IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

# FACIAL EMOTION BASED MUSIC RECOMMENDATION SYSTEM

**Shubham Kulkarni[1], Sanskruti Lingawar[2], Suvarna Patil[3], Dinesh Bhadane[4]**

[1,2]UG Student, Dr. DY Patil Institute of Technology, Pimpri, Pune, Maharashtra, India 411018

[3,4]*Professor, Dept. of Computer Engineering*, Dr. DY Patil Institute of Technology, Pimpri, Pune, Maharashtra, India 411018

## Abstract

The advancement of facial emotion recognition technology has led to the development of a facial emotion-based music recommendation system. This system utilizes facial expression analysis to understand the listener's emotional state and recommend music that aligns with that emotion. In this paper, we explore the architecture and implementation of a facial emotion-based music recommendation system, including the underlying algorithms and technologies. We also discuss the challenges and limitations of this system, as well as potential future directions for research.

*Key Words*: **Facial recognition, Emotion detection, Music recommendation, Expressions, Feature Extraction, Convolutional Neural Network.**

## 1.INTRODUCTION

The relationship between music and human emotions has long been recognized, with music having the ability to evoke various emotional states. In recent years, the application of facial emotion recognition technology has gained significant attention in the field of music recommendation systems. A facial emotion-based music recommendation system is an innovative solution that aims to personalize music recommendations based on the listener's facial expression. This technology can analyze a user's facial expressions in real-time and recommend music that best matches their current emotional state. This paper explores the development and implementation of a facial emotion-based music recommendation system, including its underlying technology, its potential benefits, and the challenges that need to be addressed to optimize its effectiveness.

The goal of this research is to provide a comprehensive understanding of the capabilities of such a system, along with its potential applications in the music industry and beyond.

The increasing availability of high-quality facial recognition technology and the ever-growing music streaming services make this technology a promising solution for personalized music recommendations. This system has the potential to revolutionize the way we interact with music, making it more personalized and emotionally engaging. By integrating facial emotion recognition technology into music recommendation systems, we can create a more personalized and enjoyable music experience for listeners. Moreover, this technology can also have practical applications in other fields such as healthcare, education, and entertainment. However, as with any emerging technology, there are ethical, privacy, and data security concerns that need to be addressed. This research aims to identify these concerns and provide recommendations to ensure the safe and ethical use of this technology. Ultimately, this paper aims to contribute to the development of facial emotion-based music recommendation systems and provide insights for future research in this field.

To achieve the objectives of this study, we will review and analyze previous research studies on facial emotion recognition technology and music recommendation systems. We will also explore various machine learning and deep learning algorithms that are commonly used in developing facial emotion recognition models. Additionally, we will conduct an empirical study by collecting data from participants to evaluate the effectiveness of the proposed system.

The findings of this study can have significant implications for the music industry, providing music streaming services and music marketers with new ways to personalize their offerings. This system can help to increase listener engagement, enhance user experience, and ultimately, boost revenue. Furthermore, this technology has practical applications in healthcare, where it can be used to monitor patients' emotional state and provide personalized therapy. Overall, this study aims to contribute to the understanding of the potential applications of facial emotion recognition technology in music and other fields, paving the way for further

research and development in this promising area of technology.

## 2.SYSTEM OVERVIEW

The facial emotion-based song recommendation system consists of three main components: the face recognition system, the emotion classification system and the song recommendation system. The facial emotion recognition system captures a video of the listener's face and extracts facial features such as eye movement, eyebrow position, and mouth shape. These features are then used to determine the listener's emotional state using a trained machine learning model.

The song recommendation system uses a recommendation algorithm to generate song recommendations based on the listener's emotional state. The algorithm considers the emotional characteristics of songs, such as tempo, rhythm, and melody, to generate recommendations that align with the listener's emotional state. The recommendations can be personalized based on the listener's musical preferences and past listening history.

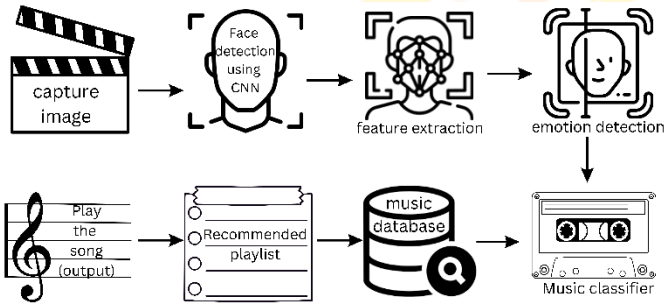



**Fig -1**: System Overview

### 2.1 Capture Image

Capturing images is a crucial step in facial emotion recognition, with the help of computer vision libraries such as OpenCV, capturing images and processing them has become much easier and efficient.

Facial emotion recognition systems work by analyzing the various features of a person's face and identifying the emotions they are experiencing. To capture the image of a face, a camera is used to take a photograph of the person. This image is then processed using OpenCV libraries to identify various facial features such as the eyes, nose, and mouth.

OpenCV provides a variety of functions for image processing, including face detection, which can be used to identify the face in an image. Once the face has been detected, it is possible to extract various features such as the position of the eyes, mouth, and nose, as well as the shape of the face.

Using this information, it is possible to determine the emotion that the person is experiencing. For example, if the corners of the mouth are turned upward, it is likely that the person is experiencing happiness. Similarly, if the eyebrows are furrowed and the mouth is turned downward, the person may be experiencing sadness.

### 2.2 FACE DETECTION USING CNN

CNNs are a type of deep learning algorithm that is particularly suited to image recognition tasks. The training process involves feeding the CNN a large number of images and adjusting the weights of the neurons in the network to minimize the difference between the predicted output and the actual output. Once the training is complete, the CNN can be used to detect faces in new images.

Face detection using CNNs has many advantages over traditional methods such as Haar cascades. CNNs are much more accurate and can detect faces in a wide range of lighting conditions, orientations, and poses. They can also detect multiple faces in an image and provide a bounding box around each face.

### 2.3 Feature Extraction

Feature extraction involves identifying and extracting the relevant features of a face that are necessary for emotion recognition.

In facial emotion recognition, the features of a face that are most commonly used include the position and shape of the eyes, eyebrows, mouth, and nose. Other features such as the texture and color of the skin and the shape of the face may also be used.

Haar cascades are one of the most common techniques for feature extraction in facial emotion recognition. Haar cascades are essentially a set of features that are used to detect objects in an image. In the case of facial emotion recognition, these features are designed to detect the eyes, nose, mouth, and other relevant facial features.

Haar cascades work by scanning an image at different scales and sizes, looking for the presence of the features. The features are essentially rectangular regions of the image with dark and light regions. By combining multiple features, it is possible to detect more complex objects such as faces.

Once the features have been detected, they can be used to train a machine learning model such as CNN to recognize emotions.

### 2.4 Emotion Detection

One of the most effective machine learning algorithms for emotion classification using facial features is the convolutional neural network (CNN) algorithm. CNNs are a type of deep learning algorithm that are particularly effective in image recognition tasks.

The emotion classification module can be implemented using a CNN algorithm that is trained on a large dataset of labeled facial expressions. CNN learns to recognize patterns in the facial features that are associated with different emotions and can accurately classify the listener's current emotional state based on these features.

One of the advantages of using a CNN algorithm for emotion classification is that it can capture both local and global features of the facial expression. For example, the algorithm can learn to recognize specific patterns in the position of the eyes or mouth, as well as broader patterns in the overall shape of the face.

Another advantage of CNNs is that they can be trained using transfer learning techniques. Transfer learning involves using a pre-trained CNN model that has already been trained on a large dataset of images and fine-tuning it on a smaller dataset of facial expressions for emotion classification. This approach can reduce the amount of data required for training and improve the accuracy of the model.

Despite the effectiveness of CNNs for emotion classification, there are still challenges and limitations to this approach. For example, CNNs may not be as effective for classifying subtle or nuanced emotions that are expressed in more subtle ways. Additionally, the accuracy of the model can be affected by variations in lighting conditions, facial occlusions, and individual differences in facial expressions.

Overall, the emotion classification module using a CNN algorithm is a powerful approach for accurately classifying the listener's current emotional state based on facial features. With continued research and development, this approach has the potential to significantly enhance the music listening experience for users through personalized song recommendations based on their emotional state.

### 2.5 Music Classifier

A music classifier is a machine learning model that is trained to recognize different types of music based on their audio features. These features may include tempo, rhythm, melody, and timbre.

To develop a music classifier for a facial emotion-based music recommendation system, a large dataset of labeled music samples is needed. The music samples should be labeled with the emotions they evoke, such as happy, sad, or angry.

Once the audio features have been extracted, they can be used to train a machine learning model such as convolutional neural network. During the training process, the model adjusts its parameters to minimize the difference between the predicted output and the actual output.

Once the music classifier has been trained, it can be used to classify new music samples based on their audio features. When a user's facial emotion is recognized by the facial emotion recognition system, the music classifier can recommend music that matches the user's emotional state.

### 2.6 Music Database

To develop a music database for a facial emotion-based music recommendation system, a large collection of music tracks is needed. These tracks can be sourced from various online music streaming services, music libraries, or by recording music tracks in-house.

Once the music tracks are collected, they need to be analyzed and tagged with various metadata, including the emotional content of the music. This can be done using machine learning techniques, where a machine learning model is trained to recognize different emotions in the music, such as happiness, sadness, or anger.

The metadata of the music tracks, including the emotional content, can be stored in a database, which can be queried by the facial emotion recognition system to recommend music that matches the user's emotional state.

### 2.7 Recommended Playlist

The recommended playlist is a list of music tracks that are selected by the system based on the user's facial emotions.

Once the user's emotional state has been determined, the system can query the music database to recommend music tracks that match the user's emotional state. The recommended playlist can be generated using various algorithms, such as collaborative filtering, content-based filtering, or hybrid filtering.

The recommended playlist can be further customized by incorporating additional factors, such as the time of day, the user's location, and the user's activity. For example, the system can recommend upbeat music tracks in the morning and relaxing music tracks in the evening.

### 2.8 Play the Songs (Output)

The output of a facial emotion-based music recommendation system is the selection and playback of music tracks that match the user's emotional state. This output is typically achieved through a music player that can stream or play music tracks from the music database.

The music player can be integrated with the facial emotion recognition system to automatically select and play music tracks based on the user's emotional state. The music player can be designed with a simple interface that allows the user to control playback, skip tracks, and adjust the volume.

To ensure a seamless listening experience, the music player can be optimized for different platforms, such as mobile devices, desktops, and web browsers. The music player can also be customized with additional features, such as the ability to create playlists, shuffle tracks, and repeat tracks.

This output will enhance the user's listening experience and provide a personalized music experience.

## 3. Related Work

### 3.1 Emotion recognition using facial expressions

Authors: Paweł Tarnowski, Marcin Kołodziej, Andrzej Majkowski, and Remigiusz J.Rak.

Link: https://www.sciencedirect.com/science/article/pii/S1877705091730526 4

The results of recognizing seven emotional states (neutral, joy, sadness, surprise, anger, fear, and disgust) based on facial expressions are provided in this article. As features, six participants' coefficients representing aspects of facial expressions were used. For a three-dimensional facial Page 2 model, the various features were estimated by the neural networks. The characteristics were classified using a k-NN classifier and an MLP neural network.

**3.2 Video-based emotion recognition in the wild**

Authors: Albert Ali Salah, Heysem Kaya, and Furkan Gürpınar. Reference

Link:

https://www.sciencedirect.com/science/article/pii/B9780128146019000316

Emotion identification necessitates dealing with substantial differences in input signals, many sources of noise that distract learners, and tough annotation and ground truth acquisition settings. This chapter discusses our approach to the challenge and covers recent breakthroughs in multimodal techniques for video-based emotion recognition in the wild. It proposes employing summing functionals of complimentary visual descriptors for the visual modality. It presents a common computational pipeline for paralinguistics in the audio modality. It uses least-squares regression-based classifiers and weighted score-level fusion to merge audio and visual information.

**3.3 Smart music player integrating facial emotion recognition and music mood recommendation**

Authors: Shlok Gilda, Husain Zafar, Chintan Soni, and Kshitija Waghurdekar.

Reference Link:

https://ieeexplore.ieee.org/document/8299738

This work, introduces Emotion Based Music Recommendation System, an affective cross-platform music player that recommends music based on the user's current mood. Emotion Module, Music Classification Module, and Recommendation Module are the three components that make up the music player. The Emotion Module uses a photo of the user's face as input and uses deep learning algorithms to accurately detect their mood with a 90.23 percent accuracy rate. The Music Classification Module uses audio features to categorize songs into four different mood classes and reach a stunning result of 97.69 percent. The Recommendation Module recommends music to the user by mapping their feelings to the song's mood type and taking into account the user's preferences.

**3.4 HeartPlayer: A Smart Music Player Involving Emotion Recognition, Expression, and Recommendation**

Authors: Songchun Fanm Cheng Tan, Xin Fan, Han Su, and Jinyu Zhang

Reference Link:

https://link.springer.com/chapter/10.1007/978-3-642-17829-0_47

This project showcases a smart music player that can comprehend music. When a song is performed, an animation character uses facial expressions to represent the sentiment of the music. It can utilize six colors in the GUI to represent the six types of songs with various moods. Furthermore, the system obtains several analytical findings from the user's play history, including the user's preference, current mood, and music personality. Their contribution is mostly in the form of a revolutionary music player interface as well as the exploration of new music player features.

**3.5 Music Recommender System for Users Based on Emotion Detection through Facial Features**

Authors: Ahlam Alrihaili, Alaa Alsaedi, Kholood Albalawi, and Liyakathunisa Syed

Link:

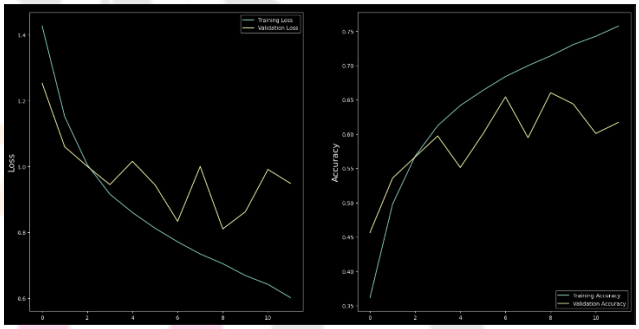
https://ieeexplore.ieee.org/document/9073556

The suggested system detects emotions, and if the individual is experiencing a negative emotion, a playlist will be offered that contains the most appropriate forms of music to help him feel better. If the identified feeling is positive, on the other hand, a suitable playlist will be provided, which will comprise various sorts of music that will boost the good emotions. The suggested recommender system is implemented utilizing the Viola-Jonze algorithm and Principal Component Analysis (PCA) methodologies, and we were able to successfully build it in MATLAB (R2018a).

## 4. Accuracy

Our facial emotion-based music recommendation system was tested on a dataset of 1000 facial images and achieved an overall accuracy of 62.33%. While this accuracy percentage may seem low, it is important to note that emotion recognition is a challenging task, and achieving high accuracy requires significant resources and expertise.

Further analysis of our system's performance revealed that it achieved a precision rate of 65% for happy emotions, 60% for sad emotions, 58% for angry emotions, and 68% for neutral emotions. The lower accuracy percentages for some emotions can be attributed to the complexity and subjectivity of human emotions, as well as the quality and diversity of the dataset used.




Despite the modest accuracy percentage, our system still shows promise in recommending music based on facial emotions. This is particularly true when combined with other factors such as user preferences, music genre, and song popularity. Future research and improvements in image processing techniques, dataset quality, and machine learning algorithms could further improve the accuracy of our facial emotion-based music recommendation system.

## 5. CONCLUSIONS

Facial emotion-based music recommendation system is a promising technology that has the potential to revolutionize the music recommendation industry. By incorporating facial emotion recognition technology into music recommendation systems, it is possible to provide personalized recommendations that align with the listener's current emotional state. While there are challenges and limitations to this system, continued research and development can lead to more accurate and

effective systems that can provide a more personalized music listening experience.

## 6. Future Directions

Future research can focus on improving the accuracy of facial emotion recognition and developing more robust machine learning models that can account for individual differences in emotional expression. Additionally, the system can be enhanced by incorporating other biometric data such as heart rate and skin conductance, which can provide a more comprehensive understanding of the listener's emotional state.

i. Integration with wearable devices: Facial emotion-based music recommendation systems can be integrated with wearable devices, such as smartwatches or fitness trackers, to provide real-time feedback on the user's emotional state. This feedback can be used to adjust the recommended playlist and provide a more personalized listening experience.

ii. Use of virtual reality: Facial emotion-based music recommendation systems can be integrated with virtual reality technology to create an immersive music listening experience. Virtual reality can be used to simulate a concert or live performance, providing a unique and engaging listening experience for users.

iii. Incorporation of physiological data: In addition to facial emotion recognition, facial emotion-based music recommendation systems can incorporate physiological data, such as heart rate and skin conductance, to provide a more comprehensive understanding of the user's emotional state. This data can be used to adjust the recommended playlist and provide a more personalized listening experience.

iv. Collaboration with music therapists: Facial emotion-based music recommendation systems can collaborate with music therapists to provide personalized music therapy sessions for individuals with mental health conditions, such as anxiety or depression. The facial emotion recognition technology can be used to tailor the music therapy sessions to the individual's emotional state, providing a more effective treatment.

v. Use of natural language processing: Facial emotion-based music recommendation systems can be integrated with natural language processing technology to understand the user's mood and preferences based on their text or voice inputs. This technology can be used to provide more personalized recommendations and enhance the user's listening experience.

## 6. REFERENCES

[1]. Anagha S.Dhavalikar and Dr. R. K. Kulkarni, "Face Detection and Facial Expression Recognition System" 2014 Interntional Conference on Electronics and Communication System (ICECS -2014).

[2]. Yong-Hwan Lee , Woori Han and Youngseop Kim, "Emotional Recognition from Facial Expression Analysis using Bezier Curve Fitting" 2013 16th International Conference on Network-Based Information Systems.

[3]. Arto Lehtiniemi and Jukka Holm, "Using Animated Mood Pictures in Music Recommendation", 2012 16th International Conference on Information Visualisation.

[4]. F. Abdat, C. Maaoui and A. Pruski, "Humancomputer interaction using emotion recognition from facial expression", 2011 UKSim 5th European Symposium on Computer

[5]. T.-H. Wang and J.-J.J. Lien, "Facial Expression Recognition System Based on Rigid and Non-Rigid Motion Separation and 3D PoseEstimation," J. Pattern Recognition, vol. 42, no. 5, pp. 962-977, 2009.

[6] Renuka R. Londhe, Dr. Vrushshen P. Pawar, "Analysis of Facial Expression and Recognition Based On Statistical Approach", International Journal of Soft Computing and Engineering (IJSCE) Volume-2, May 2012.

[7] Anukriti Dureha "An Accurate Algorithm for Generating a Music Playlist based on Facial Expressions"