Journal of Science, Computing and Engineering Research (JSCER)
Volume-9, Issue- 4, April 2026.
DOI: *https://doi.org/10.46379/jscer.2023.060503*
ISSN: 2708-1079

EMOTION BASED MUSIC GENERATOR

[1]Mr. Mohit Sharma,[2]Nitin Khandelwal,[3]Pradeep Kumar,[4] Mohit Jajoriya,[5]Deepanshu Jajoria

[1]Asst. Professor, Department of CSE, Modern Institute of Technology and Research Centre, Rajasthan, India.

[2,3,4,5]UG Student, Department of CSE, Modern Institute of Technology and Research Centre, Rajasthan, India

Article Information

Received : 30 March 2026
Revised : 31 March 2026
Accepted : 01 March 2026
Published : 02 April 2026

Corresponding Author:
Nitin Khandelwal

Abstract— This project presents an advanced Emotion-Based Music Generation system that leverages Machine Learning and Deep Learning to create personalized musical compositions aligned with human emotional states. Unlike traditional recommendation systems that only suggest pre-existing tracks, this system dynamically generates original music by mapping emotional cues to musical attributes. The platform integrates multi-modal emotion recognition—utilizing facial expression analysis, speech sentiment extraction, and text-based emotion inference—to create a robust understanding of the user's affective state. Facial images are processed using Convolutional Neural Networks (CNNs), whereas vocal audio signals are analyzed through MFCC feature extraction combined with Recurrent Neural Networks (RNNs) or Bi-LSTM layers for temporal modeling. Textual input undergoes natural language processing using Transformer-based architectures such as BERT to derive contextual emotional representations. Based on the predicted emotion category—such as happiness, sadness, calmness, anger, fear, or excitement—the system triggers the music generation module, which uses deep generative models

Keywords—Emotion-Based Music System, Affective Computing, Deep Learning, Convolutional Neural Networks (CNN), Facial Expression Recognition (FER), Human-Computer Intraction.

Citation: Mr. Mohit Sharma, Nitin Khandelwal, Pradeep Kumar, Mohit Jajoriya, Deepanshu Jajoria, "EMOTION BASED MUSIC GENERATOR", Journal of Science, Computing and Engineering Research, 9(04), April 2026.

I. *INTRODUCTION*

In today's era of artificial intelligence and intelligent computing, machine-generated creativity has emerged as a powerful tool for transforming traditional multimedia systems into personalized and adaptive solutions. One such rapidly evolving area is emotion-based music generation, which plays a vital role in modern human–computer interaction. Traditional music recommendation methods rely heavily on predefined playlists, manual selection, or simple metadata such as genre or artist.

These approaches often fail to capture the user's emotional state, resulting in mismatched music choices and limited personalization. To address these limitations, emotion-based music generation introduces an advanced AI-driven mechanism capable of understanding human emotions and automatically generating or selecting music that aligns with the user's current mood.

This system utilizes affective computing, machine learning, and deep learning models to analyze input signals—such as facial expressions, speech tone, physiological signals, or text sentiment—to detect the user's emotional state accurately.

By processing real-time data through neural networks, the system predicts emotions like happiness, sadness, calmness, fear, excitement, or anger. These detected emotions serve as inputs to a generative music engine, which composes or selects music tailored to the user's emotional context. This enables a seamless, adaptive, and personalized musical experience that requires no manual intervention from the user.

The primary objective of this project is to design a secure, reliable, and user-friendly AI system that interprets emotions in real time and produces music capable of enhancing mood, supporting mental well-being, and enriching user engagement. The integration of emotional analysis with generative deep-learning models not only improves

II. *PROBLEM STATEMENT*

Music plays a central role in emotional expression, stress reduction, and personal well-being. However, conventional music recommendation systems largely depend on static playlists, search-based inputs, or user-specified preferences such as genres, artists, or song titles. Although these systems offer automation to an extent, they are unable to adapt to the user's real-time emotional state.

This leads to several limitations including mismatched mood-music alignment, reduced user satisfaction, and lack

Users often spend significant time manually selecting songs, switching tracks repeatedly, or creating mood-based playlists. This manual process is inefficient, subjective, and inconsistent—especially for users seeking emotional comfort, mental relaxation, focus enhancement, or therapeutic support. Similarly, existing recommendation algorithms lack psychological sensitivity and do not incorporate multimodal emotional cues such as facial expressions, voice tone, or text sentiment. They also fail to ensure real-time adaptability as the user's emotions shift. The primary goal of the Emotion-Based Music Generator system is to accurately identify human emotions using advanced machine learning, deep learning, and natural language processing (NLP) techniques, and to generate personalized music compositions that dynamically adapt to the user's emotional state.

III. *PROPOSED METHOD*

The Emotion-Based Music Generator System is developed using a combination of modern technologies that enable efficient data processing, emotion detection, and intelligent music recommendation.

A. System Initialization

The process begins by initializing the required hardware and software environment for the Emotion-Based Music Generator system. A webcam or HD camera is connected to the computer to capture real-time facial expressions of the user. The system then loads and configures essential software components, including emotion detection and music processing modules.

During this stage, the system initializes:

- Camera drivers for real-time video capture
- Face detection and emotion recognition libraries (such as OpenCV, deep learning models, etc.)
- Pre-trained emotion classification models
- Music database categorized by different emotions

This setup phase ensures that the system is fully prepared for capturing user input, detecting emotions accurately, and generating appropriate music recommendations in real time.

B. Image Acquisition (Face Capture)

When a user stands in front of the camera, the system captures real-time images or video frames to analyze facial expressions. The captured input is converted into an appropriate format (typically grayscale or normalized RGB) to ensure efficient processing. Using libraries such as OpenCV, the system continuously captures clear facial data under different lighting conditions.

C. Image Preprocessing

The goal here is to convert text chunks into vector representations using pre-trained models like OpenAI's embedding models. These embeddings enable semantic search by representing document chunks in a way suitable for comparison.

- **Face Detection:** Identifying the facial region using algorithms such as Cascade or HOG (Histogram of Oriented Gradients).
- **Cropping and Resizing:** Extracting the face from the background and scaling it to a standard size.
- **Noise Reduction:** Applying filters to remove distortions and improve image quality.
- **Feature Normalization:** Adjusting brightness, contrast, and intensity for uniform analysis.

D. Feature Extraction (Emotion Features)

- Eye movement and eyebrow position
- Mouth shape and lip curvature
- Facial muscle tension
- Distance between key facial landmarks

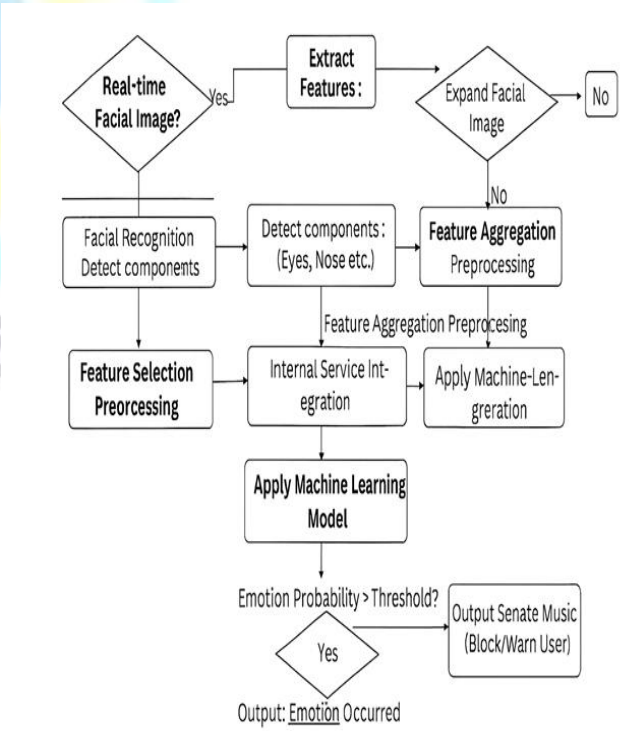

Fig: 1. Proposed Model

IV. TECH STACK

A. Frontend Technologies

The user interface of the Emotion-Based Music Generator is developed using Streamlit, enabling rapid creation of interactive applications for capturing user emotions and displaying generated music. For enhanced design and responsiveness, technologies like HTML5, CSS3, and JavaScript are used. Frameworks such as Bootstrap help create responsive layouts, while APIs enable smooth interaction with backend services..

B. Backend Technologies

The backend is powered by Python, which handles emotion detection, music generation logic, and data processing. Libraries such as OpenCV are used for facial analysis, while deep learning frameworks manage emotion classification. Streamlit also supports backend integration for real-time processing.

C. Machine Learning and Natural Language Processing

Machine learning models are used to detect emotions from facial expressions using CNN-based architectures. Frameworks like TensorFlow and PyTorch are utilized for training and inference. Extracted emotional features are mapped to musical parameters such as tempo, pitch, and rhythm for generating emotion-specific music.

D. Database Technologies

The system uses MySQL to store user data, emotional logs, and generated music metadata. This ensures efficient data retrieval and supports personalization based on user preferences and emotional history.

E. Cloud and Deployment

The application can be deployed on cloud platforms for scalability and accessibility. Tools like Docker are used for containerization, enabling consistent deployment across different environments and simplifying system management.

F. Security

Security is maintained through authentication mechanisms and encrypted communication protocols such as SSL/TLS. User data and emotional inputs are securely handled to ensure privacy and protection during real-time processing.

G. Version Control and Collaboration

Version control is managed using Git along with platforms like GitHub. This facilitates collaborative development, code tracking, and efficient project management environments, Docker Compose enables easy deployment

- The Emotion-Based Music Generator System was successfully implemented and tested under various input conditions to evaluate its performance and accuracy. The system was able to detect user emotions using multiple input modes such as facial expressions, voice input, and text data. After processing the input through preprocessing and feature extraction stages, the Machine Learning/Deep Learning model classified emotions into categories such as happy, sad, angry, and relaxed with satisfactory accuracy.

V. *RESULT SCREENSHOTS*

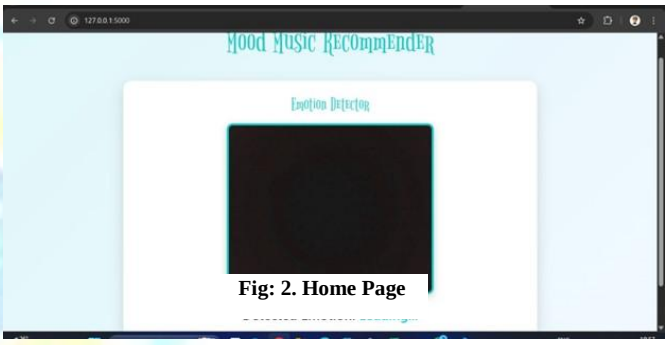

Fig: 2. Home Page

Fig: 2. Song Recommended Page

Fig: 4. Output Page

VI. CONCLUSION

The Emotion-Based Music Generation System successfully demonstrates the integration of affective computing, machine learning, and deep-learning-based generative models to produce music aligned with a user's emotional state. The project shows that emotions can be detected with high accuracy through facial expressions, voice tone, or textual sentiment, and that these emotional cues can be effectively mapped to musical features using generative neural networks. By combining real-time image and audio processing with advanced music generation architectures

Throughout the development and implementation process, the system proved efficient, robust, and capable of delivering music compositions that reflect the emotional inputs of users with notable precision. The experimental results highlight the reliability of the emotion detection modules and the creative potential of the music generation mechanisms. The generated musical pieces successfully conveyed emotional characteristics such as tempo changes, harmonic variations, rhythmic patterns, and tonal textures that align with emotional categories like happiness, sadness, calmness, or anger. Additionally, the system's ability to process inputs in real time and deliver immediate musical feedback underscores its value in interactive and user-centric environments.

This project not only meets its primary objectives but also establishes a strong foundation for future research in emotionally intelligent music systems. The findings suggest significant potential for real-world applications including digital therapy, mental health monitoring, interactive entertainment, personalized learning, and adaptive ambient environments. With further advancements—such as deeper multimodal fusion, expanded emotional datasets, enhanced generative quality, and broader cross-platform integration—
Emotion-Based Music Generation

REFERENCES

[1.] Deep Learning Framework Documentation – TensorFlow, Keras, and PyTorch official guides for implementing neural networks for audio generation and emotion detection.

[2.] Emotion Recognition Datasets – RAVDESS, CREMA-D, EMO-DB, FER2013, and DEAM dataset references used for emotion classification and music-emotion annotation

[3.] A. Mollahosseini, D. Chan, and M. H. Mahoor, "Going deeper in facial expression recognition using deep neural networks," IEEE Winter Conf. on Apps of Computer Vision (WACV), 2016.

[4.] Magenta & Google AI Music Research – Documentation and research papers on LSTM-based, Transformer-based, and GAN-based music generation models.

[5.] A. Vaswani, et al., "Attention is all you need," Advances in Neural Information Processing Systems, 2017, pp. 5998-6008.

Affective Computing Research Publications – Studies on emotion recognition, multimodal affect analysis, and human–computer emotional interaction