J. COMBIN. MATH. COMBIN. COMPUT. 127b (2025) 2161--2174

Journal of Combinatorial Mathematics and Combinatorial Computing
www.combinatorialpress.com/jcmcc

A Multi-Level Accompaniment Effect Generation Mechanism Incorporating AI Computing in Piano Art Instruction

Shuang Du[1,✉]

[1] *School of Music and Dance, Weifang University, Weifang, Shandong, 261061, China*

ABSTRACT

This paper takes the integration of AI technology into piano teaching as the starting point, generates accompaniment rhythms through AI computation, adopts deep learning model to generate accompaniment, and builds a multi-level accompaniment effect generation mechanism. Taking the MuseFlow model as the base model, the generative adversarial network and variational autoencoder are introduced to optimize the structure in a limited arithmetic environment. Quantitative and manual evaluations are used to measure the accompaniment generation effect of the proposed mechanism, and controlled experiments are designed to explore its practical application effect. The results show that the improved MuseFlow model generates accompaniment with an average pitch distance of 0.92, which is 0.15 smaller than that of MMM, and the overall score reaches 4.18. The scores of the experimental group in all six abilities are significantly higher than those of the control group, the degree of students' positive response to each ability increases to some extent, and the number of students who consider the ability of melodic creation to be at a satisfactory level is 18 more than that of the pre-experiment after the experiment.

Keywords: piano teaching, MuseFlow, generative adversarial network, variational autoencoder, musical accompaniment generation

1. Introduction

Under the immersion of the contemporary Internet environment, artificial intelligence technology is in a period of rapid development, promoting the continuous iteration and innovation of all walks of life, and the art of music has evolved into a variety of forms of expression [1-2]. As the "king of musical instruments", the piano first conformed to the pattern of the development of music, art and culture in modern society, so the smart piano appeared in the public's field of vision early and gradually applied to piano courses on various educational platforms, such as children's piano initiation classes, piano group classes, online piano sparring classes, and even some piano-related courses in professional

✉ Corresponding author.
E-mail address: WFXYYWDS@163.com (S. Du).
Received 15 September 2024; Revised 05 November 2024; Accepted 20 February 2025; Published Online 16 April 2025.
DOI: 10.61091/jcmcc127b-122

music schools [3-5]. As we all know, the piano solo performance program offered by the piano department of the Conservatory of Music is specifically designed to cultivate and train students' skills in all aspects of piano performance, and the program, which emphasizes the need for teachers and students to practice in their performance, has already been combined with the use of smart pianos in teaching, and there are numerous academic and scientific research projects related to it [6-8].

In terms of basic teaching in the piano art instruction classroom, the realization of intelligent multimedia has actually entered in the last century, and art teaching has changed relatively dramatically as a result [9-10], however, in spite of this, there is not a specific and well-established law of use for exploring new media as a means to help the implementation of teaching and learning in the piano art instruction classroom because the piano art instruction classroom is different from other subjects, which are more objective and teach students more empirically [11-12]. While the piano art instruction course is still in the traditional piano by the teacher to play for the students accompaniment, this course status quo in the current Internet situation is slightly conservative and outdated [13-14]. Therefore, if we can make full use of the function system of smart piano to replace the piano art instruction teacher to accompany the students, not only can we change the monotonous and boring classroom state, but also can solve many dilemmas faced by the teachers and students in the traditional piano art instruction course, so that this course can be carried out more efficiently and qualitatively [15-16].

This paper combs through the current use of AI technology in piano art instruction, and proposes to combine AI computation with deep learning models to generate piano accompaniment. DAW is utilized for pitch extraction, and the accompaniment rhythm is generated according to the characteristics of the accompaniment pattern. The base model, MuseFlow model, is improved by adopting a generative adversarial network as the generator of MuseFlow model and utilizing a variational autoencoder for encoding. A sliding window model is designed to assist training and generation, and the accompaniment is generated by combining the program-generated rhythms. The improved version of MuseFlow model is subjected to controlled experiments with five baseline models to compare their performance levels using quantitative and manual evaluation. The accompaniment generation mechanism proposed in this paper is put to use, and the pre- and post-experiment data comparisons are utilized to explore the effectiveness of its implementation in terms of both student ability and self-perception.

2. The use of AI technology in piano art instruction

In recent years, with the development of modern AI technology, the intelligence of piano learning is increasing day by day.The fusion of AI and music has produced a lot of intelligent composition software, like AemperMusic, OrbComposer, etc., which provides students with more ways to compose. Students can arrange different timbres and melodies so that they are presented in a certain order and frequency according to their needs. In the classroom, students can utilize the software more freely to generate and share accompaniments, thus increasing their enthusiasm and participation in learning, thus increasing their interest in learning and improving their piano learning ability. For example, with AemperMusic, after students input basic musical elements such as melody, playing instrument, rhythm, key tonality, etc., AemperMusic will be able to complete a new piece of music in a few seconds, and after that, it can also be adjusted to the timbre, harmony, and other elements to form different styles of accompaniment to satisfy the needs of students in class.

This type of music app is ideal for schools and colleges. In primary and secondary schools, the use of intelligent music software to generate accompaniment can inspire students' creativity, enrich their imagination and enhance their creativity; in colleges and universities, the use of intelligent music software to generate accompaniment can help college students better inspire their creative passion, and can also help them better utilize the intelligent software. In piano art teaching, whether in hardware or software, teaching aids have become more intelligent and diversified, but at the same time, it also puts forward a new demand for contemporary piano instructors, which is to improve their artificial intelligence literacy. In this era of rapid development of artificial intelligence, the traditional way of piano art instruction is no longer compatible with the modern educational needs, so as a contemporary piano art instructor, you should change your teaching concept, take the initiative to learn and use intelligent music software, so as to improve your level of operation of intelligent software.

3. Design of accompaniment effect generation mechanism combined with AI computing

3.1. Accompaniment tempo generation based on AI computation

3.1.1. Tone pattern characteristics:

1) Harmonic complexity. A high degree of harmonic complexity, using mostly seventh and ninth chords, and a few thirteenth chords, with the use of complex chords with omitted notes: generally switching chords every bar, but in the case of a 4/4 meter, it is possible to switch chords on the third or fourth beat.

2) Rhythmic characteristics. There are 4/4 and 2/4 beats, and in the rhythmic arrangement, the accent is always on the first beat of each measure, but there will be an eighth rest or a quarter rest on the second or fourth beat.

3) Vocal Characteristics. The accompaniment weave is dominated by columnar chord repeats, and the bass voice is a combination of octave repeats or octave superimpositions.

4) Characteristics of the register. From the point of view of the sound area, the middle and low soprano areas are distributed, but most of them are concentrated in the middle and low sound areas.

3.1.2. Pitch extraction. The process of extracting pitches means that in the environment where the DAW is running, MIDI events are transmitted when chord notes are entered from the input source MIDI keyboard or from the piano window, and the MIDI events contain various types of MIDI messages, from which we want to extract the MIDI note messages corresponding to the chord notes.

When a key is pressed on a MIDI keyboard, the keyboard sends a note on message on the MIDI OUT port. So we can get the pitch of the MIDI note from the note on message. Then we can get the pitch information of the MIDI note in the note on message, and then we can get the pitch of each value according to the MIDI Pitch Comparison Table.

After getting the MIDI pitch information of the complete chord through the MIDI keyboard input or the MIDI track of the DAW, we enter the following steps to filter and reset the pitch information, and the complete pitch extraction process is shown in Figure 1.

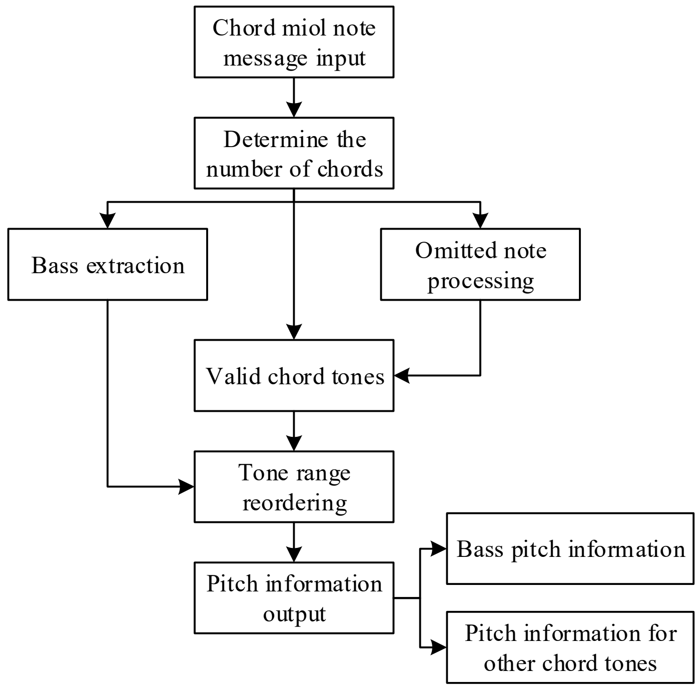

Fig. 1. Pitch extraction process

3.1.3. Rhythm generation. Rhythm generation for the piano part is accomplished by a rhythm sequencer, which is designed in the following steps:

Step 1: The minimum rhythmic unit is set to 16 cents, and the beats are 2/4 or 4/4, with a total of 4 beats in a cycle.

Step 2: Since the pitch extraction process has been divided into two parts: bass pitch information and other chord pitch information, the rhythm generation process is divided into two layers: bass rhythm generation and upper-midrange rhythm generation, which are used to output the rhythmic fills for these two parts.

3.2. Deep Learning Based Accompaniment Generation Modeling

The core of artificial intelligence music generation technology lies in the deep generative model, how to choose a suitable generative model and apply it to the field of music generation is this paper is the focus of research. Different from other deep learning models, the Flow model is highly decoupled and bi-directionally reversible. Attribute control of the generation results can be realized by vector operations on random variables.

MuseFlow is a deep learning model implemented based on the Flow model applied in the field of music generation, and the MuseFlow model has a good property: forward training is the encoding process, which maps the music material into a multi-dimensional Gaussian space through multiple encoders; the reverse operation is the decoding process, and the trained encoder is the decoder. The model generation converts randomly drawn Gaussian random variables to music through multiple decoders. Piano rolls and Midi files are easily converted to each other and can be easily interfaced with various arranging and orchestrating software. The input during model training and the output during generation use the uint8 piano roll format to represent the music.

3.2.1. Generating Adversarial Networks. In this paper, a generative adversarial network is used as a generator for the MuseFlow model. The essence of generative modeling is to train the model to describe all sample data distributions with a known probability, thus establishing a link between the two. Probability distributions are non-negative and normalized, so a neural network that can fit an arbitrary function is not directly useful for fitting probability distributions. Arbitrary probability distributions can be fitted with the Gaussian distribution integral of Equation (1), where $q(z)$ is generally a standard Gaussian distribution and $q_\theta(x \mid z)$ represents an arbitrary conditional Gaussian or Dirac distribution.

$$q(x) = \int q(z) q_\theta(x \mid z) dz \tag{1}$$

Assuming a real sample data distribution $p_{data}(x)$, the model is generated in an attempt to solve for the parameters θ of the $q_\theta(z)$ distribution as close as possible to $p_{data}(x)$. Parameter solving in probability theory is often done by the maximum likelihood approach, Eq. (2), which uses known information about the sample outcomes to back-propagate the values of the model parameters that are most likely (with maximum probability) to lead to the occurrence of those sample outcomes. The generative model is an estimate of a probability distribution with parameter θ, and the principle of maximum likelihood is to choose the parameter that maximizes that probability.

$$\theta = \arg\max \sum_{i=1}^{m} \log q\left(x^j; \theta\right) \tag{2}$$

The birth of Generative Adversarial Neural Networks (GANs) has been highly evaluated in the industry as "the most valuable idea in twenty years", and it is also a landmark technological breakthrough in the field of generation. The structure of the generative adversarial network is shown in Figure 2, the optimization objective of the GAN is equation (3), and its design idea is to use game theory to continuously optimize the generator and discriminator, so that the generated results are more and more similar to the real training set. The encoder extracts random vectors and converts them into images, and the generated results are passed to the discriminator together with the real data, and the discriminator acts as a teacher to judge the generation effect of the "student" generator, and the ideal result is achieved when the generator can generate fake images and the discriminator cannot distinguish them, i.e., the probability of 0.5. The GAN adopts the adversarial learning approach, which cleverly avoids the need to use game-theoretic methods to optimize the generator and the discriminator to achieve increasingly similar results to the real training set. The adversarial learning approach of GAN cleverly avoids probabilistic calculations, but at the same time, the absence of an encoder in the model usually prevents the data points from being directly represented in the latent space. The advantage is that it can generate clear images, but the disadvantage is that the training process is unstable and prone to collapse problems.

$$\begin{aligned} Min_G \max_D V(D,G) = \mathrm{E}_{x \sim p_{data(x)}}[\log D(x)] \\ + \mathrm{E}_{z \sim p_{z(x)}}[\log(1 - D(G(z)))] \end{aligned} \tag{3}$$

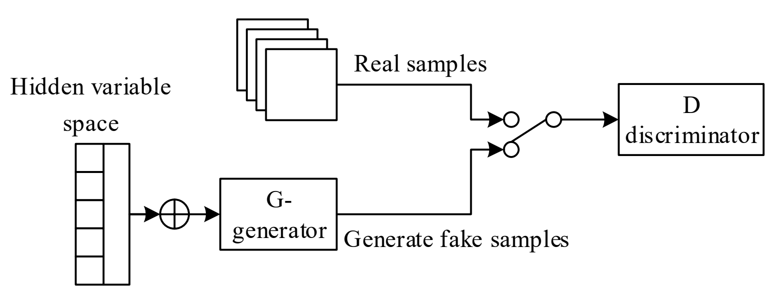

Fig. 2 Structure of GAN model

3.2.2. Variational autoencoders. The MuseFlow model maps music data into Gaussian space during training through multiple encoders, each with the same internal structure but different weights. The original Flow's encoder is suitable for image generation, but can not be directly applied to music accompaniment such as multi-channel data generation, for this reason, this paper introduces the variational autoencoder into the MuseFlow model for structural optimization.

Variational autoencoder consists of encoder and decoder, its working principle is that the encoder will map the training data to Gaussian space, the decoder can also be regarded as a generator, from the Gaussian space randomly sampled and added to the Gaussian noise to generate sequences. The structure of the variational autoencoder is shown in Fig. 3, and the difference between its structure and that of the self-encoder is that Gaussian noise is added in addition to the encoder that calculates the mean value to increase the robustness of the decoder. The dispersion of the noise from the current point depends on the variance, and when the variance is 0 it is an auto-encoder. During training, the encoder maps the real data to Gaussian space, and the decoder generates the image from the latent variable space by taking the Gaussian variables with added noise; during generation the decoder converts from extracting random variables to generating the result.

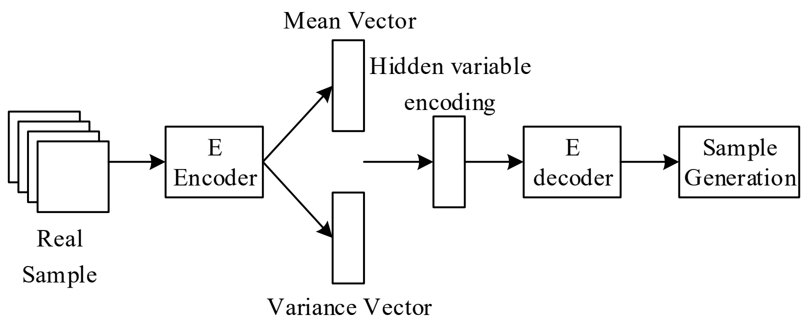

Fig. 3. Structure of the VAE model

The optimization objective of VAE is to minimize Eq. (4) by $E_{z\sim q_\lambda(z|x)}\left[\log p_\theta(x\mid z)\right]$ measuring the gap between the generated results and the real data, and $KL\left(q_\lambda(z\mid x)\|p(z)\right)$ calculating the degree of similarity between the encoder distribution $q_\lambda(z\mid x)$ and the decoder distribution $p_\theta(x\mid z)$. There is also an adversarial idea in VAE, where the reconstruction loss and the KL loss of Eq. (4) are opposites during the training process, and the two reach equilibrium when the noise points are evenly distributed around the center of the potential space.

$$ELBO = E_{z\sim q_\lambda(z|x)}\left[\log p_\theta(x\mid z)\right] - \max\left(KL\left(q_\lambda(z\mid x)\|p(z)\right) - \tau, 0\right) \tag{4}$$

3.2.3. Structural optimization in a finite arithmetic environment. The characteristics of the Flow model require equal dimensions of inputs and outputs, and a large amount of computation during training, so the computer memory space limits the length of the generated music clips.

To address this problem, this paper designs a sliding window (sliding window) mode to assist training and generation. When the data is processed to segment the music clip, the sliding window mode is used, and the window size is set to the length of the music clip, and the window size is shifted backward half a window size each time. There is continuity in the music melody itself, and the internal neural network of the encoder uses Bi-GRU to generate the result X_i, which is related to the first half of the input, in addition to the effects of the piano melody and the random variable Z_i. The sliding generation process is shown in Fig. 4, where a cell represents a bar, G represents the MuseFlow generator, and the subsequent generation process is the random variable $\left[Z_i, Z_{i+1}\right]$ of the two bars passing through the generator G to get $\left[X_i, X_{i+1}\right]$, retaining X_{i+1} in juxtaposition with the first $\left[X_1, X_2, \ldots, X_{i-1}\right]$, in addition to the X_1 and X_2 segments.

The sliding window is designed to generate longer music fragments in a limited arithmetic environment, and ideally the generation can be done directly by the length of the main melody when the computer resources are sufficiently adequate. In the experimental environment of this paper, the upper limit of the window size is set to 8 bars when increasing the length of the unit segment as much as possible, and the segmentation is moved backward by 4 bars each time. When the main theme of the music is less than or equal to 8 bars when the back end to fill the empty syllables, the model can be generated once; when the main theme of the music is greater than 8 bars, according to the 8 bars as a unit of sliding window assisted generation.

The advantage of the sliding-window generation method is that it can obtain music sequences of arbitrary length, but there are potential problems: Z The randomness of the hidden variables can easily lead to the poor articulation of the two music segments generated; if the sequence is too long, it may also produce the problem of long sequences with small correlation between the first and the last.

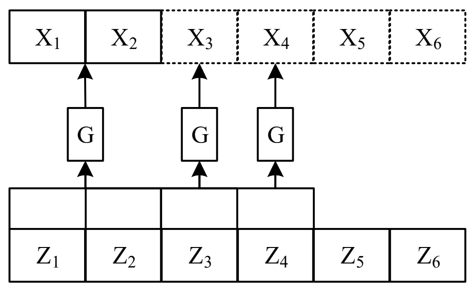

Fig. 4. Sliding window generation

4. Evaluation of the accompaniment generation effect of MuseFlow model combined with AI computation

4.1. *Improved MuseFlow Modeling Accompaniment Generation Effects*

To verify the effectiveness of the improved MuseFlow model for accompaniment generation, an example of the generated music is given in this section, and the results are measured using both quantitative and manual evaluations.

The dataset used in this section is the subset lpd-5-cleansed of the Lakh Pianoroll Dataset dataset. This subset consists of 19,746 data files in .npz format, each containing five tracks: a piano track, a guitar track, a bass track, a string track, and a drum track. This chapter performs the following preprocessing operations on the dataset: first, using pypianoroll, a Python library, each .npz file is converted into a matrix of five piano rolls (containing two dimensions: the number of time steps and the pitch range of the notes), and each piano roll matrix is split into segments of 500 time steps in length; second, the notes whose pitch is lower than C1 or higher than C8 are removed (because pitch very low and very high notes are uncommon); and finally, removing music where the piano track is empty because the piano track needs to be used as a conditional track and the condition cannot be empty. The final processing was completed to obtain 187,684 samples.

The model proposed in this paper is compared with five baseline models which are.

MuseGAN: A GAN-based neural network model for symbolic music generation that supports both generation from scratch and track-conditioned generation.

C-RNN-GAN: combines GAN and RNN in order to better generate music that matches a given condition, with bi-directional LSTMs inside the generator and discriminator.

MusicVAE: VAE-based neural network model for symbolic music generation, constructs a hierarchical decoder to model note sequences, and RNN as a bootstrap to assist in the generation of long sequences, supporting multi-track music generation.

MuseNet: A neural network model for symbolic music generation based on Transformer, trained by OpenAl's 24-head 72-layer Transformer model structure, which supports two modes of generation: simple and advanced, and can generate music of ten different instruments and styles.

MMM: Transformer-based neural network model for symbolic music generation, proposed a way to encode music data generated by 128 instruments, and trained the 8-head and 6-layer GPT2 model to generate multi-track music accompaniment.

4.1.1. Generating results. Using the improved MuseFlow model in this paper to generate the piano accompaniment, a sample piano roll is generated as shown in Figure 5, where orange indicates piano, blue indicates drums, green indicates guitar, red indicates bass, and purple indicates strings. The generation is done by using the piano track as the conditional input to the model (i.e., as the main melody) to generate the other four instrument tracks (i.e., the accompaniment to the main melody). It can be observed through the illustrations:

1) Since the piano track is the conditional track, the guitar track, bass track, and string track tend to play the same chords and the same musical scale.

2) The drum track is usually empty. If the drum track is not empty, it will be very sparse and fragmented.

3) String tracks are often long and lush.

4) The guitar track, bass track, and string track sometimes overlap (creating dark lines), which indicates a good harmonic relationship between them.

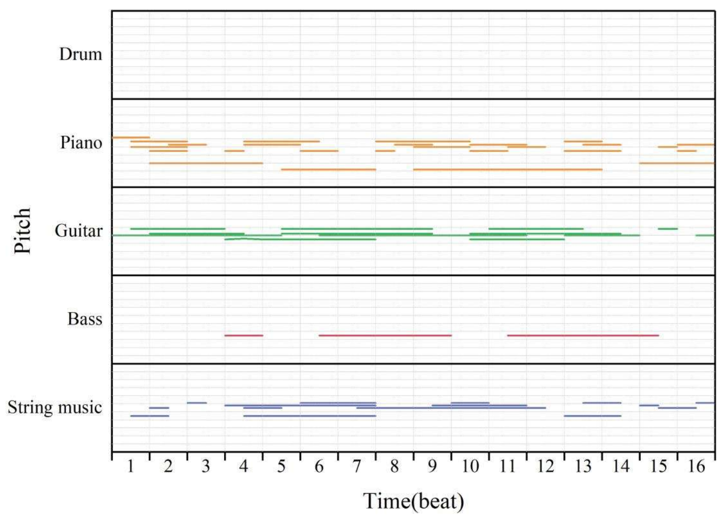

Fig. 5. Sample piano roll

4.1.2. Quantitative evaluation. In this paper, we use the pitch distance metric to evaluate whether the model learns the joint distribution of multiple tracks. The pitch distance metric is able to compute the harmonic relationship and harmony between two tracks. Specifically, it measures the distance between the chromatic features (one chromatic feature per beat) of two tracks on the pitch space. The smaller the pitch distance, the closer the harmonic relationship between the tracks and the stronger the harmony.

The pitch distances between the six pairs of tracks are shown in Table 1. In the table, P represents piano, G represents guitar, B represents bass, and S represents strings. The values in the first row of the table are the pitch distances calculated on a sample of the dataset to serve as a baseline reference. The pitch distances of the preprocessed dataset are also given in the table for comparison and reference. On the one hand, given that smaller pitch distances are better, the table shows that the pitch distances between the six pairs of tracks of the model proposed in this paper are better than the MMM, which is the best performer in the baseline model, with an average pitch distance of 0.92, which is 0.15

smaller than the MMM. It shows that the proposed model in this chapter generates music with stronger harmonic relationships between the tracks and is able to take into account the harmonic relationships between the cross-tracks, which makes it more suitable for multi-track generation. On the other hand, from the table, it can be seen that the pitch distance between the two-two tracks of the dataset and the variation of the pitch distance between the two-two tracks of the proposed model are the same, which means that the proposed model learns the characteristics of the tracks from the dataset well.

Table 1. Comparison of tonal distances between the six pairs of tracks

Data set/Method	Tone distance						
	P-G	P-B	P-S	G-B	G-S	B-S	Mean value
Dataset	0.98	1.52	0.96	1.51	1.02	1.53	1.25
MuseGAN	1.11	1.29	1.12	1.03	1.14	1.35	1.17
C-RNN-GAN	0.88	1.61	1.08	1.15	0.98	1.24	1.16
MusicVAE	0.89	1.21	1.16	1.43	1.12	1.42	1.21
MuseNet	1.37	1.19	0.89	1.43	1.25	1.33	1.24
MMM	1.07	1.18	1.03	1.27	0.92	1.19	1.11
MuseFlow	0.83	1.07	0.87	1.03	0.62	1.08	0.92

4.1.3. Manual evaluation. Although the previous analysis is useful for completeness checking, such an analysis is not sufficient to show whether the generated samples sound good or not to the human ear. Therefore, this chapter conducts another manual evaluation investigation to compare how well different models generate samples. In this chapter, a listening test investigation was conducted in which 100 participants were asked to score the proposed model and the five baseline models on the basis of the presence or absence of pleasing melodies, the presence or absence of pleasing harmonies, and an overall rating. The scoring criteria used was a 5-point scale using a Likert scale, with higher scores representing better performance. The scoring results are shown in Table 2, where the values labeled in the first row of the table are the scores of the manual evaluation of the data set samples. From the results presented in Table 2, it can be seen that the proposed model in this chapter is better than the baseline model in terms of melody and harmony, with an overall score of 4.18, which is higher than the five baseline models by 0.37, 0.29, 0.91, 1.13, and 0.13, respectively.In addition, the scores of the proposed model are closer to the scores of the dataset, which means that the samples generated by the proposed model are closer to the real data and the quality of the generated samples is superior.

Table 2. Comparison of results of manual evaluation

Data set/Method	Melody	Harmony	Overall rating
Dataset	4.19	4.18	4.21
MuseGAN	3.78	3.75	3.81
C-RNN-GAN	3.82	3.85	3.89
MusicVAE	3.08	3.22	3.27
MuseNet	2.79	2.99	3.05
MMM	4.01	4.03	4.05
MuseFlow	4.21	4.13	4.18

4.2. *Effectiveness of AI-generated accompaniment in piano art instruction*

In order to verify the effectiveness of the proposed model, this paper selects piano performance majors of a music college as the research object, and selects 50 students from the sophomore to the fourth year of college, all of whom have a certain degree of piano playing foundation, and are able to independently play piano music of a certain degree of difficulty.

The 50 students were divided into two groups, and the experimental group practiced with the accompaniment generated by the improved MuseFlow model proposed in this paper. In the experimental group, the students practiced with the accompaniment generated by the improved MuseFlow model. The improved MuseFlow model automatically generates the accompaniment based on the harmonic, rhythmic, and melodic patterns of the students' practice pieces. Students in the control group practiced with the traditional accompaniment. The experimental period was 4 weeks, at the end of which the experimental group and the control group were assessed for their competence, and the feedback from the experimental group was collected through a questionnaire.

4.2.1. Setting of experimental evaluation indicators. For both groups, the teaching experiment indicators as well as evaluation criteria are shown in Table 3.

Table 3. Teaching experiment evaluation

	Outstanding	Good	Normal	Range	Poor
Sound recognition	5	4	3	2	1
Accurate rhythm	5	4	3	2	1
Finger technique	5	4	3	2	1
Music understanding ability	5	4	3	2	1
Musical expression	5	4	3	2	1
Ability to compose melodies	5	4	3	2	1

4.2.2. Results and analysis of capacity evaluation data. After the 4-week experimental period, the two groups of students were evaluated quantitatively on their abilities in six aspects: note recognition, rhythm, technique, musical comprehension, musical expression, and melody creation ability, and the statistical data of the ability evaluation results are shown in Table 4.

Table 4. Statistical data of capability evaluation results

	Control group (number of persons)					Total score	Experimental group (number of people)					Total score
	Score						Score					
	1	2	3	4	5		1	2	3	4	5	
Sound recognition	0	0	8	12	5	97	0	0	5	11	9	104
Accurate rhythm	0	2	11	9	3	88	0	0	7	15	3	96
Finger technique	0	1	8	11	4	90	0	1	5	11	8	101
Music understanding ability	0	2	9	9	5	83	0	0	6	15	4	98
Musical expression	0	3	15	5	2	81	0	1	9	8	7	96
Ability to compose melodies	2	1	8	7	7	91	0	0	5	13	7	102

According to the above table, the individual total scores are further presented in graphs, and the results of the comparison of the total scores of each individual item between the two groups are shown in Figure 6. It is obvious from Figure 6 that all six ability scores of the experimental group are significantly higher than those of the control group, among which the difference between the two scores of musical comprehension and musical expression is the most obvious, and the single total score of the experimental group is 15 points higher than that of the control group.

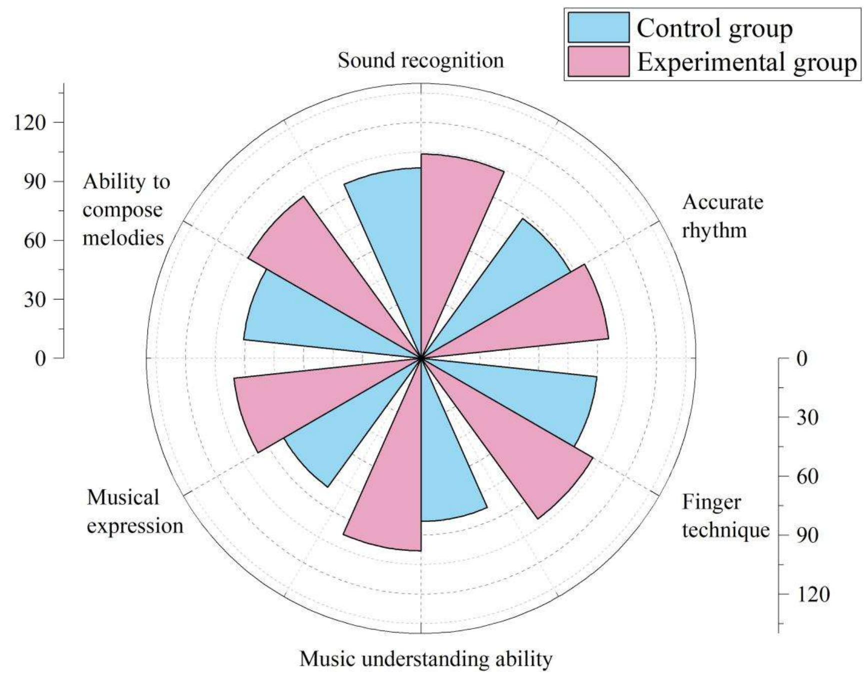

Fig. 6. Comparison results of total scores of each single item

4.2.3. Statistics and analysis of questionnaire data. In order to understand the cognitive changes acquired by the experimental group of students themselves, the students in the experimental group were interviewed by designing a questionnaire. The survey was divided into two times, the first questionnaire belongs to the pre-experiment survey (pre-test), which represents the students' situation before this piano study, and the second questionnaire belongs to the post-experiment survey (post-test), which embodies the changes that happened to the students after experiencing the piano study. The questionnaire still investigates the six aspects of sound recognition, rhythm, technique, music comprehension ability, musical expression, and melody creation ability, and asks students whether they have reached a satisfactory level in the six abilities, and the students answer the questions with "yes", "no" and "I don't know". Through the data statistics of the two questionnaire surveys, the statistical results of the number of students answering the questions are shown in Table 5.

Table 5. Statistics of the number of students answering questions

		Yes	No	Not clear
Sound recognition	Pre-test	4	16	5
	Post test	18	5	2
Accurate rhythm	Pre-test	2	18	5
	Post test	20	5	0
Finger technique	Pre-test	9	15	1
	Post test	22	3	0
Music understanding ability	Pre-test	10	13	2
	Post test	21	3	1
Musical expression	Pre-test	2	16	7
	Post test	19	5	1
Ability to compose melodies	Pre-test	3	19	3
	Post test	21	4	0

According to Table 5, the change in the number of people who answered "yes" in the questionnaire was extracted as the key information to be visualized, and the results of the comparison between the pre-test and post-test are shown in Figure 7. From the data of the questionnaire, it can be learned that through this experiment, the degree of students' positive response to each ability has increased to a certain extent, among which the most obvious change in the perception of melodic composition ability is that 21 students believe that their melodic composition ability has reached a satisfactory level after practicing with the accompaniment generated by the improved version of the MuseFlow model, which is an increase of 18 students compared with that of the pre-test. By analyzing the data from the students' pre-test and post-test questionnaires, it is possible to understand the effect of the experiment from the students' point of view, to grasp the students' evaluation of the changes in their piano learning ability during the learning process, and to understand the students' inner changes from the teachers' point of view, and to grasp the teaching effectiveness of the experiment.

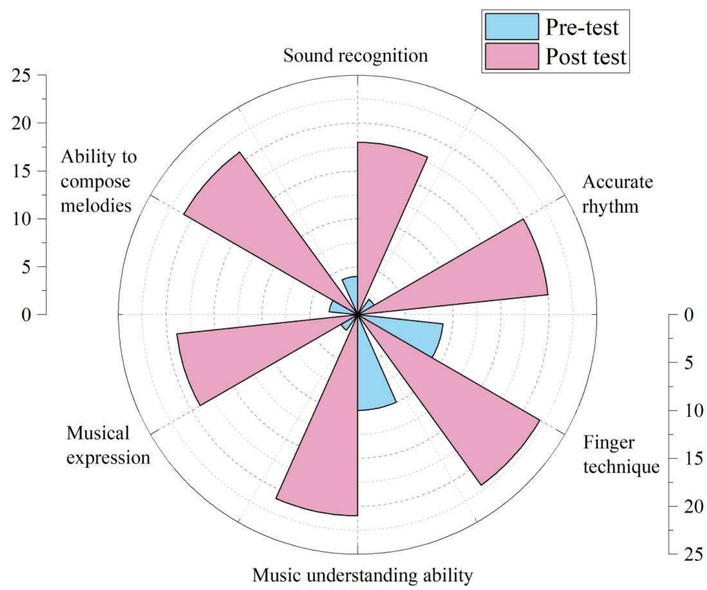

Fig. 7. Comparison of pretest and posttest results

5. Conclusion

In this paper, we design a MuseFlow model for accompaniment generation in combination with AI computation, and evaluate its generation effect with practical applications.

Comparing the improved MuseFlow model with five baseline models, the quantitative evaluation results show that the pitch distance between six pairs of tracks of the improved MuseFlow model is better than the MMM, which is the best performer among the baseline models, with an average pitch distance of 0.92, which is smaller than the MMM by 0.15, which indicates that the harmonic relationship between the tracks of the music generated by the improved MuseFlow model is stronger. The manual evaluation results show that the improved MuseFlow model is better than the baseline model in both melody and harmony, with an overall rating of 4.18, which is 0.37, 0.29, 0.91, 1.13, and 0.13 higher than the five baseline models, respectively.

The results of the students' ability evaluation showed that all six ability scores of the experimental group were significantly higher than those of the control group, among which the difference between the scores of music comprehension and music expression was the most obvious, and the total score of a single item of the experimental group was 15 points higher than that of the control group. The questionnaire data showed that the degree of students' positive response to each ability increased to a certain extent, among which the cognitive change in melodic composition ability was the most obvious, and there were 21 students who thought that the melodic composition ability reached a satisfactory level after the experiment, which proved that the model proposed in this paper was effective in practical application.

References

[1] Li, Y. (2022, January). Teaching Reform of Piano Accompaniment Art Direction Based on Data Mining Technology. In International Conference on Innovative Computing (pp. 348-356). Singapore: Springer Nature Singapore.

[2] Wang, C. S. J. (2017). Performance Practice and Overview of Selected Piano Works of Barbara Kolb. West Virginia University.

[3] Wu, Y., & Liu, C. (2023, June). Design of Piano Automatic Accompaniment System Based on Markov Model. In International Conference on Computational Finance and Business Analytics (pp. 3-10). Cham: Springer Nature Switzerland.

[4] Yang, X., & Pattananon, N. (2023). The Piano Accompaniment for Zuoquan Folk Songs in Vocal Music Teaching. Journal of Modern Learning Development, 8(11), 390-397.

[5] Zhang, X., & Liu, C. (2023, June). Design of Piano Automatic Accompaniment System Based on Artificial Intelligence Algorithm. In International Conference on Computational Finance and Business Analytics (pp. 249-258). Cham: Springer Nature Switzerland.

[6] Fraser, A. (2023). The craft of piano playing: A new approach to piano technique. Scarecrow Press.

[7] Takamori, H., Sato, H., Nakatsuka, T., & Morishima, S. (2017, July). Automatic arranging musical score for piano using important musical elements. In Proceedings of the 14th Sound and Music Computing Conference, Aalto, Finland (pp. 35-41).

[8] Xueke, Z. (2022). Analysis of the piano accompaniment characteristics of the national opera. Frontiers in Art Research, 4(2).

[9] Wang, H., Zhang, X., & Iida, F. (2024). Human-robot cooperative piano playing with learning-based real-time music accompaniment. IEEE Transactions on Robotics.

[10] Muriago, A. (2017). Re-envisioning the piano recital: an audiovisual alternative. University of Hartford.

[11] Chen, H. (2021, September). Application of Piano Automatic Accompaniment System Based on Artificial Intelligence in Piano Enlightenment Education. In 2021 4th International Conference on Information Systems and Computer Aided Education (pp. 1351-1355).

[12] Li, Y. (2020). Application of computer-based auto accompaniment in music education. International Journal of Emerging Technologies in Learning (iJET), 15(6), 140-151.

[13] Liu, H. (2022, November). Improvisational Dance Piano Accompaniment System Based on BP Neural Network. In 2022 International Conference on Computers and Artificial Intelligence Technologies (CAIT) (pp. 21-25). IEEE.

[14] Tan, X., & Li, X. (2021, October). A tutorial on AI music composition. In Proceedings of the 29th ACM international conference on multimedia (pp. 5678-5680).

[15] Rafii, Z., Liutkus, A., Stöter, F. R., Mimilakis, S. I., FitzGerald, D., & Pardo, B. (2018). An overview of lead and accompaniment separation in music. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 26(8), 1307-1335.

[16] Zhifei, S. Z. S., & Pattananon, N. (2022). Development and Benefits of Teaching Piano Accompaniment Courses in Higher Music Education, China. Journal of Modern Learning Development, 7(7), 437-444.