Capstone Project

Live Class Monitoring System[Face Emotion Recognition]

Apoorva KR

apoorvargowda1@gmail.com

Points to discuss

- Introduction
- Facial Emotion Recognition
- Problem Statement
- Data Summary
- Dependencies
- Model Creation
- Real time prediction
- deployment

INTRODUCTION

The Indian education landscape has been undergoing rapid changes for the past 10 years owing to the advancement of web-based learning services, specifically, eLearning platforms. In a physical classroom during a lecturing teacher can see the faces and assess the emotion of the class and tune their lecture accordingly, whether he is going fast or slow. He can identify students who need special attention. Digital classrooms are conducted via video telephony software program (exZoom) where it's not possible for medium scale class (25-50) to see all students and access the mood. Because of this drawback, students are not focusing on content due to lack of surveillance. • While digital platforms have limitations in terms of physical surveillance but it comes with the power of data and machines which can work for you. It provides data in the form of video, audio, and texts which can be analysed using deep learning algorithms. Deep learning backed system not only solves the surveillance issue, but it also removes the human bias from the system, and all information is no longer in the teacher's brain rather translated in numbers that can be analysed and tracked.

Facial Emotion Recognition

- Facial Emotion Recognition (FER) is the technology that analyses facial expressions from both static images and videos in order to reveal information on one's emotional state
- Facial expression recognition software is a technology which uses biometric markers to detect emotions in human faces.
- More precisely, this technology is a sentiment analysis tool and is able to automatically detect the seven basic or universal expressions: happiness, sadness, anger, neutral, surprise, fear, and disgust.
- Facial expressions and other gestures convey nonverbal communication cues that play an important role in interpersonal relations.
- Therefore, facial expression recognition, because it extracts and analyzes information from an image or video feed, it is able to deliver unfiltered, unbiased emotional responses as data.

Problem statement

We tackled the problem of recognizing the emotion of a person from an image of their facial expression. First, we built models capable of recognizing seven emotions (happy, sad, angry, afraid, surprise, disgust, and neutral). Given static, cropped headshots, our model would output a probability distribution over emotions of the pictured individual. Next, we transferred the skills learned on static datasets to implement a real-time emotion classifier. Using a webcam video feed, we built a system to continuously detect faces, extract, crop, and grayscale the face region, and classify the emotion of the person.

Data description

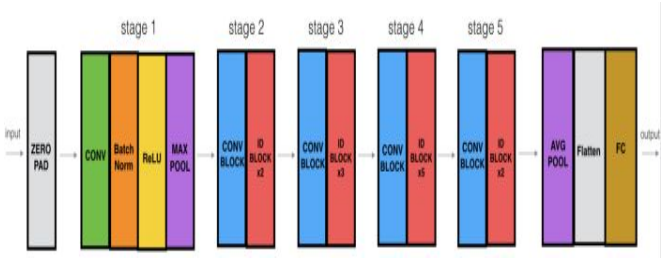
The FER-2013 dataset consists of 28,000 labeled images in the training set, 3,500 labeled images in the development set, and 3,500 images in the test set. Each image in FER-2013 is labeled as one of seven emotions: happy, sad, angry, afraid, surprise, disgust, and neutral, with happy being the most prevalent emotion, providing a baseline for random guessing of 24.4%. The images in FER-2013 consist of both posed and unposed headshots, which are in grayscale and 48x48 pixels. The FER-2013 dataset was created by gathering the results of a Google image search of each emotion and synonyms of the emotions.

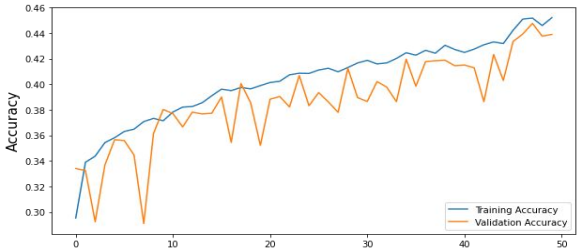
Dependencies

1. Python 3.8.8
2. Tensorflow 2.0
3. Streamlit
4. Streamlit Webrtc
5. OpenCV

ResNet-50

- ResNet-50 is a convolutional neural network that is 50 layers deep. You can load a pre-trained version of the network trained on more than a million images from the ImageNet database . The pretrained network can classify images into 1000 object categories, such as keyboard, mouse, pencil, and many animals
- training accuracy and validation accuracy is very low so our pretrained model resnet50 is not performing well

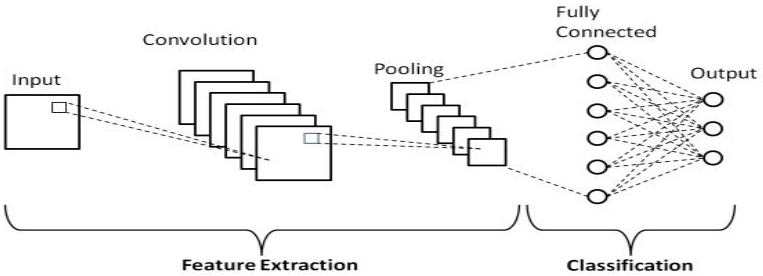




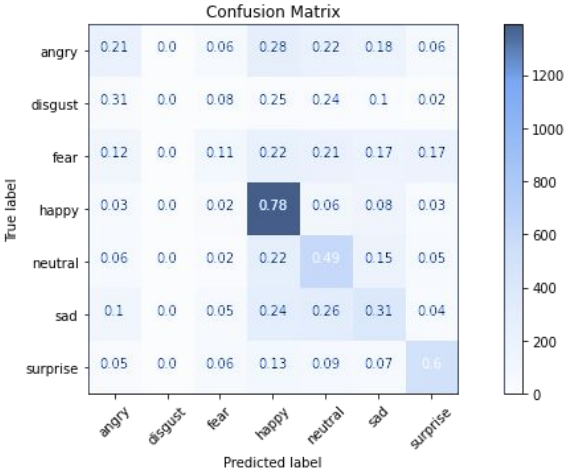
CNN model

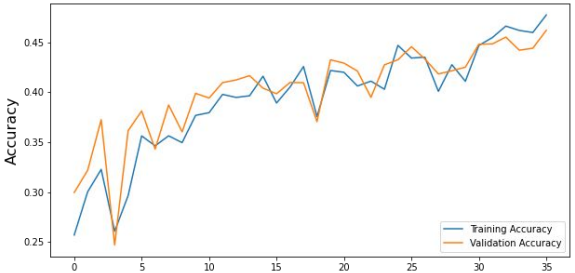
METHOD1:without dropout and batch normalization

- use batch_size of 32 and after fitting our data to our image generator, data will be generated in the batch size of 32



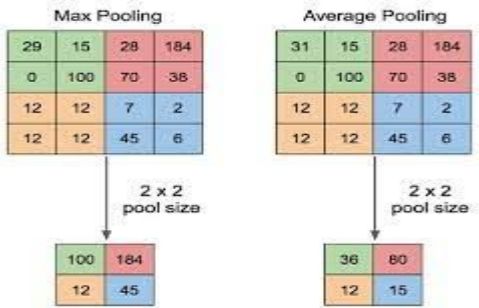
- the model did not perform well on the validation data and training data. Thus, we will move ahead and make amendments to the parameters.

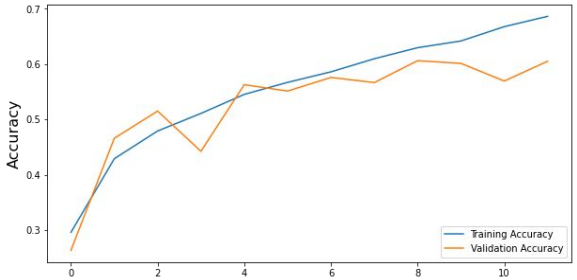




METHOD2 :CNN model with dropout and batch normalization

- Designing the CNN model for, emotion detection .creating blocks using Conv2D layer,Batch-Normalization, Max-Pooling2D, Dropout, Flatten, and then stacking them together and at the end-use Dense Layer for output





1.epochs = 100

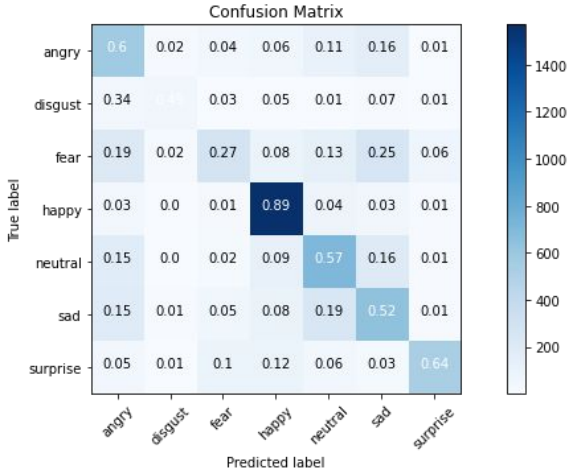
2. batch_size = 32

3. learning_rate = 0.001

- We trained the neural network and we achieved the highest training accuracy of 74.46%. After using test data to check how well our model generalize, we score an astounding 62.06% on the test set.

Confusion Matrix

- The confusion matrix clearly shows that our model is doing good job on the class happy and suprise but it's performance is low on other classes. Maybe because it confuses ,Facial expression depends on individual as well. Some person's neutral face looks like sad.



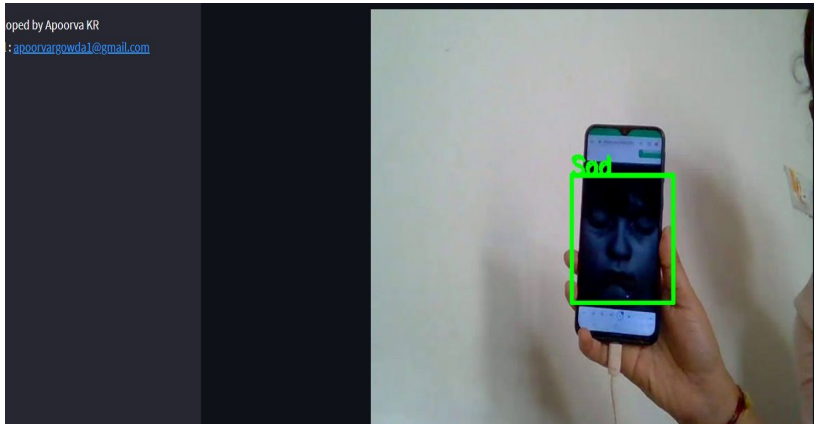
Save the Model

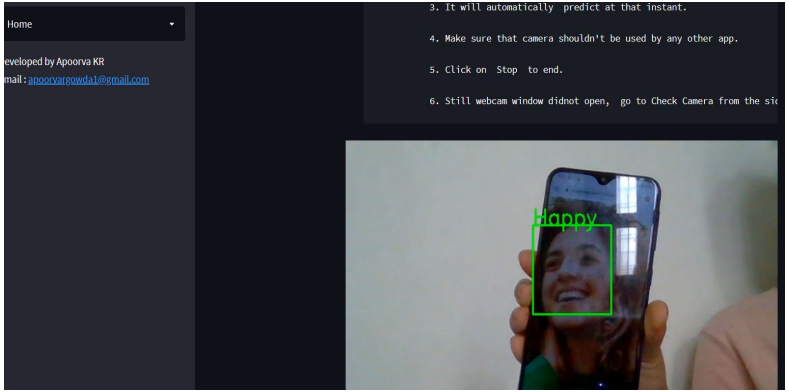
- Saving our model’s architecture into JSON and model’s weight into .h5.
- Download the saved model and weights in a directory.
- test our model in real-time using face detection.
- Loading Haar-Cascade for Face Detection We are using Haar-cascade for the detection position of faces and after getting position we will crop the faces.
- Use OpenCV to read frames and for image processing.
- Adding an overlay on the output frame and displaying the prediction with confidence gives a better look.

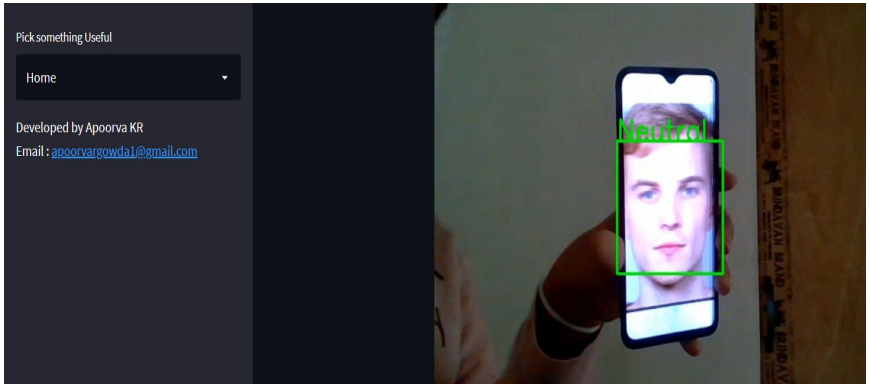
Deployment of streamlit webApp in Heroku and Streamlit

- We have created front-end using streamlit-webrtc which helped to deal with real time video streams.Image captured from the webcam is sent to Video Transformer function to detect the emotion .Then this model was deployed on heroku platform
- Streamlit is an open-source app framework for Machine Learning and Data Science teams. Create beautiful web apps in minutes.
- Heroku is a container-based cloud Platform as a Service (PaaS). Developers use Heroku to deploy, manage, and scale modern apps.
- Deployment Link for Streamlit Share:-
  https://share.streamlit.io/apoorvakr12695/face-emotion-recognition-/main/app.py
- Deployment Link for heroku :-
  https://face-recognit-apoorva.herokuapp.com/

Various prediction Images from theWebApp







conclusion

- Training accuracy and validation accuracy is very low so our pretrained model resnet50 is not performing well
- We trained the neural network and we achieved the highest training accuracy of 74.46%. After using test data to check how well our model generalize, we score an astounding 62.06% on the test set.
- Successfully deployed the streamlit webapp on heroku and streamlit share that runs on a web server
- Our model can successfully detect face and predict emotion on live feed as well as on image