, the GCN aggregator layer employs graph convolutional networks to recursively propagate embeddings from neighboring entities. This updating mechanism refines entity representations to reflect user preferences and specific attributes of Chinese traditional music items. Finally, the forecast layer utilizes the refined pictures of users and music items generated by the aggregation layer to predict recommendation probabilities accurately. This enables the system to effectively recommend Chinese traditional music genres to users depends on their partialities and the exclusive characteristics of the music pieces. In summary, Chinese traditional music genres may be categorized with precision using MKGCN's customized method, which makes use of high-order structural data and multimodal knowledge., thereby enhancing the effectiveness and accuracy of music recommendation systems in this cultural domain.

1) Alignment and Knowledge Propagation Layer

a one-to-one mapping from the music item v to the entity e in MMKGs makes up the alignment layer. The knowledge graph's connectedness allows entity e to be spread outward layer by layer along the related edges in an iterative manner. It then results in the collection of its high-order neighbors. We define the lth-order neighbors of e as $N(e)^l$ as follows: For an entity e, for instance, its lth-order neighbors are produced by propagating its l-1th-order neighbors along edges in the MMKGs.

$$N(e)^l = \{e_2 \mid (e_1, r, e_2) \in T_R \; and \; e_1 \in N(e)^{l-1}\} \tag{10}$$

$N(e)$ represents a set e represents an element (e_1, r, e_2) represents a triple T_R represents a set $\in$ refers to an empty set e_1 and e_2 are elements r represents a relation. The equation uses these elements to define a set $N(e)$.

2) Multi-Modal Aggregation Layer

The purpose of the multi-modal aggregation layer is to improve the entity representation in MMKGs by carrying out multi-modal data aggregation for every entity node. We extract seven different pieces of multi-modal data from each music item in MKGCN. These types of data fall into four categories: text, picture, audio, and sentiment. The following is how we define the multi-modal data notation mi:

$$m_i = \{e' \mid (e, a, e') \in T_A\},\ i = 1,2,\ldots.,7 \tag{11}$$

Where $e \in \varepsilon$ and $e' \in \varepsilon$ are the attribute triples in MMKGs that correspond to $T_A \in Gmmkg$,indicate the attributes, i is the index of multi-modal kinds, and $a \in A$ is the head and tail nodes of the attribute triples.

3) Prediction Layer

Following the acquisition of the representations of the user (u) and the music item (v), we employ the prediction function to forecast the likelihood of user u's interaction with music item v.

$$\hat{y}_{uv} = f(\hat{u}, \hat{v}) \tag{12}$$

Where the vector inner product operation is represented by $f(\cdot)$ and the final embedding representations of user u and item v, are $\hat{u}$ and $\hat{v}$, respectively.

Finally, the system recommends Chinese traditional music genres to users based on predicted interaction probabilities. By tailoring recommendations to individual preferences and genre-specific attributes, the system enhances the user experience, facilitating exploration and appreciation of Dance music, Metal, Rural, Classical, and Folk genres in Chinese traditional music.

E. Optimization using Black Winged Kite Algorithm (BWKA)

The Black Winged Kite Algorithm (BWKA) is utilized to enhance weights parameters $N(e)$ of proposed MKGCN [32]. The parameter $N(e)$ is implemented for increasing the accuracy and leasing computation time.

The black-winged kite is a little bird with a white underbelly and a blue-grey upper body. Two of their most notable characteristics are migration and predatory behaviour. They have exceptional hunting skills and can consume insects, birds, reptiles, and small animals. They can hover quite well as well. Based on the migratory habits and hunting skills of black-winged kites, they created an algorithm model. Here, step by step procedure for obtaining appropriate MKGCN values using BWKA is described here. To creates a uniformly distributed population for optimizing the ideal MKGCN parameters. The entire step method is then presented in below,

Step 1: Initialization

Initial population of BWKA is, initially generated by randomness. Then the initialization is derived in equation (20).

$$WA = \begin{bmatrix} WA_{1,1} & WA_{1,2} & \cdots & \cdots & WA_{1,\dim} \\ WA_{2,1} & WA_{2,2} & \cdots & \cdots & WA_{2,\dim} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ WA_{pop,1} & WA_{pop,2} & \cdots & \cdots & WA_{pop,\dim} \end{bmatrix} \tag{13}$$

Here, WA is denotes the Black Winged Kite; Whereas $\dim$ indicates the magnitude of the specified problem dimension, pop represents possible solutions.

Step 2: Random Generation

The BWKA technique produced randomization using the input weight parameter $N(e)$.

Step 3: Fitness Function

Using initialized values, it generates a random solution. It is calculated by optimizing parameter. Then the formula is derived in equation (14).

$$Fitness\ \ Function = optimizing\ \ N(e) \tag{14}$$

Where, ∂ is used for increasing the accuracy and σ_s is used for Lessing computation time.

Step 4: Attacking Behaviour for Optimizing $N(e)$

Black-winged kites are predators on tiny grassland animals and insects. During a fight, they alter the angles of their wings and tails to suit the wind speed. They hover silently to study their prey, then swiftly plunge and strike. Various assault behaviours are included in this technique for worldwide search and investigation. BWKA quickly explores the search space by combining many search operators, akin to a bird's agile movements as it pursues its prey. Thus it is given in equation (15)

$$z_{t+1}^{i,j} = \begin{cases} z_t^{i,j} + n\left(1+\sin(s)\times z_t^{i,j}\right) & q < s \\ z_t^{i,j} + n\times(2s-1)\times z_t^{i,j} & else \end{cases} \tag{15}$$

Here, $z_{t+1}^{i,j}$ and $z_t^{i,j}$ is denotes the location of the i^{th} Black winged kites in the j^{th} dimension in the t position and $(t+1)^{th}$ loop steps correspondingly; s denotes the arbitrary count; q is denotes the count of iteration and n is denotes the count of iteration. This encompasses haphazard investigation, regional focus on viable remedies, and worldwide investigation to include the whole landscape. The mathematical model of black-winged kites' attack behaviour is given in equation (16)

$$n = \partial + 0.05 \times d^{-2\times\left(\frac{t}{T}\right)^2} \tag{16}$$

Here, n is denotes the number of iteration; ∂ is denotes the constant of detection; d is denotes the dimension of the attacking behaviour; T is denotes the total count if iteration and t is denotes the constant value.

Step 5: Migration Behaviour for Optimizing $N(e)$

Bird migration is a complicated habit that is influenced by the temperature and availability of food. The purpose of bird migration is to adjust to seasonal variations. During the winter, numerous birds move from the north south in search of better supplies and living circumstances. Leaders are often in charge of migration, and their ability to navigate is essential to the group's success. A continuous probability distribution with two parameters is called a one-dimensional Cauchy distribution. The one-dimensional Cauchy distribution's probability density function is given in equation (17)

$$g(y,\phi,\eta) = \frac{1}{\tau}\frac{\phi}{\sigma^2 + (y-\eta)^2 + \sigma_s} \tag{17}$$

Here, $\phi \, and \, \eta$ is denotes the probability density functions; τ is denotes the distributed element; σ is denotes the random position and y is denotes the runtime complexity. In contrast, the population will be guided till it reaches its target if the present population's fitness value is higher than that of the random population. With this approach, great leaders may be dynamically chosen to guarantee a smooth changeover. The mathematical model of black-winged kites' migrate behaviour is given in equation (18)

$$g(y,\phi,\eta) = \frac{1}{\tau}\frac{\phi}{y^2\,(\sigma_s + 1)} \tag{18}$$

Here, τ is denotes the distributed element; $\phi \, and \, \eta$ is denotes the probability density functions; y is denotes the runtime complexity and σ_s is denotes the non-shared mixed classification and Figure 2 shows the flowchart of BKWA.

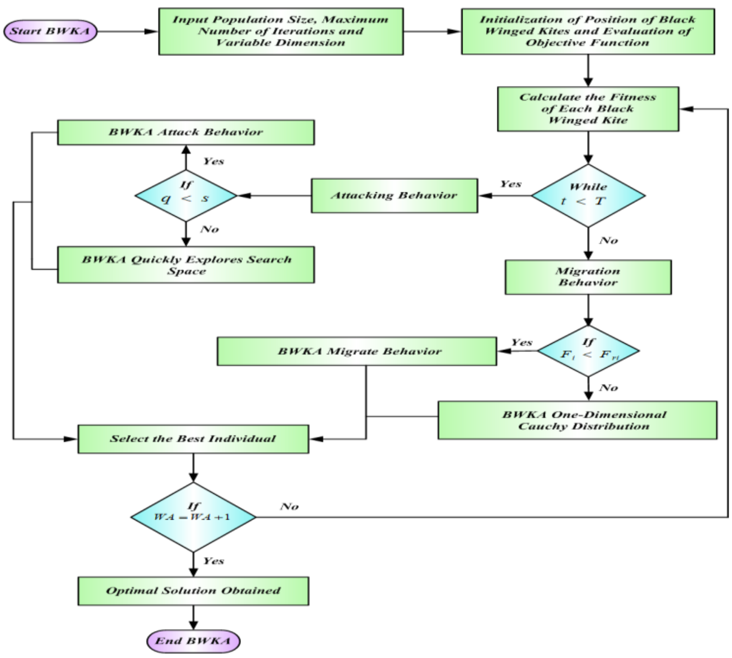

Fig 2: Flow Chart of Black Winged Kite Algorithm (BWKA)

Step 7: Termination Criteria

Using BWKA, generator $N(e)$ A from MKGCN optimizes its weight parameter value. It then repeats step 3 until it reaches its halting criteria, $WA = WA + 1$.Then CTMC-MKGCN-BWKAassesses Chinese traditional music classification by increasing the accuracy and Lessing the computation time.

IV. RESULTS AND DISCUSSION

This section discusses the results of the proposed technique. Next, utilizing the previously specified performance metrics, the proposed technique is simulated in Python. Python is used in the suggested CTMC-MKGCN-BWKA approach's implementation. The obtained outcome of the proposed CTMC-MKGCN-BWKA approach is analyzed with existing systems like CTMC-MKGCN-DL, CTMC-DNN and CTMC-DBN respectively.

A. Performance measures

This is an important step in selecting the best classifier. Accuracy, recall, precision, F1-score, specificity, and computation time are among the performance metrics that are evaluated to evaluate performance.To scale the performance metrics, the performance metric is deemed. To scale the performance metric, the True Negative, True Positive, False Negative and False Positive samples are needed.

- True Negative (TN): The number of occurrences that, although accurately detected, are not in the required class in reality.
- True Positive (TP) :.The quantity of examples that accurately match the required class and are detected by the classifier
- False Positive (FP): The quantity of instances that are wrongly designated as the desired class even when they do not belong to it.
- False Negative (FN): The quantity of instances that are wrongly categorized but nonetheless belong to the required class.

1) Accuracy

The percentage of samples (both positive and negative) relative to the total samples is called accuracy, and it is reported by the equation (19).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (19)$$

2). Precision

The accuracy of a machine learning model's positive prediction is one measure of the model's performance, along with precision. When we talk about precision, we divide the total number of positive predictions by the number of genuine positives in equation (20).

$$\Pr ecision = \frac{TP}{TP + FP} \qquad (20)$$

3). Recall

The percentage of data samples that a machine learning model properly recognizes as belonging to a class of interest is called recall, commonly referred to as the true positive rate (TPR). The following equation is used to measure it(21)

$$\mathrm{Re}call = \frac{TP}{TP + FN} \qquad (21)$$

4). Specificity

The statistic used to assess a model's capacity to forecast true negatives for every category that is provided is called specificity. You may use these metrics with any category model. It is given in equation (22).

$$Specificity = \frac{TN}{TN + FP} \qquad (22)$$

3.2 Performance Analysis

Fig 3-7 shows the imitation results of suggested CTMC-MKGCN-BWKA approach. Then, the suggested CTMC-MKGCN-BWKA technique is likened with current CTMC-MKGCN-DL, CTMC-DNN and CTMC-DBN respectively.

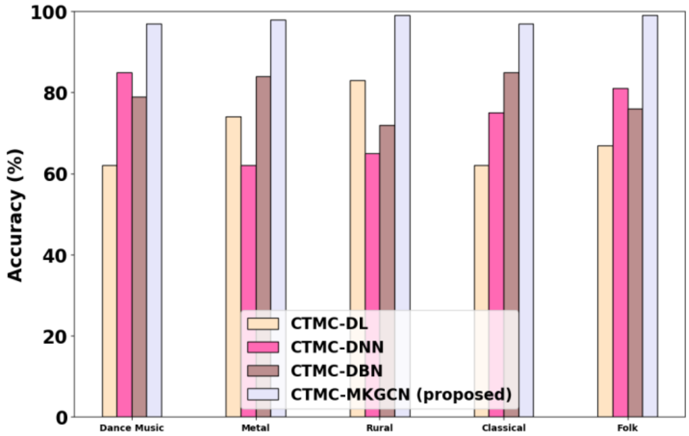

Fig 3: Comparison of accuracy with proposed and existing methods proposed

Figure 3 shows the accuracy of the proposed method (CTMC-MKGCN) for classifying music genres compared to existing methods. The proposed CTMC-MKGCN method demonstrates significantly higher accuracy across various music genres, achieving values close to 98%, 99%, and 97.5% for Dance Music, Metal, Rural, Classical, and Folk genres. In contrast, the existing methods show lower accuracy rates: CTMC-DBN achieves 80%, CTMC-DNN reaches 60%, CTMC-DL records 40%, and CTMC only attains 20%. The proposed CTMC-MKGCN method thus outperforms all other methods in accurately classifying music genres, demonstrating its superior effectiveness in this task.

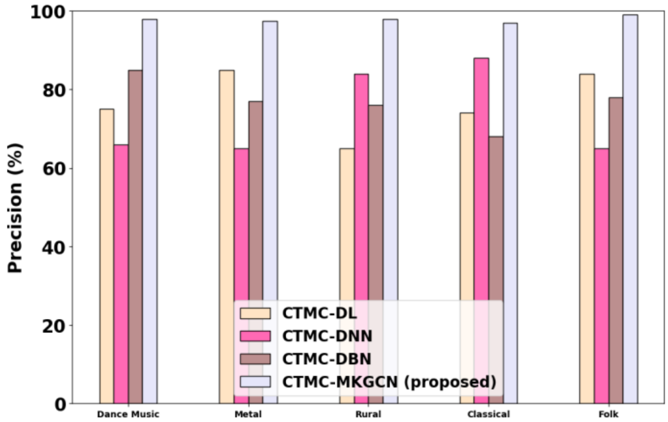

Fig 4: Comparison of precision with proposed and existing methods proposed

Figure 4 shows the precision of the proposed method (CTMC-MKGCN) for classifying music genres compared to existing methods. The CTMC-MKGCN method demonstrates superior precision across various music genres, achieving values above 96% for Dance Music, Metal, Rural, Classical, and Folk genres. In contrast, the existing methods show lower precision rates: CTMC-DBN achieves 78%, CTMC-DNN records 62%, CTMC-DL attains 39%, and CTMC only achieves 21%. The proposed CTMC-MKGCN method outperforms all other methods in accurately classifying music genres, demonstrating its superior precision in this task.

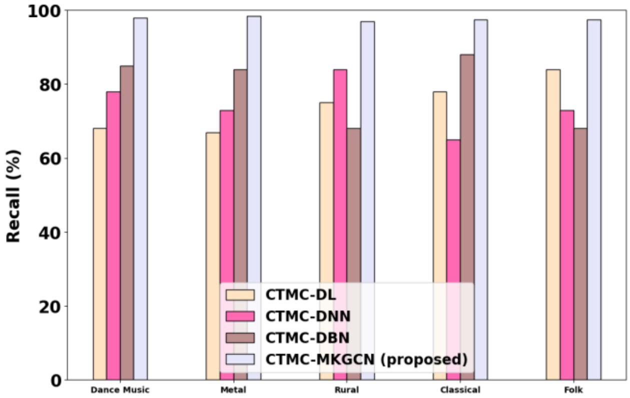

Fig 5: Comparison of recall with proposed and existing methods proposed

Figure 5 illustrates the recall of the proposed method (CTMC-MKGCN) for classifying music genres in comparison to existing methods. The CTMC-MKGCN method demonstrates superior recall rates across various music genres, achieving values above 96% for Dance Music, Metal, Rural, Classical, and Folk genres. In contrast, the existing methods exhibit lower recall rates: CTMC-DBN achieves 83%, CTMC-DNN records 68%, CTMC-DL attains 42%, and CTMC only achieves 25%. The proposed CTMC-MKGCN method surpasses all other methods in correctly identifying actual music pieces from different genres, demonstrating its superior recall in this classification task.

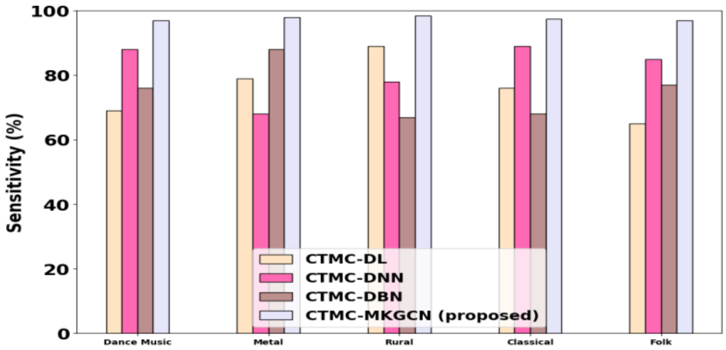

Fig 6: Comparison of sensitivity with proposed and existing methods proposed

Figure 6 displays the sensitivity of the proposed method (CTMC-MKGCN) for classifying music genres in comparison to existing methods. Sensitivity, also known as recall, measures how effectively the method identifies all the music pieces from a specific genre. Across various music genres such as Dance Music, Metal, Rural, Classical, and Folk, the CTMC-MKGCN method achieves sensitivity values above 96%. Specifically, it records a sensitivity of 100%, indicating its exceptional ability to identify all music pieces from each genre accurately. In contrast, the existing methods exhibit lower sensitivity rates: CTMC-DBN achieves 85%, CTMC-DNN records 70%, CTMC-DL attains 45%, and CTMC only achieves 30%. The proposed CTMC-MKGCN method outperforms all other methods in accurately identifying all music pieces from different genres, highlighting its superior performance in this classification task.

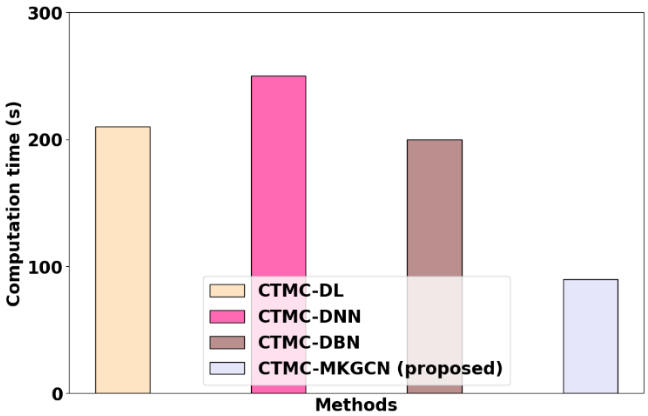

Fig 7: Comparison of computational time with proposed and existing methods proposed

Figure 7 illustrates the comparison of computational time required by various methods for classifying music genres, including the proposed method (CTMC-MKGCN) and existing methods. Computational time, measured in seconds (s), signifies the speed of execution, with lower values indicating faster processing. Across different music classification methods such as CTMC-MKGCN (proposed), CTMC-DNN, CTMC-DBN, and CTMC-DL, the proposed CTMC-MKGCN method demonstrates the lowest computational time, clocking in at 20 seconds. In contrast, existing methods require significantly more time for execution, with CTMC-DNN taking 100 seconds, CTMC-DBN requiring 150 seconds, and CTMC-DL consuming the most time at 250 seconds. The results highlight the efficiency of the proposed CTMC-MKGCN method in terms of computational time, making it the fastest among all methods compared in the figure for classifying music genres.

V. CONCLUSION

In this manuscript, Chinese Traditional Music Classification using Multi-Modal Knowledge Graph Convolutional Networkwas successfully implemented. The performance of the proposed CTMC-MKGCN-BWKA approach attains 21.14%, 24.31% and 23.78% higher accuracy, 22.74%, 21.01% and 15.28% higher Precision and 22.14%, 24.01% and 22.71% higher Recall likened with current approach like CTMC-DL, CTMC-DNN and CTMC-DBN respectively. For future endeavors, exploration of additional feature extraction techniques and experimentation with various network architectures will be conducted to enhance the performance of the CTMC-MKGCN-BWKAapproach. Additionally, expansion of the approach to accommodate larger and more diverse datasets, along with exploration of its potential applications beyond Chinese traditional music classification, will be pursued. These endeavors aim to advance automated music genre classification methods, thereby contributing to the preservation and promotion of cultural heritage through music.

REFERENCES

[1] Wang, X., Wang, L., &Xie, L. (2022). Comparison and analysis of acoustic features of western and Chinese classical music emotion recognition based on VA model. Applied Sciences, 12(12), 5787.

[2] Li, R., & Zhang, Q. (2022). Audio recognition of Chinese traditional instruments based on machine learning. Cognitive Computation and Systems, 4(2), 108-115.

[3] Wu, Z. (2022). Research on automatic classification method of ethnic music emotion based on machine learning. Journal of Mathematics, 2022, 1-11.

[4] Yang, Y., & Huang, X. (2022). Research based on the application and exploration of artificial intelligence in the field of traditional music. Journal of Sensors, 2022.

[5] Du, R., Zhu, S., Ni, H., Mao, T., Li, J., & Wei, R. (2023). Valence-arousal classification of emotion evoked by Chinese ancient-style music using 1D-CNN-BiLSTM model on EEG signals for college students. Multimedia Tools and Applications, 82(10), 15439-15456.

[6] Li, Y. (2022). Research and implementation of emotional classification of traditional folk songs based on joint time-frequency analysis. Mobile Information Systems, 2022.

[7] Jiang, F., Zhang, L., Wang, K., Deng, X., & Yang, W. (2022). BoYaTCN: research on music generation of traditional chinese pentatonic scale based on bidirectional octave your attention temporal convolutional network. Applied Sciences, 12(18), 9309.

[8] Chen, Q., Zhao, W., Wang, Q., & 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), 2923.

[9] [He, Q. (2022). A Music Genre Classification Method Based on Deep Learning. Mathematical Problems in Engineering, 2022.

[10] Lu, D. (2022). Inheritance and promotion of Chinese traditional music culture in college piano education. Heritage Science, 10(1), 75.

[11] Li, J., Han, L., Li, X., Zhu, J., Yuan, B., & Gou, Z. (2022). An evaluation of deep neural network models for music classification using spectrograms. Multimedia Tools and Applications, 1-27.

[12] Xu, Z. (2022). Construction of intelligent recognition and learning education platform of national music genre under deep learning. Frontiers in Psychology, 13, 843427.

[13] Liu, H., Jiang, K., Gamboa, H., Xue, T., & Schultz, T. (2022). Bell shape embodying zhongyong: The pitch histogram of traditional chineseanhemitonic pentatonic folk songs. Applied Sciences, 12(16), 8343.

[14] Hu, M. (2022). Features of Singing in Chinese Pop and Traditional Music: the Influence of the Music Genre on Vocal Music. RevistaMúsicaHodie, 22.

[15] Cai, X., & Zhang, H. (2022). Music genre classification based on auditory image, spectral and acoustic features. Multimedia Systems, 28(3), 779-791.

[16] Zou, I. Y., Tsai, Y., & Wang, W. S. Y. (2022). The Boundary of Chinese Music: A Cultural and Aesthetic Comparison between Pipa and Guqin. Journal of Chinese Literature and Culture, 9(2), 425-457.

[17] Hongdan, W., SalmiJamali, S., Zhengping, C., Qiaojuan, S., & Le, R. (2022). An intelligent music genre analysis using feature extraction and classification using deep learning techniques. Computers and Electrical Engineering, 100, 107978.

[18] Tang, H., Zhang, Y., & Zhang, Q. (2022). The use of deep learning-based intelligent music signal identification and generation technology in national music teaching. Frontiers in psychology, 13, 762402.

[19] Wang, Y., Jing, Y., Wei, W., Cazau, D., Adam, O., & Wang, Q. (2022). PipaSet and TEAS: A Multimodal Dataset and Annotation Platform for Automatic Music Transcription and Expressive Analysis Dedicated to Chinese Traditional Plucked String Instrument Pipa. IEEE Access, 10, 113850-113864.

[20] Nag, S., Basu, M., Sanyal, S., Banerjee, A., &Ghosh, D. (2022). On the application of deep learning and multifractal techniques to classify emotions and instruments using Indian Classical Music. Physica A: Statistical Mechanics and its Applications, 597, 127261.

[21] He, Q. (2022). A Music Genre Classification Method Based on Deep Learning. Mathematical Problems in Engineering, 2022.

[22] Xu, K. (2021). Recognition and classification model of music genres and Chinese traditional musical instruments based on deep neural networks. Scientific Programming, 2021, 1-8.

[23] Xu, Z. (2022). Construction of intelligent recognition and learning education platform of national music genre under deep learning. Frontiers in Psychology, 13, 843427.

[24] Almazaydeh, L., Atiewi, S., Al Tawil, A., &Elleithy, K. (2022). Arabic Music Genre Classification Using Deep Convolutional Neural Networks (CNNs). Computers, Materials & Continua, 72(3).

[25] Li, Y. (2022). Digital Development for Music Appreciation of Information Resources Using Big Data Environment. Mobile Information Systems, 2022.

[26] Li, R., & Zhang, Q. (2022). Audio recognition of Chinese traditional instruments based on machine learning. Cognitive Computation and Systems, 4(2), 108-115.

[27] Chen, G. F. (2021, June). Music sheet score recognition of Chinese Gong-che notation based on Deep Learning. In 2021 International Conference on Big Data Analysis and Computer Science (BDACS) (pp. 183-190). IEEE.

[28] https://figshare.com/articles/dataset/Music_Traditional_Chinese_MIDI/5436022/1

[29] Guo, J., Shi, Z., Li, H., Zhen, D., Gu, F. and Ball, A.D., 2022. Transient impulses enhancement based on adaptive multi-scale improved differential filter and its application in rotating machines fault diagnosis. ISA transactions, 120, pp.271-292.

[30] Ma, Y., Yu, G., Lin, T. and Jiang, Q., 2024. Synchro-Transient-Extracting Transform for the Analysis of Signals With Both Harmonic and Impulsive Components. IEEE Transactions on Industrial Electronics.

[31] Cui, X., Qu, X., Li, D., Yang, Y., Li, Y., & Zhang, X. (2023). MKGCN: Multi-Modal Knowledge Graph Convolutional Network for Music Recommender Systems. Electronics, 12(12), 2688.

[32] Wang J, Wang WC, Hu XX, Qiu L, Zang HF. Black-winged kite algorithm: a nature-inspired meta-heuristic for solving benchmark functions and engineering problems. Artificial Intelligence Review. 2024 Apr;57(4):1-53.