

Artificial Intelligence-Based Music Therapy Technology and Applications

Ruoting Liang

*North University of China, Taiyuan, China
2488839062@qq.com*

Abstract. Due to its non-pharmacological nature and low cost, music therapy has gained increasing attention in the fields of medicine and mental health and has been applied in the treatment of various diseases. However, traditional music therapy is limited by the lack of standardized treatment duration and frequency, as well as its inability to meet individual patient needs, hindering further development and broader application. This paper reviews relevant studies, highlighting the therapeutic effects of current music therapy methods on various diseases, and explores the application and efficacy of artificial intelligence (AI) technologies such as Random Forest (RF), Long Short-Term Memory networks (LSTM), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN) in supporting music therapy. The results show that music therapy can help improve patients' negative emotions and alleviate symptoms, while AI technologies can enhance and optimize traditional music therapy in areas such as music selection and treatment plan formulation, leading to more personalized and effective outcomes. In conclusion, music therapy demonstrates potential in clinical applications, and integrating AI technologies may offer more individualized and effective treatment plans. Future research and practice in this area should be further strengthened.

Keywords: Music therapy, clinical application of music therapy, neural networks, machine learning, deep learning

1. Introduction

Music therapy, as an emerging therapeutic approach, has received widespread attention in the fields of healthcare and mental health. With increasing emphasis on physical and mental well-being, music therapy has gradually become an important adjunctive treatment for various diseases due to its unique advantages of being non-pharmacological and cost-effective. However, the development of music therapy still faces certain challenges: there is no unified standard for treatment duration and frequency in traditional music therapy, making it difficult to meet the personalized needs of different patients; moreover, the selection of treatment plans lacks consideration of individual patient differences, resulting in difficulty in determining the optimal therapeutic approach. These issues limit the further development and wider application of music therapy.

Music therapy is a treatment method that utilizes musical interactions such as listening, singing, and playing instruments, aiming to treat diseases and improve patients' physical and mental

conditions through the interaction between music and human physiology and psychology. It has demonstrated significant effects in the treatment of various conditions. For example, in autism treatment, it can promote patients' social-emotional interactions and communication; for patients with anxiety disorders, it effectively alleviates physical symptoms and psychological distress; in depression therapy, it can reduce negative emotions such as depression and anxiety, thereby improving mental health levels. Although music therapy has been proven to have certain clinical effects and is widely used, it still suffers from insufficient personalization, meaning it cannot provide symptom-targeted treatment based on individual patient conditions. Artificial intelligence technologies, such as neural networks, can learn patients' comfort preferences, retrieve appropriate music from libraries, and generate more personalized music. Therefore, the integration of artificial intelligence into music therapy in clinical treatment holds significant importance.

This paper aims to review the application of artificial intelligence in auxiliary music therapy, with a particular focus on Random Forest (RF), Long Short-Term Memory network (LSTM), Convolutional Neural Network (CNN), and Recurrent Neural Network (RNN) (see Figure 1). It explores how AI technologies can help address the issues existing in traditional music therapy, providing new ideas and methods for the development of music therapy, promoting further advancement in this field, and offering more effective treatment options for a greater number of patients.

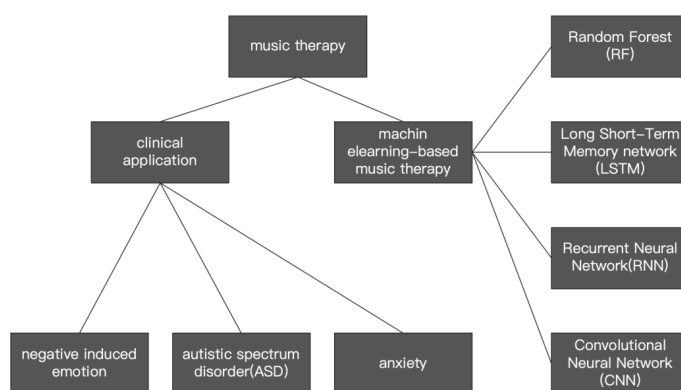


Figure 1: AI-based music therapy technologies

2. Music therapy and its clinical applications

Existing music therapy typically involves therapists establishing a connection between patients and musical rhythms, enabling patients to move in sync with the rhythm and interact with others. This approach aims to improve patients' negative emotions and even alleviate symptoms of autism. Music therapy has been widely applied in the treatment of autism and anxiety disorders. Due to its non-invasive nature, low cost, and favorable clinical feedback, it has gained acceptance among patients and their families.

Conventional treatment for depression generally includes pharmacotherapy and psychotherapy; however, medication is not always effective for some patients. As a non-pharmacological therapy, music therapy has demonstrated positive clinical effects. Ali Behzad et al. [1] conducted a randomized controlled trial to investigate the clinical efficacy of TaKeTiNa music therapy. The intervention group received 14 sessions (totaling 28 hours) of TaKeTiNa music therapy over eight weeks, while the control group underwent standard treatment during the same period. The study

assessed depression scores, blood parameters (such as LDL and cholesterol), immune cell factors, hormone levels, and heart rate variability to demonstrate that this music therapy could reduce the severity of depression, improve relevant physiological indicators, and increase heart rate variability. Additionally, psychological scales were used to evaluate patients' mental states before and after treatment, confirming that TaKeTiNa music therapy effectively improved psychological conditions, alleviated negative emotions such as depression and anxiety, enhanced mental health, and helped patients better cope with the psychological stress caused by the illness.

Children with autism spectrum disorder (ASD) often exhibit a strong preference for music, and music therapy can improve their social interaction abilities. Thomas Rabeyron et al. [2] conducted a controlled experiment to investigate whether music therapy is more effective than simply listening to music for children with ASD. The experimental group received music therapy, which included instrument playing, improvisation, and interaction with others, while the control group only listened to music without interaction. Apart from this, the treatment duration, frequency, and selected music were the same for both groups. Assessments were conducted before and after treatment using the Childhood Autism Rating Scale (CARS), Aberrant Behavior Checklist (ABC), and Clinical Global Impression scale (CGI). The results demonstrated that music therapy was superior to passive music listening in reducing somnolence and stereotyped behaviors in children with ASD, with better overall clinical impressions, supporting its use as a treatment approach for ASD. In the study by A. C. Jaschke et al. [3], children with autism were randomly assigned to two groups: the intervention group received twice-weekly, 30-minute music interaction sessions over twelve weeks, employing various improvisational techniques to stimulate social communication behaviors; the control group received standard care only, which included routine support from general practitioners, mental health, and educational professionals. Evaluations of social interaction, communication skills, and emotional expression were conducted at baseline, week 13, and week 39 using the Brief Observation of Social Communication Change (BOSCC) scale among other measures. The experiment confirmed that music therapy met the expectations of patients and their families regarding treatment methods and outcomes, and they expressed approval of this creative, non-invasive, and non-pharmacological approach.

Daniel Johnston et al. [4] indicated that individuals with autism experience developmental limitations in socio-emotional interaction and communication, and that music therapy has a positive impact on these areas. Emerging technologies such as motion sensor devices, mobile technologies, and multisensory surfaces offer new avenues for musical interaction, thereby promoting the development of motor skills and sensory cognition in individuals with autism.

Moreover, music therapy has also been shown to have certain therapeutic effects on anxiety disorders. S. Boussaid et al. [5] found that patients with chronic inflammatory rheumatic diseases often experience high anxiety levels during intravenous biologic infusions. In their study, the experimental group listened to 30 minutes of soft music during infusion, while the control group received only the infusion without music. Measures taken before and after infusion included pain (Visual Analog Scale, VAS), anxiety (State-Trait Anxiety Inventory, STAI), and vital signs (blood pressure, heart rate, respiratory rate). Results showed that, during infusion, the experimental group's STAI scores, systolic blood pressure, and heart rate were significantly lower than those of the control group, whereas diastolic blood pressure and respiratory rate showed no significant differences. This suggests that music therapy can effectively reduce anxiety, systolic blood pressure, and heart rate in patients with chronic inflammatory rheumatic diseases during biologic infusion, thereby enhancing patient comfort and providing a more humane treatment experience. Catherine Colin et al. [6], through a review and synthesis of multiple studies, pointed out that music therapy

improves physical symptoms and psychological distress, particularly stress-related parameters. In surveys on alleviating healthcare worker stress, music interventions significantly reduced stress, anxiety, psychological burden, risk of burnout, and improved psychosomatic symptoms. Furthermore, Mozart's music was shown to be especially effective in reducing anxiety and psychological burden.

3. Artificial intelligence-based music therapy

Traditional music therapy often lacks consideration of individual differences among patients, making it difficult to tailor treatment to each person's unique condition [7]. With the development of artificial intelligence (AI), its powerful capabilities in data analysis and learning have led to widespread applications across many fields. Meanwhile, electroencephalography (EEG) enables real-time, non-invasive monitoring of brain activity, facilitating advances in integrating EEG signals to recognize emotions induced by music. Classic machine learning algorithms, such as k-Nearest Neighbors (KNN) and Random Forest (RF), have demonstrated certain accuracy in recognizing musical emotions and require relatively low computational resources. However, machine learning models often suffer from limited generalization ability and face challenges in feature extraction, stability, and accuracy. With the advancement of deep learning algorithms, models such as Long Short-Term Memory networks (LSTM), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN) have gradually been applied to EEG-based recognition of music-induced emotions. These approaches provide new techniques for researching music therapy and establish new paradigms for medical fields like neuroscience [8].

3.1. Random forest

Listening to, singing, or playing music has been shown to improve quality of life and well-being among the elderly, with beneficial effects on memory, attention, expression, and anxiety. As a result, music therapy is widely employed in treating dementia. However, there remains a lack of effective tools to assist music therapists in selecting appropriate music for patients. Ingrid Bruno Nunes et al. [9] aimed to develop a recommendation system capable of automatically classifying music by genre using machine learning algorithms. This system would help music therapists more precisely choose suitable music for dementia patients, thereby enhancing the effectiveness and personalization of music therapy. Participants listened to music from the Emotify public music database, which contains 100 tracks in each of four musical genres, and categorized the music while also annotating the emotions they experienced during listening. These annotations served as ground truth for evaluating model accuracy. After preprocessing the music data, features were extracted using GNU/Octave and formatted into ARFF files as inputs for Random Forest (RF), Support Vector Machine (SVM), J48 Decision Tree, and Bayesian Network (BN) classifiers. The classifiers' performance was assessed using five metrics: accuracy, kappa statistic, sensitivity, specificity, and area under the ROC curve (AUC-ROC). Comprehensive evaluation revealed that the Random Forest algorithm performed best: with tree numbers ranging from 150 to 350, it achieved an accuracy of approximately 83%, a kappa statistic of 0.78, sensitivity of 0.96, specificity of 0.94, and AUC-ROC of 0.99. The experiment demonstrated that RF could generate playlists beneficial for patients, while requiring low memory usage and computational resources, thereby improving its practical feasibility.

3.2. Long short-term memory network

JIEMEI CHEN et al. [10] found that traditional synthetic music therapy for tinnitus patients suffers from a limited duration problem; repeated playback may trigger negative emotions in patients, which is detrimental to stress relief and neglects individual patient preferences. The Long Short-Term Memory network (LSTM) can retain historical sequential information and generate music that meets patient needs by learning pitch sequence features from existing music, thereby avoiding the repetitiveness inherent in traditional synthetic music. In their experiment, 30 tinnitus patients and 10 volunteers without tinnitus were recruited as the experimental and control groups, respectively, for a comparative study. Original music was segmented based on dominant melody pitch sequences, from which pitch sequences of each segment were extracted and the main pitch sequences identified. Two LSTM neural networks were designed and trained separately to predict the initial pitch and pitch trend of subsequent music segments. Based on these predictions, music segment labels were determined, and specific music pieces were generated. The experiment demonstrated that LSTM can accurately predict the pitch trajectory of subsequent segments based on input pitch sequence data. This enables the generated music to maintain the original musical characteristics while avoiding issues such as abrupt pitch changes and prolonged sustained notes. The music produced by this model is non-repetitive and self-similar, which helps alleviate patients' anxiety. Additionally, the absence of sudden pitch changes provides tinnitus patients with a more comfortable auditory experience and more positive emotional responses. Zixi Wang et al. [11] utilized an optimized LSTM network model to analyze music therapy data and predict treatment outcomes, providing strong support for treatment plan formulation. Their study involved bereaved family members from a city in inland China, randomly assigned to an experimental group and a control group, each comprising 240 participants. The experimental group received music therapy based on the optimized LSTM algorithm, while the control group underwent traditional music therapy. The optimized LSTM network model was constructed using dynamic time warping techniques to handle variable-length sequence data. Convolutional Neural Networks (CNNs) were employed to extract audio features, which were then integrated with psychological state data. Model parameters were tuned using random search, grid search, cross-validation, and Bayesian optimization. Additionally, L2 regularization and Dropout techniques were introduced to prevent overfitting. Treatment effects were evaluated using the Symptom Checklist-90 (SCL-90) scale. The results demonstrated that the LSTM model could learn complex patterns and long-term dependencies in the music therapy process, thereby accurately predicting treatment outcomes and providing valuable guidance for treatment plan design and adjustment.

3.3. Convolutional neural network

Traditional regression methods struggle to effectively handle the high-dimensional and nonlinear relationship between acoustic features and rhythm. The development of Convolutional Neural Networks (CNNs) has provided a more efficient approach for music selection. Rahman et al. [12] innovatively applied deep learning to process biological responses by using CNNs to assist in selecting stress relief music (Stress Relief Music, SM). Although their approach achieved over 95% test accuracy, detection required specialized equipment. Abboud et al. extracted features directly from music using the fuzzy k-Nearest Neighbors (Fuzzy k-NN) algorithm. However, k-NN's computational performance is highly influenced by the training sample size and the choice of k, making it less efficient for large datasets, especially with high-dimensional data. The advantage of CNN models lies in their ability to perform inference without the need for extensive data storage

and with higher classification efficiency for SM. Suvin Choi et al. [13] utilized CNNs to convert musical sound elements—such as frequency, amplitude, and waveform (Elements of Sound, ES)—into Mel-scaled Spectrograms (MSS), which served as key features. These sound elements act as indicators of a song’s potential to relieve stress. MSS effectively simulates human auditory sensitivity by capturing frequency and amplitude variations in audio signals. Through convolutional layers, pooling layers, and fully connected layers, CNNs automatically learn features from MSS to classify stress relief music. Their clinical study adopted a 2×2 crossover design. Participants were randomly divided into two sequence groups: Group A first experienced individual music (Individual Music, IM), followed by a washout period and then researcher-selected music (Researcher-selected Music, RM); Group B experienced the reverse sequence. Stress, happiness, and satisfaction were assessed using the Visual Analog Scale (VAS) before and after the first treatment phase and after the second phase. Results indicated that RM’s effect on emotional improvement was not inferior to IM, confirming that the CNN model could classify SM with a 98.7% test accuracy. This provides effective support for music classification and selection, advancing personalized music therapy. Jiangang Chen et al. [14] refined their dataset by removing MIDI drum loops and incomplete or unavailable tracks from the Janata 2012 study appendix and processed the remaining data using the “pydub” library. They ultimately selected 258 non-overlapping 25-second segments. The rhythms were quantified on a 128-point scale. Using Python, model parameters were optimized and trained over 1000 epochs, with performance evaluated via 5-fold cross-validation.

The results showed that this model outperformed direct CNN regression by 8.6%. Compared to models combining CNN with LSTM or Support Vector Machines (SVM) for regression tasks, it exceeded performance by 11.0% and 26.9%, respectively. Furthermore, the model required less training time, demonstrating greater accuracy and efficiency in selecting music tailored to individual patient needs.

3.4. Recurrent neural network

Existing music therapy technologies face challenges such as imprecise targeting (e.g., lack of pathological specificity and neglect of individual differences), low efficiency, limited professionalism, and privacy risks [15]. The application of Recurrent Neural Networks (RNNs) to optimize music therapy can enhance the ability to address these issues. RNNs are effective at mining temporal and semantic information from sequential data. The learning process involves forward propagation of signals and backward propagation of error signals, enabling the model output to progressively approach the expected values. Yiyao Zhang et al. [15] selected children aged 3 to 12 as subjects, ultimately focusing on 10 autistic children as the primary research group and healthy children as a control group. An online questionnaire was designed to collect information on the children’s interest in music and their preferred music types. Based on each child’s preferences, rhythmically distinct and suitably timed music pieces were selected, while avoiding extreme tonal variations and dissonant sounds. Neuroscan EEG was used to collect and analyze EEG signals from both groups during relaxation, music perception therapy, and pre-rest phases, with 10 samples each. The collected data were split into training and testing sets at a 5:1 ratio and used for recognition and analysis by the RNN algorithm. The classification accuracy of the RNN model on these signals reflects the accuracy of music perception. The experiment showed that the average music perception classification accuracy was 85% for autistic children and 84% for healthy children. These results outperformed Long Short-Term Memory (LSTM) networks, Convolutional Neural Networks (CNN), and Support Vector Machine (SVM) algorithms. Therefore, RNNs are well suited to detect the psychological state of autistic children and provide valuable references for their treatment.

4. Conclusion

This paper systematically reviews music therapy and its clinical applications, as well as the integration of artificial intelligence technologies such as machine learning and neural networks into this field. Music therapy has achieved certain therapeutic effects in clinical practice, improving negative emotions, alleviating symptoms, and reducing stress and anxiety levels in patients suffering from autism, anxiety disorders, depression, chronic inflammatory rheumatism, and relieving stress among healthcare workers. Artificial intelligence technologies, including machine learning and neural networks, have enhanced and optimized music therapy by assisting music selection, predicting treatment outcomes, and monitoring patients' psychological states. These advancements have made music therapy more personalized and precise, thereby improving therapeutic effectiveness and feasibility. Future research should further strengthen the combination of music therapy and AI technologies, establish standardized and personalized treatment protocols, explore long-term effects, and expand applications across more disease domains to promote the widespread development and clinical application of music therapy.

References

- [1] Behzad, A. , Feldmann-Schulz, C. , Lenz, B. , Clarkson, L. , Ludwig, C. , & Luttenberger, K. , et al. (2024). Taketina music therapy for outpatient treatment of depression: study protocol for a randomized clinical trial. *Journal of Clinical Medicine*, 13(9), 13.
- [2] Jaschke, A. C. , Howlin, C. , Pool, J. , Greenberg, Y. D. , Atkinson, R. , & Kovalova, A. , et al. (2024). Study protocol of a randomized control trial on the effectiveness of improvisational music therapy for autistic children. *BMC Psychiatry*, 24(1).
- [3] A, T. R. , D, J. P. R. D. C. A. , B, E. C. , C, V. B. , C, N. B. , & B, F. X. V. , et al. A randomized controlled trial of 25 sessions comparing music therapy and music listening for children with autism spectrum disorder. *Psychiatry Research*, 293.
- [4] Johnston, D. , Egermann, H. , Kearney, G. , & Särkämö, Teppo. (2018). Innovative computer technology in music based interventions for individuals with autism - moving beyond traditional interactive music therapy techniques. *Cogent Psychology*, 5(1), 1-18.
- [5] Boussaid, S. , Majdouba, M. B. , Jriri, S. , Abbes, M. , Jammali, S. , & Ajlani, H. , et al. (2020). Fri0618-hpreffects of music therapy on pain, anxiety, and vital signs in chronic inflammatory rheumatic diseases patients during biological drugs infusion. *Annals of the Rheumatic Diseases*, 79(Sup1), 1.
- [6] Colin, C., Prince, V., Bensoussan, J. L., & Picot, M. C. (2023). Music therapy for health workers to reduce stress, mental workload and anxiety: a systematic review. *Journal of Public Health*, 45(3), e532-e541.
- [7] Rajendran, T. (2022). Addressing the need for personalizing music therapy in integrative oncology. *Journal of Integrative Medicine*, 20(4), 281-283.
- [8] Su, Y., Liu, Y., Xiao, Y., Ma, J., & Li, D. (2024). A review of artificial intelligence methods enabled music-evoked EEG emotion recognition and their applications. *Frontiers in Neuroscience*, 18, 1400444.
- [9] Nunes, I. B., de Santana, M. A., Charron, N., Silva, H. S. E., de Lima Simões, C. M., Lins, C., ... & dos Santos, W. P. (2024). Automatic identification of preferred music genres: an exploratory machine learning approach to support personalized music therapy. *Multimedia Tools and Applications*, 83(35), 82515-82531.
- [10] Chen, J., Pan, F., Zhong, P., He, T., Qi, L., Lu, J., ... & Zheng, Y. (2020). An automatic method to develop music with music segment and long short term memory for tinnitus music therapy. *IEEE Access*, 8, 141860-141871.
- [11] Wang, Z., Guan, X., Li, E., & Dong, B. (2024). A study on music therapy aimed at psychological trauma recovery for bereaved families driven by artificial intelligence. *Frontiers in Psychology*, 15, 1436324.
- [12] Rahman, J. S., Gedeon, T., Caldwell, S., Jones, R., & Jin, Z. (2021). 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*, 11(1), 5-20.
- [13] Choi, S., Park, J. I., Hong, C. H., Park, S. G., & Park, S. C. (2024). Accelerated construction of stress relief music datasets using CNN and the Mel-scaled spectrogram. *PloS one*, 19(5), e0300607.
- [14] Chen, J., Han, J., Su, P., & Zhou, G. (2025). Framework for Groove Rating in Exercise-Enhancing Music Based on a CNN–TCN Architecture with Integrated Entropy Regularization and Pooling. *Entropy*, 27(3), 317.

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