iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** I **Month of publication:** January 2026

DOI: https://doi.org/10.22214/ijraset.2026.76952

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue I Jan 2026- Available at www.ijraset.com

Emotional Based Music Recommendation System for Mental Wellness

Aayush Verdhan[1], Bhoomi Namdev[2], Aryan Pandey[3], Darshani Gupta[4], Bhavana Vishwakarma[5]
CSE-AIML Department, Oriental Institute Of Science & Technology, Madhya Pradesh, India

Abstract: ***This project presents an Emotion-Based Music Recommendation System designed to enhance mental wellness through intelligent and personalized music therapy. The system uses real-time facial ex- pression recognition via a webcam and Convolutional Neural Networks (CNNs) to identify users' emotional states such as happy, sad, angry, neutral, or surprised. Based on the detected emotion, suitable music is recommended to positively influence mood, reduce stress, and support emotional balance. By integrating computer vision, deep learning, and an emotion- aware recommendation engine, the system offers a non-invasive, user-friendly, and cost-effective digital solution for mental health support and affective computing applications.***

Index Terms: Mental Wellness, Emotion Detection, Music Recommendation, Facial Expression Recognition, Deep Learning.

I. INTRODUCTION

A. Background

In today's fast-paced and digitally driven world, individuals are increasingly exposed to stress, anxiety, and emotional fatigue, making mental wellness a growing concern. According to the World Health Organization, the prevalence of mental health dis- orders continues to rise globally, highlighting the need for supportive and accessible solutions. While professional psychological therapy remains crucial, there is a significant need for everyday tools that help individuals manage their emotional well-being proactively. Music therapy has proven to be an effective, non- invasive approach for emotional regulation, stress reduction, and mood enhancement. However, selecting appropriate music during moments of emotional distress can be challenging, as individuals may lack the clarity or motivation to choose suitable tracks. With advancements in artificial intelligence and human-computer interaction, emotion-aware systems offer new opportunities to personalize wellness support. This project introduces an Emotion-Based Music Recommendation System that automatically detects a user's emotional state through facial expressions and recommends suitable music to promote emotional balance and mental well-being.

B. Problem Statements

Existing music streaming and recommendation platforms primarily depend on users' listening his- tory, genre preferences, and manually curated playlists to suggest content. While these approaches are effective for personalization, they fail to account for the user's current emotional or psychological state. As a result, recommendations may not align with the user's immediate mental wellness needs. For instance, a user experiencing sadness or emotional distress may lack the motivation to actively search for uplifting music, while history-based algorithms may continue to recommend melancholic songs, in- advertently reinforcing negative emotions. This gap highlights the need for an intelligent system capable of understanding real-time emotional cues and responding appropriately. An emotion- aware music recommendation system can bridge this gap by automatically detecting a user's mood and delivering music that supports emotional balance and mental well-being.

C. Objectives

The primary objective of this project is to develop an intelligent application that serves as a mood- aware companion.
The specific goals are:

1) To develop an intelligent, emotion-aware mu- sic recommendation system for mental wellness support.
2) To implement real-time facial detection using a webcam.
3) To apply deep learning-based Facial Expression Recognition (FER) using CNN models.
4) To classify core emotions such as Happy, Sad, Angry, Neutral, Fear, Surprise, and Disgust.
5) To map detected emotions to appropriate therapeutic music playlists.

6) To enhance user experience through a simple, intuitive, and low-interaction interface.

II. LITERATURE REVIEW

A. Music and Emotion

Psychological research has consistently demonstrated the strong relationship between music and human emotions. Music has the ability to influence mood, regulate emotional arousal, and alter psycho- logical states.

According to Thayer's model of mood [1], emotions can be categorized based on two dimensions: energy (arousal) and valence (positivity or negativity). Music can be similarly classified, enabling targeted emotional regulation. Several studies highlight the effectiveness of the iso-principle in music therapy, which involves initially matching music to an individual's current emotional state and then gradually transitioning to music that promotes a desired mood. This approach helps listeners feel emotionally understood before guiding them toward improved emotional balance.

B. Existing Systems

Several attempts have been made to integrate emotion detection with music:

1) *Wearable Sensors:* Some systems use EEG or heart rate monitors to detect stress. While accurate, these are invasive and require expensive hardware.

2) *Text-Based Sentiment Analysis:* These systems analyze social media posts or text input to deter- mine mood. However, this requires active typing from the user.

3) *Visual-Based Systems:* Early implementations used basic geometric feature extraction (distance between eyes, mouth curvature) which suffered from low accuracy in varying lighting conditions.

III. METHODOLOGY

A. Proposed System Overview

The proposed Emotion-Based Music Recommendation System is designed as a modular architecture to ensure efficiency, scalability, and real-time performance. The system is divided into the following major components:

1) *Data Acquisition Module:* Captures live video in- put using the user's webcam and extracts video frames at regular intervals. It ensures real- time interaction and handles varying lighting conditions.

2) *Preprocessing Module:* Converts captured frames into grayscale images, performs noise reduction and normalization, and resizes images to match CNN input dimensions.

3) *Face Detection Module:* Identifies and localizes the user's face, filtering irrelevant background information.

4) *Emotion Detection Module:* Utilizes a trained CNN model for Facial Expression Recognition (FER) to classify emotions (Happy, Sad, Angry, Neutral, Fear, Surprise, Disgust).

5) *Recommendation Engine:* Maps detected emotions to predefined therapeutic music categories.

6) *User Interface Module:* Displays detected emotion and recommended music with minimal user interaction.

7) *System Output:* Plays emotion-appropriate music in real time.

Figure 4.7: Flow Chart

Figure 1. Data Flow Diagram

B. Facial Expression Recognition (FER)

The core component of the system is the FER module. It leverages the FER 2013 dataset, a bench- mark dataset containing labeled facial images across multiple emotions. A Convolutional Neural Network (CNN) is employed to learn spatial features [2].

1) Face Detection: Before performing emotion analysis, it is essential to accurately locate the user's face. This system employs Haar Cascade Classifiers [3], available through OpenCV. The classifier scans the input frame and identifies facial regions based on pre-trained features. Once detected, the face region is cropped and converted to grayscale to reduce computational complexity.

2) CNN Architecture: The cropped face image (48× 48 pixels) is fed into a CNN. The architecture includes Convolutional Layers to extract features like edges and textures, and Max Pooling Layers to down- sample feature maps.

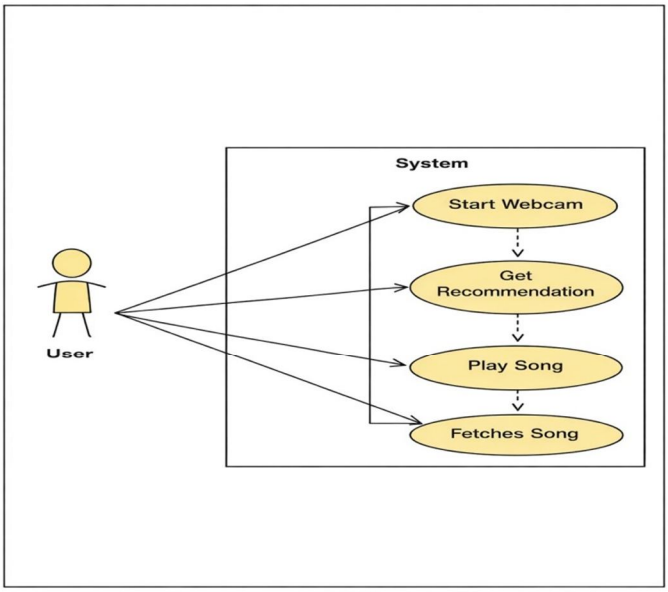


Figure 4.1: Use Case Diagram of the System

Figure 2. Overall System Architecture

C. Music Recommendation Logic

Unlike standard recommenders, this system prioritizes "Wellness." The mapping logic is as follows:

1) Detected: Angry → Recommend: Calm, Ambient, Classical (To soothe).
2) Detected: Sad → Recommend: Upbeat, Pop, Motivational (To uplift)
3) Detected: Happy → Recommend: Energetic, Dance, Pop (To maintain).
4) Detected: Neutral → Recommend: Lo-Fi, Jazz, Instrumental (To focus/relax).

IV. SYSTEM DESIGN

The system design enables real-time emotion detection and personalized music delivery. The video feed is processed frame-by-frame, passed through pre-processing (Grayscale conversion, Histogram Equalization), and then through the classifier.

A. Technology Stack

1) Language: Python
2) Computer Vision: OpenCV
3) Deep Learning Framework: TensorFlow Keras
4) GUI: Tkinter or Streamlit
5) Dataset: FER-2013

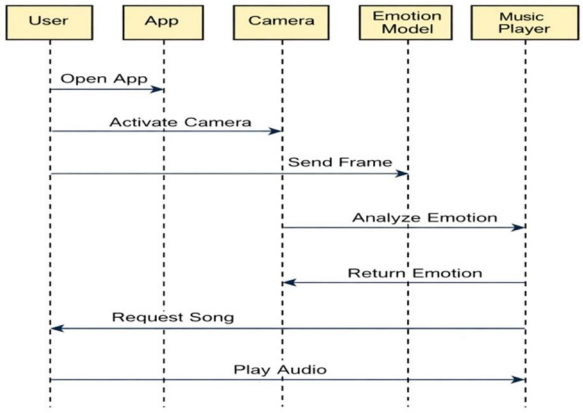


Figure 4.2: Sequence Diagram

Figure3. CNN Architecture for Facial Expression

V. RESULTS AND DISCUSSION

A. Model Performance

The CNN model was trained using the FER-2013 dataset (28,709 training images, 3,589 validation im- ages). The images are grayscale and of low resolution. The model achieved a training accuracy of approximately 65-70%, which is competitive for this dataset due to the complexity of recognizing subtle facial variations. Validation results indicated reliable classification of core emotions.

B. Qualitative Analysis

In real-time testing, the system successfully detects distinct emotions such as ”Happy” and ”Surprise” with high confidence. ”Sad” and ”Neutral” occasion- ally overlap due to subtle facial differences. The music recommendation latency is minimal.

C. Impact on Wellness

Initial user evaluations indicate that the system positively influences mental wellness. The auto- mated detection reduces decision fatigue. For in- stance, when anger is detected, calming music helps users relax. This real-time intervention demonstrates potential as a digital therapeutic aid.

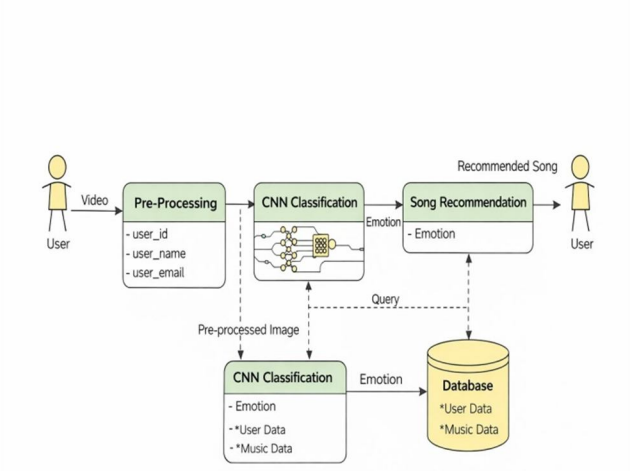


Figure 4. Emotion-to-Music Recommendation Mapping Logic

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue I Jan 2026- Available at www.ijraset.com

VI. CONCLUSION AND FUTURE SCOPE

A. Conclusion

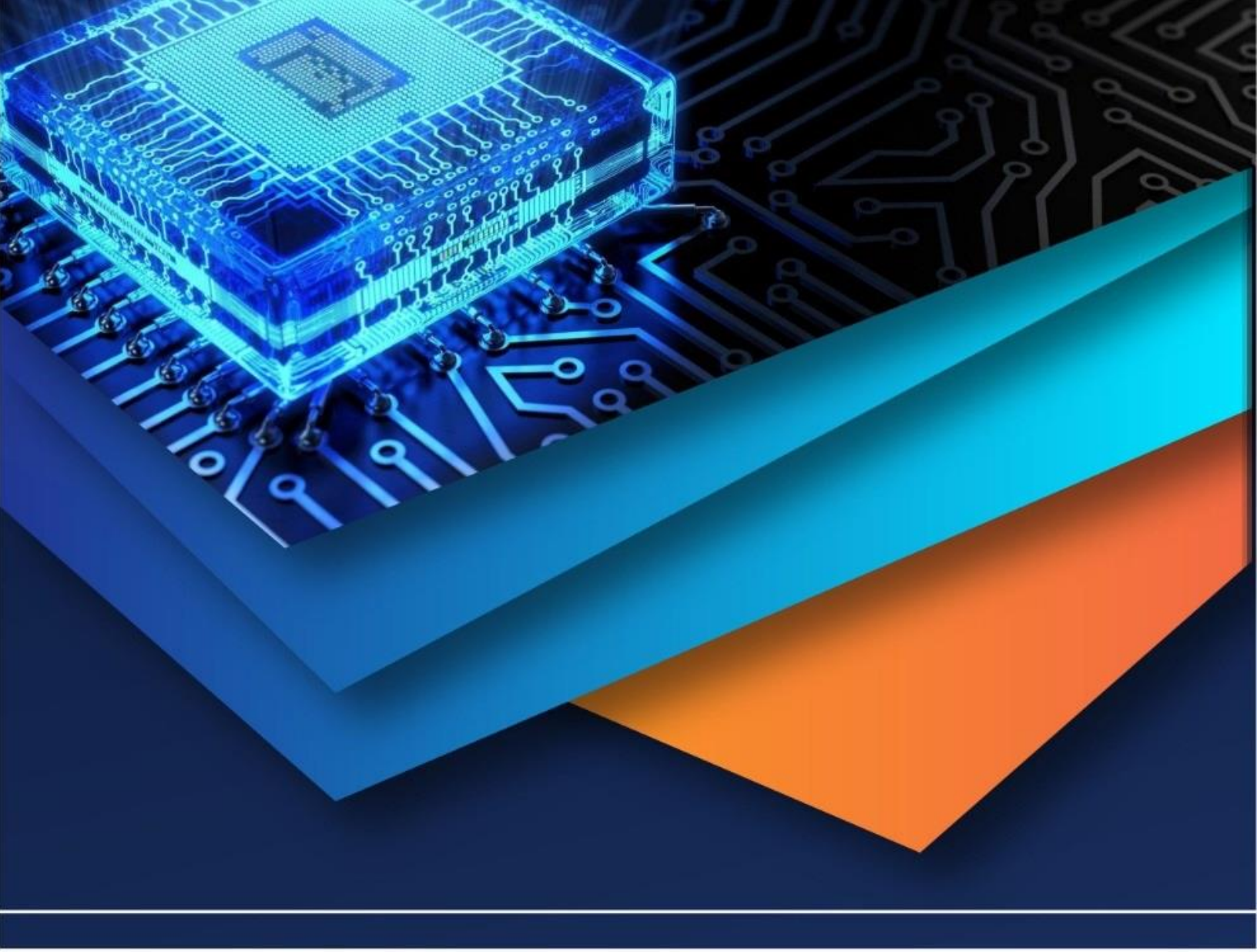
This project demonstrates the successful development of an Emotion-Based Music Recommendation System that integrates computer vision and deep learning to support mental wellness. By detecting and classifying a user's emotional state in real time, the system bridges the gap between human emotion and digital content. The use of a CNN trained on the FER-2013 dataset ensures accurate recognition, while user evaluations highlight the system's potential to reduce decision fatigue and provide immediate emotional support.

B. Future Scope

1) Multimodal Input: Combining facial expressions with voice tone and physiological signals.
2) Personalization: Learning individual user preferences over time.
3) IoT Integration: Connecting with smart home devices (lights, speakers).
4) Cross-Platform Compatibility: Extending to mobile apps.
5) Real-Time Feedback Loop: Monitoring emotional changes during playback.
6) Expanded Emotion Categories: Including stress, boredom, or relaxation.

REFERENCES

[1] R. E. Thayer, The Biopsychology of Mood and Arousal, Oxford University Press, 1989.
[2] I. Goodfellow, Y. Bengio, and A. Courville, "Challenges in representation learning: A report on three machine learning contests," Neural Networks, vol. 64, pp. 59-63, 2013.
[3] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2001, pp. 1-511-1-518.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY