

# Emotion Detection and Music Recommendation System Based on User Facial Expression

PRADEEP GOSWAMI<sup>1</sup>, RIDDDHI GARE<sup>2</sup>, SURBHI TRIPATHI<sup>3</sup>, UDDESH DWIVEDI<sup>4</sup>

<sup>1, 2, 3, 4</sup> Dept of Computer Science, Shri Ram Institute of Technology, Jabalpur, India

*Abstract— It is frequently confounding for a individual to choose which music he/she have to tune in from a massive collection of existing alternatives. There have been a few proposal frameworks available for issues like music, eating, and shopping depending upon the disposition of client. The main objective of our music proposal framework is to give proposals to the clients that fit the user's inclinations. The investigation of the facial expression/user feeling may lead to understanding the current enthusiastic or mental state of the client. Music and recordings are one region where there is a noteworthy chance to endorse plenteous choices to clients in light of their slants and moreover recorded data. It is well known that people make utilize of facial expressions to express more clearly what they need to say and the setting in which they meant their words. More than 60 percent of the clients accept that at a certain point of time the number of melodies show in their tunes library is so huge that they are incapable to figure out the song which they have to play. By creating a proposal framework, it seem help a client to make a choice with respect to which music one ought to tune in to making a difference the client to diminish his/her stress levels. The client would not have to squander any time in looking or to see up for songs and the best track coordinating the user's temperament is identified, and melodies would be appeared to the user according to his/her temperament. The picture of the client is captured with the offer assistance of a webcam. The user's picture is taken and at that point as per the mood/emotion of the client an fitting melody from the playlist of the client is appeared coordinating the user's prerequisite*

*Index Terms- Face Recognition, Feature extraction, Emotion detection, Convolutional Neural Network, Pygame, Tkinter, Music, Player, Camera.*

## I. INTRODUCTION

Individuals tend to express their feelings, primarily by their facial expressions. Music has continuously been known to change the disposition of an person[1]. Capturing and recognizing the feeling being voiced by a person and showing fitting melodies coordinating the one's disposition and can progressively calm the mind

of a client and by and large conclusion up giving a satisfying impact. The venture points to capture the emotion expressed by a individual through facial expressions. A music player is planned to capture human emotion through the web camera interface accessible on computing frameworks. The computer program captures the image of the client and at that point with the offer assistance of picture division and picture preparing techniques extracts highlights from the confront of a target human being and tries to identify the feeling that the individual is trying to express. The venture points to help the disposition of the client, by playing melodies that coordinate the requirements of the client by capturing the picture of the client. Since old times the best frame of expression examination known to mankind is facial expression acknowledgment. The best conceivable way in which individuals tend to analyze or conclude the feeling or the feeling or the contemplations that another person is attempting to express is by facial expression. In a few cases, temperament modification may too offer assistance in overcoming circumstances like sadness and pity. With the help of expression investigation, numerous health risks can be dodged, and moreover there can be steps taken that offer assistance brings the disposition of a client to a better stage.

## II. LITERATURE REVIEW

They proposed a paper which centered on the think about of changes in the curvatures of the confront and the force of the comparing pixels. The creator utilized Counterfeit Neural Networks(ANN), which was utilized to classify the feelings. The creator moreover proposed various approaches for a playlist. Zheng et al. [2]proposed two noteworthy categories for facial feature extraction, which included Appearance-based highlight extraction and geometric based feature extraction, which included extraction of a few fundamental focuses of the confront such as mouth, eyes, and determines the attitude of the client by

utilizing facial expression People often express their feeling by their expressions, hand signals, and by raising the voice of tone but mostly humans express their sentiments by their confront. Emotion-based music player decreases the time complexity of the client. For the most part, individuals have a expansive number of melodies on their playlist. Playing melodies randomly does not fulfill the temperament of the client. This framework makes a difference client to play tunes consequently concurring to their temperament. The picture of the client is captured by the web camera, and the pictures are spared. The images are to begin with changed over from RGB to parallel organize. This handle of speaking to the information is called a feature-point discovery strategy. This handle can moreover be done by utilizing Haar Cascade technology provided by Open CV. The music player is created by utilizing a java program. It oversees the database and plays the melody agreeing to the disposition of the inquired about various advances in human influence acknowledgment. He centered on different approaches that can handle sound and/or visual recordings of shows of full of feeling states. The paper gives a point by point audit of audio/visual computing strategies[3]. The impact is portrayed as a model of feeling categories which include happiness, pity, fear, outrage, appall, and shock. This paper centered on talking about the challenges in computing strategies for the improvement of programmed, unconstrained influence recognizer, which helped in feeling location. It moreover distinguished a few issues that have been missed or dodged in unimodal posed feeling acknowledgment. Parul Tambe proposed an thought which computerized the interactions between the clients and music player, which learned all the inclinations, feelings and exercises of a user and gave melody determination as a result. The different facial expressions of clients were recorded by the device to decide the feeling of the client to anticipate the class of the music[5]. Jayshree Jha et al. proposed an emotion-based music player utilizing picture preparing This appeared how various algorithms and methods that were proposed by diverse creators in their investigate may be utilized for connecting the music player along with human feelings. It has hence made a difference in lessening the endeavors of user in making and overseeing playlist and giving an great involvement to the music audience members by

bringing them the most appropriate tune concurring to the user's his/her current expression. Anukritine et al. came up with an calculation that gives a list of tunes from the user's playlist in agreement with the user's feeling. The calculation which was outlined was centered on having less computational time and too in this way decreases the fetched included in utilizing different equipment. The fundamental thought was to isolate the emotions into five categories i.e., Bliss, pitiful, outrage, astonish and fear too given a exceedingly accurate audio data recovery approach that extricated significant data from an sound flag in less time. Aditya et al. created an android application which acts as a customized music player for a user utilizing picture preparing to analyze and show melodies to client agreeing to the user's disposition. The application was created utilizing Obscure and moreover OpenCV to actualize facial acknowledgment algorithms. This paper too appeared comparison between different calculations utilized in facial discovery. The images of the client were captured utilizing the front camera of the portable gadget. It pointed to give satisfaction to music partners by extricating their feelings. A. Habibzad et al. [6] proposed a unused calculation is proposed to recognize the facial feeling, which included three stages: pre-processing, feature extraction, and classification. The to begin with portion portrays different stages in picture preparing include preprocessing, sifting utilized to extricate different facial highlights. The moment portion optimized the eye and lip circle characteristics, and in the third portion, the eye and lip ideal parameters were utilized to classify the feelings. The gotten comes about appeared that the speed of facial acknowledgment was distant superior than other usual approaches. Prof. Nutan Deshmukh et al. [7] centered on making a framework that brings the emotion of the client utilizing a camera and at that point computerizes the result utilizing the feeling detection algorithm. This calculation captures the temperament of the client after each chosen interim of time as the mood of the client may not be the same after a few time; it may or may not alter. The proposed algorithm on an normal calculated estimation takes around 0.95-1.05 sec to create an emotionbased music framework, which was way better than past existing calculations and decreases the fetched of designing. Chang Liuet al[9] portrayed a framework that makes utilize of Brain-

Computer Interfacing, also called as BCI. BCI makes utilize of gadgets to send signals to the preparing frameworks. EEG equipment is used in to screen the person's cognitive state of intellect. The downside of the plot is that they require the input from the user's brain ceaselessly to perform the classification. An calculation based on MID is utilized to ceaselessly screen and prepare the signals gotten from the brain of the client and use these signals to effectively screen and create feelings that the client is right now experiencing. Swati Vaid etal checked on EEG - Electroencephalography (EEG) is a shape of restorative science that records the electrical action from the neurons of brain cells[8]. The electrical movement of the neurons from inside the cells of the brain is enrolled. Based on the recorded action of the neurons an approximation is made, and the feeling of the individual is assessed from that investigation. This method mentioned over, in spite of the fact that serves the reason of getting the action of brain cells but comes up short to serve the purpose of movability and financial matters.

#### A. HYPOTHESIS

Here's a breakdown of hypotheses for an emotion detection and music recommendation system based on user facial expressions:

- **Emotion Detection Accuracy:** The system's accuracy in detecting facial expressions and mapping them to corresponding emotions will positively correlate with user satisfaction and engagement.
- **Music Preference Correlation:** There is a correlation between specific facial expressions and certain music genres or songs. For example, happy expressions might correlate with upbeat music genres like pop or dance, while sad expressions might correlate with slower, more melancholic music genres like blues or ballads.
- **Personalization Effectiveness:** Personalized music recommendations based on detected emotions will lead to higher user satisfaction compared to generic recommendations.
- **User Engagement:** Users will interact more frequently and spend more time using the system if the music recommendations align well with their emotional state.

- **Long-Term Impact:** Regular use of the system and receiving music recommendations aligned with emotional states could potentially impact users' mood regulation and overall emotional well-being positively.
- **Cultural Variability:** Facial expressions and their corresponding emotions might vary across cultures, impacting the effectiveness of the system's emotion detection and music recommendation algorithms.
- **Feedback Loop:** Users' feedback on the accuracy and relevance of the music recommendations will help improve the system's algorithms over time, enhancing overall performance and user satisfaction.

These hypotheses can guide the development and evaluation of the emotion detection and music recommendation system, ensuring its effectiveness and user satisfaction.

### III. PROPOSED METHODOLOGY

The proposed framework can identify the facial expressions of the client and based on his/her facial expressions extricate the facial points of interest, which would at that point be classified to get a specific feeling of the client. Once the feeling has been classified the melodies coordinating the user's feelings would be shown to the user. **Emotion Extraction Module** -The picture of the client is captured with the offer assistance of a camera/webcam. Once the picture captured, the outline of the captured picture from webcam feed is changed over to a grayscale picture to move forward the execution of the classifier, which is used to distinguish the confront display in the picture. Once the transformation is total, the picture is sent to the classifier calculation which, with the offer assistance of highlight extraction methods can extract the confront from the outline of the web camera nourish. From the extricated confront, person features are gotten and are sent to the prepared arrange to identify the feeling communicated by the user. These pictures will be utilized to prepare the classifier so that when a totally modern and unknown set of pictures is displayed to the classifier, it is able to extricate the position of facial landmarks from those pictures based on the information that it had as of now procured from

the preparing set and return the arranges of the modern facial points of interest that it recognized. The organize is trained with the offer assistance of CK broad information set[4]. This is utilized to distinguish the feeling being voiced by the user.

- Sound Extraction Module - After the feeling of the client is extricated the music/audio based on the feeling voiced by the client is shown to the client, a list of melodies based on the emotion is shown, and the client can tune in to any melody he/she would like to. Based on the regularity that the client would tune in to the tunes are shown in that arrange. This module is developed using web advances like PHP, MySQL, HTML, CSS, JAVASCRIPT.
- Feeling - Sound Integration Module - The feelings which are extricated for the melodies are stored, and the tunes based on the feeling are shown on the web page built utilizing PHP and MySQL. For case, if the feeling or the facial highlight is categorized beneath cheerful, then songs from the upbeat database are shown to the client.

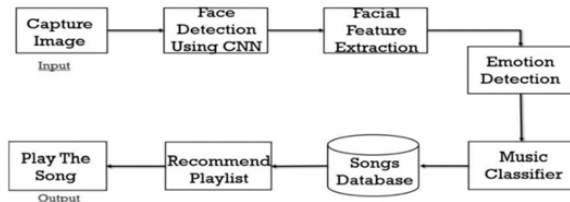


Figure 1. Methodology Flowchart

A. DATA COLLECTION AND DATA SOURCE

A data was collected from users based on 3 parameters which are,

1. What type of songs would they want to listen to when they are happy?
2. What type of songs would they want to listen to when they are sad?
3. What type of songs would they want to listen to when they are angry?

- CK+ Dataset - The dataset that is used to train the classifier is Cohn Kanade Extended dataset. The dataset consists of 593 Facial Action Coded Sequences from 123 subjects. The labeling done is to tell us about the expression that is being expressed by the subject. There are a series of images that start from the neutral expression of the

target and end with the extreme emotion expressed by the subject. In our analysis and training network, the first image and the last image are used. The other images are not used as such to train the network.

- HELEN Dataset - The dataset contains around 200 images are used for the training of the classifier. Along with the images, 164 landmarks positions are provided for each and every image present in the dataset in the form of a .txt file. These coordinates are extracted from the text files and used in the system to generate the .xml file. This XML file is further used to train the classifier. This trained classifier is used to predict the position of the landmarks in the other set of unknown images.

IV. IMPLEMENTATION

- i. Facial Expression Recognition: Use a library like OpenCV or TensorFlow to detect and recognize facial expressions in real-time or from images.
- ii. Emotion Classification: Train a machine learning model to classify the detected facial expressions into different emotions such as happiness, sadness, anger, etc. You can use deep learning techniques like Convolutional Neural Networks (CNNs) or pre-trained models like OpenFace or VGG-Face for this purpose.
- iii. Music Recommendation: Once the emotion is detected, you can map each emotion to a specific genre or mood of music. Then, recommend music from that genre or mood. You can use APIs like Spotify's Web API to fetch music recommendations based on genres or mood.

Here's a basic outline of how you could implement this in Python:

Python

```

# 1. Facial Expression Recognition
import cv2
# Load pre-trained face detector
face_cascade = cv2.CascadeClassifier('haarcascade_frontalface_default.xml')
  
```

```

# 2. Emotion Classification
# Train or load a pre-trained model for emotion
classification

# 3. Music Recommendation
# Define a function to recommend music based on
detected emotion
def recommend_music(emotion):
    if emotion == 'happy':
        # Recommend happy/upbeat music
    elif emotion == 'sad':
        # Recommend sad/melancholic music
    elif emotion == 'angry':
        # Recommend energetic/aggressive music
    # Add more conditions for other emotions

# Main loop for capturing and processing frames
while True:
    # Capture frame from camera
    ret, frame = cap.read()

    # Detect faces in the frame
    faces = face_cascade.detectMultiScale(frame,
scaleFactor=1.1, minNeighbors=5, minSize=(30, 30))

    for (x, y, w, h) in faces:
        # Crop face from frame
        face_roi = frame[y:y+h, x:x+w]

        # Perform emotion classification on the cropped
face
        emotion = classify_emotion(face_roi)

        # Recommend music based on detected emotion
recommend_music(emotion)

        # Draw bounding box around the face and display
emotion
        cv2.rectangle(frame, (x, y), (x+w, y+h), (255, 0,
0), 2)
        cv2.putText(frame, emotion, (x, y),
cv2.FONT_HERSHEY_SIMPLEX, 1, (0, 255, 0), 2,
cv2.LINE_AA)

        # Display the resulting frame
cv2.imshow('Emotion Detection and Music
Recommendation', frame)

    # Break the loop if 'q' is pressed

```

```

if cv2.waitKey(1) & 0xFF == ord('q'):
    break

```

```

# Release the capture
cap.release()
cv2.destroyAllWindows()
Remember, this is a basic outline, and you may need
to fine-tune the parameters and add more complexity
depending on your specific requirements and the
performance you're aiming for. Additionally, ensure
you handle errors and edge cases gracefully.

```

## V. RESULT

This paper proposes a music suggestion framework which extricates the picture of the client, which is captured with the offer assistance of a camera connected to the computing stage. Once the picture has been captured, the captured outline of the picture from webcam nourish is at that point being changed over to a grayscale image to progress the execution of the classifier that is utilized to distinguish the confront display in the picture. Once the transformation is total, the picture is sent to the classifier calculation which, with the help of highlight extraction strategies is able to extricate the confront from the outline of the web camera feed. Once the confront is extricated person highlights from the confront is extricated and is sent to the trained network to distinguish the feeling communicated by the client. A classifier that is utilized to distinguish or get the facial points of interest from the confront of the client is prepared on HELEN dataset. HELEN dataset contains more than 2000 pictures. The picture of the client is captured with the offer assistance of a webcam. The user's picture is taken and at that point as per the mood/emotion of the client an suitable song from the playlist of the client is played coordinating the user's requirement. The framework has effectively been able to capture the feeling of a client. It has been tried in a realtime environment for this predicate. It has to be, be that as it may, tried in distinctive lighting conditions to determine the strength of the created framework. The framework has too been able to snatch the new images of the client and suitably upgrade its classifier and preparing dataset. The framework was designed utilizing the facial points of interest plot and is tried beneath different scenarios for the result that would be gotten. It is seen that the classifier has an precision of

more than 80 percent for most of the test cases, which is beautiful great precision in terms of feeling classification. It can moreover be seen that the classifier can precisely anticipate the expression of the client in a real-time situation when tried live for a user.

VI. RESULT ANALYSIS

Experiment Results- Instructions Explained to the User. In this scenario the users were given instructions as to what is to be done to perform the prediction of the emotion expressed which provided the following results. Sometimes in cases where the inner emotion is sad and facial expression is happy it resulted in a fail case. The values are given in Table 1 and the result is shown in Figure 2.

Table 1. Instructions Explained to the User.

- User Emotion Facial Expression Accuracy
1. Happy Happy 100
  2. Sad Happy 0
  3. Happy Happy 100
  4. Sad Sad 100

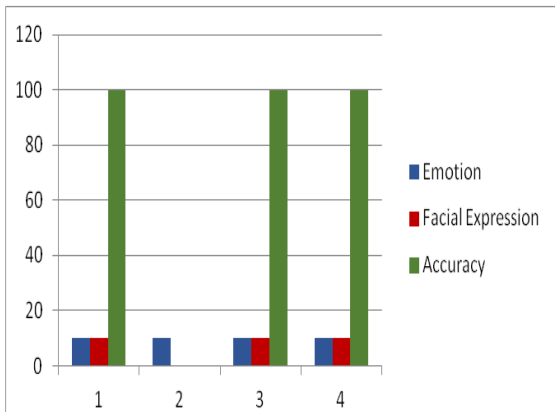


Figure 2. Instructions Explained to the user

Experiment Results - Instructions not given to the User. In this scenario the users were not given any instructions as to what is to be do and thus the inner emotions or the emotions recognized failed, there were also cases where in the emotion matched with the facial expressions of the user. The values are given in

Table 2 and the result is shown

User	Emotion	Facial Expression	Accuracy
1	Happy	Sad	0
2	Sad	Happy	0
3	Sad	Sad	100
4	Happy	Sad	0

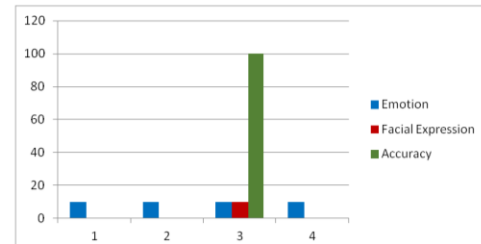


Figure 3. Experiment Results

VII. FUTURE SCOPE AND CONCLUSION

FutureWork

Diminish the time required to prepare the Feeling acknowledgment utilizing facial expressions is one of the critical themes of inquire about and has gathered much consideration in the past. It can be seen that the issue of feeling acknowledgment with the help of picture handling calculations has been expanding day by day. Analysts are continuously working on ways to resolve this by the utilize of diverse sorts of highlights and picture processing methods. The applications of picture preparing calculations in the field of both restorative science and human science are of endless significance. There are persistently modern ways and strategies being that make utilize of picture handling calculations to extricate the feeling of the client and make utilize of the extracted feeling to treat the client. Feeling acknowledgment has picked up a part of significance in all angles of life and if a strong calculation actualized which can precisely classify the feelings of the person, then a awesome bargain of progression in the industry can be accomplished with the offer assistance of this. The framework has successfully been able to capture the feeling of a client. It has been tried in a real-time environment for this predicate. In any case, it has to be tried in distinctive lighting conditions to decide the robustness of the created framework. The framework has too been able to get the unused of user and suitably upgrade its classifier and preparing dataset. The framework was outlined utilizing the facial landmarks

plot and was tried beneath different scenarios for the result that would It's seen that the classifier has an precision of more than 80 percent for most of the test cases, which is pretty great exactness in terms of feeling classification. It can too be seen that the classifier can accurately anticipate the expression of the client in a real-time situation when tried live for a user.

## VII. LIMITATIONS

- 1.The framework still is not able to record all the feelings accurately due to the less accessibility of the images in the picture dataset being used.
- 2.The picture that is encouraged into the classifier ought to be taken in a well-lit air for the classifier to give precise results.
- 3.The quality of the picture ought to be at slightest higher than 320p for the classifier to foresee the emotion of the client accurately classifier.

## REFERENCES

- [1] Londhe RR and Pawar DV 2012 Analysis of facial expression and recognition based on statistical approach International Journal of Soft Computing and Engineering 2
- [2] Kabani H, Khan S, Khan O and Tadvi S 2015 Emotion based music player International Journal of Engineering Research and General Science 3 750-6
- [3] Jyoti Rani and Kanwal Garg 2014 Emotion Detection Using Facial Expressions A Review International Journal of Advanced Research in Computer Science and Software Engineering 4
- [4] Joshi A and Kaur R 2013 A Study of speech emotion recognition methods Int. J. Comput. Sci. Mob. Comput. 2 28-31
- [5] Anwar S, Milanova M, Bigazzi A, Bocchi L and Guazzini A 2016 Real time intention recognition In IECON 2016-42nd Annual Conference of the IEEE Industrial Electronics Society 1021-1024 IEEE
- [6] A. habibzad, ninavin and Mir kamalMirnia A new algorithm to classify face emotions through eye and lip feature by using particle swarm optimization 2012 4th International Conference on Computer Modeling and Simulation
- [7] Rabashette MD, Tale MR, Hinge MA, Padale MK, Chavan MR and Deshmukh N Emotion Based Music System
- [8] Vaid S, Singh P and Kaur C 2015 EEG signal analysis for BCI interface: A review In2015 fifth international conference on advanced computing & communication technologies 143-147 IEEE
- [9] Liu C, Xie S, Xie X, Duan X, Wang W and Obermayer K 2018 Design of a video feedback SSVEP-BCI system for car control based on improved MUSIC method In2018 6th International Conference on Brain-Computer Interface (BCI) 1-4 IEEE