Emotion Recognition using Convolutional Neural Network

Praveen.R, Benjula Anbu Malar M.B

Abstract*: Emotion recognition system place the important role in many fields, particularly image processing, medical science, machine learning. As per human needs, the effect and potential use of programmed emotion recognition have been developing in a wide scope of utilizations, including human-PC communication, robot control and driver state observation. In any case, to date, vigorous acknowledgment of outward appearances from pictures and recordings is yet a testing errand because of the trouble in precisely extricating the helpful passionate highlights. These highlights are regularly spoken to in various structures, for example, static, dynamic, point-based geometric or area based appearance. Facial development highlights, which incorporate component position and shape changes, are by and large brought about by the developments of facial components and muscles on the face of enthusiastic manner. Emotion recognition system has many applications. and it plays a vital part in fault detection and in gaming application. In this project the emotion recognition is of dynamic way and not like uploading the image and finding the emotion. And this is achieved with the help of the concept of machine learning called Convolutional Neural Network. This is one of the most familiar deep learning concept. The main moto of using this concept is to maintain accuracy. The CNN consists of many intermediate state which plays the important role in producing the accurate output. The layers of CNN are input layer, hidden layer and output layer. The hidden layer is used to update weight, bias and activation function. If we use the CNN methodology the unwanted parts which is un necessary for the emotion recognition will be eliminated accurately. The CNN helps to reduce our elimination task in easier way and with minimal steps.

Keywords*: Activation Function, Bias, Conv2D, Convolutional Neural Network, Max pooling, Normalization, Weights. HR classifier.

ACRONYMS

CNN-Convolutional Neural Network
CL-Convolutional Layer
FER-Face Emotion Recognition
HR-Haar Cascade

I. INTRODUCTION

The facial components, particularly the key components will always show sign of change their position when the feelings of the person in changed. As a result, a similar element in various pictures all most part of distinctive positions. In case, the place of element may be twisted or bent out of the normal shape because of facial muscle developments.

For example there are three picture, the mouth position of the first two picture presents distinctive shapes from that in the third picture. According to the above scenario the Specific feeling, the geometric - based position and appearance-based shape typically changes starting with one picture than the next picture in the database just as in recordings.

The most recent years, the recognition system is used in the field of dynamic research with applications in a few unique field. For example human and machine association, neural science, PC illustrations. Transport security by identifying the driver weakness. Emotion recognition system has played a vital role in machine interface which helps to make communication between machine and human in efficient and easier way. Some application uses the face and thumb for the individual recognizable proof and access control. However, the execution of the face location positively influences the execution of the considerable number of uses .

Various strategies were proposed for identifying human face in pictures. Thy can be ordered into four classification. They are information based techniques, include based strategies, template based techniques and appearances-based methods. When we use individually, these strategies were not able to take care of considerable number of issues like position, appearance and impediments. Simultaneously it is smarter to work with a few progression or parallel strategies. The significant one of the outwards appearance strategies are mainly focused on five essential classes, for example angry, happy, neutral, shocked, sad.

A CNN is one of the machine learning or deep learning algorithm. In this algorithm the input file is the image and assigning some important weights and biases to the different aspect in the particular image its able to distinguish one value from another value. The needs of Pre—processing technique is really very less when compared to other image classification methods and techniques. The motivation of the domain is empower machine to see the world as people do. Likewise we use number of data for number of task. For example, image and video acknowledgment, Image Processing and Matching Pattern, Finger print matching and so on.

A. Objectives

The main objective is to prepare a solution for the outward appearance acknowledgment difficulty by separating it as sub problems and dividing those sub problems into sub groups of some particular action units.

Revised Manuscript Received on February 26, 2020.

R.Praveen, pursuing Master Of Computer Application, Vellore Instiute Of Technology,Vellore.

Prof Benjula Anbu Malar.M.B, Assistant professor, Vellore Instiute Of Technology, Vellore

This method not only focused on two class issues, which tells about the action unit is on or off, yet furthermore multi-class issues that illuminate the client about the multi action or more than one action unit in the meantime. For this aspect, we can use distinctive philosophies and methods for highlighting the extraction, standardization, determination and arrangements. Solution for these issues simply as taking the computational multifaced nature and timing problem into the idea. The project objective is to actualize face acknowledgment in a really perfect route concerning that is of run time implementation framework. Various calculation and strategies are considered for accomplishing this objective. This type of face recognition framework can be broadly utilized in our day by day life in various segments. We trust that human life can be extraordinarily assisted with this innovation.

B. Motivation

What is CNN?

A CNN, Convolutional Neural Network is one kind of neural network concept which is mainly used for analysing the image frame by frame purpose and convolutional neural network has convolutional layer, which has activation-function, fully connected layers and pooling, receptive fields and weights. By using that it performs the operation based on that functions.

Convolutional layer:

Convolutional layers are mainly used for taking the previous information from one layer to another layer and passing the outcome to the following layer. The convolutional imitated the reaction of separate neuron to see the improvements. Each convolutional neurons forms information just for its respective field. But related feed forward neural network can be used to learn just as characterized information, it isn't useful to apply this concept to pictures. An exceptionally crucial profound design that the reason of the large info sizes related with picture, where the every pixels are an important variable. For example, a complete associative layer of a(little) picture with size 10000*10000 has 100000 loads for every neuron in the second layer.

Pooling:

Convolutional neural network may be integrated with neighbourhood or globally pooling layers, clarification needed which joins the yields of a group of neuron from layer one into a unique neuron in the following layer. For instance, max pooling utilizes the most extreme incentive from every group of neurons at the earlier layer. Another type is normal pooling, which utilize the normal incentives from the each one in a group of neurons at the earlier layer.

Fully connected:

Fully associated layer interfaces with each layer of a neuron to another every layer. It is one of the basic level which is equivalent to the conventional multi-layer perceptron neural network.

Receptive field:

In neural system, every neuron gets contributed with the areas in the last layer. Fully connected layers, the every individual neuron gets contribution from each component of the previous layer. In a Convolutional layer, neurons get contribution from just a limited sub area of the previous layer. The responsive field is the previous entire layer. The CNN layer is the responsive region which is smaller than the entire previous layer.

Weights:

Every neuron in a neural network system yields an incentive by applying some capacity to the information which is originated from the open field in the previous layer. Learning in a neural system advances by making gradual changes would be the inclinations and loads. The vector of loads and the predisposition are known as a channel and speaks to some component of the info (e.g., a specific shape).

C. Problem Statement

Behaviour of sentiments through facial feelings was an object of interest since the season of Aristotle. This point became simply after 1955, when a rundown of general feelings was set up and a few parametrized frameworks were proposed. Encouraged by Deep Learning and Computer Vision, building mechanized acknowledgment frameworks has gotten a ton of consideration inside the Computer Science region. To understanding correspondence Mehran, has deduced in his investigation that in eye to eye correspondence, feelings are transmitted in extent of 55% through outward appearances. It implies that, if the PC could
catch and compare the feelings of the user, correspondence would be progressively normal and proper, particularly in the event that we consider situations where a PC would assume the job of a human

D. Proposed Solution:

As a solution to the above problem statement, enforce the concept of Convolutional Neural Network. However, such a lot of Neural Network algorithms uses back propagation due to some unique features like pooling we go for Convolutional Neural Network. The real statistics photo that is tested for highlights. This proposed system ignores the background distractions and produce accurate output. The activation function set up the highest point for each channel you use. The biggest square shape is one fix to be down sampled. The activation function consolidated by means of down sampling. Another gathering of initiation maps created by disregarding the channels the stack that is down sampled first. The second down sampling which gathers the second gathering of initiation maps. This system will displays the emotion name with their percentage level. For example if the user is happy, the happy emotion doesn't contain only happiness, there will be mixture of some other additional emotions. Like ,the happiness may be shock mixture happiness ,angry mixed happiness. According to this mixture the percentage
will bee displayed.

For example if shock mixed happiness the percentage of happy may be 77% and the shock may be 60% approximately. And its changes dynamically when the emotion of the person changes.

F. FUNCTIONAL AND NON FUNCTIONAL REQUIREMENTS

To analysing the software requirements there are two different ways are there one is functional requirements and another one non-functional requirements. In functional requirements administrator is the main actor he should have taken all the responsible of the major tasks. In non-functional requirements the administrator does not plays the main role. But it judges the operation of a system, rather than specific behaviour.

Functional Requirements:

- ✓ The module have to identify the seven different emotions accurately.
- ✓ The module should allow the users to train the all format datasets.
- ✓ The system should extract the important features from the image.
- ✓ The project should allow the users to take picture using the webcam.
- ✓ The module should easier to understand.
- ✓ The module should extract the feature dynamically.

Non-functional Requirements:

- ✓ The module should developed using C# or R language.
- ✓ The module should give graphs for the results.
- ✓ The module Graphical user Interface should be simple and clear
- ✓ The program should be platform oriented.

G. HARDWARE AND SOFTWARE REQUIREMENT

Hardware requirements

- ✓ 2.0GHz AMD A4-9425 64bit PROCESSOR
- ✓ 4GB / 8GB RAM
- ✓ WEBCAM
- ✓ 5GB HARD DISK SAPCE

Software requirements

- ✓ Python == 3.7.2
- ✓ OS == WINDOWS 8.1 / 10 64BIT
- ✓ TENSORFLOW == 1.1.0
- ✓ KERAS == 2.0.5
- ✓ PANDAS == 0.19.1
- ✓ NUMPY == 1.12.1
- ✓ OPENCV-PYTHON == 3.2.02.0

H. DATASET:

The FER2013 dataset contains more than 40,000 outward appearance pictures for seven feeling classes, including happy, sad, shock, disgust, surprised, scared and neutral. The picture mark breakdown demonstrates a skew towards joy and far from sickens. Since the FER dataset is a lot bigger, we prepared our models with 85% of its pictures and approved with the staying 15% dataset among testing. We utilized information growth strategies to improve our model's execution, connected to both datasets. These methods make an expansion in the extent of existing datasets by applying different changes to existing pictures to make new ones.

These sorts of contrasts in datasets, it is vital to cross-approve execution utilizing different datasets. It's likewise critical to take note of how the neural system performs on explicit feelings. For instance, dread will in general be misclassified more frequently than different feelings. FER2013 pictures are substantially more fluctuated. They originate from numerous sources on the web.

Outward appearances are by their tendency vague, and certain unmistakable feelings can prompt fundamentally the same as outward appearances. Because of dread, at times an individual's eyes are open all around broadly, which can be fundamentally the same as indignation. Dread can likewise deliver tight lips, which in different minutes might be a pointer of quiet. The main idea is to be keep a far away from the problem these misclassifications, especially in profound learning approaches, is to utilize the biggest dataset conceivable.

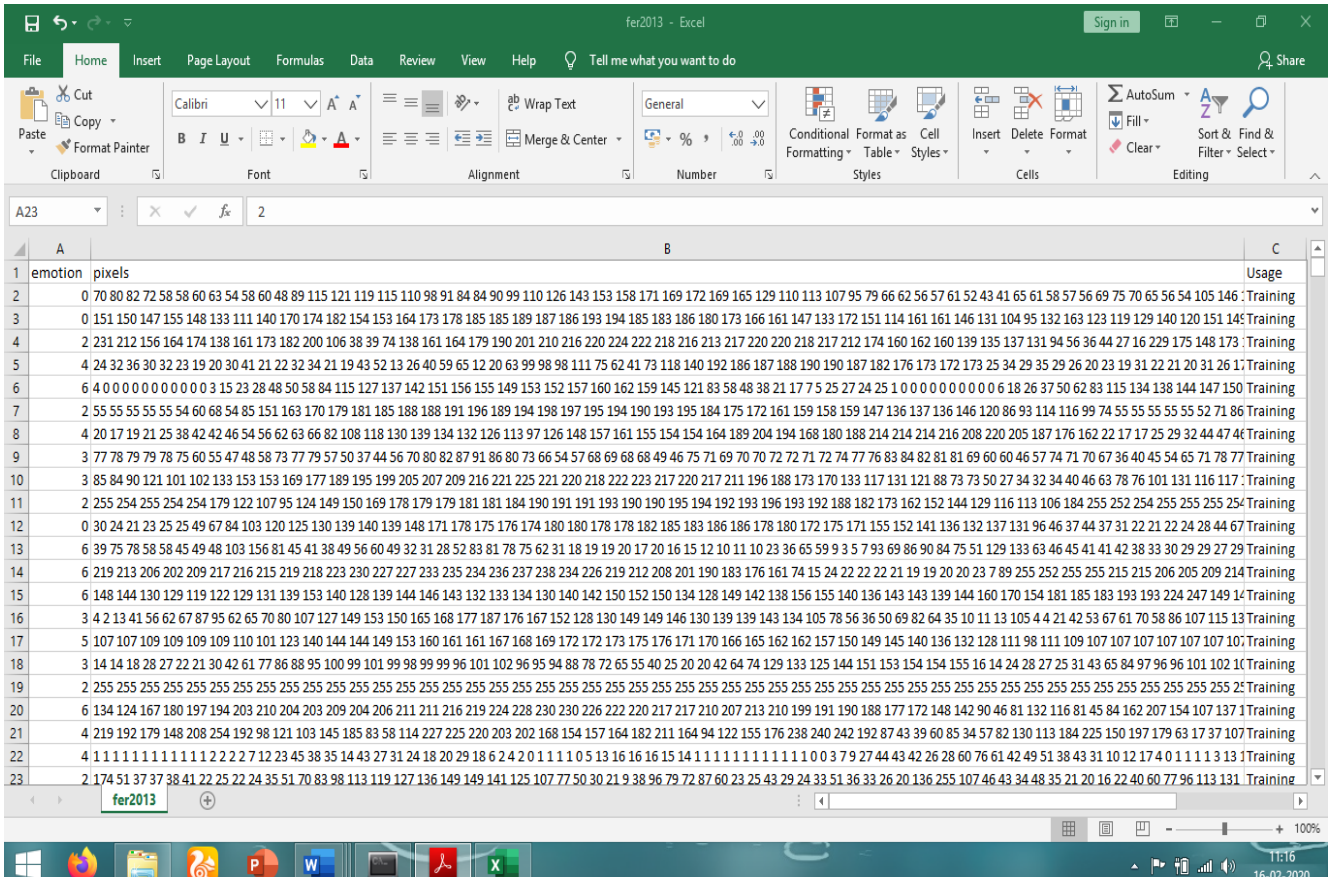

Fig 1.1 The screenshot image of FER2013 dataset

In fig 1.1 the small view of the data set is displayed. It consists of three main parameter. The emotion type,48*48 pixel values and usage. The emotion types are given in the range of 0 to 6 happy i.e neutral, happy, sad, angry, disgust, shocked and surprised. And the usage parameter consists of training data and private data. Training data are the data already trained and private are should be trained. 85% of training data and 15% of private data are present in this dataset.

II. LITERATURE SURVEY

No	Title	Method	Advantage	Disadvantage
1	FER Using Salient Features and Convolutional Neural Network	A profundity of webcam method which is used that technique is proposed here for proficient of outward appearance of the acknowledgment. For the every pixel in a profundity picture, of the eight neighbourhood directional qualities are gotten and positioned. When the position of all pixels is acquired, eight histograms are created for the eight encompassing headings.	In this method would be checked with the other traditional like all the method is the different a little storage where it showing its maximum results over other methods. An achieved the highest minimum rating is 95.20%	The weakness of this approach is CNN lays in the amount of data you provide to them. If we train with less number of images, the output would be poor results only.

2	Multiple Task CNN for Pose Invariant Face Recognition.	This method performs various tasks CNN for face acknowledgment were the personality characterize is the primary errand build up a dynamic-weighting plan to the consequently relegate the misfortune loads to the each side assignment, which takes care of the pivotal issue of adjusting between various undertakings in MTL and propose a vitality based weight of the examination technique to investigate how CNN based MTL functions.	Multi-task CNN actually in this method provides dynamic updating weighting scheme it update weights whenever the weight get lost.	Huge amount training set may reduce the processing speed and take huge amount memory the major disadvantage of multiple task approach for emotion classification. Third, both LFW and IJB-A have large variations other than pose such as the expression and blurring, etc. that cannot be well handled by the proposed method.
3	4D FER by Learning Geometric Deformations	Programmed approach for outward appearance acknowledgment from 3-D video groupings. In the proposed arrangement, the 3-D faces are spoken to by accumulations of outspread bends and a Riemannian shape examination is connected to successfully measure the disfigurements incited by the outward appearances in a given subsequence of 3-D outlines.	This approach it's can be achieved to accurately identifying with very less number of containing very different Three D frames with the maximum accuracy of 92% using HMM-based classifier and 93.21% using mean deformation-based classifier.	One limitation of the approach is, if the face contains hair fully in case of non-frontal views or the presence of occlusions.
4	Survey RGB, 3D, Thermal, and Multimodal Approaches for Facial Expression Recognition	This paper shows a general diagram of programmed RGB, 3D, warm and multimodal outward appearance investigation.	An in depth of the discussion about each step in a AFER pipeline followed, including the comprehensive taxonomy and many examples of techniques used on data captured with different video sensors (RGB, 3D, Thermal).	This paper describes about only the theoretical way gathered from various papers and resources.
5	Automatic analysis of facial affect reg, representation and recognition	Programmed influence investigation has pulled in extraordinary enthusiasm for different settings and including of the acknowledgment of activity units and essential or Non-fundamental feelings. Dissect the cutting edge arrangements by the disintegrating their pipelines into central segments, in particular face enlistment, portrayal, dimensionality decrease and the acknowledgment.	This paper provides the best approach for the (FER) or emotion recognition method to identify the exact expression that approach is contains various illumination technique's to reduce the back-ground noise and other things.	This paper survey make us to understand the problem of the emotion recognition and the future development to develop more in theoretical way.

6	FER Using Hierarchical Features with the Deep of the Comprehensive Multi-Patches Aggregation CNN	This paper has proposed a method, named Profound Complete Multi-patches Total Convolutional Neural Systems, to tackle FER issue. The proposed technique is a profound based structure, which predominantly comprises of two parts of Convolutional Neural System.	Pooling method is proposed it's to handle the if the picture in different angle like rotating the image remove the noise etc. to achieve the more pre-processing to reduce noisy data etc.	In this method gives the 89.5% accuracy but sometimes it gives the false results it cannot identify the exact results all the times.
7	High-Performance and Lightweight Real-Time Deep FER.	This work depicts a progressed pre-preparing calculation for facial pictures and an exchange learning system, two potential possibility for loosening up this prerequisite.	This method is performs well with the low end devices accurately.	In this algorithm,it has one limitation of the approach is if the face contains hair fully in case of non-frontal views and/or the presence of occlusions.
8	Micro-expression Recognition Using the Colour Spaces	A smaller scale articulation shading video cut treat the fourth-request tensor, i.e., a four-measurement exhibit. The initial two measurements are transient data, and the final one is the shading data.	The matching of independent colour and static texture of the spatial data and length of the code is very low.	In the higher dimensional huge amount of the number of neighbours provides that the maximum local information
9	Hierarchical Multi-pose FER	This model gives a bound together answer for multi-present FER, bypassing the different preparing and part of the tuning for each posture, and along these lines is adaptable to countless.	This paper provides the best approach for the (FER) or emotion recognition method to identify the exact expression that approach is contains various illumination of the techniques to the reducing the background noise and other things.	Sometimes it counts the miss detections as wrong expressions, and therefore, it will negatively affect the expression recognition performance.
10	Multi-view FER and Regression	This paper proposes an efficient method called proficient of the calculation is utilizing the estimated expanded Lagrangian multiplier of approach this works achieves 95% accuracy.	This method is identifying with very less number of containing very different Three Dimension frames with an average accuracy of 90% based classifier and 94.01% using mean deformation-based classifier.	If it's a non-frontal face images is the most common problem that have to deal with the problems of face occlusions.
11	Regression-Based Facial Expression Optimization	This paper shows a methodology for repeating ideal 3-D outward the appearances dependent on mix shape relapse.	This method is improves the fidelity of facial expressions, and thus, the user experience while maintaining the efficiency of the blend shape method.	The lack of realistic of expression are represents true facial muscle movements, thus reducing acceptance.

12	Robust FER for A Comprehensive Neuromuscular Signal Analysis	This paper introduces a far reaching study on the investigation of neuro-muscular flag exercises to perceive eleven outward appearances for Muscle Interfacing applications.	Discriminant analysis when the values of the parameter is high this method gives maximum accudracy.	Slow communication rate, detecting small number of facial gestures, being sensitive to position and lightening.
13	FER from Image Sequence based on LBP and Taylor Expansion	This paper presents programmed video-based outward appearance acknowledgment of the framework is to distinguish and order human outward from picture succession.	It handles the illumination changes of the matching of the independent colour and static texture of the spatial data and length of the code is very low.	In this method,it cannot able to be identify the faces if the person have moustache like that.
14	Human FER using Stepwise Linear Hidden Conditional Random Fields	In technique presents the accurate and strong outward appearance framework. For highlight take the feature or extracting proposes FER frame-work utilizes.	This FER method was achieved 93.5% across the 4 unquiet datasets, which is a significant improves in contrast to the existing FER methods.	This method modify the typical CNN structure with the proposed ETI-pooling strategy, which reduces not the impact of nuisance variations in classification tasks.
15	Automatic FER Using Features of Salient Facial Patches	This paper proposed a method called system for demeanor acknowledgment by utilizing appearance highlights of chosen facial patches.	This is the method working with the different resolutions provides an optimum solution for a small size images.	In this method it will give the results well only if the images are facials and doesn't contains any facial hair.

Fig 2.1 literature survey with advantage and disadvantage

III. DESIGN OF SYSTEM

A. APPROACH OF THE DESIGN

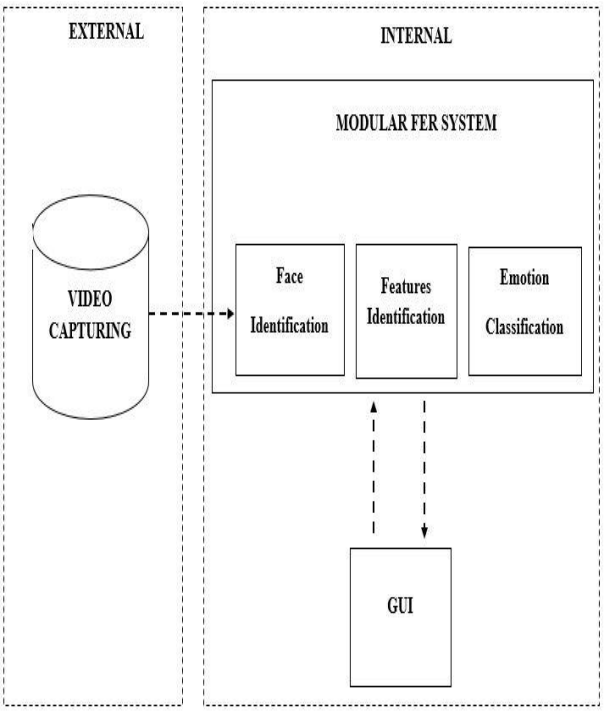

Fig 3.1 Architecture Design

In this fig 3.1 the external process describes about the task which is done externally which the user can able to see and user can control. The internal process describes about the task which is done dynamically without the user involvement .

B. FLOW CHART

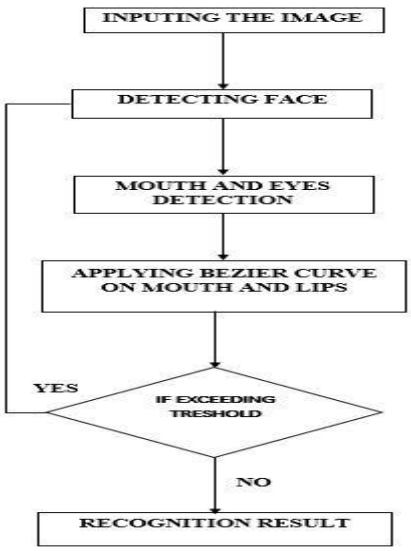

Fig 3.2 flow of the task

In fig 3.2 The threshold value is the sum of all pixel value at the time of face detection . If the sum of the pixel values doesn't match with the sum of pixel value in the FER 2013 data sheet , it seems that the threshold value is exceeded. When the threshold value is exceeded, then the process goes again for detection.

IV. DESCIPRITON OF METHODOLOGY

A. MODULE DESCRIPTION:

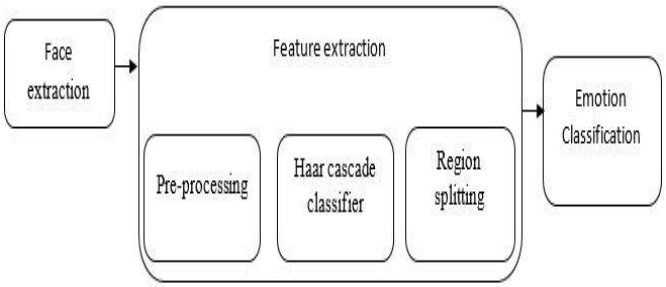

Fig 4.1 Module of the system

FACE EXTRACTION:

The human face is captured by using PC's web cam or external webcam. From that live steam the face is extracted and all other unwanted components are not considered. To achieve this efficiency and comprehensiveness I have picked the Violes-Jones Calculation for this task, so as to identify and remove the countenances. For this I have utilized the Open CV library (to be specific classifier).

FEATURE EXTRACTION:

Pre-processing:

It is a common name for operations with images at the lowest level of abstraction for both input and output are intensity images. The aim of pre-processing is an improvement of the image data that suppresses unwanted distortion or enhances some image features important for further processing.

Haar cascade classifier:

A Haar Cascade is essentially a classifier which is utilize to recognize the item for which it has been prepared for, from the source. The Haar Cascade is prepared by superimposing the positive picture over a lot of negative pictures. The training is commonly done on a server and on different stages. Better outcomes are required by utilizing top notch picture and expanding the measure of stages for which classifier is prepared. One can also use the predefined Haar Cascades which is accessible. Haar cascade classifier depends on the Haar wavelet procedure to analyse pixels in the picture into squares by work. This uses "integral picture" ideas to register the "highlights" distinguished. Haar cascades utilize the Ada-help learning calculation which chooses few significant highlights from an enormous set to give an effective result of classifier then use cascading techniques to detect face in a image.

Region splitting:

For the emotion recognition the main region of face under consideration are eyebrows and mouth. And the splitting of mouth and the eyebrows is named as region splitting.

EMOTION CLASSIFICATION:

After the sub task of feature extraction is completed the reaction of the person is produced simultaneously with their percentage level.

B. FACIAL EXPRESSION EVOLUTION:

Angry:

Includes three fundamental highlights teeth uncovering, eyebrows down and internal side fixing, squinting eyes. The capacity is clear-getting ready for assault. The teeth are prepared to nibble and undermine foes, eyes and eyebrows squinting to ensure the eyes, however not shutting altogether so as to see the adversary. The eyebrows are brought down and drawn together without the gazing eyes for example without raising the upper eyelids, at that point this articulation does not demonstrate outrage. Raised upper eyelids are an unquestionable requirement for the annoyance articulation since when we are furious we gaze at our wellspring of indignation strongly to undermine it. Also, extreme gazing is beyond the realm of imagination without raising the upper eyelids.

Fig 4.2 the angry face

Disgust:

Includes wrinkled nose and mouth. Now and again even includes tongue turning out. This articulation imitates an individual that tasted terrible nourishment and needs to spit it out, or smelling foul smell .An obvious outrageous nauseate articulation. Eyebrows are brought down shaping a 'V' over the nose and delivering wrinkles on the brow; eyes are limited to shut out the wellspring of disturb conceivable jawline is marginally pulled in reverse and a roundabout wrinkle shows.

Fig 4.3 the disgust face

Shock:

May be on the bottom that an astonishing circumstance can terrify us for a concise minute, and after that it depends whether the amazement is a decent or an awful one. In this manner the capacity is comparable.

While dread is activated by any outer data that we decipher as possibly hurtful, surprise is activated by a sudden, unforeseen occasion, paying little heed to its capability to hurt us. Surprises can be lovely as well, in contrast to fear. The outward appearances of dread and surprise are fundamentally the same as and subsequently cause incredible perplexity.

Fig 4.4 the shock face

Sad:

Includes a slight pulling down of lip corners, internal side of eyebrows is rising. Darwin clarified this articulation by smothering the will to cry. The command over the upper lip is more noteworthy than the authority over the lower lip, thus the lower lip drops. At the point when an individual shouts amid a cry, the eyes are shut so as to shield them from circulatory strain that collects in the face. In this way, when we have the inclination to cry and we need to stop it, the eyebrows are ascending to keep the eyes from shutting.

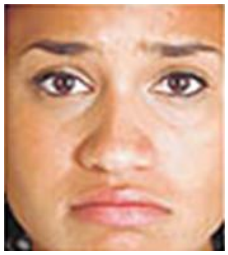

Fig 4.5 the sad face

Neutral:

It does not include in any of the reaction like happy, shock, sad, disgust, angry etc. this expression is a simple one where the lips and eyes are in normal position. Which indicates that the user is not showing any reaction. The default type of emotion is neutral. The every reaction changes starts from the neutral.

Fig 4.6 the neutral face

Scared:

Includes enlarged eyes and now and then open mouth. The capacity opening the eyes so wide is assume to help expanding the visual field to thinks about demonstrate that it doesn't really do as such and the quick eye development, which can help discovering dangers. Opening the mouth empowers to breathe discreetly and by that not being uncovered by the foe. Eyebrow is raised joined creating wrinkles on the brow; eyes have been opened to the greatest, with upper eyelids raised as high as could reasonably be expected; lips are extended on a level plane towards the ears; jaw has been pulled marginally in reverse as is obvious by the flat wrinkles on the neck.

Fig 4.7 the scared face

Happy*:*

Normally includes a grin both corner of the mouth rising, the eyes are squinting and wrinkles show up at eyes corners. The underlying practical job of the grin, which speaks to bliss, remains a riddle. A few scientists trust that grin was at first an indication of dread. Monkeys and primates gripped teeth to demonstrate predators that they are innocuous. A grin urges the mind to discharge endorphins that help reducing torment and take after a sentiment of prosperity. Those positive sentiment that one grin can create can help managing the dread. A grin can likewise create positive affections for somebody who is observer to the grin, and may even inspire him to grin as well.

C. WORKING:

How face extraction works?

OpenCV its commonly well-known library for facial extraction. OpenCV utilizes AI calculations to scan for countenances inside an image. Since countenances are so confused, there isn't one basic test that would be understand it found a face or not. Rather, there are a great many little examples and highlights that must be coordinated. The calculations break the assignment of recognizing the face into a great many littler, nibble estimated undertakings, every one of which is anything but difficult to tackle. These undertakings are likewise called classifiers.

For something like a face, you may have at least 7,000 classifiers, all of which must counterpart for a face to be identified inside blunder limits, obviously. Be that as it may, in that lies the issue: for face recognition, the calculation begins at the upper left of an image and moves down crosswise over little squares of information, taking a gander at each square. Like a progression of cascades, the OpenCV course breaks the issue of distinguishing faces into various stages. For each square, it completes an extremely harsh and speedy test. On the off chance that that passes, it completes a somewhat progressively itemized test, etc.

The calculation might had 20 to 40 this stages of falls, and it will possibly recognize the facial should clear. The preferred standpoint is that most of the image will restore a negative amid the initial couple of stages, which implies the calculation won't sit around idly testing every one of the 6,000 highlights on it. Rather than taking hours, face identification should now be possible progressively. Since face recognition is such a typical case, OpenCV accompanies various inherent falls for recognizing everything from countenances to eyes to hands to legs.

How face detection works?

Understanding the human outward appearances and the investigation of articulations has numerous perspectives, from PC examination, feeling acknowledgment, lie indicators, airplane terminal security, nonverbal correspondence and even the job of articulations in craftsmanship. Improving the abilities of perusing articulations is an essential advance towards fruitful relations. Articulations and feelings go inseparably, for example exceptional blends of face strong activities mirror a specific feeling. For specific feelings, it is exceptionally hard, and perhaps inconceivable, to stay away from it's fitting outward appearance.

For instance, an individual who is endeavouring to disregard his supervisor's irritating hostile remark by keeping an unbiased articulation may in any case demonstrate. This marvel of a short, automatic outward appearance appeared on the essence of people as per feelings experienced is called 'micro expression'. The direct opposite marvel alludes to the manner in which that some muscle developments speak to a feeling, and the contrary muscle developments speak to the contrary feeling. A great clarification for the outward appearance speaks to 'weakness' should be possible utilizing direct opposite. Weakness body motion includes hands spreading to the sides, fingers spreading and shoulders shrugging. Its outward appearances include pulling down the base lip and raising eyebrows. Darwin clarified the highlights of this articulation utilizing the direct opposite standard. He found that those developments restricting to the developments of a man who is prepared to confront something.

The developments of an individual who is setting himself up for something will resemble that: shut hands and fingers (as though he is getting ready for a battle, for instance), hands near the body for insurance and the neck is raised and tight. At a weakness circumstance the shrugging of the shoulders discharges the neck. Concerning the face: eyebrows are low (like in a method of assault or solidness), upper lip may uncover teeth. The practical wellspring of the absolute opposite can be clarified with the examination of muscles, and to be exact the rival's muscles. Each muscle has an adversary muscle that plays out the contrary development. Spreading fingers is a development done by a few muscles, and shutting the fingers is finished by the foe muscles.

For a few articulations we can't generally tell just by taking a gander at it, what is the contrary articulation, so the choice that took the gander the muscles including all the while, at that point it turns out to be exceptionally clear. A fascinating clarification to the direct opposite practical source depends on hindrance. An individual or a creature is attempting to anticipate completing a specific activity, one path is to utilize the hostile muscles. Indeed, when an improvements flag is send to a muscle, an inhibitory flag is send naturally to the rival muscle. Outward appearances that can be clarified with direct opposite all the parts of identity with hostility and maintaining a strategic distance.

How CNN works?

How a human figures out how to perceive objects, we have to demonstrate a calculation on a huge number of pictures before it is have the capacity to sum up the information and make expectations for pictures it has never observed. PCs 'see' uniquely in contrast to we do. Their reality comprises of just numbers. Each picture can be spoken to as 2-dimensional varieties of numbers, known as pixels. Any of the case have thing that they see pictures in an unexpected way, doesn't mean we can't prepare them to perceive designs, as we do. Its need to be consider what a picture is in an unexpected way.

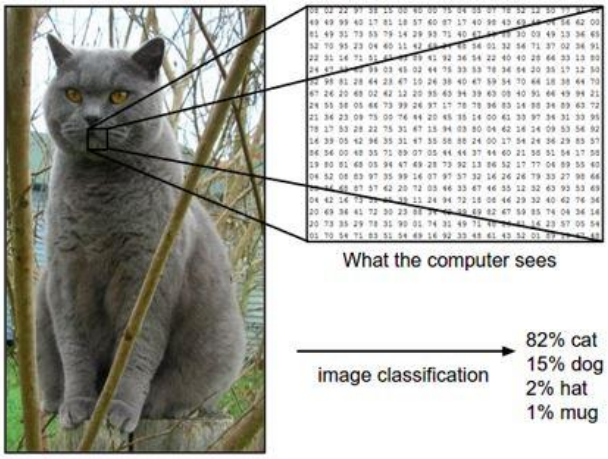

Fig 4.8 the image identification and classification (ref: MIT 6.S094 deep learning lecture)

Fig 4.8 shows the calculation how to perceive protests in pictures, we utilize a particular kind of Artificial Neural Network, a Convolutional Neural Network (CNN). The name comes from it's the most important tasks in the system called convolution. The straightforward cells enact, for instance, when they recognize fundamental fixed by lines as the shapes territory and a particular point. The unpredictable cells have bigger open fields and their yield isn't touchy to the particular position in the field. The complex cells keep on reacting to a specific improvement, despite the fact that its total direction of the eyes will change. Complex alludes to increasingly adaptable, for this situation. In vision, a responsive field of a solitary tangible area of the retina in which something will influence the terminating of that neuron (that is, will dynamic the neuron). Each tangible neuron cell has comparable open fields, and their fields are overlying.

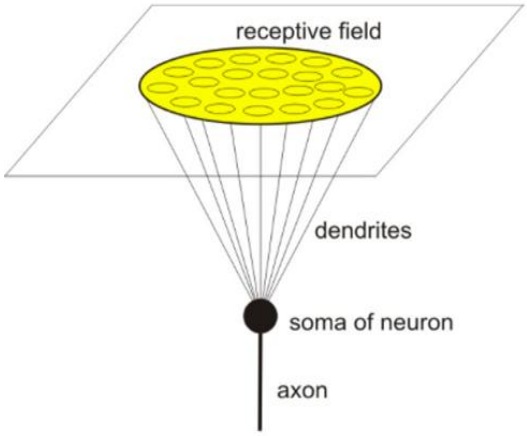

Fig 4.9 neuron and its parts

(ref: the neural model of receptive field structure and output signal of neuro)

The fig 4.9 is used to describe the stages of neural network. The level 1(input layer) for the system is receptive field the dendrites are the layer2 (hidden layer) and the soma of neuron is the layer 3 i.e output layer.

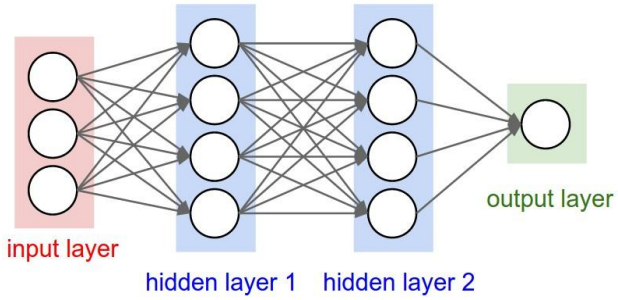

Fig 4.10 neural network diagram

(ref: coding-neural forward propagation and backward propagation)

Lets compare the fig 4.9 and 4.10 for the clear understanding the concept of neural network. The hidden layer computes the weight and the bias for the purpose of activation purpose. The bias and weights were adjusted to get the accurate output.

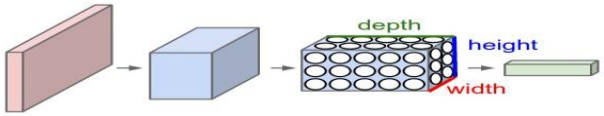

Fig 4.10 Extraction of needed parts from face

(Fig 4.10 ref:convolutional neural network introduction part 2)

In fig 4.10 it shows that the huge pictures(input) is diminished in the upcoming layers of the neural network and at last the unwanted parts are eliminated in every layer and the output is produced for our task. In this paper, the face is the input for the system. The input is sent to the next layer and at last the eyes and mouth is extracted and produced as the output.

V. RESULT

STEP 1: Find the path and paste it in the anaconda prompt

(base) C:\Users\praveenr>cd C:\Users\praveenr\Downloads\Documents\P\Face Expression Recognition using CN

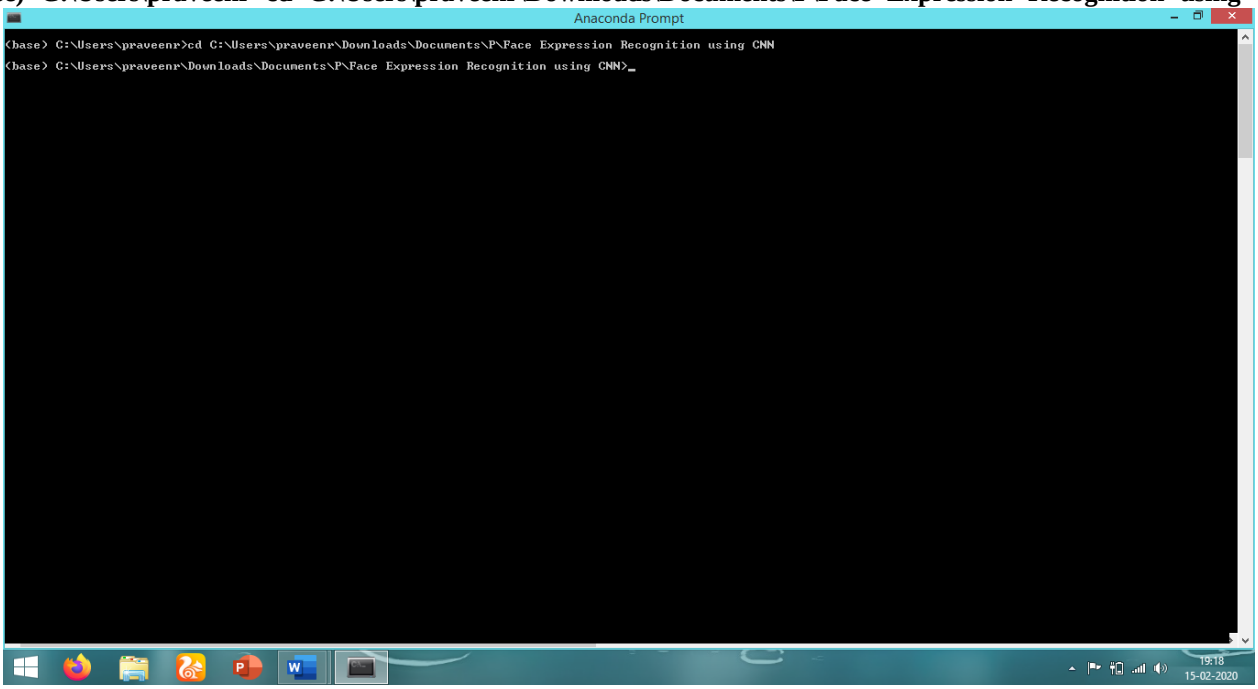

Fig 5.1 The Path of the Python File

STEP 2: Activate the tensorflow

(base) C:\Users\praveenr\Downloads\Documents\P\Face Expression Recognition using CNN>activate tensorflow

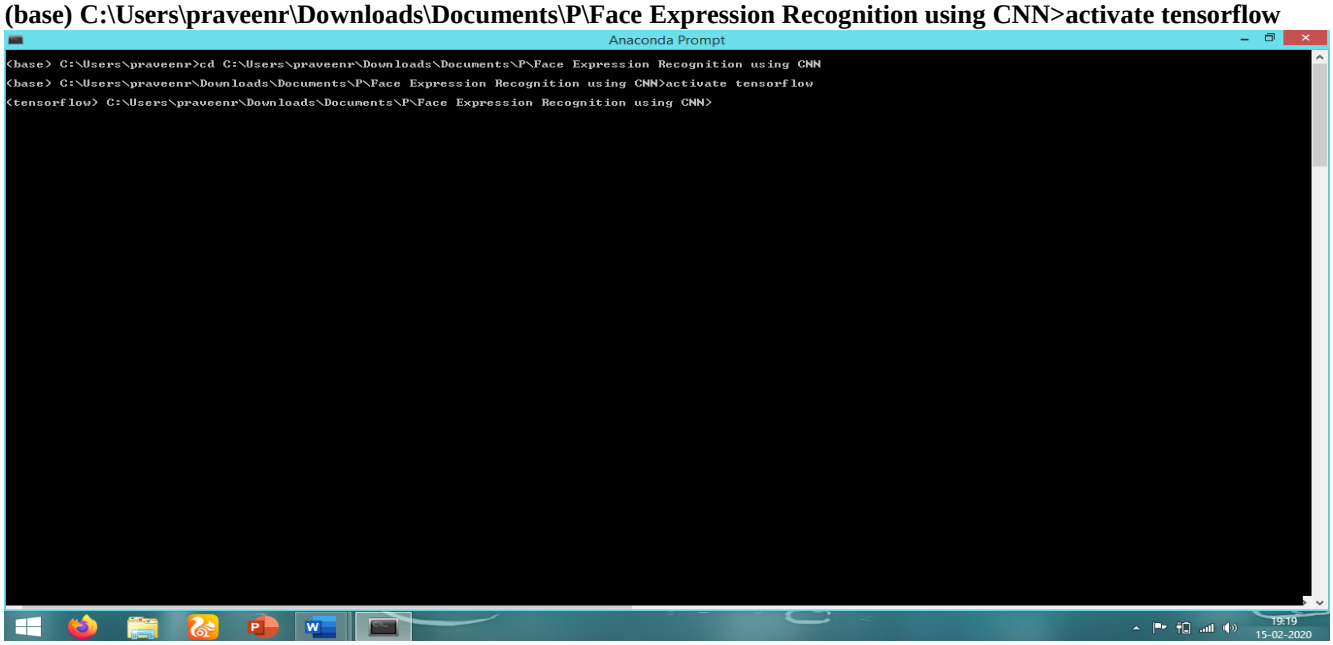

Fig 5.2 Activation of Tensorflow

STEP 3: Running the python

(tensorflow) C:\Users\praveenr\Downloads\Documents\P\Face Expression Recognition using CNN>python real_time_video.py

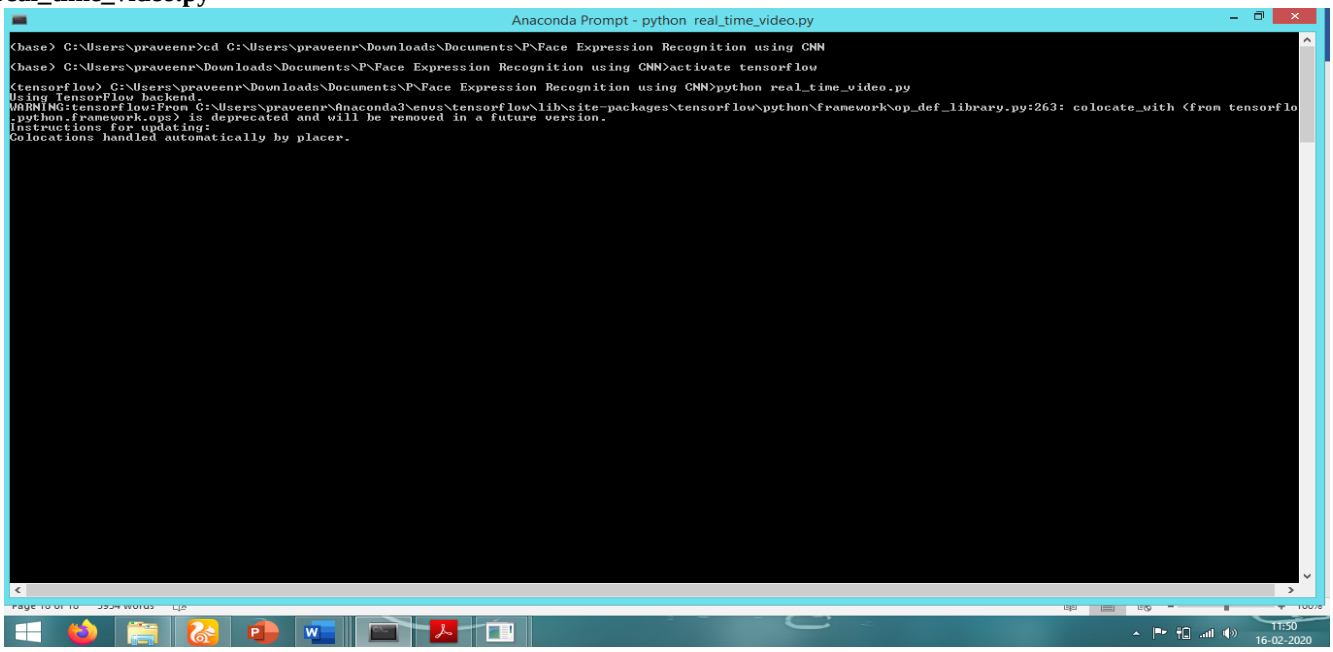

Fig 5.3 Runing the real_time_video Python File

A. NEUTRAL:

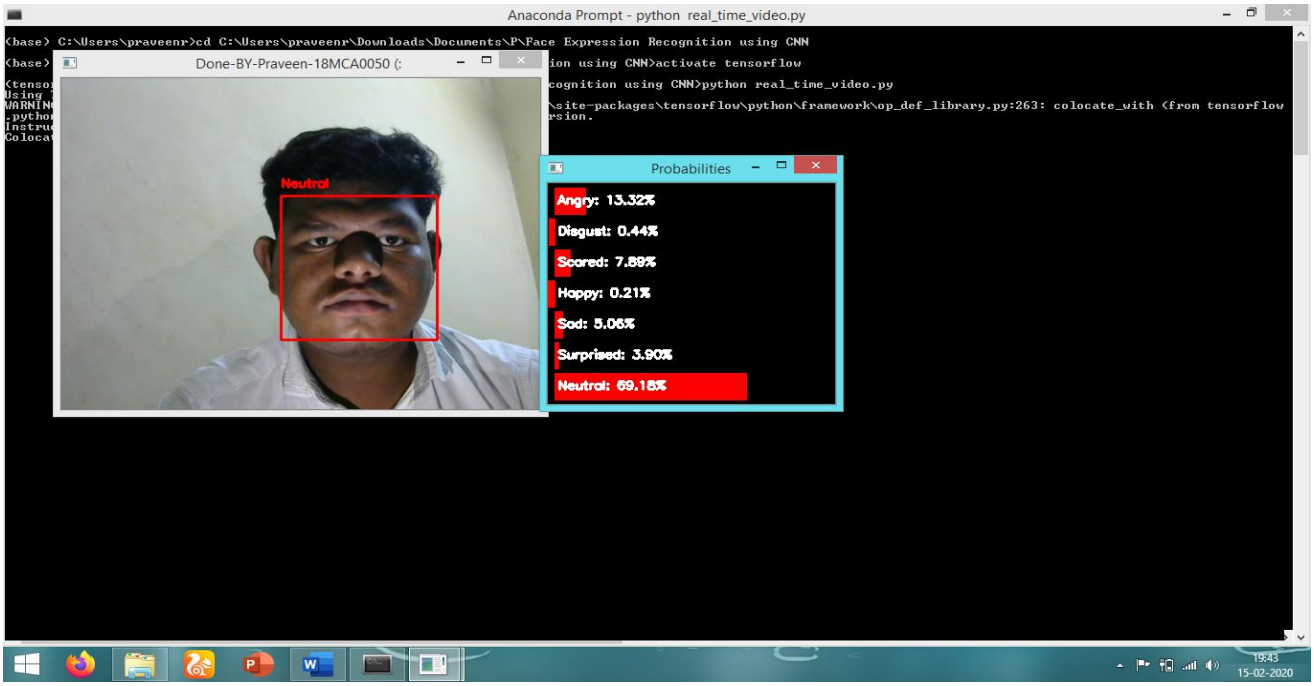

Fig 5.4 Neutral Emotion

B. HAPPY:

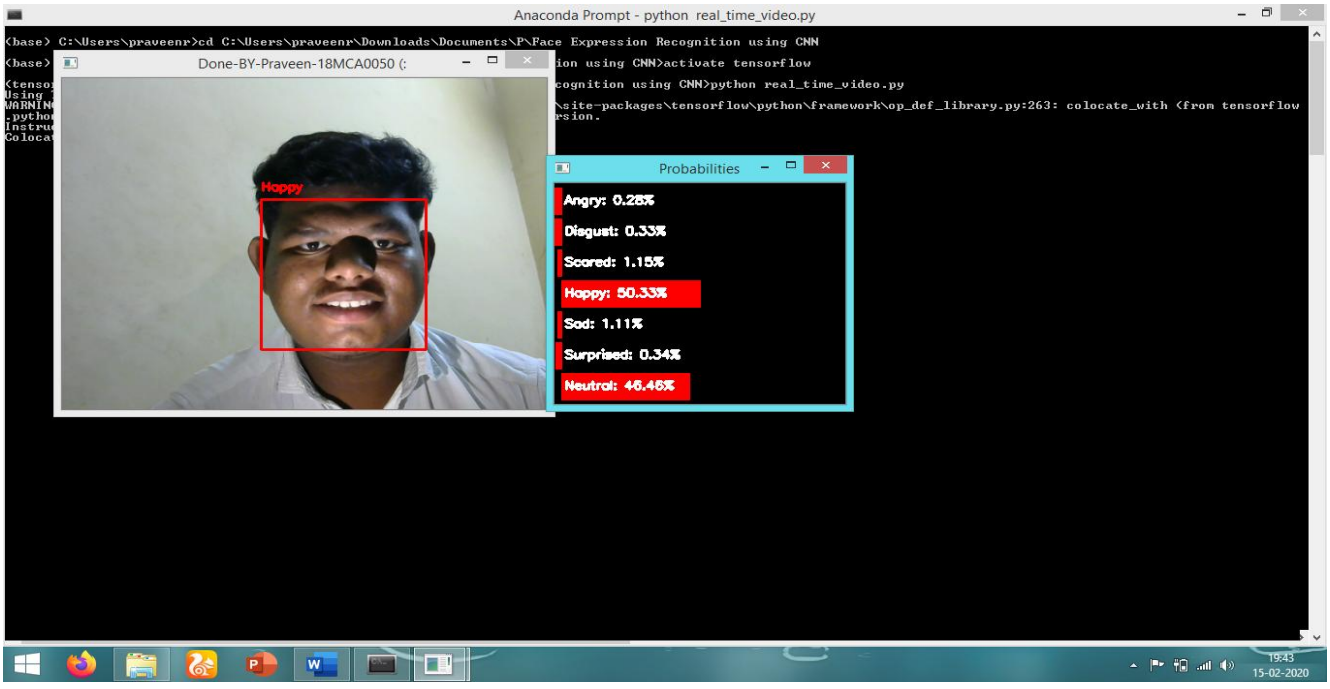

Fig 5.5 Happy Emotion

C. SAD:

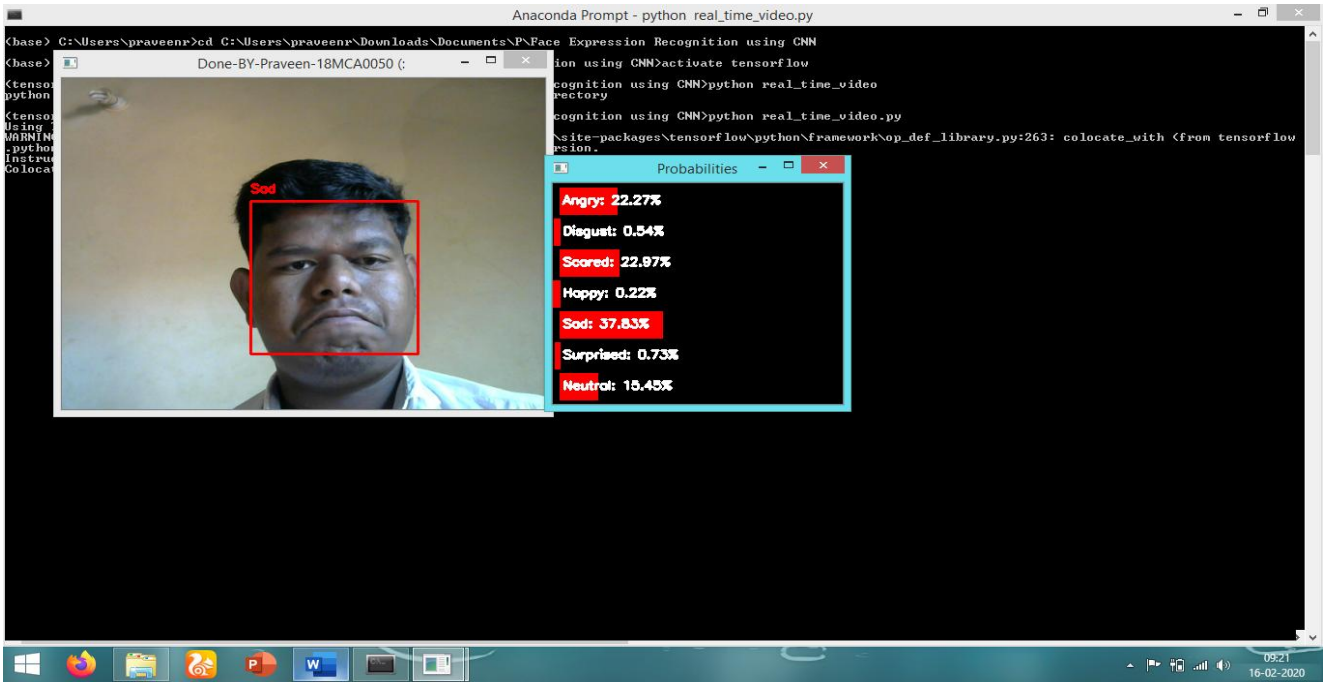

Fig 5.6 Sad Emotion

D. DISGUST:

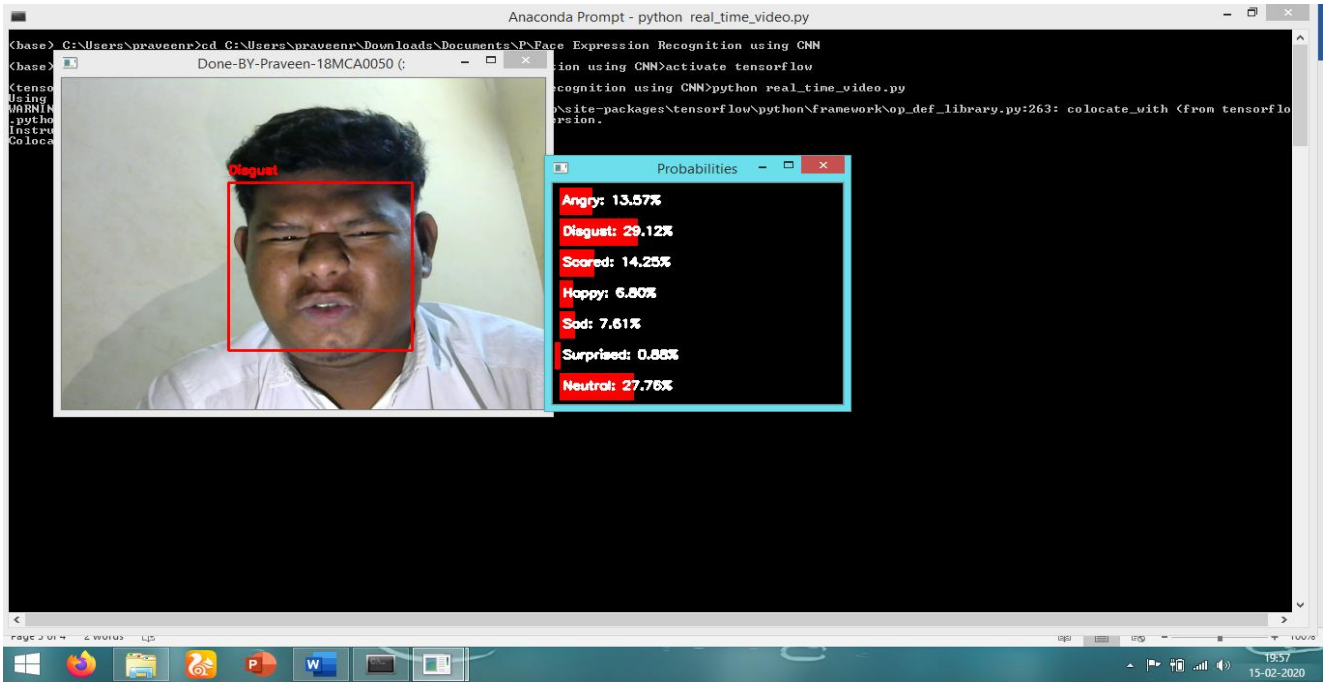

Fig 5.7 Disgust Emotion

E. ANGRY:

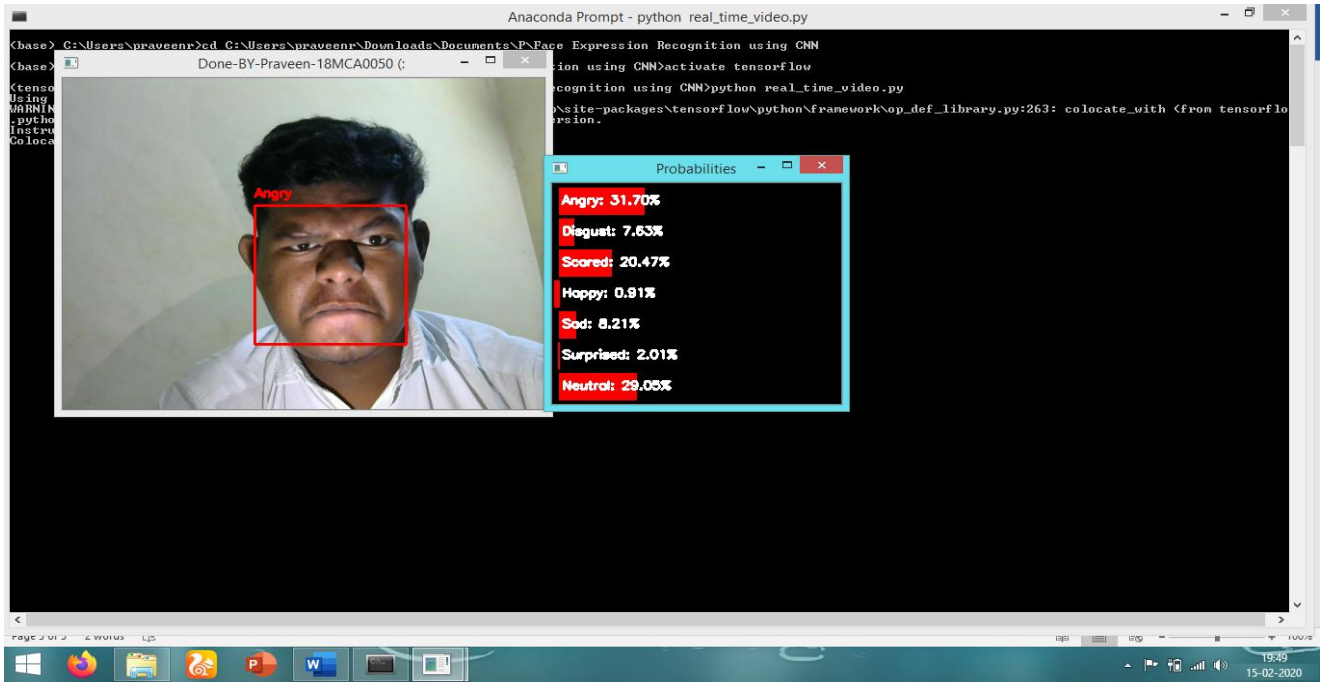

Fig 5.8 Angry Emotion

F. SURPRISED:

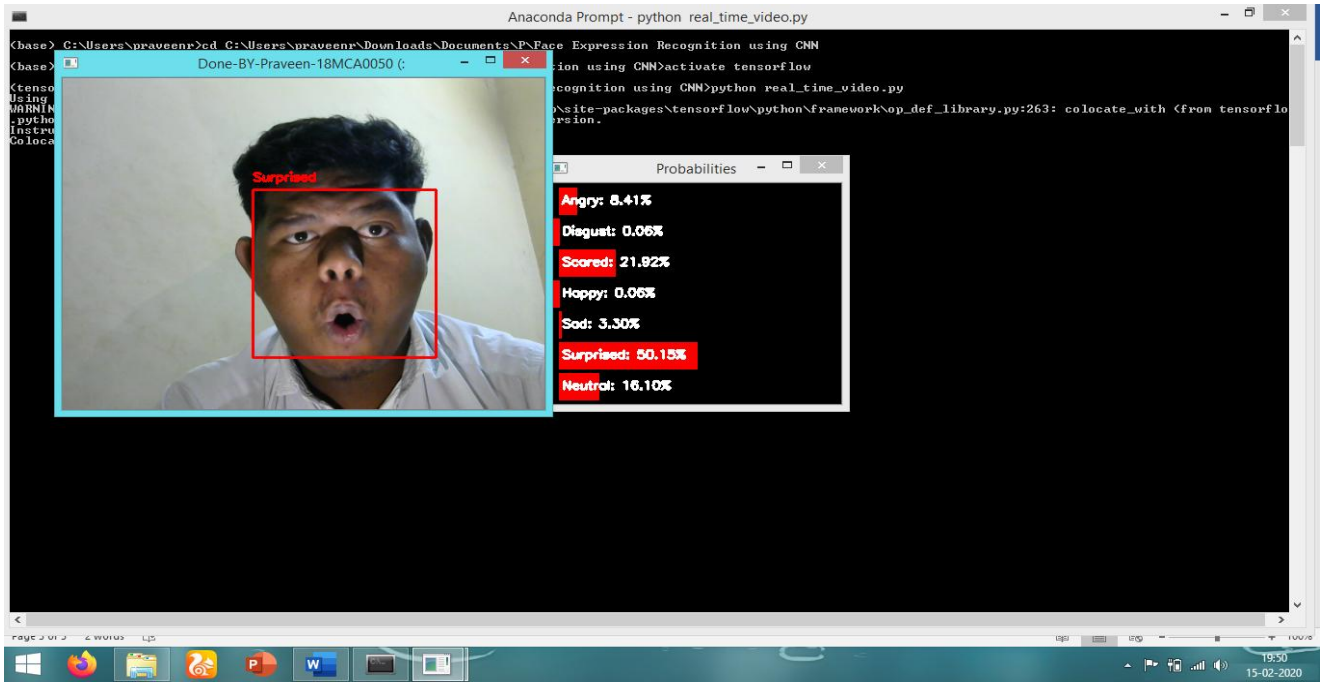

Fig 5.9 Surprised Emotion

G. SCARED:

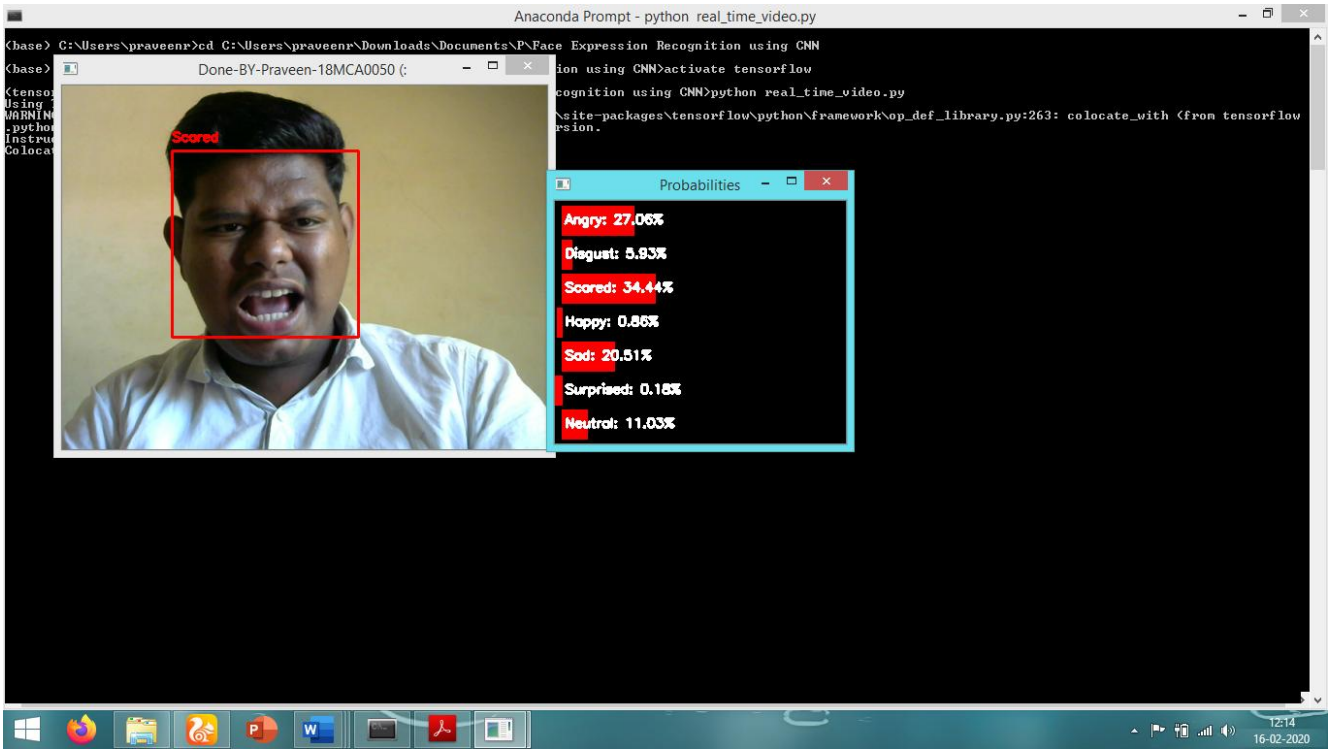

Fig 5.10 Scared Emotion

H. OTHER FEATURES:

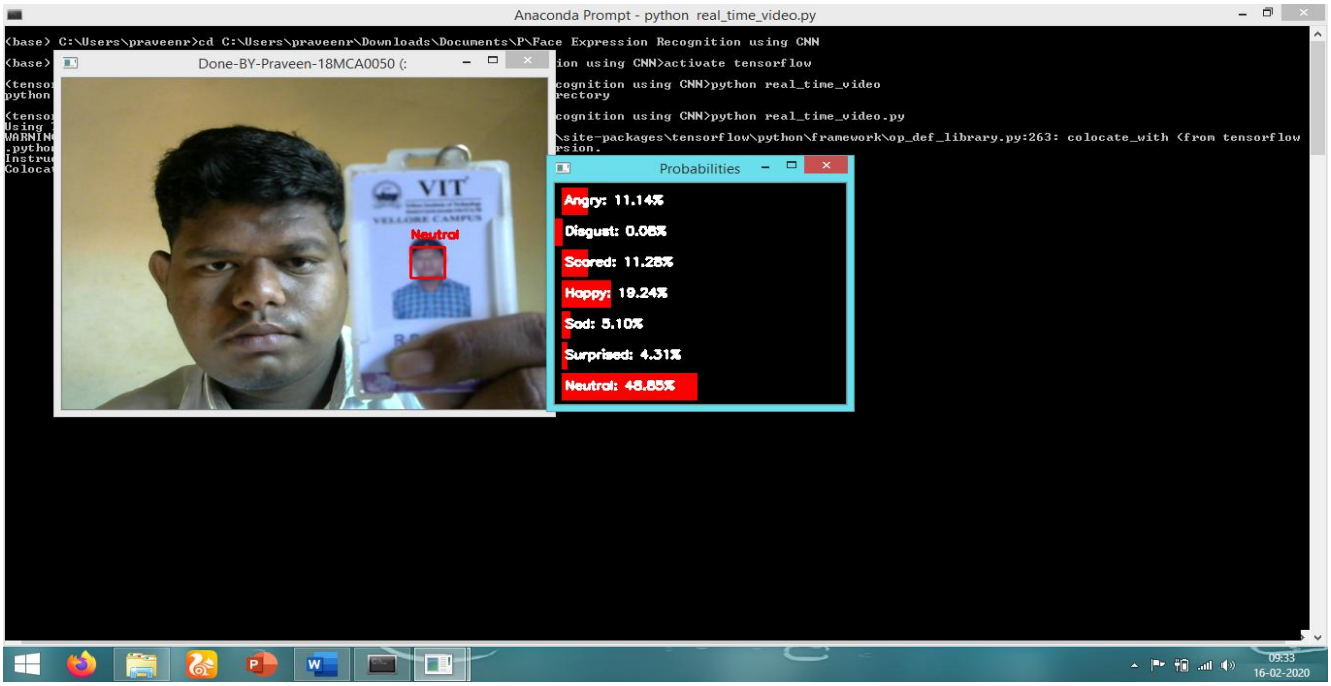

Fig 5.11 Recognition of emotion in ID Card

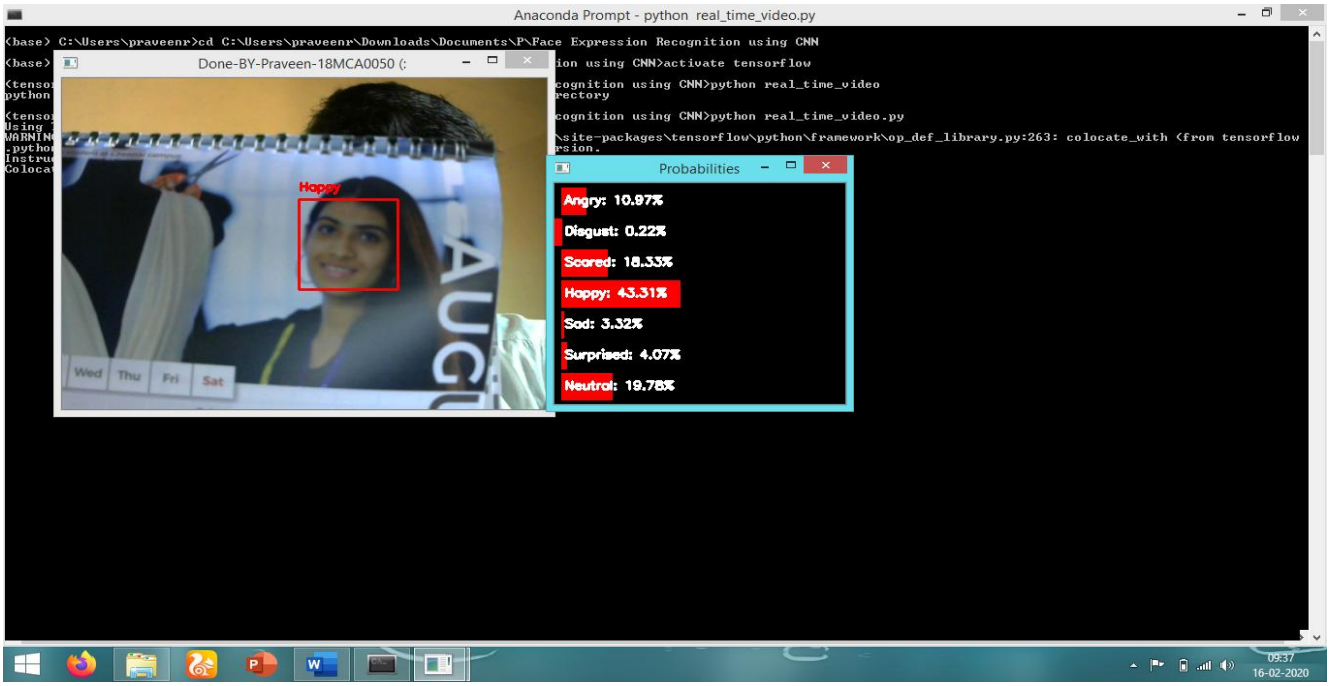

Fig 5.12 Emotion recognition of Emotion in Calendar picture

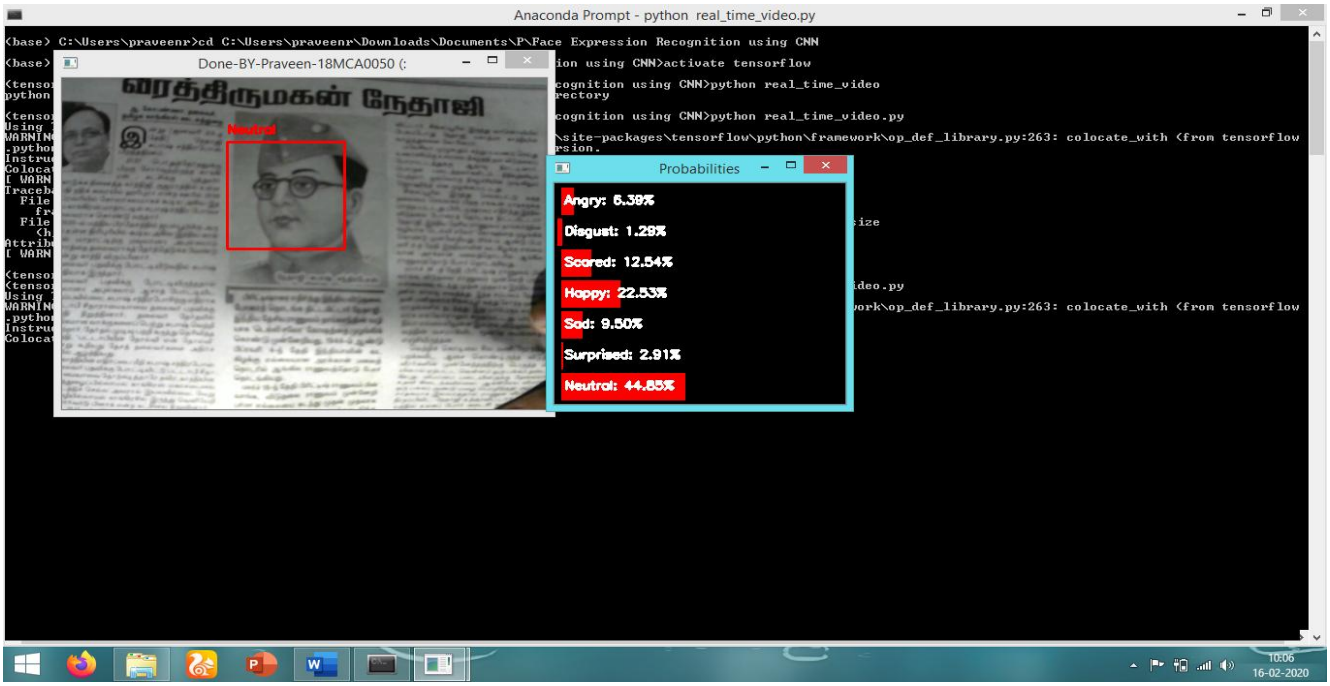

Fig 5.13 Emotion recognition of black and white image in newspaper

VI. CONCLUSION

In this project, the expressions of the faces are effectively identified by processing the dataset that consists of various facial expression which is then coded in python or classification. Our proposed architecture is recognizing the emotion of human face dynamically. Here, the main parameter consider is the position of the eyes and the mouth. The emotion is recognized according to the position change of eyes and mouth. Here in addition, it
displays the percentage of every reaction of a person dynamically and from the data, each and every single data is processed in such a way that it takes the portion of the image and keeps on cropping the image and tries to get average or maximum information out of it, which is termed as pooling.

REFERENCE

1. Siyue Xie and Haifeng Hu, “Facial Expression Recognition Using Hierarchical Features with Deep Comprehensive Multi-Patches Aggregation Convolutional Neural Networks” IEEE Transactions on Multimedia, 2019.
2. Weria Khaksar, And Jim Torresen, “Facial Expression Recognition Using Salient Features and Convolutional Neural Network.” IEEE Access December 2017.

3. Mao, Q., Rao, Q., Yu, Y., & Dong, M. "Hierarchical Bayesian Theme Models For Multipose Facial Expression Recognition." IEEE Transactions On Multimedia Vol. 19, 2017.
4. Xi Yin and Xiaoming Liu, "Multi-Task Convolutional Neural Network for Pose-Invariant Face Recognition" IEEE Transactions on Image Processing, 2018.
5. SL Happy, A Rout ray "Automatic Facial Expression Recognition Using Features of Salient Facial Patches" IEEE transactions on Affective, 2016.
6. Q Mao, Q Rao, Y Yu, M Dong "Hierarchical Bayesian Theme Models for Multi-pose Facial Expression Recognition" IEEE Transactions, 2017.
7. SJ Wang, WJ Yan, X Li, G Zhao "Micro-expression Recognition Using Color Spaces" IEEE Transactions on Image Processing, 2016.
8. MH Siddiqi, R Ali, AM Khan, "Human Facial Expression Recognition using Stepwise Linear Discriminant Analysis and Hidden Conditional Random Fieldsy" IEEE Transactions, 2017.
9. Wzheng," Multi-view Facial Expression Recognition Based on Group Sparse Reduced-rank Regression" IEEE Transactions on Affective Computing, 2017.
10. J Schwan, E Ghaleb, E Hortal, "High-Performance and Lightweight Real-Time Deep Face Emotion Recognition" IEEE Transactions on Semantic and Social, 2017.
11. Mohammad Soleymani, Sadjad Asghari-Esfeden, Yun Fu, Maja Pantic, "Analysis of ECG signals and facial expressions for continuous emotion detection", IEEE transaction on affective computing, 2019.
12. Wenming Zheng, Yuan Zong, Xiaoyan Zhou, Minghai Xin, "Cross-Domain Color Facial Expression Recognition Using Transductive Transfer Subspace Learning", IEEE transaction on affective computing, 2016.
13. Md. Zia Uddin, Mohammed Mehedi Hassan, Ahmad Almogren, Atif Alamri, Majed Alrubaian and Giancarlo Fortino, "Facial Expression Recognition Utilizing Local Direction-based Robust Features and Deep Belief Network", IEEE Access, 2016.
14. Stefanos Eleftheriadis, Ognjen Rudovic and Maja Pantic, "Discriminative Shared Gaussian Processes for Multiview and View-Invariant Facial Expression Recognition", IEEE Transactions on Image Processing, Vol – 24, Jan 2019.
15. Seung Ho Lee, Konstantinos N. (Kostas) Plataniotis and Yong Man Ro, "Intra-Class Variation Reduction Using Training Expression Images for Sparse Representation Based Facial Expression Recognition", IEEE Transactions on Affective Computing, Vol – 5, July – September, 2014.
16. Chao Qi, Min Li, Qiushi Wang, Huiquan Zhang, Jinling Xing, Zhifan Gao and Huailing Zhang, "Facial Expressions Recognition Based on Cognition and Mapped Binary Patterns", IEEE Access, 2018.
17. Ognjen Rudovic, Vladimir Pavlovic and Maja Pantic. "Context-sensitive Dynamic Ordinal Regression for Intensity Estimation of Facial Action Units", IEEE transactions on pattern analysis and machine intelligence, Vol. XX, NO. XX, XXXXXXX, 2013.
18. Mohammad Reza Mohammadi, Emad Fatemizadeh, and Mohammad H. Mahoor, "Intensity Estimation of Spontaneous Facial Action Units Based on Their Sparsity Properties", IEEE TRANSACTIONS ON CYBERNETICS, 2018.
19. https://hackernoon.com/mit-6-s094-deep-learning-for-self-driving-cars-2018-lecture-4-notes-computer-vision-f591f14b3b99
20. http://neuroclusterbrain.com/neuron_model.html
21. https://becominghuman.ai/not-just-introduction-to-convolutional-neural-networks-part-2-a7ac2723e30d
22. https://www.i2tutorials.com/technology/hidden-layers-in-neural-networks/

AUTHORS PROFILE

R.Praveen, pursuing Master Of Computer Application in VELLORE INSTIUTE OF TECHNOLOGY,VELLORE. My area of Interest is Deep Learning, Internet of Thing, Data Mining.EMAIL: praveenwin25@gmail.com

Prof Benjula Anbu Malar.M.B, working as a assistant professor in VELLORE INSTIUTE OF TECHNOLOGY, VELLORE. My area of interest is cloud computing, Internet Of Things and Networks EMAIL :benjula2929@gmail.com