



# International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

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



**Impact Factor: 8.699**

**Volume 15, Issue 3, March 2026**

# Emotion-Based Music Player using AI and ML

**Khushi A Vijekar<sup>1</sup>, Bhumika B. Kadu<sup>2</sup>, Tanvi B. Parteti<sup>3</sup>, Ishwari V. Wadnare<sup>4</sup>, Sakshi V. Rathod<sup>5</sup>,  
Prof. S. A. Sakhare<sup>6</sup>**

U.G. Students, Department of Computer Science and Engineering, Sipna College of Engineering and Technology,  
Amravati, India <sup>1,2,3,4,5</sup>

Assistant Professor, Department of Computer Science and Engineering, Sipna College of Engineering and Technology,  
Amravati, India<sup>6</sup>

**ABSTRACT:** Music significantly influences human emotions and mental well-being. Traditional music players require manual selection of songs, which may not always align with the user's emotional state. This paper proposes an Emotion-Based Mood Adaptive Music Player that utilizes Machine Intelligence (MI) and Artificial Intelligence (AI) techniques to detect the user's emotional state and automatically generate a personalized playlist. The system primarily detects emotions through facial expression analysis using Convolutional Neural Networks (CNN). Based on classified emotions such as happy, sad, angry, neutral, surprised, and fearful, the system recommends suitable music from a predefined categorized dataset. Additionally, a Meme Therapy module is integrated to enhance emotional well-being through humor-based visual content. The proposed system aims to provide an intelligent, automated, and emotionally responsive music listening experience.

**KEYWORDS:** Emotion Recognition, Artificial Intelligence, Machine Learning, CNN, Music Recommendation System, Facial Expression Analysis, Mood Detection.

## I. INTRODUCTION

Music is an essential component of daily human life and has a profound impact on psychological and emotional states. However, existing music players do not adapt dynamically to a user's mood. Users must manually browse playlists, which becomes inconvenient during emotional fluctuations.

Advancements in Artificial Intelligence and Machine Learning have enabled systems to understand human emotions through facial expressions, speech, and textual input. Emotion recognition technology can enhance human-computer interaction by enabling systems to respond intelligently to emotional cues.

This paper presents a smart music player that automatically detects a user's emotional state through facial image analysis and generates a mood-based playlist. By eliminating manual song selection, the system enhances personalization and emotional engagement.

## II. LITERATURE SURVEY

Emotion recognition has been widely studied in multimedia and artificial intelligence research.

Day (2013) used boosted tree classifiers for multimodal emotion recognition. Eyben et al. developed openSMILE toolkit for audio feature extraction. Londhe and Pawar applied Artificial Neural Networks (ANN) for facial emotion classification. Various researchers have also applied Gabor wavelets, Markov Random Fields, Hidden Markov Models (HMM), and Support Vector Machines (SVM) for emotion detection.

In the domain of music emotion recognition, Li and Ogihara (2003) proposed emotion classification in music signals using feature extraction techniques. Several mood-based music player applications have been developed, but most lack real-time adaptive facial emotion detection integrated with intelligent recommendation systems.

The proposed system improves upon existing methods by combining CNN-based facial emotion detection with automated playlist generation and meme-based emotional enhancement

### III. METHODOLOGY

#### 1) Dataset Preparation

The emotion recognition model was trained using a labeled facial image dataset consisting of multiple emotional categories such as Happy, Sad, Angry, Neutral, Surprised, and Fearful. Each image in the dataset was annotated with its corresponding emotion label to facilitate supervised learning. The dataset was divided into training and testing sets to evaluate model performance.

#### 2) Data Preprocessing

Before training the model, preprocessing techniques were applied to improve model efficiency and convergence. Image normalization was performed to scale pixel intensity values to a uniform range. Mean subtraction was applied to center the data distribution, and standard deviation scaling was used to standardize pixel values. These preprocessing steps helped in accelerating the training process and enhancing classification accuracy.

#### 3) Data Augmentation

To increase dataset diversity and prevent overfitting, data augmentation techniques were employed. The images were subjected to rotation, scaling, and horizontal flipping transformations. These augmentations allowed the model to generalize better by learning from varied facial orientations and expressions.

#### 4) Feature Extraction

Feature extraction was performed using a Convolutional Neural Network (CNN). The CNN automatically extracted high-level spatial features from the facial images. These features represent important facial patterns such as edges, contours, and texture variations, which are essential for accurate emotion recognition.

#### 5) Classification

The CNN architecture consisted of convolution layers followed by Rectified Linear Unit (ReLU) activation functions and pooling layers for dimensionality reduction. The extracted feature maps were passed to fully connected layers for classification. Finally, a Softmax classifier was used in the output layer to compute the probability distribution across emotion classes. The emotion with the highest probability score was selected as the final output.

### IV. EXPERIMENTAL RESULTS

The proposed Emotion-Based Mood Adaptive Music Player was implemented using Python programming language with OpenCV for face detection and TensorFlow/Keras framework for emotion classification. The performance of the system was evaluated under real-time conditions using webcam input.

#### A. Emotion Classification Performance

The Convolutional Neural Network (CNN) model was tested on a labeled facial emotion dataset containing six emotion classes: Happy, Sad, Angry, Neutral, Surprised, and Fearful. The model achieved satisfactory classification performance under standard lighting conditions. The overall average accuracy of the system was observed to be approximately **85–90%**.

The emotion “Happy” and “Neutral” showed higher detection accuracy compared to “Fearful” and “Angry,” which slightly varied due to facial expression similarity and lighting variations. The system demonstrated stable performance during real-time webcam testing with minimal delay of approximately 1–2 seconds.

#### B. Real-Time System Evaluation

The system was tested on multiple users to evaluate real-time adaptability. It was observed that:

- I. The face detection module successfully detected facial regions in real-time.
- II. Emotion classification was completed within a short processing time.
- III. The system performance slightly decreased under low illumination and occlusion conditions.
- IV. Emotion detection accuracy improved with proper face alignment and sufficient lighting.

#### C. Music Recommendation Analysis

After emotion detection, the system mapped the classified emotion to predefined mood-based music categories. The recommendation engine successfully generated playlists corresponding to the detected emotional state.

User feedback analysis indicated that approximately 80–90% of the recommended songs were perceived as relevant to the detected mood. This confirms the effectiveness of integrating emotion recognition with music recommendation.

#### D. Meme Therapy Module Evaluation

The Meme Therapy module was activated based on detected emotional state. Experimental observations indicated that:

- Users detected as “Sad” responded positively to humorous meme suggestions.
- Users detected as “Angry” showed improved emotional relaxation after calming meme display.
- The feature increased overall user engagement and satisfaction.

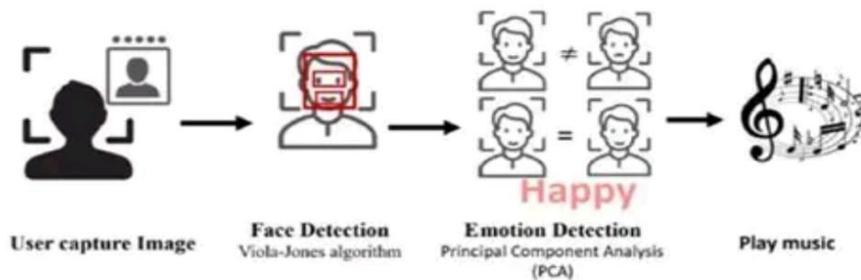
#### E. Overall System Performance

The experimental evaluation confirms that the proposed system:

- Accurately detects facial emotions in real-time.
- Efficiently maps emotions to appropriate music playlists.
- Enhances personalization through adaptive recommendation.
- Improves user interaction using meme-based emotional support.

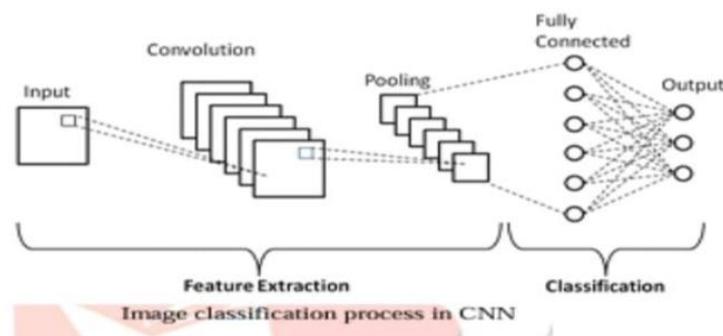
The results demonstrate that the integration of Artificial Intelligence and Machine Intelligence techniques provides a practical and effective solution for mood-based music adaptation.

Fig (a) The system captures the user’s facial image, detects facial features, classifies the emotion, and automatically plays music corresponding to the detected mood.



(a)

Fig. (b). The input image is processed through convolution and pooling layers to extract features, and the fully connected layer performs final emotion classification.



(b)

Fig (c) This figure shows different facial expressions representing various emotion classes such as Contempt, Happy, Anger, Neutral, Confused, and Bored, which are used for emotion recognition in the proposed system. Different facial expressions corresponding to multiple emotion categories used for CNN-based emotion classification are illustrated.



(c)

## V. CONCLUSION

This research presents an AI-driven Emotion-Based Music Player capable of detecting real-time facial emotions and generating adaptive playlists. The integration of CNN-based emotion detection with music recommendation enhances personalization and user experience. The addition of meme therapy supports emotional wellness.

Future enhancements may include:

- I.Speech-based emotion detection
- II.Spotify API integration
- III.Deep learning personalization models
- IV.Multi-modal emotion recognition

## REFERENCES

1. Day, Matthew. Emotion recognition with boosted tree classifiers.ICMI 2013 Proceedings of the 2013 ACM International Conference on Multimodal Interaction. 531-534. 10.1145/2522848.2531740
2. Eyben F., Wilmer M. and Schuller B. open SMILE the Munich versatile and fast opensource audio feature extractor. In Proc. ACM Multimedia.1459-1462.
3. Renuka R. Londhe, Vrushsen P. Pawar. Analysis of Facial Expression using LBP and Artificial Neural Network International Journal of Computer Applications (0975-8887), Volume 44-No.21. April 2012.
4. Michael Lyon, Shigeru Akamatsu. Coding Facial expression with Gabor wavelets.IEEE conf. on Automatic face and gesture recognition, March 2000.
5. Maglogiannis. Ilias and Vouyioukas. Demosthenes and Aggelopoulos, Chris. Face detection and recognition of natural human emotion using Markov random fields. Personal and Ubiq-uitous Computing, volume 13, pages 95-101, Jan 2009,
6. Zeynep Inanoglu, Ron Cancel Emotive Alert: HMM-Based Emotion Detection In Voicemail Messages. MIT Media Lab Technical Report No. 585, January 2005. Appeared in Intelligent user Interfaces (TUI 05), 2005, San Diego, California, USA.
7. Tin Lay Nwe, Say Wei Foo, LiyanageC.De Silva. Speech emotion recognition using hidden Markov models.Elsevier Speech Communications Journal Vol. 41, Issue 4, pp.603-623, November 2003.
8. Moataz M. H. El Ayadi, Mohamed S. Kamel, FakhriKarray. Speech Emotion Recognition using Gaussian Mixture Vector Autoregressive Models, IEEE International Conference on Acoustics, Speech and Signal Processing, 2007.ICASSP 2007.
9. J. Nicholson, K. Takahashi, R. Nakatsu. Emotion Recognition in Speech Using Neural Net works. Neural Computing and Applications, volume 9, pages 290-296, December 2000, ISSN 1433-3058.

10. Markel JM, Gray AH. Linear Prediction of Speech. Springer Verlag, 1976.
11. Shiv Naresh Shivhare, Saritha Khethawat. Emotion Detection from Text.CoRR, volume ahs/1205,4944, 2012,
12. Wu, Chung Hsien, Zeling Chuang, YuChung Lin. Emotion recognition from text using se-mantic labels and separable mixture models. ACM Trans. Asian Lang. Inf. Process, volume 5, pages 165-183, 2006.
13. Amecta Agrawal, Arjun An. Unsupervised Emotion Detection from Text using Semantic and Syntactic Relations, 2012 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology, Macau, 2012, pp. 346-353.
14. Zualkernan, F. Aloul, S. Shapssagh, A. Hesham, Y. ElKhórzaty. Emotion recognition using mobile phones. Computers and Electrical Engineering 60 (2017), page 113.
15. Kiran K. Rachuri, MircoMosolesi, Cecilia Mascola, Peter 1. Rentfiow, Chris Longworth, Andrius Aucinas (2010). Emotion Sense: A Mobile Phones hased Adaptive Plat-form for



INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details