VIT®

Vellore Institute of Technology

(Deemed to be University under section 3 of UGC Act, 1956)

Topic: Mood-Up: An Emotion Detection Technique

Course: TARP (CSE3999)
Faculty Name: GOVINDA K

GROUP MEMBERS:

Devanshi Chadha (17BCE0049)

Toshin Hyanki (17BCE0105)

Samyak Jain (17BCE0139)

Shubhankar Kulkarni (17BCE0784)

Kanishk Balakrishnan (17BCE0992)

Zisan Aalam (17BCE2361)

Yominesh Giri (17BCE2363)

Rinoj Raj Nakarmi (17BCE2364)

Shriyog Sharma (17BCE2373)

Problem Statement

Mood-Up is an emotion detection system that identifies the current mood of the user through facial expression and attempts to modify/alter the mood by giving the user some things to do based on the recognized emotion. The project aims to build a system that would try to give the users emotion-based recommendations that would help them light up their mood for good.

Facial expressions play an important role in recognition of emotions and are used in the process of non-verbal communication, as well as to identify people. They are very important in daily emotional communication, just next to the tone of voice. They are also an indicator of feelings, allowing a man to express an emotional state. People can immediately recognize the emotional state of a person. Facial expression recognition is the part of emotion recognition that is gaining more importance and needs for it increases tremendously.

Motivation

Mental health includes emotional, psychological, and social well-being. It affects how we think, feel, and act. It also help determine how we handle stress, relate to others, and make choices. Mental health is important at every stage of life, from childhood to adolescence and adulthood. Our daily life from eating, sleeping, enjoying, working, social life, everything is affected.

The ongoing global pandemic has indefinitely affected the mental health of people. It has deteriorated the health of people leading to very unstable lifestyle and mood swings, in some serious cases suicide. On a large scale, these issues affect our society, economy, and environment.

The idea behind Mood-Up is to help the user and change his mood for good and this is why the system is made very interactive. We will use the HAAR feature-based cascade classifiers along with OpenCV in Python to recognize faces and evaluate expressions and mood. The face detection is done using the HAAR Cascades. The core of the program is the recommendation system. It is a simple conditional system that would use different emotions as keywords to present users with various options he/she can choose. The system recognizes if the user is sad, angry, or happy, and based on the classification, the recommendation system presents options.

Project Outcome

Mood-Up will be a final product which each one in our society can use anywhere anytime. It is an application which has two main components of functioning, the first involves a face detection and the second component is the mood classifier. These two operations should detect three emotions, namely: Happy, Sadness, and Angry. Such a system should prove beneficial in the current scenario where people need to be helped or urged to perform more activities to ensure overall mental stability and good health.

Mood-Up can provide a small but useful service in today's fast-track world where people don't have time to comprehend as well as analyze their emotions and how to navigate them. In a world where everything has changed, that causes lives to be filled with a hectic schedule, loneliness, helplessness, and boredom, it is the technology itself that allows users to understand and deal with their feelings and emotions.

Project Plan

GANTT CHART

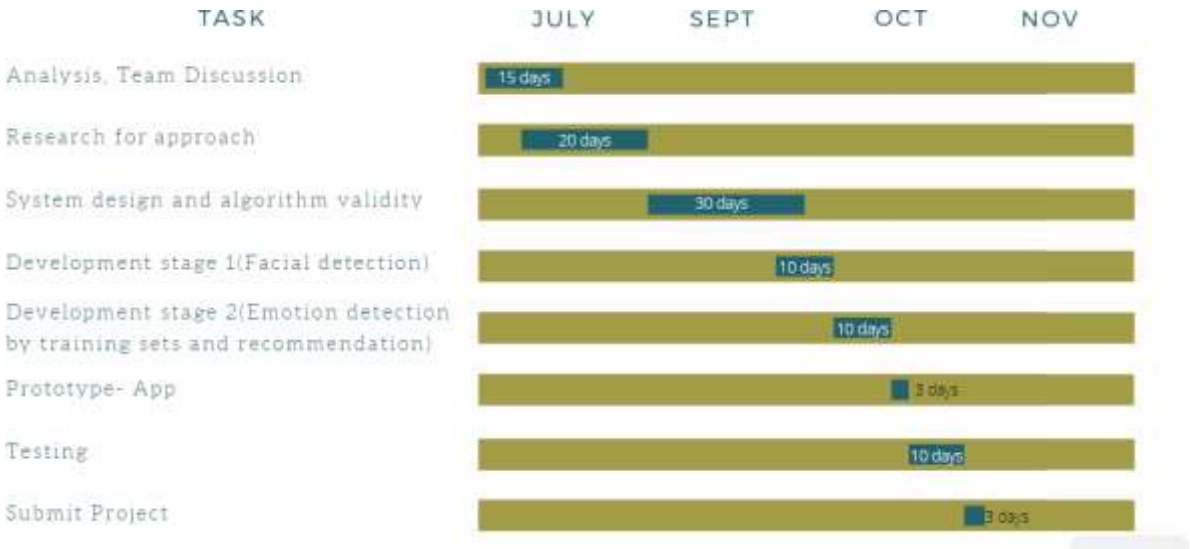

Figure1. Gantt Chart

Overview of project timeline

The analysis and team discussion part will take approximately 15 days. Here we decide the roles and responsibilities which includes designing, prototyping, coding, research etc. After that the research of the algorithm and the technologies required will take 20 days. Then we can start designing the initial part of the system and the algorithm overview. In the developmental stage, the first thing to do is the facial detection part. That will take us approx. 10 days. The next stage is to train our data for the emotion detection and the recommendation. The prototyping and the initial app design will take us 3 days. The testing of the application which includes integrity test, speed test, code test etc. will take us 10 days.

Literature Review

This exploration paper [1] plans to discover in the event that it is conceivable to dependably perceive passionate valence and excitement by utilizing non obtrusive ease sensors (EEG, GSR, and EMG) with the goal that these sensors could gather biometrics. With the assistance of music recordings from the DEAP multimodal dataset, they could effectively land at the end. They could successfully achieve a portrayal execution commensurate to the results nitty gritty in the main assessment that we deficiently copy here, even in a cross-subject course of action setting that forgoes the prerequisite for individual planning and tuning of collection models. Cerebrum related measures, for instance, electroencephalography (EEG) and skin-related measures, for instance, galvanic skin response (GSR) and electromyography (EMG) are among the most notable and extensively grasped physiological measures for impact acknowledgment, in like manner in mix with heartbeat, blood volume, and breath estimations. They found that the fringe sensor setting of the primary assessment furthermore included idea of features got from breath, blood weight and eye glimmering rate in blend in with GSR and EMG. Their models were prepared to consider features isolated from the signs recorded for all of the subjects included. This makes their technique sensible and energetic concerning target application circumstance, patients with debilitated movability and wisdom, for which getting ready and aligning classifiers are not attainable.

In this examination paper [2], they use the TensorFlow library and the Inception exhibition and apply trade learning for a particular dataset to retrain the model. They at that point separate the facial emotions, for example, delight, issue, shock and wonder. The structure can separate only the photos it is set up for essentially like individuals, seeing something they have never watched and thus they won't have the ability to remember it. It will describe the photos in the loads of rate.

They actualized Tensor Board overviews to make it more straightforward to grasp, examine, and update the retraining. They could then think of charts and estimations, for instance, how the loads or precision varied in the midst of planning.

The proposed structure [3] will display chatbot which is just a PC program. The discourse by methods for sound-related or abstract procedures is coordinated by this PC program which is nothing other than chatbot. The bot converses with individuals so it never impacts individuals to understand that it's extremely the PC that is the opposite end user. There will be a customized interface for jokes and tunes as indicated by the customers' perspective. The system will be prepared to perceive pressure and on recognizing the constraint of some inspirational proclamations to show up on the screen. What's more, besides structure, it will prepare to give a couple interfaces with site pages of inspirational talk. The data given by the system will bolster the perspective which impacts the customer to work successfully to and prompts overhaul in execution. The count and progressions which were used in the proposed system will be HAAR course figuring and man-made thinking. The aura will be distinguished on outward appearance premise by picture getting ready using HAAR course figuring.

In this examination paper [4], they have shown the results of affirmation of seven energetic states (impartial, joy, hopelessness, surprise, shock, dread, sicken) considering outward appearances. Coefficients delineating segments of outward appearances, enrolled for six subjects, were used as features. The features have been figured for three-dimensional faces. The gathering of features was performed using k-NN classifier and MLP neural framework. Their result was obtained for MLP classifier and "typical" division of data for all customers (without subject). Tests were finished under comparative conditions and at a settled situation of a customer in association with the Kinect unit. Certainly, the request accuracy was influenced by the way where customers play specific outward appearances, Microsoft Kinect was utilized for 3D stand up to showing transcendently because of its ease and straightforwardness of errand. Kinect has small separating objectives, yet a by and large high pace of picture enrolling. It has an infrared maker and two cameras. One of the cameras record indisputable light, while substitute works in infrared and is used for evaluating the profound. Infrared shafts reflected from the customer's body grant making a 3D model of a face.

This paper [5] delineates a feeling acknowledgment system, which joins mental revelations about feeling depiction with assessment and evaluation of outward appearances. The execution of the proposed structure has been inquired about with preliminary real data. Even more especially, a neurofuzzy run-based structure has been first made and used to mastermind outward appearances using an endless 2D feeling space, getting high rates in portrayal and clustering of data to quadrants of the inclination depiction space. To improve these rates for a specific customer, the fundamental course of action of rules that got them from the previous data was then balanced through a learning technique of the neurofuzzy structure, so as to get unique expressivity cases. Anyway, they plan on extending the framework subject to joined facial and movement assessment. These can give the best approach to make systems that merge examination and association of outward appearances, for giving progressively expressive and welcoming affiliations. Furthermore, progression of administer based framework affirmation gives the probability to solidify the results procured inside the arrangement of the ERMIS adventure with current learning advancements, for example in executing a MPEG-4 visual way of thinking for framework systems.

Overview of the Proposed Work

The main purpose of our project is to detect the emotion a human is feeling based on their picture and coming up with a solution to help make the person feel better. To begin with, emotion of the face is detected with HAAR classifiers. Detection performed in real time of face and interpreting different facial expressions of happiness, anger or aggression, sadness, fear, surprise etc. is based on facial characteristics and their activities. The essential features of face are studied for detection of face and expectation of expressions or emotions of face. To determine the various facial expressions, the alterations in each facial characteristic are used. For detection and classification of distinct classes of facial expressions, machine learning algorithms are used by training of distinct sets of pictures. The proposed algorithm employs open source computer vision (OpenCV) and Machine learning using python.

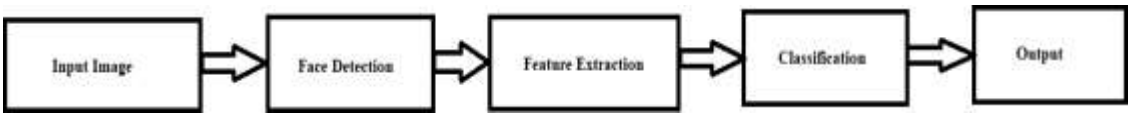

Figure2. Overview of model

We require a large database of different faces depicting different emotions. We are using our own database with thousands of images and organized them in a folder named dataset within there are 3 different subfolders (that are going to be your several classes) named anger, happy, and sad which contains the images of a particular expression. To perform this task, we must find the face on each image, convert it to gray-scale, cut and then save the picture to the dataset. We use a HAAR filter from OpenCV to automate face finding. A HAAR Cascade is fundamentally a classifier which is used to detect the object for which it has been trained from, from the source. To obtain better results one must use high-level value images and increase the amount of stages for which the classifier is killed. OpenCV provides a training method or pre-trained models. After the emotion has been detected using the above-mentioned method, we try to talk to the subject. Based on the inference made on the emotion felt, we pose a few things to do for the subject to find out the keywords in the responses from the subject that are

mapped in the predefined list that we have created. The list helps us provide accurate solutions to the subject's emotion. Suppose a keyword "angry" or "failure" or "rejection" is come across, the system will understand it and provide options for motivational quotes or songs that will help lift the mood of the person. At the moment we are focusing on emotion detection based on HAAR classifiers and we are narrowing our focus down to the 3 main emotions namely "Happiness", "Sadness" and "Anger".

References

[1] Daniela Girardi, Filippo Lanubile, Nicole Novielli. (2017).Emotion Detection Using Noninvasive LowCost Sensors.University of Bari „Aldo Moro"d.

[2] Ajay B, Anirudh C, Karthik Joshi, Keshava B, Mrs. Asha.(2017).Emotion detection using machine learning. Department of Information Science and Engineering, NIE Institute of Technology, Mysore, Karnataka, India.

[3] Prof. D.S.Thosar, Varsha Gothe, Priyanka Bhorkade, Varsha Sanap.(2018). Review on Mood Detection using Image Processing and Chatbot using Artificial Intelligence. PRES"s SVIT, Nasik, Maharashtra, India.

[4] Paweł Tarnowski, Marcin Kołodziej, Andrzej Majkowski, Remigiusz J. Rak.(2017). Emotion recognition using facial expressions.Warsaw University of Technology, Warsaw, Poland.

[5] Spiros V. Ioannou, Amaryllis T. Raouzaiou, Vasilis A. Tzouvaras, Theofilos P. Mailis, Kostas C. Karpouzis, Stefanos D. Kollias.(2005). Emotion recognition through facial expression analysis based on a neurofuzzy network. Image, Video and Multimedia Systems Laboratory, National Technical University of Athens, Zografou 15773, Greece, School of Electrical and Computer Engineering.

Proposed Methodology

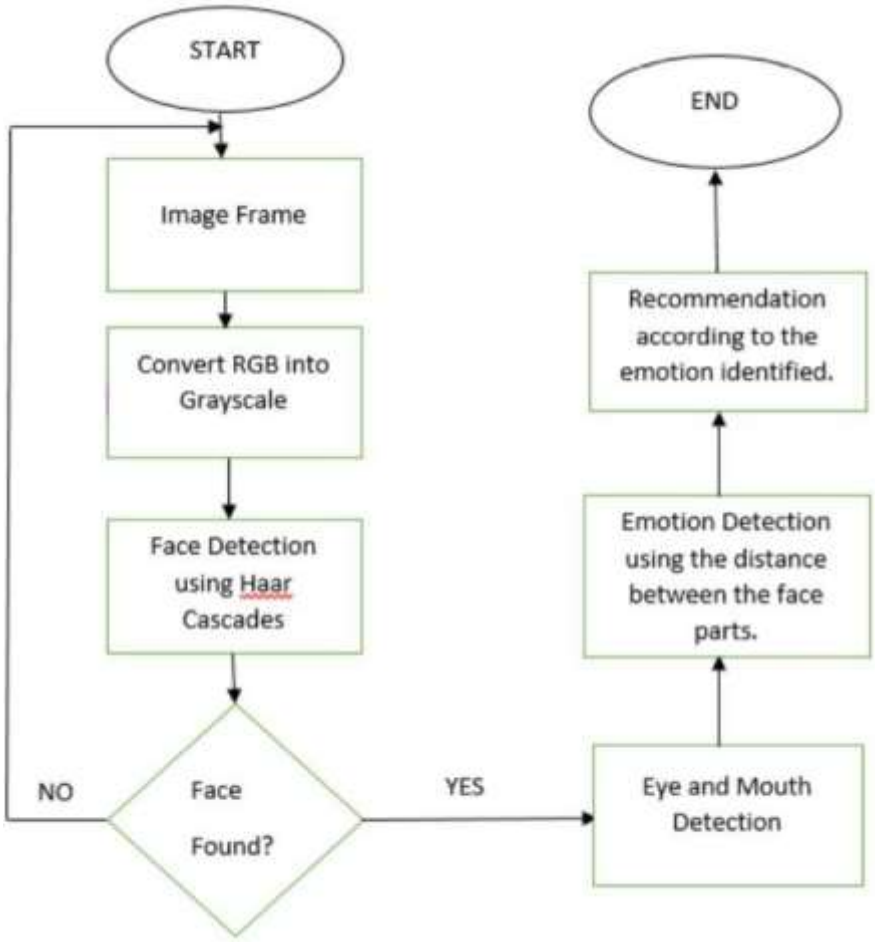

Figure 3. Flowchart of system
Image is in depth step by step procedure of the entire system

- **Face Detection** – Using HAAR Classifier, We will detect the face of the client. The image frame will get converted into Grayscale using the luminosity algorithm. For the same we will be using OpenCV and laptop's Webcam. OpenCV, it is a cross-platform library which can be used to develop real-time computer vision applications. The focus lies predominantly on image processing, video capture and analysis including features like face detection and object detection. In this project we will be using it in correspondence with the HAAR classifier.
- **Mood detection** – Detection of emotions can be obtained by extracting the features of the face. We take into the consideration histograms in order to compute the distance between the two eyebrows, eye and eyebrow, eye (left side) and nose, eye (right side) and nose

and then identify accordingly. This will be done using our trained models. Currently in this project, we are focusing on three emotions that are happy, sad and angry. In the future, we wish to use different permutations and combinations to incorporate more emotions into our project.

- **Recommendation System** - After the detection of emotion, we will be having our own recommendation based system that will be loading some questions regarding each of the emotion in the database, and our system will ask the user for the reason of the current emotion and based on the answers, we will use a data dictionary which will have keywords like failures, success, marks, exams, etc. and according to the keywords our system will try to enhance the mood of the user by performing different actions like playing songs, recommending motivational videos etc.

SAMPLE CODE

Emotions.py

emotions.py

```
1  import numpy as np
2  import argparse
3  import matplotlib.pyplot as plt
4  import cv2
5  from tensorflow.keras.models import Sequential
6  from tensorflow.keras.layers import Dense, Dropout, Flatten
7  from tensorflow.keras.layers import Conv2D
8  from tensorflow.keras.optimizers import Adam
9  from tensorflow.keras.layers import MaxPooling2D
10 from tensorflow.keras.preprocessing.image import ImageDataGenerator
11 import os
12 os.environ['TF_CPP_MIN_LOG_LEVEL'] = '2'
13
```

Figure 4. Libraries used for training and detection of emotion

```
# Create the model
model = Sequential()

model.add(Conv2D(32, kernel_size=(3, 3), activation='relu', input_shape=(48,48,1)))
model.add(Conv2D(64, kernel_size=(3, 3), activation='relu'))
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Dropout(0.25))

model.add(Conv2D(128, kernel_size=(3, 3), activation='relu'))
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Conv2D(128, kernel_size=(3, 3), activation='relu'))
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Dropout(0.25))

model.add(Flatten())
model.add(Dense(1024, activation='relu'))
model.add(Dropout(0.5))
model.add(Dense(7, activation='softmax'))

# If you want to train the same model or try other models, go for this
if mode == "train":
    model.compile(loss='categorical_crossentropy',optimizer=Adam(lr=0.0001, decay=1e-6),metrics=['accuracy'])
    model_info = model.fit_generator(
            train_generator,
            steps_per_epoch=num_train // batch_size,
            epochs=num_epoch,
            validation_data=validation_generator,
            validation_steps=num_val // batch_size)
    plot_model_history(model_info)
    model.save_weights('model.h5')
```

Figure 5. Training model

Model is trained with all the images taken in the dataset. All these images are taken in greyscale to maximize the accuracy of emotion detection

```
# emotions will be displayed on your face from the webcam feed
elif mode == "display":
    model.load_weights('model.h5')

    # prevents openCL usage and unnecessary logging messages
    cv2.ocl.setUseOpenCL(False)

    # dictionary which assigns each label an emotion (alphabetical order)
    emotion_dict = {0: "Angry", 1: "Disgusted", 2: "Fearful", 3: "Happy", 4: "Neutral", 5: "Sad", 6: "Surprised"}

    # start the webcam feed
    cap = cv2.VideoCapture(0)
    while True:
        # Find haar cascade to draw bounding box around face
        ret, frame = cap.read()
        if not ret:
            break
        facecasc = cv2.CascadeClassifier('haarcascade_frontalface_default.xml')
        gray = cv2.cvtColor(frame, cv2.COLOR_BGR2GRAY)
        faces = facecasc.detectMultiScale(gray,scaleFactor=1.3, minNeighbors=5)

        for (x, y, w, h) in faces:
            cv2.rectangle(frame, (x, y-50), (x+w, y+h+10), (255, 0, 0), 2)
            roi_gray = gray[y:y + h, x:x + w]
            cropped_img = np.expand_dims(np.expand_dims(cv2.resize(roi_gray, (48, 48)), -1), 0)
            prediction = model.predict(cropped_img)
            maxindex = int(np.argmax(prediction))
            cv2.putText(frame, emotion_dict[maxindex], (x+20, y-60), cv2.FONT_HERSHEY_SIMPLEX, 1, (255, 255, 255), 2, cv2.LINE_AA)

        cv2.imshow('Video', cv2.resize(frame,(1600,960),interpolation = cv2.INTER_CUBIC))
        if cv2.waitKey(1) & 0xFF == ord('q'):
            break

    cap.release()
    cv2.destroyAllWindows()
```

Figure 6. Detection of emotion

The system uses the model which is trained.It opens the webcam to detect the emotion of the user.

Figure 7. Recommendation system

After detection of emotion a task is recommended to the user with respect to the emotion. All recommendations are kept in an array with reference to each emotion that is detected.

OUTPUT SCREENSHOT

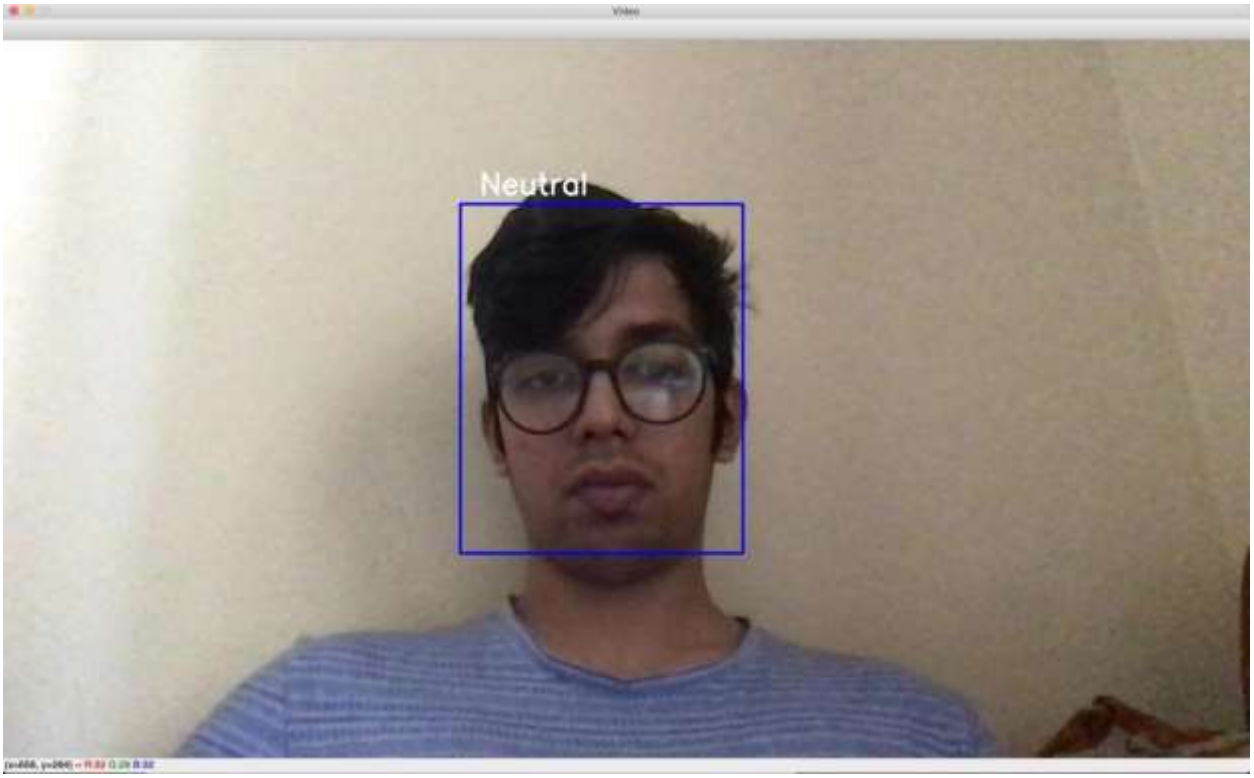

Figure 8. Emotion Detected
Neutral face has been detected by the system

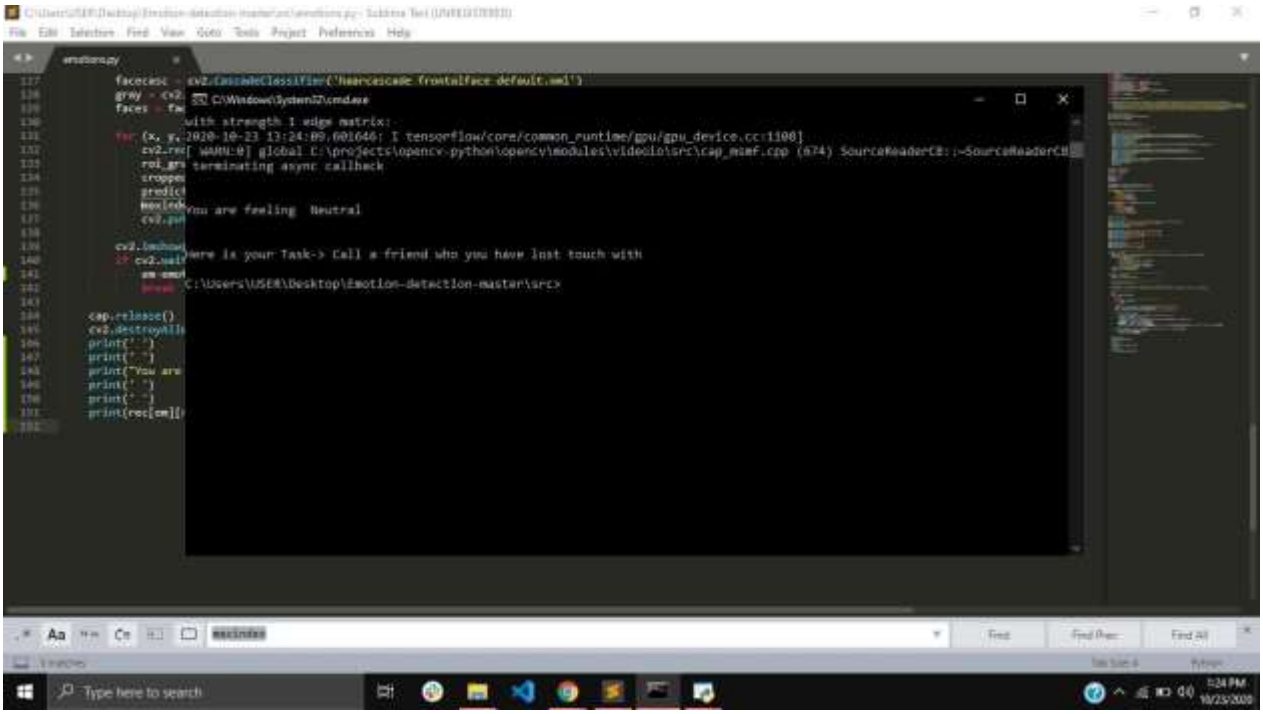

Figure 9. Task recommended
Neutral Face Detected. Task recommended here is to call an old friend whom you have lost touch with.This will make user happy

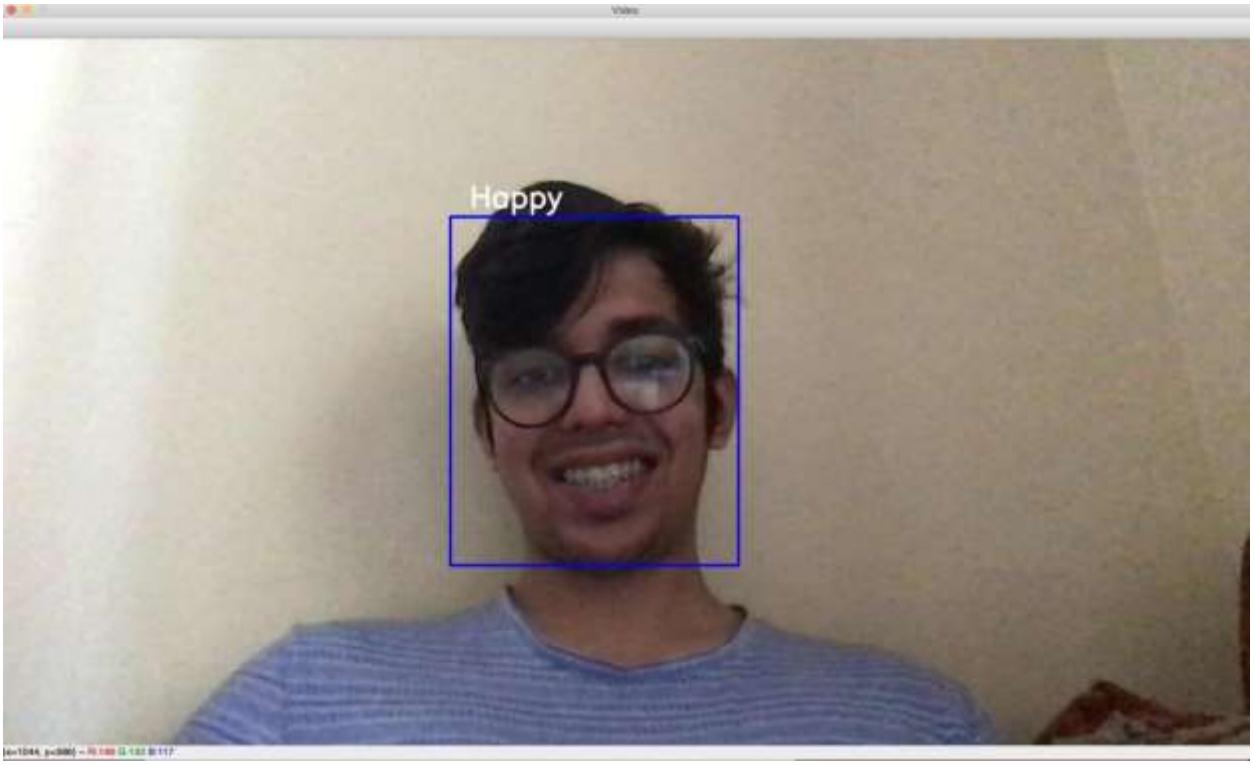

Figure 10. Emotion Detected
Happy face has been detected by the system

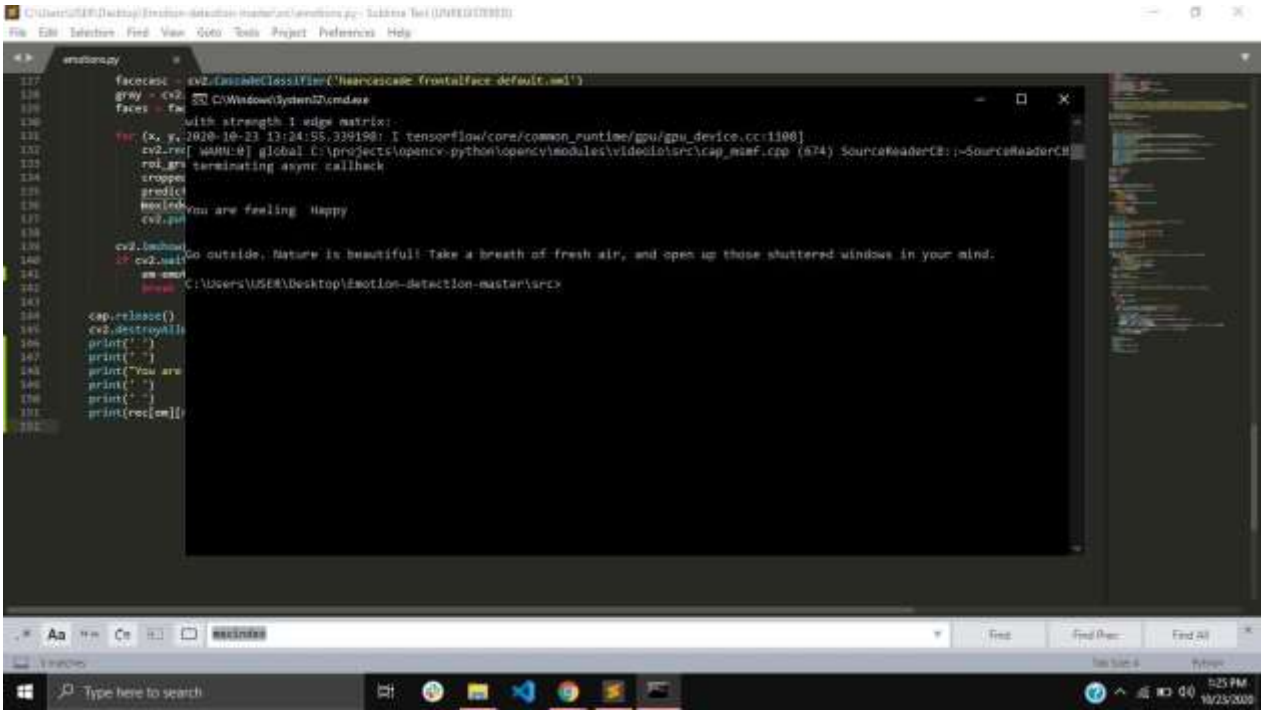

Figure 11. Task recommended
Happy Face Detected. Task recommended here is to go outside,
be glad of what life has provided you with.

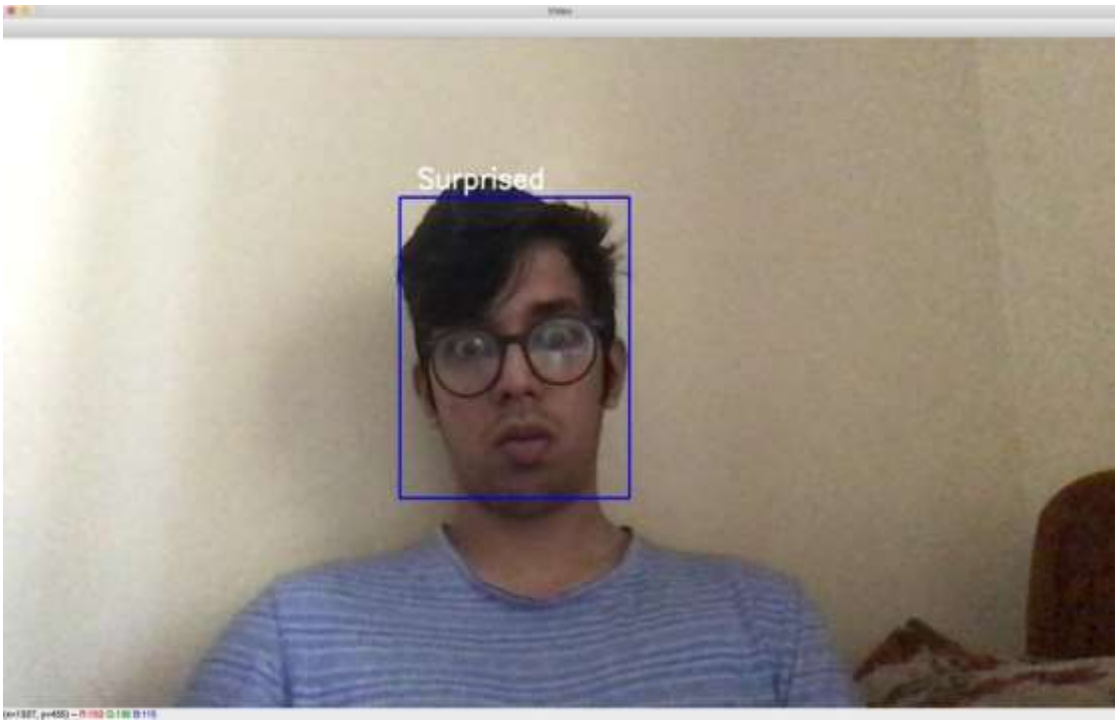

Figure 12. Emotion Detected
Surprised face has been detected by the system

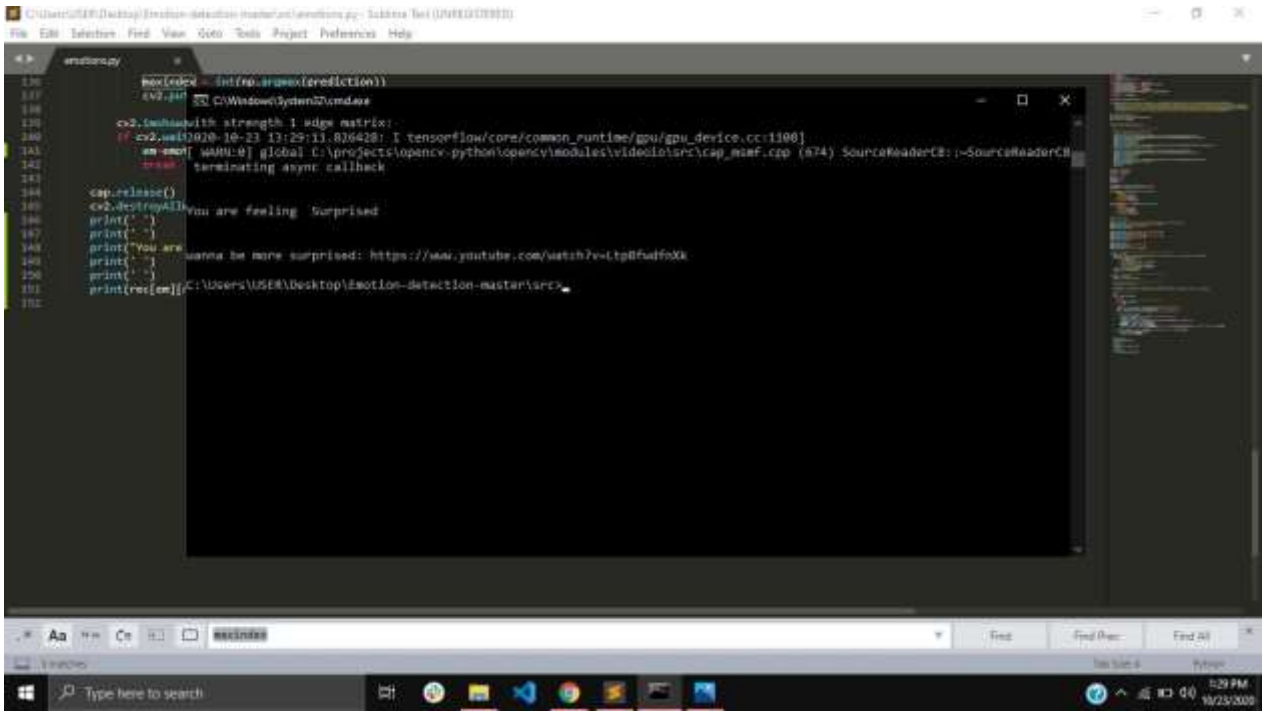

Figure 13. Task recommended
Surprised Face Detected. Task recommended here is a youtube video.

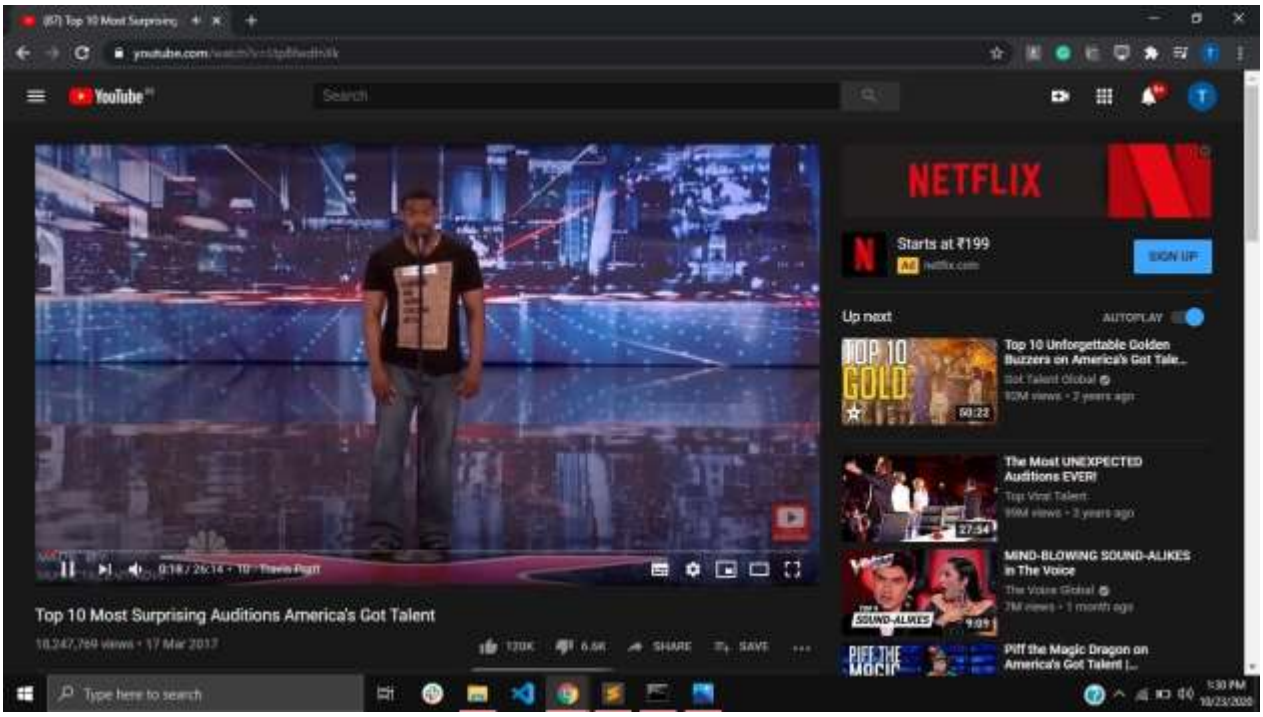

Figure 14. Example of recommendation
An astonishing video of America's got talent has been recommended

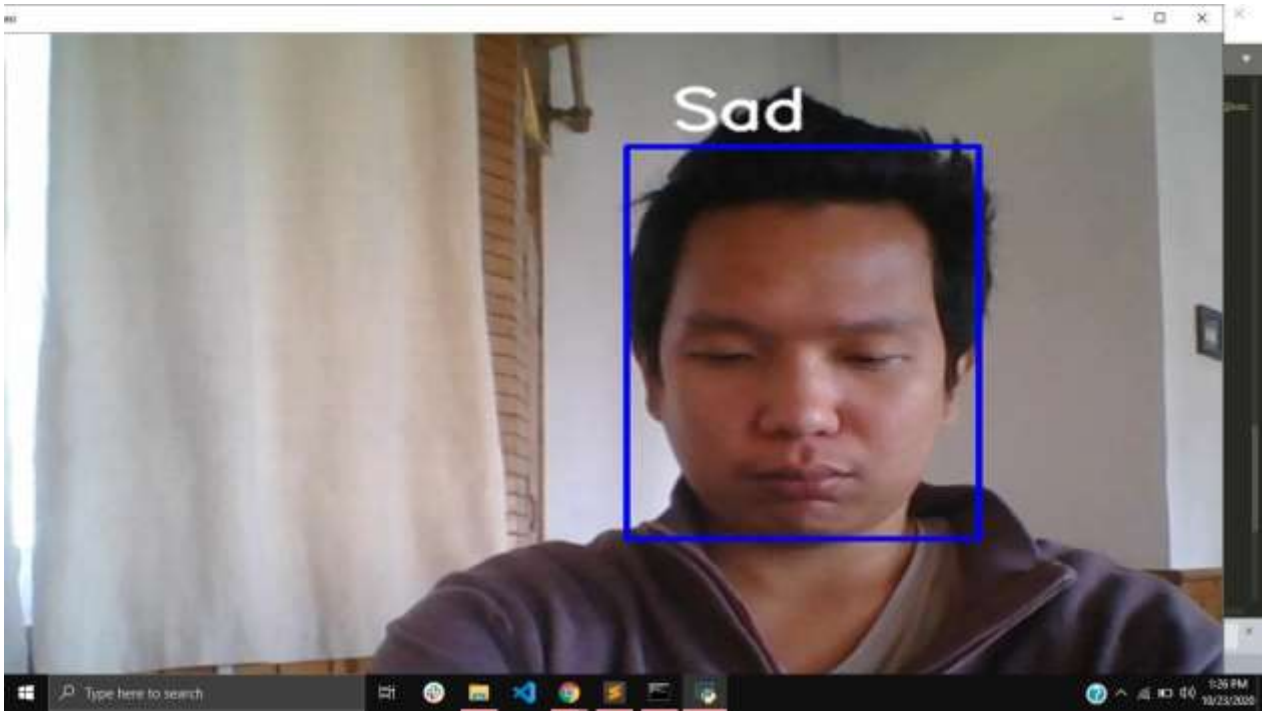

Figure 15. Emotion Detected
Sad face has been detected by the system

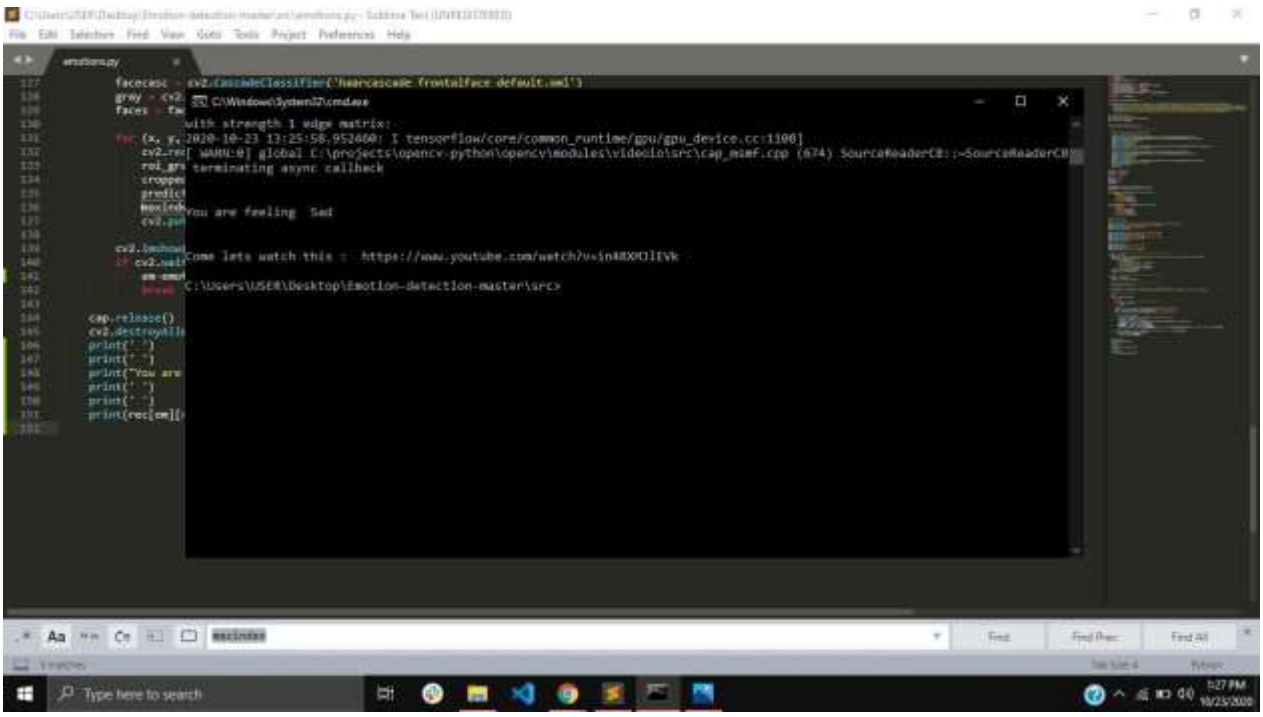

Figure 16. Task recommended

Sad Face Detected. Task recommended here is a youtube video.

Figure 17. Example of recommendation

A funny video of Zakhir khan, who is a famous standup comedian, is recommended.

Figure 18. Emotion Detected
Angry face has been detected by the system

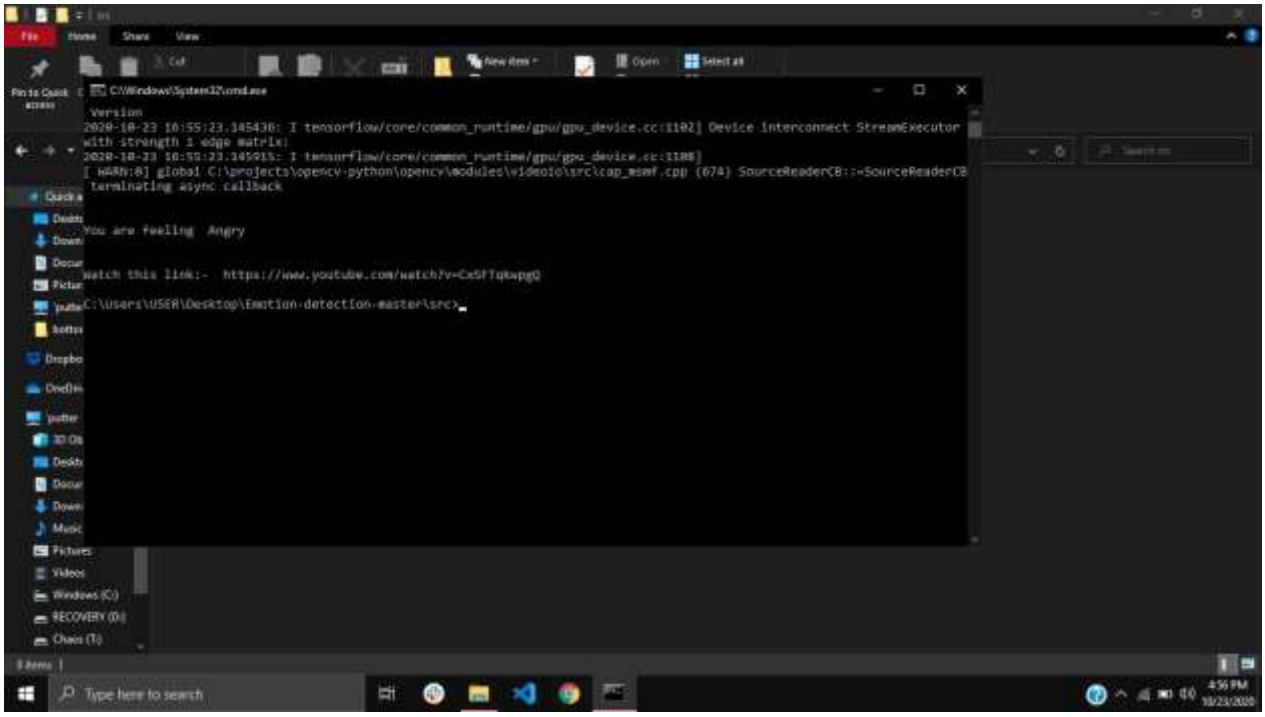

Figure 19. Task recommended
Angry Face Detected. Task recommended here is a youtube video.

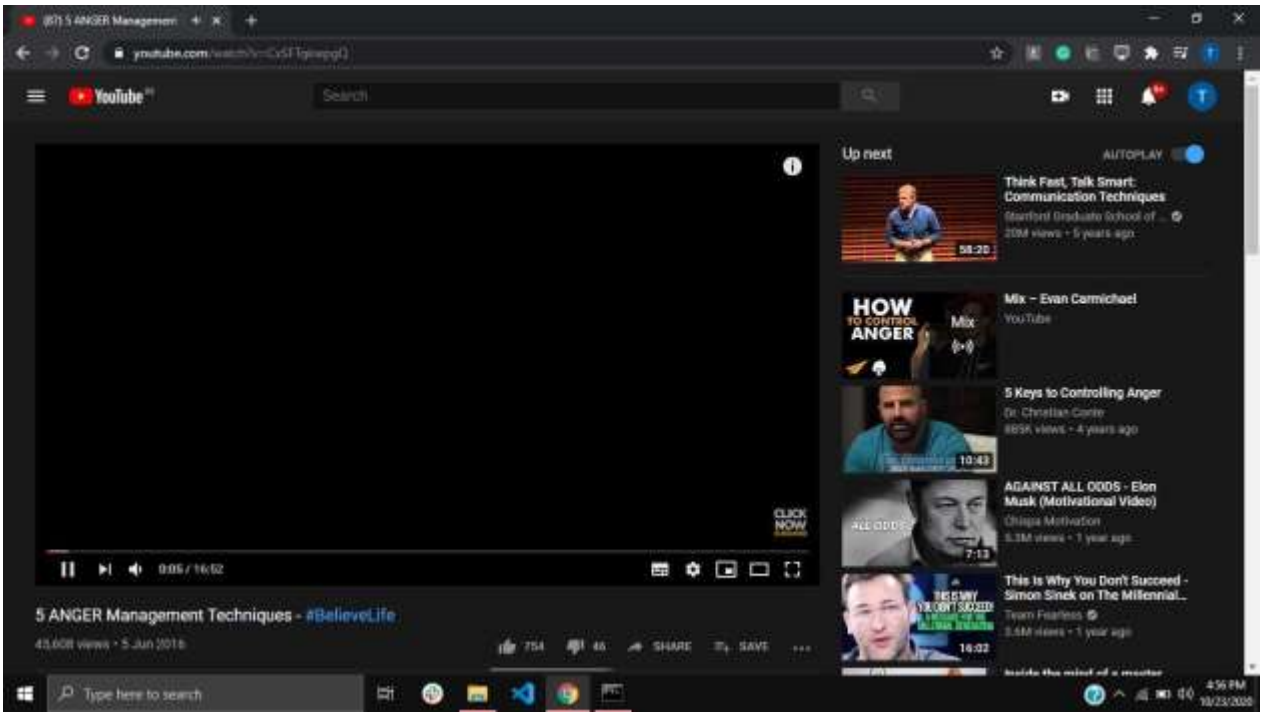

Figure 20. Example of recommendation
An anger management video has been recommended by the system.