MUSIC EMOTION RECOGNITION USING FISHER FACE CLASSIFIER AND CNN

Submitted by: -

20BCE2534- ARYAN NIKUL PATEL

20BKT0140- ANUBHAV RAJ

20BCT0328- ISHIKA BAJAJ

20BCE0704- HARSHAD SINGH RATHORE

Team ID - 1

SLOT: - A2

Artificial Intelligence – CSE 3013

Submitted to: -

Prof. Rajeshkannan R

Associate Professor Grade 1,
SCOPE

ABSTRACT

Every individual human might have completely different faces; however, their expressions tell us the same story and it notably plays a significant role in extraction of an individual's emotions and behaviour. Music is the purest form of art and a medium of expression, which is known to have a greater connection with a person's emotions. It has a novel ability to lift one's mood. This project system focuses on building an efficient music player which works on emotion of user using facial recognition techniques. In order to detect a facial expression, the system should analyse various variability of human faces like colour, posture, expression, orientation, lighting etc. Detecting facial features is a prerequisite to facial emotion recognition. The facial features extracted will generate a system thereby reducing the effort and time involved in doing it manually. Facial data is captured by employing a camera. One of the applications of this input can be for extracting the information to deduce the mood of an individual. This data can then be used to get a list of songs that comply with the "mood" derived from the input provided earlier. The emotion module makes use of deep learning techniques to spot the exact mood relative to that expression.

A brief idea about our systems working, playlist generation and emotion classification is given. This is achieved by observing the parts of the face, like eyes, lips movement etc. These are then classified and compared to trained sets of data. In this research, a human facial expression recognition system will be modelled using eigenface approach. The proposed method will use HAAR Cascade classifier (OpenCV) to detect the face in an image. Fisher Faces calculation can be utilized for decreasing the high dimensionality of the eigenspace and after that anticipating the test picture upon the eigenspace and computing the Euclidean separation between the test picture and mean of the eigen faces. The grey scale image of the face is used by the system to classify emotions such as angry, happy, neutral, sad, contempt, fear, surprise and disgust.

INTRODUCTION

In today's world, with ever increasing advancements in the field of multimedia and technology, various music players have been developed with features like fast forward, reverse, variable playback speed (seek and time compression), local playback, streaming playback with multicast streams. Although these features satisfy the users basic requirements, yet the user has to face the task of manually browsing through the playlist of songs and select songs based on his current mood and behaviour. Music plays a very important role in enhancing an individual's life as it is an important medium of entertainment for music lovers and listeners and sometimes even imparts a therapeutic approach. Facial expressions give important clues about emotions. A human can express his/her emotion through lip and eye. The work describes the development of Facial Expression Based Music Player, which is an application meant for users to minimize the efforts in managing large playlists. Generally, people have a large number of songs in their database or playlists. Thus, to avoid trouble of selecting a song, most people will just randomly select a song from their playlist and some of the songs may not be appropriate for the current mood of the user and it may disappoint the user. Facial Expression based Music Player is interactive, sophisticated and innovative mobile (Android) based application to be used as a music player in a different manner.

Importance of Emotion Recognition:

1. Human face detection plays an important role in applications such as video surveillance, human computer interface, face recognition, and face image database management.
2. Facial expressions are important cues for nonverbal communication among human beings. This is only possible because humans are able to recognize emotions quite accurately and efficiently. An automatic facial emotion recognition has many commercial uses and can be used as a biological model for cognitive processing and analysing of human brain.
3. Collectively they can enhance their applications like monitoring and surveillance analysis, biomedical image, smart rooms intelligent robots, human computer interfaces and driver's alertness system and can play a vital role in the field of security and crime investigations.

Literature Review

Paper Title Journal Details	Method/Algorithm	Challenges	Observations
Face detection and expression recognition using Haar cascade classifier and Fisherface algorithm 2019 - Springer	face detection and expression recognition using Haar classifier and using Fisherface	Major drawback include incorrect recognition in cases of uneven illumination or low-light conditions.	The dataset from which the faces were presented to the classifiers yielded a precision of 96.3% with a recognition speed of 8.2 s.

Facial recognition using Haar cascade and LBP classifiers Facial recognition using Haar cascade and LBP classifiers- 2021 - Elsevier	hierarchical face and eyes detection system based on Viola's Haar Cascade Classifiers (HCC) augmented by some knowledge-based rules.	Many of false positives could be removed by restricting the minimum number of merged detections. However, it should be pointed out that only increasing it to the number of 5 produced positive results. Further increase of required neighbours resulted in the quick deterioration of the true positive ratio without any significant change in the false positive ratio.	The tests proved that using more complex weak classifiers pays off resulting in much more reliable detections. any significant change in the FP ratio. Detectors trained on the plain background images turned out to be less efficient than those trained on rich negative set. This suggests, that in order to build a reliable classifier it is recommended to use as diversified negative set as possible.
Comparison of Viola-Jones Haar Cascade Classifier and Histogram of Oriented Gradients (HOG) for face detection 2020- IOPSCIENCE	V-J algorithm can detect frontal face very well in images, regarding of their scale, pose, makeup, expression, and illumination.	difficult to detect the face who have occlusions like using helm, eyeglass, and mask. main problem with Haar cascades is in the parameter called detect multiscale and scale factor. If the scale factor is too low, many pyramid layers will be evaluated, this will help to detect more than one faces in images. On the other hand, if scale factor is too large, it cannot detect the face in small pixel.	With six types of condition and five times of trial, obtain the accuracy by 75,33% by using V-J and 80,22% by using HOG.
Comparision of Paul Viola–Michael Jones algorithm and HOG algorithm for Face Detection - 2021 - iopscience	Histogram of Oriented Gradients (HOG)- method for detecting objects which can also be used for detecting faces. In this case, we will use HOG for detecting any front face in a picture. This method requires an image in grayscale.	The aim is to recognize any face, but too many details can only detect one specific face. Thus, only the relevant directions need to be kept.	The strength of this method is that it is not sensitive to a change in luminosity. If a picture is darker, all the pixels will be darker.

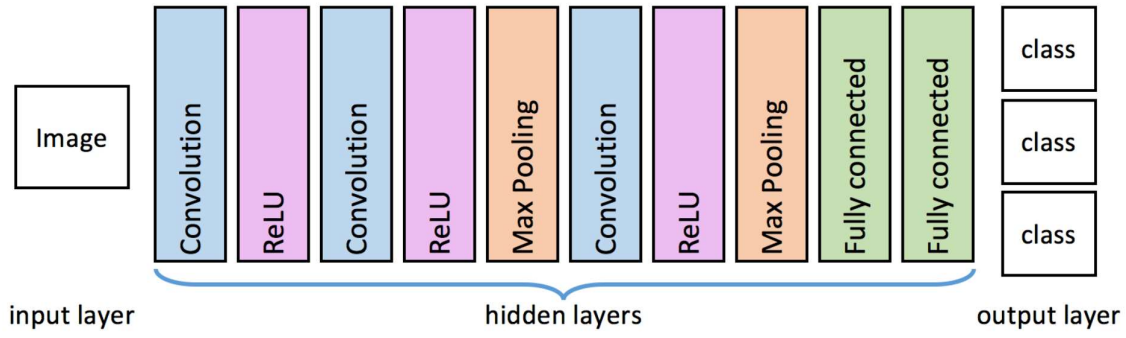
Evaluating the Performance of Eigenface, Fisherface, and Local Binary Pattern Histogram-Based Facial Recognition Methods under Various Weather Conditions- 2021 - mdpi.com	Eigenfaces is a method for performing facial recognition based on a statistical approach. The aim of this method is to extract the principal components which affect the most the variation of the images.	learning is very time-consuming, which makes it difficult to update the face database.	Eigenfaces is based on Principal Component Analysis (PCA) for reducing the number of dimensions while preserving the most important information. The training part of Eigenfaces is to calculate the eigenvectors and the related eigenvalues of the covariance matrix of the training set.
- Emotion Recognition using Fisher Face-based Viola Jones Algorithm- 2018 - ieeexplore	This algorithm is a modification of Eigenfaces, thus also uses Principal Components Analysis. The main modification is that Fisherfaces takes into consideration classes. Fisherfaces uses the method Linear Discriminant Analysis in order to make the difference between two pictures from a different class.	Fisherface is more complex than Eigenface in finding the projection of face space. Calculation of ratio of between-class scatter to within-class scatter requires a lot of processing time.	Face recognition system using fisherface methods able to recognize the image of face testing correctly with 100% percentage for the test image the same as the training image and able to recognize the image of face testing correctly with 93% when the test image different from the training image.
Developing a LBPH-based Face Recognition System for Visually Impaired People- 2021 - ieeexplore.	LBPH algorithm also requires grayscale pictures for processing the training. In contrast to the previous algorithms, this one is not a holistic approach. The aim of LBPH is to work by blocks of 3x3 pixels.	Major disadvantages of existing local binary pattern (LBP) operators are that they produce rather long histograms, which slow down the recognition speed especially on large-scale face database.	LBPH has been modified in different ways (what-when-how). One of them is called Extended LBPH. This extension is using a circular neighborhood which is composed of a radius and a number of sampling points. This approach allows a pixel to have more than eight neighbors.
A face emotion recognition	Using a convolutional neural network (CNN) is	CNN requires Large training data in order to	The input of the first layer is a picture of a

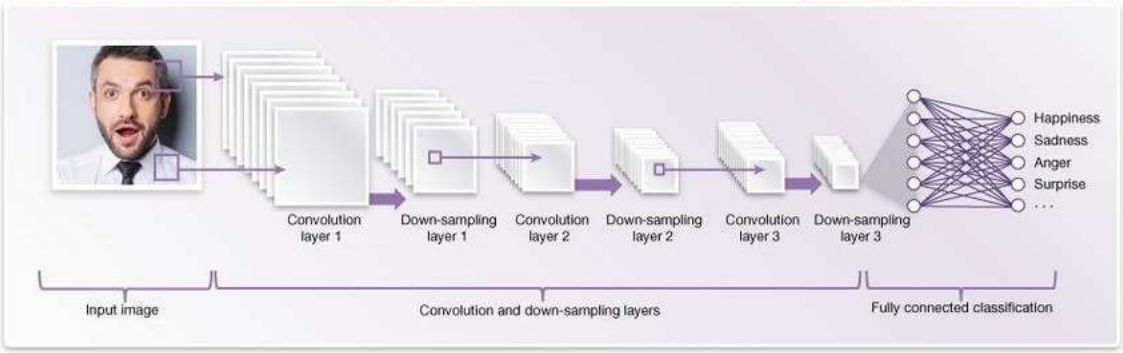
method using convolutional neural network and image edge computing H Zhang, A Jolfaei, M Alazab - IEEE Access, 2019 - ieeexplore.ieee.org	another way to perform face recognition. CNN has an architecture that enables it to use 2d pictures as inputs. A CNN consists of several layers called hidden layers. The layers are composed of several neurons. A neuron has specific weight and receives an input.	provide good accuracy as well as it does not encode the position and orientation of object.	face. The output of the last layer is the predicted class which is the person. In order to have a better comprehension of the recognition process, it is preferable to use an easy example.
Py-feat: Python facial expression analysis toolbox JH Cheong, T Xie, S Byrne, LJ Chang - arXiv preprint , 2021 - arxiv.org	OpenFace is a face recognition library. It is based on Google's FaceNet systems. OpenFace is using a deep convolutional neural network for performing a facial recognition.	A large training data set is required for model to give high accuracy and low error rate.	OpenFace was trained with 500k images. The faces are detected and extracted from the pictures with dlib's pre-trained detector that uses a HOG. Then, the faces go through a preprocessing phase and are finally used in the convolutional neural network.

Problem Statement:-

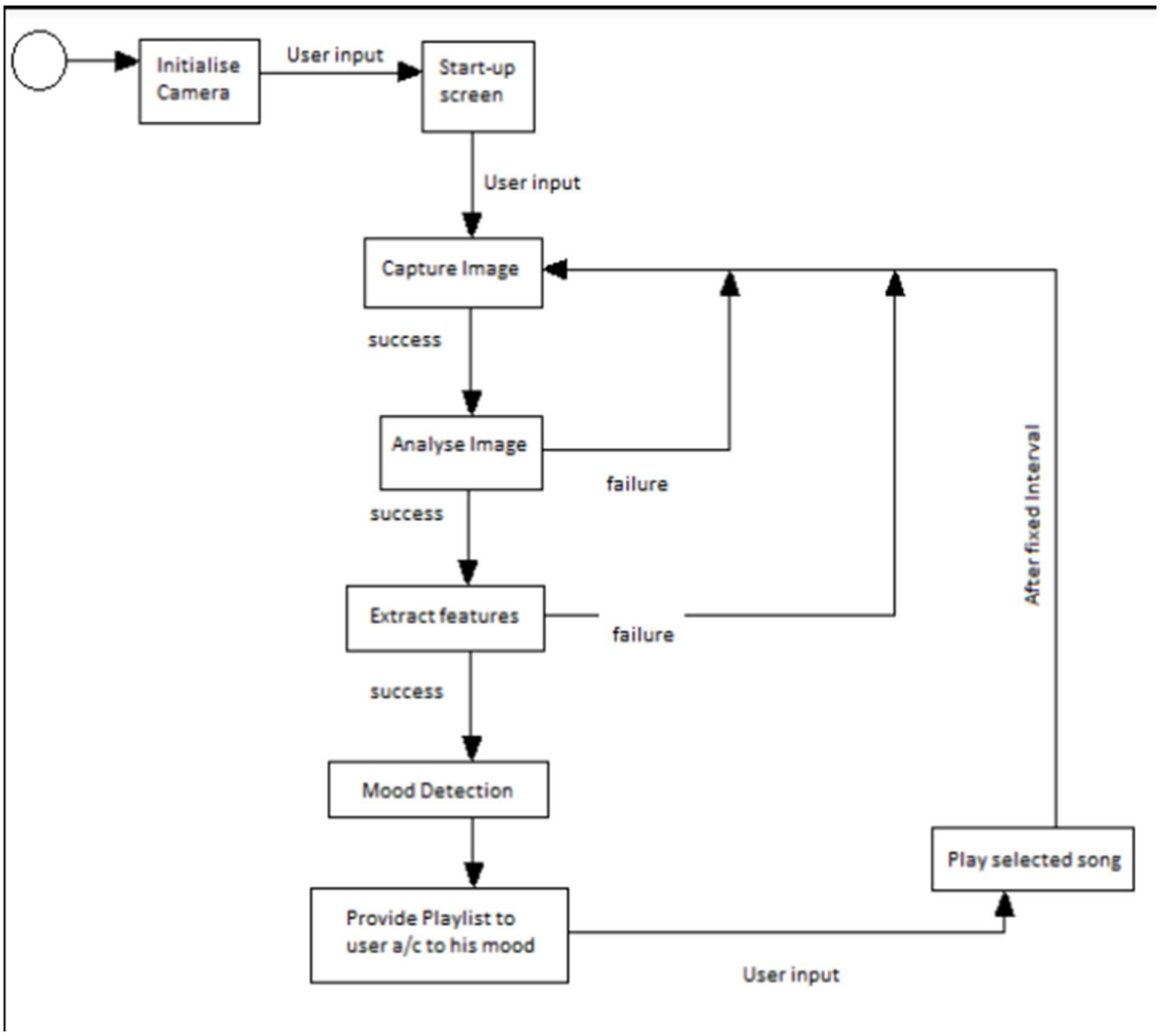
How it can save lives of vehicle drivers?
If a driver gets angry and hits his aggression on others on roadway, that aggressive and rash driving is known as road rage. Road rage distracts a driver's focus on safety and disturbs his clear-thinking mind. It can put that driver or everyone else on the road into a danger. According to National Highway Traffic Safety Administration more than 55% of the accidents around 106,729 are happened because of aggressive driving and anger might be the reason for it.
This project of music player has a special mode which makes it as a unique feature, it works on user's mood. If it finds user's mood is angry it can play soothing and calming music which can turn user's mood into calm and cool. It might reduce the anger of driver which can stop him from rash driving and can save lives. Java Script programming language is used to write the code of music player in the backend. It has three other different modes of playing songs besides mood-based feature. It has supports adding songs in the queue and songs can also be picked randomly. As we know that CSS, HTML give a great look to communicate through JavaScript and help us to interact with user which adds extra friendly nature and ease of access to user. It not only runs on the console it gives the user privilege so that user can control it manually.
Light and camera quality can be the factors which can affect the result. Good lighting in the room can give satisfactory results.

Architecture Diagram / Neural Network Diagram





FLOW DIAGRAM



Implementation and Architecture Flowchart

Pseudocode: -

Step 1 - import libraries cv2, argparse, time , os , update_Model , glob , random , eel , light.
Step 2 - SETS frequency -> 2500
SETS Duration -> 1000
Step 3 - Take emotions "angry", "happy", "sad", "neutral" , "contempt", "disgust" , “fear" and “surprise” from the directory.

Step 4 - Add Parserarguement “- - update”
Step 5 - Function crop{

for(all the dimensions) in face :
faceslice = clahe_image[increase X axis and Y axis]
faceslice = resize(faceslice(to needed dimensions))

arrayfacedict[ ] = faceslice
RETURNS faceslice
}

Step 6 - Function grab_face{
SETS frame = nolight() ;
SETS imwrite(image file , frame)
SETS imwrite(send to directory)

clahe_image = apply(gray) //for the grey scaling
RETURNS clahe_image
}
Step 7 - Function detect_face{
CALLS grab_face()
IF(length of face) >= 1 :
crop the image :

ELSE PRINT— “Multiple faces detected!!, passing over the frame , try to stay still")

}

Step 8 - Function save_face{
for(0 to 5 seconds){

takes the image of the user:
}
CALLS detect_face function

SAVES - faces to a directory

}

Step 9 — * Function UPDATE MODEL{
CHECKS (emotion directory)

For (All elements in emotion) {
CALLS save_face()
CALLS UPDATE_MODEL
}

Step 10 - Function identify_emotions{

SET prediction = [] , confidence = []

For(i in face_dictionary){

CALLS FISHFACE.predict()
SETS imwrite(images) to face_dictionary

APPEND prediction
APPEND confidence
SET output = [max(set(prediction),
key=prediction.count)]

RETURNS output:

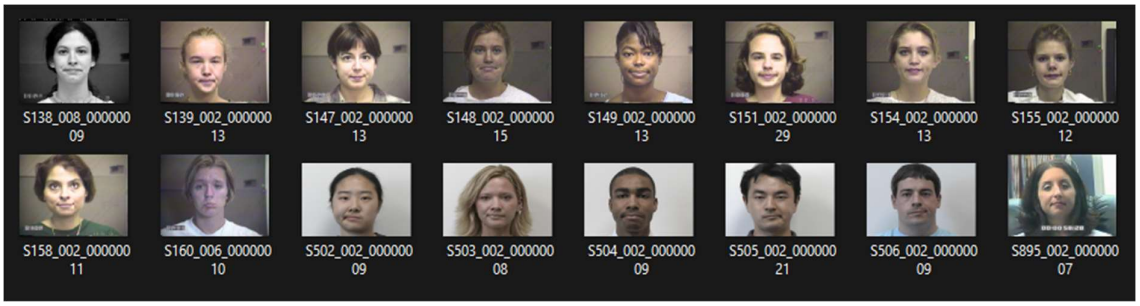
}

Experiment and Results

Dataset Sample: -

1.Anger

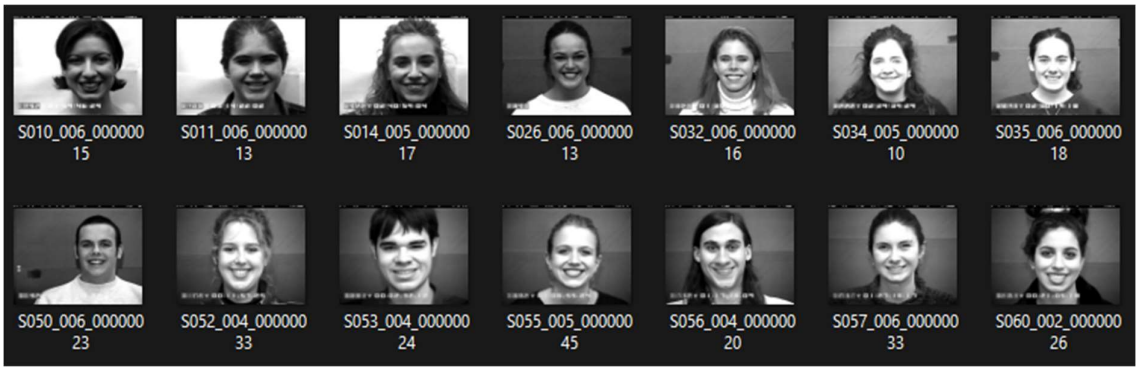
2.Contempt



3.Disgust

4.Fear

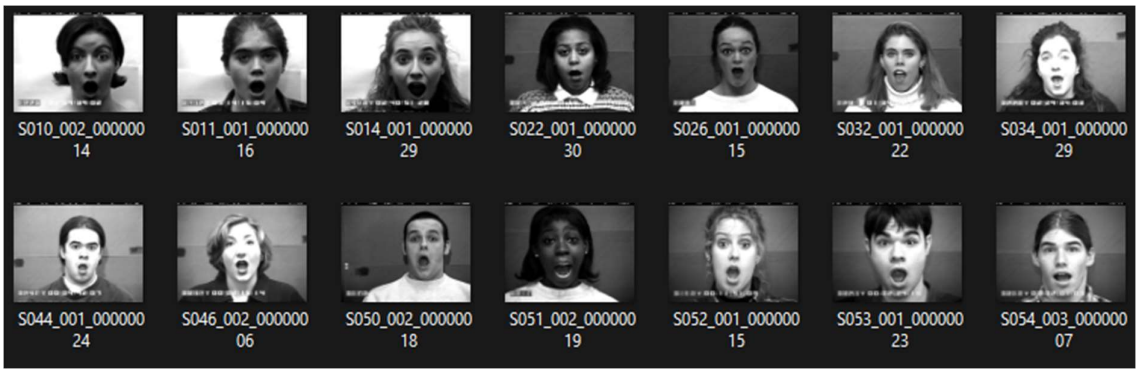
5.Happy



6.Neutral

7.Sad

8.Suprise



Dataset Interpretation: -

The proposed system is tested on datasets which consists of wide range of face images with different expressions, poses, illumination conditions and genders. We used CK and CK+ database for training of emotion detection system. Algorithm is tested on IMM database and also our own test images. Both databases are open source and our algorithm performed well on both datasets. For face recognition, we have used a webcam for capturing of faces. The implemented algorithm is capable of recognizing different persons in a single window. If the recognition environment is under proper lighting condition and less background noises then the recognition rate will be high.

The Cohn-Kanade (CK) database was released for the purpose of promoting research into automatically detecting individual facial expressions. Since then, the CK database has become one of the most widely used test-beds for algorithm development and evaluation.

During this period, three limitations have become apparent:

1) While AU codes are well validated, emotion labels are not, as they refer to what was requested rather than what was actually performed,

2) The lack of a common performance metric against which to evaluate new algorithms, and

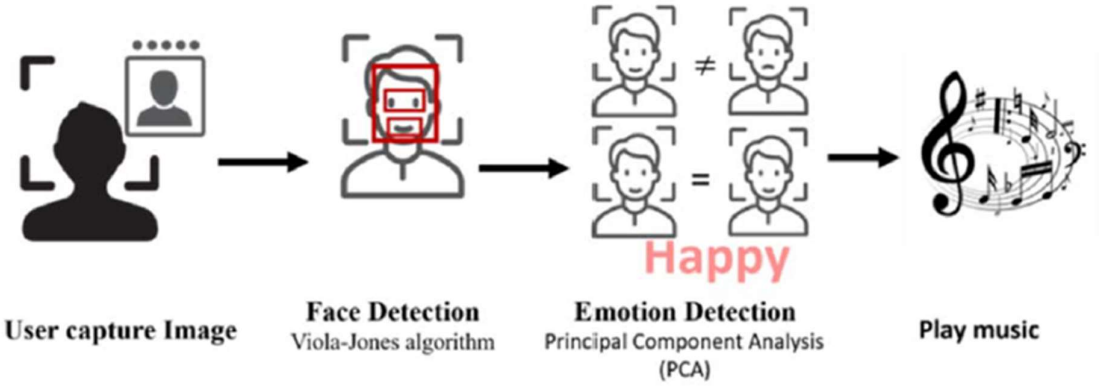
3) Standard protocols for common databases have not emerged.

As a consequence, the CK database has been used for both AU and emotion detection (even though labels for the latter have not been validated), comparison with benchmark algorithms is missing, and use of random subsets of the original database makes meta-analyses difficult.

In Extended Cohn-Kanade (CK+) database. The number of sequences is increased by 22% and the number of subjects by 27%. The target expression for each sequence is fully FACS coded and emotion labels have been revised and validated.

We downloaded the ck+ from
https://github.com/spenceryee/CS229/tree/master/CK%2B

Methodology: -



The proposed system is a music controller that is based on automatic emotion detection. A webcam is used to capture the images that will be used as input to the proposed system, then it goes to the expression detector to classify it to one of eight classes "Happy", "Natural", "Sad", "Angry", "Contempt", "Fear", "Surprise" and "Disgust".

A. Image Acquisition

As a first step in the proposed system, we begin by acquiring the image of the user's face using a built-in laptop webcam (or any external camera can be employed). The face image to be correctly processed in the proposed system must contain one face in the frontal position in a uniformly illuminated background. Also, it should not be on the user face anything that could impede the process of detection such as glasses.

B. Face Detection

After acquiring the image, the system will start to detect the face by applying the Viola-Jones algorithm. This Algorithm is considered as one of the first frameworks that recognize the objects in real time. Simply, Viola-Jones scan the images using a sub-window to detect the features of the face in the image. When the face is determined, the image is cropped to contain the face only, to enhance the proposed system performance. Also, the Viola-Jones is reused to identify and crop the left and right eyes and mouth separately. The outcome of this step is four images, face, right eye, left eye, and mouth images.

C. Emotion Detection

Next the user sentiment must be detected, we use Fisher Face method which is a well-known approach that is often used to detect face emotions. It will construct the face space and the eigenvectors that has the highest eigenvalues will be selected. Also, the acquired image will be projected over the face space.

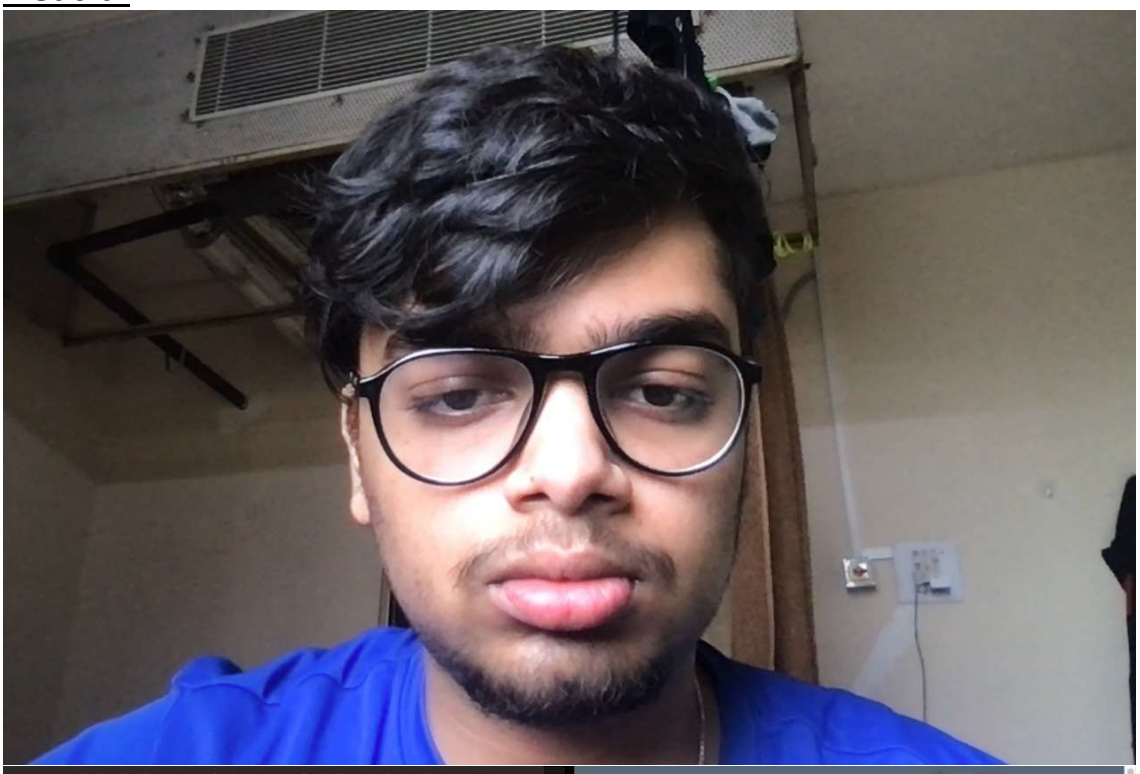
After that the emotion is detected by computing for the user image the scores for each emotion, then the emotion of the image is determined by getting the maximum score of the calculated emotion scores.

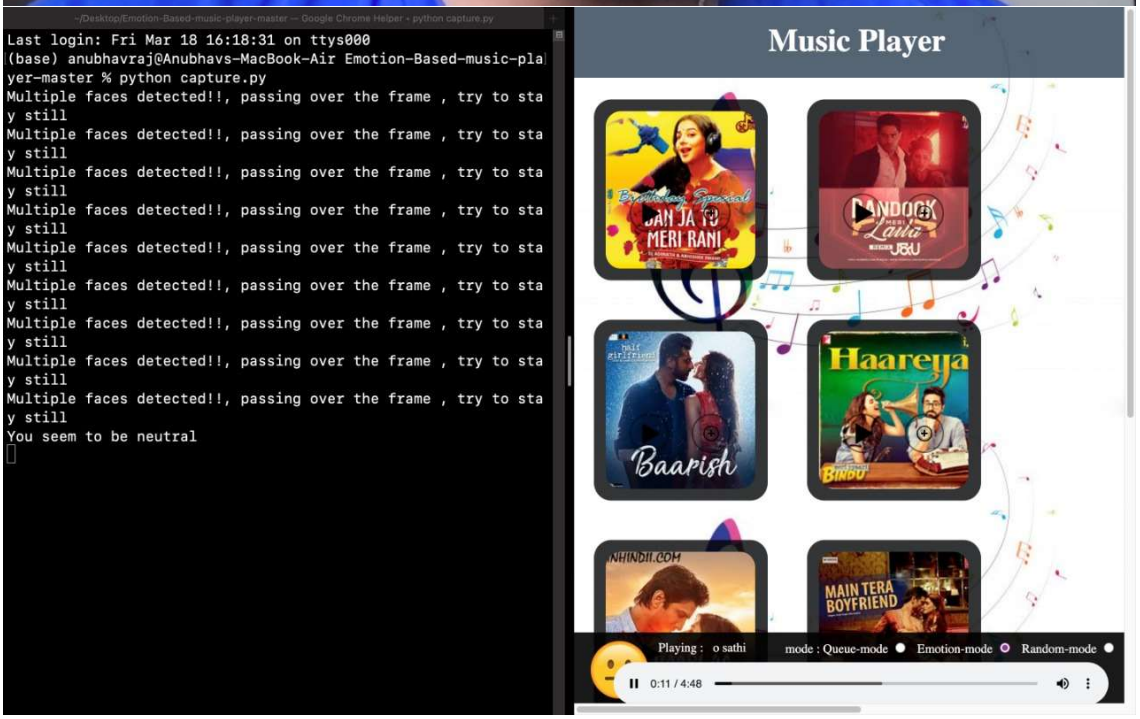
D. Enabling the correspondent Emotion playlist

The proposed system is depending on the detected emotion will present the correspondent music playlist. Since we have four emotions, we also have four playlists that offer music clips that are carefully chosen. For happy emotion, the classical music playlist will be activated, while the new age music playlist is dedicated to the natural emotion. For the negative and sad emotion the designer music playlist will be enabled to enhance the user mood to a better mood.

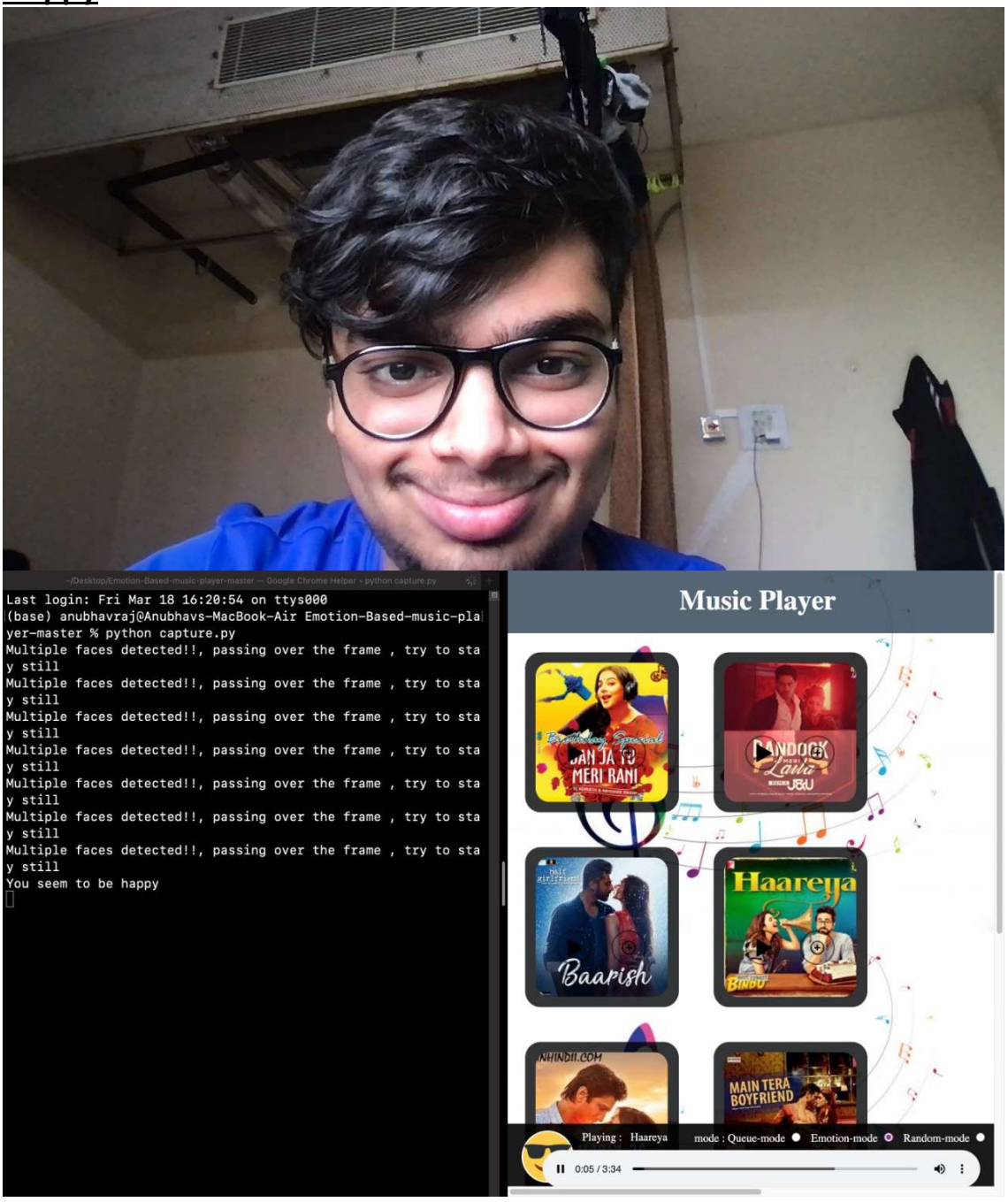
Sample Output Screen: -

Neutral

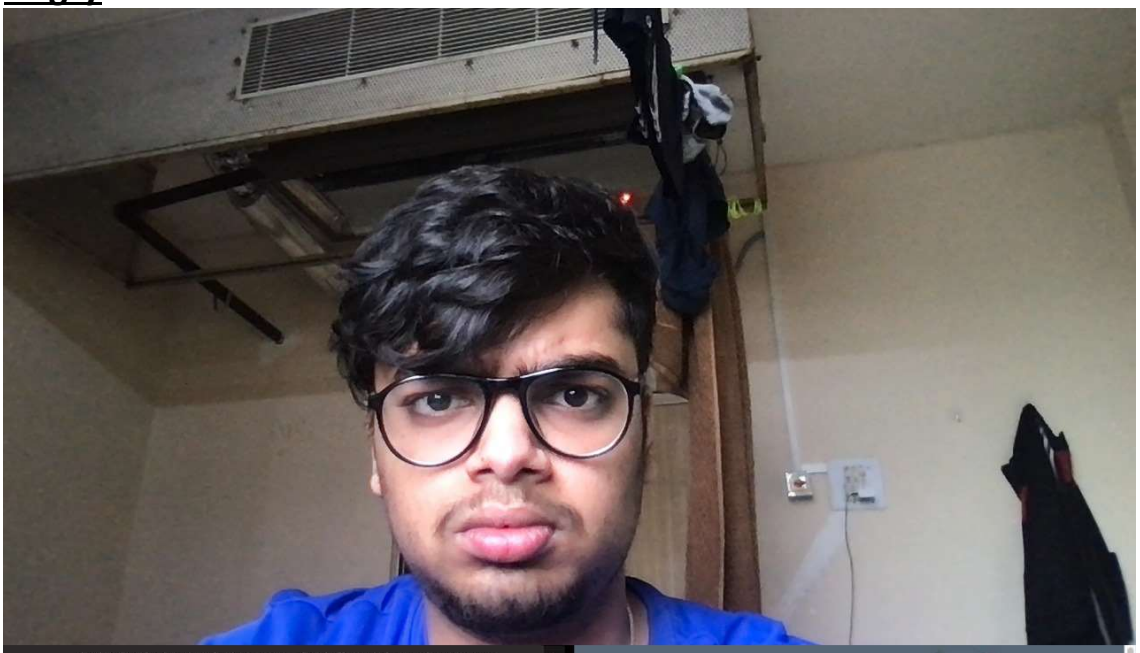




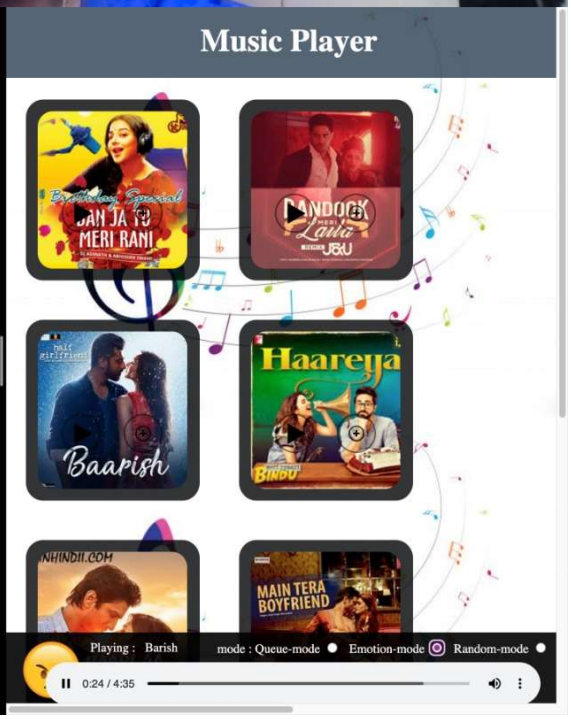
Happy


~/Desktop/Emotion-Based-music-player-master — Google Chrome Helper • python capture.py
Last login: Fri Mar 18 16:20:54 on ttys000
(base) anubhavraj@Anubhavs-MacBook-Air Emotion-Based-music-player-master % python capture.py
Multiple faces detected!!, passing over the frame , try to stay still
Multiple faces detected!!, passing over the frame , try to stay still
Multiple faces detected!!, passing over the frame , try to stay still
Multiple faces detected!!, passing over the frame , try to stay still
Multiple faces detected!!, passing over the frame , try to stay still
Multiple faces detected!!, passing over the frame , try to stay still
Multiple faces detected!!, passing over the frame , try to stay still
You seem to be happy
Music Player
MERI RANI
Baarish
Haareyja
MAIN TERA BOYFRIEND
Playing : Haareya
mode : Queue-mode
Emotion-mode
Random-mode
0:05 / 3:34

Angry



```
~/Desktop/Emotion-Based-music-player-master — Google Chrome Helper • python capture.py
Multiple faces detected!!, passing over the frame , try to sta
y still
Multiple faces detected!!, passing over the frame , try to sta
y still
Traceback (most recent call last):
  File "/Users/anubhavraj/opt/anaconda3/lib/python3.9/site-pac
kages/eel/__init__.py", line 281, in _process_message
    return_val = _exposed_functions[message['name']](*message[
'args'])
  File "/Users/anubhavraj/Desktop/Emotion-Based-music-player-m
aster/capture.py", line 126, in getEmotion
    return identify_emotions()
  File "/Users/anubhavraj/Desktop/Emotion-Based-music-player-m
aster/capture.py", line 105, in identify_emotions
    output=emotions[max(set(prediction), key=prediction.count)
]
ValueError: max() arg is an empty sequence
clear
Multiple faces detected!!, passing over the frame , try to sta
y still
Multiple faces detected!!, passing over the frame , try to sta
y still
Multiple faces detected!!, passing over the frame , try to sta
y still
Multiple faces detected!!, passing over the frame , try to sta
y still
Multiple faces detected!!, passing over the frame , try to sta
y still
Multiple faces detected!!, passing over the frame , try to sta
y still
Multiple faces detected!!, passing over the frame , try to sta
y still
Multiple faces detected!!, passing over the frame , try to sta
y still
You seem to be angry
```



Conclusion: -

The proposed system processes images of facial expressions, recognizes the actions related to basic emotions, and then plays music based on these emotions. The main advantage of this system is that it is completely independent of automation. In the future, the application can export songs to a dedicated cloud database and allows users to download desired songs, as well as to recognize complex and mixed emotions. Therefore, the developed application will provide users with the most suitable songs based on their current emotions, thereby reducing the workload of users creating and managing playlists, bringing more fun to music listeners, not only helping users, but also songs can be organized systematically. And this model also has productive future scope, facial recognition can be used for authentication purpose, an android app can be developed for mobile uses, it can also detect sleepy mood if the driver is not focused on driving. Even people who are physically challenged can use this, for them this is very better than voice-based applications.

References

1.Face detection and expression recognition using Haar cascade classifier and Fisherface algorithm
Gangopadhyay, A Chatterjee, Das - Recent Trends in Signal and Image …, 2019 - Springer

2. A Study of LBPH, Eigenface, Fisherface and Haar-like features for Face recognition using OpenCV
AM Jagtap, V Kangale, K Unune… - 2019 International …, 2019 - ieeexplore.ieee.org

3. Facial Expression Classifier Using Better Technique: FisherFace Algorithm
N Hegde, S Preetha, S Bhagwat - … International Conference on …, 2018 - ieeexplore.ieee.org

4. Mood Based Music Player
A Arora, A Kaul, V Mittal - 2019 International Conference on …, 2019 - ieeexplore.ieee.org

5. Emotion based mood enhancing music recommendation
AV Iyer, V Pasad, SR Sankhe… - 2017 2nd IEEE …, 2017 - ieeexplore.ieee.org

6. An intelligent music player based on emotion recognition
R Ramanathan, R Kumaran, RR Rohan… - 2017 2nd …, 2017 - ieeexplore.ieee.org

7. Mind Frame Music and Movie Recommendations to uplift the current mood using Deep Learning
M Gupta, SV Amirishetty, B Nithin… - … Conference on System …, 2021 - ieeexplore.ieee.org

8. Mus-Emo: An Automated Facial Emotion-Based Music Recommendation System Using Convolutional Neural Network
S Mittal, A Ranjan, B Roy, V Rathore - Advances in Communication …, 2022 - Springer

9. Comparison of Viola-Jones Haar Cascade classifier and histogram of oriented gradients (HOG) for face detection
C Rahmad, RA Asmara, DRH Putra… - IOP conference …, 2020 - iopscience.iop.org

10. Comparison of Haar-like, HOG and LBP approaches for face detection in video sequences
A Adouani, WMB Henia, Z Lachiri - 2019 16th International Multi …, 2019 - ieeexplore.ieee.org

11. Face recognition using Eigenfaces, geometrical PCA approximation and neural networks

AL Machidon, OM Machidon… - 2019 42nd International …, 2019 - ieeexplore.ieee.org
12. Face recognition using eigenfaces
GM Zafaruddin, HS Fadewar - Computing, communication and signal …, 2019 - Springer
13. Facial expression recognition by calculating euclidian distance for eigen faces using PCA
MBS Divya, NB Prajwala - 2018 International Conference on …, 2018 - ieeexplore.ieee.org
14. Face recognition using fisherface method
M Anggo, L Arapu - Journal of Physics: Conference Series, 2018 - iopscience.iop.org
15. Evaluating the Performance of Eigenface, Fisherface, and Local Binary Pattern Histogram-Based Facial Recognition Methods under Various Weather Conditions
MM Ahsan, Y Li, J Zhang, MT Ahad, KD Gupta - Technologies, 2021 - mdpi.com
16. A face emotion recognition method using convolutional neural network and image edge computing
H Zhang, A Jolfaei, M Alazab - IEEE Access, 2019 - ieeexplore.ieee.org
18. Facial emotion recognition of students using convolutional neural network
I Lasri, AR Solh, M El Belkacemi - 2019 third international …, 2019 - ieeexplore.ieee.org