American Sign Language Detection with Tensorflow

Arturo Arriaga, Hunter Carty, John Valente

Harvard Extension School

DGMD E-14 Section 1 (16693)

Professor Jose Luis Ramirez Herran

December 14, 2021

Project Motivation:

The motivation for this project came out of an interest to merge the ideas of gesture recognition and computer vision. Two members of our group know American Sign Language and all three members expressed a desire to learn the fundamentals of computer vision. After considering variations of the project, we ultimately settled on a system that would meet the following group interests.

- Address the needs of American Sign Language users which total about 500K in the US, and over 70 million worldwide (with other variations of signed languages).
- Explore the fundamentals of computer vision.
- Incorporate object detection and gesture recognition in real time.

Literature Research:

Visual Object Recognition - Kristen Grauman & Bastian Leibe
Jumpstarting Raspberry Pi Vision - Sandy Antunes & James West
https://towardsdatascience.com/sign-language-recognition-using-deep-learning-6549268c60bd

Object Detection: Current and Future Directions
https://www.frontiersin.org/articles/10.3389/frobt.2015.00029/full

Keras API documentation
https://keras.io/api/

MediaPipe documentation
https://docs.mediapipe.dev

Project Category

Computer Vision
Gesture Recognition
Motion Detection
Object Tracking

Project Description

Our group project trains a system to recognize 5 distinct signs that are used in American Sign Language in real time and displays the results to the user.

If we are successful at training a model with an accuracy above 90% we will consider moving that model to a raspberry pi or similar device.

Objectives

Train 5 signs with an accuracy of above 50%
Detect in real time the sign that the user presents.
Detect when a user has change a sign.
Detect 2 word sign combinations

Equipment

Windows or Mac computer

Web camera

shorturl.at/fnvD0

Proof of Concept

We used this design as a starting point for designing out system.

Following the Real Time Sign Language Detection with TensorFlow Object Detection and Python | Deep Learning SSD tutorial on YouTube to generate a minimum viable product using TensorFlow (Renotte).

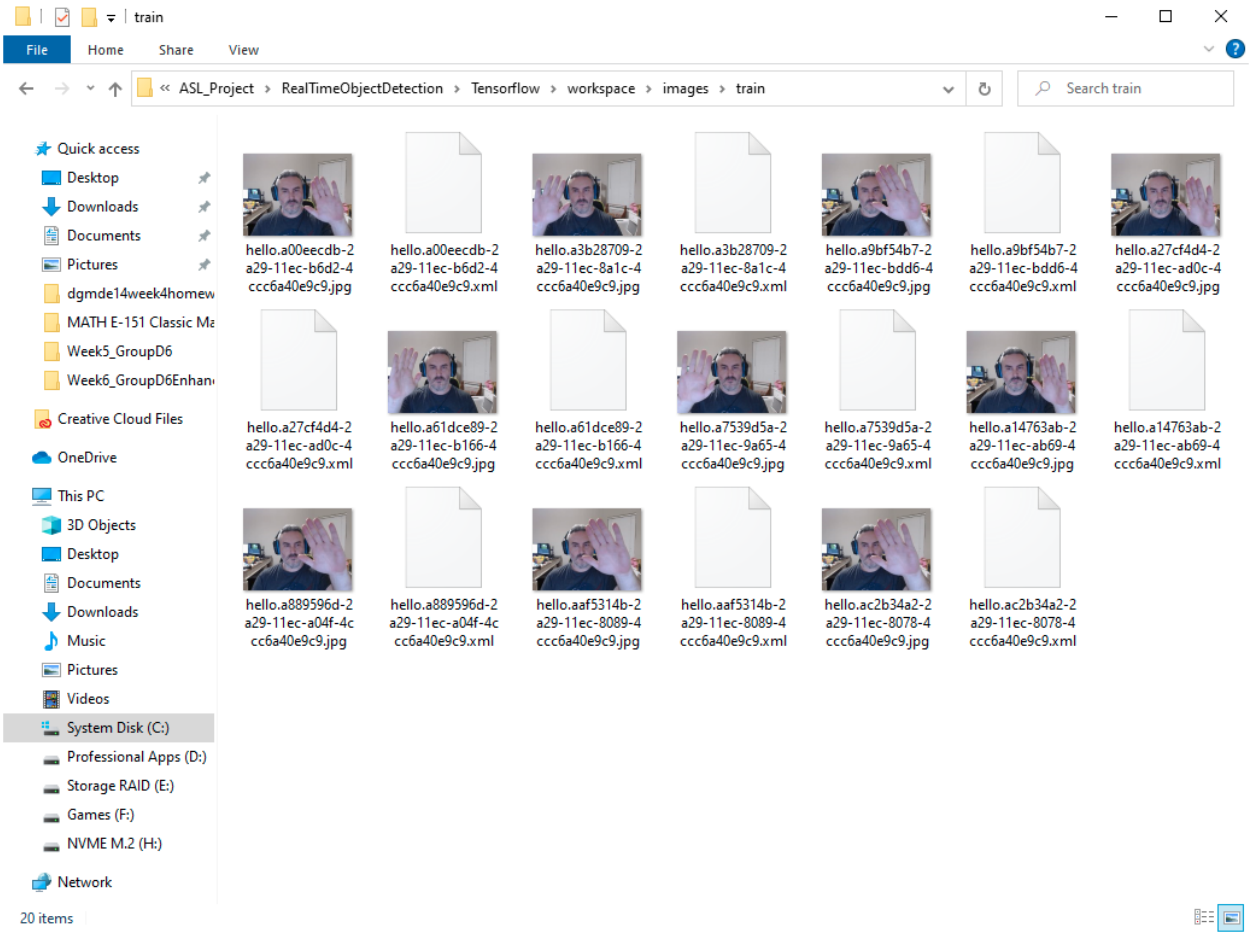


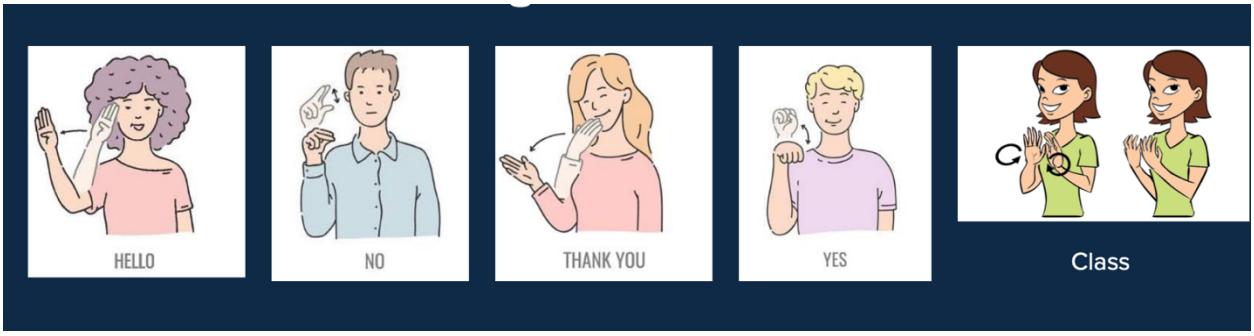
Figure 1. Collecting Several Images to Train the ASL sign for hello.

Implementation:
Our system is designed using 5 linearly depend scripts and implemented with the following tools:
Tensorflow
Keras – a deep learning framework using a sequenial model
Mediapipe – holistic detection of hands, face, and pose positions
Open CV- Computer vision libraries

Python Scripts:
Dependency Installation
Camera/OpenCV testing
MediaPipe landmark collection
Keras Sequential Model training
Real-time Sign detection

Our data model:
Our system was trained to detect the following 5 signs:

1. Hello
2. No
3. Thank you
4. Yes
5. Class



Testing results:
We obtained a wide range of accuracy during our testing. Our system is able to detect a user's sign with an accuracy between 26-90%.

We encountered the following challenges with this system:

- Similarity between signs decreased accuracy.
- Frame rate distorted changes between signs.
- Sign location and movement decreased accuracy.

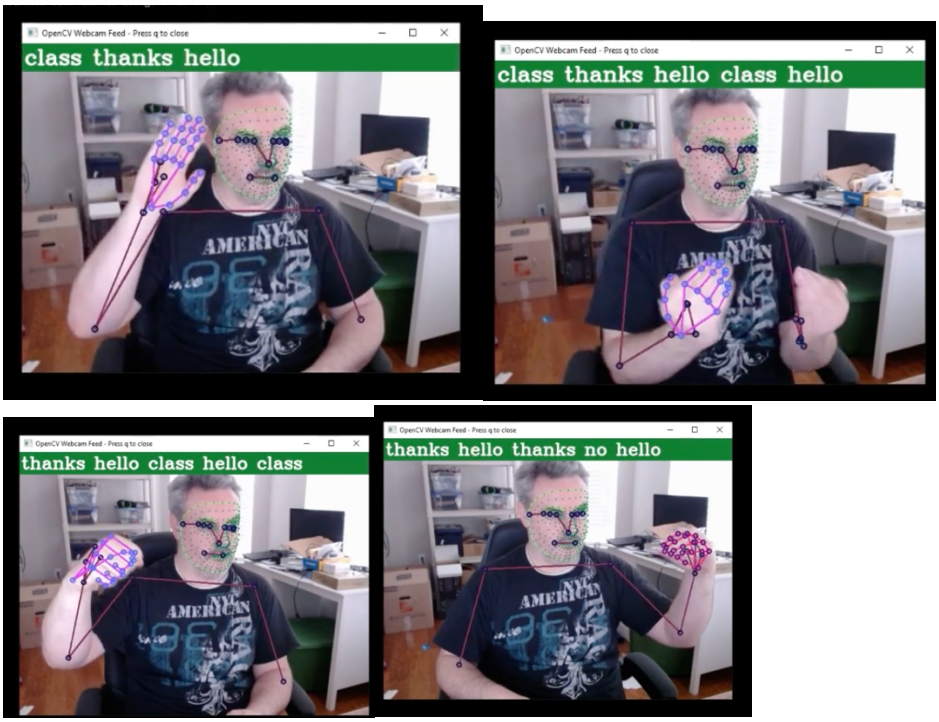
Final Product Demo:

The following demo shows the three primary scripts.

- **Capture**
- **Training**
- **Detection**

https://www.youtube.com/watch?v=XH5PCTVzV70

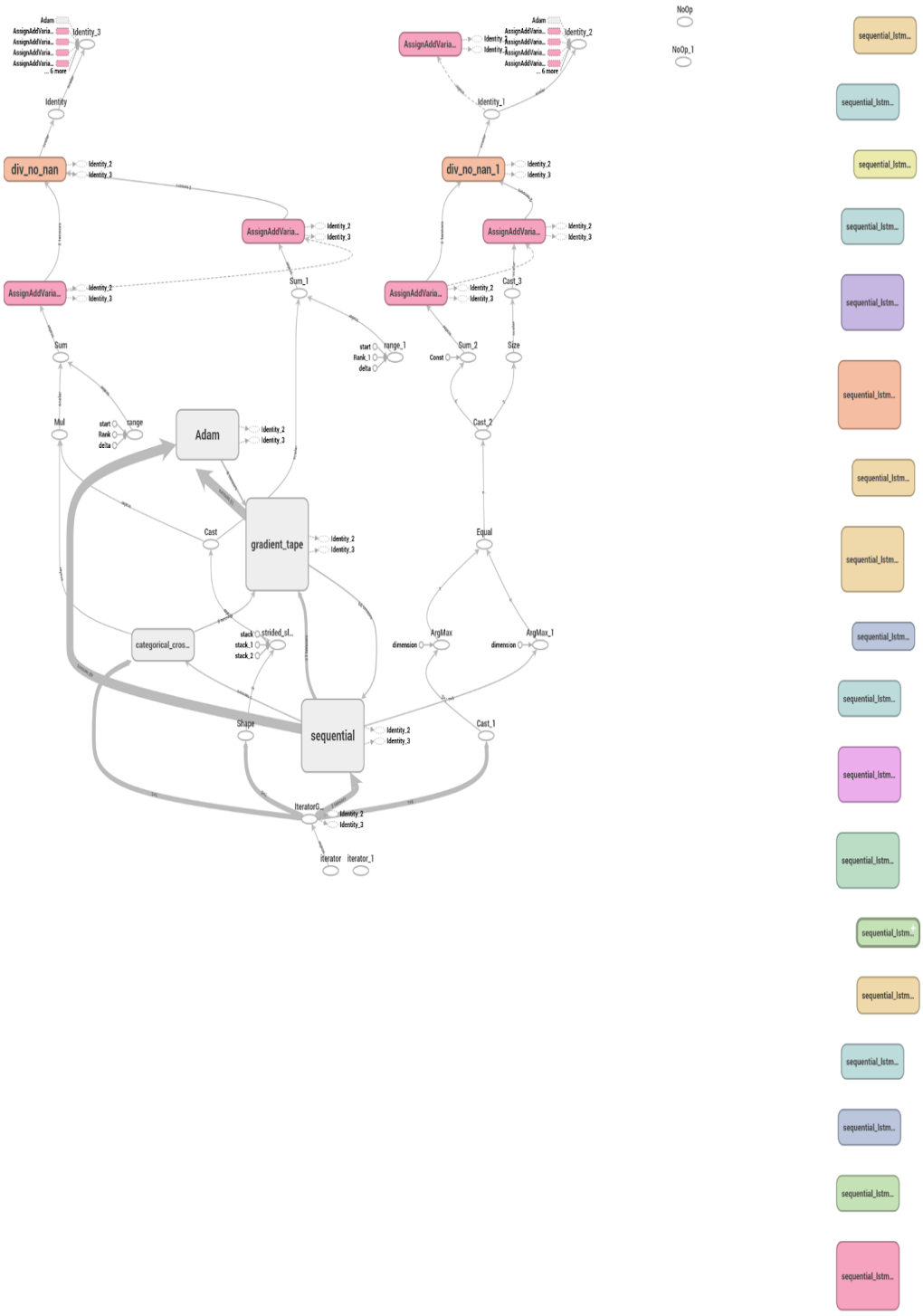
This demo is a rough capture, with limited training a lot of false positives were present.



Visualing the flow of data

The following figures show the results of our Tensorboard scalar graph depicting the changes in loss and accuracy after every epoch. The results show some inconsistencies and the accuracies of our model varied from 25%-96%.

Tensorboard Graph
This graph shows our model's design.



Conclusions:

Our project was a success. We were able to train the model to recognize the 5 signs that we intended on training. We accomplished our goal of learning the fundamentals of computer vision and adapt to the constraints of working with a raspberry pi when training our models.

Given the limitations of the raspberry pi, specially it's low frame rate, we were unable to use our trained model on the system. An implementation with high accuracy would require a higher model and/or the inclusion of a Google Coral to improve the frame rate of the device.

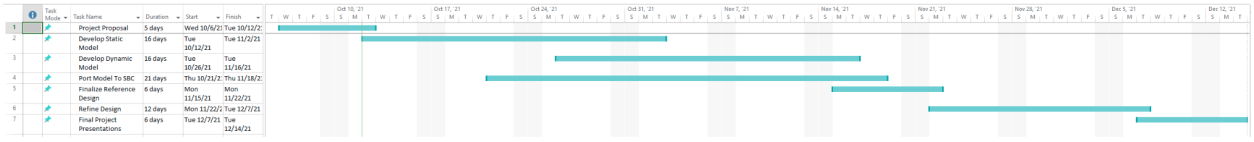
Future Expansion

- Our system could be enhanced and expanded by improving the following:
- Accuracy of the signs.
- Accuracy detection sign motion.
- Increase the number of signs trained.
- Add the ability to detection multi-word signs.
- Adapt to the ability to recognize facial gestures in addition to signs.
- Adapted to run on a smartphone to take advantage of the faster processing and frame rate speeds.

Work Plan and Milestones

Tentative Plan:

- Brainstorm, design, and develop proposal
    - 10/6 - 10/12
- Develop Static Model
    - 10/12 - 11/2
- Develop Dynamic Model
    - 10/26 - 11/16
- Complete Port/Regeneration of Models On SBC
    - 10/21 - 11/18
- Finalize Reference Design
    - 11/15 - 11/22
- Refine Design
    - 11/22 - 12/7
- Final Project Presentations
    - 12/7 - 12/14



Milestone dates:

- October 12, 2021 - Project Proposal

- October 19, 2021 - Proof of Concept (working model detecting one sign)
- November 2, 2021 - Develop Static Model (Simple Signs)
- November 16, 2021 - Develop Dynamic Model (Hand Gestures)
- November 18, 2021 - Finalize Porting Model From PC/Mac To SBC
- November 22, 2021 - Reference Design Finalized
- December 7, 2021 - Refined Design Completed
- December 14, 2021 - Final Project Presentation

References

Renotte, N. [Nicholas Renotte]. (2020, November 5). *Real Time Sign Language Detection with Tensorflow Object Detection and Python | Deep Learning SSD* [Video]. YouTube. https://www.youtube.com/watch?v=pDXdlXlaCco

Renotte, N. [Nicholas Renotte]. (2021, June 19). *Sign Language Detection using ACTION RECOGNITION with Python | LSTM Deep Learning Model* [Video]. YouTube. https://www.youtube.com/watch?v=doDUihpj6ro