Dr. B.R. Ambedkar National Institute of Technology Jalandhar

Bachelors of Technology
in
Computer Science & Engineering Department

Project Report on Emojify

Submitted by
Harshita Garg(19103046)
Kartik(19103054)

Mentored By -
Dr. Geeta Sikka

Department of Computer Science & Engineering
NATIONAL INSTITUTE OF TECHNOLOGY
Jalandhar, India
Academic year : 2021-22

CONTENTS

ACKNOWLEDGEMENT

It’s rightly said that there are hundreds of people working behind the curtain to make a Play successful. The final result of my project- Emojify required a lot of guidance and help from many people and we ,Group , are extremely fortunate to have gotten this throughout the journey of our project.

Whatever this project is today is only due to such supervision and assistance and We would not forget to thank them from the bottom of my heart.

We are extremely thankful to Professor A L Sangal, Head, Department of Computer Science & Engineering who has provided all his direct and indirect support.

We are extremely thankful to my mentor Dr.Geeta Sikka , Associate Professor, who believed in my idea and suggested new ideas wherever required. She has supported us fully in giving solutions to my problems.

We are extremely thankful to all my batchmates who helped me in developing the dataset and testing my model when I needed them. It could have been impossible to make this project without them.

We are extremely thankful to have received constant encouragement and guidance from all the Faculties of the Department of Computer Science & Engineering(CSE) who gave me their time and suggestions and thence helped us in successfully completing our project work. Also, We would like to extend our sincere esteem to all staff in the laboratory for their timely support.

Thank you
Group

ABSTRACT

Emotions play a vital role in every second of a person's life. From the moment of their birth till their last breath, every emotion can be observed (for example, anger, disgust, fear, happiness, sadness, surprise and neutral) emotions.

One way to indicate nonverbal cues is by sending emoji(e.g.,), which requires users to make a selection from large lists. Given the growing number of emojis, this can in-cur user frustration, and instead we propose Emojify ,where we use a user's facial emotional expression to filterout the relevant set of emoji by emotion category.

INTRODUCTION

Nonverbal behavior conveys affective and emotional information, to communicate ideas, manage interactions, and disambiguate meaning to improve the efficiency of conversations . One way to indicate nonverbal cues is by sending emoji, which are graphic icons managed by the Unicode Consortium that are identified by unicode characters and rendered according to a platform's font package.

We propose Emojify, a system and method to use users' facial emotional expressions as system input to filter emojis by emotional category and also detect the gender of the user and outputs the corresponding emoji. Suppose if User is a male model will output a emoji of a guy and if user is a Female it will output a emoji of a girl having same expression also tells the expression.

Emojis are basically graphical representation of emotion. Our software will be using these in order to make the working more interactive and fun.

Our software when run, is capable of detecting seven of the many human emotions which are visible on their faces. The output of our software will include two types of data as an output:-

1. Name of the emotion detected.
2. Respective emoji of the emotion detected.
3. Also predict the gender of the person and outputs the respective Emoji.

MOTIVATION

Face2Emoji is motivated by two findings from the literature:that a primary function of emojis is to express emotion, andthat most emojis used are face emojis. In research it was found that 60% (139/228) of their analyzed message by US participants were emoji used for expressing emotion. In an Instagram emoji study, faces accounted for 6 of the top 10 emojis used, providing further evidence that people frequently use emoji to express emotion. Furthermore, according to a 2015 SwiftKey report, faces accounted forclose to 60 percent of emoji use in their analysis of billionsof messages.

The study of non verbal communication via emotions originated with Darwin's claim that emotion expressions evolved in humans from prehuman non verbal displays.

Further-more, according to Ekman , there are seven basic emotions which have acquired a special status among the scientific community: Anger, Disgust, Fear, Happiness, Sadness, Happiness and Surprise.

LITERATURE REVIEW

Emotion Recognition System involves the process of acquiring the images, processing the images, detection of faces then extracting the expression features and mapped it onto corresponding emoji.

In recent years there has been tremendous research done on Face Emotion Recoginition .

With the help of a literature survey, it was realized that the basic steps in Face Emotion recognition are:-

- Face Detection
- Preprocessing of an Image
- Feature extraction
- Classification

Face Detection:-

Face detection process is to extract the face region from the background. Input images having various illumination conditions and complex backgrounds can be confusing to identify the expressions. It involves segmentation and extraction of facial features from uncontrolled background.

Pre Processing of an Image:-

The image pre-processing step is important in facial expression recognition. Raw images may be corrupted with noise and unwanted effects. If the test image has different illumination conditions to that of the training set images the facial expression recognition may fail.To minimize the environmental and illumination changes in the image the image is pre-processed. The motive of image preprocessing step is to obtain images of normalized intensity so that illumination and environmental changes will have no effect.

Feature Extraction:-

The process of translating the features from the input image into some set of features is termed as feature extraction.Feature extraction helps in reducing the huge amount of data to small data which increases the computational efficiency. A set of points are selected which represent the characteristics of human face. The face parts like mouth corners, eyebrows are identified using features such as pixel intensities.

The two main categories of feature extraction methods used in Facial Expression Recognition system are:

- Geometric Based
- Appearnce Based

1) Geometric Based:- In this, geometric feature vectors are formed considering geometric relationships such as angles and positions between different facial parts like nose, eyes, ears. The facial expression is determined by the movement of the facial points.

2)Appearance Based:-The features focus on facial appearances like bulges, wrinkles around mouth and eyes. This method applies image filters on the whole image to extract a feature vector.

Classification

The features extracted from previous block tries to classify the features based on the similarities between the feature data. Classification is done by supervised learning. The classifier has to be trained first and then the input test data is given to recognize the expression of the image

FEASIBILITY

Now, we will discuss the feasibility of this project.

1.TECHNICAL AND HARDWARE

- Python should be supported in the system
- Webcam should be there if prediction using live camera feed is required
- Internet connectivity is not required

2. ECONOMICAL FEASIBILITY

- This project doesn't require much cost in development.
- The only requirement is a camera supporting device.

Scope

It can be used when we require to insert an emoji instantly just by our expressions without needing to search for a particular one emoji It saves time and make the working of apps fun.

Can be used in product review, comments, online chatting. It can be used in Various chatting apps like whatsapp ,Instagram,SnapChat,Telegram E-learning,Monitoring etc.

Every stream in itself has numerous applications for this technology. Like, in the case of medicine, when AI has to deal with a patient of depression, it demands a very accurate use of emotion recognition system to come up with a required cure and process of improvement.

KEYWORD AND DEFINITIONS

Neural networks:

They are a set of algorithms which attempt to recognize Underlying relationships in batch of data using methods which mimic the Working of human brain.Like in brain we have cells here we have nodes and dendroites are the inputs .Each coonection in neuron transfers information to other neuron. And in neural network each layer transfers the input which is dimensionally reduced

Convolutional Neural Networks:

They are a type of artificial neural network which are mainly used to classify images. We cannot use ANN (Artificial Neural Networks) for image classification problems because two- dimensional images have to be translated to 1-dimensional vectors when we use ANN . So the classification task becomes more challenging as the number of training parameters are increased rapidly and requires great amount of storage and processing capability.

Convolution Layer:

In the convolution layer, we take a small window size [typically of length 5*5] that extends to the depth of the input matrix. The layer consists of learnable filters of window size.

We have different layers like

- Pooling layer
- Fully Connected Layer
- Activation Layer
- Final Output layer

DATASETS AND FEATURE

Datsets play a vital role in training ,testing and validation of facial expression recognition systems.Here we have used two types of datasets one is pose based and other one is spontaneous based. In Pose based datasets different users are asked to display different facial expression and we have captured those expressions. Spontaneous based are the natural expression which are directly input to the modal.

We have used Posed database for training and for validation and testing we have used Spontaneous as well as posed database.
Here we have Code two models one is gender detection and emotion detection . To train both models we have used two different datasets for training .

For Gender Detection we have used Kaggle UTK Face Datasets.

For Face emotion Analysis we have used FER 2013 from Kaggle.

Kaggle UTK Face

UTKFace is a large scale datasets having images over 23 k with long age span ranges from(0-116) . This datasets have images with annotations of age ,gender,ethinicity. These all annotations are stored in the name string of file separating with '_'. For Ex name of file is 100_0_0 xyz… here 100 denotes age and 0 denotes gender

The images in datasets is of wide variety of expressions ,illumination,resolution etc. Provide a wide range to test the model. And datasets is already clean .

We have split this huge dataset into train and test for detecting gender of a person. And we have acheived the accuracy arround 88% in gender model.

Facial Emotion Recoginition

The datasets contain images around 35k captured at differnt angle . All the images are of 48X48 pixels gray scale images of different facial expressions.The datasets is already divided into test and train .And in train and test datasets are divided into 7 folders of different expressions Anger, Happy, Sad, Fear, Neutral,Disgust, Surprise.

METHODOLOGIES

1. Recognizing emotion using CNN

a. Build a CNN architecture

First, we import libraries which are required for the model and then the training and validation generators are set up. To train the model, we rescale the images.

b. The training and validation generators are initialized

c. Finally, the CNN architecture is built

1. In this there are two 2-D convolutional layers with kernel size of 3*3 and the input are the images of 48*48 .
2. Next, max pooling layer of 2*2 is added and then a dropout of 0.2.
3. Then again a 2-d convolutional layer and a max pooling layer is added .
4. This step is repeated again and after this a dropout of 0.25 is added
5. Finally, we flatten the outputs and pass it through dense which has 1024 outputs and the we add a dropout of 0.5.
6. The activation function used in all above layers is ReLu (Rectified Linear Unit)
7. Finally the dense layer, which maps to the 7 outputs is there and the classifier used is the softmax classifier.

2. Model Training on the FER-2013 dataset

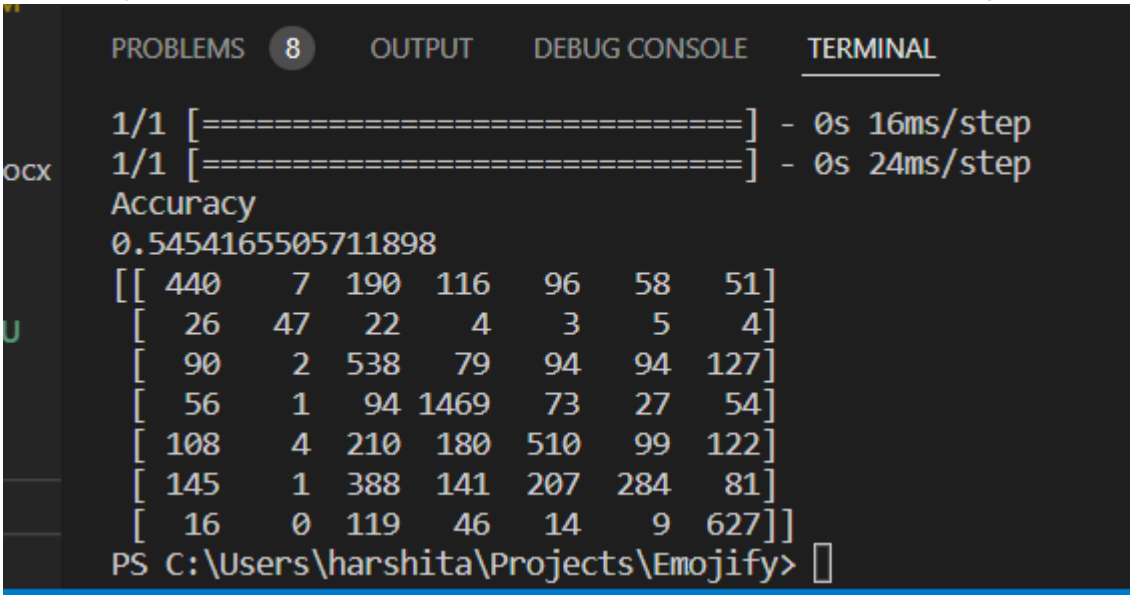
We are going to train our network on all the images that we have, which is the FER 2013 dataset, and then we will save the weights in the model for future predictions.

To detect the bounding boxes of a face in camera, we use OpenCV Harrascade XML and then predict the emotions. Adam optimizer is used as it is generally better than other optimization algorithms

3. Predicting emotions

Finally, we predict the emotion either through web cam or a given Image .We make an emojis folder and save emojis for each of the datasets, Seven emotions.

Accuracy we have achieved is 54% and the confusion Matrix We got



Gender Detection Model

1. Convolutional Neural Network used for Gender detection:

First, import the required libraries. The input to the first layer are images of dimensions 100*100. First layer is a 2-d convolutional layer with 3*3 kernel size and a stride of

L2 regularizer is used in kernel. ReLu is used as activation function

Dropout of 0.1 is further performed and then a 2-d max pooling is provide ReLu is used as activation function.

Second layer is again a 2-d convolutional layer with 1 as its stride and Using L2 regularizer in kernel and further has a 2-d max pooling layer . ReLu is used as activation function

This configuration is again repeated with next 2 layers

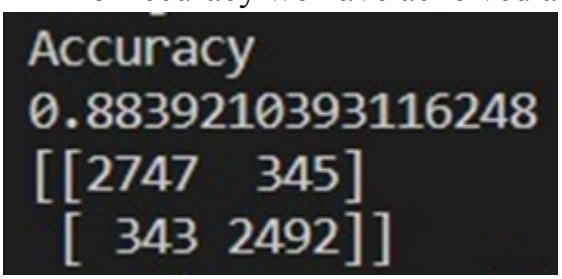
Finally, the output is flattened and passed through a dense layer and a Dropout of 0.2 is done. Sigmoid function is used as activation function in this layer.

2. Model Training on the UTK dataset

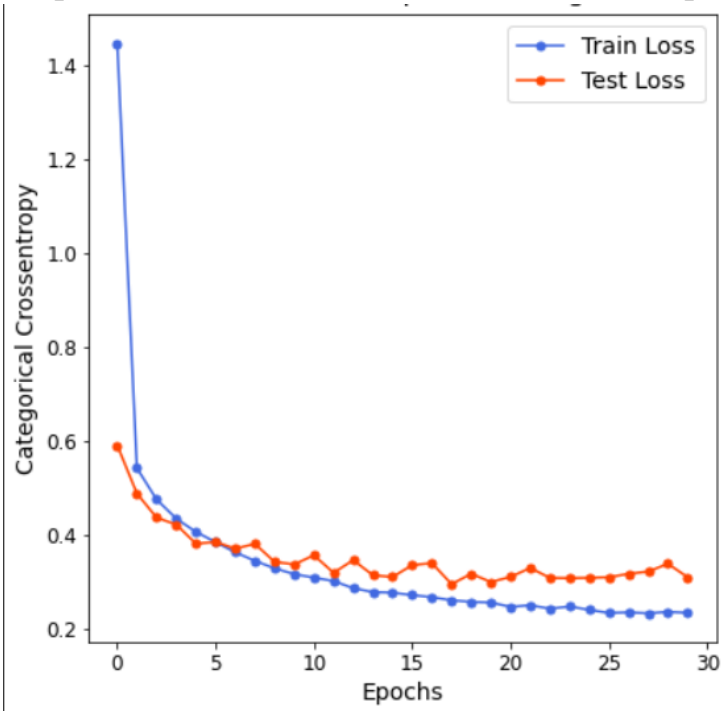
We are going to train our network on all the images that we have, which is the UTK dataset, and then we will save the weights in the model for future predictions. Adam optimizer is use as it is generally better than other optimization algorithms

3. Accuracy Aand the Confusion Matrix

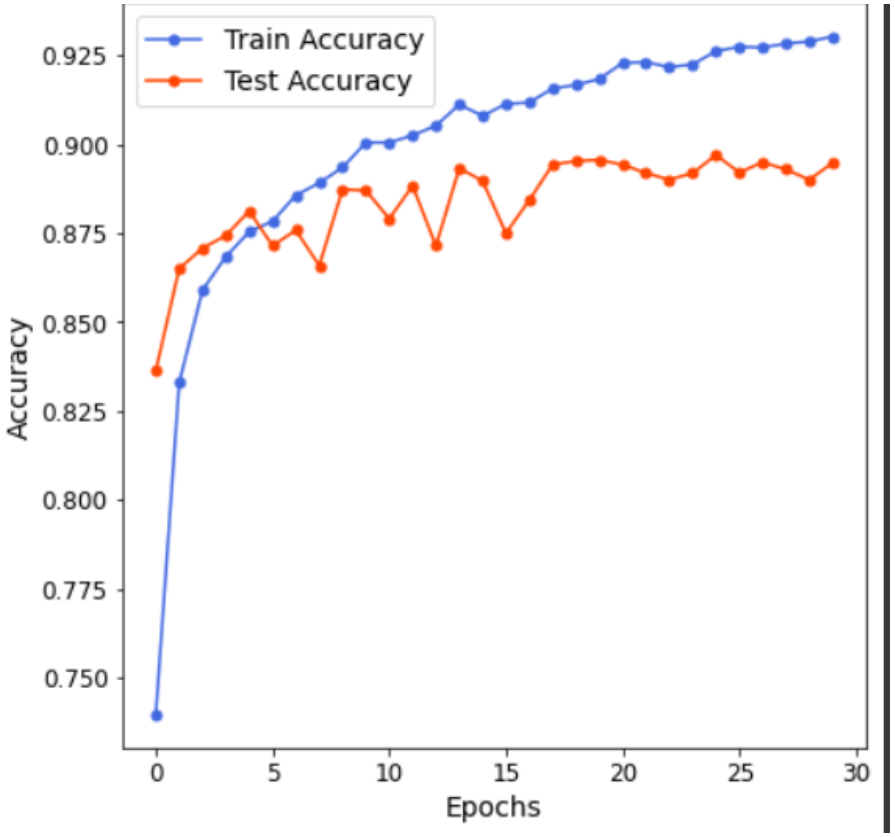
The Accuracy we have acheived and the confusion matrix:-



Graph of Train Loss Vs test loss at different Epochs

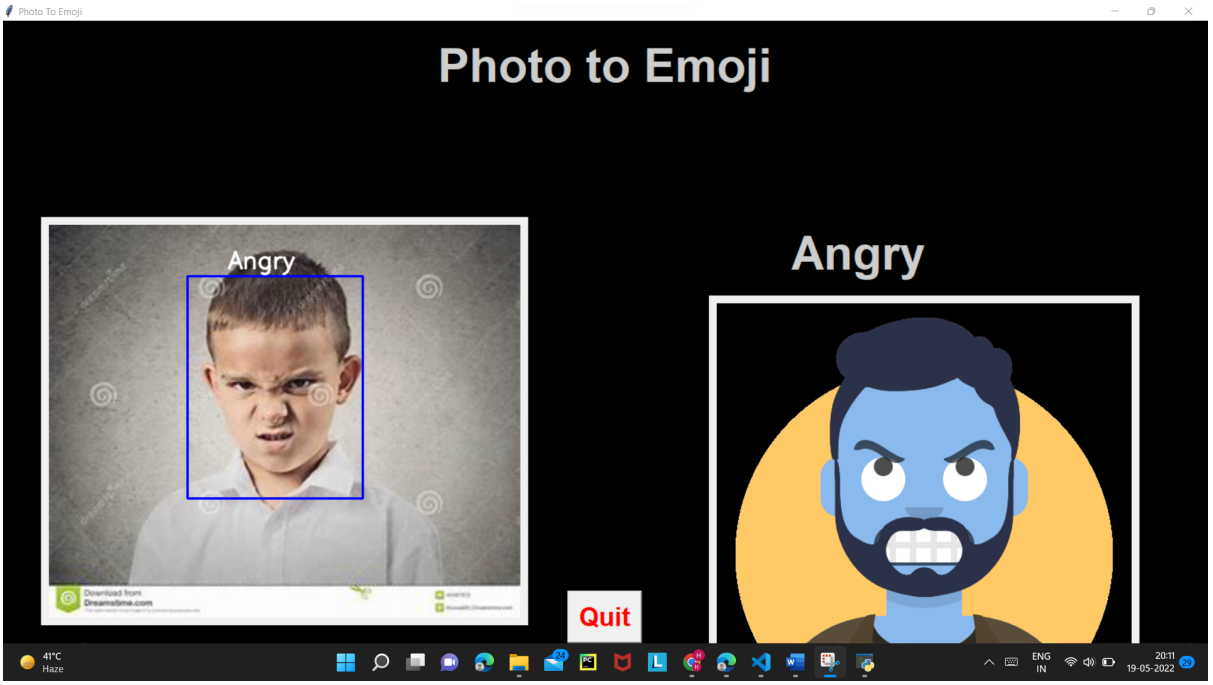


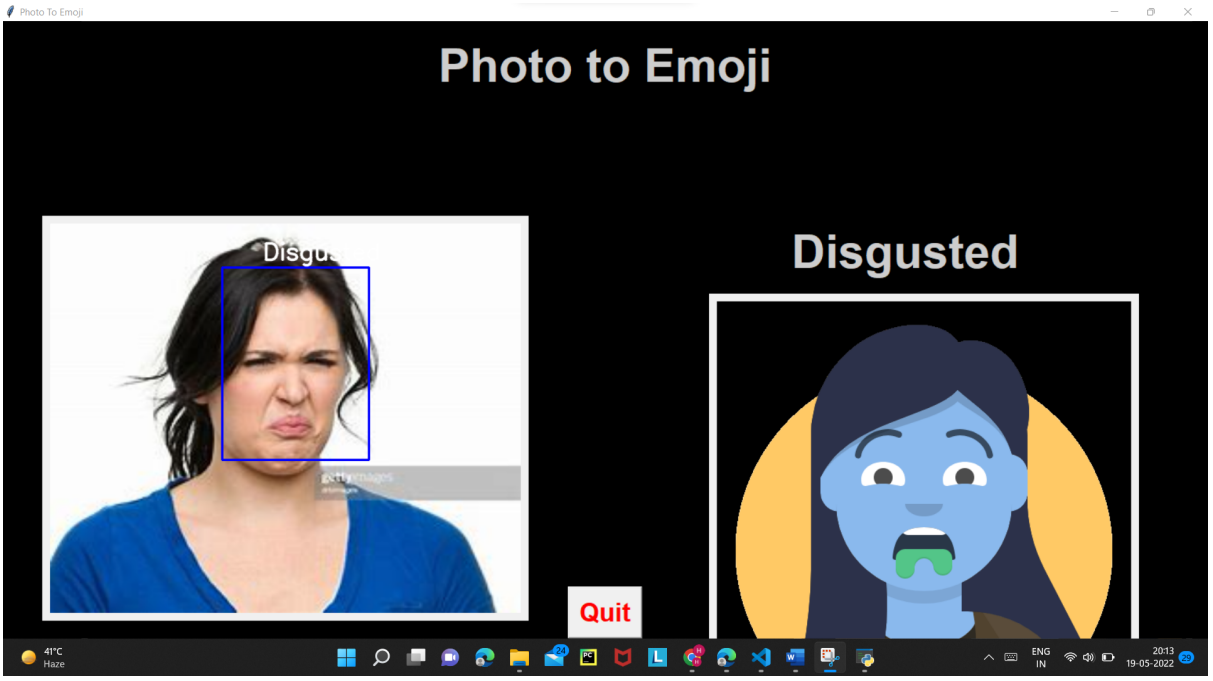
Test Accuracy vs Train Accuracy at different Epochs



RESULTS

Model is successfully able to distinguish between gender and the different facial expression.
We have also added a quit button to move out of the window





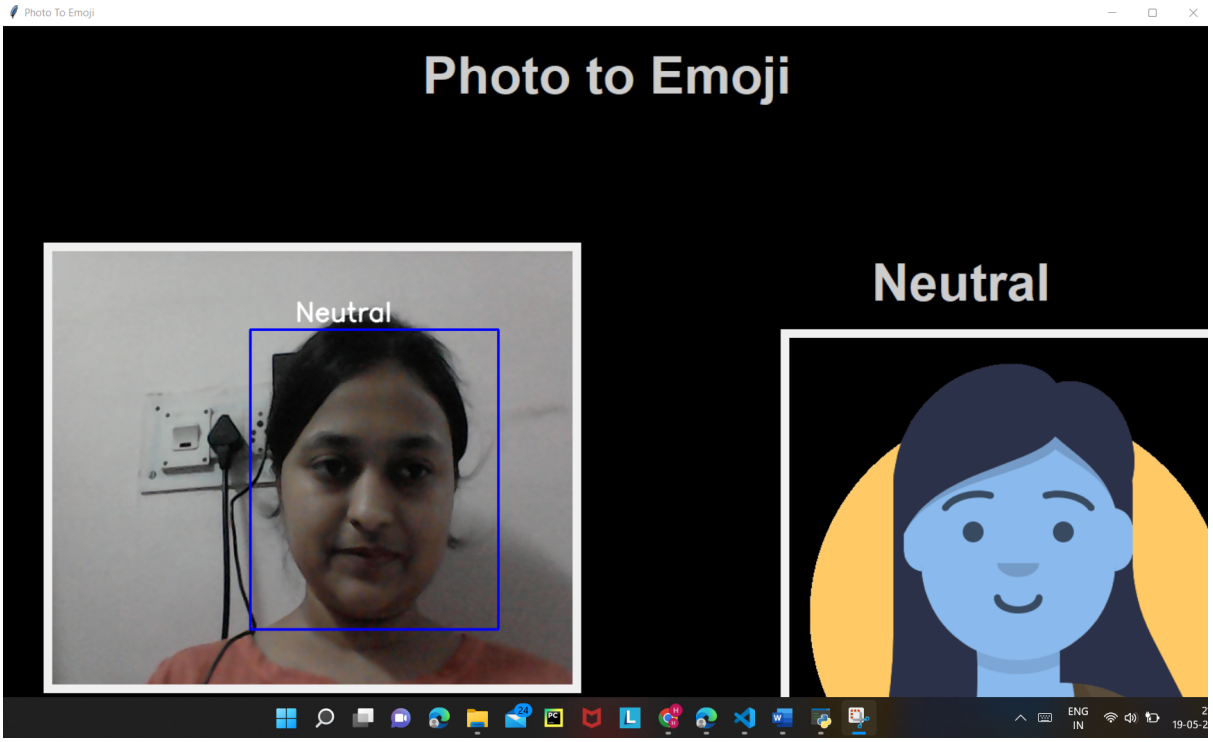


Photo To Emoji
Photo to Emoji
Neutral
Neutral

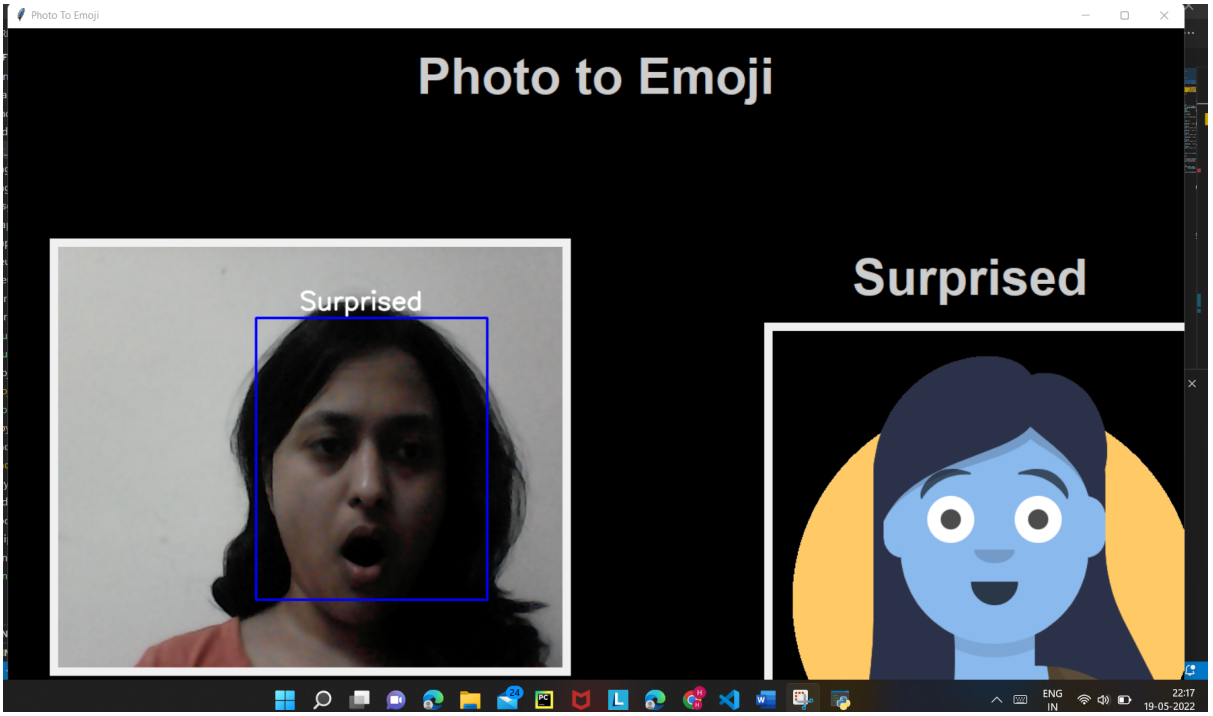


Photo To Emoji
Photo to Emoji
Surprised
Surprised

FUTURE SCOPE AND IMPROVEMENTS

1. The emotion model now we have has an accuracy of just 64.7% and gender model has accuracy of 88.7% which is not a very high accuracy we can work on the accuracy of the model.
2. Here are some expressions in which model could’nt differentiate correctly like Surprised and Fearfull
3. Sometimes Model could’nt able to correct the gender correctly there’s need of some improvement
4. At time of Analysis on Real video it takes a little bit time to predict the expression correctly.
5. Face should be held high in Camera and should be clearly visible

CONCLUSIONS

So concluding this report here in this project we have developed a CNN based emotion detection model which works in both real-time and post based videos and pictures using two models which are Gender Detection and Emotion Detection. We are really happy with the outcome of model and hoping to make more improvements with added functionalities with improved accuracy in future.

REFERENCES

(PDF) Face2Emoji: Using Facial Emotional Expressions to Filter Emojis (researchgate.net)

26647690.pdf (stanford.edu)

https://www.ijercse.com/specissue/april-2018/26.pdf

https://towardsdatascience.com/building-your-own-artificial-neural-network-from-scratch-on-churn-modeling-dataset-using-keras-in-690782f7d051

https://opencv.org/

https://en.wikipedia.org/wiki/TensorFlow

https://en.wikipedia.org/wiki/Convolutional_neural_network