# COUNTENANCE BASED BOOM BOX

**A Project Report**

***submitted***

*in partial fulfillment of the requirements for*

*the award of the degree of*

**Bachelor of Technology**

in

**Computer Science & Engineering**

(JNTUA, Anantapuramu)

by


| | |
|---|---|
| **Mohammed Imdaad** | **162N1A0559** |
| **Machani Sunil Kumar** | **162N1A0552** |
| **Methuku Vineeth** | **162N1A0551** |


Under the guidance of

**Prof P.Rama Bayappa Reddy**

HOD, Dept. of CSE

BITS,Kurnool

**Department of Computer Science & Engineering**

**BRINDAVAN**

**INSTITUTE OF TECHNOLOGY & SCIENCE (BITS - KNL)**

**(Approved by AICTE & Affiliated to JNTUA)**

**NH-7, PEDDATEKUR, KURNOOL-518218**

**Jan-2020**

# BRINDAVAN
## INSTITUTE OF TECHNOLOGY & SCIENCE (BITS - KNL)
## NH-7, PEDDATEKUR, KURNOOL-518218

## CERTIFICATE

This is to certify that the Project Report entitled **COUNTENANCE BASED BOOM BOX** is the bonafide record of the Project Work carried out under our Guidance and supervision by

| | |
|---|---|
| Mohammed Imdaad | 162N1A0559 |
| Machani Sunil Kumar | 162N1A0552 |
| Methuku Vineeth | 162N1A0551 |

in partial fulfillment of the requirements for the award of Bachelor of Technology in Computer Science & Engineering and, is submitted to the Department of Computer Science & Engineering, Brindavan Institute of Technology & Science, NH-7, Peddatekur, Kurnool.

**Head of the Dept.**

**Prof.P.Rama Bayapa Reddy**
**HOD, Department of CSE**
**Brindavan**
**Institute of Technology & Science**
**NH-7, Peddatekur, Kurnool**

**Guide**

**Prof.P.Rama Bayapa Reddy**
**HOD, Dept. of CSE**
**Brindavan**
**Institute of Technology & science**
**NH-7, Peddatekur, Kurnool**

**Signature of the External Examiner**

# ACKNOWLEDGMENTS

Our hard work never shines if we do not convey my heartfelt gratitude to those people from whom we got considerable support and encourage during this Project report.

We would like to express our gratitude to our Guide **Prof. P. Rama Bayapa Reddy HOD** of Computer Science & Engineering, **BRINDAVAN INSTITUTE OF TECHNOLOGY & SCIENCE,** for his constant support and guidance throughout the project report.

We would like to thank **Prof. P. Rama Bayapa Reddy**, Head of the Department, Computer Science & Engineering, **BRINDAVAN INSTITUTE OF TECHNOLOGY & SCIENCE**, Kurnool for his valuable suggestions from time to time during this Project report.

We also express our special thanks **to Prof. N. Siva Prasad Reddy Academic Director and Dr. A. SreeKanth Gowd Principal**, **BRINDAVAN INSTITUTE OF TECHNOLOGY & SCIENCE**, for their support and encouragement to complete our Project report.

We are happy to express our sincere thanks to all teaching and non-teaching staff, Department of Computer Science & Engineering for their help. Finally, We thankful to all those who helped directly or indirectly in making endeavor a success.

**Project Associatives**

# ABSTRACT

The human face is an important organ of an individuals body and it especially plays an important role in extraction of an individuals behaviour and emotional state. Recent studies confirm that humans respond and react to music and that music has a high impact on person's brain activity. People tend to listen to music based on there mood and interests. Extracting the required input from the human face can now be done directly using a camera. Human emotions are meant for mutual understanding and sharing feelings and intentions. The emotions are manifested in verbal and facial expressions. Facial expression is a form of nonverbal communication. This project suggest songs for user based on their expression by capturing facial expressions. In this system, computer vision components are used to determine the user's emotion through facial expressions.

Once the emotion is recognized, the system suggests a play-list for that emotion, this eliminates the time-consuming and tedious task of manually Segregating or grouping songs into different lists and helps in generating an appropriate playlist based on an individual's emotional features. The project is developed in such a way that it can manage content accessed by user, analyse the image properties and determine the expression of the user based on mp3 file properties so that they can be added into appropriate play lists according to the expression. The softwares used in this project are OpenCV (Open source computer vision), AWS (Amazon Web Services) and python language. An inbuilt camera is used to capture the facial expressions such as happy, sad, anger, surprise, fear, confused, disgust and neutral of a person which reduces the designing cost of the system as compared to other methods. The proposed is not a normal system, in this if the person's expression is sad, fear, anger or disgust then it will play songs quite opposite to that expressions as to overcome their expression from the present expression by listening pleasant music. If the person expression is happy, surprise, neutral, confused it will play songs accordingly.

The main theme of the project is to detect the emotions that appear to be expressed on the face, and the confidence level in the determination. It is only making a determination of the physical appearance of a person's face. It is not a determination of the person’s internal emotional state and should not be used in such a way. For example, a person pretending to have a sad face might not be sad emotionally. The best possible and effortless pleasure of music, Facial Expression Recognition (FER) based systems have been adopted as they provide more fast, accurate and efficient results with less effort.

With the world moving towards fields like Artificial Intelligence(AI) and Machine Learning(ML), our aim is to provide the users a platform through which on their current expression, music is played using Facial Expression Recognition.

# INDEX

## TABLE OF CONTENTS

# LIST OF FIGURES

# CHAPTER 1

# 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 user's 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 behavior. 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.

Music has been proven to be an integral part of everyone's life. It acts as a source for entertainment and also used for various medical needs as it is proven to be a Stress Reliever. There are numerous high-end music players available with the latest features of handling the volume, modulation, pitch, sound, genre, etc. Though these features are very useful for the users but sometimes it becomes quite irritating and time-consuming to manually browse through the playlist for the intended song which user wants to play based on his/her mood and emotional state. For the purpose of providing the users with the best possible and effortless pleasure of music, Facial Expression Recognition (FER) based systems have been adopted as they provide more fast, accurate and efficient results with less effort. With the world moving towards fields like Artificial Intelligence(AI) and Machine Learning(ML), our aim is to provide the users a platform through which on their current mood, music is played using Facial Expression Recognition.

Facial expressions give important clues about emotions. Computer systems based on affective interaction could play an important role in the next generation of computer vision systems. Face emotion can be used in areas of security, entertainment. A human can express his/her emotion through lip and eye. The work describes the development of Countenance Based Boom Box, 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. Countenance Based Boom Box is interactive, sophisticated and innovative web based application to be used as a music player in a different manner.

The application works in a different manner from the traditional software as it scans and classifies the audio files present on the device and according to the predefined parameters (Audio Features) present on the application in order to produce a set of mood based playlists. The real time graphical input provided to the application is classified (Facial expression recognition) to produce a mood which will then be used to select the required playlist from the earlier set. The main objective of the paper is to design and generate a playlist based on current emotional state and behavior of the user. Face detection and facial feature extraction from image is the first step in emotion based music player. For the face detection to work effectively, user needs to provide an input image which should not be blur and tilted.

The main objective of this work is to develop an intelligent system that can easily recognize the facial expression and accordingly play a music track based on that particular expression/emotion recognized. The seven universally classified emotions are Happy, Sad, Anger, Disgust, Fear, Surprise and Neutral.

# CHAPTER 2

# LITERATURE SURVEY

The potential abilities of humans to be able to provide inputs to any system in various ways has caught the attentions of various learners, scientists, engineers, etc from all over the world. The paper by Hafeez Kabini et al [1] addressed the problem of the existing methods typically handle only deliberately displayed and exaggerated expressions of prototypical emotions despite the fact that deliberate behavior differs in visual appearance, audio profile, and timing from spontaneously occurring behavior, by taking efforts to develop algorithms that can process naturally occurring human affective behavior have recently emerged. They introduced and surveyed these recent advances and discussed human emotion perception from a psychological perspective. They examined available approaches to solving the problem of machine understanding of human affective behavior, and discuss important issues like the collection and availability of training and test data.

The "mind" is a term that has always attracted scientists towards understanding it in a wholesome manner. The most natural way to express emotions is using facial expressions. We humans, often use nonverbal cues such as hand gestures, facial expressions, and tone of the voice to express feelings in interpersonal communications. Nikhil Zaware et al [2] stated that it is very time consuming and difficult to create and manage large playlists and to select songs from these playlists.

Thus, it would be very helpful if the music player itself selects a song according to the current mood of the user using an application to minimize the efforts of managing playlists. In their paper, they stated a way to automatically detect the mood of the user and generate playlist of songs which is suitable for the current mood. The image is captured using webcam and that image is passed under different stages to detect the mood or emotion of the user. The application is thus developed in such a way that it can manage content accessed by user, analyze the image properties and determine the mood of the user. The application also includes the facility of sorting songs based on mp3 file properties so that they can be added into appropriate playlists according to the mood.

To detect facial expression as indicator to cast a music playlist is one task carried out by Setiawardhana et al [3] in their technical paper. They work by doing

facial expression detection system input performed offline by taking photograph of a subject with nearest position from the camera where facial position should not be tilted. The image is identified as a combination of color and feature extraction is performed based on location of eyebrow, eye, and mouth. They use Artificial Neural Network Back propagation method for facial expression detection. The output data is an index, which automatically select and play the music. In this way, the music is modified according to the changes of facial expression. They designed a system to detect three facial expressions: normal, angry, and happy expression. The similarity between features values from each expression influence the ability to differentiate each expression. Offline system evaluation is performed with back propagation neural network method, for learning process, it reaches convergent value with lowest error value when using 10 unit neuron on hidden layer, learning rate value is 0.0625325 and mean square error value is 0.0135.

The paper by Henal Shah et al [4] conveys our proposed intelligent music player using sentimental or emotion analysis. The Emotions are a basic part of human nature. They play a vital role throughout the life. Human emotions are meant for mutual understanding and sharing feelings and intentions. The emotions are represented in verbal and through facial expressions. One can also express his emotions through written text. This paper mainly focus on the methodologies available for detecting human emotions for developing emotion based music player, the approaches used by available music players to detect emotions, the approach a music player follows to detect human emotions and how it is better to use the proposed system for emotion detection. It also gives brief idea about our system‟s working, playlist generation and emotion classification.

Anukriti Dureha [5] suggested manual segregation of a playlist and annotation of songs, in accordance with the current emotional state of a user, as a labour intensive and time consuming job. Numerous algorithms have been proposed to automate this process. However the existing algorithms are slow, increase the overall cost of the system by using additional hardware (e.g. EEG systems and sensors) and have less accuracy. This paper presents an algorithm that automates the process of generating an audio playlist, based on the facial expressions of a user, for rendering

salvage of time and labour, invested in performing the process manually. The algorithm proposed in this paper aims at reducing the overall computational time and the cost of the designed system. It also aims at increasing the accuracy of the designed system. The facial expression recognition module of the proposed algorithm is validated by testing the system against user dependent and user independent dataset. Experimental results of the process indicate that the user dependent results give 100% accuracy, while user independent results for joy and surprise are 100 %, but for sad, anger and fear are 84.3 %, 80 % and 66% respectively. The overall accuracy of the emotion recognition algorithm, for user independent dataset is 86%. In audio, 100 % recognition rates are obtained for sad, sad-anger and joy-anger but for joy and anger, recognition rates obtained are 95.4% and 90 % respectively. The overall accuracy of the audio emotion recognition algorithm is 98%. Implementation and testing of the proposed algorithm is carried out using an inbuilt camera. Hence, the proposed algorithm reduces the overall cost of the system successfully. Also, on average, the proposed algorithm takes 1.10 sec to generate a playlist based on facial expression.

Currently there are different methodology proposed by researchers to classify the emotional state of human behaviour. We have only focused on some of the basic emotion of human.

A precise and efficient approach for examine the extracted facial expression was developed by Renuka R. Londhe et al. These document mainly focused on the study of changes in the facial curve and it also focus on the intensity of the corresponding pixels. The artificial neural networks (ANN) were used to classify the characteristics extracted in 6 main universal emotions such as anger, disgust, fear, happiness, sadness and surprise. A scaled conjugate gradient back propagation algorithm correlated with a two-layer neural network was used and achieved a detection rate of 92.2%. In order to reduce the human effort and time required to manually separate songs from a playlist, different approaches have been proposed in correlation with different classes of emotions and moods.

Thayer [16] proposed a very useful two-dimensional model (stress energy v / s), plotted on two axes and whose emotions are represented by a two-dimensional

coordinate system based on two axes or on the four quadrants, which is represented by the two-dimensional diagram be formed. , The musical mood names and AV values of a total of 20 subjects were tested and analyzed in the work of Jung Hyun Kim [7]. Based on the results of the analysis, the aircraft AV was divided into 8 regions (clusters), which illustrate the mood using an efficient data mining algorithm for k-means clusters.

Galen Chuang et al. [2] made emotisphere, which is an intelligent sensor based device, which produce music depend upon client's present anxious state. This device translate physiological marker for emotion which are detected by a sensor called galvanic skin sensor and an impulse sensor, into composition of sound and light.This instrument works when user puts their hands on the ball to start recording and can hear the result immediately.

W. Amelia et al. [3] created a hybrid method that uses a combination of a keyword recognition technique and a learning method. Emotion recognition is based on Paul Ekman's basic emotions, which are anger, disgust, fear, happiness, sadness and surprise. The learning-based method used three algorithms: the multinomial logistic regression, the support vector machine (SVM) and the multinomial naive Bayes. Here the entry is a short story and the system determines the type of emotion that would be induced in the reader. This technique uses several learning methods to deduce emotions, which makes the model expensive to calculate.

A. Metallinou et al. [4] investigated how emotional information is conveyed through facial and voice modalities and how these modalities can be used effectively to improve the accuracy of emotion recognition. Markers are placed on different areas of the actors' faces, and the data for each marker is summed, and the net value is the GMM for the face. Based on the GMM, the respective emotion is retrieved from the database. This technique requires that these markers be used whenever the emotion needs to be extrapolated.

Chien Hung Chen et al. [7] classify the songs by dividing each song into voice clips, which are then divided into 2 parts to know names and to abstain. The functions

are calculated for the name and abstentions. The problem with this technique is that features are computed twice for the same song because all features are used by both the principal and the chorus.

Jung Hyun Kim et al. [1] Creation of a musical ambience model based on probabilities and implementation of a music recommendation system using the ambience model. Their pattern could express the complex mood of a song and generate a list of similar songs for multiple entries, mood tags, a song, and a value for valence excitation. Problems with their style that make it difficult to classify and express a song in a mood day or region of the Valencia excitement plan, different people feel different after listening to the same song and they haven't thought about the mood model adapt to the music and track the mood model of mood changes of users.

Many approaches have been developed to extract the facial and audio properties of an audio signal. There are very few systems available that have the ability to create an emotionbased songs playlist using human emotions. The few existing system designed which can create an playlist automatically. But they used an additional devices. The devices such as sensors or EEG systems. Using such devices additionally increases the overall cost of the proposed design. Some of the disadvantages of the existing system are as follows. Existing systems are very complex, it is complex in terms of time to extract facial features in real time. Existing systems can create a playlist but with less accuracy.

Thus, it yields better performance, in terms of computational time, as compared to the algorithms in the existing literature.

# CHAPTER 3

# PROPOSED SYSTEM

## 3.1 ABOUT PROPOSED SYSTEM

The proposed system tries to provide an interactive way for the user to carry out the task of creating a playlist. the working is based on different mechanisms carrying out their function in a pre-defined order to get the desired output. The working can be stated as follows:

1. The proposed System works by first providing a simple enough interface which prompts the user to scan the memory for audio files when the application is opened.

2. Then after the files are detected, they are scanned for audio features and these features are extracted.

3. Then the extracted feature values are subjected to classification according to the parameters provided.

4. These parameters include a limited set of genre types based on which the audio feature values will be processed.

5. After this, the songs are segregated into different playlists based on the feature extraction process. Hence lists of similar sounding songs or songs belonging to similar genres are generated.

6. In the next step, the user camera is invoked with proper permissions and a real time graphical input (image) is provided to the system.

7. The system first checks for the presence of a face in the input using the face detection process , then classifies the input and generates an output which is an emotion (mood) based on the expression extracted from the real time graphical input.

8. After this the classified expression acts as an input and is used to select an appropriate playlist from the initially generated playlists and the songs from the playlists are played.

# CHAPTER 4

# SYSTEM OVERVIEW

The system requests for the user to provide it with his/her image for which the mood can be recognized. After receiving the image, the pre-processing of image is carried out where the image is enhanced because the face detection phase needs a clear image without any noise or blurriness for processing. Face detection is done with the help of the Viola-Jones algorithm alongside Bounding Box technique for detecting the essential facial features.

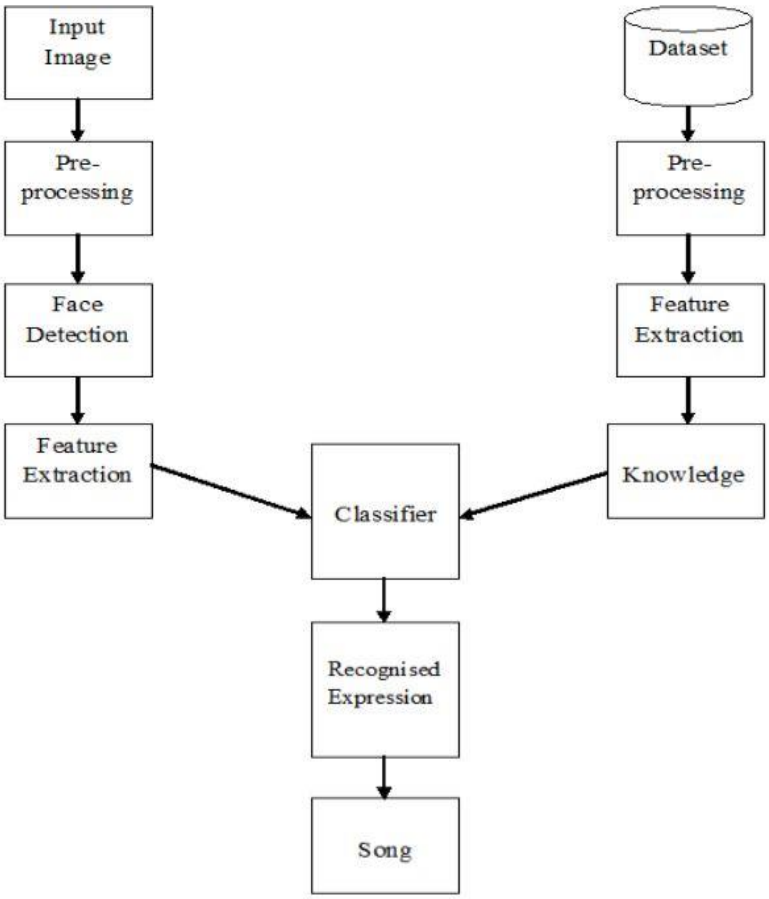



**Figure : 4.1** System Overview

A image dataset is used for the purpose of training phase which consists of various images labeled in an appropriate manner for every emotion making up for a large collection of images. The user provided test image is then trained with the dataset images and with the help of minimum Euclidean distance to classify the mood of the user in action.

## 4.1 SYSTEM REQUIREMENTS

For the system, Amazon Web Services (AWS) is used as the platform for entire development purpose as it supports numerical computation, matrix based calculations, algorithm implementation and creating user interface and provides the system developer various tools for user interface designing.

The Computer Vision System Toolbox (OpenCV) provides algorithms and functions to design as well as simulate computer vision and video processing systems. Feature detection and extraction, matching along with object detection and tracking are also supported.

The users will need to have well functioning computer system ensuring approximate RAM, processor, Hard Drive and Operating System requirements.

## 4.2 SYSTEM ARCHITECTURE

The system architecture for the proposed system takes the input image is loaded into the system in .jpg format. Then each image undergoes preprocessing i.e. removal of unwanted information like background colour, illumination and resizing of the images. Then the required features are extracted from the image and stored as useful information. These features are later added to the classifier where the expression is recognised with the help of AWS Rekognition technique. Minimum the value of the distance calculated, the nearest the match will be found. Finally, a music track will be played based on the emotion detected of the user.

## 4.3 STEPS INVOLVED TO DESIGN THE SYSTEM

To design the system, training dataset and test images are considered for which the following procedures are applied to get the desired results. The training set is the raw data which has large amount of data stored in it and the test set is the input given for recognition purpose. The whole system is designed in 3 steps:

### 4.3.1. Image Acquisition

In any of the image processing techniques, the first task is to acquire the image

from the source. These images can be acquired either through camera or through standard datasets that are available online. The images should be in .jpg format. The images considered here are user dependent i.e. dynamic images. The number of sample training images considered here is 1 as AWS contains its own sample images in its system.

### 4.3.2. Pre-processing

Pre-processing is mainly done to eliminate the unwanted information from the image acquired and fix some values for it, so that the value remains same throughout. In the pre-processing phase, the images are converted from RGB to Gray-scale and are resized to 256*256 pixels. The images considered are in .jpg format, any other formats will not be considered for further processing. During pre-processing, eyes, nose and mouth are considered to be the region of interest. It is detected by the cascade object detector which utilizes AWS Rekognition technique.

### 4.3.3. Facial Feature Extraction

After pre-processing, the next step is feature extraction. The extracted facial features are stored as the useful information in the form of vectors during training phase and testing phase. The following facial features can be considered “Mouth, forehead, eyes, complexion of skin, cheek and chin dimple, eyebrows, nose and wrinkles on the face”. In this work, eyes, nose, mouth and forehead are considered for feature extraction purpose for the reason that these depict the most appealing expressions. With the wrinkles on the forehead or the mouth being opened one can easily recognise that the person is either surprised or is fearful. But with a person’s complexion it can never be depicted. To extract the facial features AWS Rekognition technique is used.

# CHAPTER 5

# METHODOLOGY

## 5.1 BLOCK DIAGRAM

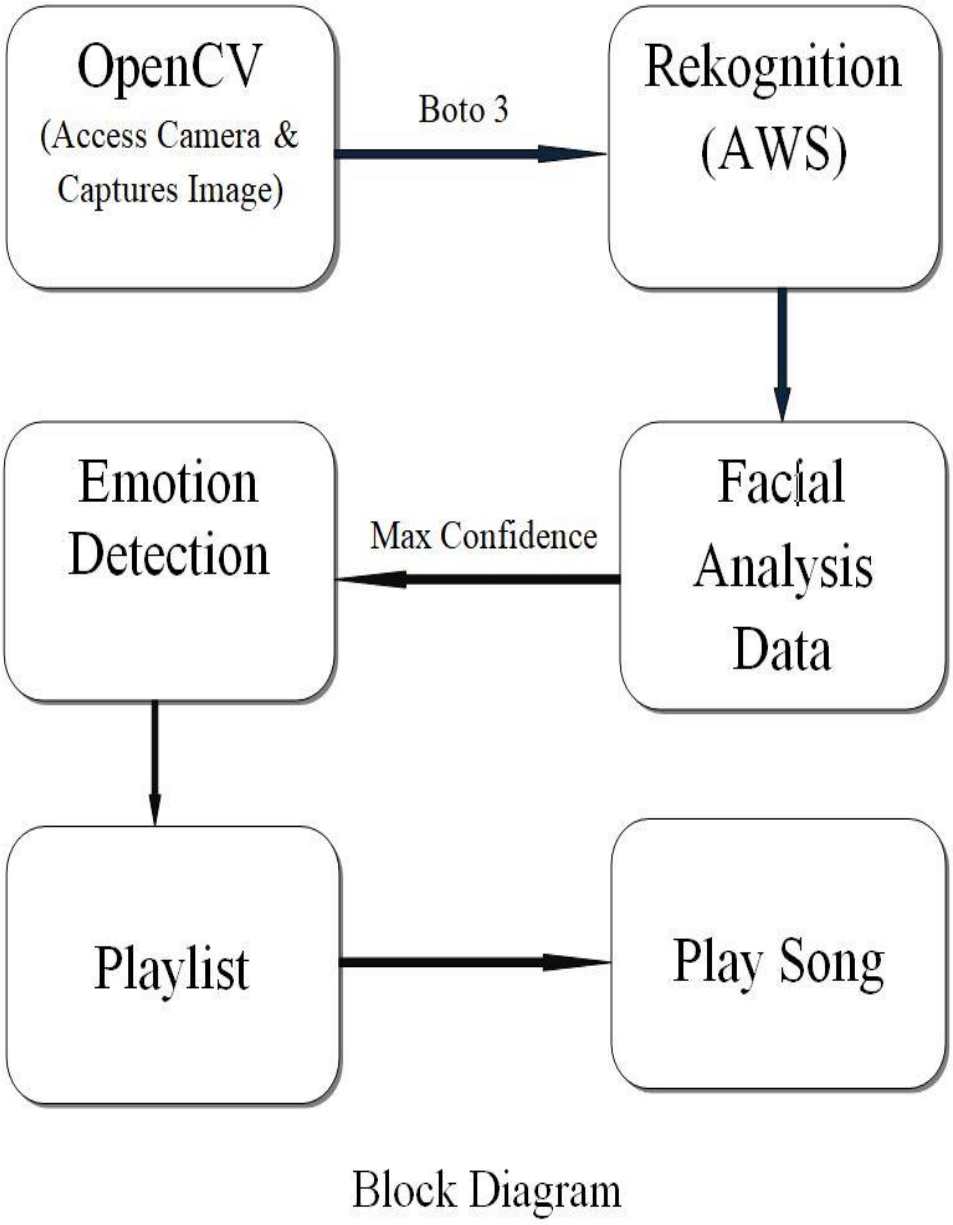



**Figure : 5.1** Block Diagram

## 5.2 OPENCV

OpenCV was started at Intel in 1999 by Gary Bradsky, and the first release came out in 2000. Vadim Pisarevsky joined Gary Bradsky to manage Intel's Russian software OpenCV team. OpenCV is a cross-platform library using which we can develop real-time computer vision applications. It mainly focuses on image

processing, video capture and analysis including features like face detection and object detection. OpenCV now supports a multitude of algorithms related to Computer Vision and Machine Learning and is expanding day by day. OpenCV supports a wide variety of programming languages such as C++, Python, Java, etc., and is available on different platforms including Windows, Linux, OS X, Android, and IOS. OpenCV-Python is the Python API for OpenCV, combining the best qualities of the OpenCV C++ API and the Python language.

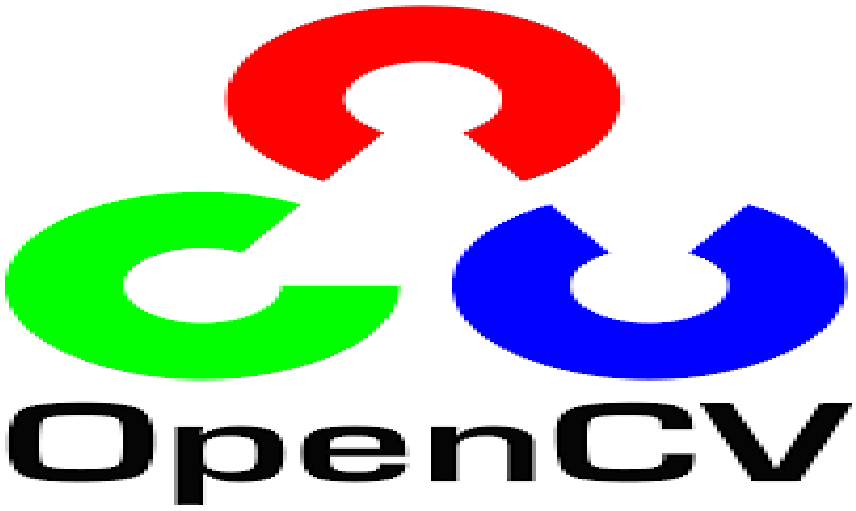



**Figure : 5.2** OpenCV

Open CV is a library of programming functions mainly aim at real time computer vision. It is a C++ implementation library. There is a javacv library which is derived from OpenCV using this we will implement viola and Jones face detection algorithm. Face detection is important as it will classify only if face is present. Expression recognition is also done using graphical based classification method. Audio files will be scanned and features will be extracted from them and according to the mood we get the playlist

OpenCV is the huge open-source library for the computer vision, machine learning, and image processing and now it plays a major role in real-time operation which is very important in today's systems. By using it, one can process images and vedios to identify objects, faces, or even handwriting of a human.

### 5.2.1 Applications

- Face recognition
- Automated inspection and surveillance
- number of people – count (foot traffic in a mall, etc)
- Vehicle counting on highways along with their speeds
- Interactive art installations
- Anamoly (defect) detection in the manufacturing process (the odd defective products)
- Street view image stitching
- Video/image search and retrieval
- Robot and driver-less car navigation and control
- object recognition
- Medical image analysis
- Movies – 3D structure from motion
- TV Channels advertisement recognition.

### 5.2.2 OpenCV Functionality

- Image/video I/O, processing, display (core, imgproc, highgui)
- Object/feature detection (objdetect, features2d, nonfree)

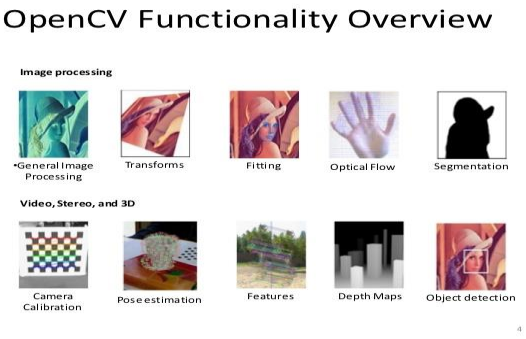



**Figure : 5.3** OpenCV Funtionality

- Geometry-based monocular or stereo computer vision (calib3d, stitching, videostab)
- Computational photography (photo, video, superres)
- Machine learning & clustering (ml, flann)

## 5.3 Boto 3 Library

Boto3 is the Amazon Web Services (AWS) Software Development Kit (SDK) for Python, which allows Python developers to write software that makes use of services like Amazon S3 and Amazon EC2. You can find the latest, most up to date, documentation at our doc site, including a list of services that are supported.


**Figure : 5.4** Boto 3

Boto derives its name from the Portuguese name given to types of dolphins native to the Amazon River. Boto is a Python package that provides interfaces to Amazon Web Services. Currently, all features work with Python 2.6 and 2.7. Work is under way to support Python 3.3+ in the same codebase.

Boto3 is acts as mediator between Open CV and Amazon web services. OpenCV is used to detect the images and vedios, Boto3 is the library which helps to send the captured image by OpenCV to the Amazon Web Services.

## 5. 4 AMAZON WEB SERVICES

Amazon web service is a platform that offers flexible, reliable, scalable, easy-to-use and cost-effective cloud computing solutions. AWS is a comprehensive, easy to use computing platform offered Amazon. The platform is developed with a combination of infrastructure as a service (IaaS), platform as a service (PaaS) and packaged software as a service (SaaS) offerings.

**Figure : 5.5** AWS

Amazon Web Services offers a wide range of different business purpose global cloud-based products. The products include storage, databases, analytics, networking, mobile, development tools, enterprise applications, with a pay-as-you-go pricing model.


**Figure : 5.6** AWS Functions

There are so many services are given by the amazon web services among them Amazon recognition is one of the service. It is used to recognize the image.


**Figure : 5.7** AWS Services

# CHAPTER 6

# AMAZON REKOGNITION

Amazon Rekognition makes it easy to add image and video analysis to your applications. You just provide an image or video to the Amazon Rekognition API, and the service can identify objects, people, text, scenes, and activities. It can detect any inappropriate content as well. Amazon Rekognition also provides highly accurate facial analysis, face comparison, and face search capabilities. You can detect, analyze, and compare faces for a wide variety of use cases, including user verification, cataloging, people counting, and public safety.

Amazon Rekognition is based on the same proven, highly scalable, deep learning technology developed by Amazon's computer vision scientists to analyze billions of images and videos daily. It requires no machine learning expertise to use. Amazon Rekognition includes a simple, easy-to-use API that can quickly analyze any image or video file that's stored in Amazon S3. Amazon Rekognition is always learning from new data, and we're continually adding new labels and facial comparison features to the service.

Common use cases for using Amazon Rekognition include the following:

• Searchable image and video libraries – Amazon Rekognition makes images and stored videos searchable so you can discover objects and scenes that appear within them.

• Face-based user verification – Amazon Rekognition enables your applications to confirm user identities by comparing their live image with a reference image.

• Sentiment and demographic analysis – Amazon Rekognition interprets emotional expressions such as happy, sad, or surprise, and demographic information such as gender from facial images. Amazon Rekognition can analyze images, and send the emotion and demographic attributes to Amazon Redshift for periodic reporting on trends such as in store locations and similar scenarios. Note that a prediction of an emotional expression is based on the physical appearance of a person's face only. It is not indicative of a person's internal emotional state, and Rekognition should not be used to make such a determination.

• Facial Search – With Amazon Rekognition, you can search images, stored videos, and streaming videos for faces that match those stored in a container known as a face collection. A face collection is an index of faces that you own and manage. Searching for people based on their faces requires two major steps in Amazon Rekognition:

Index the faces and Search the faces are two steps.


**Figure : 6.1** AWS Rekognition

• Unsafe content detection – Amazon Rekognition can detect adult and violent content in images and in stored videos. Developers can use the returned metadata to filter inappropriate content based on their business needs. Beyond flagging an image based on the presence of unsafe content, the API also returns a hierarchical list of labels with confidence scores. These labels indicate specific categories of unsafe content, which enables granular filtering and management of large volumes of user-generated content (UGC). Examples include social and dating sites, photo sharing platforms, blogs and forums, apps for children, ecommerce sites, entertainment, and online advertising services. Amazon Rekognition Developer Guide

• Celebrity recognition – Amazon Rekognition can recognize celebrities within supplied images and in videos. Amazon Rekognition can recognize thousands of celebrities across a number of categories, such as politics, sports, business, entertainment, and media.

• Text detection – Amazon Rekognition Text in Image enables you to recognize and extract textual content from images. Text in Image supports most fonts, including highly stylized ones. It detects text and numbers in different orientations, such as those commonly found in banners and posters. In image sharing and social media applications, you can use it to enable visual search based on an index of images that

contain the same keywords. In media and entertainment applications, you can catalog videos based on relevant text on screen, such as ads, news, sport scores, and captions. Finally, in public safety applications, you can identify vehicles based on license plate numbers from images taken by street cameras.

• Custom labels– With Amazon Rekognition Custom Labels, you can identify the objects and scenes in images that are specific to your business needs. For example, you can find your logo in social media posts, identify your products on store shelves, classify machine parts in an assembly line, distinguish healthy and infected plants, or detect animated characters in videos.

## 6.1 How Amazon Rekognition Works

Amazon Rekognition provides two API sets. You use Amazon Rekognition Image for analyzing images, and Amazon Rekognition Video for analyzing videos. Both APIs analyze images and videos to provide insights you can use in your applications. For example, you could use Amazon Rekognition Image to enhance the customer experience for a photo management application. When a customer uploads a photo, your application can use Amazon Rekognition Image to detect real-world objects or faces in the image.

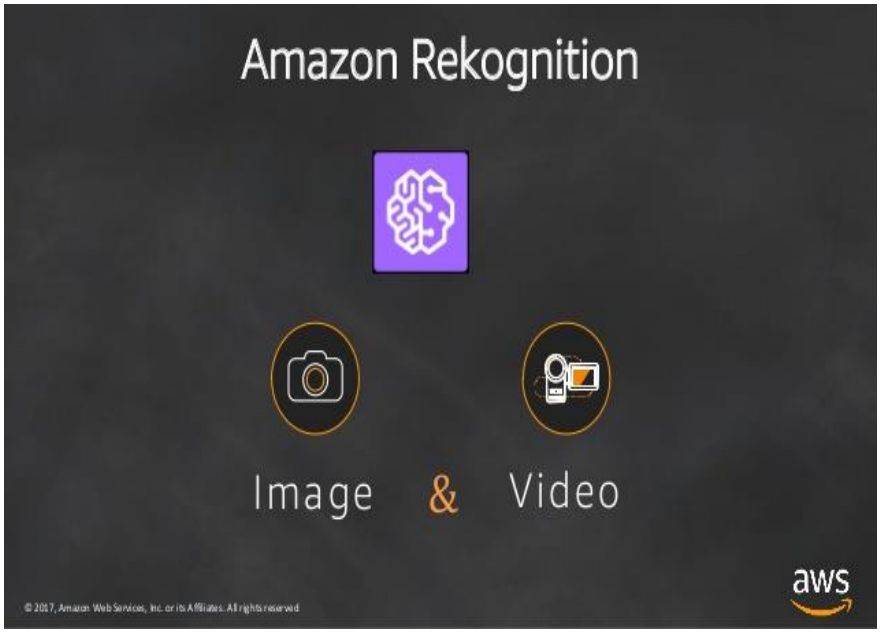



**Figure : 6.2** AWS Rekognition formats

After your application stores the information returned from Amazon Rekognition Image, the user could then query their photo collection for photos with a specific object or face. Deeper querying is possible. For example, the user could query for faces that are smiling or query for faces that are a certain age.

You can use Amazon Rekognition Video to track the path of people in a stored video. Alternatively, you can use Amazon Rekognition Video to search a streaming video for persons whose facial descriptions match facial descriptions already stored by Amazon Rekognition. The Amazon Rekognition API makes deep learning image analysis easy to use. For example, RecognizeCelebrities (p. 401) returns information for up to 100 celebrities detected in an image. This includes information about where celebrity faces are detected on the image and where to get further information about the celebrity.

Amazon Rekognition returns the following facial attributes for each face detected, along with a bounding box and confidence score for each attribute:

Gender

Smile

Emotions

Eyeglasses

Sunglasses

Eyes open

Mouth open

Mustache

Beard

Pose

Quality

Face landmarks

From all these facial attributes we need only emotion attribute to detect emotions

The seven universally classified emotions are

1. Happy
2. Sad
3. Anger
4. Disgust
5. Fear
6. Surprise
7. Neutral


**Figure : 6.3** Different Facial Expressions

## 6.2 Face Detection

In the past few years, face recognition owned significant consideration and appreciated as one of the most promising applications in the field of image analysis. Face detection can consider a substantial part of face recognition operations. According to its strength to focus computational resources on the section of an image holding a face. The method of face detection in pictures is complicated because of

variability present across human faces such as pose, expression, position and orientation, skin colour, the presence of glasses or facial hair, differences in camera gain, lighting conditions, and image resolution.

Face Detection is the first and essential step for face recognition, and it is used to detect faces in the images. It is a part of object detection and can use in many areas such as security, bio-metrics, law enforcement, entertainment, personal safety, etc.
It is used to detect faces in real time for surveillance and tracking of person or objects. It is widely used in cameras to identify multiple appearances in the frame Ex- Mobile cameras and DSLR's. Facebook is also using face detection algorithm to detect faces in the images and recognise them.

## 6.3 Applications:

- Facial motion capture
- Facial recognition
- Photography
- Marketing
- Emotional Inference
- Lip Reading

## 6.4 Face Detection Methods

These methods divided into four categories:

- *Knowledge based*
- *Feature based*
- *Template matching*
- *Appearance based*

**Knowledge Based :**

The knowledge-based method depends on the set of rules, and it is based on human knowledge to detect the faces. Ex- A face must have a nose, eyes, and mouth within certain distances and positions with each other.

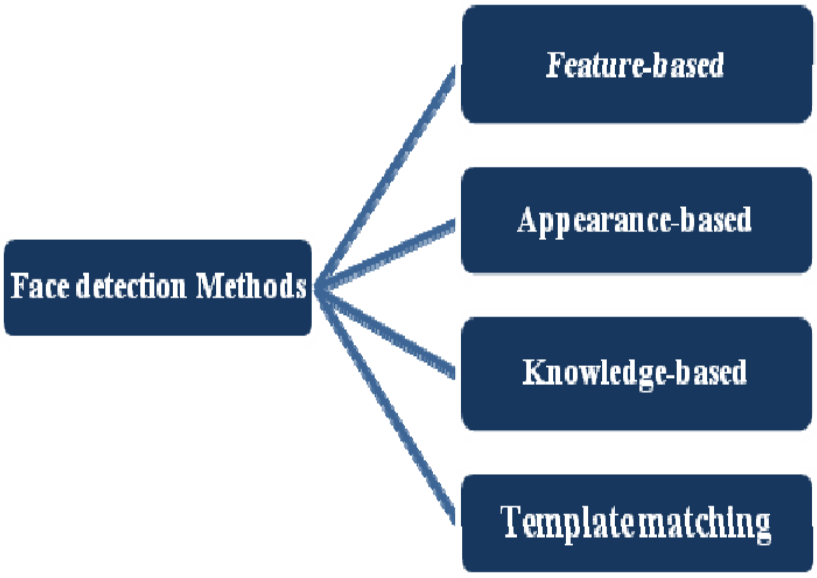



**Figure : 6.4** Face Dectection Methods

**Feature Based:**

The feature-based method is to locate faces by extracting structural features of the face. It is first trained as a classifier and then used to differentiate between facial and non-facial regions. The idea is to overcome the limits of our instinctive knowledge of faces. This approach divided into several steps and even photos with many faces they report a success rate of 94 percent.

**Template Matching:**

Template Matching method uses pre-defined or parameterised face templates to locate or detect the faces by the correlation between the templates and input images. Ex- a human face can be divided into eyes, face contour, nose, and mouth.

**Appearance Based:**

The appearance-based method depends on a set of delegate training face images to find out face models. The appearance-based approach is better than other ways of performance. In general appearance-based method rely on techniques from statistical analysis and machine learning to find the relevant characteristics of face images. This method also used in feature extraction for face recognition.

## 6.5 How the Face Detection Works:-

There are many techniques to detect faces, with the help of these techniques, we can identify faces with higher accuracy. These techniques have an almost same procedure for Face Detection such as OpenCV, Neural Networks, Matlab, etc. The face detection work as to detect multiple faces in an image.


**Figure : 6.5** Detection of Facial Expression with accuracy level

Here we work on OpenCV for Face Detection, and there are some steps that how face detection operates, which are as follows

Firstly the image is imported by providing the location of the image. Then the picture is transformed from RGB to Grayscale because it is easy to detect faces in the grayscale. After that, the image manipulation used, in which the resizing, cropping, blurring and sharpening of the images done if needed. The next step is image segmentation, which is used for contour detection or segments the multiple objects in a single image so that the classifier can quickly detect the objects and faces in the picture. The next step is algorithm, which is proposed by AWS Rekognition for face detection. This algorithm used for finding the location of the human faces in a frame or image. All human faces shares some universal properties of the human face like the eyes region is darker than its neighbour pixels and nose region is brighter than eye region.

# CHAPTER 7

# EXPERIMENTAL RESULT

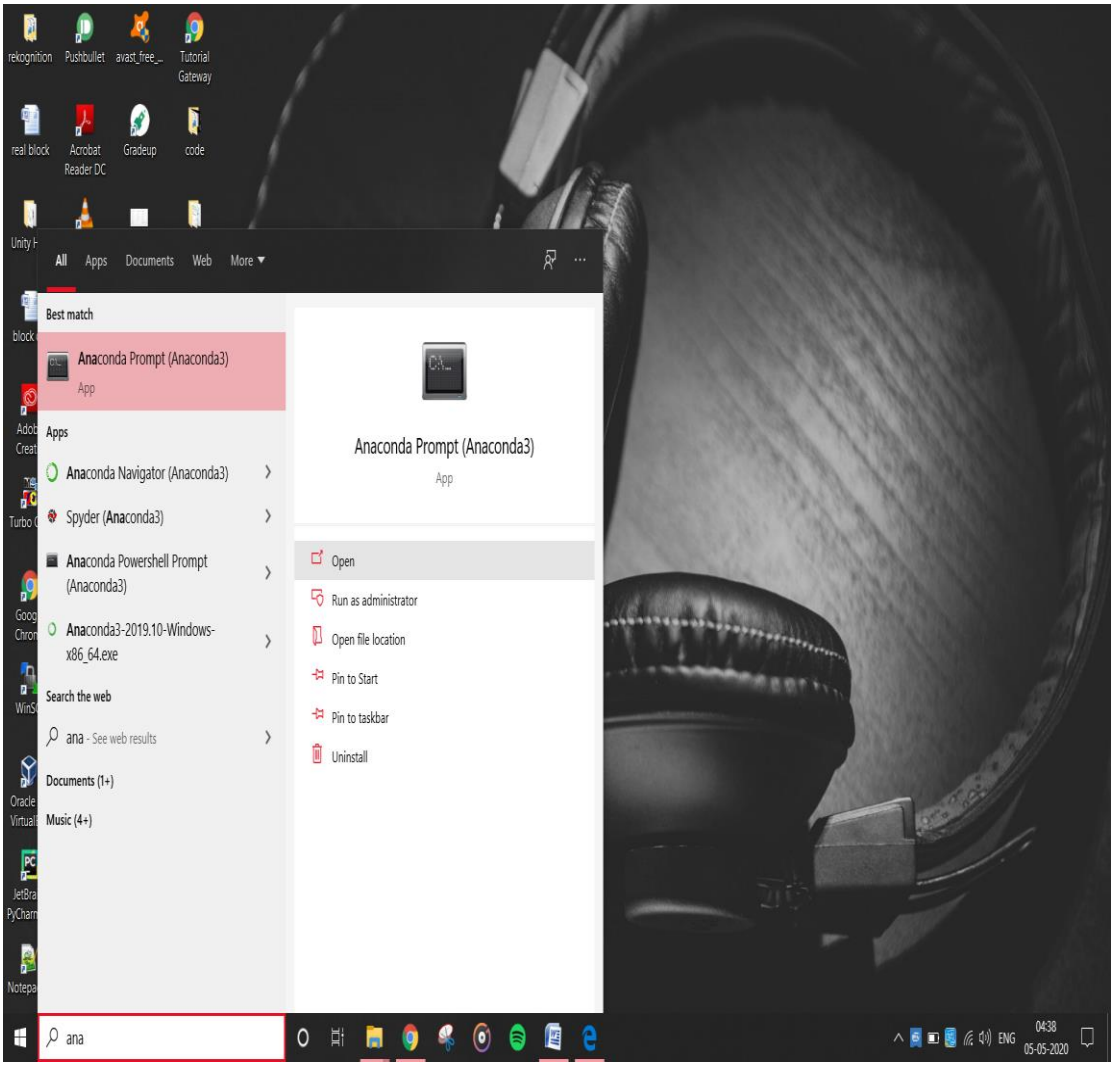
The system uses the following technologies to develop Countenance Based Boom Box. AWS Rekognition service in combination with OpenCV and Boto3 library is used in the development of the Countenance Based Boom Box.

## 7.1 TESTING AND EVALUATION

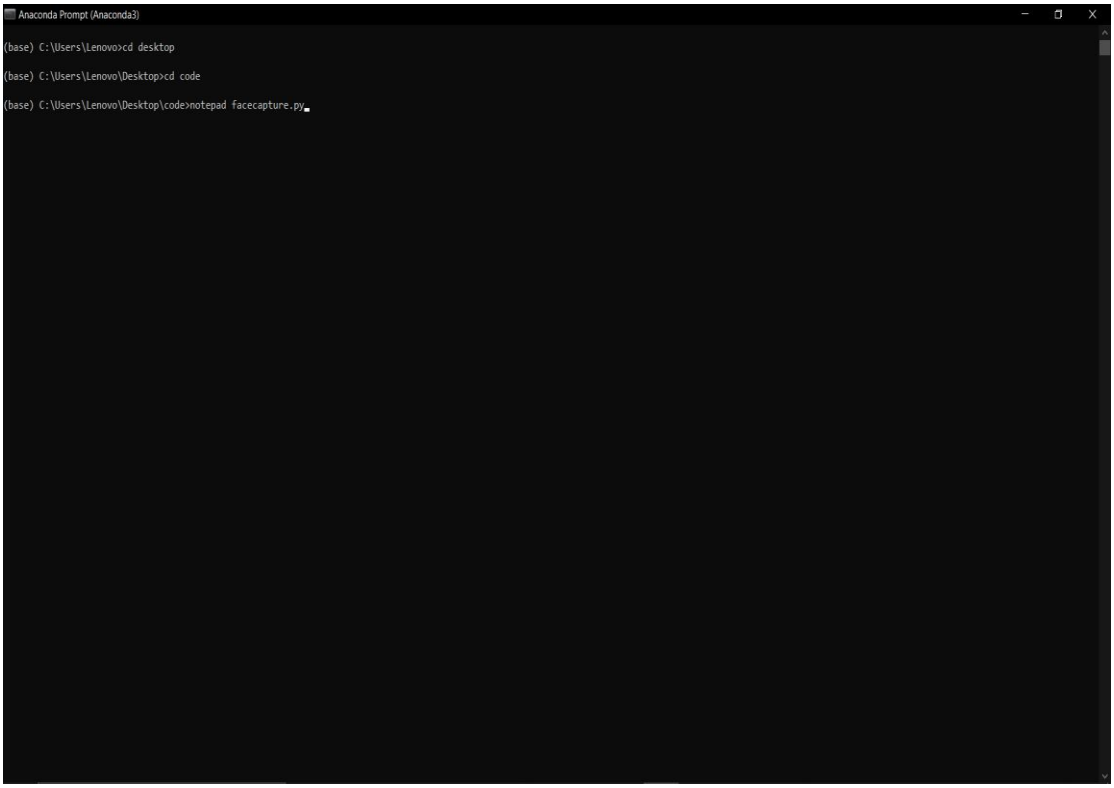
To test the software there are some basic features that should be tested, start by running the normally without any errors, to check the camera is working properly, and the most important part is to test the emotion detection accuracy functionalities, starting from capturing an image that can be used in the testing and detect the emotion and select the right song.

Here is some steps with figures for the software testing:

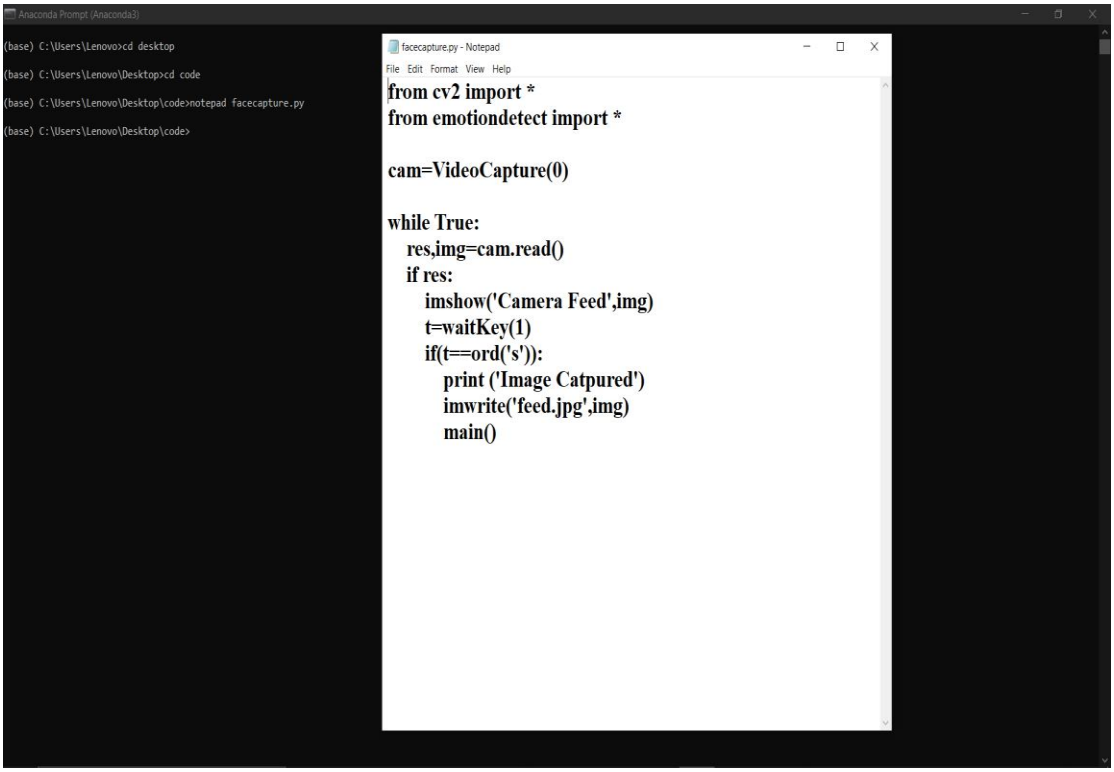
1. Open Anaconda prompt and activate the environment.



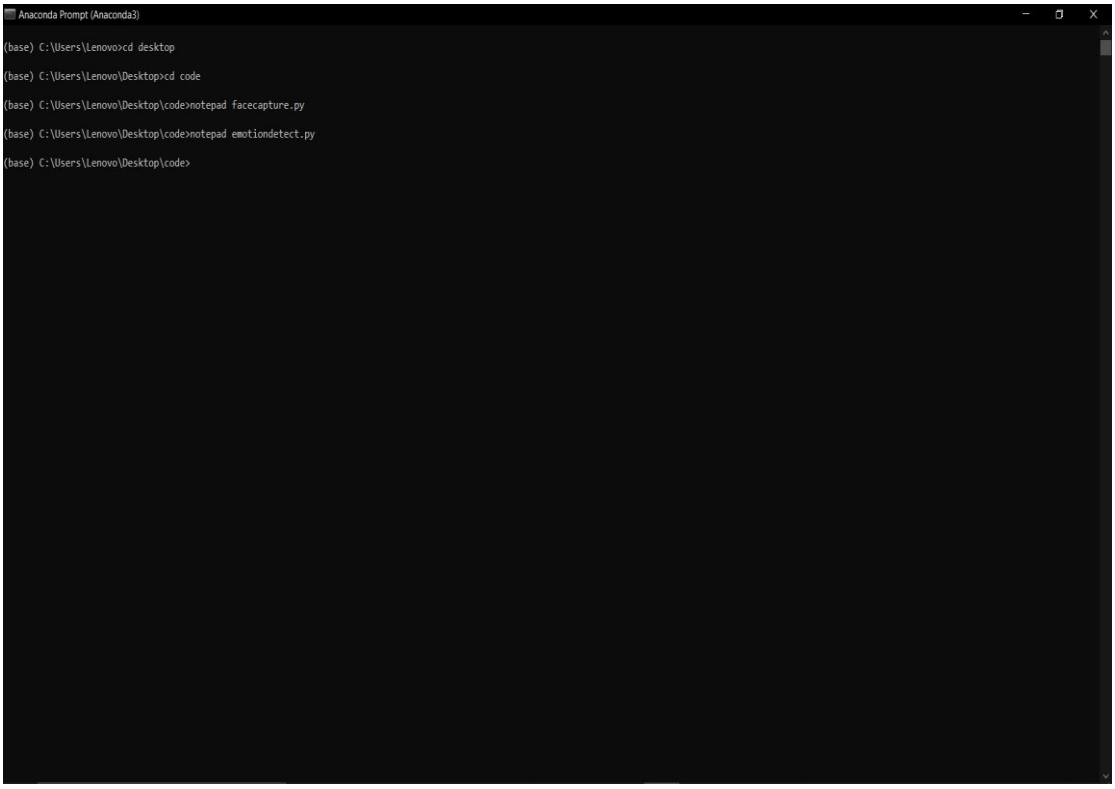
2. Open face capture file in anaconda prompt by navigating directories.



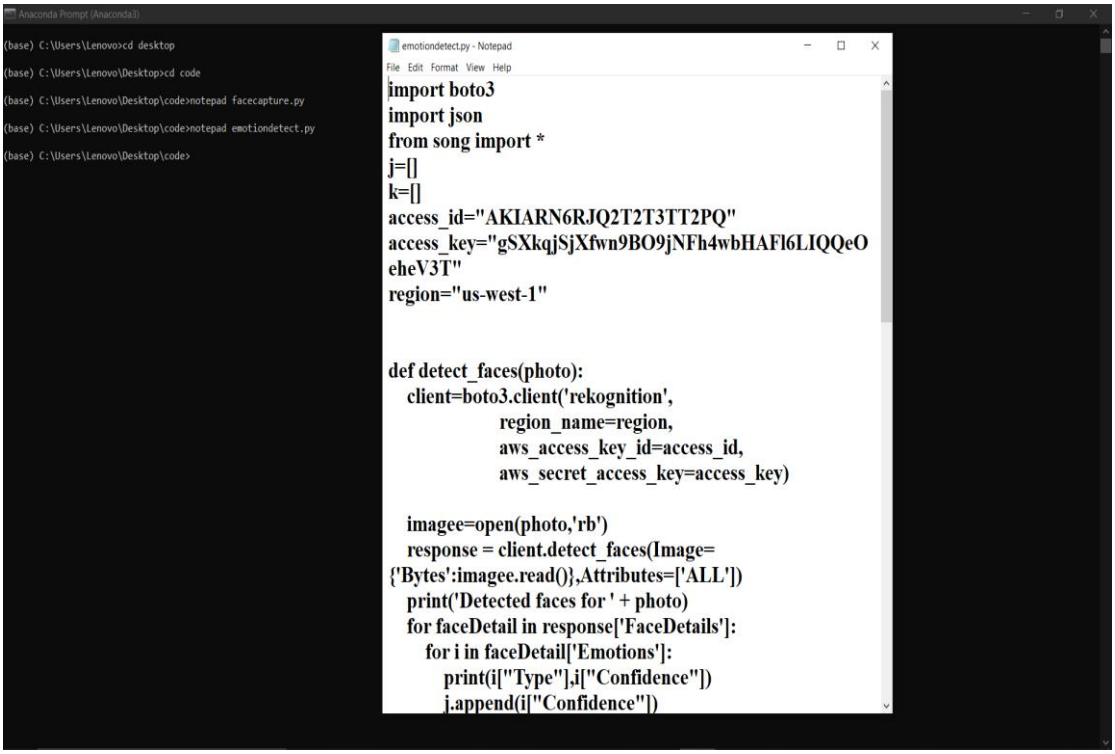

3. Code for face capture in python.

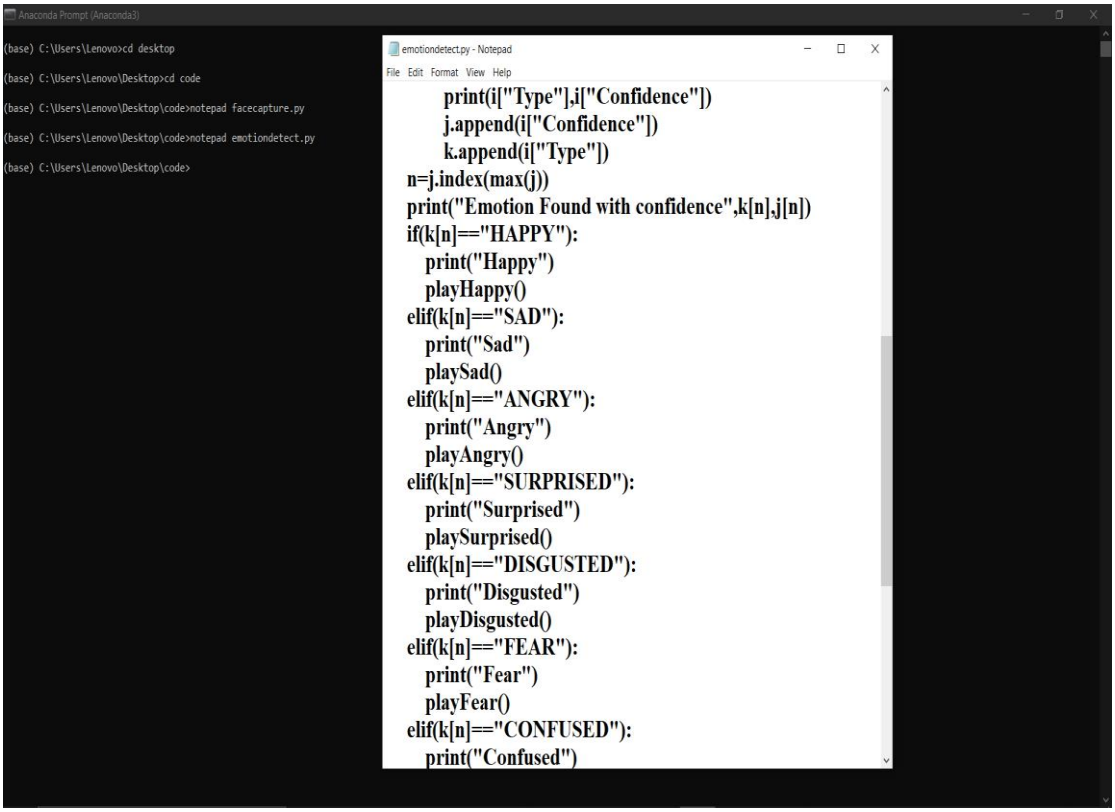


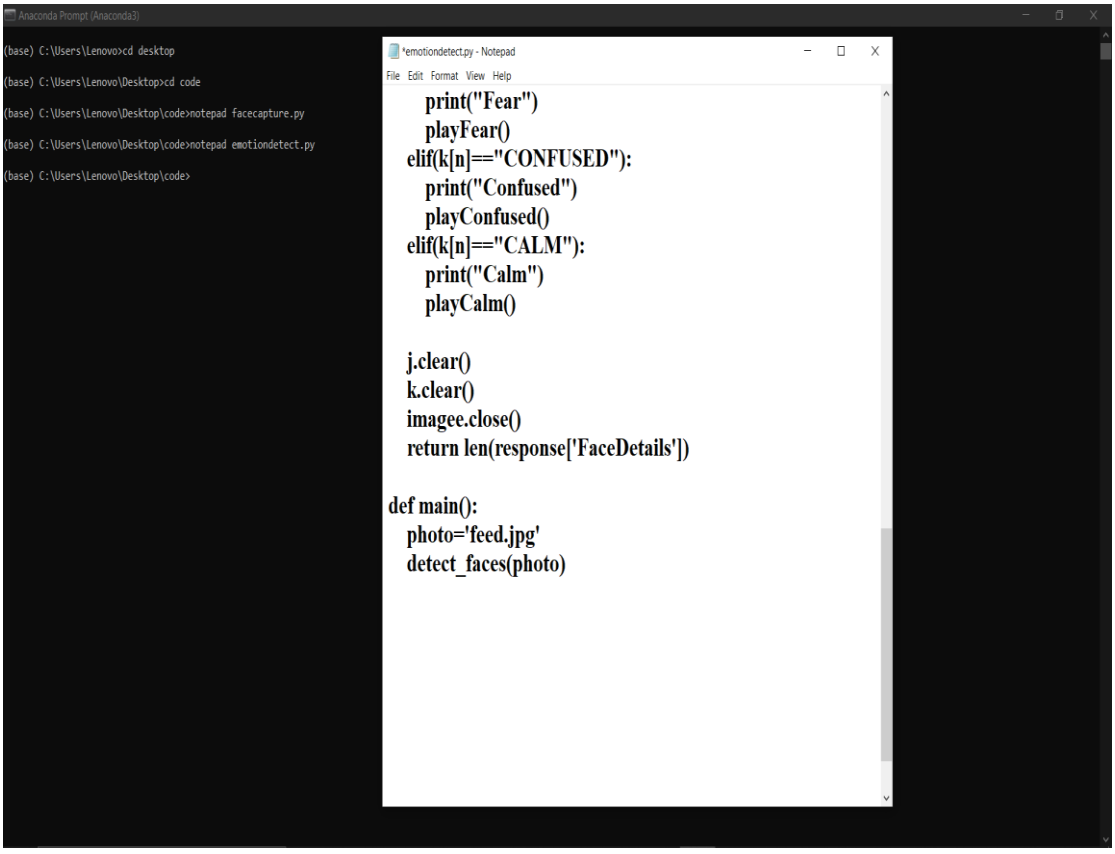
4. Open face detection file in anaconda prompt by navigating directories.



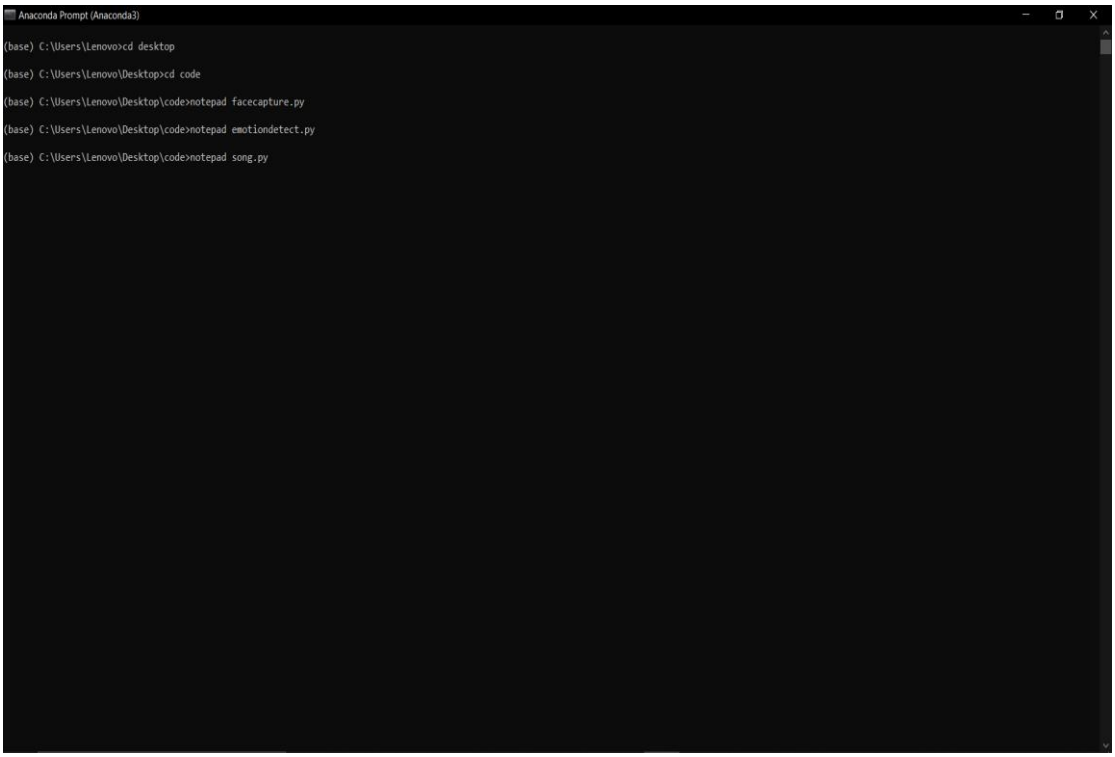

5. Code for emotion detection in python.




Anaconda Prompt (Anaconda3)
(base) C:\Users\Lenovo>cd desktop
(base) C:\Users\Lenovo\Desktop>cd code
(base) C:\Users\Lenovo\Desktop\code>notepad facecapture.py
(base) C:\Users\Lenovo\Desktop\code>notepad emotiondetect.py
(base) C:\Users\Lenovo\Desktop\code>
emotiondetect.py - Notepad
File Edit Format View Help
print(i["Type"],i["Confidence"])
j.append(i["Confidence"])
k.append(i["Type"])
n=j.index(max(j))
print("Emotion Found with confidence",k[n],j[n])
if(k[n]=="HAPPY"):
print("Happy")
playHappy()
elif(k[n]=="SAD"):
print("Sad")
playSad()
elif(k[n]=="ANGRY"):
print("Angry")
playAngry()
elif(k[n]=="SURPRISED"):
print("Surprised")
playSurprised()
elif(k[n]=="DISGUSTED"):
print("Disgusted")
playDisgusted()
elif(k[n]=="FEAR"):
print("Fear")
playFear()
elif(k[n]=="CONFUSED"):
print("Confused")


Anaconda Prompt (Anaconda3)
(base) C:\Users\Lenovo>cd desktop
(base) C:\Users\Lenovo\Desktop>cd code
(base) C:\Users\Lenovo\Desktop\code>notepad facecapture.py
(base) C:\Users\Lenovo\Desktop\code>notepad emotiondetect.py
(base) C:\Users\Lenovo\Desktop\code>
*emotiondetect.py - Notepad
File Edit Format View Help
print("Fear")
playFear()
elif(k[n]=="CONFUSED"):
print("Confused")
playConfused()
elif(k[n]=="CALM"):
print("Calm")
playCalm()
j.clear()
k.clear()
imagee.close()
return len(response['FaceDetails'])
def main():
photo='feed.jpg'
detect_faces(photo)

6. Open song file in anaconda prompt by navigating directories.




7. Code for song palying in python.

song.py - Notepad

File Edit Format View Help

```
def stop_music():
    mixer.music.stop()

def show_details(play_song):
    file_data = os.path.splitext(play_song)

    if file_data[1] == '.mp3':
        audio = MP3(play_song)
        total_length = audio.info.length
    else:
        a = mixer.Sound(play_song)
        total_length = a.get_length()

    # div - total_length/60, mod - total_length % 60
    mins, secs = divmod(total_length, 60)
    mins = round(mins)
    secs = round(secs)
    timeformat = '{:02d}:{:02d}'.format(mins, secs)
    print ("Total Length" + ' - ' + timeformat)

    t1 = threading.Thread(target=start_count, args=(total_length,))
    t1.start()

def start_count(t):
    # mixer.music.get_busy(): - Returns FALSE when we press the stop button (music stop playing)
    # Continue - Ignores all of the statements below it. We check if music is paused or not.
    current_time = 0
```

song.py - Notepad

File Edit Format View Help

```
    # Continue - Ignores all of the statements below it. We check if music is paused or not.
    current_time = 0
    while current_time <= t and mixer.music.get_busy():
        if 0:
            continue
        else:
            mins, secs = divmod(current_time, 60)
            mins = round(mins)
            secs = round(secs)
            timeformat = '{:02d}:{:02d}'.format(mins, secs)
            print ("Current Time" + ' - ' + timeformat)
            time.sleep(1)
            current_time += 1

def play_music(play_it):
    if 0:
        mixer.music.unpause()
    else:
        stop_music()
        time.sleep(1)
        mixer.music.load(play_it)
        mixer.music.play()
        show_details(play_it)


def playHappy():
 print ('Playing Happy Songs Playlist')
 k=random.choice(happy_playlist)
```

song.py - Notepad

File Edit Format View Help

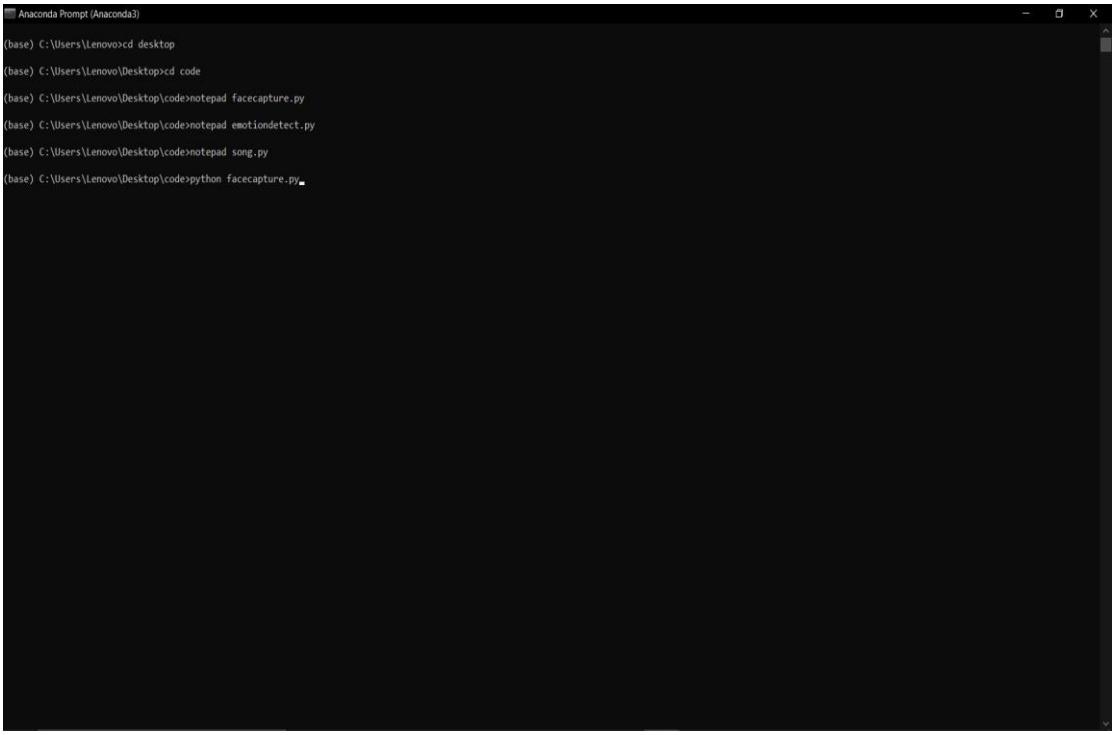
```
def playHappy():
 print ('Playing Happy Songs Playlist')
 k=random.choice(happy_playlist)
 play_music(k)

def playSad():
 print ('Playing Sad Songs Playlist')
 k=random.choice(sad_playlist)
 play_music(k)

def playAngry():
 print ('Playing Angry Songs Playlist')
 k=random.choice(angry_playlist)
 play_music(k)

def playSurprised():
 print ('Playing Surprised Songs Playlist')
 k=random.choice(surprised_playlist)
 play_music(k)

def playDisgusted():
 print ('Playing Disgusted Songs Playlist')
 k=random.choice(disgusted_playlist)
 play_music(k)

def playFear():
 print ('Playing Fear Songs Playlist')
```

*song.py - Notepad

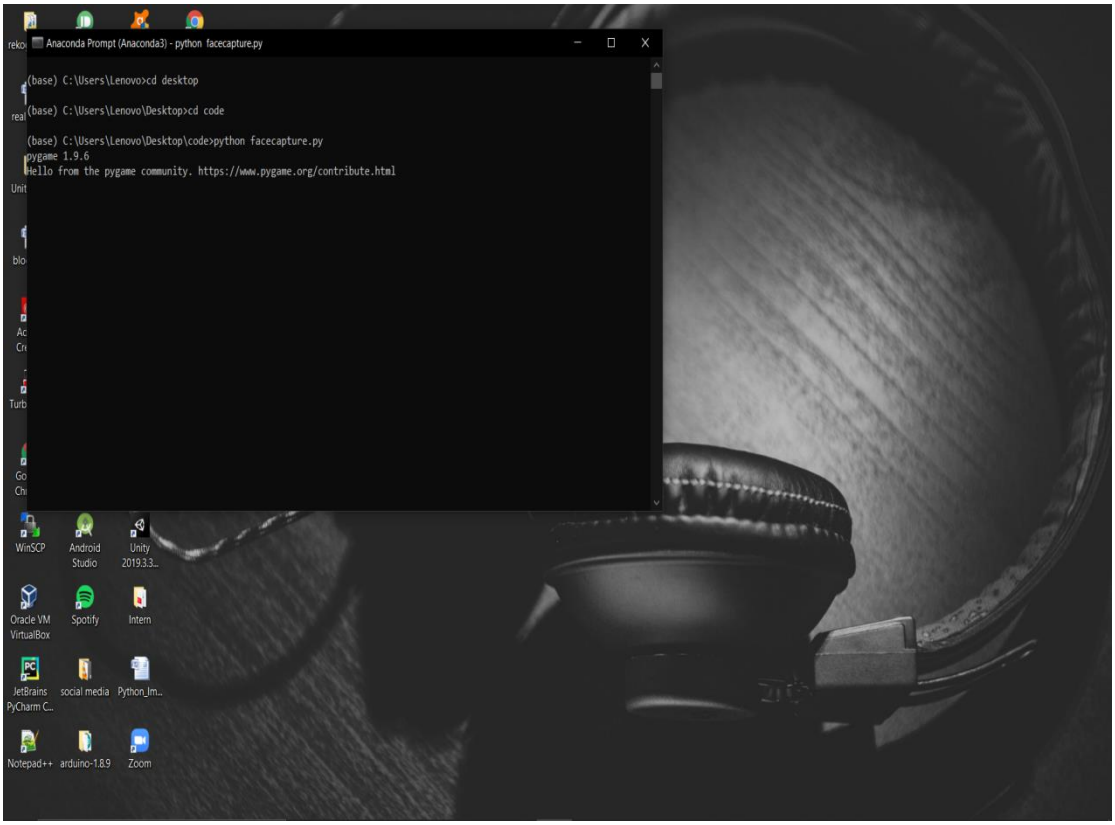
File Edit Format View Help

```
def playFear():
 print ('Playing Fear Songs Playlist')
 k=random.choice(fear_playlist)
 play_music(k)

def playConfused():
 print ('Playing Confused Songs Playlist')
 k=random.choice(confused_playlist)
 play_music(k)

def playCalm():
 print ('Playing Calm Songs Playlist')
 k=random.choice(calm_playlist)
 play_music(k)

#playHappy()
```

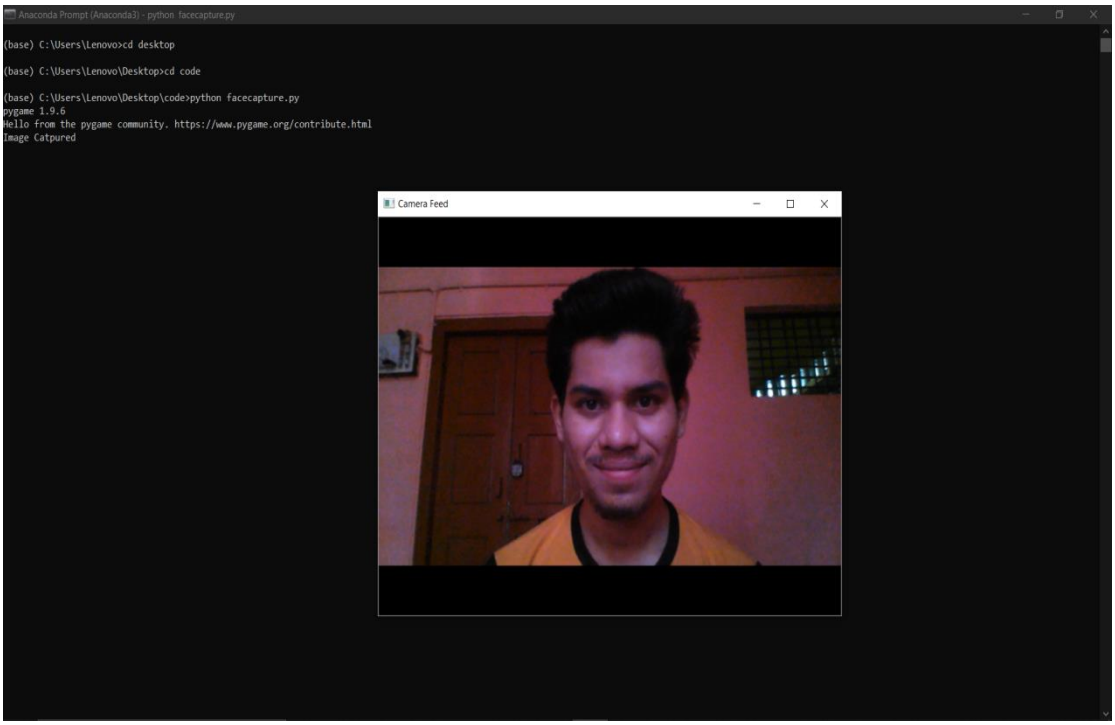
8. Run facecapture.py program.



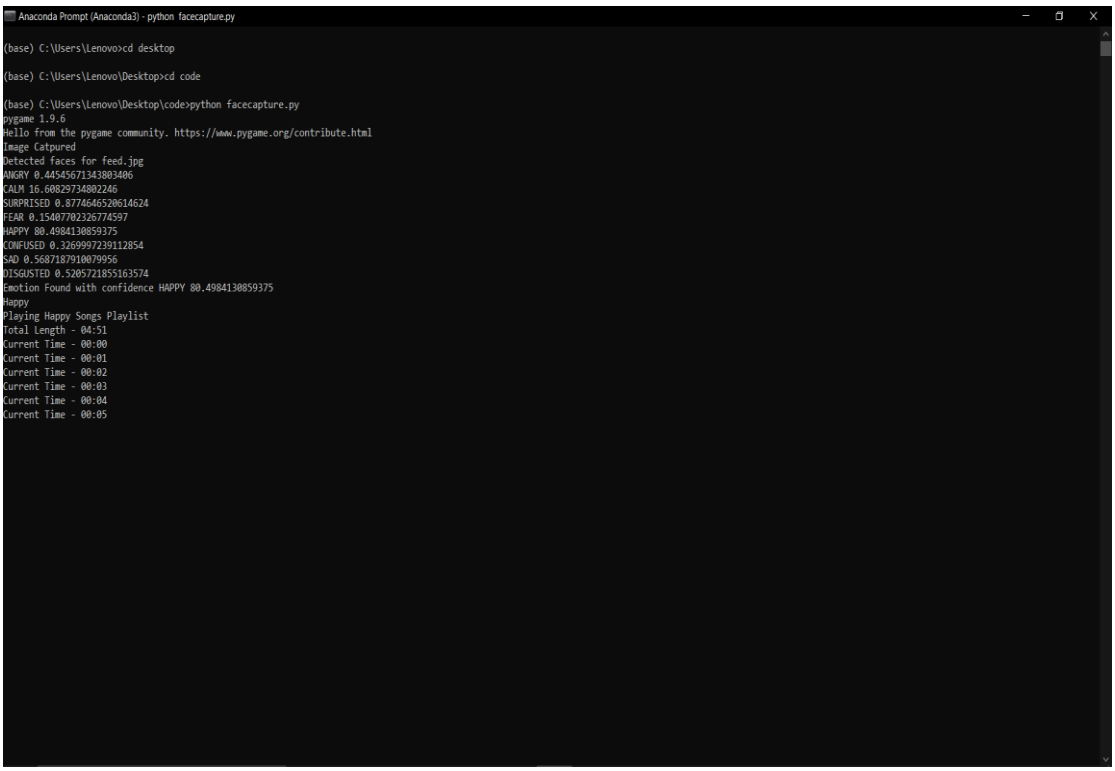

9. While Running facecapture.py program.



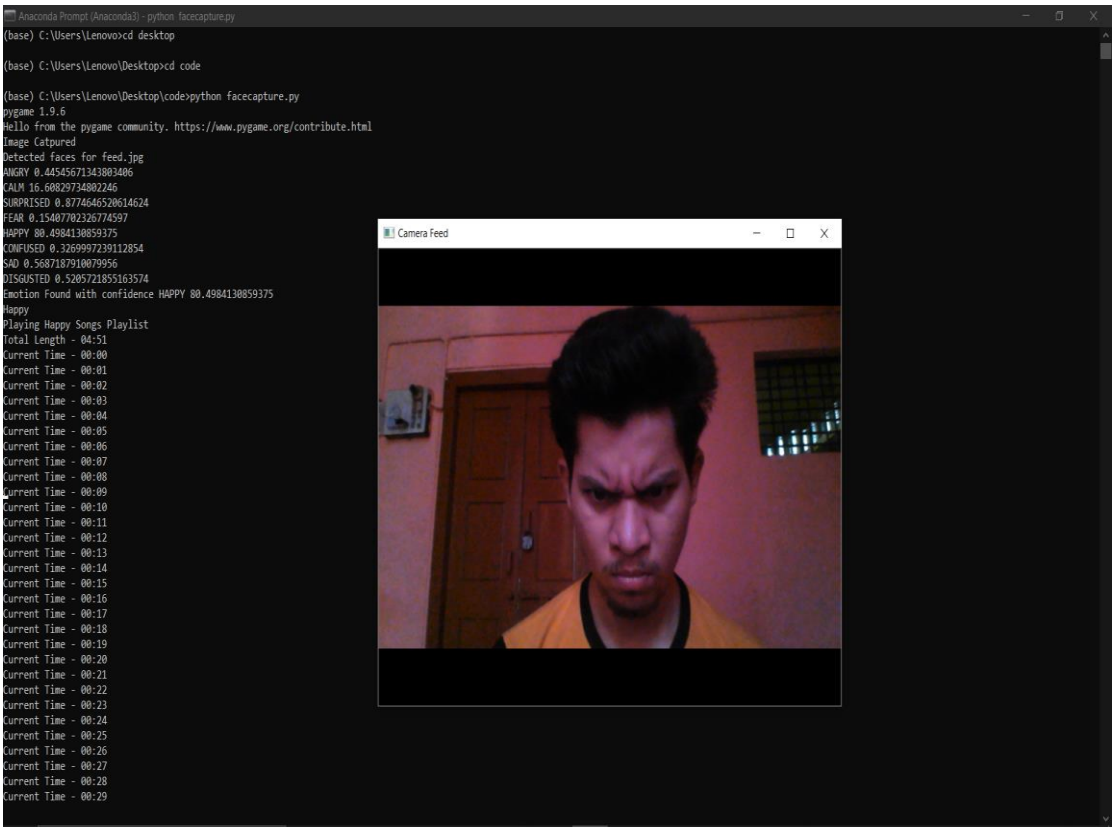
10. Camera opens for live image capturing.



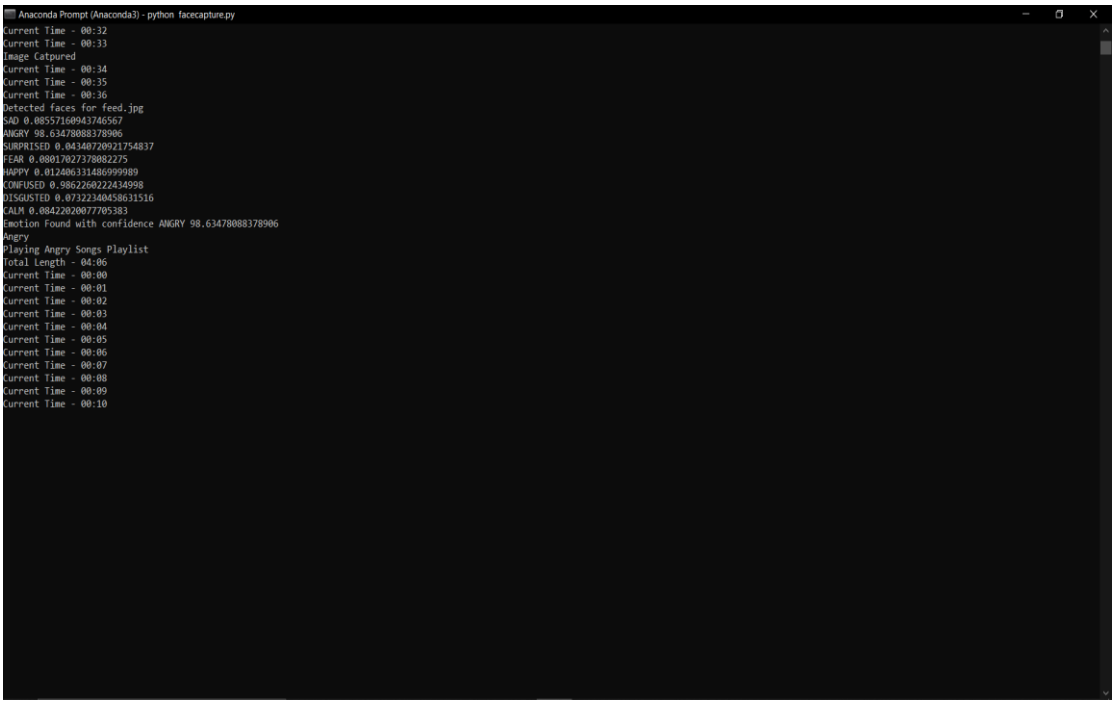

11. Detected happy emotion from the image and plays emotion based song.



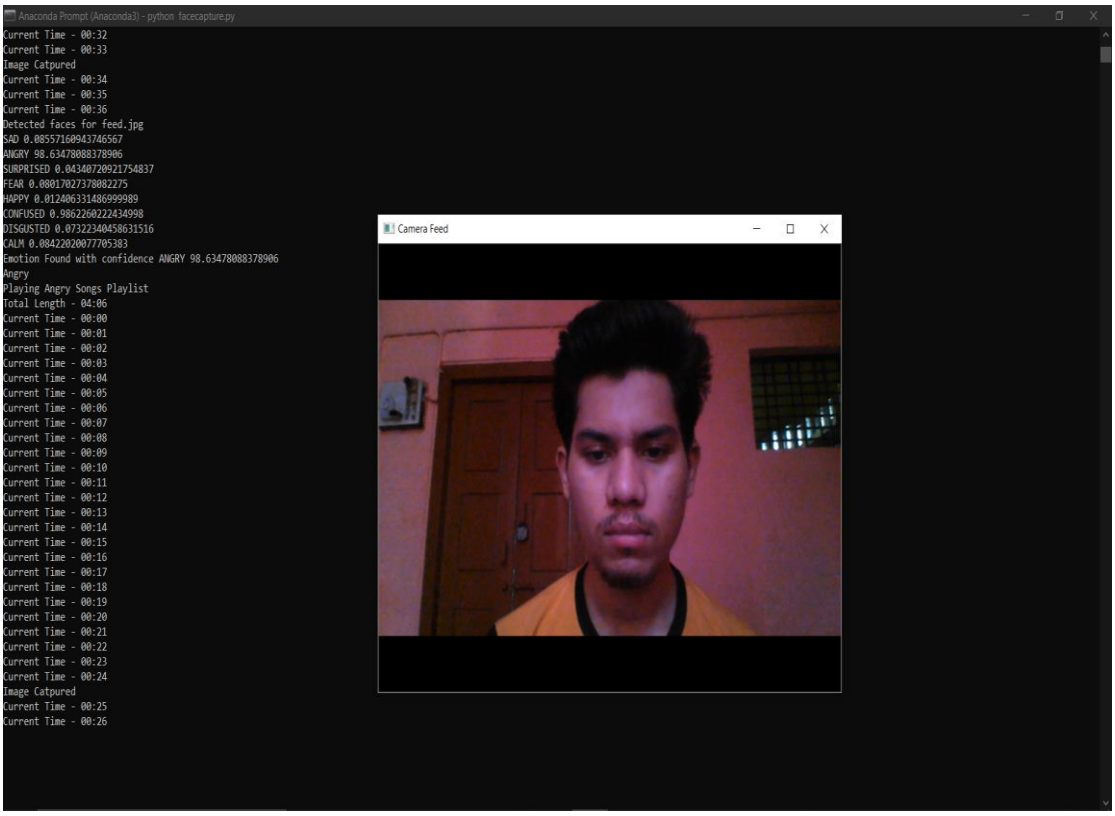
12. Camera opens for live image capturing.



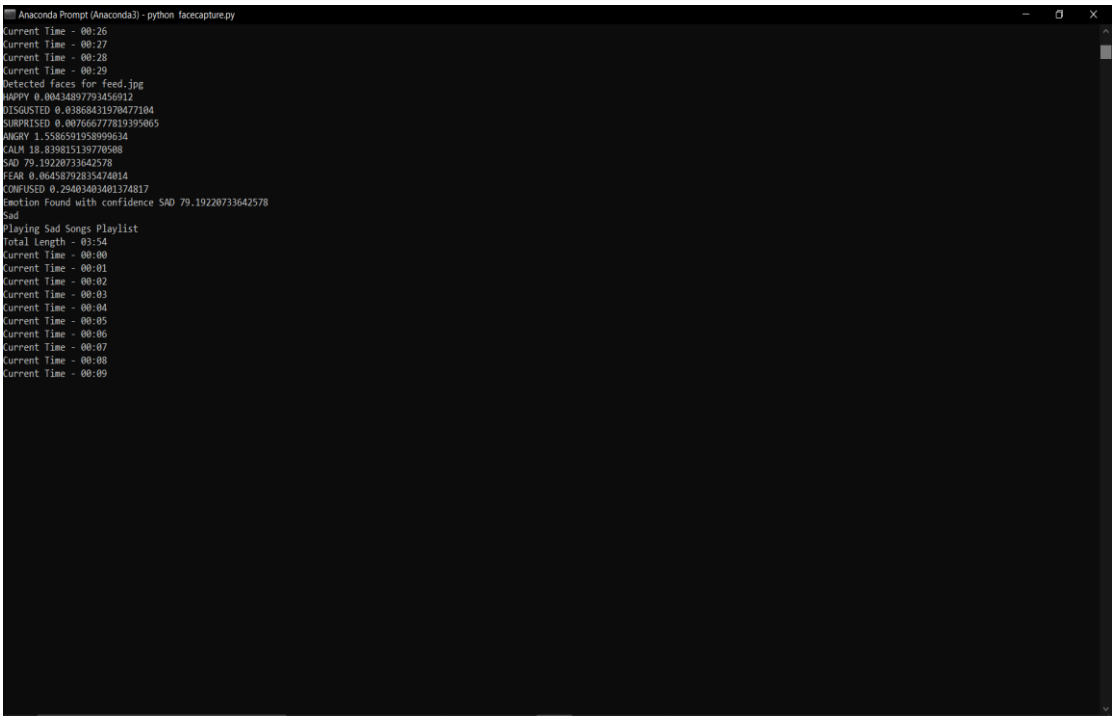

13. Detected angry emotion from the image and plays emotion based song.



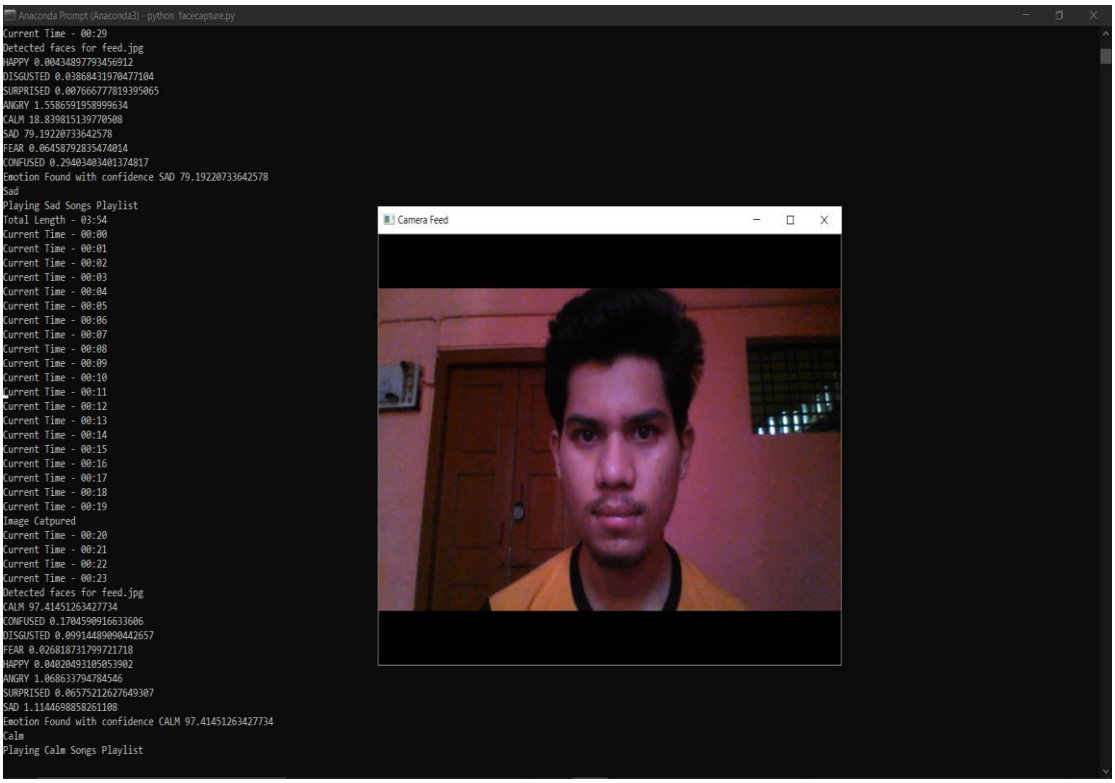
14. Camera opens for live image capturing.



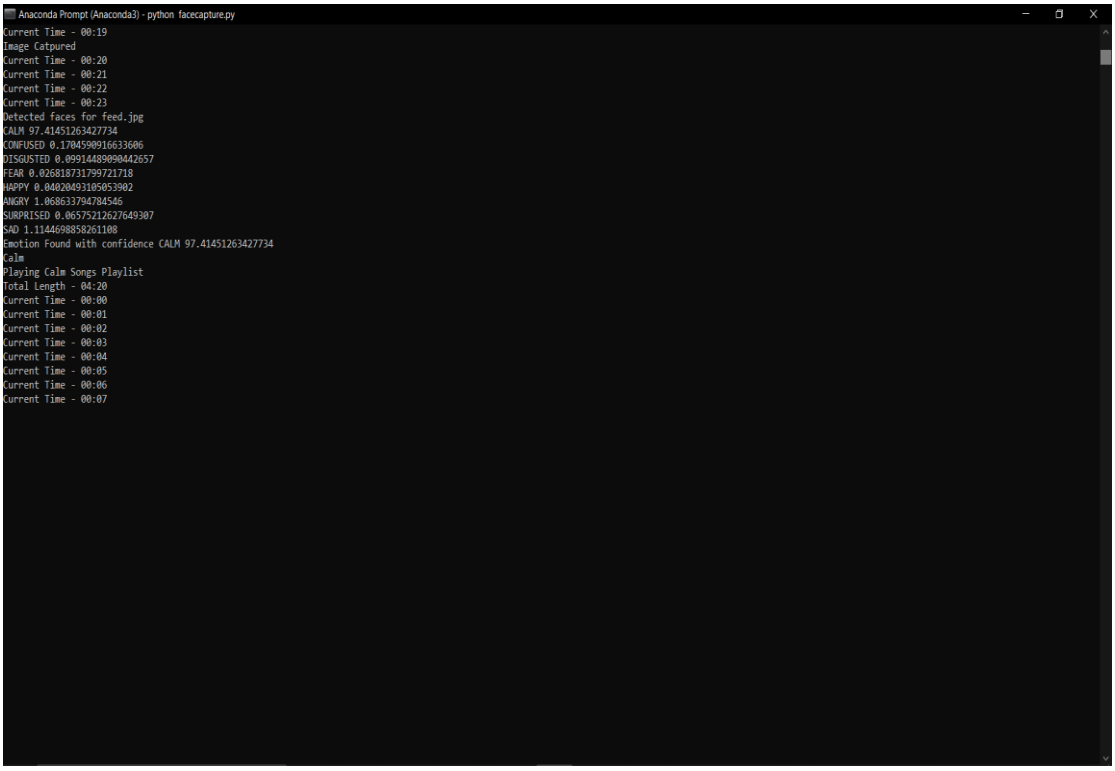

15. Detected sad emotion from the image and plays emotion based song.



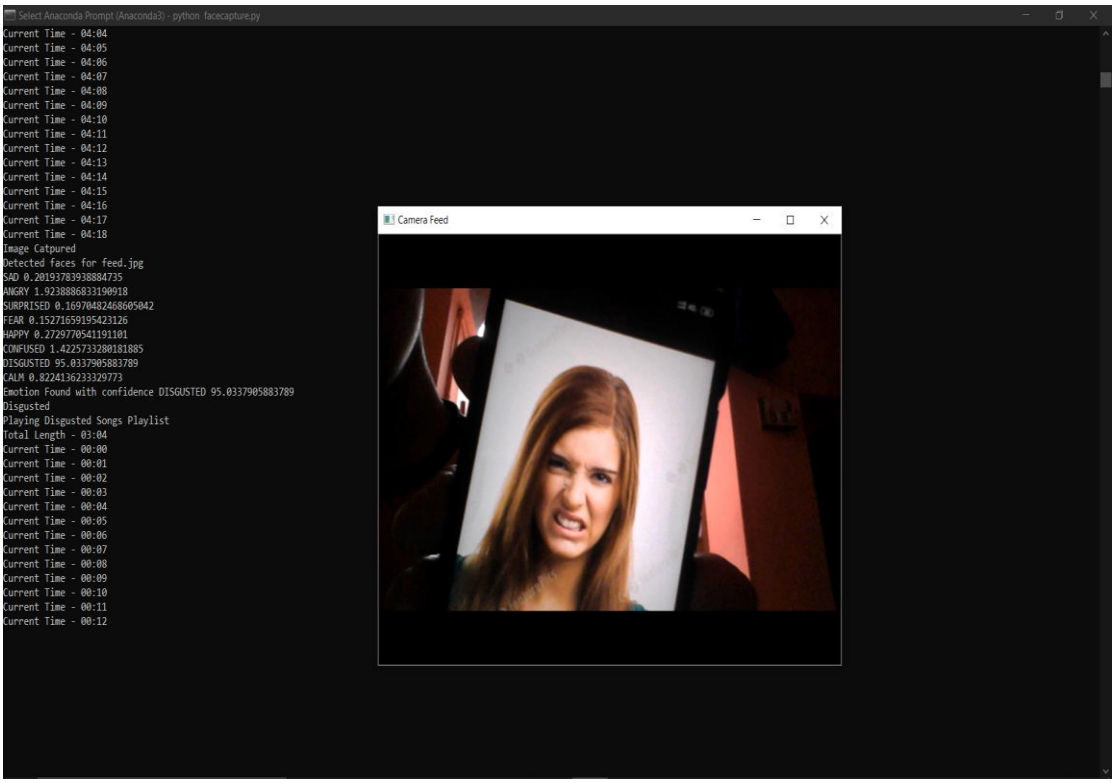
16. Camera opens for live image capturing.



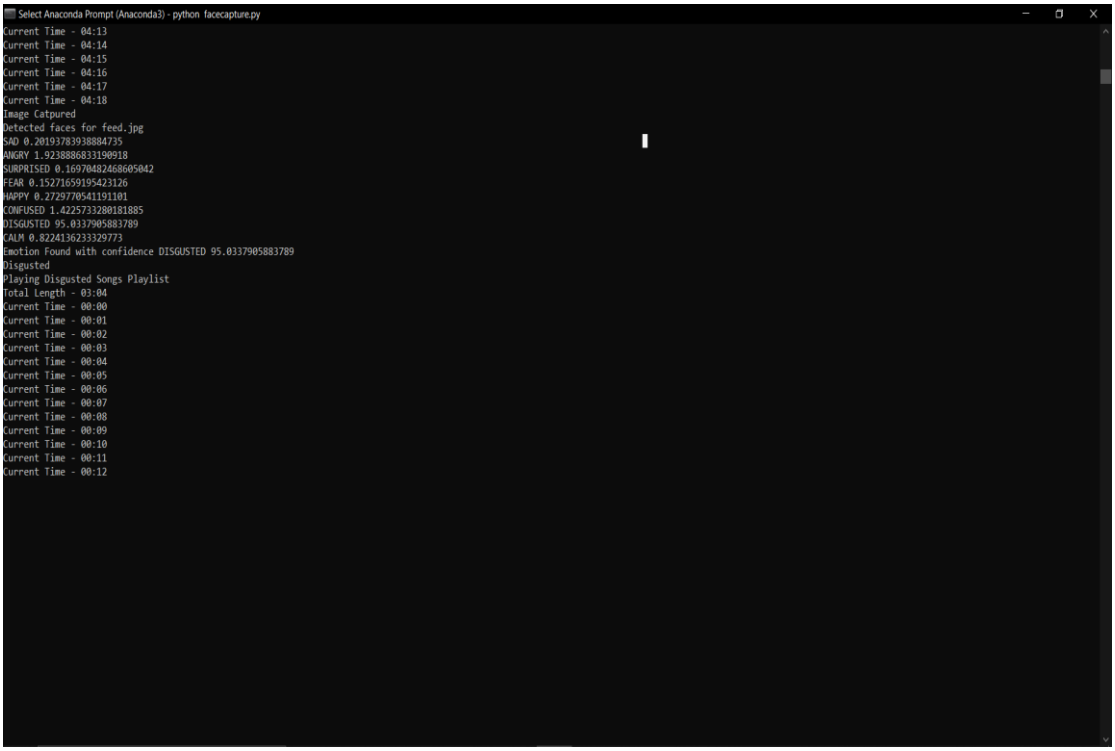

17. Detected calm emotion from the image and plays emotion based song.



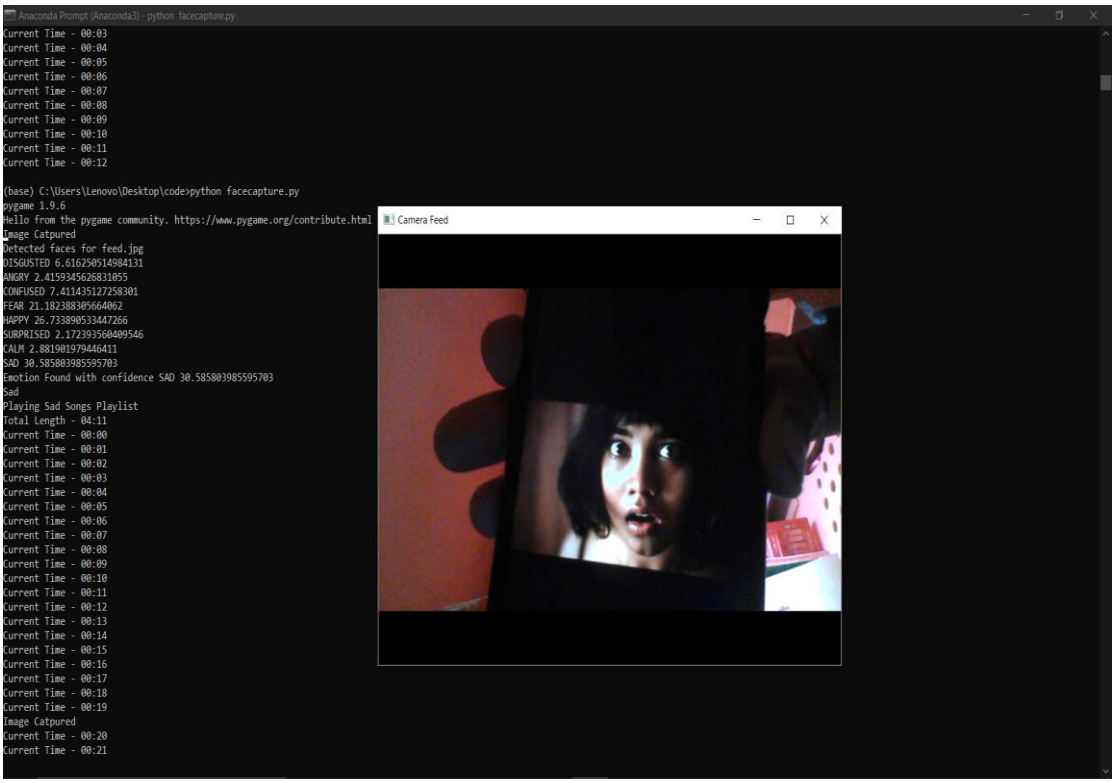
18. Camera opens for live image capturing.



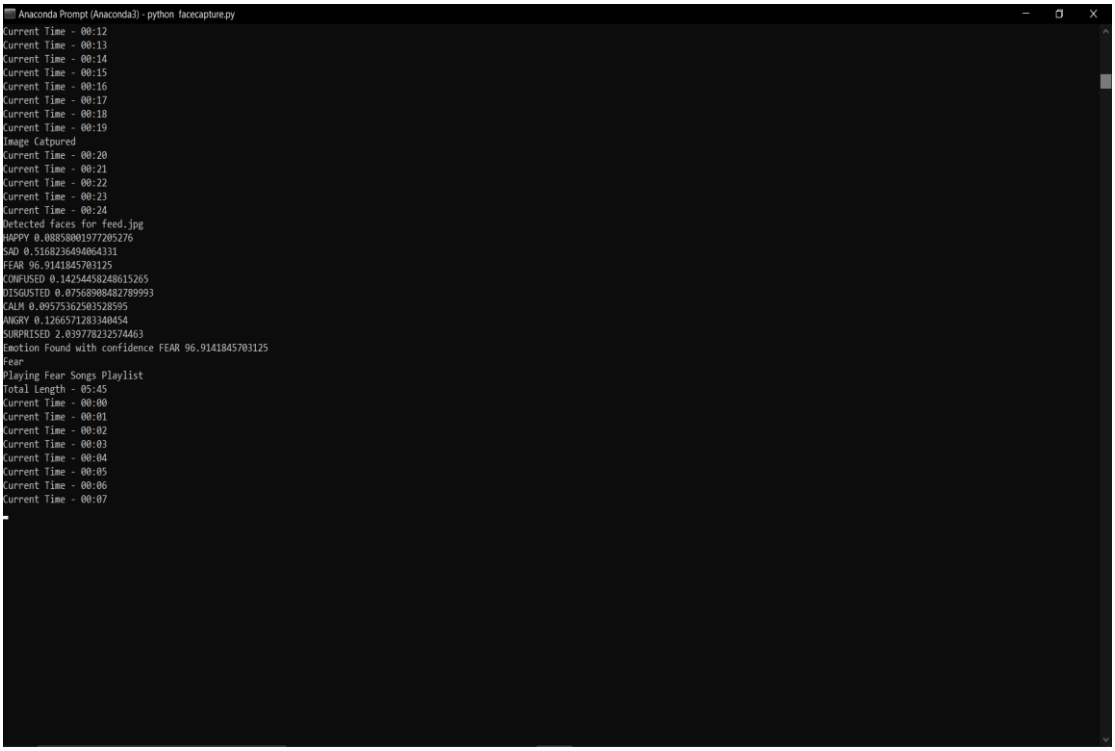

19. Detected disgusted emotion from the image and plays emotion based song.



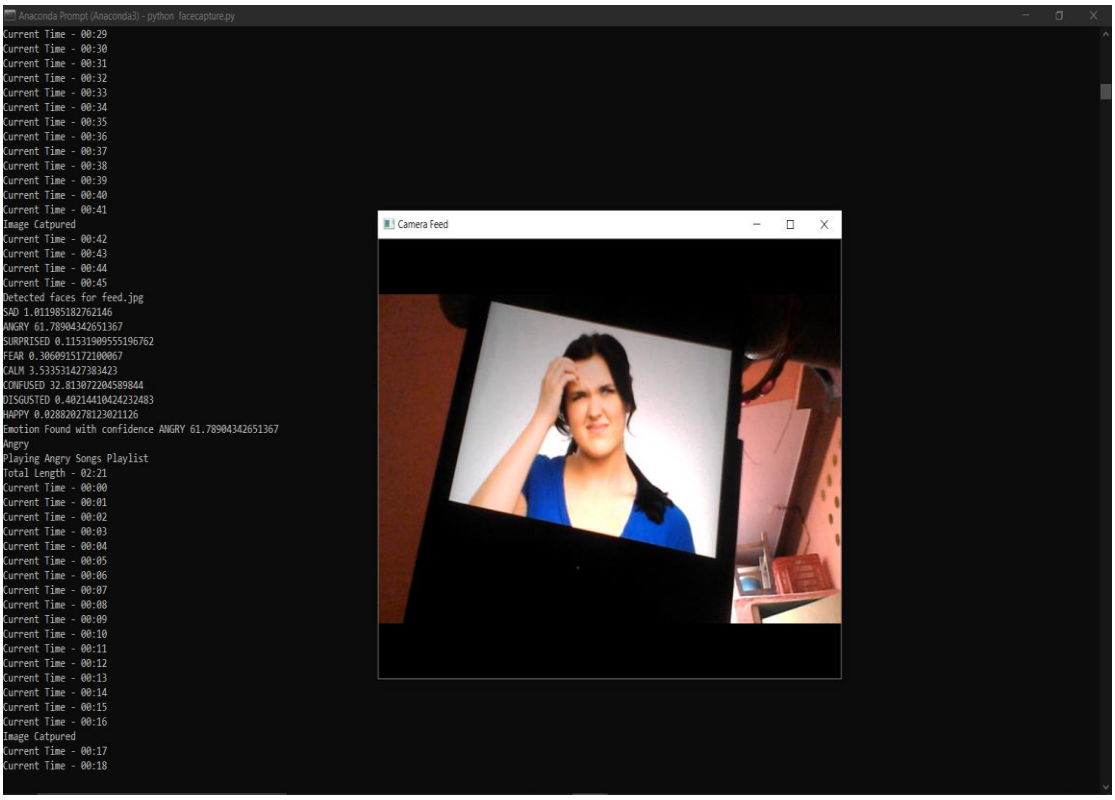
20. Camera opens for live image capturing.



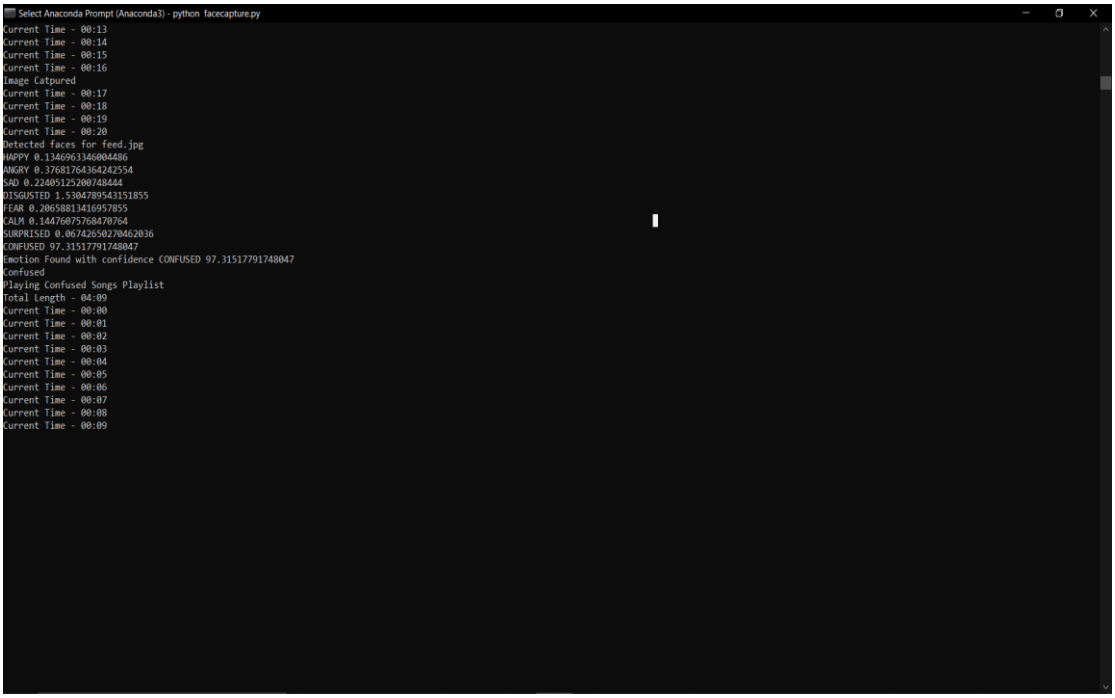

21. Detected fear emotion from the image and plays emotion based song.



22. Camera opens for live image capturing.




23. Detected confused emotion from the image and plays emotion based song.



In the above all sreenshots it is clearly showned the step by step implementation process of image capturing and detecting emotion from the captured image which is stored as feed.jpg in the local storage by using AWS Rekognition service with the help of boto3 library and openCV.

# CHAPTER 8

# ADVANTAGES AND DISADVANTAGES

## 8.1 ADVANTAGES

- Extremely fast feature computation.
- Efficient feature selection
- Scale and location invariant detector
- User friendly
- To reveal from stress
- Based On Facial Expressions
- Easy of use
- Mixed mood detection
- Reduced computational time
- Web Application
- No trouble of troublesome selection of songs
- Can be used in vehicles
- Users don’t want to select song manually.
- No need of playlist.
- Users don’t want to classify the songs based on the emotions.

## 8.2 DISADVANTAGES

- Shuffled songs may not match the mood of the user.
- Time consuming.
- Sensitive to lighting conditions.
- User has to classify the songs into various emotions and then for playing the songs user has to manually select a particular emotion.

# CHAPTER 9

# FUTURE SCOPE

The future scope for the proposed system would be to implement it on mobiles. To design a mechanism that would help in the music therapy treatment for the music therapists to treat the patients suffering from mental stress, acute depression and trauma. It can also be used to determine the mood of a physically challenged person. In the proposed work, only one emotion can be detected at a time so it can be extended to mixed mood detection by continuously recording the face of the user and also this system would be to implement it on Mobile platforms like Android or iOS, as there has been increase in the number of Mobile phone application users worldwide. Also in the work, at a time only one emotion is recognized so system design can be extended to a mixed mood recognition system which will further enhance the functionality and accuracy of the overall system.

This system is applicable in various fields like Medical Science and Psychology as it will be helpful in Music Therapy treatment for music therapists to treat their patients suffering from depression, mental stress, fatigue and trauma. This can be also useful to identify the mood of a physically disabled person as we can understand their needs and help them accordingly.

# CHAPTER 10

# CONCLUSION

Music Player has changed in many different ways since it was first introduced. Now-a-days people like to get more out of different applications, so the designing of applications and the thought process behind it has changed. The users prefer more interactive & sophisticated yet simple to use applications.

The proposed system (Facial Expression based Music Player) presents a music player capable of generating a playlist from the songs‟ audio features and thereby providing the user with an easy way to get the playlist. The proposed model makes use of the Viola-Jones algorithm implemented using java and OpenCV (javacv) to carry out one phase of its functioning and svm classifier is used to carry out the audio feature extraction and classification.

The references used provide us with vital information about the different techniques and strategies followed to carry out their respective individual systems. Based on the knowledge obtained the above content not only provides an in-depth knowledge of the proposed Software development project system but also tries to incorporate the information from various sources in order to execute the tasks using open-source resources. The various aspects of the project have been presented in the above pages in an adequate manner

The proposed work presents facial expression recognition system to play a song according to the expression detected. It uses PCA approach to extract features, and Euclidean distance classifier classifies these expressions. In this work, real images i.e. user dependent images are captured utilizing the in-built camera

# CHAPTER 11

# REFERENCES

1. [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, “Human-computer interaction using emotion recognition from facial expression”, 2011 UKSim 5th European Symposium on Computer Modelling and Simulation.
5. T.-H. Wang and J.-J.J. Lien, “Facial Expression Recognition System Based on Rigid and Non-Rigid Motion Separation and 3D
6. “Generating Music Playlist Based on Facial Expression” US 8094891 B2 Jan 10, 2012.
7. “System Apparatus and Method for Sorting Music Files Based on Moods” US 2012/0226706 A1 Sep 6, 2012.
8. Wilson, P and Fernandez, J. Establishing a face recognition research environment using open source software. ASEE GulfSouthwest Annual Conference, March,2005
9. Bill, David “Personalizing content based on emotions”, US2010/0321519, San Francisco: Dec. 23, 2010.
10. Menezes, P., Barreto, J.C. and Dias, J. Face tracking based on Haar-like features and eigenfaces. 5th IFAC Symposium on
11. Facial feature detection using Haar classifiers, Phillip Ian Wilson, Dr. John Fernandez Open CV reference Manual Hironori Takimoto, Yasue Mitsukura, Minoru Fu kumi and Norio Akamatsu, “F ace Detection and Emotional Extraction
12. https://thecodacus.com/opencv-face-recognitionpython-part1/
13. http://www.paulvangent.com/2016/04/01/emotionrecognition-with-python-opencv-and-a-face-dataset/

14. https://pythonprogramming.net/haar-cascade-face-eyedetection-python-opencv-tutorial
15. Hafeez Kabini, Sharik Khan, Omar Khan, Shabana Tadvi "EMOTION BASED MUSIC PLAYER" International Journal of Engineering Research and General Science, Volume 3, Issue 1, 2015.
16. Nikhil Zaware, Tejas Rajgure, Amey Bhadang, D.D. Sakpal "EMOTION BASED MUSIC PLAYER" International Journal of Innovative Research & Development, Volume 3, Issue 3, 2014.
17. Setiawardhana, Nana Ramadijanti, Peni Rahayu "FACIAL EXPRESSIONS RECOGNITION USING BACKPROPAGATION NEURAL NETWORK FOR MUSIC PLAYLIST ELECTIONS" Jurnal Ilmiah Kursor, Volume 6, Issue 3, 2012.
18. Henal Shah, Tejas Magar, Purav Shah and Kailas Devadkar "AN INTELLIGENT MUSIC PLAYER USING SENTIMENTAL ANALYSIS" International Journal of Innovative and Emerging Research in Engineering, Volume 2, Issue 4, 2015.
19. Anukriti Dureha "AN ACCURATE ALGORITHM FOR GENERATING A MUSIC PLAYLIST BASED ON FACIAL EXPRESSIONS" International Journal of Computer Applications, Volume 100-No.9, 2014.
20. Byeong-jun Han, Seungmin Rho, Roger B. Dannenberg, Eenjun Hwang "SMERS: MUSIC EMOTION RECOGNITION USING SUPPORT VECTOR REGRESSION" 10th International Society for Music Information Retrieval Conference (ISMIR 2009).
21. Benoit Mathieu, Slim Essid, Thomas Fillon, Jacques Prado, Ga¨el Richard "YAAFE, AN EASY TO USE AND EFFICIENT AUDIO FEATURE EXTRACTION SOFTWARE" Institut Telecom, Telecom ParisTech, CNRS/LTCI.
22. Debasmita Chakrabartia and Debtanu Duttab, "Facial Expression Recognition using Eigenspaces," in Int. Conf. on Computational Intell.: Modeling Techniques and Applicat., 2013, pp. 755-761.

23. Müge ÇarÕkçÕ and Figen Özen, “A Face Recognition System based on Eigenfaces method,” INSODE, 2011, pp. 118-123.

24. Kevin C. Tseng, Yu-Te Wang, Bor-Shing Lin and Ping Han Hsieh, “Brain Computer Interface-based Multimedia,” in 8th Int. Conf. on Intelligent Inform. Hiding and Multimedia Signal Process., Piraeus, 2012, pp. 277 – 280.

25. A. habibzad, ninavin, Mir kamalMirnia,‖ A new algorithm to classify face emotions through eye and lip feature by using particle swarm optimization.

26. Byeong-jun Han, Seungmin Rho, Roger B. Dannenberg and Eenjun Hwang, ―SMERS: music emotion recognition using support vector regression‖, 10thISMIR , 2009.

27. Alvin I. Goldmana, b.Chandra and SekharSripadab, ―Simulationist models of face-based emotion recognition‖.

28. Carlos A. Cervantes and Kai-Tai Song , ―Embedded Design of an Emotion-Aware Music Player‖, IEEE International Conference on Systems, Man, and Cybernetics, pp 2528-2533 ,2013.

29. Fatma Guney, ‖Emotion Recognition using Face Images‖, Bogazici University, Istanbul, Turkey 34342.

30. Jia-Jun Wong, Siu-Yeung Cho,‖ Facial emotion recognition by adaptive processing of tree structures‖.

31. K.Hevener,‖The affective Character of the major and minor modes in music‖, The American Journal of Psychology,Vol 47(1) pp 103-118,1935.

32. Samuel Strupp, Norbert Schmitz, and KarstenBerns, ―Visual-Based Emotion Detection for Natural Man-Machine Interaction‖.

33. Michael lyon and Shigeru Akamatsu, ―Coding Facial expression with Gabor wavelets.‖, IEEE conf. on Automatic face and gesture recognition, March 2000

34. Kuan-Chieh Huang, Yau-Hwang Kuo, Mong-Fong Horng ,‖Emotion Recognition by a novel triangular facial feature extraction method‖.

35. S. Dornbush, K. Fisher, K. McKay, A. Prikhodko and Z. Segall ‖Xpod- A Human Activity and Emotion Aware Mobile Music Player‖, UMBC Ebiquity, November 2005.

36. 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.

37. Russell, ―A circumplex model of affect‖, Journal of Personality and Social Psychology Vol-39(6), pp1161–1178 , 1980.

38. Sanghoon Jun, Seungmin Rho, Byeong-jun Han and Eenjun Hwang, ―A fuzzy inference-based music emotion recognition system‖,VIE,2008.

39. Thayer ‖ The biopsychology of mood & arousal‖, Oxford University Press ,1989.

40. Z. Zeng ―A survey of affect recognition methods: Audio, visual, and spontaneous expressions,‖IEEE. Transaction Pattern Analysis, vol 31, January 200