Proceedings of the 6th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/64/20241425

Composing music under certain conditions based on neural network

Zhen Chen

School of Art and Science, University of Miami, Miami FL 33124, USA

zxc510@miami.edu

Abstract. As a matter of fact, composing music based on machine leaning has been a common trend contemporarily. This study delves into the emerging field of AI music composition, tracing its development from early stage to contemporary. Moreover, significant models such as recurrent neural networks (RNNs), long short-term memory networks (LSTMs), and MusicVAE are examined in depth. This study also explores a wide range of applications of AI music, including its role in areas such as education and therapy. Despite these advances, the paper still highlights the limitations of AI music, particularly its struggles to match human creativity and emotional depth. In addition, the paper discusses the ethical and social issues surrounding AI music, such as copyright disputes and employment implications. Looking forward, the future of AI music is the development of intelligent creative tools, personalized music recommendations, and therapeutic interventions. This research not only provides a comprehensive overview of AI music but also underscores its potential to transform how one creates, consume, and interact with music in the digital age.

Keywords: AI music, computer music, music generation, AI-based music composition.

1. Introduction

Suno, which was launched on December 20, 2023, is an AI-powered music composition platform that uses generative algorithms to create songs blending vocals and instruments [1]. After discovering the new AI music generator, one user from the reddit community reviewed that it was truly shocked by the initial outcome [2]. With the development of technology, artificial intelligence (AI) has become increasingly mature and widely applied in various fields, including music composition, which has attracted widespread attention and interest. Unlike traditional music composition, which relies on the creative skills of human beings, AI music automates music generation through learning and imitation of human music composition. Therefore, its history is relatively short. Dillon divided the timeline of AI music development into three major phases: the early stage (1950 - 1970), the transitional period (1980 - 1990) and the current period (2000 - present) [3]. During the early stage, Lejaren Hiller's Illiac Suite marked a pivotal moment in the development of AI music [4]. Moreover, in the transitional period, David Cope from the University of California devised the Experiments in Musical Intelligence (EMI) system [4]. This system employed generative models to analyze pre-existing music and generate novel compositions, representing a noteworthy advancement in AI music research [4]. In 2002, Francois Pachet developed the "Continuator" algorithm at the Sony Computer Science Laboratory, enabling interactive music composition with live musicians by continuing music from where they left off [4].

DOI: 10.54254/2755-2721/64/20241425

In recent years, music composition systems based on AI or deep learning have achieved more remarkable advancements. The recently published album “Hello World” had its formal debut in early January 2018. This is Flow Machine’s first album to be released commercially, which sets it apart from their earlier work and heralds a new era in pop music. Introduced in the early part of 2016, Flow Machines are sophisticated creative tools designed to boost user creativity [5]. It uses a unique Lead Sheet Database (LSDB) instead of a conventional dataset, differing from other tools in their training approach [6]. Moreover, Amper Music is a cloud-based platform designed to simplify the process of creating soundtracks for movies and video games, as it generates AI algorithms that assist users in composing music of various genres [7]. It also uses its own dataset library, which currently has hundreds of specially designed instruments and over a million individual samples. Every aspect of their system is constructed from scratch by the developers themselves, who meticulously sculpt the sound to achieve an organic quality distinct from synthetic alternatives [8]. The developers build every component of their system from the ground up, carefully crafting the audio to provide a natural sound that sets it apart from artificial substitutes.

The purpose of this study is to examine the state and trends of artificial intelligence (AI) in music composition and to talk about ways to enhance the functionality and efficiency of AI-based systems for composition. The proposed research framework includes defining basic concepts and classifications of computer music composition; exploring prevalent models commonly used in computer music composition; analyzing the practical applications of computer music composition as well as discussing the challenges and limitations existing in current AI-based music composition systems, and proposing solutions and improvements.

2. Basic Concepts

Before doing research on the study of AI music composition, it is important to understand the basic concepts and classification of computer music composition. Computer music composition refers to the process of using computer technology to assist or entirely replace humans in the process of music composition [9]. It is primarily applied in four areas: music generation, music synthesis, music analysis, and music performance. In the field of music generation, publications such as Google, OpenAI, and Sony indicate a global interest in this field [10]. Computer music generation uses computational methods for composing music [11]. It can occur through prediction (such as forecasting the pitch of the next note in a melody) or classification (like identifying the chord associated with a melody), using the distribution and correlations acquired by the deep model [12]. Consider the Mozart sonata generation C-RBM. In the C-RBM design by Lattner et al., a limited Boltzmann machine (RBM) is utilized to learn the local structure, which is interpreted as the musical texture, from a set of musical compositions, which in practice are mostly Mozart sonatas [13]. Moreover, the concept of "music synthesis" includes the methods for generating the auditory elements of music, including musical instruments and vocalizations. Furthermore, the term "music synthesis" refers to the techniques used to create the aural components of music, such as vocalizations and musical instruments. Many digital musical instruments, such as electric pianos and synthesizers, presently use this technology to create a variety of soundscapes, from electronic to authentic [14]. Spectrum Modeling Synthesis (SMS) is used to track the harmonic series in the sound and assumes the remaining part as noise [15]. By decomposing harmonics and noise, SMS enables detailed reconstruction and manipulation of instrument sounds, including non-harmonic instruments [16]. In the domain of music analysis, the concept of music and artificial intelligence revolves around using platforms as the core of music theory and employing AI analysis for studying the transmission of music media, including music recording and audio recognition. Music education based on developmental principles can serve as a crucial theoretical foundation for fostering creativity and facilitating growth in information analysis [17]. This teaching model has the potential to significantly enhance teaching efficiency and performance, freeing both teachers and students from burdensome educational responsibilities [18]. Furthermore, AI is increasingly being used to perform music. It enables a machine musician to perform alongside human musicians in a musical ensemble [14]. The computer system SAXEX, based on Case-Based Reasoning (CBR) technology, is capable of performing

DOI: 10.54254/2755-2721/64/20241425

expressive monophonic music, like human performance. CBR has proven to be a highly potent technique, enabling the direct utilization of implicit information extracted from recordings of human performers [15].

3. Models and applications

The types and categories of AI music models are diverse, covering various application scenarios and technical approaches, with sequence models being particularly suited for handling music sequence data, such as note sequences or time series. Among them, recurrent neural networks (RNNs) and long short-term memory networks (LSTMs) are commonly used sequence models. An RNN can be conceptualized as a conventional neural network enhanced by incorporating the sequence's items as inputs and dedicated outputs for encoding and retaining information about the preceding items processed by the network [19]. Fig. 1 displays its logic diagram. Sun's DeepHear architecture is intended to produce ragtime jazz tunes [20]. It is made up of hierarchically nested 4-layer stacked autoencoders with 16 hidden units, which are progressively fewer in number. After being trained on a corpus of 600 ragtime music measures by Scott Joplin, divided into 4-measure segments, the model produces output by first feeding random data into each of the 16 bottleneck hidden-layer units. The output is then produced in the same 4-measure format as the training samples after being sent through the series of decoders [21].

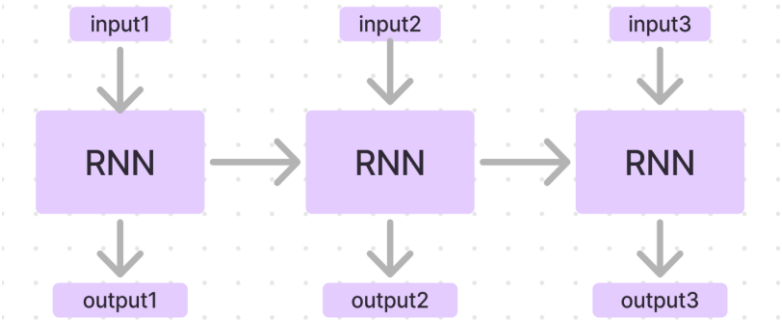

Figure 1. RNN loop [19].

However, the basic RNN has a drawback “vanishing gradient problem” [22]. This issue arises when earlier inputs diminish in influence with the introduction of subsequent inputs to the network. LSTMs were specifically devised to deal with this problem by ensuring the significance of information to retain and what to discard. Initially introduced in 1997 by Sepp Hochreiter and Jürgen Schmidhuber, LSTMs have undergone modifications over time to enhance their efficacy for specific tasks. In 2015, Google researchers tried to further refine the established LSTM structure, exploring potential enhancements through empirical analysis. Their research, documented in "An Empirical Exploration of Recurrent Network Architectures," concluded that, for most applications, the existing LSTM structure remains unmatched following a straightforward architecture search [23]. There is another model. Roberts et al.'s MusicVAE [24] creates musical scores via a hierarchical recurrent variational autoencoder. Its main purpose is to encode melodies into a condensed representation and subsequently reconstruct them. The decoder is a hierarchical two-level RNN that consists of a high-level RNN called the Conductor that generates a sequence of embeddings and a lower-level RNN that uses each embedding as an initial state and additional input concatenated to previously generated tokens to create each subsequent subsequence [25]. It adheres to the principles of a variational autoencoder encapsulating recurrent networks (RNNs).

Computer music has diverse applications, such as education as shown in Fig. 2. Firstly, AI can automatically evaluate students' musical performance, including playing skills, rhythm accuracy, and musical understanding. It provides instant feedback and suggestions to help students improve their playing skills and deepen their understanding of music. For example, during music lessons, AI systems can select the most appropriate material for teaching based on their learning needs and personal capabilities [26]. In addition, AI has the capability to create diverse educational tools like music theory

DOI: 10.54254/2755-2721/64/20241425

learning apps, rhythm training programs, and audiovisual materials, aiding students to better understand and master music knowledge and skills. For example, Google's AI composing tool Magenta Studio can automatically include and grow new melodies and drum beats based on user input, creating a fully realized musical composition. Similar AI music composition software is getting more complex and intuitive, with the potential to be an essential tool for students to write creative music in the classroom of the future [26]. Furthermore, AI music applications can increase interest through AI education. Presently, music education in Chinese schools typically covers music appreciation, theoretical studies, and practical skills. However, despite students' fondness for music, they may not necessarily enjoy music classes. In response to this frequent occurrence, music educators are continuously innovating new instructional approaches. In order to maximize student participation, this entails utilizing contemporary network technologies and multimedia instruction more extensively, as well as investigating instructional strategies like Kodály, Orff, and Dalcroze systems [27]. Apart from its successful applications in education, AI music also shows significant potential in therapy and mental well-being which is shown in Fig. 3.

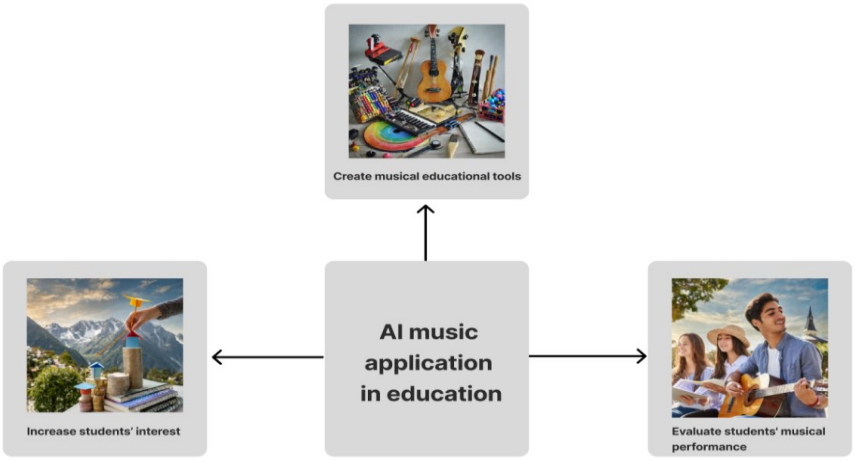

Figure 2. AI music application in education [26].

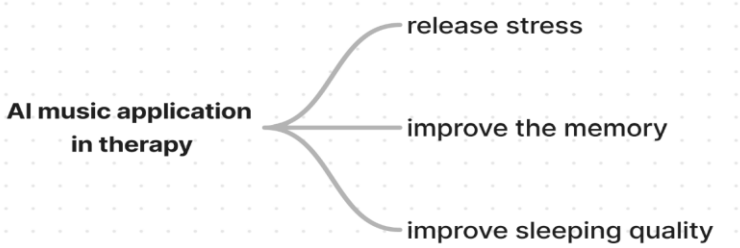

Figure 3. AI music application in therapy flow [27].

AI-generated music, customized to an individual's emotional state, helps in releasing stress, anxiety, and depression. Listening to specific music genres allows people to manage their emotions and improve their mental health. Music therapists usually assist clients in understanding their emotions. The initial stage focuses on emotional awareness, where therapist's guide clients to recognize, experience, and express their negative thoughts and sentiments. Participants do point out that various client personalities could call for marginally varied therapy modalities. For example, participants suggest employing improvisational tactics to help an introverted client who is having low arousal emotions or who is unable to communicate their emotions and not aware of them. Considering the clients' prior musical background, suggestions include using music genre generators and melody/harmony generators to simplify the composition process and help clients to create accompanying music [28]. For those with cognitive

DOI: 10.54254/2755-2721/64/20241425

impairments or in need of rehabilitation, AI music can act as an adjunctive therapy. For example, music therapy can aid in memory retrieval, enhance concentration, and help patients in reconnecting with their surroundings. A clinical research survey by Cohen-Mansfield, Marx, and Werner found that upon entering institutions, most elderly individuals with dementia and depression did not receive any activity arrangements, with over half of institutional caregivers primarily focusing on providing physical care, neglecting the spiritual needs of elderly individuals with dementia. Research reports suggest that the more structured activities elderly individuals with dementia participate in, such as the AI bedtime music therapy courses in this study, not only significantly reduce problematic behaviors but also improve daily living functions, sleep quality, and reduce anxiety and depression. Furthermore, these activities promote self-expression, social interaction, and positive behavior [29]. Research on the benefits of music therapy for improving social skills in kids—especially those with minor intellectual disabilities—was presented by Duffy and Fuller. Pre- and post-intervention scores on five target competencies were compared using a short social skills exam created especially for the analysis [29]. AI-generated music provides a relaxing method for stress-related conditions or sleep disturbances. By listening to specific music genres, people can relax both mentally and physically, easing anxiety and stress while improving sleep quality. AI sound therapy before bedtime can help patients calm their emotions and relieve stress and enhance the subconscious' ability to produce α(alpha) brain waves + θ(theta) brain waves; decrease the production of negatively generated β(beta) brain waves energy [30].

4. Limitations and prospects

Although AI music has been applied in many aspects, it still has many limitations. While AI can generate musical compositions, they may lack human creativity and emotional expression. Music produced by AI may lack depth, personalization, or emotional resonance, making it challenging to compare with human creations. AI music can never fully satisfy the craving for creativity and embodiment. Although AI may be capable of producing acceptable synthwave music, it will always struggle with creating songs because songs are valued as a form of personal expression. But there's more to AI music than meets the eye: the storyline has been misleading us. Contrary to being solely the result of algorithms, AI-generated music is a hybrid creativity that relies on the collaboration between humans and computers – a form of cyborg creativity that merges human creativity with the capabilities of machine learning, challenging the traditional dichotomy between nature and technology [31]. The development of AI music may raise ethical and social issues that require thorough consideration and resolution. 'Automated lyrical song-writing application' (ALYSIA) is an acronym for the program that was released in 2019 by WaveAI, a Silicon Valley business. Those songs created by ALYSIA illustrate the concept referred to by roboticist Masahiro Mori from Japan as the 'uncanny valley' [32], the discomfort one tends to experience when faced with something not quite human enough.

The future of AI music is very promising, with numerous potential directions for development and applications. From a market perspective, AI music software and tools provide music producers with smarter and more efficient creative instruments, helping them quickly generate music materials, adjust, and improve their works. Jukedeck and Amper Music specialize in producing royalty-free soundtracks tailored for content creators like videographers, offering features for adjusting the music's duration, style, and climax to suit the video or presentation being produced [33]. Besides, AI technology has the capability to analyze users' music preferences and behaviors, offer personalized music recommendations and services, and present music content that matches their tastes and moods. Recommendation engines use machine learning techniques, similar to those found in music services like Apple Music and Spotify to analyze the listening patterns and preferences of users. These algorithms identify patterns in users' listening behaviors and apply this data to suggest new tracks, artists, and soundtracks likely to captivate customers' interest [34].

5. Conclusion

In conclusion, AI music composition is at a crossroads where it must strike a balance between incredible achievements and constraints. While systems like Suno showcase AI's potential in generating melodies

DOI: 10.54254/2755-2721/64/20241425

and compositions, they often fail to capture the depth and emotional resonance of human-created music. Despite this, AI's applications in education, therapy, and content creation demonstrate its benefits, from helping students learn to enhancing mental health. As AI-generated music continues to evolve, ethical and social considerations must be carefully navigated. Looking ahead, AI music has promise in providing more intelligent creative tools for musicians, offering personalized music experiences for listeners, and serving as a treatment for various health conditions. The future of AI music is not just about algorithms but also about the collaborative creativity between humans and machines, challenging the traditional boundaries of music creation.

References

[1] Cowen T 2024 My review of Suno, AI-generated music Retrieved from https://policycommonsnet/artifacts/11765165/my-review-of-suno-ai-generated-music/12656560/ on 24 Apr 2024 CID: 2050012592/31zcxqz
[2] Reddit-Dive into Anything Redditcom, 2024 Retrieved from: wwwredditcom/r/SunoAI/comments/1bwmez4/my_thoughts_about_suno_ai_and_ai_music/ Accessed 8 Apr 2024
[3] Ranwala D 2024 The Evolution of Music and AI Technology Watt AI, 23 July 2020, Retrieved from: watt-aigithubio/blog/music_ai_evolution Accessed 8 Apr 2024
[4] Verma S 2021 International Journal of Applied Research vol 7(2) pp 272–75.
[5] Avdeeff M 2019 Arts (Basel) vol 8(4) pp 130–30.
[6] Francois P, Pierre R and Benoit C 2021 Assisted Music Creation with Flow Machines: Towards New Categories of New Springer International Publishing, Cham po 485–520.
[7] Arango B 2024 Top Best 10 AI Music Composers in 2024 Retrieved from: filmorawondersharecom/audio-editing/best-ai-music-composerhtml.
[8] Amper. Ai composition with real instruments URL, retrieved from: https://www ampermusiccom/music/
[9] Cook P 2017 A NIME Reader: Fifteen years of new interfaces for musical expression pp 1-13.
[10] Civit M, Civit-Masot J, Cuadrado F and Escalona M J 2022 Expert Systems with Applications vol 209 p 118190.
[11] Filippo C and Antonio R 2024 Frontiers in Artificial Intelligence vol 3 p 14.
[12] Jean-Pierre B and François P 2024 Neural Computing & Applications vol 32(4) pp 981–993.
[13] Lattner S, Grachten M and Widmer G 2018 Journal of Creative Music Systems vol 2 pp 1-31.
[14] Research and Development - AI Sound Synthesis Technology - Yamaha Corporation Yamahacom, 2019 Retrieved from: wwwyamahacom/en/tech-design/research/technologies/aisynth/.
[15] Lopez R and Josep L A 2002 AI magazine vol 23(3) pp 43-43.
[16] Dai S 2023 Towards Artificial Musicians: Modeling Style for Music Composition, Performance, and Synthesis via Machine Learning (Doctoral dissertation, Stanford University).
[17] Wei J, Karuppiah M and Prathik A 2022 Computers and Electrical Engineering vol 100 p 107851.
[18] Zulić H 2019 INSAM Journal of Contemporary Music, Art and Technology vol 2 pp 100-114.
[19] Conner M, Gral L, Adams K, Hunger D, Strelow R and Neuwirth A 2022 Music generation using an lstm. arXiv preprint arXiv:2203.12105.
[20] Sun F 2017 DeepHear—composing and harmonizing music with neural networks Retrieved from: https://fephsungithubio/2015/09/01/neural-music.html
[21] Briot J P and Pachet F 2020 Neural Comput & Applic vol 32 pp 981–993.
[22] Michael P 2018 Illustrated Guide to LSTM's and GRU's: A Step by Step Explanation Retrieved from: https://towardsdatasciencecom/illustrated-guide-to-lstms-and-gru-s-a-step-by-stepexplanation-44e9eb85bf21 (2018)
[23] Conner M, Gral L, Adams K, et al 2022 Music generation using an lstm arxiv preprint arxiv:2203.12105.
[24] Roberts A, Engel J, Raffel C, Hawthorne C and Eck D 2018 International conference on machine learning pp 4364-4373.
[25] Briot J P and Pachet F 2020 Neural Computing & Applications vol 32(4) pp 981-993.

DOI: 10.54254/2755-2721/64/20241425

[26] Qin J 2022 Mathematical Problems in Engineering vol 2022 p 2138059.
[27] Lin Y and Ding J 2020 IOP Conf Ser: Mater Sci Eng vol 750 p 012101.
[28] Sun J, Yang J, Zhou G, Jin Y and Gong J 2024 arXiv preprint arXiv:2402.14503.
[29] Duffy B and Fuller R 2000 Journal of Applied Research Intellect Disability vol 13(2) pp 77–89.
[30] Tsai Y C and David C H 2023 International Journal of Innovative Application on Social Science and Engineering Technology vol 41 pp 1-24.
[31] Cole R 2020 Popular Music vol 39(2) pp 332–338.
[32] Mori M 1970 IEEE Spectrum vol 6 pp 1-6.
[33] Casini L, Marfia G and Roccetti M 2018 IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC) pp 27-31.
[34] Olayeni S 2023 The impact of artificial intelligence (AI) in music business industry Doctoral thesis of Centria University of Applied Science.