Facial Expression Recognition System

Advisor: Professor Peter Milder

John Boccio, Muhammad M. Hussain, Yi-Cheng Cheng

Stony Brook University
Electrical and Computer Engineering

ABSTRACT

Facial expression recognition(FER) aims to extract additional information out of human images. The expressions used often consist of the following: happy, sad, fear, disgust, neutral, angry, surprise. Finding these expressions is a complicated and non-trivial task, which has led to the use of convolutional neural networks. By using large, labeled datasets, these neural networks can be trained to predict the expressions shown on a face.

INTRODUCTION

We have developed a practical FER algorithm, and applied it to our internet of things device using a Jetson Nano and camera, resulting in a real-time high accuracy system.

Goal of front-end device: Can be placed anywhere for a user can gather data on the facial expressions of people by the camera.

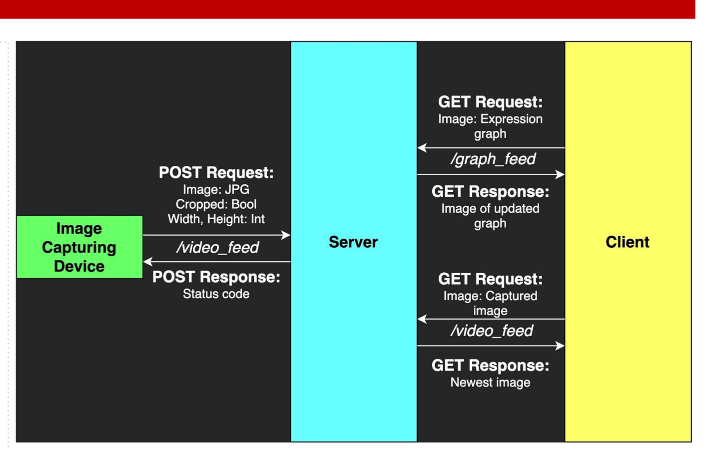
Use cases: customer satisfaction evaluation, psychology experiments, helping people with autism, advertisement signs, interviews, dating & ride-hailing applications, to observe movie watchers, application users, drivers, etc.

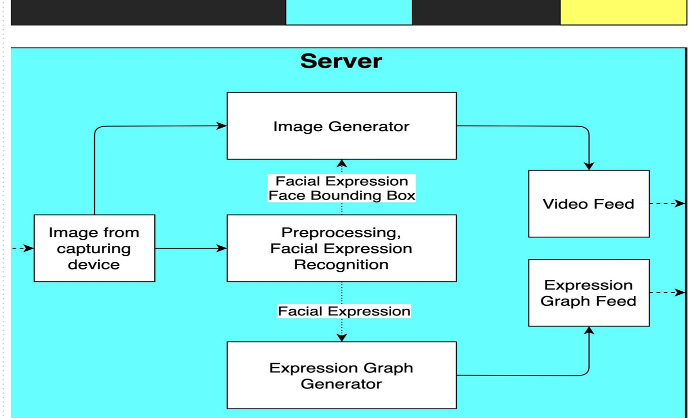
The device connects to a server where various statistics of the expressions seen can be stored and analyzed. We have successfully created an end-to-end internet-of-things (IoT) device that includes an image capturing device, server, and client so that the front-end display can be accessed from any device.

Implementation using multiple convolutional neural networks: **VGGFace** for facial expression recognition and utilizing transfer learning using modern models such as **EfficientNet.**

METHODS

This entire project can be viewed as a data pipeline. At the input, our image capturing device is capturing a stream of images that get pre-processed using OpenCV computer vision algorithms. These frames can then be sent to a server that does the heavy computation for determining the facial expression, and finally, the results are handed off to the front-end display to be viewed in a web browser.





RESULTS

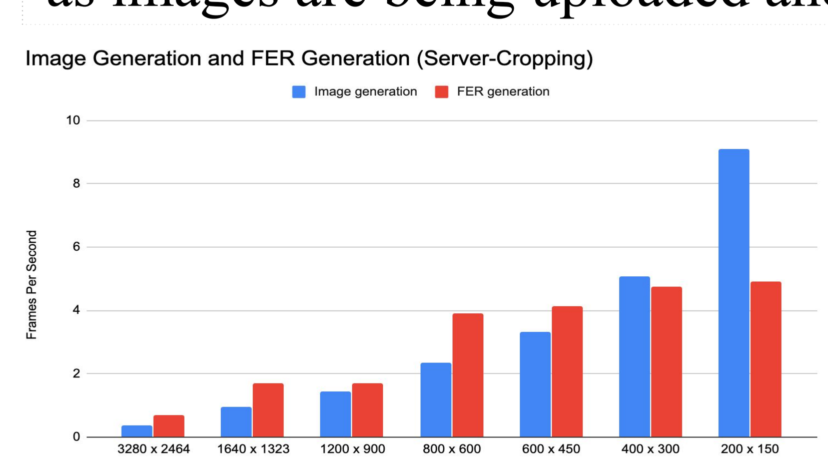
Image Processing Algorithms: Face Detection and Image Crop using Haar Cascade Classifier and OpenCV. Illumination normalization through Histogram Equalization and Gamma Correction.

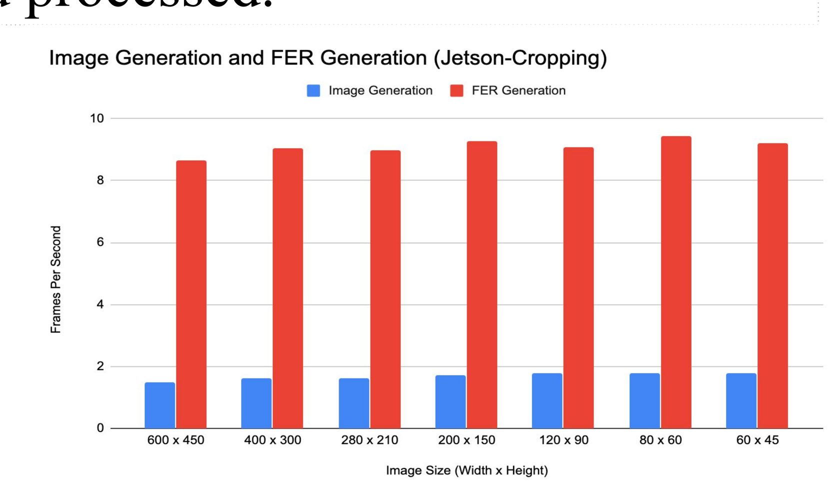
FPS of Image Generation & Expression Generation: The facial recognition and cropping on the Jetson could not process images fast enough to keep the image sending fully saturated. Due to this, we decided to transfer full images of size 400x300 to the server and perform all the processing on the server-side for optimal FPS.

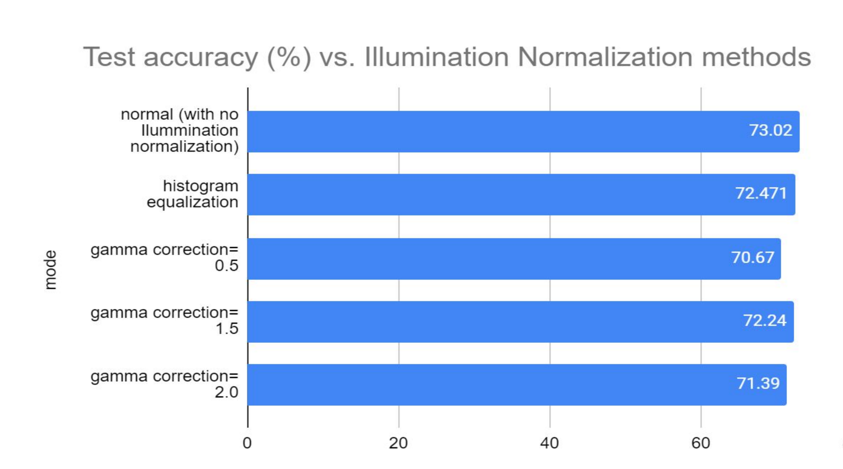
VGG-Face: A CNN investigating how the depth of a CNN affects its accuracy in a large scale image recognition setting. 160M parameters for detecting expressions, that has an accuracy of 73% accuracy on FER2013's test set.

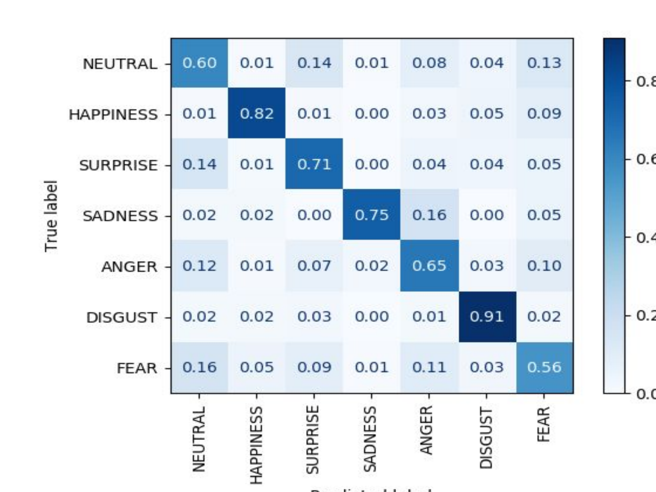
EfficientNet Transfer Learning on FER2013: We used EfficientNet for transfer learning due to its promising results on other datasets. After training EfficientNet-b2, we were able to achieve 64% accuracy while being ~14x smaller than VGGFace.

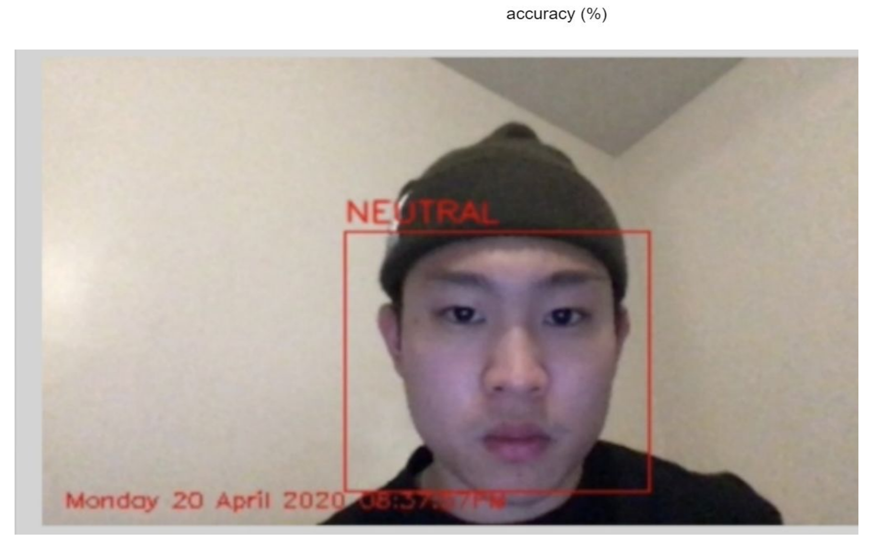
Front-end Display: The video data from Jetson's camera is streamed to a browser by rendering on an HTML page along with an expression probability graph generated with Matplotlib. The front-end display is dynamically updated by the server in real-time as images are being uploaded and processed.

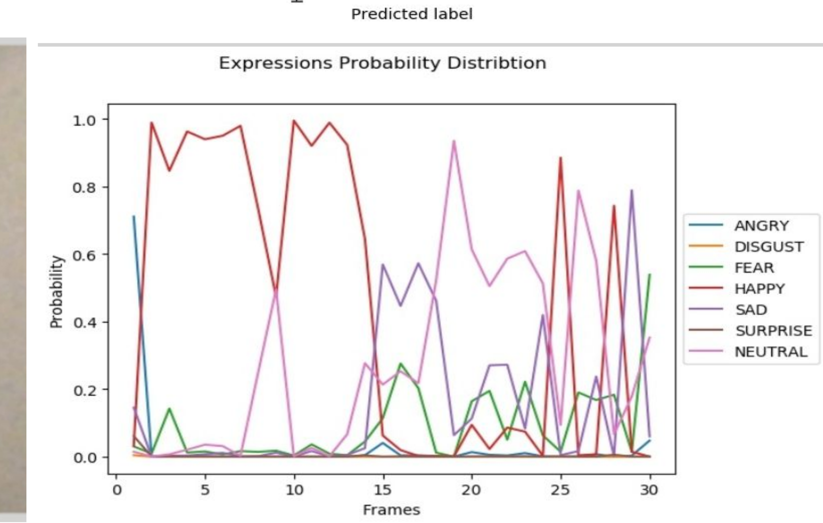












DISCUSSION & CONCLUSION

Our Facial Expression Recognition system can detect facial expressions using VGG-Face with 73% accuracy with FER2013's test set.

We have successfully created an end-to-end IoT device that includes an image capturing device, server, and client. The front-end can be accessed from any device.

Implemented CNNs for facial expression recognition, including modern models like EfficientNet, along with Computer Vision and Deep Learning concepts with Python Libraries & frameworks.

IMPLICATIONS

Our Facial Expression Recognition System will detect facial expressions which could be implemented in any situation where human reactions are important or necessary.

Environmental Concerns: Camera on standby mode in off-hours

Ethical and Legal Concerns: Terms of Use Agreement

Privacy & Security Concern: Server over https so that the data is encrypted as it's sent, & use a secure AWS server so that the servers cannot be hacked.

KEY REFERENCES

★ Image Processing: The analysis and manipulation of a digital image, in order to improve its quality for intended purpose.

★ Convolutional Neural Network (CNN): A specialized neural network used for data which has spatial relationships.

★ Histogram Equalization improves contrast in images by spreading out the intensity range of the image.

★ Gamma Correction is a function that maps luminance levels to compensate for the non-linear luminance effect of display devices, or sync it to human perceptual bias on brightness.

★ FPS: Frames per Second

★ IoT: Internet-of-Things

★ FER: Facial Expression Recognition

ACKNOWLEDGEMENTS

Professor Peter Milder

Tony Olivo, Senior Design Lab

Team Name: Oizys