Implementation Of Machine Learning and Artificial Intelligence for Music Composition

Xiaohan Yang

Nanjing Foreign Language School, Nanjing, 21000, China

Yang.bo4@zte.com.cn

Abstract. With the rapid developemnt of machine learning techiniques, it has been widely adopyed in various fields. This paper explores the models and applications of machine learning in music composition. It traces the evolution of computer music from its early days with electronic music pioneers to its contemporary implementations, showcasing the integration of cutting-edge technologies. Various models and techniques employed in music composition, including rule-based systems, deep learning, and stochastic composition methods, are thoroughly examined. Furthermore, significant challenges posed by copyright issues are addressed, along with potential solutions to overcome them. Looking ahead, one envisions a future where computer music achieves greater emotional depth, accessibility, and customization despite inherent complexities. Advances in copyright management and licensing systems are expected to diversify the pool of musical compositions available for training machine learning models, resulting in more emotionally resonant and diverse computer-generated music. This study sheds light on the potential of combining computer music with machine learning, providing valuable insights into the intricate world of music composition and technological innovation for future music creators and enthusiasts.

Keywords: machine learning, music composition, copyright management, licensing system.

1. Introduction

With the rapid development of the digital economy and the continuous innovation of new media technologies, the production and creation of music have gradually integrated with digital concepts, giving rise to the field of computer music. Computer music is a novel form of artistic creation that effectively combines modern computer technology with various electronic audio processing techniques to produce music. Through the use of computer music systems, individuals can engage in a series of operations including music notation, composition, experimentation, and digital recording with greater convenience. This has significantly transformed the conventional methods and inherent patterns of music creation [1]. Furthermore, computer music has lowered the costs of music creation, opened up new frontiers in the realm of music, and catalyzed the emergence of fresh musical and creative concepts.

With the emergence of computers in the late 1945s in the United States, pioneers of electronic music exploration such as Matthews and Hiller keenly sensed that this novel technological apparatus held the potential to produce enchanting and marvelous music. They believed it would exert a profound influence on future human musical endeavors. This anticipation became a reality in 1957 when Max Mathews, a scientist at Bell Labs in New Jersey, USA, designed the world's first computer-generated music software series known as Music [2], utilizing computer and related hardware components. Though the inaugural piece of computer-generated music, "The Sliver Scale," composed by Mathews in collaboration with software and hardware, lasted a mere seven seconds, its creation signaled two significant developments. Firstly, it marked the inception of computer music production systems. Secondly, it heralded the dawn of the era of computer-generated synthetic music. Nowadays, computer music can achieve automated composition and assistive composition. Automated composition refers to computers composing music through specific algorithms, while assistive composition involves simplifying the creative process using computer assistance. For instance, algorithmic composition, as demonstrated by Cope, involves using this method to emulate the compositional style of specific composers [3]. Moreover, computer music can aid in the storage, retrieval, and dissemination of music [4].

Machine learning plays a pivotal role in automating the music composition process within computer music. One prominent approach involves the creation of systems that autonomously generate high-quality music with minimal user intervention. Such systems leverage an input corpus of music and incorporate machine-learning techniques, including genetic algorithms and Markov chains, to craft music that sounds musically coherent. An exemplary illustration of this approach is the Automated Composer of Style-Sensitive Music II (ACSSM II) system [5].

ACSSM II exemplifies how machine learning can be harnessed to generate music sequences that fulfill various constraints, such as pitch range, harmonic structure, and adherence to a probabilistic model of a composer's style. It exploits a clustering space formed by grouping similar music segments, employing a genetic algorithm to derive the final music output. This method is especially effective when symbolic matching falls short due to unreliable data, as statistical techniques allow integration of available information. Machine learning also aids in understanding performers' behavior and information sources by characterizing them, even when data is inconsistent or noisy. Moreover, it uncovers how musicians convey emotions through their performances, enabling modifications of MIDI files to convey the desired emotions. Ultimately, machine-readable notations, enhanced with stylistic and emotional annotations, will guide synthesizers in generating expressive performances. This is particularly valuable in complex computer music systems where poorly defined problems and data variation are commonplace [6]. Cheng-Zhi Anna Huang and her colleagues made music composition more approachable by designing the first AI-powered Google Doodle, the Bach Doodle, allowing users to create their own melodies harmonized in the style of Bach using the machine learning model Coconet. To achieve this, a simplified sheet music-based interface was devised to enable users to easily input monophonic melodies. Subsequently, to support a scalable interactive experience, the team re-implemented Coconet and integrated it into TensorFlow.js to run in the browser. By introducing dilated depth-wise separable convolutions and operation fusion, they reduced Coconet's runtime from 40 seconds to just 2 seconds. Furthermore, post-training weight quantization was employed to shrink the model's download size to approximately 400KB. n the process of composition, users input their melodies and then, upon clicking the "Harmonize" button, send the harmonization request to the machine learning model. This model, trained on Bach's musical style and harmony rules, analyzes the input melodies and generates harmonized voices matching the input. The generated harmonized voices are presented one by one to the users, color-coded to denote their relation to the input melodies. The work by Cheng-Zhi Anna Huang and her team vividly showcases the potential of machine learning in the domain of music composition. By integrating machine learning models with user-friendly interfaces, they enabled individuals to participate in music creation in unprecedented ways, merging machine learning and music to craft a unique composition experience [7].

The motivation behind this article is to organize and supplement the existing knowledge about machine learning and artificial intelligence in music composition. By providing a clear overview, this study aims to explore how these advanced technologies impact and alter the ways of music composition. The subsequent sections of this article will delve into Basic Descriptions of machine learning, Principles for music composition, various Models employed, Applications, and also touch upon Limitations & Future outlook in this domain.

2. Principle for music composing

This discipline entails the development of practical software solutions tailored to a wide array of applications, including computer vision, speech recognition, natural language processing, and robot control. In recent years, machine learning has emerged as the preferred approach for many AI system developers. This preference arises from the recognition that training a system by presenting it with examples of desired input-output behavior is often far more feasible and efficient than manual programming, which would necessitate accounting for every conceivable input scenario. One significant catalyst for the widespread adoption of machine learning has been the proliferation of networked and mobile computing systems. These systems possess the capability to collect and transmit massive volumes of data. Machine learning techniques are leveraged to extract valuable insights, make predictions, and facilitate data-driven decision-making from these extensive datasets. This application of machine learning has ushered in a new era of innovation and problem-solving across various domains [8]. Machine learning includes various types, with supervised learning, where algorithms learn from labeled data, and unsupervised learning, where they find patterns in unlabeled data, being notable examples. Supervised learning teaches models to make predictions or classifications based on examples, while unsupervised learning discovers inherent data structures and relationships. These foundational approaches serve as the basis for solving diverse problems in fields like computer vision, speech recognition, and more. Certainly, machine learning encompasses additional types such as Semi-Supervised Learning, Reinforcement Learning, and multi-task learning. These variations offer distinct approaches to solving specific challenges and expanding the capabilities of AI systems [9]. Certainly, the applications of machine learning are incredibly diverse and continue to expand rapidly. In genetics and genomics, machine learning helps identify patterns in DNA sequences, predict disease risks, and advance personalized medicine. In cancer prognosis and prediction, it aids in early detection, treatment optimization, and outcome forecasting, potentially saving lives. Epilepsy research benefits from machine learning by analyzing brain activity data to better understand seizure patterns and improve treatment strategies. In minerals processing, it optimizes extraction and refining processes, reducing costs and environmental impact. Machine learning also plays a vital role in cell image analysis, automating the identification of cellular structures and abnormalities, expediting medical diagnoses and drug development. These applications represent just a fraction of the vast landscape of machine learning, demonstrating its transformative impact on various industries and scientific disciplines. Its versatility and adaptability continue to drive innovation and push the boundaries of what is achievable in the data-driven era [10].

3. Principle for music composing

Music composition using computer technology operates on fundamental principles. It begins with data input, which can be manual or sourced from existing libraries. Algorithms govern the generation of musical elements, dictating note arrangements, harmonies, and melodies. Music analysis may precede, allowing computers to identify patterns. Some systems enable real-time interaction, enhancing creativity. Music synthesis converts elements into audible music, and the final composition can be shared in various formats. Computer technology for music composition indeed involves the implementation and manipulation of various musical concepts like frequency, pulse, and form. This process often relies on mathematical models to achieve the desired musical outcomes. Several software tools and frameworks have emerged to facilitate music composition, such as Nyquist and OpenMusic. These tools provide composers and musicians with the means to experiment with and realize their creative ideas, incorporating mathematical and computational principles into the music composition process [11]. The principle of music composing using computer technology involves the use of algorithms and programming to generate or assist in the generation of new pieces of music. This process has been of interest since the 1980s and has grown in popularity with the development of AI and Deep Learning techniques. In this process, a composer starts with a basic melodic or harmonic idea, which is then developed using different harmonic progressions and structured in sections. The melodic and harmonic parts of a piece of music can be played by different instruments, which are combined using Instrumentation and Orchestration techniques. The use of computer technology can assist in generating new melodies, harmonies, and rhythms, and can be used to analyze existing pieces of music to identify patterns and structures [12].

People also utilizing generative systems to generate and compose music. The principles governing music composition by generative systems are multifaceted and contingent upon the system's design and objectives. They often draw inspiration from existing musical compositions, employing techniques of imitation and learning. These systems analyze extensive datasets of pre-existing music, discerning underlying structures, harmonies, melodies, and other essential elements. Subsequently, they endeavor to craft novel musical pieces that exhibit resemblance to these elements. For example, the general principle of the Multi-Track Music Machine (MMM) is to use a generative system based on the Transformer architecture to create multi-track music. The MMM's approach is based on a novel representation for multi-track musical material, which creates a time-ordered sequence of musical events for each track and concatenates several tracks into a single sequence. This approach takes advantage of the Transformer's attention mechanism, which can adeptly handle long-term dependencies [13].

There are several types of models used in AI-based algorithmic music composition, including heuristic composition methods, deep learning composition methods, stochastic composition methods, agent composition methods, declarative programming composition methods, and grammar composition methods. These models encompass a wide range of techniques and technologies for creating music through artificial intelligence algorithms [14]. The Anticipation-RNN, a variant of Recurrent Neural Network (RNN), is specifically engineered for the generation of symbolic music sequences [15]. Its unique characteristic centers around the concept of "anticipation," allowing the model to predict the transitions between low-pitched and high-pitched notes seamlessly, even within a left-to-right generation process. This anticipation mechanism is underpinned by a melodico-rhythmic encoding approach, which permits the model to specify when a note should be played without explicitly defining its rhythm. This encoding hinges on the notion that melody and rhythm in music are inherently interconnected, with rhythm derivable from the melody itself. The Anticipation-RNN undergoes training using a dataset comprising symbolic music sequences, learning to forecast the subsequent note in the sequence based on the preceding ones. Throughout the training phase, the model is fine-tuned to minimize the disparity between its predicted notes and the actual notes within the training dataset. Once trained, the model can generate fresh music sequences. To initiate this process, an initial note is provided, from which the model proceeds to generate subsequent notes, building upon the previous ones. The model's versatility enables it to generate sequences of varying lengths and diverse musical styles. Moreover, the Anticipation-RNN offers the capability to enforce user-defined positional constraints by adjusting the probability distributions of notes at each time step. Remarkably, the model can produce captivating musical phrases while adhering to these constraints, demonstrating proficiency in generating convincing melodies akin to those found in Bach-like chorales. DeepBach is a graphical model designed to generate polyphonic music in the style of Johann Sebastian Bach [16]. The model is trained on Bach's chorale harmonizations and utilizes pseudo-Gibbs sampling and a unique representation of musical data to generate convincing chorales that closely resemble Bach's style. Unlike other automatic music composition approaches, DeepBach is capable of generating music in a non-sequential manner and allows users to impose positional constraints on the generated score. The model is steerable, meaning that users can constrain the generation by imposing constraints such as notes, rhythms, or cadences in the generated score. This feature allows for interactive composition and enhances creativity. The architecture of DeepBach includes a neural network with one hidden layer for local interactions between a note and its context. The model also includes two stacked LSTM units for capturing musical dependencies. Dropout techniques are applied to improve the model's performance. To facilitate interaction with DeepBach, a plugin for the MuseScore music editor was developed, allowing users to easily call DeepBach on any rectangular region and generate or reharmonize chorales. DeepBach presents a versatile and efficient solution for automatic music composition in the spirit of Johann Sebastian Bach, granting users the capability to create and interact with polyphonic music reminiscent of Bach's distinctive style.

4. Application

Computer music finds diverse applications in contemporary music composition, harnessing algorithms, software, and computational models to foster original musical expressions and artistic creations. Algorithmic composition represents a facet of computer music, employing programming languages and software to automate music generation. Notable software like GarageBand, Chordbot, and TonePad exemplify this approach [17]. Recent years have witnessed substantial advancements in computer music, propelled by disruptive technologies like artificial intelligence. Machine learning, neural networks, and evolutionary computation have been instrumental in augmenting creativity and catalyzing innovation in music composition. Genetic algorithms, for instance, have been deployed to discover harmonious counterpoints for music samples and craft accompaniments adhering to predefined rules. Furthermore, interactive genetic algorithms have emerged to evolve chord templates based on feedback from listeners. Computer music has permeated every facet of music composition, encompassing rhythm, melody, and accompaniment. Computational tools such as genetic algorithms and neural networks have been instrumental in evaluating and generating these elements within computer-generated music. The evolution of visual programming languages and artificial intelligence software has further propelled the development of sophisticated and intelligent tools in this domain. Practical applications of computer music span data-driven interactive systems and digital audio technology. Data-driven interactive systems employ both rule-based and non-adaptive approaches, diligently analyzing and categorizing musical data to discern critical patterns and unveil underlying statistical behaviors. Within this realm, probabilistic generative modeling represents a powerful tool, elucidating and interpreting musical features for training future models or crafting music styles with precision. Meanwhile, digital audio technology has revolutionized music production, empowering composers to establish their studios for comprehensive post-production tasks. These include multi-track recording, waveform manipulation, and mastering tape processing, culminating in a paradigm shift in contemporary music production. This evolution empowers composers with greater autonomy and control over the creative process.

One classic application is probabilistic generative modeling, which leverages data-driven methodologies to capture and explain various musical features, such as chord occurrence probabilities or note transition frequencies [18]. These findings serve as valuable training data for future models designed to produce specific music styles. Summary statistics extracted from these models also contribute to the generation of new music, combining extracted features in diverse settings. Deep learning, while starting relatively late in its application to music composition, has had a significant impact. Deep learning integrates a vast number of computational units to effectively discern the latent features of musical melodies. Deep learning models can not only reconstruct the input melodies but also comprehend the long-term structure of music with minimal human intervention. For instance, Google's Magenta project, leveraging music and image data, trains machine learning models to create new content, and Music Transformer serves as an example. In this project, a group of musicians enhances the coherence of a song by modeling the connections between individual notes and making modifications [19].

5. Limitations & Future outlook

Copyright issues pose a series of significant challenges in computer music, with the most prominent being limitations and lag in machine learning databases. Due to copyright restrictions, many classical music pieces and modern tracks cannot be included in training datasets, resulting in a lack of diversity and breadth in machine learning models' databases. This can make it more challenging for models to capture different music styles and elements, thus reducing their effectiveness in generating music. Furthermore, copyright problems can lead to data incompleteness and inaccuracy. Some music data may be restricted by copyright, making it impossible to obtain complete information when training models. This can affect the model's understanding and analysis of music, as they cannot fully grasp the music's inherent structure and characteristics. This can result in significant gaps between the generated music and original compositions, diminishing the quality and creativity of the music. Additionally, computer music itself faces several challenges. For example, it's difficult to fully capture the emotions and creativity in music through algorithms, as music often involves deep emotional expression. Moreover, accurately simulating complex music styles can be challenging, as generating music similar to existing works may involve potential copyright issues. The user interfaces of computer music software can also be complex, limiting some people's usage. Furthermore, computer music often requires extensive data and models, making it resource intensive. The creative process can also be constrained by algorithm limitations, affecting the uniqueness of compositions. Finally, music is highly subjective, meaning that people may have varying opinions on what constitutes beautiful music. Despite the formidable challenges posed by copyright issues and the inherent complexities of computer music, there are promising avenues for the future of this innovative field. Advancements in copyright management and licensing systems could facilitate greater access to a broader range of musical compositions for training machine learning models. This would address the issue of limited diversity in training datasets and enable models to better capture various music styles and elements. Improved copyright frameworks may also encourage artists and rights holders to participate more actively in making their work available for creative AI applications. As machine learning algorithms continue to evolve, there is potential for greater emotional and creative expression in computer-generated music. Emotion-aware algorithms and deep learning techniques may better capture the nuanced emotional content of music, allowing for more emotionally resonant compositions. Addressing the challenge of accurately simulating complex music styles while avoiding copyright infringement can be achieved through the development of AI models that emphasize creativity and originality. These models could generate music inspired by various styles without directly copying copyrighted material, thus avoiding legal issues. Improvements in user interfaces for computer music software will likely make the field more accessible to a broader audience. User-friendly tools and platforms could democratize music creation, allowing individuals with varying levels of musical expertise to engage in the creative process. Advancements in data collection and the availability of large, diverse datasets will continue to drive progress in computer music. As more musical data becomes accessible, models will benefit from a richer source of inspiration, leading to more innovative compositions. While algorithmic constraints can impact the uniqueness of compositions, ongoing research will likely address these limitations. Customizable AI music composition tools will empower musicians and composers to work collaboratively with AI, striking a balance between human creativity and machine-generated assistance. Lastly, the subjectivity of music will remain a challenge, but it also offers opportunities for customization. AI systems that adapt to individual preferences and styles may enhance the listener's experience by tailoring music to their unique tastes. In conclusion, despite the hurdles, the future of computer music appears promising. By overcoming copyright challenges, improving algorithms, enhancing user interfaces, and embracing the subjectivity of music, computer music has the potential to become an invaluable tool for musicians and composers, fostering creativity and expanding the boundaries of musical expression.

6. Conclusion

The implementation of machine learning and artificial intelligence for music composition represents a transformative leap in the creative possibilities of the music industry, ushering in an era of innovative compositions and personalized musical experiences. This paper delves into the evolution of computer music, tracing its roots from early electronic music pioneers to contemporary applications of machine learning and AI. It explores various models and techniques used in computer music, including rule-based systems, deep learning, and stochastic composition methods. The article also highlights the significant challenges posed by copyright issues and the potential solutions to overcome them. Looking to the future, it envisions computer music offering greater emotional depth, improved accessibility, and customization, despite the hurdles it faces. Despite the formidable challenges posed by copyright issues and the inherent complexities of computer music, there are promising avenues for the future of this innovative field. Advancements in copyright management and licensing systems could facilitate greater access to a broader range of musical compositions for training machine learning models, leading to more diverse and emotionally resonant computer-generated music.The significance of this study lies in its exploration of the potential of computer music combined with machine learning, highlighting the challenges of copyright and the intricacies of music composition, providing insights for future music creation and technological innovation.

References

[1] C. Roads and M, Mathews, Computer Music J., **4(4)**, 15-22 (1980).

[2] M. Edi, *The Oxford Handbook of Algorithmic Music* (Oxford University Press, Oxford, 2018).

[3] C. David, Retrieved from: http://artsites.ucsc.edu/faculty/cope/.

[4] C. Roads, *The computer music tutorial,* (MIT press, Matur, 1996).

[5] M. Chan, J. Potter, and E. Schubert, *Improving algorithmic music composition with machine learning*, In 9th International Conference on Music Perception and Cognition (2006), pp. 1848-1854.

[6] R. B. Dannenberg, *Artificial intelligence, machine learning, and music understanding*, In Proceedings of the 2000 Brazilian symposium on computer music: arquivos do simpsio brasileiro de computao musical (2000, SBCM).

[7] C. Z. A. Huang, C. Hawthorne, A. Roberts, et al., *The bach doodle: Approachable music composition with machine learning at scale*, arXiv preprint arXiv:1907.06637 (2019).

[8] M. I. Jordan and T. M. Mitchell, Science, **349(6245)**, 255-260 (2015).

[9] B. Mahesh, Inter. J. of Sci.& Res., **9(1),** 381-386 (2020).

[10] P. P. Shinde and S. Shah, A review of machine learning and deep learning applications. 2018 Fourth international conference on computing communication control and automation (ICCUBEA) (2018), pp. 1-6).

[11] E. Miranda, *Composing music with computers*, (CRC Press, London, 2001).

[12] C. Hernandez-Olivan, and J. R. Beltran, Adv. in Speech a music technology: computational aspects and applications, **11**, 25-50 (2022).

[13] J. Ens and P. Pasquier, *Mmm: Exploring conditional multi-track music generation with the transformer*, arXiv preprint arXiv:2008.06048 (2020).

[14] O. Lopez-Rincon, O. Starostenko, and G. Ayala-San Martín, *Algoritmic music composition based on artificial intelligence: A survey*. 2018 International Conference on Electronics, Communications and Computers (CONIELECOMP) (2018), 187-193.

[15] G. Hadjeres, and F. Nielsen, Interactive music generation with positional constraints using anticipation-rnns. arXiv preprint arXiv:1709.06404.

[16] G. Hadjeres, F. Pachet, and F. Nielsen, *Deepbach: a steerable model for bach chorales generation*, international conference on machine learning (PMLR, 2017), pp. 1362-1371.

[17] Q. Fu, J. of Phys. Conf. Series, **1820(1)**, 012153 (2021).

[18] Understanding and Mining Patterns of Audience Engagement and Creative Collaboration in Largescale Crowdsourced Music Performances. MIDAS, 2020 Retrieved from: https://midas.umich.edu/research-hubs/music/koutra/. [Accessed: 21- Dec- 2020].

[19] Cs224d.stanford.edu, 2020. Retrieved from: http://cs224d.stanford.edu/reports/allenh.pdf. [Accessed: 21- Dec- 2020].