IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 22 May 2024, accepted 19 June 2024, date of publication 24 June 2024, date of current version 5 July 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3419050

TOPICAL REVIEW

Development Status, Frontier Hotspots, and Technical Evaluations in the Field of AI Music Composition Since the 21st Century: A Systematic Review

WEIJIA YANG [1,4], (Member, IEEE), LIN SHEN [2], (Member, IEEE), CHIH-FANG HUANG [3], (Member, IEEE), JOHEE LEE [4], AND XIAN ZHAO [5]

[1]College of Humanities, Arts and Education, Shandong Xiehe University, Jinan 250109, China
[2]Music Technology Department, Music College, Shanghai Normal University, Shanghai 200234, China
[3]Department of Health and Marketing, Kainan University, Taoyuan City 33857, Taiwan
[4]Department of Performance Music, Kyonggi University, Seoul 03746, Republic of Korea
[5]Film-Television and Communication College, Shanghai Normal University, Shanghai 200234, China

Corresponding author: Lin Shen (shenlin@shnu.edu.cn)

This work was supported by the 2023 Youth Fund Project of the Ministry of Education (MOE) Foundation on Humanities and Social Sciences, under the grant titled "AI Music Application Research in Animation Sound Design Education in Universities" with phased results under Project 23YJC760169.

ABSTRACT In recent years, "Artificial Intelligence (AI)" has become a focal point of discussion. AI music composition, an interdisciplinary field blending computer science and musicology, has emerged as a prominent area of research. Despite rapid advancements in AI music creation technology, there remains a dearth of comprehensive surveys addressing the core technologies within this domain. To address this gap, this study conducted a comprehensive search across multiple databases spanning a 23-year period (2000–2023) on the topic of "AI music composition." Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) standard for literature screening, the study systematically organized the development status, frontier hotspots, and technical evaluations of the field. Drawing from literature data, the study verified Price's Law, Lotka's Law, and Bradford's Law—three scientific productivity laws—while summarizing the current landscape from four perspectives: authors, organizations, countries, and journals. Subsequently, utilizing VOSviewer and CiteSpace, two technical software tools, the study conducted an in-depth analysis consisting of four steps: clustering, time zone, burst words, and high-frequency referenced literature. The study presented the evolution trajectory of frontiers and hotspots through visualization maps. Finally, building upon quantitative statistical insights, the study qualitatively expanded research efforts by organizing and evaluating the latest AI music generation algorithm technologies. The systematic literature analysis, both quantitative and qualitative, aims to furnish researchers and practitioners in related fields with systematic references.

INDEX TERMS Artificial intelligence, automatic music composition, bibliometrics, PRISMA.

I. INTRODUCTION

In the early 21st century, the advancement of science and technology ushered in a new era symbolizing the maturation of AI technology, characterized by automated productivity and intelligent creativity [1]. The emergence of the global "Industry 4.0" wave in 2013, with AI as its cornerstone, further accelerated the development of AI. By the end of 2022, the disruptive global popularity of OpenAI's chatbot model "ChatGPT" garnered widespread attention,

The associate editor coordinating the review of this manuscript and approving it for publication was Mingbo Zhao.

showcasing AI's cross-disciplinary achievements. The convergence of AI with diverse disciplines and fields, such as AI music therapy [2], AI intangible cultural heritage [3], and AI metaverse concerts [4], has garnered global interest. Among these intersections, AI Music Composition emerges as a fundamental technology, effectively expanding the realm of music composition to encompass music generation and application [5].

The fundamental principle of AI Music Composition lies in utilizing computer language to delineate the framework of music theory, establishing diverse libraries of musical phrase materials, training corresponding music style models through machine learning, and ultimately generating novel digital music content [6]. This research avenue has persisted for over four decades [7], predominantly from the standpoint of computer science, with a focus on topics like algorithm models, system design, and industrial applications. For instance, endeavors include spatial model training for drum rhythms [8], the development of a song generation system based on lyrics [9], and the creation of an AI music generation system tailored for automotive safety [10], typically employing case analysis and empirical research methodologies.

However, there remains a scarcity of studies conducting bibliometric analyses from a statistical perspective. Despite a few review studies [11], [12], there is still a gap in systematically evaluating the technology. Given the interdisciplinary nature of AI Music Composition, diverse interdisciplinary research is crucial for its development [13]. Quantitative analysis and qualitative evaluation of relevant publications [14] can aid researchers in discerning the field's primary developmental trajectory and technical frontiers amidst its complexity [15]. Therefore, this paper undertakes a systematic bibliometric analysis of the AI Music Composition field, encompassing its overall development status, frontier hotspots, and technical evaluation. The main research questions (RQs) addressed in this article are:

- RQ1: About development status. Who are the main authors of papers in this field? What are their affiliated organizations and countries? Which journals feature these research papers prominently?
- RQ2: About frontier hotspots. What are the most frequent keyword hotspots in this field's papers? What are some notable research papers? What temporal distribution and development trends do they exhibit?
- RQ3: About technical evaluation. What are the mainstream algorithm technologies in this field? How can they be categorized? What are some representative case studies?

II. METHODS

A. RESEARCH METHODOLIGY

1) ABOUT BIBLIOMETRIC

Bibliometrics serves as a widely utilized quantitative research method for literature analysis, focusing on the quantitative examination of information such as authors, keywords, affiliated journals, and references, thereby offering a comprehensive overview of the research landscape through specific data [16]. Additionally, presenting data results through visual charts offers the advantage of clarity and ease of understanding [17]. Leveraging distinctive algorithm modules such as "clustering," "time zone," and "burst" [14], [18], this study swiftly presents various literature data in complex analyses, providing crucial data references for the quantitative analysis process [19]. Hence, this study employs bibliometric visualization maps to conduct an analysis of the AI music composition field.

2) BIBLIOMETRIC TOOLS

In bibliometric research, the crafting of visualization maps necessitates the use of specialized software tools. VOSviewer [20] and CiteSpace [21] stand as the primary visualization tools in bibliometric research [22], capable of directly importing data from databases such as Web of Science (WOS), Scopus, IEEE Xplore, among others, to generate visualization maps [23].

VOSviewer relies on probabilistic methods, featuring Multi-Dimensional Scaling (MDS) and Random Walk (RW) algorithms at its core. It particularly excels in key author collaboration network analysis, key journal distribution analysis, and co-word clustering analysis, showcasing a prominent advantage in these areas [24].

On the other hand, CiteSpace operates on set theory methods, incorporating Latent Semantic Indexing (LSI), Log-Likelihood Ratio (LLR), and Mutual Information (MI) algorithms. These algorithms are highly advantageous in conducting topic evolution analysis, keyword burst intensity analysis, and literature co-citation network analysis [25], [26].

Leveraging the distinctive strengths of each software tool, they are employed collaboratively to provide a comprehensive analytical perspective. Therefore, this study utilizes both VOSviewer and CiteSpace to generate quantitative visualization maps of information such as authors, keywords, and co-citations within the literature. To ensure standardized map presentations, this study also employs Microsoft Excel [27], Adobe Photoshop [28], Scimago Graphica [29], and RAW Graphs [30] for visual adjustments.

3) PRISMA GUIDELINES

The data source for visual bibliometric analysis necessitates rigorous screening. Directly conducting visual bibliometric analysis with search data entails certain errors, and the varying quality of literature can lead to inaccuracies in keyword extraction and co-citation analysis by software tools. Hence, further precise screening of the search data is essential [18].

In light of this, this study adopts the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement to systematically complete the literature screening process, gradually identifying effective literature [31]. PRISMA serves as an authoritative reporting standard for systematic literature reviews, with the latest 2020 statement providing detailed explanations and reporting recommendations for 27 checklist items [32]. Particularly in the literature screening phase of the database, PRISMA offers a standardized process and clearly delineates the inclusion and exclusion criteria for literature [33], ensuring the completeness and accuracy of literature statistics.

Consequently, the research methodology of this paper is as follows: Utilize databases to search for literature in the AI Music Composition field since the 21st century, adhering to PRISMA systematic literature screening standards for data screening. Report the development status by organizing information such as authors, publications, institutions, and countries. Delineate frontier hotspots through clustering analysis of co-occurring keywords, co-cited literature, etc. Present technical evaluations by selecting representative AI music generation algorithm cases.

B. DATA RETRIEVAL

1) DATA SOURCES

This study selects Web of Science (WOS), IEEE Xplore, ACM Digital Library, and Scopus as the data sources. These databases are widely recognized by experts and scholars as the most suitable high-quality digital literature resource databases for bibliometric analysis [22]. They offer a range of data including literature titles, authors, references, citation counts, impact factors, etc. [34].

To ensure a more accurate grasp of the high-quality research context, the scope of indexed articles in this study first focuses on the core collection of Web of Science. This database predominantly hosts many of the highly cited quality research articles. However, given the interdisciplinary nature of AI Music Composition, this study also incorporates IEEE Xplore, ACM Digital Library, and Scopus databases in the search plan [35]. Therefore, selecting these four high-quality databases as the data source is reasonable, as it enables the retrieval of a large number of high-quality articles for screening and analysis.

2) SEARCH STRING

This study references several published journals to establish the keyword search criteria. Ultimately, three keywords were selected for indexing in combination: "Music," "Generation," and "AI" [11], [36], [37]. Considering the diversity of keywords, this study also incorporates alternative terms. These derivative terms primarily focus on the AI and Generation aspects: "artificial intelligence," "machine learning," "deep learning," "algorithm," and "automatic." These five sets of vocabulary collectively form the derivative expressions for the "AI" component. Additionally, the truncated spelling of the two words "generation" and "composition" constitute the derivative expressions for the "Generation" component. The search content in each database remains consistent, with slight adjustments made according to the characteristics of each database. The specific search strategy used in WOS is:

TS = (("Music") AND ("Generat*" OR "Compos*") AND ("Artificial Intelligence" OR "Deep learning" OR "Machine Learning" OR "Algorithmic" OR "Automatic"))

3) INCLUSION/EXCLUSION CRITERIA

To ensure the generation of high-standard bibliometric statistics results, this study has delineated detailed inclusion criteria for articles across three aspects: "search period," "search language," and "document type." Given that the pivotal developmental period of artificial intelligence spans the 21st century, this study selects January 1, 2000, as the starting date for the literature search, with the end date set as the time of data organization for this study, i.e., December 31, 2023. In screening for high-quality journal articles, this study limits the document type to Articles and Review Articles, which necessitate peer review. Additionally, the search language for the literature is confined to English. The specific search details are outlined in Table 1.

TABLE 1. Summary of Inclusion/exclusion criteria.

Category	Specific Standard Requirements
Research database	WOS, IEEE, ACM, Scoups
Search keywords	("Music") AND ("Generat*" OR "Compos*") AND ("Artificial Intelligence" OR "Deep learning" OR "Machine Learning" OR "Algorithmic" OR "Automatic")
Search period	2000-01-01 to 2023-12-01
Language	English
Document types	"Articles" or "Review Articles"

4) ARTICLE SCREENING RESULTS

A total of 1519 documents (comprising 1467 papers and 52 review papers) were retrieved through the search, with the search results from the four databases detailed in Table 2.

TABLE 2. Search results.

Database	Quantity
WOS	929
IEEE	229
ACM	24
Scopus	337

In order to obtain more accurate and high-quality literature data, this study adheres to the PRISMA standard and conducts literature screening. The screening process is illustrated in Figure 1. Zotero literature management software [38] was utilized for four rounds of excluding retrieved literature data. Three authors independently reviewed the papers based on the title, abstract, and full-text content, with papers under dispute being reviewed by the team and consulted with relevant experts to determine whether they should be eliminated. The screening process comprises four parts: initial processing via

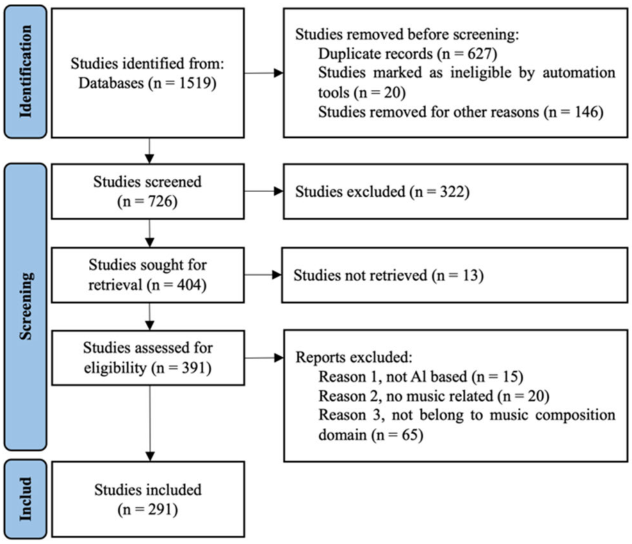

FIGURE 1. PRISMA article screening methodology.

literature management software, preliminary screening based on weak relevance of keywords, screening for effective literature retrieval, and thorough evaluation of full-text relevance.

5) DATA ABSTRACTS

To maintain consistency and rectify potential discrepancies in the data from the selected articles, we adopted a standardized data processing approach. Special attention was directed towards ambiguities arising from non-standard information in areas such as author names, country designations, and keywords. Pursuant to this, the following corrective measures were undertaken:

- Author Name Verification: Our approach took into account various nuances, including identical names, previously used names, abbreviations, and other relevant variations.
- Standardization of Country Designations: To ensure consistency, we unified China's three Special Administrative Regions into a single "China" designation. Additionally, we grouped the four Commonwealth nations under the "UK" label. Additionally, we adopted consistent abbreviations for all countries.
- Keyword Standardization: We streamlined the representation of keywords by combining similar terms, balancing keyword lengths, and categorizing both high and low-frequency keywords appropriately.

Following a thorough analysis of a carefully curated selection of papers, a wealth of valuable data was uncovered (see Table 3).

C. LIMITATIONS

This study has three main limitations. Firstly, while the four mainstream databases utilized in this study generally cover a wide range of research literature, there are inherent limitations. Despite WOS, IEEE, and other databases being high-quality sources that encompass much of the research on AI music generation, the possibility of missing articles remains. Secondly, the study is constrained by the inclusion criterion restricting articles to English language only. Although the search identified some high-quality articles in other languages (such as Chinese, Korean, or Spanish), they were excluded based on the language criterion. Thirdly, despite employing a combination of quantitative and qualitative analysis methods, the review is still limited by the statistical measurement capabilities of VOSviewer and CiteSpace, two analysis software tools. While these tools are representative statistical instruments, there is room for improvement in terms of information comprehensiveness and algorithm consistency. Hence, future researchers may consider incorporating multiple quantitative and qualitative indicators to augment this aspect.

TABLE 3. Data abstracts of screening articles.

Category	Quantity
Publications	291
Authors	556
Organizations	285
Countries	39
Journals	99
Co-occurrence keywords	784
Co-cited references	7266

III. RESULT VERIFICATION

Referring to the laws of scientific productivity aids in interpreting research results [39]. This study utilized two analytical software tools, VOSviewer and CiteSpace, to conduct a quantitative analysis of all the literature's authors, organizations, countries, and journal data, and described the field's development status in accordance with Price's Law, Lotka's Law, and Bradford's Law.

A. PUBLICATION OVERVIEW

The annual distribution of journal numbers serves as a reflection of the research level and development degree of this academic field. Figure 2 illustrates the temporal distribution of 291 papers in the AI Music Composition research field. Overall, the number of publications in this field has experienced a sharp rise over the past decade. Prior to 2016, the annual publication number remained at a very low level, averaging about 3 papers per year, with no relevant papers published in 2000 and 2002. However, post-2016, there has been a rapid increase in the annual publication number, with last year witnessing explosive growth, reaching 94 papers annually. This trend indicates that the research field has entered a mature development stage.

B. AUTHOR STATUS

Quantitative analysis of paper authors serves as the primary reporting indicator for describing the current status. The study employed Price's Law to verify the retrieved paper

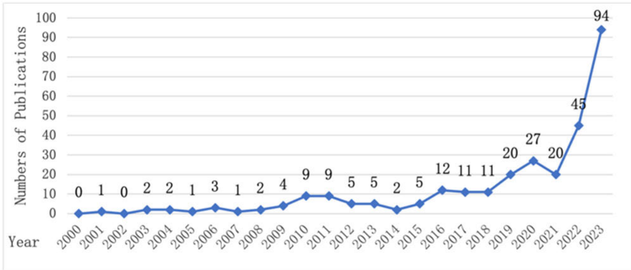

FIGURE 2. Distribution of publications from 2000 to 2023.

author data, aiming to ascertain whether a scientifically stable author collaboration group has emerged in the AI Music Composition field and to identify the current development status.

The distinguished scholar Price observed that within the same subject, half of the papers are authored by a group of highly productive authors, and the total number of papers produced by this core group of authors is approximately equal to the square root of the total number of papers [40]. If these numbers are indeed equal, it demonstrates the formation of a scientifically stable author collaboration group. The formula is:

$$\sum_{m+1}^{I} n\left(x\right) = \sqrt{N} \tag{1}$$

- "N" represents the total number of papers published on the subject.
- "n(x)" represents the cumulative number of papers by authors who have published "x" papers.
- "m+1" is the initial value of the cumulative sum.
- "I" is the final value of the cumulative sum.

Determining "m+1" and "I" facilitates the verification of the formula. Within this calculation framework, "I" is synonymous with "n", both representing the number of papers by the most prolific author in the field. According to the statistics from VOSviewer, the most prolific author has published 7 papers, thus I = 7. "m" denotes the threshold value for the number of papers published by core authors. Price, drawing on Lotka's Law [41], outlined the minimum number of papers "m" published by core authors.

$$m = 0.749 \times \sqrt{n_{max}} \tag{2}$$

After calculation, the threshold value of the number of papers published by core authors in this study is approximately 1.98. Hence, authors who have published 2 or more papers (including 2) are classified as core authors in this field. Substituting the aforementioned values into the formula, a total of 66 core authors have published 162 papers, constituting 75% of the total number of papers. This figure meets Price's standard of half of the papers (50%). Consequently, it can be inferred that the AI Music Composition research field has already established a relatively scientific and stable author collaboration group, indicating that the research is in a relatively mature development stage.

C. ORGANIZATION STATUS

The quantitative analysis data of authors' institutions serves as the second reporting indicator for describing the current status. The relationship between authors and their institutions is integral, with 285 institutions from various countries contributing to the AI music composition research field. Table 4 outlines the top ten institutions based on the number of papers published. Primarily, these institutions consist of professional departments within universities.

TABLE 4. Top 10 organizations in the AI music composition field.

Rank	Organizations	Documents	Citations	Average Citation /Publication
1	Dongguk University of Seoul, Korea	7	31	4.43
2	Queen Marys University of London, England	6	123	20.5
3	University of Plymouth , England	6	62	10.33
4	Communication University of China	5	41	8.2
5	University of Malaga, Spain	4	152	38
6	Singapore University of Technology and Design	4	118	29.5
7	University of the Basque Country - UPV/EHU, Spain	4	35	8.75
8	Pompeu Fabra University, Spain	4	24	6
9	University of Sussex, England	4	15	3.75
10	National Chiao Tung University of Taiwan, China	4	6	1.5

Dongguk University of Seoul (Korea, Seoul) leads in the number of papers published in this research field, with 7 papers, notably attributed to the Yunsick Sun and Shuyu Li team mentioned earlier. Additionally, among the top ten institutions, there are 3 from Spain and the UK each, and 2 from China, all demonstrating high citation counts. This trend indicates, to some extent, that institutions in Europe and Asia hold prominent positions in this field. Particularly noteworthy is the University of Malaga in Spain, with the highest number of citations (152 papers) and the highest average number of citations per paper (38 times), closely associated with Professor Francisco Jose Vico Vela and his Melomics automatic music creation system. Furthermore, the institution with the second-highest average number of citations (29.5 times) is the Singapore University of Technology and Design, where Professor Dorien Herremans and his AMAAI laboratory have made significant contributions.

D. COUNTRY STATUS

Regional analysis data constitutes the third reporting indicator for the current status description. By assessing the country of origin across 39 countries, the number of publications is measured to analyze the contribution of different countries to research in the field of AI Music Composition. Price's Law can be leveraged not only to analyze core authors in a field but also to determine the threshold value of the number of papers published by high-yield countries [42]. According to statistics from VOSviewer, the highest number of papers in the field is 57. By substituting this value into the previously mentioned formula (2), the minimum number of papers for high-yield countries can be calculated as: $m \approx 6$. Consequently, countries with a publication volume of $\geqslant$6 papers are classified as high-yield countries in this field, and the results are depicted in Figure 3.

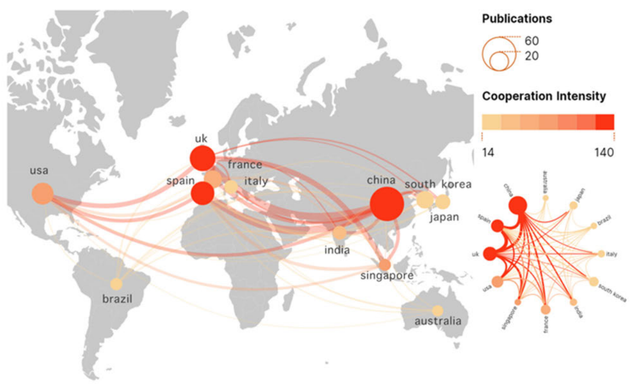

FIGURE 3. Geographic distribution and cooperation intensity map of high-yielding countries.

In Figure 3, the size of the circles corresponds to the number of publications, while the thickness of the lines reflects the strength of cooperation between countries in terms of publications. From the figure, it's evident that the distribution of publishing countries in this field is not balanced, with a significant top-heavy effect. The majority of papers are authored by scholars from a few countries, such as China, the UK, Spain, and the USA, predominantly located in Asia, Europe, and North America. This observation aligns with the conclusions drawn from the analysis of key institutions mentioned earlier.

E. JOURNAL STATUS

The quantitative analysis data of the journals hosting the articles serves as the fourth reporting indicator for describing the current status. Utilizing Bradford's Law, the 99 journals obtained from the search are categorized into zones to investigate whether the AI Music Composition field adheres to the scientific distribution pattern of the number of articles published in journals and to ascertain whether the field has established a stable system of journals hosting articles.

Bradford's Law facilitates the division of journals obtained from the search into zones. Bradford observed that in a disordered collection of scientific literature, the distribution of literature and corresponding journals is highly asymmetrical or skewed, and this distribution follows a certain quantitative relationship [43]. By ranking journals in descending order based on the number of papers published in a specific discipline, it's possible to delineate the core area and subsequent continuous zones. The number of papers published in each zone follows a ratio of $1 : a : a^2 : \ldots$

According to Bradford's Law, papers and journals are categorized into three levels of zones, as depicted in Table 5. The First Zone comprises journals with a publication volume of $\geqslant$6 papers, totaling 8 journals with 74 papers published. The Second Zone includes journals with a publication volume of 2 to 5 papers, totaling 29 journals with 80 papers published. The Third Zone consists of journals with a publication volume of 1 paper, totaling 62 journals with 62 papers published. The number of papers in the three zones fluctuates around 72, and the ratio of the number of journals is approximately $1 : 3 : 3^2$. This suggests that research papers in the AI Music Composition field generally adhere to the scientific distribution pattern of the number of articles published in journals.

TABLE 5. Data abstracts of screening articles.

Zone	Publications /Journal	Journals	Publications
First Zone	≥6	8	74
Second Zone	2-5	29	80
Third Zone	1	62	62

The journal zoning in this field is visualized through a heat map. Figure 4 demonstrates the regional distribution of the 99 journal nodes, categorized by publication volume, offering a clear depiction of the three-tiered zoning of journals. Each journal is distinguished by font size, reflecting its publication volume. The deep color cluster in the central area represents the First Zone, while the outer ring of the seven-color cluster area represents the Second Zone. The outermost area, characterized by weak connections and light color, signifies the Third Zone. The results derived from Bradford's Law regarding the distribution of the number of publications in journals, coupled with the comprehensive analysis of the visualization map, indicate that the field has indeed established a stable system of journals hosting articles.

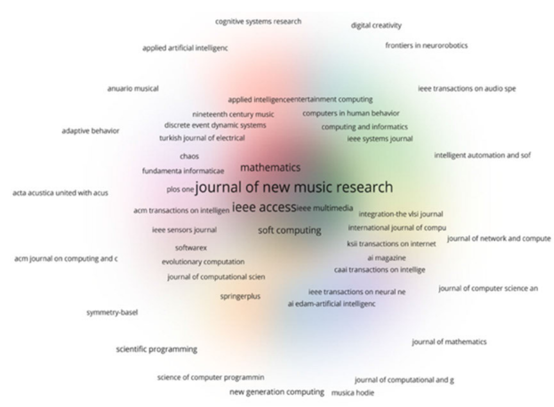

FIGURE 4. Regional distribution map of papers and journals.

In summary, this study has conducted a comprehensive analysis of literature data across four dimensions: authors, organizations, countries, and journals, presenting the research status through data results. Firstly, an analysis of 556 authors was conducted, verifying the author data of the papers using Price's Law, which revealed the formation of a scientifically stable author collaboration group within the field. Next, an analysis of 285 organizations was performed, indicating that these institutions are predominantly professional departments of universities, with collaboration between institutions being common. Subsequently, an analysis of 39 countries revealed that Europe and Asia are the main contributing regions, with representative countries including Spain, the UK, China, and South Korea. Finally, an analysis of 99 journals was conducted, and it was verified using Bradford's Law that the field adheres to a scientific distribution pattern in terms of the number of articles published in journals, thereby indicating the establishment of a stable system of journals hosting articles.

IV. HOTSPOT ANALYSIS

A. ANALYSIS OF CO-OCCURRENCE KEYWORDS

Keywords serve as highly condensed representations of the core essence of papers. Analyzing co-occurring keywords enables a depiction of the broader context within the research field and offers insights into research hotspots. Co-occurring keyword analysis primarily focuses on clustering high-frequency keywords [44]. In this study, Price's Law was employed to analyze core authors and prolific countries, and this method can also be used to determine the threshold for the number of high-frequency keywords [42]. According to VOSviewer statistics, there are 41 keywords with the highest frequency. By applying this value to formula (2), the threshold value for the number of high-frequency keywords, referred to as "m," is approximately 4.8. Consequently, keywords with a frequency of $\geqslant 5$ are deemed high-frequency keywords in this field, totaling 32. Visual co-occurrence network maps for these high-frequency keywords were generated using VOSviewer, resulting in Figure 5, which illustrates 32 nodes (high-frequency keywords) divided into 3 clusters (keyword clusters).

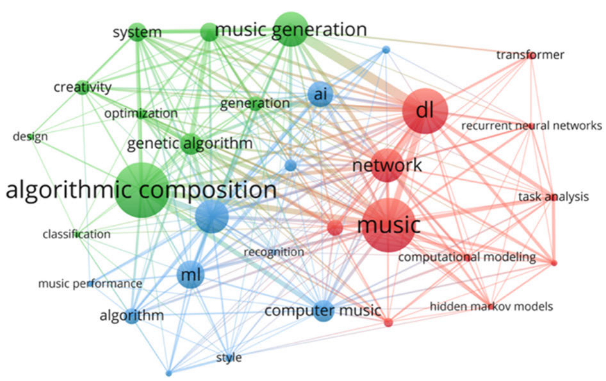

FIGURE 5. The cluster map of co-occurrence keywords.

Analyzing keyword clusters enables researchers to gain deeper insights into the main research trends in a specific field, aiding in understanding the specific aspects of each trend. In the figure, the node size reflects the frequency of occurrence of the keywords; larger nodes indicate higher frequencies. The thickness of the connecting lines between nodes indicates the strength of their correlation; thicker lines suggest more frequent co-occurrences in the same document. Different colors of clusters represent distinct research topics. As depicted in Table 6, this study summarized the vocabulary of the three clusters based on their occurrence frequency. Subsequently, we will analyze these three keyword clusters individually.

TABLE 6. Cluster of keywords in the AI music composition field.

Cluster	Color	Keywords
1		Algorithmic Composition, Music Generation, Genetic Algorithm, Music Composition, System, Creativity, Generation, Optimization, Classification, Design
2		Model, ML, AI, Computer Music, Algorithm, Markov Models, Perception, Style, Multiple Viewpoint Systems, Music Performance, Recognition
3		Music, DL, Network, Melody, Rhythm, Computational Modeling, Task Analysis, Hidden Markov Models, Music Information Retrieval, Recurrent Neural Networks, Training, Transformer

- Cluster 1: The green cluster of keywords centers on overall system design. Within this cluster, the design of composition systems is a significant research area, encompassing studies on structured music sequence design [45], music creativity mining [46], and the application of genetic algorithms [47]. Algorithm optimization for composition is also crucial, with topics like multi-objective optimization model design [48], evolutionary computing [49], and adaptive music classification [50] gaining popularity.
- Cluster 2: The blue cluster of keywords focuses on machine learning algorithms. Here, the Markov chain probability model holds a pivotal role as a classic algorithm model for music generation in machine learning, with its transition probabilities forming the core of its research technique [51]. Markov chains are continuously optimized through research, with topics like non-homogeneous Markov chains [52], hidden Markov chains [53], and Markov decision processes (MDPs) [54] gradually emerging as focal points, predominantly applied in automatic harmony arrangement. Hidden Markov chains also witness significant developments in the realm of deep learning. Style transfer for personalized music performance [55] is another vital research direction in this area, aiming to achieve highly realistic simulation of AI-generated traditional music, covering instrument sounds to overall note performance [56]. Instrument sound modeling [57], [58], [59], vocal synthesis [60], emotional recognition perception [61], and related topics represent active research areas in this domain.

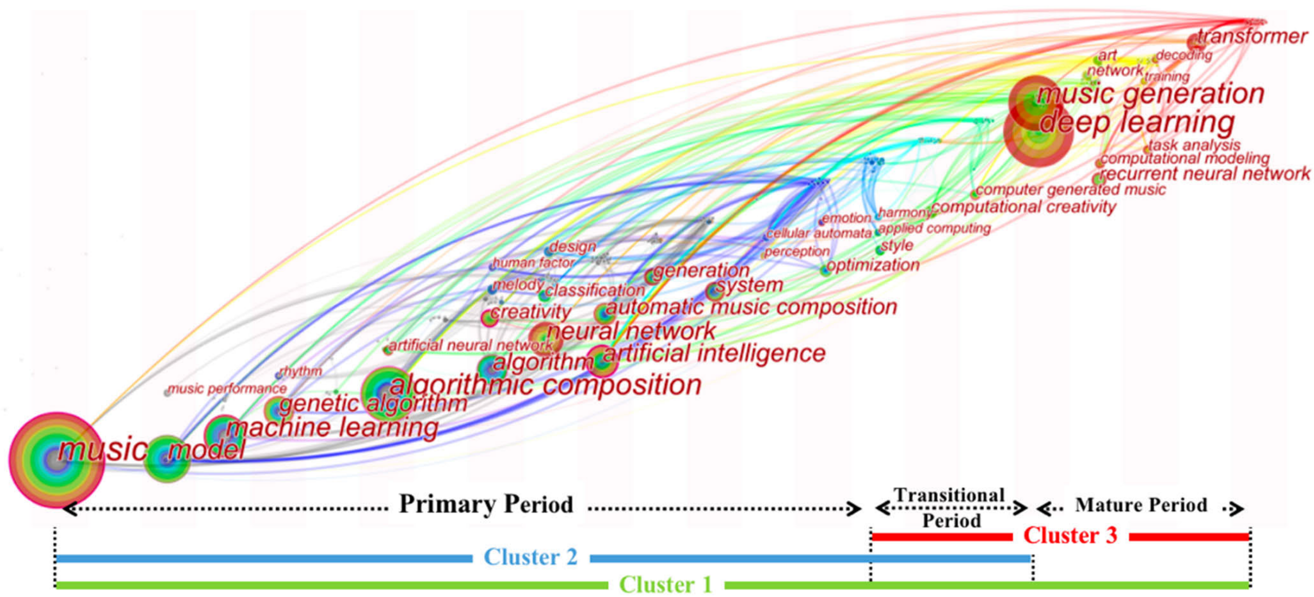

FIGURE 6. The time-zone view of keywords.

- Cluster 3: The red cluster of keywords delves into deep learning neural networks. Here, traditional neural networks remain significant, including Convolutional Neural Network (CNN) [62], Recurrent Neural Network (RNN) [63], Temporal Convolutional Network (TCN) [64], Long Short-Term Memory (LSTM) [65], Generative Adversarial Network (GAN) [66], [67], largely applied in melody and harmony prediction [68], [69], [70]. New neural network models play an essential role in this cluster, such as the Groove2Groove style transfer neural network model [55], the GenoMus artificial music creativity model [71], and the Least Squares SeqGAN (LS-SeqGAN) automatic music generation model [72]. The neural networks studied within this cluster can be broadly categorized as Artificial Neural Networks (ANN) [73], primarily employed for modeling and training [74] in music data programming, tagging, classification, etc. [75].

Further refinement of the content of the above clusters reveals an overview of research hotspots in this field. These three cluster themes demonstrate that while grounded in the theoretical knowledge of music composition, research in this area predominantly leans towards computer science. From the visualization graph, it's evident that the themes of "machine learning algorithms" (Cluster 2) and "deep learning neural networks" (Cluster 3) exist in parallel, both serving as crucial technical supports for the theme of "overall system design" (Cluster 1).

B. TIME-ZONE ANALYSIS OF CLUSTERS

The co-occurrence analysis of keywords serves to illuminate research themes, while their temporal analysis can further elucidate the developmental stages of research topics within the field. In this study, CiteSpace was utilized to generate a temporal map of keywords (Figure 6), facilitating an evolutionary analysis spanning from 2000 to 2023. The background bars in the temporal map denote the years, with the size of keywords indicating their frequency level of occurrence. Lines between keywords represent their co-occurrence relationships. Below the graph, the number of publications is listed by year, while above, corresponding analyses of the distribution of the three cluster themes over the years and the distribution of the three developmental stages in this section are provided.

With the assistance of temporal maps, research topics can be positioned within their corresponding developmental stages. From Figure 6, it's evident that the field can be categorized into three developmental stages: the primary stage, transitional period, and mature stage, based on the distribution of high-frequency keywords across different years. These stages align with the three cluster themes discussed earlier. Cluster 1 (Figure 5) represents a consistent research theme across all stages, manifesting in various synonyms such as algorithmic composition [76], automatic music composition [77], and music generation [78]. Despite variations in keyword presentation over time, the focal point remains on "AI music composition system development." After filtering out keywords related to this constant theme, it's observed that the majority of keywords are technical in nature, with research transitioning from machine learning (Cluster 2 in Figure 5) to deep learning (Cluster 3 in Figure 5).The developmental stages are outlined as follows:

- Primary Period (2001–2015): Machine learning takes center stage during this period, with a particular emphasis on algorithmic research. Initially, research focused on developing auxiliary system tools for music performance [79], gradually shifting towards the creation of virtual performers to replace human performances [80]. Rhythm and melody generation remained continuous topics, with genetic algorithms [81] and probability models [82] serving as representatives, though they are no longer at the forefront of research.

- Transitional Period (2016–2019): The focus gradually transitions from machine learning to deep learning. Artificial neural networks, which are part of machine learning and central to deep learning [83], are prominent during this period, blurring the distinction between the two. Notable keywords during this period include style transfer [55] and optimization models [48].
- Mature Period (2020–2023): Deep learning emerges as the primary focus during this period, with research concentrating on neural network training, including RNN [84], Training[67], Transformer[85], etc. This period also corresponds to the findings in the previous section regarding the current state of research.

C. ANALYSIS OF BURST KEYWORDS

The burst analysis of keywords can refine and pinpoint new trends in research during a certain period in the field, and to some extent, it can also illustrate the evolutionary path of frontier hotspots. Figure 7 displays the top 10 keywords by burst strength, and based on the burst strength and the time series, these burst keywords can be classified into the three periods discussed in the previous section:

FIGURE 7. Top 10 keywords with the strongest citation bursts.

- Primary Period: This period represents the foundational phase of keywords. It can be observed from the relatively low burst strength of "music performance" that during this time, AI composition mainly involved the presentation of electronic music works, falling within the scope of Electronic music and Computer music research [86]. AI technology was in a supportive role for composition and performance.
- Transitional Period: This period signifies a rapid development phase for keywords. Most keywords in the graph appeared after 2016, indicating the rapid development of the field of AI composition post-2016. Analysis of keyword content reveals that it mainly revolves around the transition from machine learning to deep learning technologies, with burst strengths not particularly high, indicating rapid iteration from "machine learning" to "deep learning."
- Mature Period: This period marks a period of significant increase in high-quality burst keywords. The keyword with the highest burst strength is "deep learning," with a burst strength of 6.85, once again confirming its recent research popularity. Other burst keywords in the past three years include "neural network" and "task analysis," further confirming the frontier hotspot themes summarized in the previous section.

D. ANALYSIS OF HIGH-CITED REFERENCES

High-cited references serve as concrete indicators of the research trends within a particular discipline. Table 7 provides an overview of the top 10 most cited references. Notably, "MuseGAN" by Hao-Wen Dong and colleagues [87] emerges as the most frequently cited study. This research exemplifies technical exploration, employing deep learning-based interference, composer, and hybrid models to generate multi-track music within a generative adversarial network framework. Delving into the highly cited articles, it becomes evident that they predominantly explore four research avenues: the "construction of composition models" [70], [87], [88], [89], [90], the "design of system approaches" [78], [91], "music performance rendering" [92], and the "investigation of music generation" [12], [93]. Upon scrutinizing the journals associated with these references, it becomes apparent that conference papers constitute a significant portion, with ACM and AAAI conferences being particularly prominent. This underscores the leading role of computer technology at the forefront of research in this field.

Overall, this study has systematically identified frontiers and hotspots through four steps: cluster analysis, temporal analysis, burst analysis, and analysis of highly cited references. Firstly, high-frequency co-occurring keywords underwent cluster analysis, leading to the identification of three frontier thematic backgrounds. Next, the temporal distribution of keyword themes was examined to position them within three developmental periods. Subsequently, burst words were quantitatively analyzed and categorized into three partitions, further refining the themes and hotspots of each period. Finally, through the analysis of highly cited references, four categories of frontier research directions were outlined. Through this comprehensive approach, the study has clearly delineated the evolutionary trajectory of frontiers and hotspots in this field.

V. COMMENTARY OF DEEP LEARNING COMPOSITION ALGORITHMS

After reviewing the frontiers and hotspots in the field of AI music composition, the focus shifts to the domain of deep learning algorithms. Next, qualitative evaluations and introductions of mainstream algorithm technologies will be conducted based on case studies found in the literature. While numerous algorithms are utilized in AI music generation, they can be categorized into several main classes. The subdivision criteria used in this study are organized based on previous

TABLE 7. Top 10 articles with the highest citations.

Rank	References	Authors	Journals (Year)	Citations
1	MuseGAN: Multi-track Sequential Generative Adversarial Networks for Symbolic Music Generation and Accompaniment	Hao-Wen Dong et al.	Proceedings of the AAAI Conference on Artificial Intelligence(2018)	28
2	DeepBach: a Steerable Model for Bach Chorales Generation	Gaetan Hadjeres et al.	Proceedings of the 34th International Conference on Machine Learning(2017)	20
3	A Hierarchical Latent Vector Model for Learning Long-Term Structure in Music	Adam Roberts et al.	Proceedings of the 35th International Conference on Machine Learning(2018)	12
4	A Functional Taxonomy of Music Generation Systems	Dorien Herremans et al.	ACM Computing Surveys(2017)	11
5	Pop Music Transformer: Beat-based Modeling and Generation of Expressive Pop Piano Compositions	Yu-Siang Huang et al.	Proceedings of the 28th ACM International Conference on Multimedia(2020)	10
6	This time with feeling: learning expressive musical performance	Sageev Oore et al.	Neural Computing and Applications (2020)	9
7	AI Methods in Algorithmic Composition: A Comprehensive Survey	J.D.Fernandez et al.	Journal of Artificial Intelligence Research(2013)	9
8	Automatic melody composition based on a probabilistic model of music style and harmonic rules	Carles Roig et al.	Knowledge-Based Systems(2014)	8
9	Deep learning for music generation: challenges and directions	Jean-Pierre Briot et al.	Neural Computing and Applications(2020)	8
10	Generative adversarial networks	Ian Goodfellow et al.	Communications of the ACM(2020)	8

comprehensive reviews [12], [67], [94] and are divided into three categories: Neural Networks (NNs), Variational Auto-Encoder (VAE), and Transformer.

A. NEURAL NETWORKS

In the field of music generation, Convolutional neural networks (CNN), Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), Generative Adversarial Networks (GAN), and other mainstream artificial neural networks are the core technologies.

1) CNN

CNNs are primarily utilized for processing image data, but in music generation, they can be employed to discern local patterns and features within musical signals. Currently, there are relatively few music generation models based on CNNs, as processing music data with temporal sequence characteristics presents challenges. WaveNet [95], developed by Google DeepMind, stands out as an advanced audio generation model that forecasts the probability distribution of audio data using dilated causal convolutions (Dilated Casual Convolutions) to produce high-fidelity music clips. Despite WaveNet's adeptness in handling long-span temporal dependencies, the music it generates often lacks significant repetitive structures and musicality. Furthermore, WaveNet finds application in other music generation tasks, such as TimbreTron [96], which achieves timbre style transfer in the log-CQT domain by integrating the WaveNet encoder and decoder for audio reconstruction and timbre transformation.

2) RNN AND LSTM

RNNs demonstrate excellent performance in handling time-series data like music segments, but they face challenges with long-term dependencies, often leading to issues like vanishing or exploding gradients. LSTM networks, as an improvement over RNNs, significantly enhance their ability to handle long-term dependencies through unique cell states and gate mechanisms. Both have become the mainstay in music generation systems and are widely applied to generate melodies and polyphonic music. For example, Song From PI [97] employs a hierarchical RNN model to generate melodies, drums, and chords, while Jambot [98] combines chord LSTM and polyphonic LSTM to generate polyphonic music. Other systems like Anticipation-RNN [99], DeepBach [88], and Performance RNN [92] showcase the diversity and effectiveness of RNNs and LSTMs in music generation.

3) GAN

GANs leverage a generator and discriminator to achieve the goal of generating high-quality music through adversarial training. Despite facing challenges in time-series processing, multi-track music generation, and style transfer, researchers have enhanced the quality and diversity of music generation by innovating models such as C-RNN-GAN [100]. This is achieved by effectively utilizing RNNs and conditional generation mechanisms. MuseGAN [87] and BinaryMuseGAN [101], with their improved generator architectures, produce highly correlated multi-track music, while LeadSheetGAN [102] utilizes functional lead sheets as conditional inputs to generate piano rolls with rich information. In terms of style transfer and audio generation, models like CycleGAN [103] and CycleBEGAN [104] achieve effective music style conversion and efficient audio signal synthesis through innovative loss functions and generation strategies. However, GANs encounter challenges in training difficulty and lack interpretability in modeling text data or musical scores, which are issues requiring further resolution in current research.

B. VAE

VAE, as a compression algorithm, has been successfully applied to analyze and generate pitch dynamics and instrument performance information in polyphonic music, primarily addressing music restructuring and prediction problems. Faced with the challenge of processing multimodal data, VAE combined with a hybrid encoding model of Recurrent Neural Networks (RNNs) demonstrates its powerful capabilities in music style transfer and sequence modeling. For instance, MIDI-VAE [105] utilizes triple-stacked encoders and decoders for music style transfer, while Music-VAE [89] employs hierarchical decoders and bidirectional RNN encoders, optimizing the modeling of long-sequence music.

Derived models from VAE show strong potential. MG-VAE [106] focuses on separating style and content latent spaces for generating Eastern folk songs. MIDI-Sandwich2 [107] achieves breakthroughs in polyphonic music simulation and multi-track music generation through conditional VAE and RNN-based multimodal fusion VAE networks. MuseAE [108], by introducing adversarial autoencoders, provides greater flexibility in selecting the prior distribution of latent variables for music generation. These advancements underscore the broad application prospects of VAE and its derivative models in music generation, although further exploration is needed to ensure training stability and a deeper understanding of generation mechanisms.

C. TRANSFORMER

Google's Transformer architecture has revolutionized music generation with its attention mechanism. Compared to traditional CNNs or RNNs, it efficiently handles long-term dependencies, supports data parallelism, and provides self-attention visualization capabilities. While Transformers face challenges with high spatial complexity in music generation, the introduction of Multitrack Music Transformer [109] significantly reduces this complexity, enhancing its suitability for music composition. LakhNES [110] leverages Transformer-XL for transfer learning, generating complex multi-instrument music compositions from a small database.

Transformer-based music generation frameworks promise high-quality improvements. Transformer VAE [111] combines Transformer with VAE to overcome VAE's limitations in handling time-series structures. MusIAC [112], employing multi-level control, enhances scalability and controllability to improve music's structural integrity and expressiveness. The Transformer-based framework MELONS [113] produces high-quality multi-bar melodies using structure generation and melody generation networks, enhancing music style fidelity, melody richness, and sequence generation efficiency. These studies offer new possibilities and tools for music composition, significantly elevating music style fidelity, melody richness, and sequence generation efficiency.

In summary, the three categories of deep learning algorithms used for music generation serve as the technological foundation for the current hotspots in the field. The evolving relationship of current hotspots indicates that "deep learning neural networks" remain an absolute frontier topic, while technologies such as "GAN," "Transformer," and "VAE" [66], [67] continue to evolve in recent years. These discussions on such technological directions supplement the qualitative research on popular themes summarized earlier.

VI. CONCLUSION

AI music composition is part of the emerging interdisciplinary field of computational musicology, which has been evolving rapidly in recent years due to ongoing innovations in computer technology. In this study, we utilized VOSviewer and CiteSpace to conduct a bibliometric analysis of 291 documents in the AI Music Composition field spanning from 2000 to 2023. This analysis encompassed quantitative statistics and qualitative evaluations of 556 authors, 285 organizations, 39 countries, 99 journals, 784 co-occurring keywords, and 7266 co-cited references.

- Development Status (RQ1, Section III): While validating Price's Law, Lotka's Law, and Bradford's Law, this section provides a summary of literature data from four perspectives: authors, organizations, countries, and journals. Presently, the field of AI music composition has established a scientific journal publishing system and a stable network of author collaborations. Authors are predominantly affiliated with academic departments in universities, and there is a noticeable trend towards international collaboration among author teams, with Europe and Asia emerging as the primary contributing regions.
- Frontier Hotspots (RQ2, Section IV): Leveraging the VOSviewer and CiteSpace software, this section systematically identifies frontiers and hotspots through four analytical steps: cluster analysis, temporal analysis, burst analysis, and examination of highly cited references. The evolutionary trajectory is visually depicted through maps. Currently, neural network algorithms and models related to deep learning dominate the mainstream topics.
- Technical Evaluation (RQ3, Section V): The qualitative assessment based on technical directions presented in this study complements the preceding quantitative statistics. Neural Networks (NNs), Variational Auto-Encoder (VAE), and Transformer represent three key types of deep learning algorithms employed in music generation, serving as the technical cornerstone for current hotspots. Integration of various algorithms is ongoing, and both big data models and small data algorithms are avenues for continuous technological advancement.

AI music composition, as an emerging field, possesses significant potential despite certain research gaps. This study, initiated from bibliometrics, systematically reviews and organizes the domain, offering substantial value. Its goal is to aid scholars in delineating research areas, identifying hotspots and frontiers, and tackling technical hurdles. It is anticipated that AI music composition, being a youthful interdisciplinary domain, can craft splendid melodies amidst the ongoing technological evolution.

ACKNOWLEDGMENT

The authors express their deep appreciation to Shandong Xiehe University in China, Shanghai Normal University in China, Kainan University in Taiwan, and Kyonggi University in South Korea, for their invaluable support throughout this project. Special thanks are also extended to Wang Chunmei, Yan Yuchao, and all the dedicated researchers who made substantial contributions to this research.

DISCLOSURE STATEMENT

The authors affirm that the research was carried out without any commercial or financial affiliations that could be interpreted as potential conflicts of interest.

DATA AVAILABILITY STATEMENT

The data generated for this study are available in the article materials. For further information or inquiries, please contact the corresponding author.

REFERENCES

[1] C.-Y. Chang and Y.-P. Chen, "AntsOMG: A framework aiming to automate creativity and intelligent behavior with a showcase on cantus firmus composition and style development," *Electronics*, vol. 9, no. 8, p. 1212, Jul. 2020, doi: 10.3390/electronics9081212.

[2] S. K. Pandey and R. R. Janghel, "Recent deep learning techniques, challenges and its applications for medical healthcare system: A review," *Neural Process. Lett.*, vol. 50, no. 2, pp. 1907–1935, Oct. 2019, doi: 10.1007/s11063-018-09976-2.

[3] Q. Chen, W. Zhao, Q. Wang, and Y. Zhao, "The sustainable development of intangible cultural heritage with AI: Cantonese Opera singing genre classification based on CoGCNet model in China," *Sustainability*, vol. 14, no. 5, p. 2923, Mar. 2022, doi: 10.3390/su14052923.

[4] C. Jin, F. Wu, J. Wang, Y. Liu, Z. Guan, and Z. Han, "MetaMGC: A music generation framework for concerts in metaverse," *EURASIP J. Audio, Speech, Music Process.*, vol. 2022, no. 1, pp. 1–15, Dec. 2022, doi: 10.1186/s13636-022-00261-8.

[5] M. Hamman, "On technology and art: Xenakis at work," *J. New Music Res.*, vol. 33, no. 2, pp. 115–123, Jun. 2004, doi: 10.1080/0929821042000310595.

[6] C. Roads, "Research in music and artificial intelligence," *ACM Comput. Surv.*, vol. 17, no. 2, pp. 163–190, Jun. 1985, doi: 10.1145/4468.4469.

[7] J.-P. Briot, "From artificial neural networks to deep learning for music generation: History, concepts and trends," *Neural Comput. Appl.*, vol. 33, no. 1, pp. 39–65, Jan. 2021, doi: 10.1007/s00521-020-05399-0.

[8] D. Gómez-Marín, S. Jordà, and P. Herrera, "Drum rhythm spaces: From polyphonic similarity to generative maps," *J. New Music Res.*, vol. 49, no. 5, pp. 438–456, Oct. 2020, doi: 10.1080/09298215.2020.1806887.

[9] C.-F. Huang and S.-H. Yao, "Algorithmic composition for pop songs based on lyrics emotion retrieval," *Multimedia Tools Appl.*, vol. 81, no. 9, pp. 12421–12440, Apr. 2022, doi: 10.1007/s11042-022-12408-y.

[10] C.-F. Huang and C.-Y. Huang, "CVAE-GAN emotional AI music system for car driving safety," *Intell. Autom. Soft Comput.*, vol. 32, no. 3, pp. 1939–1953, 2022, doi: 10.32604/iasc.2022.017559.

[11] M. Civit, J. Civit-Masot, F. Cuadrado, and M. J. Escalona, "A systematic review of artificial intelligence-based music generation: Scope, applications, and future trends," *Expert Syst. Appl.*, vol. 209, Dec. 2022, Art. no. 118190, doi: 10.1016/j.eswa.2022.118190.

[12] J. D. Fernandez and F. Vico, "AI methods in algorithmic composition: A comprehensive survey," *J. Artif. Intell. Res.*, vol. 48, pp. 513–582, Nov. 2013, doi: 10.1613/jair.3908.

[13] A. Field, "Mining the ambient commons: Building interdisciplinary connections between environmental knowledge, AI and creative practice research," *Interdiscipl. Sci. Rev.*, vol. 47, no. 2, pp. 185–198, Apr. 2022, doi: 10.1080/03080188.2022.2036408.

[14] O. Ellegaard and J. A. Wallin, "The bibliometric analysis of scholarly production: How great is the impact?" *Scientometrics*, vol. 105, no. 3, pp. 1809–1831, Dec. 2015, doi: 10.1007/s11192-015-1645-z.

[15] R. N. Broadus, "Toward a definition of 'bibliometrics,'" *Scientometrics*, vol. 12, nos. 5–6, pp. 373–379, Nov. 1987, doi: 10.1007/bf02016680.

[16] P. Mayr and A. Scharnhorst, "Scientometrics and information retrieval: Weak-links revitalized," *Scientometrics*, vol. 102, no. 3, pp. 2193–2199, Mar. 2015, doi: 10.1007/s11192-014-1484-3.

[17] G. Abramo, C. A. D'Angelo, and F. Viel, "The field-standardized average impact of national research systems compared to world average: The case of Italy," *Scientometrics*, vol. 88, no. 2, pp. 599–615, Aug. 2011, doi: 10.1007/s11192-011-0406-x.

[18] M. Gutiérrez-Salcedo, M. Á. Martínez, J. A. Moral-Munoz, E. Herrera-Viedma, and M. J. Cobo, "Some bibliometric procedures for analyzing and evaluating research fields," *Int. J. Speech Technol.*, vol. 48, pp. 1275–1287, Dec. 2017, doi: 10.1007/s10489-017-1105-y.

[19] A. Diem and S. C. Wolter, "The use of bibliometrics to measure research performance in education sciences," *Res. Higher Educ.*, vol. 54, no. 1, pp. 86–114, Feb. 2013, doi: 10.1007/s11162-012-9264-5.

[20] *VOSviewer—Visualizing Scientific Landscapes*. Accessed: Aug. 9, 2023. [Online]. Available: https://www.vosviewer.com

[21] CiteSpace. *CiteSpace Home*. Accessed: Aug. 9, 2023. [Online]. Available: https://citespace.podia.com/

[22] X. Ding and Z. Yang, "Knowledge mapping of platform research: A visual analysis using VOSviewer and CiteSpace," *Electron. Commerce Res.*, vol. 22, no. 3, pp. 787–809, Sep. 2022, doi: 10.1007/s10660-020-09410-7.

[23] X. Pan, E. Yan, M. Cui, and W. Hua, "Examining the usage, citation, and diffusion patterns of bibliometric mapping software: A comparative study of three tools," *J. Informetrics*, vol. 12, no. 2, pp. 481–493, May 2018, doi: 10.1016/j.joi.2018.03.005.

[24] N. J. van Eck and L. Waltman, "Software survey: VOSviewer, a computer program for bibliometric mapping," *Scientometrics*, vol. 84, no. 2, pp. 523–538, Aug. 2010, doi: 10.1007/s11192-009-0146-3.

[25] C. Chen, "Searching for intellectual turning points: Progressive knowledge domain visualization," *Proc. Nat. Acad. Sci. USA*, vol. 101, no. 1, pp. 5303–5310, Apr. 2004, doi: 10.1073/pnas.0307513100.

[26] C. Chen, "CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature," *J. Amer. Soc. Inf. Sci. Technol.*, vol. 57, no. 3, pp. 359–377, Feb. 2006, doi: 10.1002/asi.20317.

[27] Microsoft 365. *Microsoft Excel Spreadsheet Software*. Accessed: Aug. 9, 2023. [Online]. Available: https://www.microsoft.com/en-us/microsoft-365/excel

[28] *Official Adobe Photoshop—Photo & Design Software*. Accessed: Aug. 9, 2023. [Online]. Available: https://www.adobe.com/products/

[29] *Scimago Graphica*. Accessed: Aug. 9, 2023. [Online]. Available: https://www.graphica.app/

[30] *RAWGraphs*. Accessed: Aug. 9, 2023. [Online]. Available: https://www.rawgraphs.io/

[31] M. J. Page, J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, L. Shamseer, J. M. Tetzlaff, and D. Moher, "Updating guidance for reporting systematic reviews: Development of the PRISMA 2020 statement," *J. Clin. Epidemiol.*, vol. 134, pp. 103–112, Jun. 2021, doi: 10.1016/j.jclinepi.2021.02.003.

[32] M. J. Page et al., "PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews," *BMJ*, vol. 372, p. n160, Mar. 2021, doi: 10.1136/bmj.n160.

[33] D. Pati and L. N. Lorusso, "How to write a systematic review of the literature," *HERD, Health Environ. Res. Design J.*, vol. 11, no. 1, pp. 15–30, Jan. 2018, doi: 10.1177/1937586717747384.

[34] M. Gaviria-Marin, J. M. Merigó, and H. Baier-Fuentes, "Knowledge management: A global examination based on bibliometric analysis," *Technol. Forecasting Social Change*, vol. 140, pp. 194–220, Mar. 2019, doi: 10.1016/j.techfore.2018.07.006.

[35] C.-H. Liu and C.-K. Ting, "Computational intelligence in music composition: A survey," *IEEE Trans. Emerg. Topics Comput. Intell.*, vol. 1, no. 1, pp. 2–15, Feb. 2017, doi: 10.1109/TETCI.2016.2642200.

[36] I. Choi, "Computation and semiotic practice as compositional process," *Comput. Math. Appl.*, vol. 32, no. 1, pp. 17–35, Jul. 1996, doi: 10.1016/0898-1221(96)00084-3.

[37] I. Santos, L. Castro, N. Rodriguez-Fernandez, Á. Torrente-Patiño, and A. Carballal, "Artificial neural networks and deep learning in the visual arts: A review," *Neural Comput. Appl.*, vol. 33, no. 1, pp. 121–157, Jan. 2021, doi: 10.1007/s00521-020-05565-4.

[38] J. T. Coar and J. P. Sewell, "Zotero: Harnessing the power of a personal bibliographic manager," *Nurse Educator*, vol. 35, no. 5, pp. 205–207, Sep. 2010, doi: 10.1097/nne.0b013e3181ed81e4.

[39] M. M. Carvalho, A. Fleury, and A. P. Lopes, "An overview of the literature on technology roadmapping (TRM): Contributions and trends," *Technol. Forecasting Social Change*, vol. 80, no. 7, pp. 1418–1437, Sep. 2013, doi: 10.1016/j.techfore.2012.11.008.

[40] D. J. D. S. Price, "Little science, big science," in *Little Science, Big Science,*. New York, NY, USA: Columbia Univ. Press, 2019, pp. 59–70, doi: 10.7312/pric91844.

[41] R. C. Coile, "Lotka's frequency distribution of scientific productivity," *J. Amer. Soc. Inf. Sci.*, vol. 28, no. 6, pp. 366–370, Nov. 1977, doi: 10.1002/asi.4630280610.

[42] W. Wei, J. Ge, S. Xu, M. Li, Z. Zhao, X. Li, and J. Zheng, "Knowledge maps of disaster medicine in China based on co-word analysis," *Disaster Med. Public Health Preparedness*, vol. 13, no. 3, pp. 405–409, Jun. 2019, doi: 10.1017/dmp.2018.63.

[43] B. C. Brookes, "'Sources of information on specific subjects' by S. C. Bradford," *J. Inf. Sci.*, vol. 10, no. 4, pp. 173–175, Apr. 1985, doi: 10.1177/016555158501000406.

[44] J. Chen, C. Wang, and Y. Tang, "Knowledge mapping of volunteer motivation: A bibliometric analysis and cross-cultural comparative study," *Frontiers Psychol.*, vol. 13, Jun. 2022, Art. no. 833150, doi: 10.3389/fpsyg.2022.883150.

[45] T. Zeng and F. C. M. Lau, "Automatic melody harmonization via reinforcement learning by exploring structured representations for melody sequences," *Electronics*, vol. 10, no. 20, p. 2469, Oct. 2021, doi: 10.3390/electronics10202469.

[46] R. De Prisco and R. Zaccagnino, "Creative DNA computing: Splicing systems for music composition," *Soft Comput.*, vol. 26, no. 18, pp. 9689–9706, Sep. 2022, doi: 10.1007/s00500-022-06828-z.

[47] R. Zamani, "Combining evolutionary computation with the variable neighbourhood search in creating an artificial music composer," *Connection Sci.*, vol. 31, no. 3, pp. 267–293, Jul. 2019, doi: 10.1080/09540091.2019.1603200.

[48] P. J. P. de León, J. M. Iñesta, J. Calvo-Zaragoza, and D. Rizo, "Data-based melody generation through multi-objective evolutionary computation," *J. Math. Music*, vol. 10, no. 2, pp. 173–192, May 2016, doi: 10.1080/17459737.2016.1188171.

[49] R. Loughran and M. O'Neill, "Evolutionary music: Applying evolutionary computation to the art of creating music," *Genetic Program. Evolvable Mach.*, vol. 21, nos. 1–2, pp. 55–85, Jun. 2020, doi: 10.1007/s10710-020-09380-7.

[50] P.-H. Kuo, T. S. Li, Y.-F. Ho, and C.-J. Lin, "Development of an automatic emotional music accompaniment system by fuzzy logic and adaptive partition evolutionary genetic algorithm," *IEEE Access*, vol. 3, pp. 815–824, 2015, doi: 10.1109/ACCESS.2015.2443985.

[51] K. Verbeurgt, M. Dinolfo, and M. Fayer, "Extracting patterns in music for composition via Markov chains," in *Innovations in Applied Artificial Intelligence*, B. Orchard, C. Yang, and M. Ali, Eds. Berlin, Germany: Springer, 2004, pp. 1123–1132, doi: 10.1007/978-3-540-24677-0_115.

[52] C. Roig, L. J. Tardón, I. Barbancho, and A. M. Barbancho, "A non-homogeneous beat-based harmony Markov model," *Knowl.-Based Syst.*, vol. 142, pp. 85–94, Feb. 2018, doi: 10.1016/j.knosys.2017.11.027.

[53] S. Shi, S. Xi, and S.-B. Tsai, "Research on autoarrangement system of accompaniment chords based on hidden Markov model with machine learning," *Math. Problems Eng.*, vol. 2021, pp. 1–10, Oct. 2021, doi: 10.1155/2021/6551493.

[54] W. Huang, Y. Xue, Z. Xu, G. Peng, and Y. Wu, "Polyphonic music generation generative adversarial network with Markov decision process," *Multimedia Tools Appl.*, vol. 81, no. 21, pp. 29865–29885, Sep. 2022, doi: 10.1007/s11042-022-12925-w.

[55] O. Cífka, U. Simsekli, and G. Richard, "Groove2Groove: One-shot music style transfer with supervision from synthetic data," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 28, pp. 2638–2650, 2020, doi: 10.1109/TASLP.2020.3019642.

[56] S. Chakraborty, S. Dutta, and J. Timoney, "The cyborg philharmonic: Synchronizing interactive musical performances between humans and machines," *Humanities Social Sci. Commun.*, vol. 8, no. 1, pp. 1–9, Mar. 2021, doi: 10.1057/s41599-021-00751-8.

[57] R. Ramirez, E. Gomez, V. Vicente, M. Puiggros, A. Hazan, and E. Maestre, "Modeling expressive music performance in bassoon audio recordings," in *Intelligent Computing in Signal Processing and Pattern Recognition*, D. S. Huang, K. Li, and G. W. Irwin, Eds. Berlin, Germany: Springer, 2006, pp. 951–957, doi: 10.1007/978-3-540-37258-5_119.

[58] R. Ramirez, E. Maestre, and X. Serra, "A rule-based evolutionary approach to music performance modeling," *IEEE Trans. Evol. Comput.*, vol. 16, no. 1, pp. 96–107, Feb. 2012, doi: 10.1109/TEVC.2010.2077299.

[59] F. J. M. Ortega, S. I. Giraldo, A. Perez, and R. Ramírez, "Phrase-level modeling of expression in violin performances," *Frontiers Psychol.*, vol. 10, p. 776, Apr. 2019, doi: 10.3389/fpsyg.2019.00776.

[60] E. R. Miranda and B. N. Siegelwax, "Teaching qubits to sing: Mission impossible?" *Int. J. Unconventional Comput.*, vol. 17, no. 4, pp. 303–331, Aug. 2022, doi: 10.48550/arXiv.2207.08225.

[61] L. Cai and Q. Cai, "Music creation and emotional recognition using neural network analysis," *J. Ambient Intell. Humanized Comput.*, vol. 10, pp. 1–10, Dec. 2019, doi: 10.1007/s12652-019-01614-6.

[62] Y. Singh and A. Biswas, "Lightweight convolutional neural network architecture design for music genre classification using evolutionary stochastic hyperparameter selection," *Expert Syst.*, vol. 40, no. 6, pp. 1–18, Jul. 2023, doi: 10.1111/exsy.13241.

[63] G. C. Sergio and M. Lee, "Scene2Wav: A deep convolutional sequence-to-conditional SampleRNN for emotional scene musicalization," *Multimedia Tools Appl.*, vol. 80, no. 2, pp. 1793–1812, Jan. 2021, doi: 10.1007/s11042-020-09636-5.

[64] F. Jiang, L. Zhang, K. Wang, X. Deng, and W. Yang, "BoYaTCN: Research on music generation of traditional Chinese pentatonic scale based on bidirectional octave your attention temporal convolutional network," *Appl. Sci.*, vol. 12, no. 18, p. 9309, Sep. 2022, doi: 10.3390/app12189309.

[65] S. Ji and X. Yang "Emotion-conditioned melody harmonization with hierarchical variational autoencoder," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Honolulu, Oahu, HI, USA, Oct. 2023, pp. 228–233, doi: 10.1109/SMC53992.2023.10394545.

[66] P. L. T. Neves, J. Fornari, and J. B. Florindo, "Self-attention generative adversarial networks applied to conditional music generation," *Multimedia Tools Appl.*, vol. 81, no. 17, pp. 24419–24430, Jul. 2022, doi: 10.1007/s11042-022-12116-7.

[67] J. Min, Z. Liu, L. Wang, D. Li, M. Zhang, and Y. Huang, "Music generation system for adversarial training based on deep learning," *Processes*, vol. 10, no. 12, p. 2515, Nov. 2022, doi: 10.3390/pr10122515.

[68] J.-F. Paiement, Y. Grandvalet, and S. Bengio, "Predictive models for music," *Connection Sci.*, vol. 21, nos. 2–3, pp. 253–272, Nov. 2010, doi: 10.1080/09540090902733806.

[69] Y. Feng, K. Chen, and X.-B. Liu, "Harmonizing melody with meta-structure of piano accompaniment figure," *J. Comput. Sci. Technol.*, vol. 26, no. 6, pp. 1041–1060, Nov. 2011, doi: 10.1007/s11390-011-1200-1.

[70] C. Roig, L. J. Tardón, I. Barbancho, and A. M. Barbancho, "Automatic melody composition based on a probabilistic model of music style and harmonic rules," *Knowl.-Based Syst.*, vol. 71, pp. 419–434, Nov. 2014, doi: 10.1016/j.knosys.2014.08.018.

[71] J. López-Montes, M. Molina-Solana, and W. Fajardo, "GenoMus: Representing procedural musical structures with an encoded functional grammar optimized for metaprogramming and machine learning," *Appl. Sci.*, vol. 12, no. 16, p. 8322, Aug. 2022, doi: 10.3390/app12168322.

[72] M.-Y. Chen, W. Wei, H.-C. Chao, and Y.-F. Li, "Robotic musicianship based on least squares and sequence generative adversarial networks," *IEEE Sensors J.*, vol. 22, no. 18, pp. 17646–17654, Sep. 2022, doi: 10.1109/JSEN.2021.3066200.

[73] S. Latif, H. Cuayáhuitl, F. Pervez, F. Shamshad, H. S. Ali, and E. Cambria, "A survey on deep reinforcement learning for audio-based applications," *Artif. Intell. Rev.*, vol. 56, no. 3, pp. 2193–2240, Mar. 2023, doi: 10.1007/s10462-022-10224-2.

[74] X. Fu, H. Deng, and J. Hu, "Automatic label calibration for singing annotation using fully convolutional neural network," *IEEJ Trans. Electr. Electron. Eng.*, vol. 18, no. 6, pp. 945–952, May 2023, doi: 10.1002/tee.23804.

[75] P. Lisena, A. Meroño-Peñuela, and R. Troncy, "MIDI2vec: Learning MIDI embeddings for reliable prediction of symbolic music metadata," *Semantic Web*, vol. 13, no. 3, pp. 357–377, Apr. 2022, doi: 10.3233/sw-210446.

[76] Y. Feng, "Advances in algorithmic composition," *J. Softw.*, vol. 17, no. 1, pp. 209–215, 2006, doi: 10.1360/jos170209.

[77] G. Boenn, M. Brain, M. De Vos, and J. Ffitch, "Automatic music composition using answer set programming," *Theory Pract. Log. Program.*, vol. 11, nos. 2–3, pp. 397–427, Mar. 2011, doi: 10.1017/s1471068410000530.

[78] D. Herremans, C.-H. Chuan, and E. Chew, "A functional taxonomy of music generation systems," *ACM Comput. Surv.*, vol. 50, no. 5, pp. 1–30, Sep. 2018, doi: 10.1145/3108242.

[79] E. Chew, J. Liu, and A. R. J. François, "ESP: Roadmaps as constructed interpretations and guides to expressive performance," in *Proc. 1st ACM Workshop Audio Music Comput. Multimedia (AMCMM)*, vol. 6, Oct. 2006, pp. 137–145, doi: 10.1145/1178723.1178744.

[80] A. Botros, J. Smith, and J. Wolfe, "The virtual flute: An advanced fingering guide generated via machine intelligence," *J. New Music Res.*, vol. 35, no. 3, pp. 183–196, Sep. 2006, doi: 10.1080/09298210601045559.

[81] A. Gartland-Jones, "MusicBlox: A real-time algorithmic composition system incorporating a distributed interactive genetic algorithm," in *Applications of Evolutionary Computing* (Lecture Notes in Computer Science), G. Raidl, S. Cagnoni, J. Cardalda, D. Corne, J. Gottlieb, A. Guillot, E. Hart, C. Johnson, E. Marchiori, J. Meyer, and M. Middendorf, Eds. Berlin, Germany: Springer, 2003, pp. 490–501.

[82] J.-F. Paiement, S. Bengio, and D. Eck, "Probabilistic models for melodic prediction," *Artif. Intell.*, vol. 173, no. 14, pp. 1266–1274, Sep. 2009, doi: 10.1016/j.artint.2009.06.001.

[83] A. Bonnici, R. B. Dannenberg, S. Kemper, and K. P. Camilleri, "Editorial: Music and AI," *Frontiers Artif. Intell.*, vol. 4, pp. 1–2, Feb. 2021, doi: 10.3389/frai.2021.651446.

[84] R. Mira, E. Coutinho, E. Parada-Cabaleiro, and B. W. Schuller, "Automated composition of Galician Xota—Tuning RNN-based composers for specific musical styles using deep Q-learning," *PeerJ Comput. Sci.*, vol. 9, p. e1356, May 2023, doi: 10.7717/peerj-cs.1356.

[85] K. Choi, J. Park, W. Heo, S. Jeon, and J. Park, "Chord conditioned melody generation with transformer based decoders," *IEEE Access*, vol. 9, pp. 42071–42080, 2021, doi: 10.1109/ACCESS.2021.3065831.

[86] K. Li, Y. Qian, and W. Zhao, "An auxiliary function approach for lasso in music composition using cellular automata," *J. Appl. Statist.*, vol. 41, no. 5, pp. 989–997, May 2014, doi: 10.1080/02664763.2013.859233.

[87] H.-W. Dong, W.-Y. Hsiao, L.-C. Yang, and Y.-H. Yang, "MuseGAN: Multi-track sequential generative adversarial networks for symbolic music generation and accompaniment," in *Proc. AAAI Conf. Artif. Intell.* Washington, DC, USA: AAAI Press, Apr. 2018, pp. 34–41, doi: 10.1609/aaai.v32i1.11312.

[88] G. Hadjeres, F. Pachet, and F. Nielsen, "DeepBach: A steerable model for Bach chorales generation," in *Proc. 34th Int. Conf. Mach. Learn.*, vol. 70, Cambridge, MA, USA, Jul. 2017, pp. 1362–1371, doi: 10.48550/arXiv.1612.01010.

[89] A. Roberts, J. Engel, C. Raffel, C. Hawthorne, and D. Eck, "A hierarchical latent vector model for learning long-term structure in music," in *Proc. 35th Int. Conf. Mach. Learn.*, vol. 80, Cambridge, MA, USA, Jul. 2018, pp. 4364–4373, doi: 10.48550/arXiv.1803.05428.

[90] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *Commun. ACM*, vol. 63, no. 11, pp. 139–144, Oct. 2020, doi: 10.1145/3422622.

[91] Y.-S. Huang and Y.-H. Yang, "Pop music transformer: Beat-based modeling and generation of expressive pop piano compositions," in *Proc. 28th ACM Int. Conf. Multimedia*. New York, NY, USA: ACM, Oct. 2020, pp. 1180–1188, doi: 10.1145/3394171.3413671.

[92] S. Oore, I. Simon, S. Dieleman, D. Eck, and K. Simonyan, "This time with feeling: Learning expressive musical performance," *Neural Comput. Appl.*, vol. 32, no. 4, pp. 955–967, Feb. 2020, doi: 10.1007/s00521-018-3758-9.

[93] J.-P. Briot and F. Pachet, "Deep learning for music generation: Challenges and directions," *Neural Comput. Appl.*, vol. 32, no. 4, pp. 981–993, Feb. 2020, doi: 10.1007/s00521-018-3813-6.

[94] G. Nierhaus, *Algorithmic Composition: Paradigms of Automated Music Generation*. Vienna, Austria: Springer, Aug. 2009, doi: 10.1007/978-3-211-75540-2.

[95] A. van den Oord, S. Dieleman, H. Zen, K. Simonyan, O. Vinyals, A. Graves, N. Kalchbrenner, A. Senior, and K. Kavukcuoglu, "WaveNet: A generative model for raw audio," 2016, *arXiv:1609.03499*.

[96] S. Huang, Q. Li, C. Anil, X. Bao, S. Oore, and R. B. Grosse, "TimbreTron: A WaveNet(CycleGAN(CQT(Audio))) pipeline for musical timbre transfer," 2018, *arXiv:1811.09620*.

[97] H. Chu, R. Urtasun, and S. Fidler, "Song from PI: A musically plausible network for pop music generation," 2016, *arXiv:1611.03477*.

[98] G. Brunner, Y. Wang, R. Wattenhofer, and J. Wiesendanger, "JamBot: Music theory aware chord based generation of polyphonic music with LSTMs," in *Proc. IEEE 29th Int. Conf. Tools Artif. Intell. (ICTAI)*, Nov. 2017, pp. 519–526, doi: 10.1109/ICTAI.2017.00085.

[99] G. Hadjeres and F. Nielsen, "Interactive music generation with positional constraints using anticipation-RNNs," 2017, *arXiv:1709.06404*.

[100] O. Mogren, "C-RNN-GAN: Continuous recurrent neural networks with adversarial training," 2016, *arXiv:1611.09904*.

[101] H.-W. Dong and Y.-H. Yang, "Convolutional generative adversarial networks with binary neurons for polyphonic music generation," 2018, *arXiv:1804.09399*.

[102] H.-M. Liu and Y.-H. Yang, "Lead sheet generation and arrangement by conditional generative adversarial network," in *Proc. 17th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Dec. 2018, pp. 722–727, doi: 10.1109/ICMLA.2018.00114.

[103] G. Brunner, Y. Wang, R. Wattenhofer, and S. Zhao, "Symbolic music genre transfer with CycleGAN," in *Proc. IEEE 30th Int. Conf. Tools Artif. Intell. (ICTAI)*, Nov. 2018, pp. 786–793, doi: 10.1109/ICTAI.2018.00123.

[104] C.-W. Wu, J.-Y. Liu, Y.-H. Yang, and J.-S. R. Jang, "Singing style transfer using cycle-consistent boundary equilibrium generative adversarial networks," 2018, *arXiv:1807.02254*.

[105] G. Brunner, A. Konrad, Y. Wang, and R. Wattenhofer, "MIDI-VAE: Modeling dynamics and instrumentation of music with applications to style transfer," 2018, *arXiv:1809.07600*.

[106] J. Luo, X. Yang, S. Ji, and J. Li, "MG-VAE: Deep Chinese folk songs generation with specific regional styles," in *Proc. 7th Conf. Sound Music Technol. (CSMT)*, H. Li, S. Li, L. Ma, C. Fang, and Y. Zhu, Eds. Singapore: Springer, 2020, pp. 93–106, doi: 10.1007/978-981-15-2756-2_8.

[107] X. Liang, J. Wu, and J. Cao, "MIDI-sandwich2: RNN-based hierarchical multi-modal fusion generation VAE networks for multi-track symbolic music generation," 2019, *arXiv:1909.03522*.

[108] A. Valenti, A. Carta, and D. Bacciu, "Learning style-aware symbolic music representations by adversarial autoencoders," 2020, *arXiv:2001.05494*.

[109] H.-W. Dong, K. Chen, S. Dubnov, J. McAuley, and T. Berg-Kirkpatrick, "Multitrack music transformer," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Jun. 2023, pp. 1–5, doi: 10.1109/ICASSP49357.2023.10094628.

[110] C. Donahue, H. H. Mao, Y. Ethan Li, G. W. Cottrell, and J. McAuley, "LakhNES: Improving multi-instrumental music generation with cross-domain pre-training," 2019, *arXiv:1907.04868*.

[111] J. Jiang, G. G. Xia, D. B. Carlton, C. N. Anderson, and R. H. Miyakawa, "Transformer VAE: A hierarchical model for structure-aware and interpretable music representation learning," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2020, pp. 516–520, doi: 10.1109/ICASSP40776.2020.9054554.

[112] R. Guo, I. Simpson, C. Kiefer, T. Magnusson, and D. Herremans, "MusIAC: An extensible generative framework for music infilling applications with multi-level control," in *Artificial Intelligence in Music, Sound, Art and Design*, T. Martins, N. Rodríguez-Fernández, and S. M. Rebelo, Eds. Cham, Switzerland: Springer, 2022, pp. 341–356, doi: 10.1007/978-3-031-03789-4_22.

[113] Y. Zou, P. Zou, Y. Zhao, K. Zhang, R. Zhang, and X. Wang, "Melons: Generating melody with long-term structure using transformers and structure graph," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2022, pp. 191–195, doi: 10.1109/ICASSP43922.2022.9747802.

WEIJIA YANG (Member, IEEE) is currently pursuing the Ph.D. degree in philosophy (musicology) with Kyonggi University, South Korea, with a specialization in music technology. He is an Associate Professor with the College of Humanities, Arts and Education, Shandong Xiehe University, China. His research interests include electronic music, AI music composition, and interactive sound installations. He is recognized as a Distinguished Member of the Chinese Association for Artificial Intelligence (CAAI), the International Computer Music Association (ICMA), and the Institute of Electrical and Electronics Engineers (IEEE), with a particular emphasis on the AI and art domain.

LIN SHEN (Member, IEEE) holds the position of an Associate Professor, a Master's Supervisor, and the Director of the Music Technology Department, School of Music, Shanghai Normal University. He also has a significant international presence as a Visiting Scholar with Bangor University, U.K. His research areas are diverse, spanning electronic music composition, AI composition, and music technology. He holds key positions, such as the Executive Vice Chairperson of Shanghai Computer Music Association and the Vice Chairperson of the Sound Working Committee of the China Film Association and the Future Music Committee of the China Musical Instrument Association. He is also a member of the IEEE Computational Intelligence Society.

CHIH-FANG HUANG (Member, IEEE) is a Professor with the Department of Health and Marketing, Kainan University, Taiwan. With an unwavering dedication to the domain of AI sound therapy systems, his expertise, cultivated over two decades, spans AI automated composition, singing voice synthesis, audio-visual processing, music emotional computing, and sound and music psychology. In addition to his academic role, he holds the chairmanship of the Taiwan Computer Music Association (TCMA). As an Active Member of both IEEE and the International Computer Music Conference (ICMC), he contributes as a Paper Session Reviewer and a Music Session Reviewer.

JOHEE LEE received the master's and Ph.D. degrees from Kyonggi University. He is an Assistant Professor with the Department of Performance Music, Kyonggi University, South Korea. His research covers a variety of analysis in composition, including film music and computer music. His research interests include film music, computer music, music analysis, sound design, and music education.

XIAN ZHAO received the Ph.D. degree in sound art from the Sound School of Beijing Film Academy. She is an Associate Professor and a Master's Tutor with the Film-Television and Communication College, Shanghai Normal University. She is also a Visiting Scholar of the Music Department, University of Indianapolis. Her research interests include film sound creation and sound art theoretical research. She is a member of the Film Sound Art Working Committee of the China Film Association, the Film Recording Professional Committee of China Sound Engineers Association, and the American Choral Directors Association; and a Researcher of SHNU Intangible Cultural Heritage Inheritance Research Center.