

# The investigation of artificial intelligence-based applications in music education

**Ge Wang**

Department of Software Engineering, Dalian Maritime University, Dalian, 116026, China

weiyanyxv@mail.ustc.edu.cn

**Abstract.** The field of music education has garnered significant attention due to the burgeoning advancements in artificial intelligence technology. Currently, there are many works on applying Artificial Intelligence (AI) technology to the field of music education, and this review aims to comprehensively summarize the application of AI methods in multiple subfields of music education. In this paper, according to the focus of AI technology, the areas of music education to which it has been applied are divided into four modules, namely personalized learning, automatic assessment, composition assistance and interactive teaching. Firstly, representative AI methods applied to each of these four subfields are introduced, including but not limited to: the principle of the AI method, the improvement brought by the application. The text then describes the current status of the application of these methods in their domains, discusses the limitations of the application, and makes brief conjectures on how to break through these limitations. The research in this review paper is important for the development of AI technology applied to the field of music education. At the same time, this review also provides directions and suggestions for future research to promote the further development of AI in the field of music education.

**Keywords:** music education, artificial intelligence, machine learning, deep learning.

## 1. Introduction

Music an integral component of human culture and society with a rich history spanning millennia, serves as a significant medium for expressing emotions, fostering social cohesion, and enriching human life. In contemporary times, the advent of computer technology has a particularly significant impact on music. Computer technology has opened up new possibilities for music creation and performance, e.g., digital audio processing [1], electronic musical instruments [2], music synthesis and arranging [3], and music analysis and categorization [4]. Artificial Intelligence (AI) technology, as a rising star in the field of computer technology, also brings more possibilities to the field of music, especially in music education. The combination of artificial intelligence and music education can be regarded as a breakthrough innovation. First of all, AI can provide students with a personalized learning experience based on their learning speed, ability and interest. Additionally, AI can automatically adjust the teaching content and methods according to students' learning progress and feedback. Moreover, AI technology can help music education collaborate across boundaries with other fields to promote creative thinking and cross-disciplinary learning.

AI, as a burgeoning scientific discipline, focuses on the study, development, and application of theories, methods, techniques, and systems that aim to model and expand human intelligence. It is a branch of computer science that involves a range of disciplines. It is also a challenging discipline. And machine learning, as the core of it, can be categorized according to schools of thought, which include symbolism, behaviorism, and connectionism. Regarding connectionism, it argues that the goal of machine learning is to construct a neural network-based model that is capable of learning non-linear mapping relationships between inputs and outputs. Symbolism emphasizes the symbolic processing power of human intelligence, arguing that high levels of intelligent behavior can be achieved through symbolic representation and reasoning. Common connectionist approaches include deep neural networks [5], convolutional neural networks [6], and recurrent neural networks [7].

The combination of artificial intelligence and music education is the application of artificial intelligence methods to the field of music education, which can help students better learn and master music theory and techniques. It can be divided into four main sections. Firstly, personalized learning uses machine learning and data mining techniques to provide personalized learning plans and suggestions for each student based on their learning performance and feedback. For example, based on students' music preferences and abilities, appropriate music pieces and practice methods are recommended for them to improve their learning. Additionally automated assessment is a method that can assess students' rhythm, intonation, and performance skills by analyzing the audio and video data from the video of their performances, thereby helping them to identify and correct problems in a timely manner. Moreover, composition assistance uses machine learning and natural language processing technologies to automatically generate and recommend music clips and chord progressions based on the music material provided by students, helping them to better understand and master the basics and techniques of music composition. Finally interactive teaching utilizing virtual reality technology and robotics can provide students with a more realistic and vivid music teaching experience. For example, utilizing virtual reality technology, students can participate in virtual music sessions and performance activities to improve their performance skills and musical communication abilities.

Since the combination of Artificial Intelligence and Music Education as a novel interdisciplinary subject is very important and promising, it is necessary to write a review that summarizes recent results in this field. The rest of the paper is organized as follows: section 2 will present artificial intelligence methods applied in the field of personalized learning, automatic assessment, compositional assistance and interactive teaching. The application of artificial intelligence in the field of music education will be discussed in Section 3. Finally, section 4 provides the conclusion.

## 2. Method

### 2.1. *Personalized learning*

Personalized learning is about providing students with a personalized music learning experience through AI technology, tailoring course content, difficulty and teaching methods to students' abilities and needs in order to improve their learning effectiveness and interest [8].

*2.1.1. Violin music transcriber for personalized learning.* Jie et al. introduced an enhanced version of the violin music transcriber aimed at facilitating personalized learning [1]. Their proposed approach focuses on identifying duo-pitch occurrences, where two strings are bowed simultaneously, in real-world violin audio recordings captured in a home environment. To accomplish this, the researchers utilized a semitone band spectrogram, which is a spectral depiction of the signal that directly relates to music. By leveraging the inherent constraints of violin sound, this method improves both the accuracy and speed of transcription compared to existing techniques. The investigators performed comprehensive evaluations, which encompassed the analysis of individual pitch notes and duo-phonetic pitch samples within the playable range of the violin (G3-B6), as well as the assessment of music excerpts.

## 2.2. Automatic assessment

Automated Assessment is an automated assessment of music performance, including sound quality, rhythm, and technique, through AI technology to improve the accuracy and efficiency of the assessment.

*2.2.1. Violin tone quality assessment.* Chen et al. proposed a machine learning-based system for assessing violin tone quality. The system uses a recording of a violin performance as input and evaluates the sound quality of the performance by analyzing the features and patterns of the audio signal, which can help teachers to more accurately evaluate the performance level and progress of their students.

The automatic assessment of tone quality in violin music performance system is based on machine learning algorithms that are trained to recognize patterns in the audio signal of violin performances. The system analyzes various features of the audio signal, such as pitch, volume, and vibrato, to determine the quality of the tone produced by the violinist. And the system uses a large datasets of violin performances to train the machine learning algorithms, which allows it to accurately recognize and evaluate different aspects of violin tone quality. Once the system has analyzed a performance, it provides a detailed report on the various aspects of the performance, including the overall tone quality, the pitch accuracy, the rhythm, and the articulation. Furthermore, this system is especially useful for violin teachers, who can use it to provide more accurate feedback to their students. The system gives teachers a more objective measure of the tone quality of their students' performances, which can help them to identify areas where the student needs to improve. Additionally, the system can be used to track the progress of a student over time, allowing the teacher to see how the student's tone quality is improving as they practice and develop their skills.

*2.2.2. Video soundtrack evaluation with machine learning.* Touros et al. introduced a machine learning approach for the evaluation of video audio tracks. The method uses audio and video data to classify and evaluate audio tracks, which can help video and television producers to better evaluate the quality and adaptability of audio tracks and improve the efficiency and quality of audio track production.

## 2.3. Composition assistance

Composition assistance aims to help music creators with composing and arranging through AI technology, including music generation, harmonic analysis, score generation, and automatic arranging.

*2.3.1. Generative adversarial network for music Composition.* Dong et al. proposes a multi-track music generation model based on Generative Adversarial Network (GAN). The model is capable of generating both the main melody and accompaniment simultaneously and can be personalized according to the user's needs and guidance, thereby assisting music creators in realizing the creation and arrangement of music.

Specifically, the model uses the GAN framework to generate multi-track music by learning from a large datasets of music pieces. Unlike traditional music generation models, this model can generate multiple tracks of music simultaneously and can be personalized according to the user's needs and guidance. For example, users can influence the model's generation results by providing specific melody or harmony segments, thereby achieving more personalized music creation and arrangement.

This paper provides a new music generation approach for music creators, allowing them to generate personalized music based on their needs and guidance. Additionally, the GAN framework used in the model provides new ideas and methods for research in the fields of music and artificial intelligence. The research results from this paper have broad application prospects in the music industry and music education, providing innovative solutions for these fields.

*2.3.2. Deep neural networks for music composition.* Mimitakis introduced a method for audio separation and mixing based on deep neural networks. The method can separate the audio tracks of different instruments and mix them according to the user's needs, providing music creators with more options for sound and creative space. This paper provides a new method for music creation that uses

deep neural networks for audio separation and mixing, allowing for more diverse sound and musical effects. The method can also assist music creators in adjusting and processing music in post-production, achieving better musical results. Additionally, the visual interface of the method makes it easier for music creators to interactively adjust the audio separation and mixing results.

#### *2.4. Interactive teaching*

Interactive teaching is about providing a more interactive and personalized teaching experience through AI technology, including intelligent instructional design, virtual teaching assistants, and adaptive teaching.

*2.4.1. AI assisted music teaching design.* Dai et al. proposed a design method based on artificial intelligence technology to assist music teaching. The method combines the characteristics of music teaching and the needs of students and uses artificial intelligence technology to provide personalized teaching content and methods, which can help teachers to better design and implement teaching.

*2.4.2. VR and AI technology.* Yan et al. developed an interactive teaching mode utilizing VR and AI technology and conducted experimental analysis to validate the integration of VR and AI in music education. The experimental results showcased impressive recognition rates for the bass signal, alto signal, and treble signal, achieving 100%, 90%, and 100% accuracy respectively.

### **3. Applications and discussion**

#### *3.1. Music education and performance evaluation*

The techniques proposed by Chen et al. are primarily used in the fields of music education and performance evaluation [9]. This approach, for instance, holds potential for the facilitation of violin teachers in evaluating tone quality among a large number of students, or help judges at music competitions make fair and consistent scores. However, there are some limitations to this technique. First, the judgment of tone quality usually depends on the subjective preferences of individuals, which engendering a possible variance between the scoring outcomes of the automated evaluation mechanism and those rendered by human assessors. Additionally, training efficient tone quality evaluation models requires a large amount of labeled data, which may be challenging to obtain in practical applications. Furthermore, such systems may not be able to take into account some non-audio aspects of musical performances, such as the body language and emotional expression of the performer.

#### *3.2. Music composition*

MuseGAN is a music generation model based on GANs that generates multi-track symbolic music. Dong et al. applied this model to the field of music composition and accompaniment generation [10]. For example, it can help music composers generate new musical material or generate suitable accompaniments for an existing music clip. However, MuseGAN also suffers from some limitations. First, the generated music may lack long-term structural coherence, since GANs typically struggle to maintain consistency when generating longer sequences. In addition, GANs require large amounts of training data and computational resources, which may limit their application in low-resource environments. Furthermore, MuseGANs may have difficulty capturing some of the subtle expressive and emotional details in music. In forthcoming endeavors, it may be possible to solve the aforementioned issues by improving the structure and training algorithms of GANs.

Mimilakis proposed a technique on the use of deep neural networks for the separation and mixing of jazz recordings [11]. The application of this technique is mainly in the field of music composition. For example, it can be used to extract specific instrument sounds from a recording, or to automatically adjust the balance of individual instrument sounds during mixing. However, there are some limitations of this technique, for example, the effect of music separation may be affected by noise and reverberation, deep neural networks require a large amount of training data and computational resources as well as the fact

that such a system may not be able to fully understand the artistic expression of music. In the future, this technique can be further improved to enhance the effectiveness of music separation and mixing. Moreover, its application can be further extended by combining it with other music analysis and processing techniques.

#### 4. Conclusion

This article focuses on the application of AI technology in the field of music education. The article introduces the contribution of AI in music education from four aspects, namely Personalized learning, Automatic assessment, Composition assistance and Interactive teaching. These techniques employ advanced AI technological tools such as machine learning, deep neural networks, generative adversarial networks, and AI-based virtual reality. In general, although the application of AI technology to actual music education still faces numerous difficulties to be solved. This paper covers a relatively limited area of music education, which can be followed by a wider range of areas. The content of each field can also be enriched by introducing more AI technology tools. This paper also does not compare the advantages and disadvantages of different methods in detail, which can be added in the future studies.

#### References

- [1] Roads C 1996 The computer music tutorial (MIT Press)
- [2] Miranda E R 2012 Musical creativity and the affordances of technology Contemporary (Music Review 31(1)) p 17-35
- [3] Rowe R 2001 Machine Musicianship (MIT Press)
- [4] Li T Ogiwara M and Li Q 2008 A survey of content-based music retrieval systems (Journal of Intelligent Information Systems, 30(1)) p 1-40
- [5] Hinton G E and Salakhutdinov R R 2006 Reducing the dimensionality of data with neural networks (Science, 313(5786)) p 504-507.
- [6] LeCun Y Bottou L Bengio Y and Haffner P 1998 Gradient-based learning applied to document recognition (Proceedings of the IEEE, 86(11)) p 2278-2324.
- [7] Hochreiter S Bengio Y Frasconi P and Schmidhuber J 2001 Gradient flow in recurrent nets: The difficulty of learning long-term dependencies. In J. F. Kolen & S. C. Kremer (Eds.), A Field Guide to Dynamical Recurrent Neural Networks pp. 237-244 IEEE Press
- [8] Jie W Boo J Wang Y and Loscos A A 2006 Violin Music Transcriber for Personalized Learning. (In 2006 IEEE International Conference on Multimedia and Expo ) pp 2081-2084 Toronto ON Canada: IEEE. doi 10.1109/ICME.2006.262644
- [9] Chen K Y and Lin Y C 2019 Automatic assessment of tone quality in violin music performance IEEE Transactions on Audio, Speech, and Language Processing, 25(5), p1050-1061.
- [10] Touros G and Giannakopoulos T 2023 Video Soundtrack Evaluation with Machine Learning: Data Availability, Feature Extraction, and Classification. In A. Biswas, E. Wennekes, A. Wiczorkowska, & R. H. Laskar (Eds.), Advances in Speech and Music Technology (pp. 137-157). Signals and Communication Technology. Springer, Cham
- [11] Dong H.-W Hsiao W.-Y Yang L.-C and Yang Y.-H 2018 MuseGAN: Multi-track Sequential Generative Adversarial Networks for Symbolic Music Generation and Accompaniment. Proceedings of the AAAI Conference on Artificial Intelligence, 32(1). <https://doi.org/10.1609/aaai.v32i1.11312>