Capstone Project -5
Face Emotion Recognition

By

Govind Wakure

Content

Problem Statement

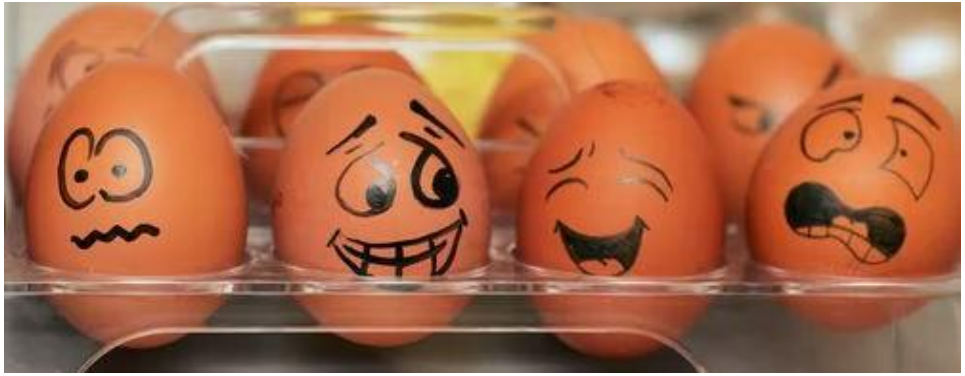
Key Point:

One of the many challenges in web based learning is how to ensure quality learning for students. Digital platforms might overpower physical classrooms in terms of content quality but when it comes to understanding whether students are able to grasp the content in a live class scenario is yet an open-end challenge.

Introduction

What is FER?

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.



Data Summary

FER2013 (Facial Expression Recognition 2013) Dataset

The data consists of 35887 grayscale images of faces at a resolution of 48x48 pixels. The faces have been automatically registered such that they are more or less centred in each image and take up around the same amount of area.

It has seven categories based on the emotion expressed in the facial expression (0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral). There are 28,709 examples in the training set and 3,589 examples in the public and private test sets.

Approach

Exploratory Data Analysis

↓

Data Preprocessing

↓

Image Data Augmentation

↓

Model Training

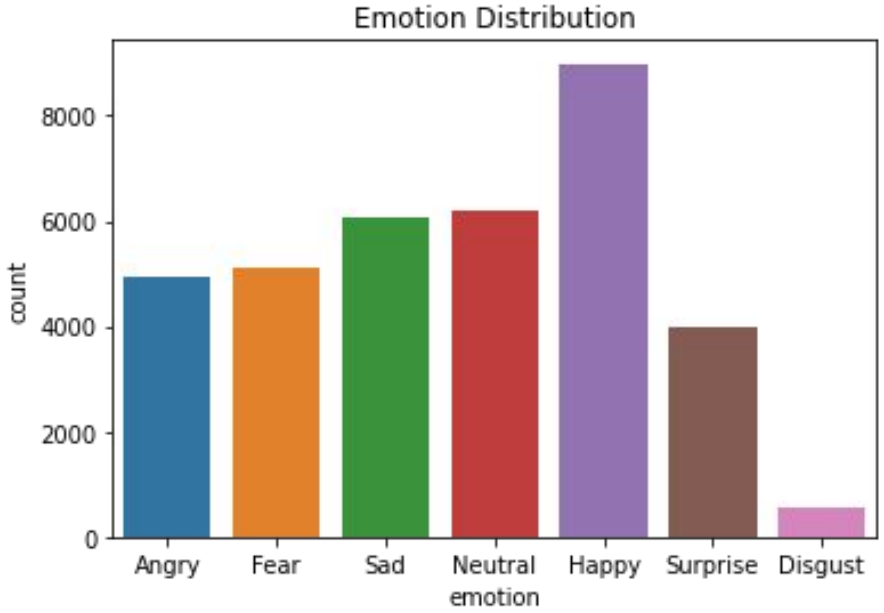
↓

Model Performance and Evaluation

↓

Model Deployment

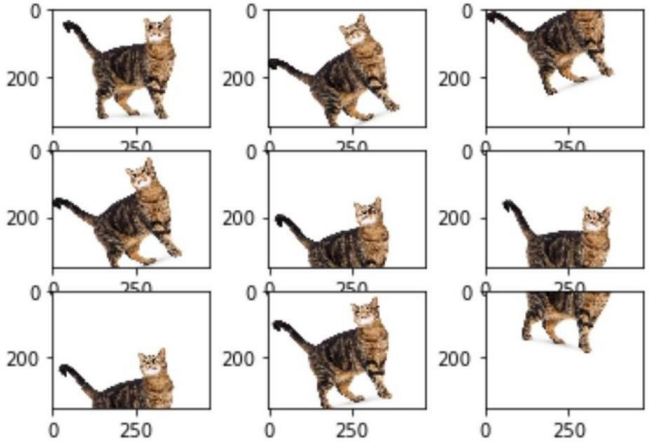
Exploratory Data Analysis



0	Angry	4953
1	Disgust	547
2	Fear	5121
3	Happy	8989
4	Sad	6077
5	Surprise	4002
6	Neutral	6198

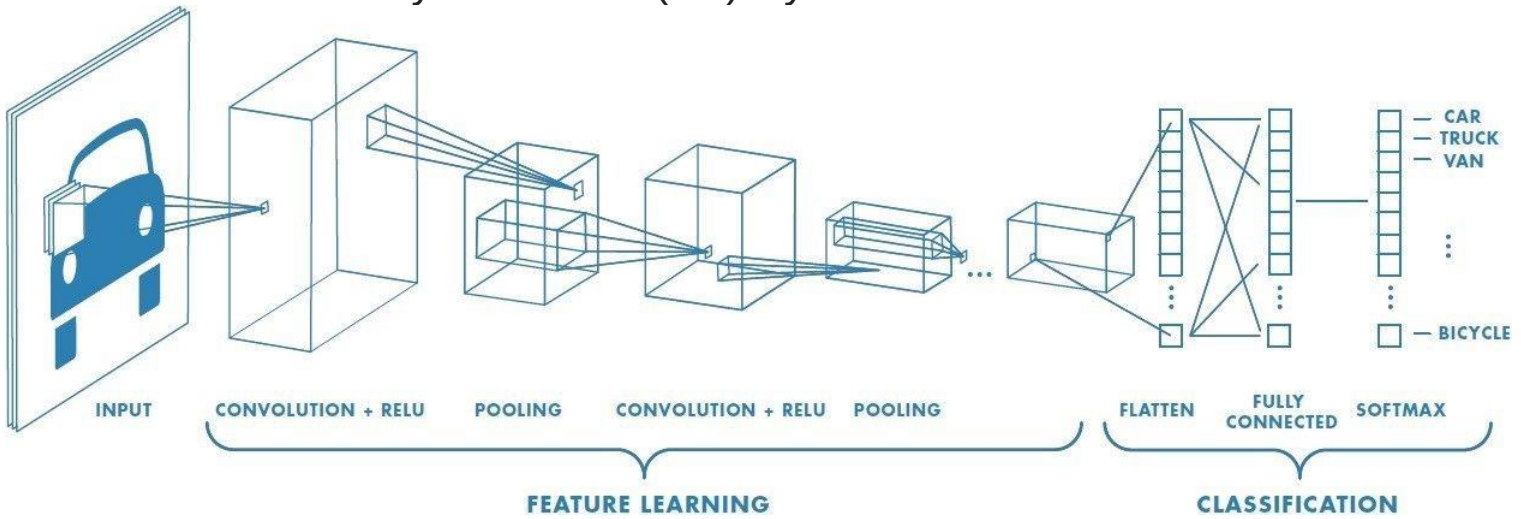
Data Augmentation

Image augmentation is a method of altering original images by applying various transformations to them, resulting in many altered copies of the same image. Depending on the augmentation techniques used, such as shifting, rotating, flipping, and so on, each copy is unique in certain ways.

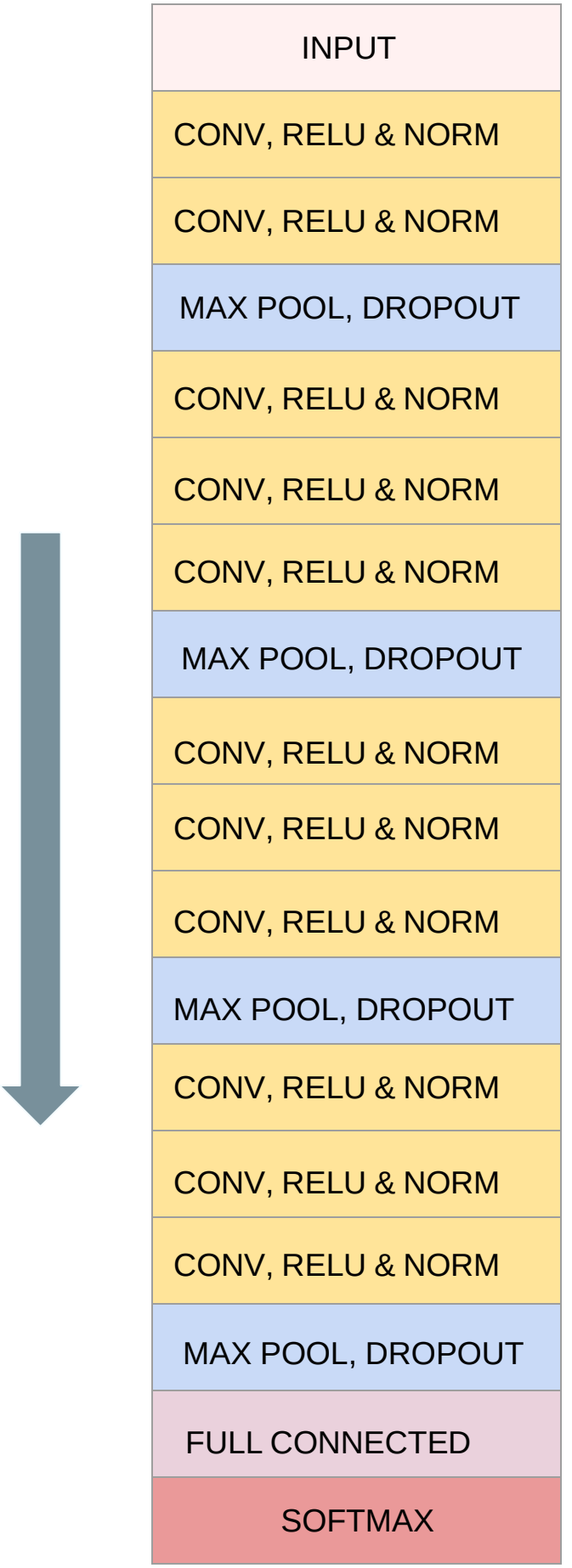


Convolutional Neural Networks

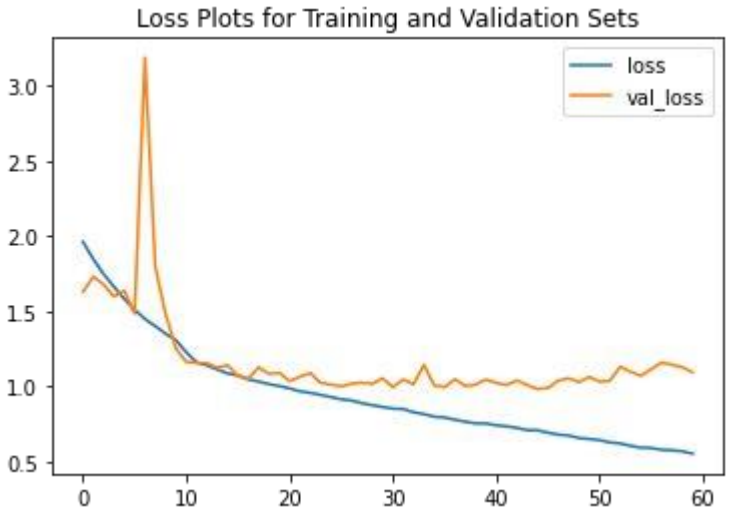
- Convolutional neural networks use principles from linear algebra, notably matrix multiplication, to discover patterns inside an image, making them more scalable for image classification and object recognition tasks.
- They have three main types of layers, which are:
  - Convolutional layer
  - Pooling layer
  - Fully-connected (FC) layer

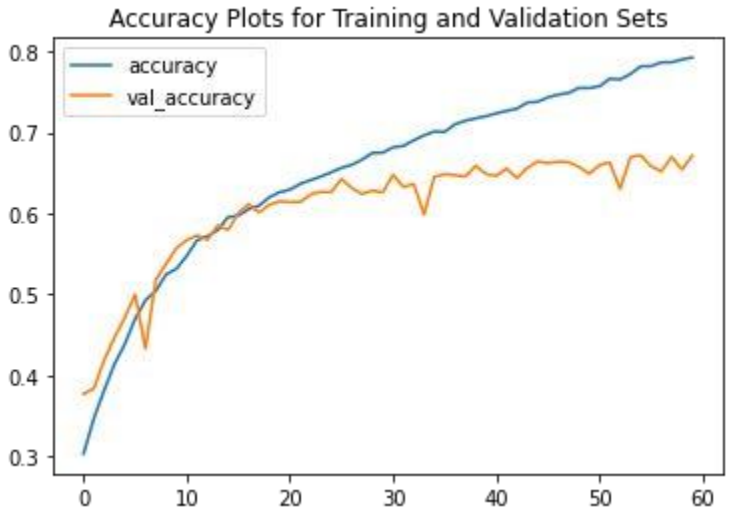


Model Architecture



Model Performance and Evaluation





The best model weights were restored, which gave the following results for training and validation datasets.
loss: 0.7075 - accuracy: 0.7381 - val_loss: 0.9825 - val_accuracy: 0.6643

Model Performance and Evaluation

	precision	recall	f1-score	support
0	0.64	0.56	0.60	491
1	0.83	0.62	0.71	55
2	0.53	0.49	0.51	528
3	0.90	0.86	0.88	879
4	0.51	0.58	0.54	594
5	0.78	0.75	0.77	416
6	0.62	0.72	0.67	626
accuracy			0.68	3589
macro avg	0.69	0.65	0.67	3589
weighted avg	0.68	0.68	0.68	3589

Here are the results of the model on test dataset:

It gave an overall accuracy of 0.68 which is close to the results for validation sets, it means the model is working similarly for the unseen data as well.

Real Time Emotion Detection

- With the help of OpenCV-Python, a real time emotion detection was conducted locally.
- Next a web app was created using Streamlit framework for the same.

Web App and Deployment

Heroku

https://face-emo-recog1.herokuapp.com/

Challenges:

Heroku provides free services to deploy models on it’s cloud platform but reducing the size of the slug generated by all the dependencies and model itself was a challenge. Since the slug size is large than soft limit provided by Heroku which is 300 MB, the booting time is really slow.

AWS EC2 Instance

http://54.84.63.103:8501/

Amazon Web Services is not a free platform but provides free tier services for twelve months, the model was deployed on a Ubuntu virtual machine instance. The computing is really fast.

Conclusion

- Face Emotion Recognition is a crucial application of deep learning algorithms which can be extended to every industry.

- Future work in relation to this project can include tracking and analyzing the emotions of the students. For example If a student is continuously predicted to be sad for a class of an hour, he/she could be flagged and a report of all the students could be generated at the end of the lecture for better analysis and further customized lesson plans.

- Another important point to conclude is CNN models could achieve extraordinary results if appropriate and good amount of training data is provided. For example for this particular case, the training data should include images of students while studying.

- The model gave 71% accuracy for training data and 66% for validation data. On the other hand it gave 68% accuracy for the test.

References

- Analytics Vidhya
- GeeksforGeeks
- Towards Data Science
- Built in Data Science
- Analytics India Magazine
- https://edps.europa.eu/system/files/2021-05/21-05-26_techdispatch-facial-emotion-recognition_ref_en.pdf