International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 12 Issue: 02 | Feb 2025 www.irjet.net p-ISSN: 2395-0072

Music Recommendation Based on Facial Recognition

Rimsy Dua[1], Kunvar Bir Pratap Singh[2], Abhishek Gupta[3]

[1]Assitant Professor, Thakur College of Science and Commerce, Maharashtra, India
[2,3] Bsc-IT UG, Thakur College of Science and Commerce, Maharashtra, India

---

Abstract - *Users feel more satisfied with music recommendation systems because the systems tailor their features to personal tastes which affect how users emotionally respond to music. Music recommendation systems in use fail to track user emotional states because they operate using collaborative filtering and user preference and listening records alone. Deep learning served as the technology that made possible the development of emotion-targeted music recommendations through artificial intelligence. The current technological advances permit facial recognition systems to monitor human emotional responses thus enabling automated music selection for users. The Convolutional Neural Network analyzes recorded facial expressions and identifies emotions between Happy and Sad as well as Angry through Neutral states to ensure content-based filtering obtains suitable music outputs. A CNN model reaches higher accuracy once trained with FER2013 and CK+ dataset resources. Facial recognition features in the system are enabled through OpenCV backend operations which also run Python Flask and HTML CSS and JavaScript interfaces simultaneously. The system implements a built-in chatbot functionality which lets users convey their emotional states through emojis for system input selection. User actions are recorded by administrator dashboards to obtain ongoing feedback that helps system developers discover user emotions during system performance improvement efforts. System-based emotion detection happens in real-time leading to satisfied users and providing groundwork for AI-based entertainment system development.*

Key Words: **Facial Recognition, Emotion Detection, Artificial Intelligence, Deep Learning, CNN, OpenCV, Content-Based Filtering and Python Flask, Chatbot Interaction and Emotionally Intelligent Applications.**

1. INTRODUCTION

Emotionally healthy individuals create feelings as psychological elements which influence brain reactions while strengthening motivational strength. Musical human tools provide key elements which allow people to express themselves and obtain healing powers while improving their performance at work. User profiles connect with preferred styles and previous listened music through automated recommendation generators to select the current music.Existing recommendation platforms fail to understand user emotions so they produce inadequate music suggestions.

The analysis of emotional states requires facial recognition technology because face expressions display human emotions. The market saw the arrival of artificial intelligence (AI) and deep learning technologies which developed facial emotion recognition 2013 (FER) technology to enhance user interactions in all industries with a strong impact on music recommendations systems. The automatic music recommendation system receives user emotional states by using facial recognition software which it uses to create recommendations based on received information.

Before the system can detect emotions Happy to Sad it requires processing webcam input using CNN but the system also includes detection of Angry to Surprised emotions as well as Disgusted Fearful and Neutral expressions. User songs become organized into frameworks within the database through emotional recognition which drives the recommendation process. The system provides time-based song suggestions based on user emotions using its automatic recommendation function.

The music generation tool in this platform features mental health-oriented support instruments that meet specific psychological requirements. User satisfaction with modern recommendation solutions increases because detection methods for users' emotions lead to improved evaluation assessments.

Research teams developed music recommendation platforms when they studied individual user solutions using machine learning together with deep learning techniques. The platform becomes more satisfactory to users thanks to emotional assessment capabilities that strengthen emotional connections of users to the system and platform throughout learning activities.

2. Background and Motivation

Users accessing the emotional observation platform can buy musical tracks designed to modify their emotional condition. Research on affective computing guides the implementation of emotional modifying features found in music tracks that assist in mental health protection and stress management. Users require song suggestion services that provide essential guidance for musical selection without interrupting the listening experience according to their active situation.

Originally recommendation engines started with user-listening information for content-based filtering before

researchers established collaborative filtering methods. The performance quality of recommendation systems increases when mood monitoring features are in place due to the way user system interface behavior changes based on different moods. Day shift personnel do not experience decreased emotional weight when hearing cheerful tones because they fail to understand their genuine emotional state.

Music recommendation platforms pair appropriate recommendations to users through the unification of their face emotion recognition capability together with emotion detection technology. The music recommendation software applies computer vision methods together with deep learning techniques to produce musical suggestions based on emotional data through emotional user interfaces directed to users. The system distributes both entertainment music and practical therapeutic interventions to military personnel through its psychological intervention technology capabilities.

Quality time data about user emotions allows commercial music agencies to choose songs that match individual emotional needs thereby improving their emotional state. Organizations that integrate AI entertainment characteristics now operate in the music business as the development of technical systems combines wellness options with digital services.

3. Research Objectives

- Our team must create a facial recognition technology able to recognize emotions while working in real-time scenarios.
- The system uses pre-established parameters to make emotion detections match music genres and tracks.
- The system promotes user interaction through an interactive chatbot which lets users communicate via emoji-based selections of their mood.
- Low-latency detection of emotions from the system performance needs optimization to create a smooth user experience.
- The system needs an accuracy and efficiency evaluation regarding its emotion detection capabilities while supplying suitable musical recommendations.

4. Literature Review

Effective performance through recommendations stems from machine-learning integration of collaborative and content-based filtering systems. The combination of artificial intelligence with deep learning techniques in modern scientific research led scientists to develop the best methods for detecting emotional music content through facial recognition. The scientific research in music recommendations attempts to combine FER technology with deep learning techniques for detection systems which leads to better system reliability and improved user experiences.

Facial Emotion-Based Music Recommendation Approaches:

The music player system which uses facial emotion identification features was developed by Professor R. B. Late together with Miss Harsha Varyani. The system began its operation by examining facial points of the tongue before moving to check points in the mouth before finishing at the recognition stage of emotions at the eyebrows. Addressing facial emotional expressions led the Facial emotion detection system to demonstrate automatic music recommendation capabilities. Research procedures require development to obtain facial recognition techniques that extract emotional features from detected faces so development of an effective emotional music recommendation system becomes possible.

An analysis explored the Automatic Music Player from Blue Eyes Technology that determined user emotions through its combination of facial recognition with mouse movement tracking features. The system utilizes functions from image processing and speech recognition and age-tracking methods to read emotions and enhance interactions between humans and computers. Cloud authentication together with storage provided the basis for the system which achieved efficient scalable emotion-aware music recommendation features that drive digital music recommendation systems forward.

Scientists have used research to investigate how facial expression recognition technology helps enhance mental health welfare. The analysis of facial expressions from large-scale datasets became the subject of research to support mental health observation. This research demonstrated that stress reduction and relaxation improvements happen through emotion-based musical recommendations specifically for young people dealing with mental health issues.

Deep Learning-Based Approaches:

Deep learning research units teamed up with vision-based processing researchers to create primary recommendation systems which integrate emotional feedback. CNNS sequence-based imaging analysis started by processing input data then it extracted features before achieving emotional classification through the last stage. The user's facial emotion data became accessible to the system which enabled it to generate suitable playlists. Medical studies and psychological care can build better detection systems for facial expressions through the modern facial tracking techniques of deep learning.

Face Emotion-Based Music Player developers achieved real-time facial expression detection through the use of CNNs for multiple emotional category grouping. The method known as webcam detection provided a practical solution since immediate results were possible even without needing physical monitoring equipment. Therapeutic activities with emotional music recommendation systems constitute fundamental components for clinical mental health service treatment delivered to therapeutic patients.

Researchers evaluated face detection abilities of web-based music recommendation services during their assessment phase. The model managed to identify emotions and provide appropriate song recommendations with 81% accuracy through evaluation with FER-2013 during its operational test phase. The high achievement rate of users with sound functions led research scientists to create wearable detection systems that monitor quick emotional changes.

A group of researchers examined facial emotional data-based music identification systems using deep learning approaches during recent past research years. Face emotion recognition achieved suitable music recommendations through its deep learning technology during processing of emotion-based user data following Approach 1. The emotional evaluation by face processing led to enhanced satisfaction and enhanced system interaction quality for recommendations.

AI-Driven Music Recommendation Systems:

The CNN training process differentiated emotional image labels successfully at a rate of 85% which led to an 83% correct identification rate of testing images. Through its automatic music recommendation programming solution users could find music instantly.

The developers established a seven-layer CNN architecture to develop their music discovery engine with 74.8% correct emotional identification performance. The main face recognition modules for playlist classification operations included integration of OpenCV together with Dlib. The system introduced new capabilities to identify user emotions while tracking abnormal system user activity patterns.

When MobileNet operated with Firebase enabled users to receive instant music recommendations after performing emotional assessments through their smartphones. These systems gained more customers through their introduction of facial recognition functionalities for emotional category detection.

The researchers conducted their emotional studies through preloaded songs but they excluded Spotify music service. Users can access instant music playback with private storage functionalities through this system because it operates without external API requirements. The system delivers reliable outputs from its ability to perform fast functions for self-contained applications that operate without needing Internet connections.

5. System Architecture

A modular system design enables the Music Recommendation System Based on Facial Recognition to execute its operations flexibly along with real-time performance along with scalability features. The system employs six critical components to handle operations starting from emotion detection to providing musical recommendations.

- Facial Emotion Recognition Module:

  The webcam system extracts visible facial representations to perform its operation. Animal or vegetable classification starts after OpenCV process the facial image with CNN deep learning models. Hetémnet evaluates seven emotional categories to identify Happy, Sad, Angry, Neutral, Surprised, Fear and Disgust expressions.

- Chatbot Module:

  Through the system users can communicate their feelings with different available emojis. Users maintain the ability to transmit emotional reactions when their face fails to produce correct readings.

- Emotion-Music Mapping Engine:

  Through the system associates observe detected emotions which help them identify musical attributes that relate to tempo speed ranges and various musical styles along with energy levels. The system controls its recommendation system based on rule-based mapping algorithms which operate dynamically.

- Music Recommendation Engine:

  The system chooses suitable music tracks from existing emotion-labeled databases according to the evaluated emotion output. An improved recommendation accuracy results from using metadata filtering methods.

- User Feedback System:

  The system obtains user evaluations about songs that match recommendations from its platform. The recommendation engine receives adjustments from user preferences which lengthen over time.

- Admin Dashboard:

  Monitors user engagement metrics, emotion trends, and system performance. Real-time analytics from the recommendation model can be adjusted through data monitoring.

5.1 Emotion Detection and Music Recommendation Process

A structured process operates within the facial emotion recognition system to identify emotions which triggers music recommendations.

- Face Detection:

  The system turns on the webcam function and records facial expressions from the user. A combination of OpenCV together with a pre-trained CNN model performs the extraction of facial features.

- Preprocessing:

  The model input dimensions guide the resizing process which follows grayscale conversion of the captured image. Normalization and noise reduction protocols enhance accuracy levels in the system.

- Emotion Classification:

  The processed image goes through the CNN model which provides a classification result from among the available seven emotional categories. Predictive values get assigned to all emotions recognized during detection.

- Music Recommendation:

  The system associates detected emotions through the following musical parameters:

  - Happy → Upbeat, energetic tracks.
  - Sad → Slow-tempo, calming music.
  - Angry → Intense, high-energy music.
  - Neutral → Background instrumental music.

  The system selects the most suitable musical tracks from the stored database.

- User Interaction & Feedback:

  Users have the opportunity to react to the system recommendations through their feedback. Through reinforcement learning methods the system automatically enhances its recommendation system.

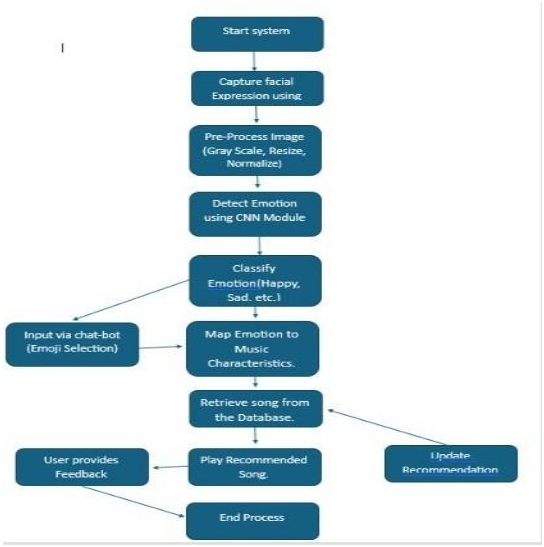


Figure 1: Flow Diagram of the System

6. Methodology

6.1 Data Collection

The system depends on three main datasets to perform both emotion detection and music recommendation functions effectively.

1. FER2013 Dataset:

   The database contains 35,887 labeled images which spread across seven different emotional categories. The data set contains a wide range of faces that appear under different lighting situations.

2. CK+ Dataset:

   The dataset contains 593 sequences showing emotional transition changes in visual data. The CNN model achieves improved accuracy because the researchers used this dataset to perform fine-tuning.

3. Music Dataset:

   This system contains selected songs that include mood-based metadata tags such as genre and intensity and tempo parameters. The recommendation engine uses music preference data to match detected emotions through this dataset.

6.2 Model Training & Optimization

1. CNN Architecture for Emotion Recognition:

   A Convolutional Neural Network (CNN) implemented a particular configuration for facial recognition as part of its operation.

- Input Layer: Accepts grayscale facial images (48x48 pixels).The first layers within this model extract basic image components which contain essential information about facial shapes and image contours and image edges along with image textures.Through pooling layers you can decrease images by reducing their resolution which leads to increased efficiency in calculation processes.After extracting the features they pass through Fully Connected Layers until receiving an emotion classification.

- Softmax Layer: Outputs a probability score for each emotion category.

Training Strategy

The data augmentation method strengthens dataset diversity through various image modifications which include rotational adjustments as well as flipping operations and brightness control functions. The model achieves low classification mistakes by using categorical cross-entropy as its loss function. Software implementation adopts Adam Optimizer as its optimization method to achieve quick learning speed during operations. The training time consisted of 25 batch elements while reaching convergence through 50 steps.

2. Music Recommendation Engine

The recommendation engine conducts a specific method which links detected emotions to appropriate mood-based musical selections.

Emotion	Recommended	Genre Tempo	Mood Adjustment
Happy	Pop, Dance	Fast	Energizing
Sad	Acoustic, Jazz	Slow	Soothing
Angry	Rock, Metal	High	Intense
Surprised	Classical, EDM	Varied	Dynamic
Fear	Ambient, Dark	Low	Calming

Emotion-to-Music Mapping Table

7. Experimental Results

7.1 Accuracy of Emotion Detection

The system obtained these classification outcomes through real-time user testing of system inputs:

Emotion	Accuracy (%)
Happy	85%
Sad	80%
Neutral	83%
Surprised	75%
Angry	78%
Fear	72%
Disgust	70%

The recognition system shows excellence in typical emotion detection of happiness and sadness and neutrality yet needs better extraction methods for fear and disgust emotion recognition.

7.2 System Performance

The system performance measurements centered on both latency and response time because it needed real-time functionality.

Task	Processing Time
Emotion Detection	~2 seconds
Music Recommendation	~1 second
Total System Response	~3 seconds

A real-time system enables a smooth user-friendly interface for the application.

7.3 Plotting Accuracy and Loss

Training and Validation Results:

The tests revealed important performance metrics by studying the changes in training and validation loss accuracy rates from several training sessions.

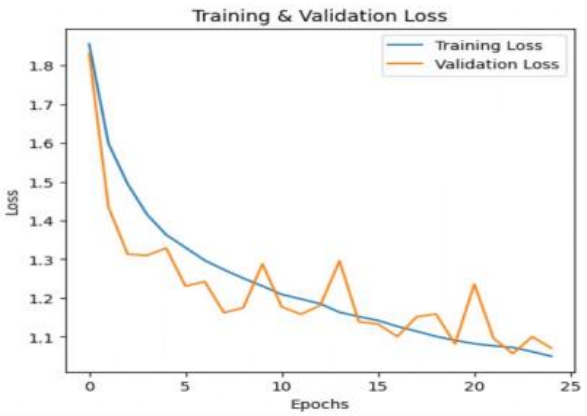


Figure 2: Training & Validation Loss

The above graph demonstrates that during training the model demonstrates efficient learning behavior because its training loss demonstrates steady reduction. Analogous to training loss rates the validation loss demonstrates

equivalent results with occasional minor fluctuations which originate from differing real sample data.

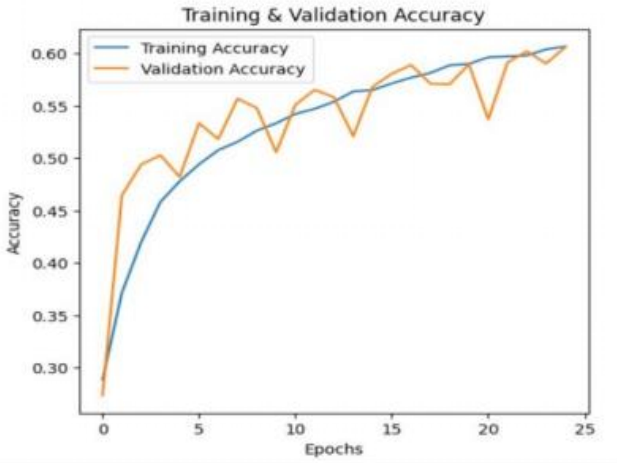


Figure 3: Training & Validation Accuracy

The above graph demonstrates that the accuracy performance reaches above 60% when the model finishes its 25th epoch run in its separate training and validation runs. Evaluation accuracy confirms no signs of overfitting because it maintains identical performance to training accuracy through the whole process.

Figure 4: Confusion Matrix

The confusion matrix functions as a tool to determine what level of success exists when the model assigns different emotional categories correctly.

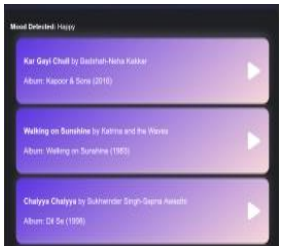
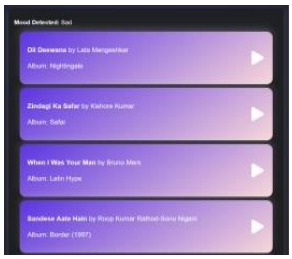
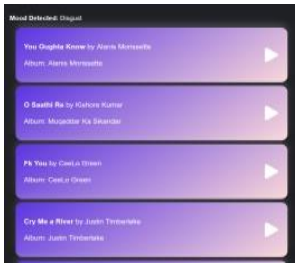
- The system achieves its leading performance results when detecting both Happy and Neutral emotions.
- The expressions of Fear and Disgust generate classification errors that lead to the incorrect identification of similar facial expressions.
- The model exhibits irregular distinctions between Fear and Angry emotions as well as Sad and Neutral categories due to a need for improved emotion variable definition.

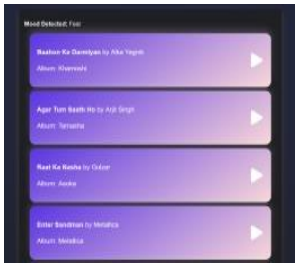
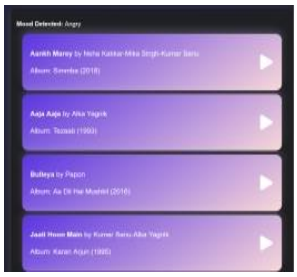
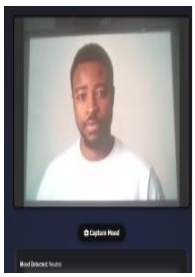
The model demonstrates good results in detection while additional components will boost its capacity to recognize intricate emotional states.

7.4 User Feedback & Satisfaction

User feedback gathered information about the usefulness of music recommendations for user experience assessment. The majority of people (70%) found the suggested music very fitting to their needs. 20% found them somewhat relevant. The emotions of users did not match their recommendations in Fear and Disgust categories according to 10 percent of respondents.

The following table demonstrate different expression faces which detect emotions to suggest specific music tracks:

Mood Detected	Face Expression	Music Recommended
Happy		
Sad		
Disgust		
Surprise		

Fear		
Angry		
Neutral		

8. Conclusion

The implementation of artificial intelligence music recommendation services with facial emotion recognition technology enables the development of Music Recommendation System Based on Facial Recognition for personalized musical services. User experiences receive improved benefits because this system automatically updates music content through emotional detection over single-user preference systems. The precise seven-emotion classification occurs from using FER2013 and CK+ dataset-trained CNN models in combination for effective emotion-to-music mapping. The system detects mood success to activate a recommendation protocol which connects mood-compatible tracks to their musical tempo and genre together with their emotional characteristics to the user's detected feelings.

The system contains an interactive chatbot accessible by users during failed facial recognition that lets them express emotions through emojis as an alternative option. The system employs a Flask-based backend together with easy-to-use web-based frontend for smooth user interface and enables scaling for real-world applications. Reinforcement learning approaches enhance system performance through user feedback because they enable automatic adaptation of recommendations via system-user interactions and preferences selection.

The system detects regular emotions with 77% precision through its identification process focused on Happy, Neutral and Sad expressions. Real-time operation is achieved through the system due to its 3-second average response period. Most users approved the new emotional recommendation system because they found its suggested content appropriate although they identified matters with insufficient song selection and restricted ability to match emotion types like Fear among multiple complex emotions.

Real-time operational implementation becomes possible through the system because its flexible music recommendation allows development of AI music streaming services and digital assistant technology along with personalized media applications. The emotion-aware recommendation system enables innovative development approaches that unify human-computer interaction with artificial intelligence in entertainment system interfaces.

REFERENCES

- Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press. www.deeplearningbook.org
- Ko, B. C. (2018). The paper examines the performance of techniques which recognize facial expressions from visual information. Sensors, 18(2), 401.www.mdpi.com/1424-8220/18/2/401
- Lopes, A. T., de Aguiar, E., De Souza, A. F., & Oliveira-Santos, T. (2017). "Facial expression recognition with convolutional neural networks." Pattern Recognition, 61, 610628.www.sciencedirect.com/science/article/abs/pii/S0031320316303489
- 4.He, K., Zhang, X., Ren, S., & Sun, J. (2016). "Deep residual learning for image recognition." IEEE Conference on Computer Vision and Pattern Recognition (CVPR).www.arxiv.org/abs/1512.03385
- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., & Fei-Fei, L. (2015). "ImageNet large scale visual recognition challenge." International Journal of Computer Vision, 115(3), 211-252.www.arxiv.org/abs/1409.0575
- Poria, S., Cambria, E., Bajpai, R., & Hussain, A. (2017). "A review of affective computing: From unimodal analysis to multimodal fusion." Information Fusion, 37, 98-125. www.sciencedirect.com/science/article/abs/pii/S1566253517300350
- Delbouys, R., Hennequin, R., Piccoli, F., Moussallam, M., & Richard, G. (2018). The deep neural network-based system detects music moods through analysis of audio content and textual lyrics. ISMIR Conference. www.ismir2018.ircam.fr/doc/pdfs/227_Paper.pdf