www.ijiarec.com

ISSN:2348-2079

Volume-9 Issue-2

International Journal of Intellectual Advancements and Research in Engineering Computations

Emotion detection using CNN

Mrs.A. Thamaraiselvi M.E[1]., A.Anushree[2], D. Kalaiarasi[3], P.Kirubashini[4], J.P Megna Walter[5]

Assistant Professor/CSE Department, Vivekanandha College Of Engineering For Women (Autonomous).

UG-Students, CSE Department, Vivekanandha College Of Engineering For Women (Autonomous**)**.

ABSTRACT

Emotion Recognition is a method to assess human facial expression and classify it into positive emotion categories. Such assignment usually requires the function extractor to discover the feature, and the skilled classifier produces the label primarily based on the feature. The hassle is that the extraction of characteristic might also be distorted by way of variance of vicinity of object and lights circumstance in the image. Emotion Detection Using CNN project, gives answer for this hassle via the use of a deep studying algorithm known as Convolution Neural Network (CNN) to tackle the troubles. By the use of CNN algorithm, the characteristic of photo can be extracted besides user-defined feature-engineering, and classifier mannequin is built-in with characteristic extractor to produce the end result when enter is given. In this way, this technique produces a feature-location-invariant picture classifier that achieves greater accuracy than normal linear classifier when the variance such as lighting fixtures noise and history surroundings seems in the entered image.

Keywords: Emotion recognition, Convolutional neural network (CNN), Feature extraction, Deep learning.

INTRODUCTION

Human beings communicate with one another in the forms of speech, gestures and emotions. A pc will be capable to engage with human beings an awful lot extra naturally if they are successful of perception human emotion. It would additionally assist in the course of counseling and different fitness care associated fields. In an E-Learning system, the presentation fashion may also be assorted relying on the students state. However, in many cases, static emotion detection is no longer very useful. It is integral to recognize the user emotions over a length of time in a stable environment. Thus, our paper proposes an algorithm that is aimed at real-time facial emotion recognition. For real-time purposes, facial emotion focus has a wide variety of applications. Facial emotion consciousness ought to be used in conjunction with different structures to furnish a structure of safety. For instance, ATMs may want to be set up such that they won't dispense cash when the person is scared. In the gaming industry, emotion-aware video games can be developed which ought to fluctuate the subject of a degree relying on the player's emotions. At present, gamers generally provide some shape of verbal or written feedback. Using facial emotion recognition, the wide variety of testers can be multiplied to accommodate human beings who use unique languages or human beings who are unable to cohesively convey their opinions on the game. By judging their expressions at some point of special factors of the game, a time-honored appreciation of the game's sturdy and susceptible factors can be discerned.

Emotions can additionally be gauged whilst a viewer watches advertisements to see how they react to them. This is specifically beneficial when you consider that advertisements do now not

Author for correspondence:
Assistant Professor/C.Se Department, Vivekanandha College Of Engineering For Women (Autonomous)

normally have remarks mechanisms aside from monitoring whether or not the advert was once watched and whether or not there used to be any person interaction. Software for cameras can use emotion awareness to take pictures each time a consumer smile. The goal of this mission is to classify human faces into one of the six ordinary thoughts or a seventh impartial emotion.

The following two strategies are used for respective noted duties in face awareness system.Haar cascade is feature-based cascade classifier: It detects frontal face in a picture well. It is actual time and quicker in evaluation to different face detector. Xception CNN Model (Mini_Xception, 2017): We will teach a classification CNN mannequin structure which takes bounded face (48*48 pixels) as entered and predicts chances of 7 thoughts in the output layer.

Objective

The facial expression of human emotion is one of the predominant subjects in facial recognition. This projection constructs a device of deep gaining knowledge model to classify a given photo of human facial emotion into one of the seven simple human emotion. The method we take to construct the mannequin is thru switch gaining knowledge of an present pre-trained model, and the end result will be evaluated primarily based on accuracy of the model. The duties in this task consist of preprocessing the picture data, augmentation to increase the present small dataset, take a look at earlier than coaching the model, education process, and predication with evaluation. The baseline of the take a look at will be round 14.7% (1 in7), and our goal is to reap end result higher than baseline accuracy. There are seven exclusive emotions: happy, angry, sad, fear, surprise, neutral, and disgust. The input will be uncooked picture of the expression, and output will show the emotion of the image.

Existing System

VGG is a convolutional community layer that is sixteen layer deep. It is a pretrained network educated with tens of millions of pix from ImageNet Database. It can classify photographs into a thousand object categories. They are utterly related feed ahead neural networks.

VGG16

The VGG community structure was once delivered by way of Simonyan and Zisserman in their 2014 paper, Very Deep Convolution Networks for Large Scale Image Recognition. This community is characterized via its simplicity, the usage of solely 3×3 convolutional layers stacked on pinnacleof every different in growing depth. Reducing quantity dimension is treated by using max pooling. Two fully-connected layers, every with 4,096 nodes are then observed by means of a softmax classifier. The **“16”** and **“19”** stand for the quantity of weight layers in the network. VGG16- a Convolutional Neural Network model which was proposed by K .Simonyan and A .Zisserman. This VGG16 model was one of the famous CNN algorithm which was submitted in ILSVRC-2014.

VGG16 Architecture

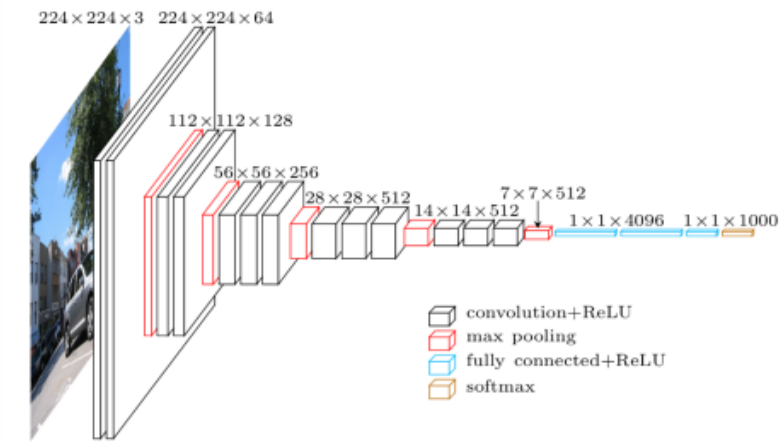

Figure1. VGG16 architecture

- The network architecture weights themselves are quite large (connecting disk/bandwidth). Not the best when comes to practical approach in comparison to with other CNN types.
- Three fully-connected (FC) layers follows a stack of convolutional layers : first two have 4906 channels each, third performs 1000-way ILSVRC classification.
- The final layer is the soft-max layer.
- Spatial pooling is consist of five max-pooling layers, max-pooling is performed over a 2x2 pixel window.

Proposed System

In this project, the algorithm used is Xception CNN in the location of VGG16. It replaces different modules with depth clever separable convolutions. It is an extension of different convolution models. Exception in a convolution layer is seventy one layers deep. It is a mixture of two most profitable layers specifically residual modules and depth smart separable convolutions. It is additionally a pretrained mannequin for photograph classification.

Xception CNN

Xception is an adaptation from Inception, where the Inception modules have been replaced with depthwise separable convolutions. It has also roughly the same number of parameters as Inception-v1 (23M). Xception is an extension of the Inception architecture which replaces the standard Inception modules with depthwise separable convolutions. Xception sports the smallest weight serialization at the size of only 91MB.

Advantages

- It automatically detects the errors without human supervision.
- It helps us overcoming two important problems by reducing the number of parameters.
- One is the entities cannot be multiplied without necessity, another is use of small CNN alleviate us from slow performances in hardware constrained system such robot platforms.

System Flow Design

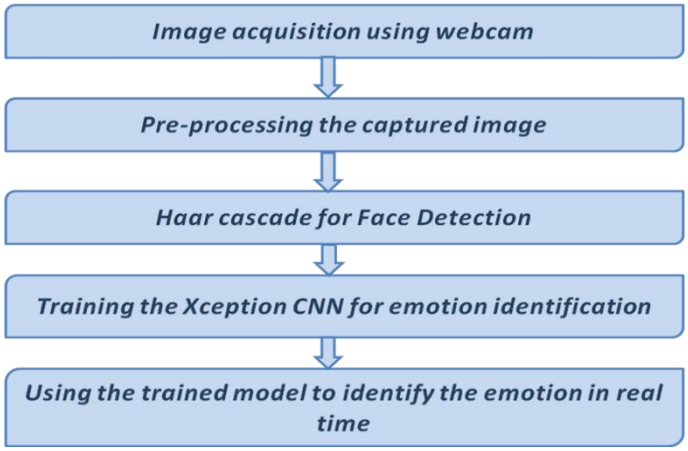

Figure 2. System flow diagram.

This system is divided into different modules and each performs specific tasks in an operating sequence. At first, the system detects the face of the user using the webcam. Then, the captured image is pre-processed. The face detection is done using Haar Cascade classifier. Haar Cascade classifier is a machine learning based approach for object detection approach where a cascade function is trained from a lot of images. Based on the training it is then used to detect the objects in the other images.

The captured image is given as input to CNN which learns features directly. A Convolutional Neural Network is a Deep Learning algorithm which takes an input image, assign importance to various aspects/objects in the image and be able to differentiate one from the other. The pre-processing required in a CNN is much lower as compared to other classification agorithms.

While in primitive methods filters are hand-engineered, with enough training, CNN have the ability to learn these filters/characteristics automatically. Xception CNN model is used for

predicting the emotion. Xception is a deep convolutional neural network architecture that involves Depth-wise Separable Convolutions.

Xception architecture has over performed VGG-16, Res Net and Inception V3 in most classical classification challenges. Each song is assigned with an emotion tag. When the emotion is predicted, songs with those tags are played. For the emotion recognition module, the system is trained using Kaggle FER2013 dataset containing images of happy, anger, sad, neutral, disgust, fear and surprise emotions

Algorithms Used In Emotion Detection

Haar Cascade

Haar Cascade is a machine learning object detection algorithm proposed by Paul Viola and Michael Jones in their paper "Rapid Object Detection using a Boosted Cascade of Simple Features" in 2001.

It is a machine learning based approach where a cascade function is trained from a lot of positive and negative images. It is then used to detect objects in other images. OpenCV offers pre-trained Haar cascade algorithms, organized into categories (faces, eyes and so forth), depending on the images they have been trained on.

Xecption CNN

Xception is a very deep model of convolutional neural networks. A deep CNN architecture that involves depth-wise separable convolutions. Developed by google researchers. Over performed VGG-16, Res Net and Inception V3 in most classical problems. Xception is the combination of two most successful experimental assumptions namely residual modules and depth-wise separable convolutions.

Xception Model

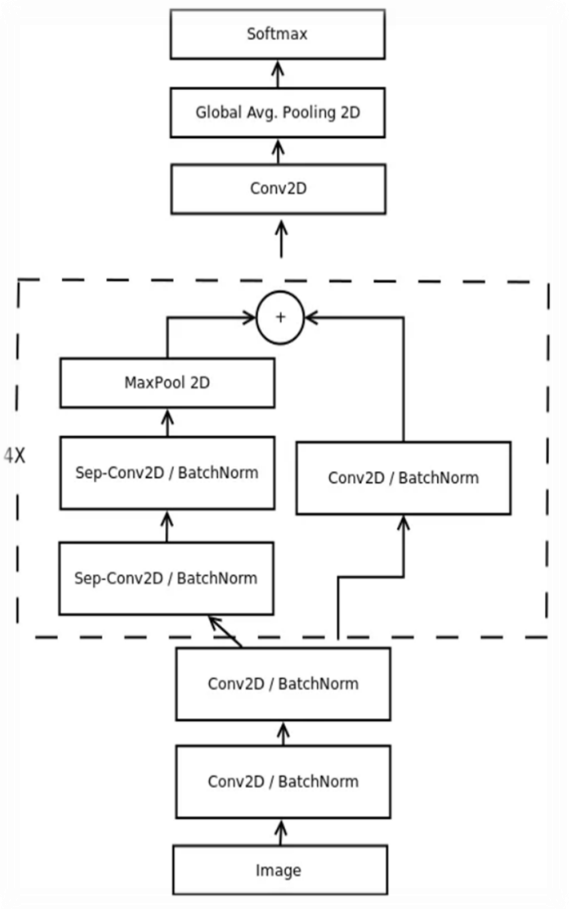

Figure 3. Xception CNN architecture

Implementation Of CNN Algorithm

Tested the model that we build for emotion detection in real-time using OpenCV and webcam The Jupyter notebook is used in our local system to make use of a webcam.

After importing all the required libraries, load the model weights that we saved earlier after training. After importing the model weights we have imported a haar cascade file that is designed by open CV to detect the frontal face. After importing the haar cascade file, execute the code written to detect faces and classify the desired emotions. We have assigned the labels for different emotions like angry, happy, sad, surprise, neutral. As soon as you run the code a new window will pop up and your webcam will turn on.

It will then detect the face of the person, draw a bounding box over the detected person, and then convert the RGB image into grayscale & classify it in real-time.

Experimental Results

As an emotion classification task is a binary classification, the classification accuracies of valence are shown in figure 4. According to the expression in face, the emotion can be detected and provides the emotion of the detected image.

Output

Dataset Trained Performance

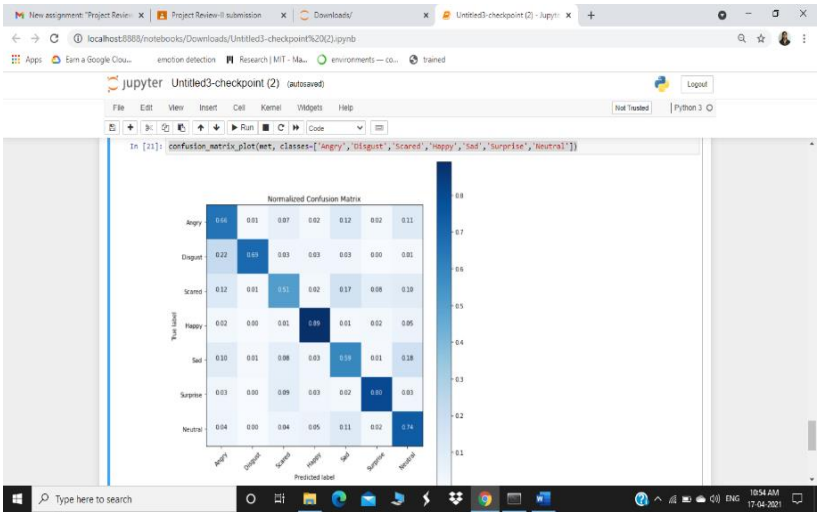

Figure 4. Person's correlation result between emotion and features.

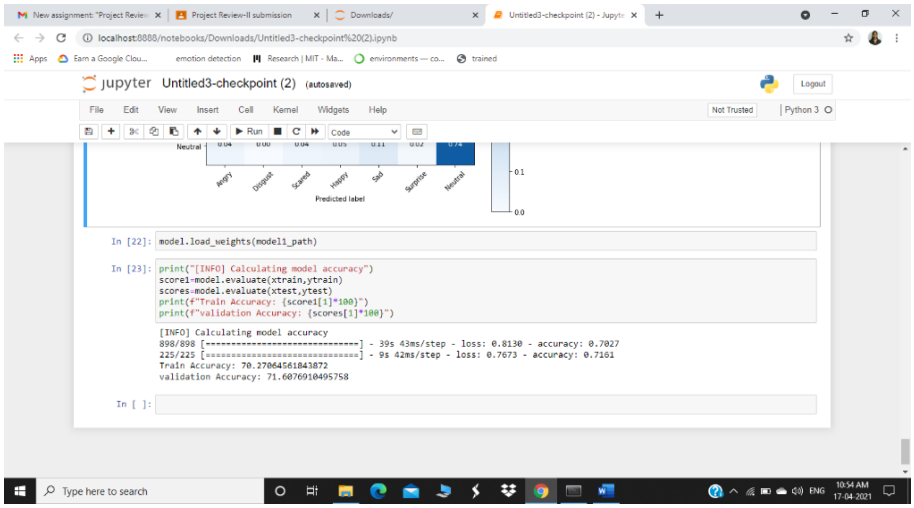

Figure 5. Result of executed Dataset (FER2013)

Face Detection Output

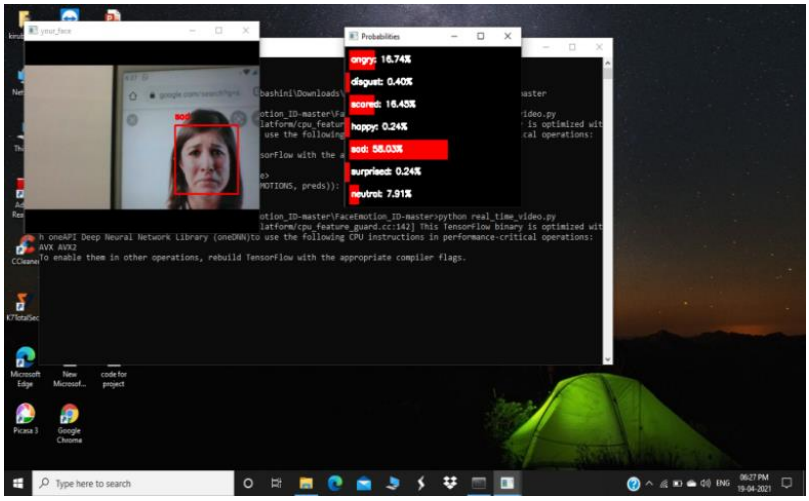

Figure 6. Detecting the sad face.

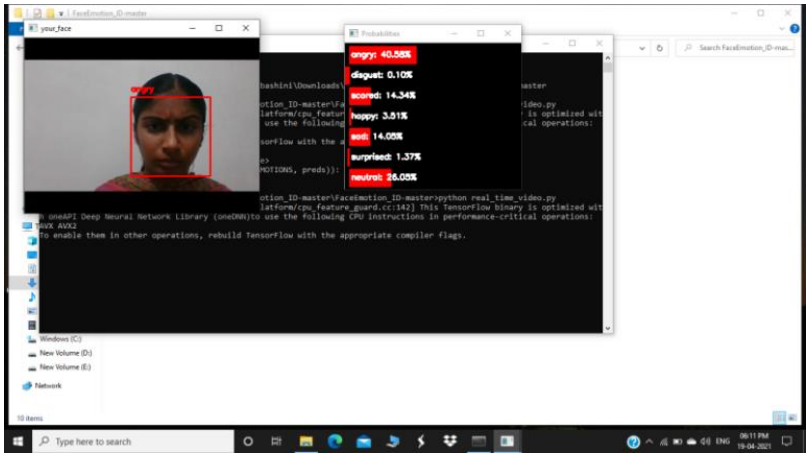

Figure 7.Detecting the angry face.

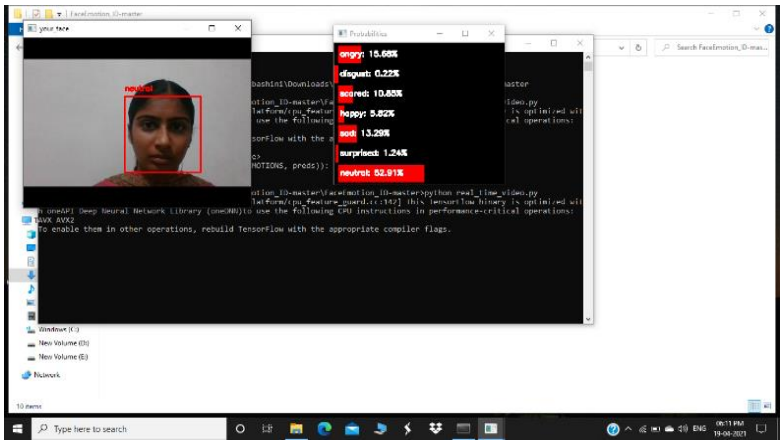

Figure 8.Detecting the neutral face.

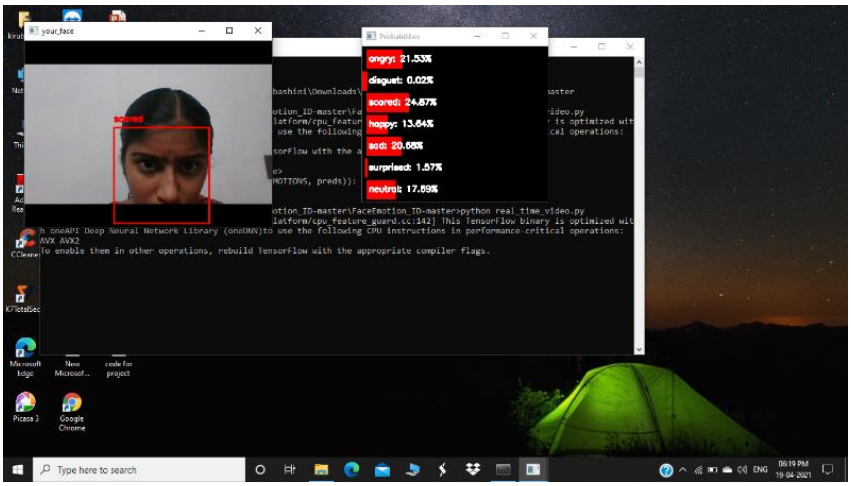

Figure 9.Detecting the scared face.

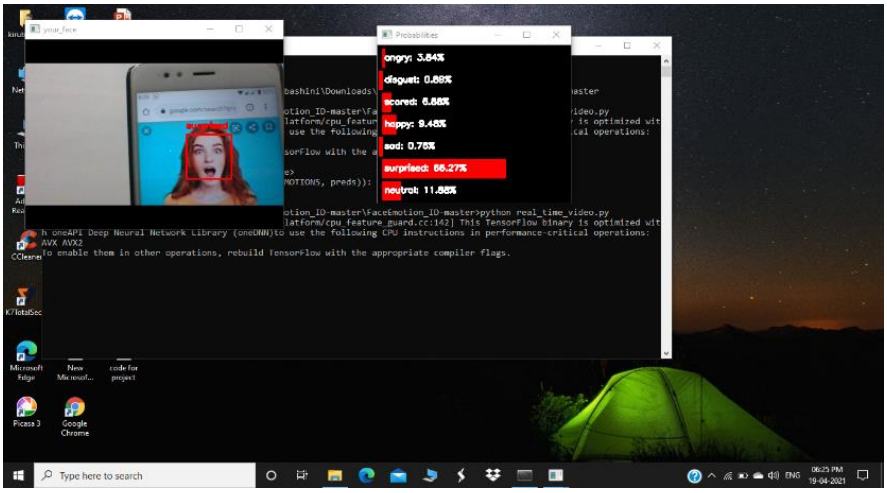

Figure 10.Detecting the surprise face.

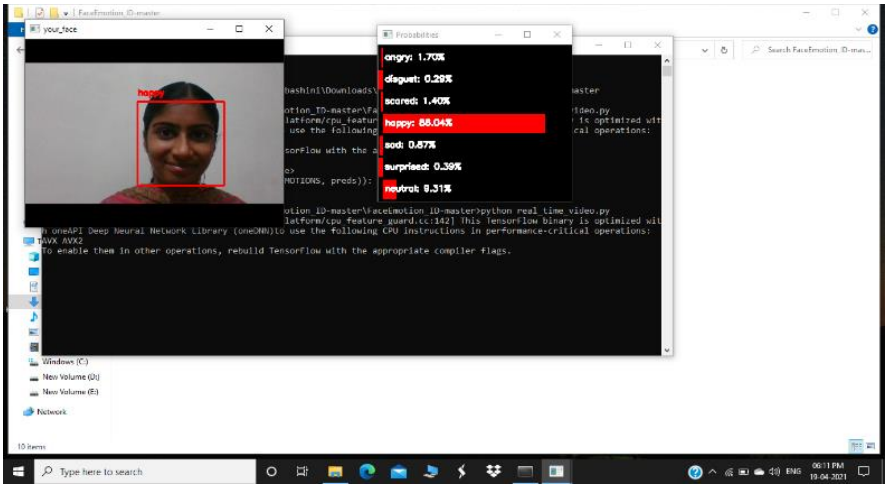

Figure 11.Detecting the happy face.

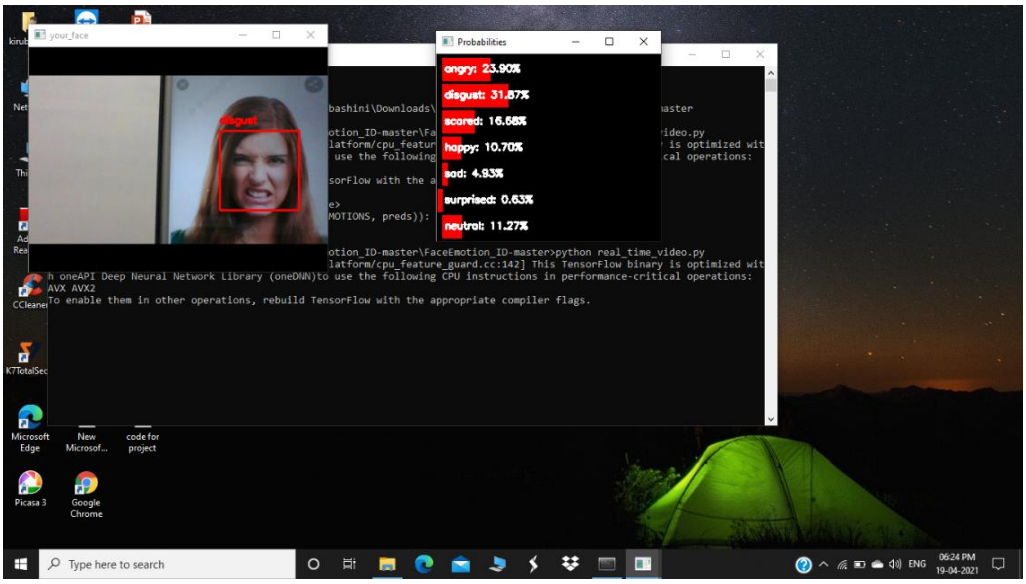

Figure 12. Detecting the disgust face.

RESULTS

The Xception CNN model is trained without preprocessing on FER2013 dataset. The mini-Xception model are more interpretable than the ones learned from ours sequential fully-CNN. This model shows the best result on FER2013 dataset, where the performance is 94.5%.

Figure 13. Performance rate comparison

	Top-1 accuracy	Top-5 accuracy
VGG-16	0.715	0.901
ResNet-152	0.770	0.933
Inception V3	0.782	0.941
Xception	**0.790**	**0.945**

Conclusion And Future Enhancement

We understood the whole pipeline on how to build an emotion detection model using pre-trained architecture like Xception CNN.

We initially defined network and trained it so that it is capable of classifying the correct emotion and then we made use of that trained model to classify the emotions in real-time.

We proposed an emotion recognition model using xception CNN for the short recognition interval. Using FER-2013 database gives better performance than the existing method.

The accuracy obtained for emotion expression and recognition using Xception is 95.60% and precision and recall rate is 93% and 90% respectively.

Future Work

In the future, the model will be trained and tested on broader sets of data to not only fit the training dataset or lab condition images but also to achieve better accuracy on wild emotion recognition.

In addition to this, the model is also expected to deliver the real life application in education to improve the learning and interactions in class.

REFERENCES

[1]. CH. Sadhvika, Gutta.Abigna, P. Srinivas Reddy, "Emotion Based Music Recommendation System", Vol 7, April 2020.

[2]. Subhash Rathod, Akshay Talekar , Swapnil Patil , Prajakta Gaikwad , Ruchita Gunjal, "Face Emotion Based Music Detection System", International Journal Of Advance Scientific Research And Engineering Trends Volume-4, 2019

[3]. Octavio Arriaga, Paul G, Matias Valdenegro, "Real-time Convolutional Neural Networks for Emotion and Gender Classification", Oct 2017

[4]. R. Patel, A. Vollal, P. B. Kadam, S. Yadav, R. M. Samant,"Moody Player: A Mood based Music Player", June 2018

[5]. Priyanka , T.R. Revanth Kumar, "Real-time Facial Expression Recognition System using Raspberry Pi", International Journal of Advanced Engineering Research and Science (IJAERS), 2017.

[6]. Gupte A, Naganarayanan A and Krishnan M "Emotion Based Music Player- XBeats" International Journal of Advanced Engineering Research and Science 3 236854

[7]. 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.

[8]. Hafeez Kabani , Sharik Khan , Omar Khan , Shabana Tadvi, "Emotion Based Music Player", International Journal of Engineering Research and General Science Volume 3, January-February, 2015

[9]. Setiawardhana, Nana Ramadijanti, Peni Rahayu "Facial Expressions Recognition Using Backpropagation Neural Network For Music Playlist Elections" Jurnal Ilmiah Kursor, Volume 6, Issue 3, 2012.

[10]. Gokul Krishnan K, Parthasarathy M, Sasidhar D, Venitha E, "Emotion Detection And Music Recommendation System Using Machine Learning", International Journal of Pure and Applied Mathematics.