Automatic Music Creation Based on Bayesian Networks

Cundi Fang
Institute of Information Science and Engineering
Hunan University
Hunan Changsha, China
86+18890023897
qdfcd@hnu.edu.cn

Zhiyong Li
Institute of Information Science and Engineering
Hunan University
Hunan Changsha, China
86+13607436411
zhiyong.li@hnu.edu.cn

Zhihao Ye
Institute of Information Science and Engineering
Hunan University
Hunan Changsha, China
86+15626124528
1079345916@qq.com

ABSTRACT
Many models of music arrangement network based on neural network have been proposed so far. But most use deep learning or GAN confrontation network, which is difficult to solve the problem of high randomness of music editing. Therefore, this paper attempts to use Bayesian network to learn the music characteristics of different styles of music. And then to create new music more in line with different styles of music, in order to replace manpower to complete the work of artistic creation. Through the design of Bayesian network model and genetic algorithm, the network can effectively learn the music style characteristics and music characteristics of a certain piece of music or a certain type of music. Then using the data the model had learned to recreate the music. To explain the implementation of the model, multiple sets of examples of different wind characteristics are used. Through this experiment, it is concluded that the combination of Bayesian network and genetic algorithm can produce a more effective music arrangement network model.

CCS Concepts
• Networks→Network properties

Keywords
Music, Neural Network, Bayesian networks

ACM Reference format:
Cundi Fang, Zhiyong Li and Zhihao Ye. 2020. Automatic Music Creation Based on Bayesian Networks. In Proceedings of *2020 4th International Conference on Vision, Image and Signal Processing(ICVISP 2020), December 9-11, 2020, Bangkok, Thailand.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3448823.3448871

1. INTRODUCTION
As a frontier science, artificial intelligence can be applied to many fields, including music, conversation and other fields of artistic creation. A variety of ways to process music files using deep learning are given in [1]. But because of the high randomness and innovation of artistic creation itself, it is difficult to produce more innovative music by using general deep learning or neural network algorithm. Therefore, the purpose of this paper is to solve the problem of low innovation when using neural network algorithm to create music. So in this paper, we try to use Bayesian network to build a model to describe the connection probability between music context.

ICVISP 2020, December 9–11, 2020, Bangkok, Thailand

ACM ISBN 978-1-4503-8953-2/20/12…$15.00
https://doi.org/10.1145/3448823.3448871

Because music itself has certain regularity and periodicity, it is difficult to find out the definite rules and requirements artificially(Figure 1). This requires the network not only to be able to learn the characteristics of data, but also to have a strong randomness and innovation. While learning the trend of music, the network does not stick to the music in the data set, but produce their own new music. The general neural network is more suitable for the deterministic data classification problem, but it is difficult to deal with and produce better results in the face of innovative problems.

Based on Bayesian network[8], this paper makes a lot of experiments on the problem of "feature extraction and regeneration of music melody based on artificial intelligence" in various specific ways, using its characteristics of high randomness in neuron selection, and designs and trains a variety of neural network models that can write music, and obtains better results. meanwhile, as a control experiment as well as different directions try, and take CNN convolutional neural network, GAN adversarial neural network as the main method, carry out secondary experiments in a variety of specific ways, design and train a variety of neural network models, and also achieve certain results.

2. RELATED WORKS
Abstract characteristics of music are demonstrated in [3]. A specific focus of music processing proposed in [4] is the audio field. they discuss some of the challenges that arise in establishing semantic relationships across various representations and point out how the established cross-relationships open new avenues for music analysis, navigation and retrieval.

Different methods to generate music content using deep learning (deep artificial neural network) are investigated and analyzed in [2]. They provide guidance on the use of deep learning in music editing from the perspective of objectives, uses, concepts, formats, coding methods, deep neural network types, challenges, strategies.

An end-to-end generative model is proposed in [5], which is able to create music under specific mixing conditions of composer style. It is also an expansion and innovation in deep neural network.

Different methods are used in the [6], trying to use GAN adversarial neural networks to process the multi-track order of symbolic music and accompaniment. Also similar in [7].

3. PRINCIPLES AND METHODS

3.1 Music Norms

A piece of music can consist of one or more instruments. To facilitate analysis, this paper uses midi music files. In midi music files, each instrument occupies one track in the midi music file, and any two tracks do not interfere with each other. This paper divides the main song into two types, the main song has only one, no specific rules, indicating the melody part of the music; while the slave track can have multiple, regular and periodic, with certain rules, such as the commonly used trend 6-4-2-1, representing the chords of the music. Therefore, a piece of music can be divided into two parts: melody and chord.

Although the melody also has certain regularity and periodicity compared with the chord, it has no very specific requirements, no strict general rules, is the soul part of a piece of music, but also the main creative part of a piece of music, the chord part has certain rules, its own regularity and periodicity is very strong, it is difficult to use a relatively simple number of music symbols to express. hence, this paper only extracts and regenerates the main track of midi music files. Also, the new music created by the network only describes the melody, contains only the main music, is a MIDI format of the music file, do not handle chords.

In addition, according to the listening sense and rhythm of music, the paper has established music libraries such as "soothing "," radical "," lyric "," Chinese style" and so on, while "Chinese style" includes music with different rhythms and different rhythms such as " Blue and white porcelain "," Fen Rao" and "Cao Cao", which can basically represent more classic" Chinese style "types of songs, and the resulting network effect is relatively good.

3.2 Based on Bayesian Networks

Bayesian network, also known as belief network, or directed acyclic graph model, is a probabilistic graph model first proposed by Judea Pearl in 1985. It is an uncertainty processing model that simulates causality in human reasoning. each of its nodes can represent different random variables, and the results generated by each use are random and therefore have better randomness than the general neural network[8].

In this paper, because the music itself has a strong randomness, it is difficult to describe the specific trend and connection mode of its notes with a set of regular rules, but there must be some correlation and similarity in the same type of songs, that is, there must be similar notes combination fragments, so we can use Bayesian network to learn and describe, through the intensive collection of a piece of music notes combination, and with the rules of contextual connection between notes as a label, you can learn the characteristics of different notes combination in a song. By learning a certain number of songs of the same type, like a song written by an author in the same period, or a large number of music original songs integrated by a music player, we can learn the combination of notes of a certain type of music, and we can create it independently through the characteristics of the relevance learned, so as to achieve the effect of intelligent composition.

In this paper, according to the basic principle of network, a set of Bayesian network function library is designed and encapsulated. Including basic training data processing, data label automatic generation, network model building, network model training, network model preservation, network model import, music melody claim, music melody packaging and other main function methods, can meet the basic needs of use.

By the time this paper is written, the designed Bayesian network is a three-layer tree-like network structure(Figure 2).

The first layer is the input layer, which inputs either a starting rhythm type or a musical score type, or a starting waveform of n term Fourier series, as the data set, and its label is the weight stored in the lower layer network.

The second layer is the connection layer, which analyzes the type of input and makes different classification. Because of the low probability that the note combination appears exactly the same, it is necessary to design the similarity rate function to judge which kind or which kind. Similar to the learning rate, the higher similarity rate will provide more accurate classification results, but the horizontal width of the network will be limited, and the randomness and innovation will be reduced when the music is generated; the lower similarity rate will reduce the classification accuracy, while the horizontal widening of the whole network can provide more random re-creation results. the recommended learning rates obtained through experimental tests are "0.9-note-type music spectrum network "and "0.6-fourier series network ", which is determined by the characteristics of the data set and the way the data is described when sampling. At the same time, the design weight of each bit in the similarity rate function will also affect the accuracy of the classification.

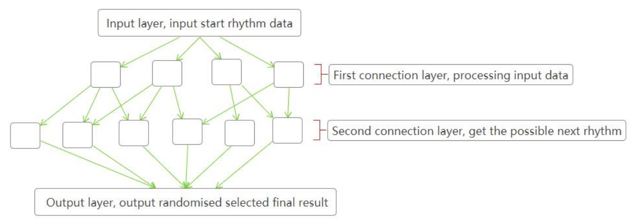

Figure 1. Schematic diagram of network structure. the Bayesian network we designed is a three-layer tree-like network structure. First layer is input layer, input is a starting rhythm or music type, or a starting waveform of n term Fourier series. The second layer is the connection layer, which analyzes the type of input and makes different classification. A third layer is a connecting layer, and the connected child node is the next associated rhythm type or music pattern learned. Fourth layer is output layer, output is a result of random prediction of rhythm and music patterns, or as n Fourier series of the result waveform.

The third layer is a connecting layer, and the connected child node is the next associated rhythm type or music pattern learned. The parameter is the number of times the rhythm type or music pattern appears; or the n term Fourier series of the next associated waveform, the parameter is the number of times the waveform appears. Because the waveform can not be exactly the same, the similarity rate function is also used here, and the learning rate parameter consistent with the upper network is used to control the width of the whole network, so that the context setting of a certain note combination can be obtained. By learning a song or a certain type of song and increasing the data set, we can get the rhythm and the music trend of a song or a certain type of song, or the general trend of the waveform.

Fourth layer is output layer, output is a result of random prediction of rhythm and music patterns, or as n Fourier series of the result waveform. the similarity rate function is also set in the result prediction to select the branch closest to the above, and if there are multiple similar branches, random selection is made again from it. After determining above, one or more of the following are randomly selected from the child nodes above, that is, the distributed probability of the following stored, and multiple new following are generated again by genetic algorithm. Finally select one from the recorded following and the newly generated following of the genetic algorithm as the result output.

Through three random selection, the randomness of music creation is preserved to the greatest extent, and because of the addition of genetic algorithm, a new combination of notes can be generated, which does not exist in the original data set but roughly accords with the characteristics of the original data set, which ensures the innovation of music creation.

3.3 Based on GAN Countering Neural Networks (Table 1,2)

GAN adversarial neural network is mainly to apply the idea of zero-sum game in game theory to the neural network. By setting up a generative network and a discriminant network, the two confront each other and optimize the two networks through the results of the game, and finally get a generative network that can generate as real a result as possible. Now mostly used to generate pictures [9].

Table 1. GAN generation network composition. Total params: 1,545,345. Trainable params: 1,544,833. Non-trainable params: 512.

dense_1	(Dense)	(None, 1024)	1025024
activation_1	(Activation)	(None, 1024)	0
dense_2	(Dense)	(None, 256)	262400
batch_normalization_1	(Batch	(None, 256)	1024
activation_2	(Activation)	(None, 256)	0
reshape_1	(Reshape)	(None, 2, 1, 128)	0
up_sampling2d_1	(UpSampling2	(None, 4, 5, 128)	0
conv2d_1	(Conv2D)	(None, 4, 5, 64)	204864
activation_3	(Activation)	(None, 4, 5, 64)	0
up_sampling2d_2	(UpSampling2	(None, 8, 5, 64)	0
conv2d_2	(Conv2D)	(None, 8, 5, 32)	51232
activation_4	(Activation)	(None, 8, 5, 32)	0
up_sampling2d_3	(UpSampling2	(None, 16, 5, 32)	0
conv2d_3	(Conv2D)	(None, 16, 5, 1)	801
activation_5	(Activation)	(None, 16, 5, 1)	0

Table 2. GAN judging network composition. Total params: 247,809. Trainable params: 0. Non-trainable params: 247,809.

conv2d_4 (Conv2D)	(None, 16, 5, 64)	1664
activation_6 (Activation)	(None, 16, 5, 64)	0
max_pooling2d_1 (MaxPooling2	(None, 8, 5, 64)	0
conv2d_5 (Conv2D)	(None, 4, 1, 128)	204928
activation_7 (Activation)	(None, 4, 1, 128)	0
max_pooling2d_2 (MaxPooling2	(None, 2, 1, 128)	0
flatten_1 (Flatten)	(None, 256)	0
dense_3 (Dense)	(None, 128)	32896
activation_8 (Activation)	(None, 128)	0
dense_4 (Dense)	(None, 64)	8256
activation_9 (Activation)	(None, 64)	0
dense_5 (Dense)	(None, 1)	65
activation_10 (Activation)	(None, 1)	0

The first half is the generation network, the second half is the discrimination network, the generation network generates the new song, then sends into the discrimination network to carry on the judgment. Because of the use of one-dimensional input, the final result is not bad, there is hope to continue to optimize, especially when considering the music chord in the later stage can generate a more complex music rhythm.

3.4 Based on CNN Convolutional Neural networks (Table 3)

CNN convolutional neural networks, similar to common neural networks, are analog biological midbrain neural networks, which consist of neurons similar to biological neurons that are interconnected with constants with learnable weights and biases. Each neuron has the function of receiving multiple input signals, signal processing, and processing the output of results [10].

It can be seen that although there are more layers of convolution neural network, it is more to force the one-dimensional input to the two-dimensional level, and then treat it as a two-dimensional picture to operate. This operation is not very effective, but it is the most direct way to deal with a dimension of songs in the current situation. However, if you need to consider the chord and melody of the music at the same time, not just the melody of the music, you can treat a song as two-dimensional or even three-dimensional data.

Table 3. CNN network composition. Total params: 125,096. Trainable params: 125,096. Non-trainable params: 0.

conv2d_1	(Conv2D)	(None, 5, 1, 1)	3
up_sampling2d_1	(UpSampling2	(None, 10, 4, 1)	0
conv2d_2	(Conv2D)	(None, 10, 4, 3)	9
activation_1	(Activation)	(None, 10, 4, 3)	0
up_sampling2d_2	(UpSampling2	(None, 20, 16, 3)	0
conv2d_3	(Conv2D)	(None, 20, 16, 8)	104
activation_2	(Activation)	(None, 20, 16, 8)	0
max_pooling2d_1	(MaxPooling2	(None, 10, 8, 8)	0
conv2d_4	(Conv2D)	(None, 10, 8, 16)	528
activation_3	(Activation)	(None, 10, 8, 16)	0
max_pooling2d_2	(MaxPooling2	(None, 5, 4, 16)	0
flatten_1	(Flatten)	(None, 320)	0
dense_1	(Dense)	(None, 256)	82176
activation_4	(Activation)	(None, 256)	0
dense_2	(Dense)	(None, 128)	32896
activation_5	(Activation)	(None, 128)	0
dense_3	(Dense)	(None, 64)	8256
activation_6	(Activation)	(None, 64)	0
dense_4	(Dense)	(None, 16)	1040
activation_7	(Activation)	(None, 16)	0
dense_5	(Dense)	(None, 5)	85
reshape_1	(Reshape)	(None, 5, 1, 1)	0

4. EXPERIMENTS

4.1 Data Format

There are two different forms of segmentation according to notes and according to waveforms.

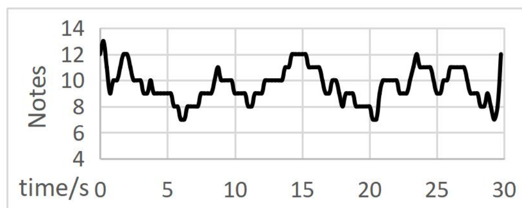

Figure 2. Music waveforms that the model needs to process. It can be seen from the music waveform that a piece of music itself has periodicity and regularity, but the amplitude and wavelength of each period are different, only the trend is similar and the shape is different, which brings great challenges to the learning of our neural network.

1) *Partitioned by note.* Taking the five notes as a group, the combination of the types of notes is defined as the rhythm type, and the combination of the time of the notes is defined as the music spectrum type. By directly learning and recording the connection law of the combination of any adjacent two groups of notes, the characteristic law of the melody trend of a certain piece or a certain type of music is obtained.

2) *Segmentation by waveform.* The song is regarded as a square wave waveform every 30 s, and then the Fourier series of certain n items is calculated by Fourier transform. By learning the connection law of any adjacent two groups of waveforms, the characteristic law of the melody trend of a certain piece or a certain kind of music is obtained.

4.2 Based on Bayesian Networks

About the characteristics of network models, first of all, using the dense collection of data, that is, each time moving a note or a waveform, each note of each song, or each of the 30 s of the waveform are collected, retaining the vast majority of the possible context of the combination of notes, to ensure the preservation of music details.

Secondly, unlike the conventional neural network, this neural network draws the probability curve by learning the occurrence probability of each possible rhythm type, and predicts the next possible rhythm type according to the curve, and the prediction result of each time is uncertain(Figure 3). As far as possible in order to ensure the periodicity and stability of music, produced in the original data set similar to the composition of notes, and in the network design to increase the randomness of the composition generation, so that music is no longer like the general neural network, there are easy to fall into too random and too mechanical rigid two extreme problems.

Finally, the whole network model is modularized to realize the interface of basic functions such as network data acquisition, network training, network prediction, network preservation, network import and so on.

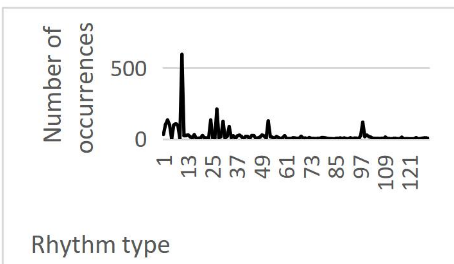

Figure 3. Probability of Rhythm type. It can be seen from the picture that there are many different rhythmic patterns in a piece of music, but most of the rhythmic patterns may appear only once or several times; and there is also one or more kinds of music, which appear much more frequently than other types, which shows that this rhythm type is the "basic rhythm type" of the music.

4.3 Based on GAN Countering Neural Networks

In this paper, because the fundamental purpose is also to make the computer by learning a certain number of music, from which to sum up the rules and generate new music, similar to the idea of GAN network, so there is the possibility of implementation according to the GAN network. A design generation network accepts a 1000-dimensional random array, and then outputs a 16x5x1 three-dimensional array data representing the composition of notes, including rhythm and music patterns. Design discriminant network accepts a 16x5x1 three-dimensional data, the output is 1 or 0, representing whether it meets the requirements.

First pass in the generated network, generate the network to generate music fragments, the music fragments into the discriminant network, the label is now passed in 1, that is, on behalf of the "real" music, but in fact is generated by the network to produce "false" music, now the discriminant network will be judged as false.

After that, the model will constantly adjust the generated network parameters, and the parameters of the discriminative network are set to be unadjustable, so in order to continuously reduce the loss value, the model will always adjust the parameters of the generated network until the discriminative network thinks it is "real" music.

At the moment, the discrimination network and the generation network reach a balance. That is to generate "false" music generated by the network, discrimination network has been unable to distinguish. So continue to iterate, improve the accuracy of the discrimination network, so reciprocating cycle, until the generation of music that even people can not distinguish.

4.4 Based on CNN Convolutional Neural networks

During this paper, by using CNN neural network, we study the probability and time of the occurrence of all the notes in a certain music library, and use big data to obtain the probability of the occurrence of a particular note in a certain type of music style, and carry out random combination to get new music.

The basic method is using CNN neural network, divides music into multiple groups of rhythmic fragments, and then learns the connection probability between rhythm type and rhythm type. The existing rhythm type is connected by probability, and the notes are obtained by big data to recombine to obtain the music.

In this regard, four main methods have been tried. Including:

3) *get the beat type of the song, modify the note with big data to get the new song with the same original beat type;*
4) *get the rise and fall of the note with the same original beat type, get the new song at random;*
5) *get the rhythm type probability through the beat type of the song, get the new song through the big data note;*
6) *get the rhythm type and rhythm type connection probability, connect the existing rhythm type through the probability, get the new song through the big data note.*

5. EXPERIMENT(FIGURE 6)

5.1 Experimental Results of Bayesian Network

Up to the writing of this paper, there are five kinds of model design methods, and there is a certain progressive between the different design methods.

7) *Learn how to connect the rhythm type and get the probability of the rhythm type.* The next rhythm type is obtained by giving a starting rhythm type in advance, and then the new rhythm type is used as input to obtain the next rhythm type, which forms a chain structure and is finally packaged into a new piece of music. The experimental results show that the newly formed music is easy to fall into some fixed rhythm cycle, and can only be reconnected according to the existing rhythm type to get the music, and can not produce the new rhythm type.

8) *Learn the connection between the rhythm type and the music type to obtain the probability distribution of each bit of the next note combination.* It is basically the same as method 1, but improves the circulation mode, and uses the probability of occurrence and time length of each note to learn, and regenerates the rhythm type and the music spectrum type respectively in the newly formed note combination, thus solving the problem that can not produce the new rhythm type and easily fall into the fixed note combination cycle.

9) *Learn how to connect rhythm and music patterns, set the probability function, and get the connection probability of the next*

note combination. on the basis of method 2, the importance parameters of each bit are adjusted. When predicting the next rhythm type, because the probability of exactly the same rhythm type is very low, it is necessary to set a similar function to judge the similarity rate of two rhythm types, so as to determine the next rhythm type while maintaining a high randomness, and the influence of different notes on the similarity rate is different. Since the first and last notes are involved in the link between the two rhythmic patterns, the design parameters are large because of the high degree of influence on the similarity rate, and the middle notes represent the randomness and the central trend of the combination of the notes. here the final choice is to use quadratic functions to approximate simulate its importance parameters. Experimental results show that the new music is more natural, with the lowest sense of Mechanical sense and violation and higher originality, and is the best model in all models.

10) Learn the connection mode between rhythm type and music spectrum type, set the probability function and add genetic algorithm to obtain the connection probability of the next note combination. On the basis of method 3, genetic algorithm[12] is added to generate the result of three times model for the same input rhythm type. the newly generated note combination is taken as the parent class, and the simulated genetic algorithm is used to combine two groups to generate new subclasses, and a note combination is randomly selected from the parent class and the subclass as the final result. the genetic algorithm is added to improve the randomness of the whole model. Experimental results show that the result is not as good as method 3 because the randomness is too high.

11) Learn the connection mode between 30 s waveforms to obtain the connection probability of the next waveform. Instead of using the "rhythmic note type" combination to describe the note combination, the Fourier transform is used to obtain the expression description of a certain 30 s music waveform. In the model, the probability of the connection law of the waveform is studied and recorded, and the new results are obtained by genetic algorithm.

By Fourier transform, a periodic function waveform can be described by the sum of a finite sine expression and a cosine expression. Since this experiment uses 30 s of music waveform, it can be expanded into a period function of 30 s to describe the 30 s waveform.

$$\begin{cases} a_n = \dfrac{2}{b-a}\displaystyle\int_n^b f(x)\cos\left(\dfrac{2\pi nx}{b-a}\right)dx \\ b_n = \dfrac{2}{b-a}\displaystyle\int_n^b f(x)\sin\left(\dfrac{2\pi nx}{b-a}\right)dx \\ f(x) = \dfrac{a_0}{2} + \displaystyle\sum_{n=1}^{\infty}[a_n \cos(\dfrac{2\pi nx}{b-a}) + b_n \sin(\dfrac{2\pi nx}{b-a})] \end{cases} \quad [11]$$

where a is the interval starting point, b is the interval terminating point, $b-a$ is the function period, here is the set 30 s, $f(x)$ is the function to be transformed, here is the sound wave function.

Because of the midi music file, the sound wave function can be regarded as a complex square wave format, and the waveform can be described by different number of sampling points.

The main factors affecting the results of Fourier transform are the number of Fourier series terms and the number of sampling points. The fitting relationship between the number of different Fourier series terms and the waveform is obtained by experiments(Figure 4).

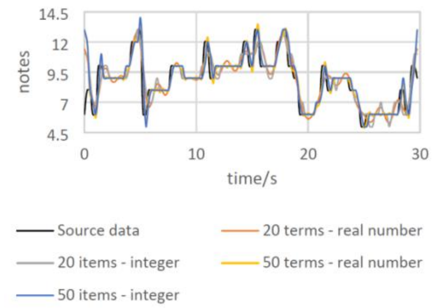

Figure 4. Waveform description of different number of terms. The black waveform in the figure is the original music data, while the other four waveforms are fitted by the Fourier series variation under different terms. It can be seen that the waveform obtained by 50 Fourier series is obviously closer to the original waveform than that obtained by 20 Fourier series; the waveform after rounding is closer to the original waveform than that in the case of real number. The results show that the waveform with the largest number of Fourier series can be used and the obtained waveform can be rounded.

Black is metadata, red and yellow are waveform descriptions for different items, gray and blue are waveform descriptions after integerization.

This figure represents the comparison between the waveform described under different Fourier series terms and the original waveform. it can be seen that as the number of terms increases, the function waveform described is closer to the waveform obtained from the original sampling. At the same time, because the note type of the audio file processed in this experiment can be represented by integer, the original waveform to be described is square wave in integer format, so the Fourier series expression can be integerized, and the final calculation function waveform can be obtained.

Although the similarity with the original waveform increases with the increase of the number of terms, it will also increase the running time, so the relationship curve between the number of terms and time and the similarity rate can be obtained by experiments on different Fourier series terms with the number of uniform sampling points. The relationship between the number of different Fourier series terms and the running time and the fitting effect is obtained(Figure 5).

It can be seen from the image that there are two main inflection points, that is, the first inflection point appears when the number of terms is 35, and the change tends to smooth after this inflection point; at the second inflection point when the number of terms is 90, the similarity rate will have a sharp drop and then continue to smooth. the optimal number of terms and the number of sampling points is determined by experiments to be 35-600 or 90-600.

Because there is no clear scalar to divide the advantages and disadvantages of music, this paper adjusts the parameters of model design by randomly selecting a certain number of people to evaluate the advantages and disadvantages of the generated music.

After the model generation training, the obtained music waveform has been close to the general music waveform, and after testing the music has a higher sense of listening, basically meet the needs of the arranger.

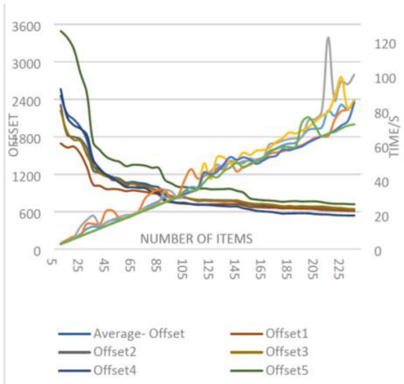

Figure 5. Relationship between number of terms and time and similarity rate. The curves of offset and processing time for different songs under different Fourier series terms are shown respectively. It can be seen that with the increase of the number of Fourier series terms, the processing time required basically shows a linear rise, while the offset of the original waveform obtained shows a downward trend, and there are two obvious inflection points when the number of terms is 30 and the number of terms is 90.

5.2 Experimental Results of GAN Neural Network

The results obtained are in the middle level in all the network models of this paper, and the resulting music has the disadvantages of high randomness. During the experiment, the loss oscillation occurred in the later stage of network training. The Bayesian networks can be embedded into the generating network part of the GAN network in the following experiments because of the characteristics of the self-learning network of confrontation network.

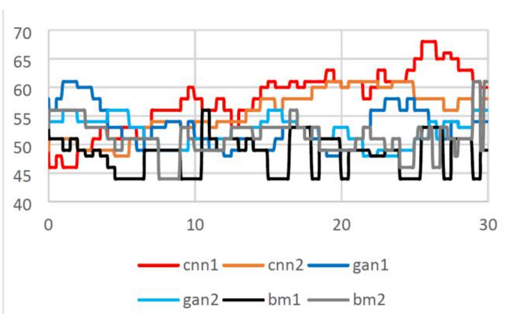

Figure 6. Summary of final experimental results. This graph shows that in the final results, although CNN and GAN can achieve certain results, but the results of the waveform is disorganized, showing that the music is rapid and changeable. The results obtained by Bayesian network method are relatively close to square wave and more stable.

5.3 CNN Convolutional Neural Network

Because the notes are obtained by big data, it is almost impossible to have the same rhythm as the existing music library, and the resulting music is more innovative and random, and the experimental results show that the response of the music results is surprisingly good. However, there is a problem that the feature extraction and learning of the music are relatively poor, and the model generated by the training of different musical media is less differentiated.

6. CONCLUSIONS

This paper adopts three ways of Bayesian network, GAN confrontation neural network and CNN convolution neural network respectively, and makes various attempts on artificial intelligence orchestration direction, and designs, implements and trains twelve different neural networks.

Experiments show that artificial intelligence, as a frontier science, can not only be used for classification tasks and data processing tasks, but also can be applied to artistic creation fields such as music and painting. the Bayesian network algorithm can be applied to artificial intelligence music creation, and the results obtained are more stable and innovative than CNN convolutional neural network and GAN confrontation network, and the results are better.

However, there are still "deadlocks" in this experiment, that is, because the Bayesian network is learning the probability of rhythm type, it may fall into some rhythm type cycle in the generation of music. Although the introduction of genetic algorithms for mitigation, but still can not be completely avoided.

REFERENCES

[1] Briot J P , Hadjeres, Gaëtan, Pachet, François-David. Deep Learning Techniques for Music Generation -- A Survey[J]. 2017.

[2] Hinton G E . Boltzmann machine[J]. encyclopedia of machine learning, 2007.

[3] Wiggins G A , Daniel Müllensiefen, Pearce M T . On the Non-Existence of Music: Why Music Theory is a Figment of the Imagination[J]. Musicae Scientiae, 2010, 14(1 suppl):231-255.

[4] Muller, Mathias, Pratzlich, Thomas, Bohl, Benjamin, & Veit, Joachim. (2013). Freischutz digital: a multimodal scenario for informed music processing.

[5] Mao H H , Shin T , Cottrell G W . DeepJ: Style-Specific Music Generation[J]. 2018.

[6] Dong H W , Hsiao W Y , Yang L C , et al. MuseGAN: Multi-track Sequential Generative Adversarial Networks for Symbolic Music Generation and Accompaniment[J]. 2017.

[7] Dong H W , Yang Y H . Convolutional Generative Adversarial Networks with Binary Neurons for Polyphonic Music Generation[J]. 2018.

[8] Kyburg H E . Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference by Judea Pearl[J]. journal of philosophy, 1991, 88(8):434-437.

[9] Goodfellow I J , Pouget-Abadie J , Mirza M , et al. Generative Adversarial Networks[J]. Advances in Neural Information Processing Systems, 2014, 3:2672-2680.

[10]]Krizhevsky A , Sutskever I , Hinton G . ImageNet Classification with Deep Convolutional Neural Networks[J]. Advances in neural information processing systems, 2012, 25(2).

[11] Fourier, Jean Baptiste Joseph. The Analytical Theory of Heat[J]. 1906.

[12] HOLLAND,J. Adaptation in natural and artificial systems : an introductory analysis with application to biology[J]. control & artificial intelligence, 1975.