Capstone Project
Facial Emotion Recognition

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.

Problem Statement

- We will solve the before-mentioned challenge by applying deep learning algorithms to live video data. The solution to this problem is by recognizing facial emotions.

What is Facial Emotion Recognition?

- Facial emotion recognition is the process of detecting human emotions from facial expressions. The human brain recognizes emotions automatically, and software has now been developed that can recognize emotions as well. This technology is becoming more accurate all the time, and will eventually be able to read emotions as well as our brains do.

- AI can detect emotions by learning what each facial expression means and applying that knowledge to the new information presented to it. Emotional artificial intelligence, or emotion AI, is a technology that is capable of reading, imitating, interpreting, and responding to human facial expressions and emotions.

Other Applications of Facial Emotion Recognition

- **Market Research:** Companies have traditionally done market research by conducting surveys to find out about what consumers want and need. This method however, assumes that the preferences stated are correct and reflect future actions. But this is not always the case. Another popular approach in market research is to employ behavioral methods where user's reactions are observed, while interacting with a brand or a product. Although effective, such techniques can quickly become very labor intensive as the sample size increases. In such circumstances, facial expression recognition technology can save the day by allowing companies to conduct market research and measure moment-by-moment facial expressions of emotions automatically, making it easy aggregate the results

- **VideoGame Testing:** Facial expression recognition can also be used in the video game testing phase. In this phase, usually a focus group of users is asked to play a game for a given amount of time and their behavior and emotions are monitored. By using facial expression recognition, game developers can gain insights and draw conclusions about the emotions experienced during game play and incorporate that feedback in the making of the final product. Facial expression analysis is a practical means of going beyond the typical survey approach. It is a way of appreciating what the user is experiencing, all while getting feedback. When feedback is taken in this format, it becomes genuinely non-intrusive when it comes to user experience

About our DataSet

- 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.

Problems in FER 2013

- **Imbalance Dataset**
- **Intraclass Variation**
- **Contrast**
- **Eye-Glasses**
- **Outlier/Null Values**

Dependencies

- Python
- Tensorflow
- Keras
- Streamlit
- Streamlit-Webrtc
- Opencv

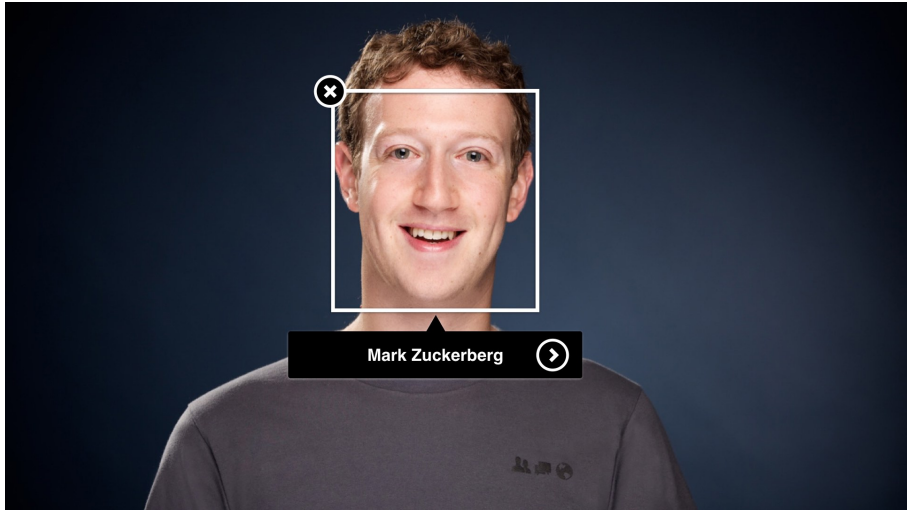
Streamlit

Model Training

Using a Pre-trained Model

- In past, there have been many research in feild of computer vision and so topic of emotion recognition was very popular. Some researchers published their own model with various facial features detection techniques which can be used. Since our aim is built a proper live class monitoring system it is necessary to check how the model works.
- We will use DeepFace. DeepFace is a deep learning facial recognition system created by a research group at Facebook. It identifies human faces in digital images. The program employs a nine-layer neural network with over 120 million connection weights and was trained on four million images uploaded by Facebook users.

What is DeepFace?

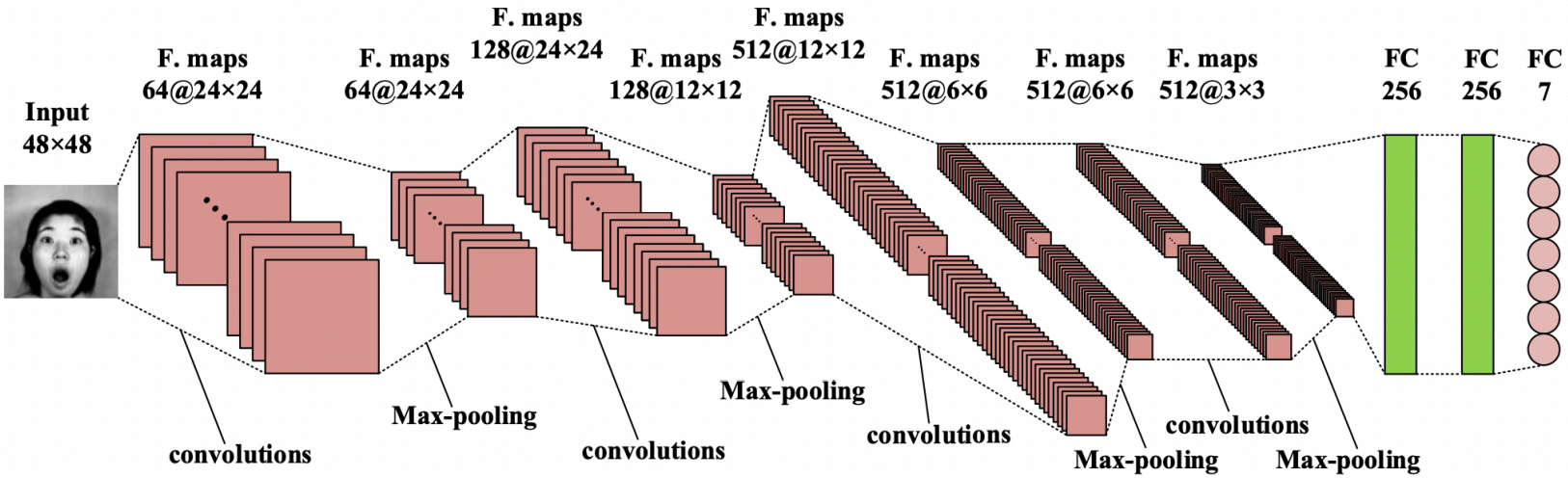


Deep-face is a lightweight face recognition and facial attribute analysis (age, gender, emotion and race) framework for python. The library is mainly based on Keras and TensorFlow.

Comparison of Different Pre-trained Models

Architecture	Parameters	Accuracy	Run Time (ms)
VGG16	16296K	**0.671**	1136.473
VGG19	21606K	0.666	1357.670
ResNet50	33034K	0.618	1643.625
DenseNet121	9143K	0.493	1921.401
DenseNet169	15404K	0.292	2953.563
MobileNet	5335K	0.479	**603.593**
MobileNetV2	8558K	0.533	782.538
From Scratch	N/A	0.648	684.758

CNN Model


Input
48×48
F. maps
64@24×24
F. maps
64@24×24
F. maps
128@24×24
F. maps
128@12×12
F. maps
512@12×12
F. maps
512@6×6
F. maps
512@6×6
F. maps
512@3×3
FC
256
FC
256
FC
7
convolutions
Max-pooling
convolutions
Max-pooling
convolutions
Max-pooling
convolutions
Max-pooling

Pre-Trained Model vs CNN Model

- Improper Prediction
- Accuracy
- Experimented with less Data
- Tuning of parameters
- Enhanced Baseline

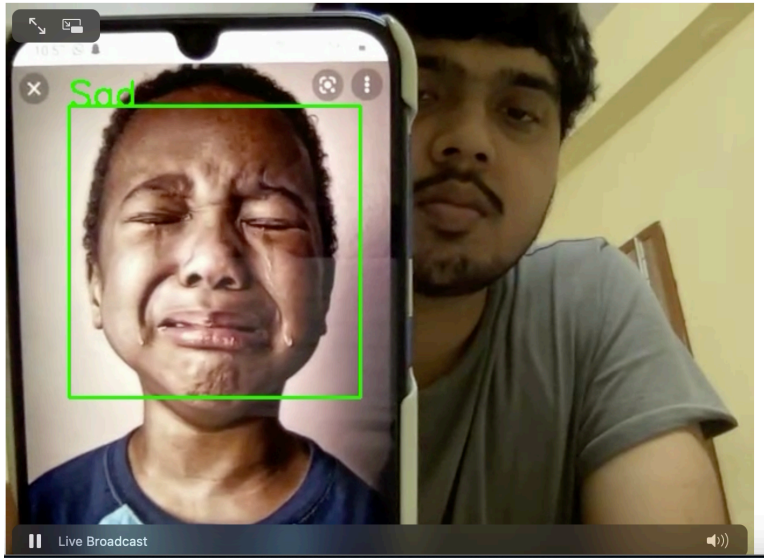
Real time Emotion Detection - Local server

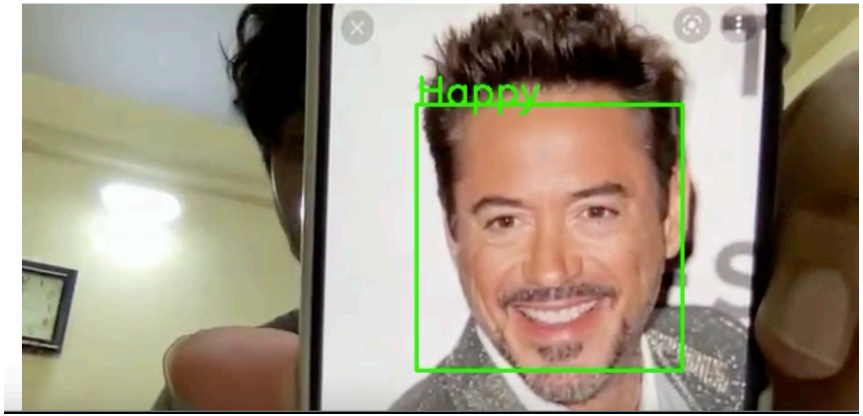
- The front end to deploy model as web app is done using streamlit. As opencv-python cannot be used for cloud servers, we used streamlit-webrtc for that purpose.
- To access video of web app running on local server click on link given: https://github.com/SajalSinha/Facial-Emotion-Recognition/blob/main/Sample_Local.mp4

Deployment Platform

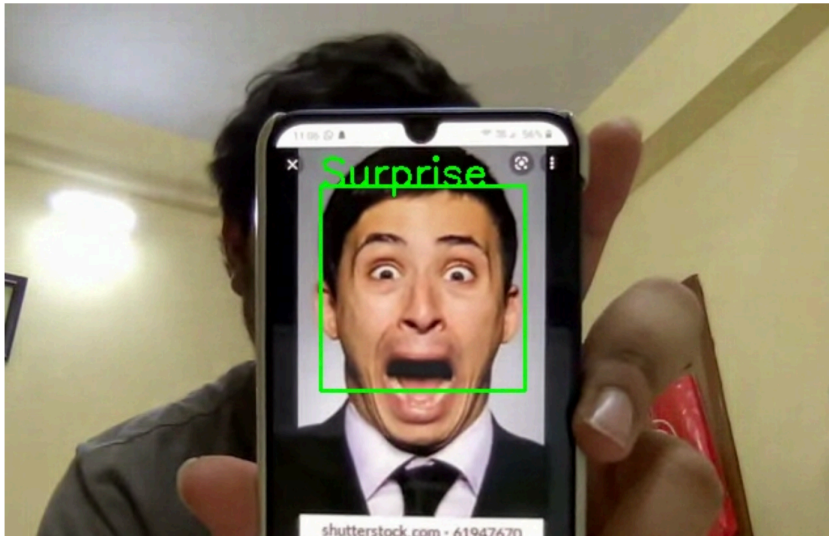
- Heroku is a cloud platform as a service supporting several programming languages. One of the first cloud platforms, Heroku has been in development since June 2007, when it supported only the Ruby programming language, but now supports Java, Node.js, Scala, Clojure, Python, PHP, and Go.
- To access web app please click on this link (Heroku): https://face-emotion-recognition-webap.herokuapp.com/
- To access web app please click on this link (streamlit share) : https://share.streamlit.io/sajalsinha/facial-emotion-recognition/main/app.py

Snaps of local server working

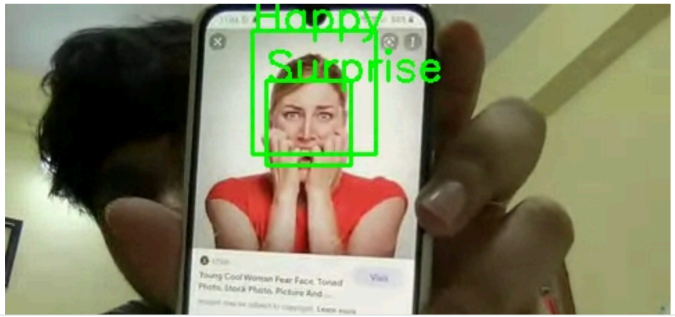




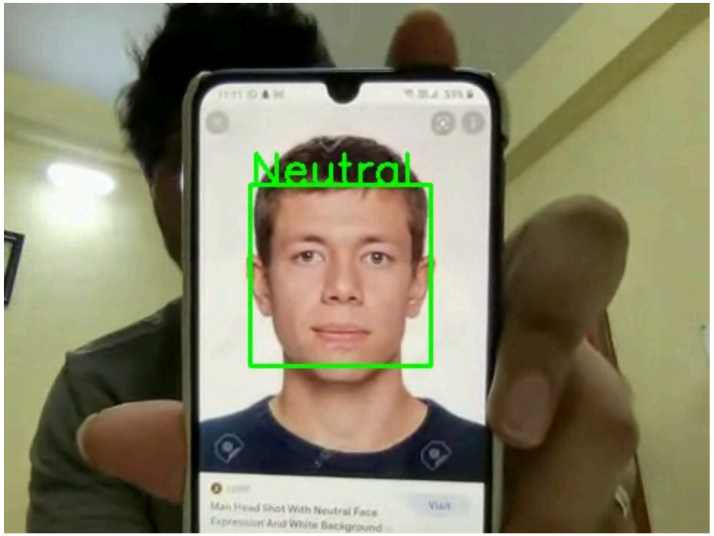

Angry
Live Broadcast


Surprise

6. Still Webcam window did not open, go to check camera from the sidebar.



streamlit_logo.png | tf-logo-card-16x9.png | python-program....png



Conclusion

- Our model gave validation accuracy of ~62% and train accuracy of 90%.
- Model is successfully able to detect emotions and faces.
- It got easily deployed on Heroku and Streamlit Share
- Frontend Work has been done using Streamlit

Challenges Faced

- While Making CNN model, its training was hard n Colab as system was crashing often.
- Running high epoch was stressful
- Access of webcam in google colab as well as in streamlit was tough.
- While deployment, version compatibility has to be properly managed

Thank-you