Closed-Loop Emotion-Regulation Through Dynamic Music Generation: An AI Framework for Stress and Mood Adaptation

Shilpa Bhandari[1], Dr. Soumya[2]
[1]*Research Scholer, ICIS, Srinivas University*
doi.org/10.64643/IJIRTV12I12-206820-459

***Abstract*—The growing levels of stress and mental health challenges in modern society have increased the need for intelligent and personalized emotional regulation tools. This paper presents a closed-loop artificial intelligence framework that dynamically generates music to help regulate users' emotional states in real time. The proposed system combines emotion recognition techniques with generative music models to create personalized soundscapes that support stress reduction and mood stabilization. The framework continuously monitors emotional signals using physiological and behavioral data obtained from wearable sensors and user interactions. Based on this information, the system adapts musical attributes such as tempo, harmony, rhythm, and timbre to influence the listener's emotional state. Machine learning models analyze biometric indicators including heart rate variability, facial expressions, and vocal characteristics to estimate emotional conditions, which are widely used in affective computing systems [2], [7].**

A generative music engine then produces adaptive musical compositions tailored to the detected mood using deep learning–based music generation models [3], [16]. In addition, the architecture incorporates reinforcement learning to refine music generation strategies based on user responses, allowing the system to gradually improve its effectiveness over time [10]. Experimental evaluation indicates that dynamically generated AI music provides better relaxation and mood stabilization compared to conventional static playlists. The proposed framework offers a scalable and personalized solution for digital mental wellbeing applications such as therapeutic support systems, meditation platforms, and stress-management tools.

***Index Terms*—Emotion Recognition, Generative AI, Music Generation, Reinforcement Learning, Stress Detection, Human-AI Interaction, Adaptive Music Systems.**

I. INTRODUCTION

Mental health and emotional wellbeing have become increasingly important global concerns. Factors such as demanding lifestyles, constant digital engagement, and workplace pressure have contributed to rising levels of stress and emotional imbalance. Music therapy has long been recognized as an effective method for managing emotions, reducing stress, and improving mood [1], [8]. However, traditional music therapy typically relies on preselected playlists and human supervision, which limits its ability to adapt to individual emotional changes in real time.

Recent advances in artificial intelligence, affective computing, and generative models have opened new possibilities for building intelligent emotional support systems. AI-based music generation models can create dynamic musical compositions that adapt to user preferences and emotional contexts [3], [16]. When combined with emotion recognition technologies, these systems can produce personalized sound environments that respond directly to a user's emotional state [7], [15].

Closed-loop systems offer a promising solution for achieving this type of adaptive emotional regulation. In a closed-loop system, emotional signals are continuously monitored, analyzed, and used to modify the system's output. In the context of music generation, this involves adjusting musical characteristics such as tempo, melody, harmony, and sound textures to guide listeners toward relaxation or emotional balance [13].

The framework proposed in this study integrates emotion recognition, generative music models, and reinforcement learning within a unified closed-loop architecture. By continuously analyzing physiological signals such as heart rate variability, facial expressions, and vocal patterns, the system estimates the user's emotional condition and generates music designed to reduce stress or promote positive emotions. Reinforcement learning techniques further enable the system to improve its performance through continuous feedback and adaptation [10]. Such intelligent systems have potential applications in digital mental health platforms, workplace wellbeing programs, personalized meditation tools, and therapeutic support environments.

II. OBJECTIVES

The primary objectives of this research are as follows:

- To design a closed-loop AI framework capable of generating music dynamically based on real-time emotional feedback from users.
- To implement multimodal emotion recognition techniques using physiological signals, facial expression analysis, and voice-based emotion detection.
- To develop a generative music engine that adjusts musical elements such as tempo, melody, rhythm, and harmony to influence emotional states.
- To incorporate reinforcement learning mechanisms that improve music generation strategies by learning from user emotional responses.
- To evaluate the effectiveness of the proposed system in reducing stress and improving emotional balance through controlled experimental studies.

III. METHODOLOGY

This research introduces an integrated artificial intelligence framework that combines emotion sensing, generative music modeling, and adaptive feedback mechanisms to regulate emotional states dynamically. The system operates through several interconnected modules that together form a continuous feedback loop between emotion detection and music generation.

System Architecture

The proposed architecture consists of four major components:

• Emotion Detection Module
• Music Generation Engine
• Reinforcement Learning Feedback System
• User Interaction Interface

These components work together to enable real-time emotional monitoring and adaptive music generation.

Emotion Detection Module

The emotion detection module identifies the user's current emotional state using data collected from multiple sources. These include physiological signals, facial expressions, and vocal characteristics.

Physiological signals such as heart rate variability, electrodermal activity, and respiration patterns can be captured through wearable devices. Facial expressions are analyzed using computer vision techniques that identify emotional cues from facial movements. Similarly, voice-based emotion recognition analyzes acoustic properties such as pitch, tone, and speech rhythm.

Machine learning models, including convolutional neural networks and recurrent neural networks, process these signals to classify emotional states into categories such as stress, calmness, happiness, or sadness.

Because emotional states can change rapidly, the system continuously updates these predictions to track emotional variations over time.

Dynamic Music Generation Engine

The music generation engine creates adaptive musical compositions using generative artificial intelligence models. Instead of selecting songs from a predefined library, the system dynamically generates music by adjusting parameters known to influence emotional perception.

These parameters include:

• Tempo and rhythm
• Harmonic progression
• Melody structure
• Instrumentation and timbre
• Sound intensity and dynamics

Generative models such as transformer-based music models or recurrent neural networks produce music sequences in real time.

Emotional states are mapped to specific musical characteristics based on findings from music psychology research.

For example:

• When stress levels are high, the system generates slower tempos, softer harmonies, and ambient sound textures.
• When a low mood is detected, the system produces uplifting melodies and brighter harmonic progressions.
• When the user is calm, the system maintains relaxation through gentle rhythms and soothing soundscapes.

Through these adjustments, the system attempts to guide emotional states toward balance and relaxation.

Reinforcement Learning Feedback Loop

To continuously improve the effectiveness of the generated music, the framework incorporates reinforcement learning. In this approach, the system evaluates how the user's emotional state changes after listening to generated music.

The process operates as follows:

First, the system generates music based on the detected emotional condition.

Next, physiological sensors monitor the user's emotional response during playback.

If stress indicators decrease or positive emotional signals increase, the system receives a positive reward signal.

The reinforcement learning agent then updates its policy to generate more effective musical patterns in future sessions.

Over time, this learning process allows the system to personalize music generation strategies for each individual user.

User Interface and Interaction

The system also includes an interactive interface that enables users to interact with the platform easily. Through the interface, users can view emotional indicators, start relaxation sessions, and provide feedback on the generated music.

Users can also adjust preferences such as music style or genre, allowing the system to incorporate personal taste into its adaptive generation process.

This interaction layer improves user engagement and helps the system learn more accurate personalization patterns.

IV. ANALYSIS AND OBSERVATIONS

To evaluate the proposed framework, experimental studies were conducted in which participants performed stress-inducing tasks followed by relaxation sessions using the adaptive AI music system. Emotional responses were monitored using physiological signals as well as self-reported mood assessments.

The results revealed several important observations.

First, the closed-loop system successfully adapted music based on changing emotional conditions. When stress indicators increased, the system gradually shifted toward calmer musical characteristics, including slower tempos and softer harmonic structures.

Second, the reinforcement learning mechanism improved performance across multiple sessions. As the system learned users' physiological responses and music preferences, the generated music became more effective at reducing stress.

Third, participants reported greater emotional comfort and satisfaction when listening to dynamically generated music compared to traditional static playlists. The personalized and adaptive nature of the music contributed to a more immersive relaxation experience.

Despite these positive outcomes, some limitations were observed. Emotion recognition accuracy depended heavily on the quality of sensors and environmental conditions. Additionally, differences in individual musical preferences highlighted the need for stronger personalization mechanisms.

Overall, the results suggest that AI-generated adaptive music has strong potential as a tool for emotional regulation and stress management.

V. CONCLUSION

This study presents a novel artificial intelligence framework for closed-loop emotional regulation through dynamic music generation. By integrating emotion recognition techniques, generative music models, and reinforcement learning mechanisms, the proposed system offers a personalized and adaptive approach to managing stress and emotional wellbeing. Unlike traditional music recommendation systems that rely on fixed playlists, the proposed framework continuously monitors user emotions and dynamically modifies musical output to guide emotional transitions toward relaxation and positive mood states.

Experimental findings demonstrate the potential of adaptive AI-generated music for various mental wellbeing applications, including digital therapy systems, meditation platforms, and workplace stress-management tools.

Future research may explore the integration of brain–computer interfaces, advanced emotion recognition models, and large-scale clinical validation to further enhance the effectiveness of adaptive music-based emotional regulation systems.

REFERENCES

[1] P. N. Juslin and J. Sloboda, Handbook of Music and Emotion: Theory, Research, Applications, Oxford University Press, 2010.

[2] R. Picard, Affective Computing, MIT Press, 1997.

[3] A. Huang et al., “Music transformer: Generating music with long-term structure,” arXiv preprint arXiv:1809.04281, 2018.

[4] I. Goodfellow et al., “Generative adversarial nets,” Advances in Neural Information Processing Systems (NeurIPS), 2014.

[5] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, pp. 436–444, 2015.

[6] D. Herremans, C. Chuan, and E. Chew, “A functional taxonomy of music generation systems,” ACM Computing Surveys, 2017.

[7] M. Soleymani et al., “A survey of multimodal sentiment analysis,” Image and Vision Computing, 2017.

[8] S. Koelsch, “Brain correlates of music-evoked emotions,” Nature Reviews Neuroscience, 2014.

[9] T. Nakano and M. Goto, “Emotion recognition from music signals,” IEEE Transactions on Audio, Speech and Language Processing, 2011.

[10] R. Sutton and A. Barto, Reinforcement Learning: An Introduction, MIT Press, 2018.

[11] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning, MIT Press, Cambridge, MA, 2016.

[12] J. A. Russell, “A circumplex model of affect,” Journal of Personality and Social Psychology, vol. 39, no. 6, pp. 1161–1178, 1980.

[13] P. N. Juslin and D. Västfjäll, “Emotional responses to music: The need to consider underlying mechanisms,” Behavioral and Brain Sciences, vol. 31, no. 5, pp. 559–575, 2008.

[14] S. Koelsch, “Toward a neural basis of music-evoked emotions,” Trends in Cognitive Sciences, vol. 14, no. 3, pp. 131–137, 2010.

[15] E. Gómez, M. Haro, and J. Herrera, “Music emotion recognition: A survey,” IEEE Transactions on Affective Computing, vol. 5, no. 3, pp. 293–307, 2014.

[16] D. Herremans, C. Chuan, and E. Chew, “A functional taxonomy of music generation systems,” ACM Computing Surveys, vol. 50, no. 5, pp. 1–30, 2017.

[17] H. Dong, W. Hsiao, L. Yang, and Y. Yang, “MuseGAN: Multi-track sequential generative adversarial networks for symbolic music generation,” in Proc. AAAI Conference on Artificial Intelligence, 2018.

[18] A. Vaswani et al., “Attention is all you need,” in Advances in Neural Information Processing Systems (NeurIPS), pp. 5998–6008, 2017.

[19] J. Huang et al., “Music Transformer: Generating music with long-term structure,” arXiv preprint arXiv:1809.04281, 2018.

[20] M. Soleymani, S. Asghari-Esfeden, Y. Fu, and M. Pantic, “Analysis of EEG signals and facial expressions for continuous emotion detection,” IEEE Transactions on Affective Computing, vol. 7, no. 1, pp. 17–28, 2016.