Digital Signal Processing and Artificial Intelligence for Automatic Learning

DLINE JOURNALS

Online ISSN: 2583-5009

DSP 2024; 3 (3)

https://doi.org/10.6025/dspaial/2024/3/3/83-89

Digital Music Creation System Based on Recursive Neural Network Using AI Algorithm

TaoZhang
Henan Vocational Institute of Arts
450000, Zhengzhou, Henan, China
76525360@qq.com

ABSTRACT

In recent years, the advancement of artificial intelligence technology has become a significant driving force for digital music creation, attracting widespread attention. Compared to the high cost and uncertain timing of manual creation, the extreme cost-effectiveness of AI music can effectively help developers reduce costs and increase efficiency. Therefore, the automation of digital music creation has broad applications and significance. The recursive neural network is an artificial intelligence algorithm based on neural networks, mainly with a loop structure, that can adaptively process sequence data. In a digital music creation system, recursive neural networks can predict and generate note sequences, thereby achieving automated music creation. Various factors, such as the quality of training data, the settings of model parameters, and the stylistic characteristics of the music influence the predictive results of recursive neural networks. Therefore, meticulous parameter tuning and optimization need to be conducted for different application scenarios.

Received: 11 May 2024
Revised: 2 August 2024
Accepted: 14 August 2024

Keywords: Artificial Intelligence Algorithm, Recursive Neural Network, Music Creation, System

1. Introduction

With the continuous development of artificial intelligence technology, more and more fields have begun to explore and apply AI technologies. In the digital music creation field, AI technology applications are already widespread. Digital music creation includes music using a computer, including all aspects, such as melody, harmony, rhythm, and arrangement [1,2]. Digital music creation can not only greatly improve the efficiency and quality of music creation but also provide more creativity and possibilities for music creation. However, there are some challenges and problems in digital music creation. The main issue is the complexity and uncertainty of digital music creation. Digital music creation must consider many factors, such as music theory, harmonic changes, melodic lines, rhythm patterns, etc. The interactions among these factors are complex and require creators to have high musical literacy and skills [3]. At the same time, digital music creation also needs to consider computer programming and algorithm design knowledge, which brings more challenges to digital music creation. To solve the problems of complexity and uncertainty in digital music creation, the recursive neural network algorithm can be used to learn and simulate input music data to generate new digital music works [4]. Through this system, users can input a piece of music data, and the system will automatically learn and simulate according to the input music data and generate new digital music works. The traditional method of digital music creation requires manual work, which is inefficient and easily influenced by the subjectivity of the creators. By adopting the recursive neural network artificial intelligence algorithm, our digital music creation system can automatically complete music creation without human intervention. This can significantly improve the efficiency and quality of music creation. At the same time, this system can also provide users with more personalized digital music works, meeting the different needs and preferences of users. In addition, the digital music creation system based on the recursive neural network artificial intelligence algorithm can bring more creativity and possibilities to the field of digital music. For example, it can be used for automatic arrangement, rhythm generation, harmony matching, etc. These applications can greatly promote the development and innovation of the digital music field, providing more possibilities and choices for music creation. The recursive neural network (RNN) is an artificial intelligence algorithm based on neural networks, mainly with a loop structure, that can adaptively process sequence data [5]. In a digital music creation system, recursive neural networks can predict and generate note sequences, thereby achieving automated music creation. Various factors, such as the quality of training data, the settings of model parameters, and the stylistic characteristics of the music influence the predictive results of recursive neural networks. Therefore, meticulous parameter tuning and optimization need to be conducted for different application scenarios.

2. Related Work

In the digital age, the development and application of digital music creation systems have become one of the hotspots in the music field. The emergence of digital music creation systems provides a more convenient, efficient, and free way for music creators to create music. Also, it provides more creative possibilities and space for music lovers and professionals. Against this background, the digital music creation system based on the recursive neural network artificial intelligence algorithm has gradually attracted the attention of researchers. Gong and others discussed using deep learning technology to build a music recommendation system to recommend personalized music based on users' listening history and preferences [6]. The authors proposed a neural network-based model that can learn the feature representation of music and classify and recommend music based on user preferences. Phuoc and others focused on how to use machine learning algorithms to identify emotions in music [7]. The authors first introduced the basic concepts and methods of emotion recognition and then discussed some current emotion recognition algorithms, such as deep learning-based emotion recognition algorithms, sound feature-based emotion recognition algorithms, etc., and discussed their pros and cons. Marceau and others discussed using recursive neural networks (RNN) to generate music [8]. The authors proposed a music generation model based on LSTM (Long Short-Term Memory) networks, which can learn the feature representation of music from the input music data and generate new music pieces with similar features. Yan and others discussed using collaborative filtering algorithms to build a personalized music recommendation system [9]. Collaborative filtering is a recommendation algorithm based on user behavior. It can recommend other music similar to the user's interests based on the user's historical behavior and preferences. The authors proposed a recommendation model based on collaborative filtering, which can make personalized recommendations based on the user's preferences and historical behavior. Dao and others focused on the progress of research on digital music copyright protection technology [10]. The authors introduced the basic concepts and methods of digital music copyright protection, such as watermark technology, encryption technology, digital signature, etc. They discussed some current copyright protection technologies, such as blockchain-based copyright protection technology etc. In addition, the study also explored how to use deep learning technology to classify music styles. A style classification model based on convolutional neural networks was proposed. This model can learn the feature representation of music from audio data and automatically classify music into different style categories, such as pop, rock, classical, etc. The design and implementation of the digital music education platform were also focused on. The basic concepts and methods of digital music education, such as the design principles of online education platforms, the characteristics of music education, etc., were introduced, and how to use modern technology to build a digital music education platform to provide efficient and personalized music education services was discussed. This work involves all aspects of digital music, including basic concepts, melody construction, harmony processing, rhythm control, arrangement techniques, etc. At the same time, it also focuses on research progress in the recommendation system, emotion recognition, copyright protection, etc., of digital music.

3. Material and Method

3.1. Neural Network Algorithms

The structure of a neural network goes from the input layer to the hidden layer and then to the output layer. The nodes between each layer are disconnected, with each layer being fully connected. In a recursive neural network, the electrical output of each layer's nodes is related to the input and the output at the previous moment. Specifically, the input to the network's hidden layer includes not only the output of the input layer but also the hidden layer at the previous moment. The recursive neural network model is a biologically inspired intelligent model based on research into biological neural mechanisms. Neurons in a recursive neural network can be roughly divided into three types: input neurons, output neurons, and hidden neurons. The function of input neurons is to obtain data or other information from the external environment directly. The function of output neurons is to reflect the constructed neural system to the external environment. Both types of neurons are directly related to the external environment. Hidden neurons are not directly related to the external environment and exist within the neural network system. Their important role is to obtain input information from within the recursive neural network and then apply the obtained output information to other neurons in the recursive neural network system after calculation by formula 1.

$$J = \frac{\sum_{1}^{i} \frac{S_i}{100} * X_i}{\sum_{1}^{i} X_i} \qquad \textbf{(1)}$$

3.2. Neural Network Algorithm Process

When constructing a recursive neural network model, it is necessary first to collect enough training samples, known as training cycles. After several training cycles, the neural network's performance will be significantly improved, eventually achieving a stable output result. Measurement indicators are needed as input parameters for the network to select the number of features. Some of these input parameters are quantitative values, and some are qualitative values. A sublevel scoring method is used for the quantification of qualitative values. For quantitative values, the standardization method of engineering economics can be used for processing. The weight settings or initial threshold values can be set randomly. However, suppose the weight value is relatively large. In that case, the result may be that the network quickly saturates, and the weight's initial value will affect the network's convergence speed to a certain extent. The neural network's learning process is realized by using the given sample data, and the metric reflecting the learning result is the root mean square error of the network. Generally, when the root mean square error of the recursive neural network is less than 0.1, it indicates that the training results of the given samples meet the requirements. The flow of the recursive neural network algorithm is shown in Figure 1.

3.3. Activation Function

The weight adjustment equation of the recursive algorithm is calculated according to the learning speed of each subset, the output results, and the required dataset. These subsets include data obtained from the source and data processed through the network. The number of its nodes is the number of input variables. The output layer outputs the information processing results to the outside world. The number of nodes is the number of output variables. In the output layer of the network, the error signal is determined by the difference between the expected and actual results, which can directly reflect the accuracy of the output. In the hidden layer, however, the error signals are generated from the results of the previous layers, opposite to the results of the output layer. The most commonly used activation function in neural networks is the Sigmoid function, as shown in formulas 2 and 3. Using this function to "squeeze" the input of the previous layer of neurons, the output is compressed into an interval. During the construction process, the structure of the recursive neural network is determined according to the features of the function. Applying the recursive neural network makes it possible to effectively combine the input sample data with the output data, thereby constructing a model that can accurately predict the output results.

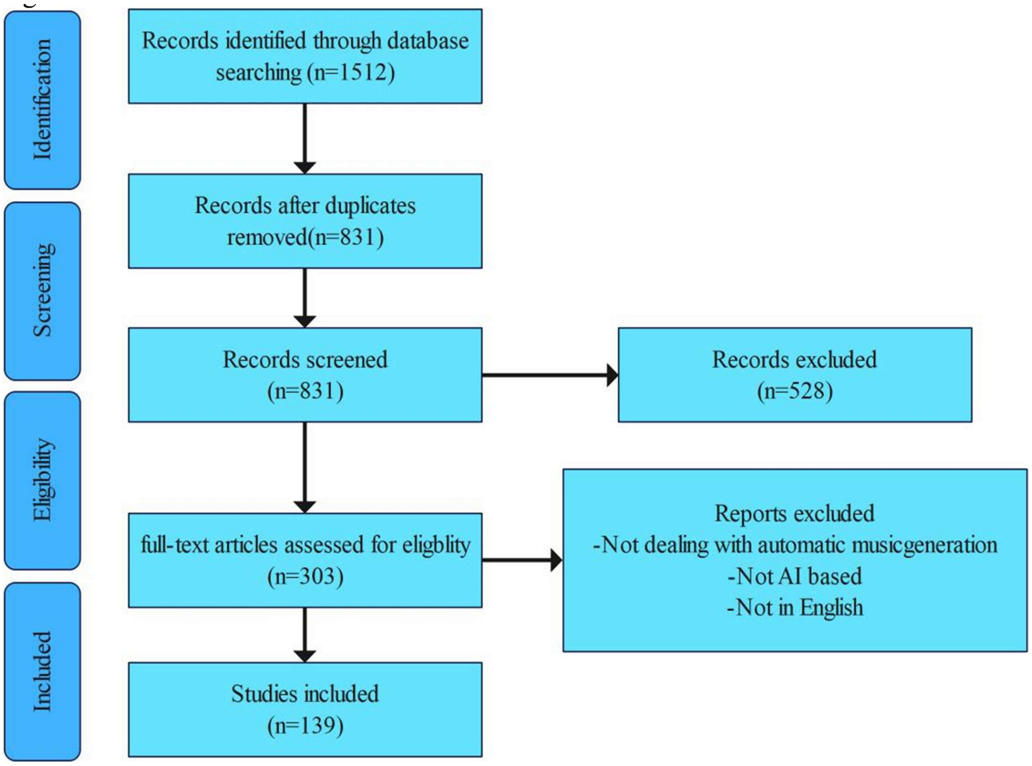

Figure 1. Recursive Neural Network Flowchart

$$\Delta w_{ij}^{n} = \eta \delta_j^{n} y_i^{n-1} = \eta y_i^{n-1} \sum_{k=1}^{m_{n+1}} \delta_k^{n+1} w_{jk}^{n+1} f'_{n-1}\left(\quad {}_{n-1} \right) \tag{2}$$

$$y_i = f\left(\sum_{k=1}^{n} K_{ki} v_k + v_{i0} \right) \tag{3}$$

4. Results

4.1. Experimental Design

Regarding system design, this study plans to collect many MIDI files to train the neural network model. Then, a music generation model will be established based on the preprocessed MIDI data. In addition, to improve the quality of the generated music, researchers will design some adaptive adjustment strategies, such as completing system optimization and user interaction in the mixing process according to user-specified music features and mixing parameters such as volume, balance, and spatial sense, to improve the efficiency and quality of music composition. We will also conduct subjective and objective evaluations of the generated music to assess its quality and potential advantages and disadvantages. By exploring the above research ideas, we will establish an AI-based digital music creation system to achieve the automatic creation of digital music. With the application needs of music teaching, high-end live concerts, and other occasions, the future presents a relatively rapid growth trend. By further integrating elements such as lighting and scene setting, emphasizing the movement of time and space, as well as the color, aesthetic, and size of objects corresponding to the height, timbre and intensity of the music, we characterize the emotional features of music to achieve better audiovisual effects and aesthetic effects of music. Since the selected feature set largely reflects the corresponding classification standards, by analyzing the relationship between the feature sets of two different classification standards in the standard music sample library and the user classification standard feature set, we can obtain the relationship between the user classification and the two standard classifications. According to the correlation analysis results of the classification criteria, a new feature set is constructed to reflect both the standard and user classifications. Using this new feature set, a reasonable user classifier is constructed. Finally, music that meets user needs is selected from the standard music library using the new classifier.

4.2. Experimental Result Analysis

The recursive neural network is used as an independent detection technology to obtain and analyze data from the neural network directly. By adopting the method of algorithm decomposition, we can preprocess most of the information in the original model and only retain a small part of the model, and then make decisions based on the output of the preprocessing results, thereby improving the model's performance. This method can greatly improve the accuracy and reliability of the model. However, training with just one set of samples often comes with a lot of noise, which makes it impossible to get satisfactory results with any parameters in each round of iteration, thus making the effect of gradient reduction techniques relatively poor. We can effectively solve linearly inseparable classification problems through the recursive neural network model. In addition, when the number of hidden layers is sufficient, this model can classify various patterns. However, since the learning method of each node is unique, their outputs often fail to achieve the expected results, as shown in Figure 2.

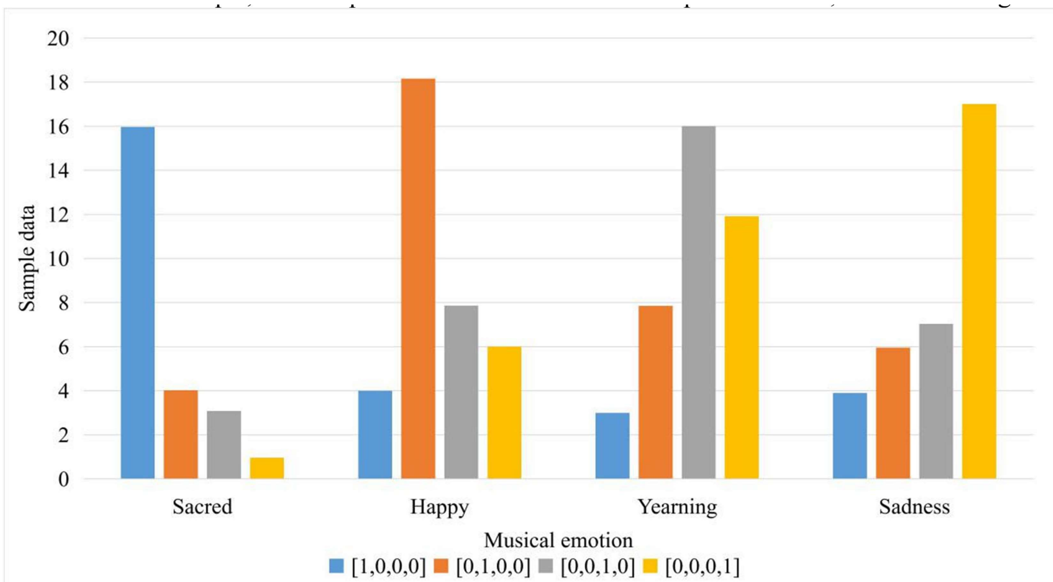

Figure 2. Recognition Result of the Neural Network

From the results, it can be concluded that the classification result of the neural network after data processing (with an accuracy rate of 81.2%) is significantly better than the result before data processing (73%). However, although the data has been processed, the actual error is still inevitable, so the computational accuracy of 81.2% is a very good recognition result. By comparing two different music style classification methods, we found that the method based on paragraph features is more effective. We established two different models and evaluated their classification results. We found that the accuracy rate of this method in the test set is as high as 90%, proving its effectiveness. As the dimensionality of the features continues to increase, the training time for the two modes will correspondingly become longer, especially when the dimensionality exceeds 4000. However, the research found that even during the training process, the improved feature selection mode can obtain higher accuracy at a lower cost; therefore, it can still provide high efficiency. To achieve better accuracy, we should try to reduce the input for training, which is also a reasonable choice. Removing some redundant features within a proper range can improve computational efficiency without much impact on accuracy. When the dimensionality is below 4000, the time changes are not too large. Therefore, sacrificing some training time to improve accuracy is acceptable. When the training set data changes between 1% and 10%, the accuracy of the proposed method is higher, fluctuating between 85% and 99%, with little fluctuation. The adoption of text serialization techniques and the attention mechanism has increased the model's accuracy from 40% to 90%, significantly improving the model's performance. When the training data is small, the text vector feature representation has high-dimensional sparsity, and the learning ability of the model is relatively weak. Despite using two different modes to handle large amounts of text information, due to the lack of sufficient redundancy, their classification effects still have uncertain factors, which cause the final results to be biased. To solve this problem, we adopted the method of embedding perturbation elements into the input layer, making the model more robust, as shown in Figure 3.

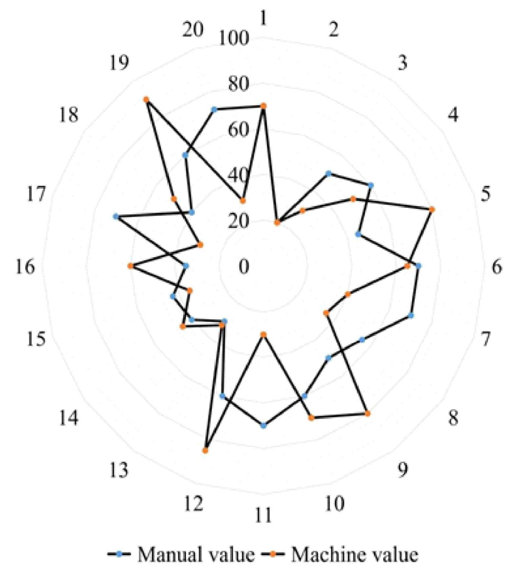

Figure 3. Changes in Music Evaluation Values

The probability estimate of the logistic regression model is used as the final style authenticity index. The logistic regression model achieves comparable performance in style classification compared to SVM, and the prediction has a probability interpretation. The index ranges from 0 to 1, with higher values representing the L1 writing style and vice versa. We map the style index to a bipolar color axis to make it easier for users to understand the trend of writing styles. Cross-validation results show that in the worst case, where the sample length is 70 sentences, the logistic regression model's accuracy and recall rates are higher than 10%. We obtain a multinomial logistic regression model for each of the five selected features and train them simultaneously, using the prediction output of the model as a measure of the writing style of user text in the five categories of features. The overall range is from 0 to 1, with higher values indicating closer to the L1 writing style and vice versa.

5. Conclusions

The rapid development of artificial intelligence technology provides more possibilities for innovation in digital music. In the digital music creation system, artificial intelligence can assist in directly creating music. This paper introduces the digital music creation system's main work and innovation points based on artificial intelligence and points out existing problems and future research directions. However, compared with traditional music creation, digital music based on artificial intelligence still faces challenges and problems. For example, existing artificial intelligence algorithms find it difficult to express and convey emotion when generating music, lacking the connotation and meaning of human music creation. Moreover, due to the limitations of large amounts of training data, digital music creation systems may lack sufficient diversity and creativity. This direction also needs to be improved and optimized in the digital music creation system. Therefore, future research aims to improve music style classification and emotion analysis, improve the performance of music generation algorithms, increase the diversity and creativity of generated music, make the generated music closer to human music creation, and convey more emotions and meanings. It can be combined with ChatGPT to empower virtual people and robots and transform them into singers who can create and perform music improvisationally according to human requirements. We need to optimize the digital music creation system based on artificial intelligence while also researching how the system can be integrated and empowered with other products.

References

[1] Ku, H. H., Liu, C. H and Wang, W. C. (2022). Design of an artificial intelligence of things-based indoor planting model for *Mentha spicata*. *Processes, 10*(1), 116.

[2] Ichimasa, K., Nakahara, K., Kudo, S. E., et al. (2022). Novel "resect and analysis" approach for T2 colorectal cancer with use of artificial intelligence. *Gastrointestinal Endoscopy, 96*(4), 665-672.

[3] Crowell, B., Cug, J and Michalikova, K. F. (2022). Smart wearable internet of medical things technologies, artificial intelligence-based diagnostic algorithms, and real-time healthcare monitoring systems in COVID-19 detection and treatment. *American Journal of Medical Research, 9*(1), 17-32.

[4] Amann, J., Vetter, D and Blomberg, S. N., et al. (2022). To explain or not to explain?—Artificial intelligence explainability in clinical decision support systems. *PLOS Digital Health, 1*(2), e0000016.

[5] Junejo, A. R., Li, X and Madiha, H., et al. (2021). Molecular communication networks: Drug target scalability based on artificial intelligence prediction techniques. *Journal of Nanoparticle Research, 23,* 1-23.

[6] Gong, X and Li, X. (2023). Human–robot interactive communication and cognitive psychol ogy intelligent decision system based on artificial intelligence—Case study. *International Journal of Humanoid Robotics, 20*(02n03), 2240005.

[7] Phuoc, N. V. (2022). The critical factors impacting artificial intelligence applications adoption in Vietnam: A structural equation modeling analysis. *Economies, 10*(6), 129.

[8] Marceau, G and Daniele, F. (2022). Artificial intelligence and international economic law by Shin-yi Peng, Ching-Fu Lin, and Thomas Streinz. *World Trade Review, 21*(4), 516-520.

[9] Yan, L., Song, H and Guo, Y., et al. (2022). HLDnet: Novel deep learning-based artificial intelligence tool fuses acetic acid and Lugol's iodine cervicograms for accurate pre-cancer screening. *Biomedical Signal Processing and Control, 71,* 103163.

[10] Van Dao, D., Bui, Q. A. T and Nguyen, D. D., et al. (2022). Prediction of interlayer shear strength of double-layer asphalt using novel hybrid artificial intelligence models of ANFIS and metaheuristic optimizations. *Construction and Building Materials, 323,* 126595.