A Scoping Review of Emerging AI Technologies in Mental Health Care: Towards Personalized Music Therapy

Natália Santos
Faculty of Engineering, University of Porto
up202100737@fe.up.pt

Gilberto Bernardes
INESC TEC, Faculty of Engineering,
University of Porto
gba@fe.up.pt

ABSTRACT

Music therapy has emerged as a promising approach to support various mental health conditions, offering non-pharmacological therapies with evidence of improved well-being. Rapid advancements in artificial intelligence (AI) have recently opened new possibilities for 'personalized' musical interventions in mental health care. This article explores the application of AI in the context of mental health, focusing on the use of machine learning (ML), deep learning (DL), and generative music (GM) to personalize musical interventions. The methodology included a scoping review in the Scopus and PubMed databases, using keywords denoting emerging AI technologies, music-related contexts, and application domains within mental health and well-being. Identified research lines encompass the analysis and generation of emotional patterns in music using ML, DL, and GM techniques to create musical experiences adapted to user needs. The results highlight that these technologies effectively promote emotional and cognitive well-being, enabling personalized interventions that expand mental health therapies.

1. INTRODUCTION

Music as a form of therapy is a historical practice, evidenced by historical records dating back to various ancient cultures, highlighting its healing properties [1]. The relationship between music and mental health has garnered increasing interest in recent decades and has been the subject of extensive research [2]. Music therapy (MT), at the intersection of medicine and music, has emerged in recent years as a promising approach to treating a variety of pathological conditions, including anxiety, depression, substance abuse, Alzheimer's, eating disorders, sleep disorders, Autism Spectrum Disorder, Down Syndrome, among others [3–5]. However, the demands of such mental pathologies and the subjective nature of the musical experience present challenges in formulating universally effective interventions [2]. Advancements in artificial intelligence (AI) have directed attention to the field of machine learning (ML) and, more recently, deep learning (DL), which has been established as a robust and versatile computational tool. Consequently, in recent years, this tool has emerged as a valuable resource in processing audio signals and music [6].

ML is a subset of AI that encompasses systems that learn patterns from data and make predictions or decisions without explicit programming, differing from classical statistical methods that rely on predefined equations and fixed models. Instead, ML algorithms automatically adjust their parameters based on data patterns, enhancing scalability and adaptability [7, 8]. DL, a specialized subfield of ML, expands on these principles by utilizing artificial neural networks with multiple layers (deep neural networks) to model and extract complex features from large datasets [9].

ML and DL applied to music encompasses two main research lines within music information retrieval (MIR) – focusing on extracting information from music data for various applications, such as genre classification, music recommendation, instrument recognition, sound source separation, sung voice detection, emotions recognition, and transcription – and generative music (GM) – related to the automatic creation of musical content [5, 6, 10–12].

The discussion around ML and DL applied to music has raised questions about the potential impact of this tool on MT, as it can be resourceful in detecting emotional responses to sound and music and in generating automatic music adapted to the particularities of each individual [6, 13]. Thus, it can become a complement to mental health professionals and a valuable asset to support therapists and patients, adding value to traditional treatments [5, 6, 14]. This article explores the applications of ML and DL technologies for music processing in mental health by reviewing existing studies across both domains.

The remainder of this paper is organized as follows. Section 2 detailed the methods adopted to retrieve the relevant articles from the literature. Section 3 explores the connection between artificial intelligence and music, emphasizing the use of ML and DL. Section 4 focuses on the relationship between ML, DL, and GM in personalizing music therapy, presenting a new perspective on mental health treatments.

2. METHODOLOGY

The methodology adopted for our scoping review can be divided into three distinct stages. The first stage involved identifying keywords relevant to the search in scientific

DOI: https://doi.org/10.5281/zenodo.15032577

repositories (Section 2.1). The second stage focused on establishing methods for filtering the retrieved articles (Section 2.2). Finally, the third stage analyzed emerging research lines within the scoped domains (Section 2.3).

2.1 Selection and Keywords Refinement

To retrieve relevant literature for the scoping review, we defined three keyword sets based on the primary study topics. First, we considered two main technologies: ('machine learning' OR 'deep learning'). Second, we defined two music-related contexts: ('Music' OR 'Music Therapy'). Third, we established three application domains: ('Therapy' OR 'Mental Health' OR 'Mindfulness'). After retrieving the results, we observed a significant incidence of certain musical contexts, which were not initially included in the keywords set and could hinder important results: 'generative music,' 'automatic music generation,' and 'music recommendation,'emotion-based music recommendation' and 'emotion-integrated music recommendation.'

The final set of threefold keyword sets were combined, resulting in the following logical expression: (deep learning OR machine learning) AND (music OR music therapy OR generative music OR automatic music generation OR music recommendation OR emotion-based music recommendation OR emotion integrated music recommendation) AND (therapy OR mental health OR mindfulness).

2.2 Search and Screening Process

The search was conducted in two leading platforms: PubMed, due to its focus on the health field, and Scopus, chosen for its extensive coverage of reliable citations and well-known databases such as Elsevier, Springer Nature, Wiley-Blackwell, Taylor & Francis, SAGE Publications, Oxford University Press, Cambridge University Press, arXiv, DOAJ, SciELO, IEEE, ACM, European Patent Office (EPO) and United States Patent and Trademark Office (USPTO). The search was performed in October 2024, targeting articles published between January 2000 and October 2024.

Our initial search yielded 169 articles (114 in PubMed and 55 in Scopus). After removing 37 duplicates, 132 articles were considered. In the first screening stage, the articles' titles, abstracts, keywords, and conclusions were manually reviewed, excluding those whose primary focus was not on the application of ML or GM in mental health. As a result, 96 studies were excluded, leaving 36 studies eligible for further analysis.

In the final stage, the full texts of the selected studies were reviewed, from which nineteen articles were additionally excluded due to (n = 13) [do not focus directly on mental health or therapy] and (n = 6) [not having enough data on the ML technologies used]. After these steps, a total of 17 studies were included in the final analysis.

2.3 Organization and Categorization Process

From the final set of articles, we categorized them into two thematic research lines, summarized in Table 1. 1) The correlation between emotions and music, specifically the modulation of listeners' emotional states and music emotion recognition and 2) the intersection of machine learning (ML), generative music (GM), and the personalization of music therapy.

The first line, discussed in Section 3, explores the application of ML and DL models to understanding emotions in music. This theme is examined from two perspectives: 1) the modulation of emotional responses in users by mapping physiological measurements to personalized recommendations, and 2) the automatic recognition and generation of music with controlled emotional content. The second line, covered in Section 4, delves into the interplay between ML, GM, and the personalization of music therapy, emphasizing a novel paradigm for mental health treatment.

3. DEEPENING THE UNDERSTANDING OF EMOTIONS AND MUSIC

Music has long been recognized as a powerful force that can profoundly affect a person's emotions and mental state [6, 22]. However, understanding how music interacts with the human brain and influences mental health is a complex and constantly evolving field [2]. The human brain's response to music is multifaceted and involves a complex network of neural processes. Since the early days of neuroscience, researchers have been interested in musical perception and its implications for cognition and emotion. Music comprises various elements, including harmony, rhythm, melody, timbre, and dynamics. Each of these elements uniquely influences how music is perceived and experienced emotionally. For example, musical rhythms can modulate the autonomic nervous system, affecting heart rate, blood pressure, and even breathing [2, 6].

Recent advances in DL technologies applied to music offer innovative approaches to understanding and analyzing music. Unlike traditional techniques of musical analysis, which often rely on simplified categorizations of musical elements, such as genre or style, DL can extract complex, non-linear patterns from music that may influence a person's emotional response. For instance, DL models can identify subtle variations in tempo, harmonic progressions, and dynamic shifts that are not easily captured by traditional linear statistical methods. These patterns may include syncopation, polyphonic textures, or microtonal variations, which can evoke specific emotional states such as calmness or tension [5, 6, 29]. This capability is further enhanced by the ability of DL to identify emotional features based on large datasets where human listeners have reported emotional responses to musical elements. For example, a DL model could learn associations between minor chord progressions and feelings of sadness or rapid tempo fluctuations with excitement based on datasets such as the EmoMusic Database or the DEAM dataset [16, 27, 30]. What makes DL particularly unique in this context is its capacity to simultaneously process multiple layers of musical information (e.g., melody, harmony, rhythm, and timbre) in a hierarchical manner, uncovering interactions between features that classical statistical approaches might overlook. This multi-layered analysis can be especially valu-

Research Line	Technologies	Tasks	Condition or Study Population	References
Modulating Listeners' Emotions	Machine Learning /	Biophysical measurements	Depression	[15] (EEG)
	Deep Learning		Anxiety	[15] (EEG), [16] (GSR, HR)
			Stress	[17] (EEG), [16] (GSR, HR)
			General Population	[18] (GSR), [19] (EDA, BVP, ST) [15] (EEG)
		Listener's emotional response	Anxiety	[20], [21]
			Depression	[21]
			General Population	[14]
Music Emotion Recognition	Machine Learning /	Listener's emotional response	General Population	[22]
	Deep Learning		Dementia patients	[23]
Personalized Music Therapy	Machine Learning /	Listener's emotional response	ASD patients	[24], [25]
	Deep Learning /		Stress	[26], [27]
	Generative Music		General Population	[5]
		Biophysical measurements	Stress	[28] (HR, GSR)

Table 1. Categorization of the reviewed articles according to three main research lines. The table presents articles grouped into the areas of modulating listeners' emotions through music, music emotion recognition using ML techniques, and personalized music therapy. For each research line, the table lists the technologies used, the tasks analyzed, the studied populations or conditions, and the relevant references associated with each study. Biophysical measurements are indicated when adopted: electroencephalogram (EEG), heart rate (HR), galvanic skin response (GSR), electrodermal activity (EDA), blood volume pulse (BVP), and skin temperature (ST).

able in personalized music therapy, where tailoring music to an individual's emotional state requires a nuanced understanding of how complex musical patterns influence affective responses [5, 6, 29].

3.1 Understanding and Modulating Listeners' Emotional State

Several studies have explored ML combined with biophysical measurements and psychological assessments to create personalized musical systems that induce emotions and promote emotional well-being. These studies use various approaches to map listeners' emotional responses and generate music tailored to evoke desired emotional states, such as mindfulness, calmness, and relief from stress, fear, and anxiety.

Williams et al. [18] developed a real-time generative music system synchronized with the listener's galvanic skin response (GSR), using ML to create emotionally congruent compositions, focusing on emotional regulation and mindfulness promotion. Initially, a musical corpus was generated using Hidden Markov Models (HMM) and labeled with emotional tags based on perceptual questionnaires. Musical features were extracted to train a supervised learning algorithm with listeners' emotional responses. During the experimental phase, participants listened to the compositions generated while their physiological responses were monitored through GSR, followed by self-reported emotional assessments through questionnaires. The analysis revealed a correlation between the GSR readings and the perceived emotions, indicating that the generated music consistently influenced the listeners' emotional states. The system employed biofeedback as a control signal to adjust the generation of new compositions tailored to the listener's emotional state. The findings suggest promising applications in music therapy, mindfulness, and automated soundtrack generation, highlighting the potential of ML-based approaches for personalized emotional interventions.

Building on the work mentioned above, Williams et al. [14] trained a ML system to improve mental health by inducing affective states. This study focused on detecting emotions in music and creating musical sequences based on the probability that specific emotional states occurred in the listener. The results showed that listeners could easily discriminate between emotional states in the musical sequences and that the use of synthesized timbres was effective, indicating that this approach could be beneficial in therapeutic contexts.

Three studies [19–21] have explored the use of convolutional neural networks (CNNs) and ML to analyze listeners' emotional responses and improve the effectiveness of music as a therapeutic tool. Rahman et al. [19] employed CNNs to classify musical genres and associated emotions, achieving high accuracy (99.2% for genre classification and 98.5% for emotion classification). The results emphasize the importance of physiological signals such as electrodermal activity (EDA), blood volume pulse (BVP), and skin temperature (ST) in identifying emotions triggered by music. They suggested the use of genetic algorithms to improve the accuracy of classification.

Wang et al. [20] used DL (CNNs and autoencoders) to process emotional and behavioral data, identifying patterns between music and emotional responses. They proposed personalized music recommendations to improve psychological well-being by reducing anxiety, increasing mental energy, and improving resilience. The results highlighted the potential of DL to integrate empirical data with psychological theories for personalized therapeutic and educational interventions. Joy et al. [21] utilized DL (namely, CNNs) in a mood recognition system, combining facial recognition and musical features (rhythm, melody) to create personalized playlists based on the user's emotional state. They show that CNNs improve the accuracy of music recommendations, contributing to emotional regulation and the treatment of conditions such as depression and anxiety.

Studies [31] [15] and [17] align in their exploration of EEG signals to recognize emotions and offer personalized musical interventions. Sanker et al. [31] proposed a music recommendation system utilizing Russell's valence-arousal model of emotions [32], achieving high accuracy in emotional classification (93.2% for valence and 95.3% for arousal). The research suggested improvements such as the integration of real-time EEG and the development of audio-focused databases. Complementarily, [15] reports advances in affective brain-computer music interfaces (aBCMI), which adjusts real-time music based on EEG signals. Using ML algorithms such as SVM, k-NN, and Random Forest, this system was designed to interpret and respond appropriately to users' emotional states, highlighting therapeutic applications for conditions such as depression and anxiety. Both studies underscore the potential of combining EEG and ML techniques to create personalized music for emotional regulation and well-being, suggesting that real-time integration could further enhance these interventions. In turn, study [17] proposed a system using EEG data and ML to detect stress levels and recommend personalized music to alleviate it. Based on K-Means and SVM, the system classified stress into three levels (low, medium, and high) with 95% accuracy. This approach personalized musical recommendations based on the detected stress level, adjusting song tempos to promote emotional well-being.

Parallel to this, Guo et al. [16] focused on an application called EMO-Music, which utilizes physiological signals from wearable devices, such as heart rate and skin conductance, to identify the user's emotional state and recommend therapeutic music. By employing the DL BERT model for data analysis of these physiological metrics, the system outperformed traditional methods in accuracy, demonstrating that this type of technology can effectively address conditions such as stress and anxiety, enhancing the personalization of therapeutic experiences.

These studies demonstrate the significant potential of ML/DL in creating personalized musical systems to induce specific emotions and promote well-being. Combining biophysical measurements, such as EEG, GSR, BVP, and ST, with ML offers a promising approach to developing more effective and targeted interventions, especially in therapeutic contexts. Using DL models to detect and classify emotions and recommend personalized music is a valuable tool for promoting mental health and emotional regulation.

3.2 From Analysis to Prediction: ML and GM in Musical Emotional Recognition

While some studies applying ML and GM aim to understand music by analyzing the listener's emotional responses, others focus on examining the emotional nature of music itself to predict which emotions it may evoke in listeners. Both approaches aim to use music therapeutically, but they follow opposite processes, creating a complementary field of research.

Emotion-driven music recommendations have gained attention in recent studies, including the work of Panwar et al. [22], which emphasizes the use of social media for enhancing emotional models. By integrating information from platforms like Twitter, the authors sought to enrich the emotional model, offering a more contextualized analysis of listeners' emotions in different regions. The effectiveness of the model, combining CNN and recurrent neural networks (RNN), highlighted not only the ability to identify emotions in music but also suggested future applications, such as inducing specific emotions through music, particularly in a therapeutic context.

Many of the recent advances in MER, such as those presented by this study, can be seen as a continuation of the initial efforts of Panwar et al. [22] with greater sophistication in analysis and therapeutic context. Adopting more advanced technology, such as DL, increased the accuracy of emotional recognition and expanded therapeutic possibilities, providing a better adaptation of musical interventions to an individual's emotional state. Additionally, the study highlights the importance of incorporating contextual and physiological data to enhance the precision of emotion recognition, offering a more robust framework for personalized music interventions, particularly in mental health therapies.

On the other hand, the personalization of music therapy has also become a key objective of other studies applying ML to support the treatment of specific conditions, such as dementia. Nunes et al. [23] use ML to classify music genres and support personalized therapy for dementia patients. The study demonstrated that personalized playlists could improve quality of life and reduce neuropsychiatric symptoms by classifying music based on its association with positive emotional memories, particularly those from the patients' youth. The model's accuracy, exceeding 83%, reinforces the viability of such an approach but also underscores the need for therapist intervention to adjust musical choices based on individual patients' emotional responses. While ML can automate music classification and recommendation, the therapist's presence remains essential to maximize the therapy's effectiveness.

These innovative approaches in ML, DL, and music studies highlight the effectiveness of emotion recognition and recommendation techniques and open doors for using music as a therapeutic tool capable of inducing or regulating emotions precisely and personally. The combination of auditory, physiological, and social data, as observed in the various studies reviewed, offers new possibilities for personalizing music therapy, ensuring that interventions are more aligned with the emotional needs of patients, especially in therapeutic and well-being contexts. The ongoing integration of these approaches, with a focus on personalization and cultural context, reflects the transformative potential of ML in the field of music therapy and music emotion recognition.

4. MACHINE (DEEP) LEARNING, GENERATIVE MUSIC, AND PERSONALIZATION OF MUSIC THERAPY: A NEW PARADIGM IN MENTAL HEALTH

One of the most exciting promises of ML is the ability to personalize the musical experience for each person. Advanced algorithms can analyze a person's emotional and psychological profile, adapting the musical selection. Personalizing the musical experience involves adjusting this selection based on individual preferences, emotional state, and specific goals of each person. In the context of mental health, this approach can be powerful in creating tailor-made playlists aimed at relieving stress, promoting calmness, increasing motivation, or inducing feelings of happiness and well-being [5, 6, 14, 18].

DL algorithms use various techniques and computational models to analyze and understand music. It includes extracting musical features such as harmony, rhythm, melody, timbre, density, and others and identifying complex and subtle patterns that can influence a person's emotional response to music. Combined with an individual's emotional data and psychological profile, algorithms can generate highly personalized music recommendations tailored to their needs and specific preferences [5,6,29]. Studies have been developed to analyze how musical characteristics and parameters can influence individuals' musical experience and how GM can be an ally in creating personalized music. These studies can potentially revolutionize the fields of MT and psychology, offering new approaches to improving people's emotional and cognitive well-being.

4.1 Personalized Music and Mental Health: The Future of Music Therapy

The application of technologies such as ML and DL has proven promising in MT, allowing for greater personalization and effectiveness in therapeutic interventions. Raglio et al. [26] use decision trees to predict the efficacy of music therapy in promoting relaxation. Factors such as age, gender, education, musical training, and frequency of music listening were analyzed, and the model achieved 79% accuracy, identifying key variables such as initial relaxation levels and the combination of education and musical training. Furthermore, the study reinforced the relevance of ML-based methods, rather than DL, for identifying personalized predictive factors that maximize the therapeutic benefits of music.

Zhang et al. [24] focused on treating children with Autism Spectrum Disorder (ASD) using DL-based interactive robots to improve musical perception and social interaction. The research employed gesture recognition and EEGs to assess the children's musical responses, achieving an average accuracy of 85%. These results highlight the role of music as a powerful therapeutic tool for promoting social skills and exploring the potential of technology in teaching and treating neurodevelopmental conditions. Santos et al. [25] introduced an GM application to capture the musical preferences of children with ASD. By allowing the children to manipulate musical parameters such as pitch, tempo, timbre, intensity, density, and percussiveness to create personalized music. The study aimed to understand which musical elements were most effective for each individual. Although the technological approach differs from the interactive robots, both studies share the goal of personalizing music interventions. While the robot study focuses on social interaction and musical perception through gestures and physical responses, the generative app allows for a more individualized and exploratory musical creation, offering a way for the participants themselves to shape their therapeutic experience. These two studies, though distinct in the technology employed, converge on the importance of personalization and the child's active involvement in the therapeutic process.

Modran et al. [5] investigated how emotional and musical characteristics can predict the therapeutic effects of music through a multi-class neural network. Using a subset of the Million Dataset and techniques such as K-Fold cross validation, the model categorized emotions into four distinct groups, achieving an accuracy of over 94%. Furthermore, integrating attention mechanisms and contextual data contributed to a more refined recognition of music's emotional and therapeutic effects. They highlight gaps in previous research and propose an innovative solution to personalize music interventions according to each individual's emotional and musical profile, representing a significant advancement for music therapists and patients seeking personalized therapies.

Regarding stress reduction, two recent studies adopt DL to optimize the personalization of therapeutic music. Choi et al. [27] explored CNNs and Mel-scaled spectrograms to build optimized music datasets for relaxation. The model achieved an impressive accuracy of 98.7%, demonstrating that the music selected by the system was as effective as options validated by experts in reducing stress and promoting emotional well-being. This approach reduces the costs associated with the biological validation of playlists and expands the possibilities of creating more accessible and effective music interventions. Zhang [28] integrated emerging technologies such as the Internet of Things (IoT), big data, and mixed-density neural networks (MDNN) to monitor and reduce stress in university students. Wearable and environmental sensors collected real-time data – such as heart rate, skin conductivity, volume, and environmental temperature – and were processed through ML to enable immediate adjustments in therapeutic music to alleviate stress as detected. The approach emphasized the relevance of playlist personalization and demonstrated how combining music and technology can offer practical and dynamic solutions to enhance mental well-being.

The combined advancements from these studies strengthen the growing intersection between technology, music, and mental health. Integrating ML, DL, and emerging tools such as IoT and big data expands the capacity to create personalized interventions, maximizing the therapeutic effects of music. This evidence points to a promising future where music and technology unite to transform mental health approaches, promoting greater access and effectiveness for diverse populations and needs.

5. DISCUSSION AND CONCLUSIONS

This reviewed the intersection between MT and emerging AI technologies, focusing on ML, DL, and GM. It highlights current research lines on applying these technologies to transform the therapeutic experience, enabling highly personalized and effective interventions in the mental health field.

From this scoping review, two major lines of research emerged. The first examines musical emotions from a twofold perspective: modulation of listeners' emotional states in response to music, exploring the interplay between emotional, cognitive, and musical aspects, and application of ML, DL, and GM technologies to recognize and predict emotional patterns in musical contexts. The second research line highlights the potential of personalized music for mental health interventions by adapting musical contents to individual needs and enabling more effective therapeutic experiences. While ML and DL have contributed to a deeper understanding of emotional and cognitive interactions with music, GM has made it possible to create musical content tailored to the specific needs of each individual.

The studies analyzed show notable advances in identifying emotional patterns in music and the personalization of music recommendations based on biological, psychological, and contextual data. The evidence demonstrates the effectiveness of these technologies in promoting emotional and cognitive well-being and opens up new possibilities for the application of music in the treatment of a wide range of mental health conditions, ranging from emotional and cognitive disorders to more complex psychological disorders.

Although the results are encouraging, ethical, technical, and operational challenges must be addressed. The identified challenges point to the need for greater standardization in research protocols, given the use of divergent methods and metrics between studies, making direct comparisons difficult [6, 16–18, 31]. In addition, the limitation of small, demographically specific samples, such as autistic children [24, 25] or university students [28], compromises the generalizability of the results. Cultural gaps in understanding musical preferences and their emotional implications also highlight the importance of more inclusive and diverse approaches [23]. Another significant obstacle lies in oversimplifying human emotional responses through limited classifications, underestimating the complexity and dynamics of emotional states [5, 22].

From an ethical and technical point of view, the challenges related to the privacy and security of biometric data still require greater attention, especially in systems that use sensitive information, such as EEG and cardiac variation, to personalize therapies. The lack of validation in real clinical contexts also limits the applicability of technological solutions, which often remain confined to controlled environments or prototypes [18,26]. Added to this is the lack of robust exploration into the cumulative and long-term impacts of GM and the reliance on subjective tools, such as questionnaires, to assess therapeutic effects [14, 15].

Furthermore, something concerning in the studies analyzed in the context of emotion detection is the tendency to rely predominantly on objective measures such as EEG, GSR, BVP, and ST. While these biophysical signals can provide valuable insights, they risk neglecting the subjective experience of the individual. Emotional responses are inherently personal, and subjective assessments can offer critical perspectives on how emotions are felt and experienced. Over-reliance on objective data could lead to an exaggerated perception of the effects of music on mental health. Therefore, a balanced approach that integrates both subjective evaluations and objective measures is essential to capture a more holistic understanding of emotional states and their modulation through music therapy.

Emerging and underexplored areas offer ways to overcome these limitations and expand the therapeutic benefits of GM. Real-time dynamic adaptation, supported by immersive technologies such as virtual and augmented reality, is among the most promising fronts [16, 18, 24, 31]. Likewise, integrating multimodal stimuli – such as sonic, visual, and tactile – can further enrich therapeutic interventions. Future research should also explore the impact of underrepresented musical genres [24, 26], develop tools that simplify GM for therapists without technical training, and consider cultural and ethnographic nuances in individualizing interventions [29].

Other critical areas of research involve the use of advanced DL models to create music tailored to individual emotional states and psychological traits. Among these models are GANs, which use a generator and a discriminator to produce realistic, personalized music [33]. Probabilistic diffusion models, which start with random noise and refine the sound until detailed results are obtained, are also promising for music generation [34]. Transformer-based networks, capture long-term dependencies in musical sequences, generating more cohesive compositions. In addition, quantized variational autoencoders (VQ-VAE) create discretized representations of the music, facilitating the generation of new pieces with diversity and complexity [35]. StyleGAN, originally used for images [36], can be adapted to control stylistic aspects of music, such as rhythm and timbre, while multimodal approaches, such as CLIP, make it possible to generate music from emotional or psychological descriptions. The application of AI in longer therapeutic processes, with the ability to monitor and progressively adapt interventions based on emotional progress, also emerges as a relevant field that has yet to be investigated.

6. REFERENCES

[1] M. H. Thaut, "Music as therapy in early history," *Progress in brain research*, vol. 217, pp. 143–158, 2015.

[2] D. L. Bowling, "Biological principles for music and mental health," *Translational Psychiatry*, vol. 13, no. 1, 2023, 374.

[3] J. H. Wakim, S. Smith, and C. Guinn, "The efficacy of music therapy," *Journal of perianesthesia nursing*, vol. 25, no. 4, pp. 226–232, 2010.

[4] L. Gassner, M. Geretsegger, and J. Mayer-Ferbas, "Effectiveness of music therapy for autism spectrum disorder, dementia, depression, insomnia and schizophrenia: update of systematic reviews," *European journal of public health*, vol. 32, no. 1, pp. 27–34, 2022.

[5] H. A. Modran, T. Chamunorwa, D. Ursuțiu, C. Samoilă, and H. Hedeșiu, "Using deep learning to recognize therapeutic effects of music based on emotions," *Sensors*, vol. 23, no. 2, 2023, 986.

[6] J. Liang, "Harmonizing minds and machines: survey on transformative power of machine learning in music," *Frontiers in Neurorobotics*, vol. 17, 2023, 1267561.

[7] N. Kühl, M. Goutier, R. Hirt, and G. Satzger, "Machine learning in artificial intelligence: Towards a common understanding," *arXiv preprint arXiv:2004.04686*, 2020.

[8] G. M. d. M. Paixão, B. C. Santos, R. M. d. Araujo, M. H. Ribeiro, J. L. d. Moraes, and A. L. Ribeiro, "Machine learning na medicina: Revisão e aplicabilidade," *Arquivos brasileiros de cardiologia*, vol. 118, pp. 95–102, 2022.

[9] J.-P. Briot, G. Hadjeres, and F.-D. Pachet, "Deep learning techniques for music generation–a survey," *arXiv preprint arXiv:1709.01620*, 2017.

[10] F.-F. Kuo, M.-F. Chiang, M.-K. Shan, and S.-Y. Lee, "Emotion-based music recommendation by association discovery from film music," in *Proceedings of the 13th annual ACM international conference on Multimedia*, 2005, pp. 507–510.

[11] J. Bevington and D. Knox, "Cognitive factors in generative music systems," in *Proceedings of the 9th Audio Mostly: A Conference on Interaction With Sound*, 2014, pp. 1–8.

[12] X. Wang and Y. Wang, "Improving content-based and hybrid music recommendation using deep learning," in *Proceedings of the 22nd ACM international conference on Multimedia*, 2014, pp. 627–636.

[13] M. Bhaumik, P. U. Attah, and F. Javed, "Emotion integrated music recommendation system using generative adversarial networks," *SMU Data Science Review*, vol. 5, no. 3, 2021, 4.

[14] D. Williams, V. J. Hodge, and C.-Y. Wu, "On the use of ai for generation of functional music to improve mental health," *Frontiers in Artificial Intelligence*, vol. 3, 2020, 497864.

[15] H. Glassman, D. Dwyer, N. John, D. Laesker, and M. So, "Affective brain-computer music interface in emotion regulation and neurofeedback: A research protocol," *Undergraduate Research in Natural and Clinical Science and Technology Journal*, vol. 6, pp. 1–9, 2022.

[16] H. Guo, J. Zhang, Y. Jiang, Y. Qi, S. Chen, Z. Chen, W. Lin, J. Cao, and S. hou Li, "Emo-music: Emotion recognition based music therapy with deep learning on physiological signals," in *2024 IEEE First International Conference on Artificial Intelligence for Medicine, Health and Care (AIMHC)*. IEEE, 2024, pp. 10–13.

[17] J. Patil, N. M. Ranjan, P. Dange, A. Patil, D. Salunke, and S. Borde, "Personalized stress mitigation through eeg based stress classification and music recommendation," in *International Conference on Computing and Machine Learning*. Springer, 2024, pp. 181–191.

[18] D. Williams, V. J. Hodge, L. Gega, D. Murphy, P. I. Cowling, and A. Drachen, "Ai and automatic music generation for mindfulness," in *2019 AES International Conference on Immersive and Interactive Audio: Creating the Next Dimension of Sound Experience*. York, 2019.

[19] J. S. Rahman, T. Gedeon, S. Caldwell, R. Jones, and Z. Jin, "Towards effective music therapy for mental health care using machine learning tools: human affective reasoning and music genres," *Journal of Artificial Intelligence and Soft Computing Research*, vol. 11, no. 1, pp. 5–20, 2021.

[20] T. Wang, Y. Zhao, and M. Yin, "Analysis and research on the influence of music on students' mental health under the background of deep learning," *Frontiers in Psychology*, vol. 13, 2022, 998451.

[21] R. Joy, M. Thanka, J. Dhas, and E. Edwin, "Music mood based recognition system based on machine learning and deep learning," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 11, no. 2, pp. 904–911, 2023.

[22] S. Panwar, P. Rad, K.-K. R. Choo, and M. Roopaei, "Are you emotional or depressed? learning about your emotional state from your music using machine learning," *The Journal of Supercomputing*, vol. 75, pp. 2986–3009, 2019.

[23] I. B. Nunes, M. A. de Santana, N. Charron, H. S. e. Silva, C. M. de Lima Simões, C. Lins, A. B. de Souza Sampaio, A. M. N. de Melo, T. C. V. da Silva, C. Tiodista *et al.*, "Automatic identification of preferred music genres: an exploratory machine learning approach to support personalized music therapy," *Multimedia Tools and Applications*, pp. 1–17, 2024.

[24] Y. Zhang, C. Zhang, L. Cheng, and M. Qi, "The use of deep learning-based gesture interactive robot in the treatment of autistic children under music perception education," *Frontiers in Psychology*, vol. 13, 2022, 762701.

[25] N. Santos, G. Bernardes, R. Cotta, N. Coelho, and A. Baganha, "Assessing musical preferences of children on the autistic spectrum: Implications for therapy," *Sound and Music Computing*, Jul 2024.

[26] A. Raglio, M. Imbriani, C. Imbriani, P. Baiardi, S. Manzoni, M. Gianotti, M. Castelli, L. Vanneschi, F. Vico, and L. Manzoni, "Machine learning techniques to predict the effectiveness of music therapy: A randomized controlled trial," *Computer Methods and Programs in Biomedicine*, vol. 185, 2020, 105160.

[27] S. Choi, J.-I. Park, C.-H. Hong, S.-G. Park, and S.-C. Park, "Accelerated construction of stress relief music datasets using cnn and the mel-scaled spectrogram," *PLoS ONE*, vol. 19, no. 5, 2024, 0300607.

[28] J. Zhang, "Iot-enabled musical therapy to alleviate physiological stress in college students using big data and mixed-density neural networks," *Mobile Networks and Applications*, pp. 1–21, 2024.

[29] L. Moysis, L. A. Iliadis, S. P. Sotiroudis, A. D. Boursianis, M. S. Papadopoulou, K.-I. D. Kokkinidis, C. Volos, P. Sarigiannidis, S. Nikolaidis, and S. K. Goudos, "Music deep learning: deep learning methods for music signal processing—a review of the state-of-the-art," *IEEE Access*, vol. 11, pp. 17 031–17 052, 2023.

[30] S. Shelke and M. Patil, "Exploring machine learning techniques for music emotion classification: A comprehensive review," in *2024 11th International Conference on Computing for Sustainable Global Development (INDIACom)*. IEEE, 2024, pp. 1188–1195.

[31] S. V. Sanker, N. B. R. S. Bilakanti, A. Thomas, V. P. Gopi, and P. Palanisamy, "Emotion-recognition-based music therapy system using electroencephalography signals," in *Edge-of-Things in Personalized Healthcare Support Systems*. Elsevier, 2022, pp. 217–235.

[32] J. A. Russell, "A circumplex model of affect." *Journal of personality and social psychology*, vol. 39, no. 6, pp. 1161–1178, 1980.

[33] J. Engel, K. K. Agrawal, S. Chen, I. Gulrajani, C. Donahue, and A. Roberts, "Gansynth: Adversarial neural audio synthesis," 2019. [Online]. Available: https://arxiv.org/abs/1902.08710

[34] M. Plasser, S. Peter, and G. Widmer, "Discrete diffusion probabilistic models for symbolic music generation," 2023. [Online]. Available: https://arxiv.org/abs/2305.09489

[35] S. Han, H. Ihm, and W. Lim, "Symbolic music loop generation with VQ-VAE," *CoRR*, vol. abs/2111.07657, 2021. [Online]. Available: https://arxiv.org/abs/2111.07657

[36] D. Jeong, S. Doh, and T. Kwon, "Träumerai: Dreaming music with stylegan," 2021. [Online]. Available: https://arxiv.org/abs/2102.04680