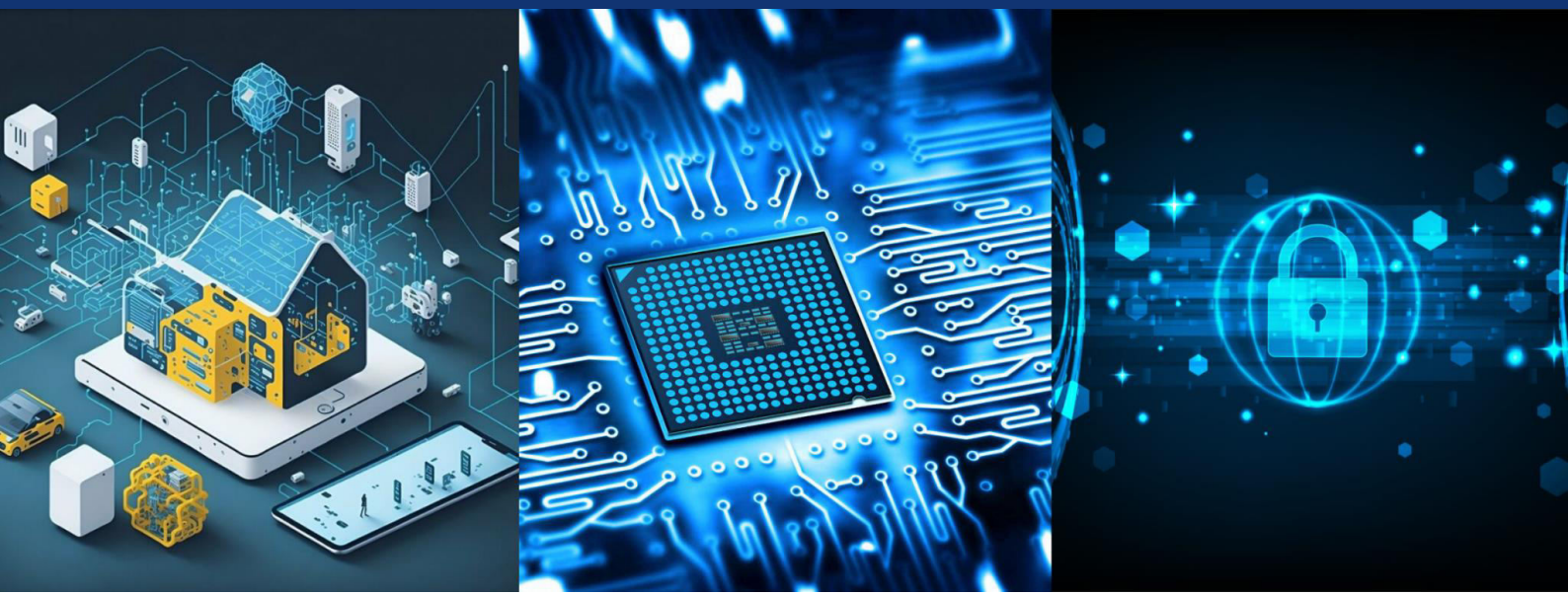
IJIRCCE

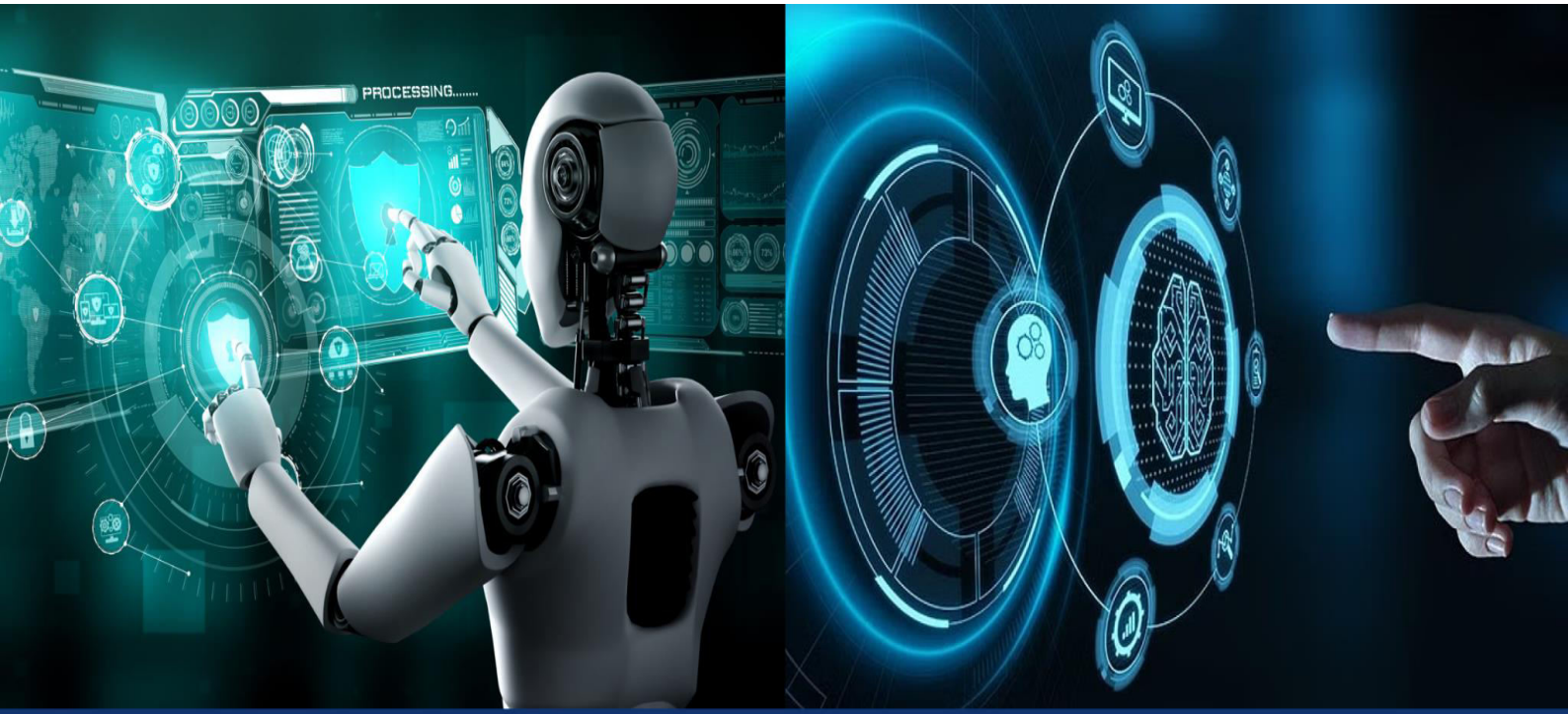
ISSN(O): 2320-9801

ISSN(P): 2320-9798



International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.771

Volume 14, Issue 3, March 2026

© 2026 IJIRCCE | Volume 14, Issue 3, March 2026 DOI: 10.15680/IJIRCCE.2026.1403054

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Emotion-Based Music Therapy System

Y. Sunaina[1], Y. Gayatri[2], Kodamanchili Kusuma Kumari[3], Kakumanu Lokeshwar[4], Mukala Kuladeep[5], Lopinti Thrushitha[6], Chirra Nikhil Reddy[7]

Assistant Professor, Department of CSE (Data Science), NSRIT, Vishakhapatnam, India[1,2]

Student of Department of CSE (Data Science), Nadimpalli Satyanarayana Raju Institute of Technology, Visakhapatnam, Andhra Pradesh, India[3,4,5,6,7]

ABSTRACT: With the rapid advancement of technology, music has emerged as a powerful medium for influencing emotions and promoting mental well-being. This project presents an Emotion-Based Music Therapy System that applies artificial intelligence and machine learning techniques to detect user emotions in real time and recommend suitable music accordingly. The system analyzes facial expressions captured through a webcam using computer vision and deep learning models, particularly Convolutional Neural Networks (CNNs), to accurately classify emotional states. Based on the detected emotion, the application suggests songs intended to enhance mood, such as calming tracks for stress, uplifting melodies for sadness, and energetic music for low motivation. To strengthen personalization, users can select their preferred language, ensuring culturally relevant and emotionally appropriate recommendations. Unlike traditional music recommendation systems that depend mainly on historical listening patterns, this system dynamically adapts to the user's current emotional condition, providing a more interactive and responsive experience. Integration with the Spotify API enables real-time access to curated playlists and tracks. Experimental observations indicate that emotion-aware music recommendations can contribute positively to mood regulation and user engagement. This project highlights the potential of AI-driven emotional intelligence in creating innovative, adaptive, and therapeutic digital music applications.

KEYWORDS: Emotion Recognition, Facial Expression Analysis, Deep Learning, Convolutional Neural Networks (CNN), Real-Time Detection, Spotify API, Music Recommendation, OpenCV, Flask, Personalized Music Therapy, FER2013 Dataset.

I. INTRODUCTION

Music has long been acknowledged as a powerful medium for influencing human emotions and supporting mental well-being. While conventional music recommendation systems primarily depend on historical listening behavior and user preferences, they often fail to respond effectively to the user's immediate emotional state. Emotions are dynamic, and a static recommendation approach may not always align with the user's current mood or psychological needs.

This project presents an **Emotion-Based Music Therapy System** that leverages **Artificial Intelligence (AI)** and **Machine Learning (ML)** to detect and interpret a user's real-time emotional condition and recommend appropriate music accordingly. The system captures live input through a webcam and analyzes facial expressions using **computer vision techniques** and **Convolutional Neural Networks (CNNs)** to classify emotions such as happiness, sadness, anger, and neutrality. By integrating intelligent emotion recognition with a music recommendation engine, the system aims to deliver a more adaptive and personalized user experience.

Developed using **OpenCV**, **deep learning models**, and a **Flask-based web application**, the system provides an interactive and user-friendly interface designed with a Neumorphism-inspired UI. Unlike traditional entertainment-focused platforms, this AI-driven solution dynamically adapts to the user's emotional changes, transforming music consumption into a responsive and therapeutic interaction.

This project emphasizes the potential of AI-powered emotional intelligence in enhancing personalized digital experiences. By bridging the gap between technology, music, and mental wellness, the system offers a novel approach to mood regulation, stress reduction, and emotional support. Future improvements, including advanced architectures such as **Vision Transformers (ViTs)** and scalable cloud deployment, can further enhance system accuracy, efficiency, and accessibility.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

II. RESEARCH OBJECTIVES AND METHODOLOGY

The primary and secondary objectives of this research and project are to develop an AI-driven Emotion-Based Music Recommendation System and evaluate its effectiveness in enhancing emotional well-being through personalized music therapy. The study aims to:

1. Analyze the role of emotion recognition and artificial intelligence in creating a personalized and responsive user experience.
2. Evaluate the accuracy and efficiency of real-time facial expression detection and its impact on music recommendation quality.
3. Identify key deep learning models and techniques (e.g., CNN, RNN) applicable for emotion detection from visual and audio inputs.
4. Examine the effectiveness of AI-based music recommendation in improving mood, engagement, and mental wellness.
5. Integrate and assess the use of Spotify API for emotion-specific playlist curation.

III. LITERATURE SURVEY

Conventional music recommendation systems predominantly rely on user listening history, genre preferences, and collaborative filtering techniques to suggest songs. Although these methods provide personalization based on past behavior, they frequently overlook the user's real-time emotional state, limiting their ability to adapt to dynamic mood variations. Since emotions fluctuate continuously, static recommendation mechanisms may fail to deliver contextually relevant music experiences. Recent advancements in Artificial Intelligence (AI) and Machine Learning (ML) have facilitated the development of emotion-aware recommendation systems capable of adapting to users' current psychological conditions. Numerous studies have investigated emotion detection through modalities such as facial expression analysis, speech signals, and physiological responses, employing techniques including Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Support Vector Machines (SVMs), with deep learning models demonstrating superior performance due to their ability to capture complex emotional patterns. An extensive review of existing literature was conducted, analyzing over fifty research articles, from which key studies were selected based on relevance to real-time emotion detection and AI-driven personalization. The review indicates that benchmark datasets such as FER2013, DEAM (Database for Emotional Analysis of Music), and the Million Song Dataset are widely used for model training and evaluation, while some studies also utilize custom datasets derived from user feedback and sensor-based data. The findings suggest that emotion- based music recommendation systems offer more adaptive and engaging user experiences compared to traditional playlist-based approaches. Building upon these insights, the proposed system integrates real-time facial emotion recognition with language-aware music recommendation to deliver emotionally aligned and therapeutically beneficial listening experiences.

IV. METHODOLOGY

The emotion-based music recommendation system integrates real-time emotion detection using deep learning with Spotify API-based song recommendations to enhance user experience. The system is designed to detect facial expressions using a webcam, classify the detected emotion, and suggest a relevant Spotify playlist that matches the user's mood. The methodology consists of the following steps:

1. **Data Collection and Preprocessing:**

To develop an accurate emotion recognition model, facial expression images are processed and standardized.

1.1 Facial Expression Data Collection

The **FER2013 dataset** is utilized, consisting of 48×48 grayscale facial images categorized into seven emotions: Angry, Disgusted, Fearful, Happy, Neutral, Sad, and Surprised. In addition to the dataset, live webcam input is captured and processed in real time using OpenCV.

1.2 Image Preprocessing

Image preprocessing is performed using Keras' **ImageDataGenerator**, including pixel rescaling (1/255 normalization) and data augmentation techniques. Images are converted to grayscale, resized to 48×48 pixels, and organized into batches of 64 for efficient training.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

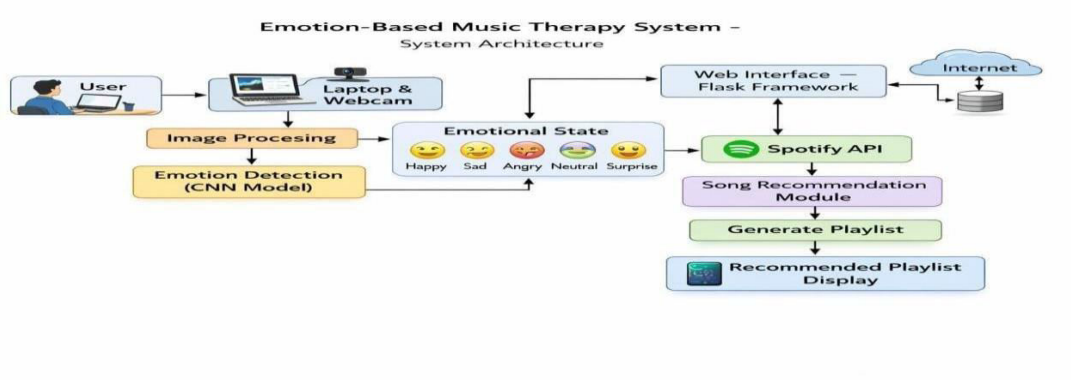
2. **Emotion Detection and Classification:**

2.1 Deep Learning Model for Emotion Recognition

A Convolutional Neural Network (CNN) model is employed to classify facial emotions. The architecture includes multiple Conv2D layers with ReLU activation for feature extraction, followed by MaxPooling layers for dimensionality reduction. Dropout layers (0.25–0.5) are applied to minimize overfitting. The flattened output is connected to Dense layers, with a final Softmax layer for seven-class emotion classification. The model is trained using the Adam optimizer and Categorical Crossentropy loss, achieving approximately 86% accuracy after 75 epochs

2.2 Real-Time Emotion Detection Using Webcam

Real-time video frames are captured using OpenCV. Faces are detected with a Haarcascade classifier, and the detected face region is preprocessed before being passed to the trained CNN model. A multithreaded webcam stream improves frame rate and system responsiveness.



3. **Music Recommendation Engine:**

Once an emotion is detected, the system fetches a playlist from Spotify API based on the recognized emotion.

3.1 Spotify API Integration

The Spotipy library is used for Spotify Web API authentication and music retrieval. Functions such as getTrackIDs() extract track identifiers from emotion-specific playlists, while getTrackFeatures() retrieves metadata including song name, album, and artist. The data is stored using Pandas DataFrames and optionally saved as CSV for offline analysis

4. **Real-Time Feedback and Playlist Adjustment:**

The system continuously monitors user expressions to dynamically adjust recommendations. If the detected emotion remains unchanged after music playback, the system updates the playlist to maintain engagement and emotional alignment.

5. **Deployment and User Interface:**

5.1 Flask Web Application

The interface allows users to view detected emotions, select language preferences, and interact with music recommendations.The system is implemented as a Flask-based web application featuring a Neumorphism-inspired UI.

5.2 Storage and Model Management

Trained model weights are stored in model.h5 for efficient loading. Spotify track information is preserved in CSV files for quick access and reduced API calls

V. HARDWARE AND SOFTWARE SETUP

To build and test the Emotion-Based Music Recommendation System, specific hardware and software components were required for real-time emotion detection and Spotify playlist integration.

Hardware Requirements:

Webcam – Captures real-time facial expressions of the user.

Computer/Laptop – Runs the Python-based application, processes video frames, and fetches playlists.

Software & Libraries:

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

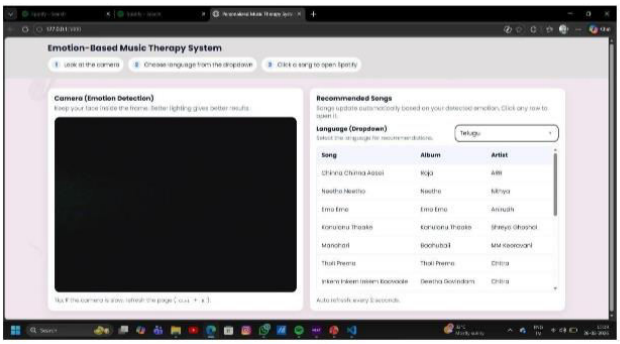
International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Python – Core programming language used for development. OpenCV – Handles webcam access and real-time video processing.
TensorFlow & Keras – Used to train and run the CNN model for emotion recognition. Flask – Backend web framework for running the interface.
Spotipy – A Python wrapper for the Spotify Web API, used to fetch music based on detected emotions.
Pandas – Used for organizing and storing song data in CSV format.
Haarcascade Classifier – Used for initial face detection before emotion prediction.

VI. SYSTEM IMPLEMENTATION

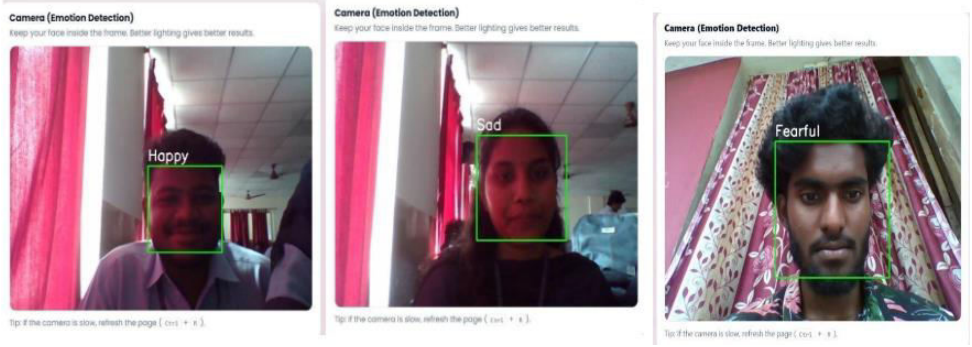
The implementation was divided into multiple steps, ensuring each component worked effectively before integrating everything into a functional system.
Step 1: Capturing and Processing Video Feed
The webcam is initialized using OpenCV to capture continuous video frames. Each frame is converted to grayscale and resized to 48×48 pixels to match the CNN input dimensions. The processed frame is then passed to the trained CNN model for real-time emotion prediction.



Step 2: Emotion Recognition using CNN:

The Convolutional Neural Network (CNN) model was trained using the FER2013 dataset, which contains labeled facial images representing seven distinct emotions: Angry, Disgusted, Fearful, Happy, Neutral, Sad, and Surprised. The model architecture was designed with multiple Conv2D layers for deep feature extraction, along with MaxPooling layers for dimensionality reduction. Dropout layers were incorporated to prevent overfitting, and Dense layers with Softmax activation were used for multi-class emotion classification.

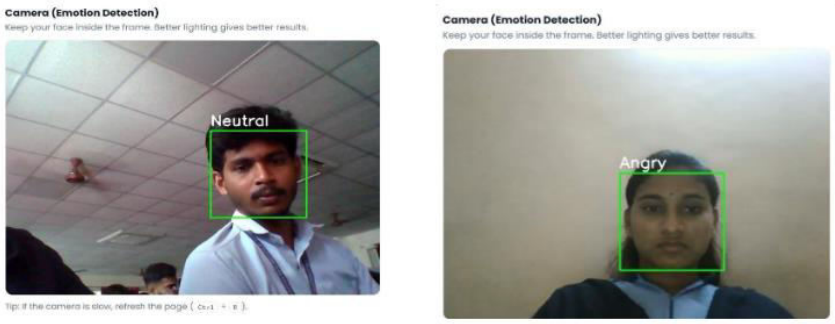
Upon processing each frame captured from the webcam, the CNN model predicts one of the seven emotional categories and returns the corresponding emotion label. This detected emotion is then utilized by the system to trigger the music recommendation module, which selects songs based on the identified emotional state and the user's preferred language, ensuring a personalized and mood-aligned listening experience.



www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

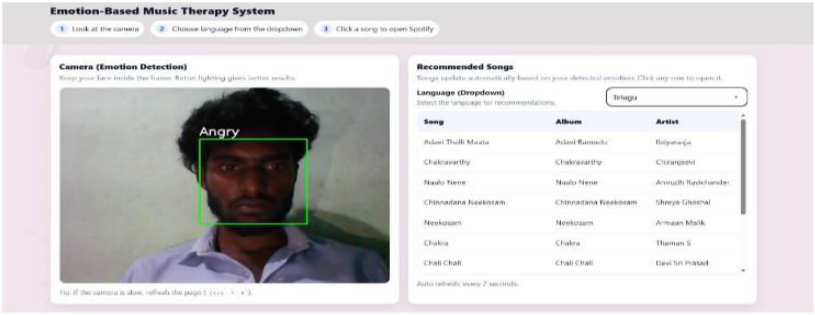
(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Step 3: Music Recommendation with Spotify API

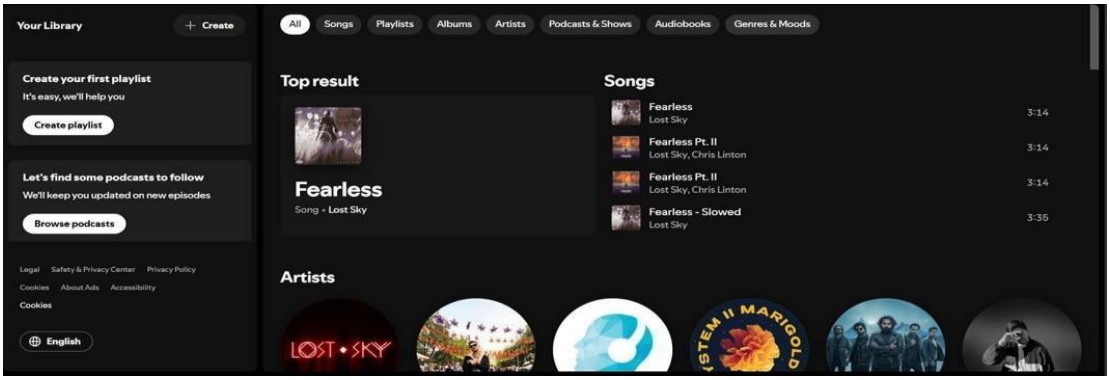
The detected emotion, along with the user's selected preferred language, was passed to a mapping dictionary (music_dist) that associates each emotion with a corresponding Spotify playlist. This mapping ensures that the recommended songs are both emotionally relevant and language- specific, enhancing personalization.

Using the Spotipy library, the system retrieved tracks from the selected playlist through Spotify's Web API. The fetched track details, including song name, album, and artist, were organized into a Pandas DataFrame and either displayed within the application interface or exported as CSV files for each emotion–language combination.



Step 4: Real-Time Playlist Switching:

As the user's facial expressions changed, the CNN model continuously monitored and updated the detected emotional state. The system performed real-time emotion tracking by analyzing each processed webcam frame. When a new emotion was identified, the application automatically retrieved and presented an updated set of music recommendations. This enabled dynamic mood-based playlist switching, ensuring that the suggested songs remained aligned with the user's current emotional condition.To prevent excessive Spotify API requests and reduce rapid playlist fluctuations, emotion predictions were stabilized using a short buffering mechanism before triggering playlist transitions.



www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Step 5: User Interface

A Flask-based web interface was created using HTML and CSS with Neumorphism design, allowing users to interact visually with the application.

The interface displayed:

- Current detected emotion
- Suggested Spotify playlist
- Track details like song name, album, and artist
- Users could optionally click on songs to be redirected to Spotify for full playback. The application continued running until manually stopped by the user.

VII. CONCLUSION

The Personalized Music Therapy System based on Facial Expression is an intelligent application developed to enhance the music listening experience through real-time emotion-aware recommendations. By utilizing facial expression recognition, machine learning, and music recommendation techniques, the system dynamically selects songs that correspond to the user's detected emotional state. The project effectively integrates computer vision, AI-based emotion classification, and external services such as the Spotify API to deliver an automated and user-centric solution. The implementation demonstrates how emotion recognition can be applied to personalize digital interactions and improve user engagement. Additionally, the feedback mechanism contributes to refining recommendation accuracy, enabling better alignment between emotional states and music selection. Overall, the system highlights the potential of combining artificial intelligence with music therapy concepts to create adaptive, interactive, and emotionally responsive applications.

REFERENCES

1. Keerthana M., Shruthi M., and Aravind Kumar S., "Emotion Based Music Recommendation System," International Journal of Creative Research Thoughts (IJCRT), vol. 9, no. 6, pp. e356–e360, June 2021.
2. Anand R., Sabeenian R.S., Gurang D., and Kirthika R., "AI-based Music Recommendation System using Deep Learning Algorithms," International Journal of Creative Research Thoughts (IJCRT), vol. 8, no. 5, pp. 1234–1240, May 2020.
3. Katkuri S., Chegoor M., Sreedhar K.C., and Sathyanarayana M., "Emotion Based Music Recommendation System," International Journal of Advanced Research in Computer and Communication Engineering, vol. 9, no. 5, pp. 45–50, May 2020.
4. Joshi S., Jain T., and Nair N., "Emotion Based Music Recommendation System Using LSTM-CNN Architecture," International Journal of Creative Research Thoughts (IJCRT), vol. 11, no. 2, pp. 678–685, February 2023.
5. Babu T., Nair R.R., and Geetha A., "Emotion-Aware Music Recommendation System: Enhancing User Experience Through Real-Time Emotional Context," arXiv preprint, arXiv:2311.10796, November 2023.
6. Chang X., Zhang X., Zhang H., and Ran Y., "Music Emotion Prediction Using Recurrent Neural Networks," arXiv preprint, arXiv:2405.06747, May 2024.
7. Mammadli R., Bilgin H., and Karaca A.C., "Music Recommendation System based on Emotion, Age and Ethnicity," arXiv preprint, arXiv:2212.04782, December 2022.
8. Jing E., Liu Y., Chai Y., Yu S., Liu L., Jiang Y., and Wang Y., "Emotion-aware Personalized Music Recommendation with a Heterogeneity-aware Deep Bayesian Network," arXiv preprint, arXiv:2406.14090, June 2024.
9. Florence and Uma, "Emotion-Based Music Recommendation System Using Facial Expression Analysis," International Journal of Research Publication and Reviews, vol. 5, no. 11, pp. 7876–7882, November 2024.
10. Patel et al., "Emotion-Based Music Player Using CNN for Facial Emotion Recognition," International Journal of Research Publication and Reviews, vol. 5, no. 11, pp. 7883–7890, November 2024.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details