



Runarljod

Rune Song

Erdos Data Science Bootcamp Project
Fall 2022

Group Members (alphabetical order):

Akash Banerjee

Ethan Farber

Balin Fleming

Lauren Keyes

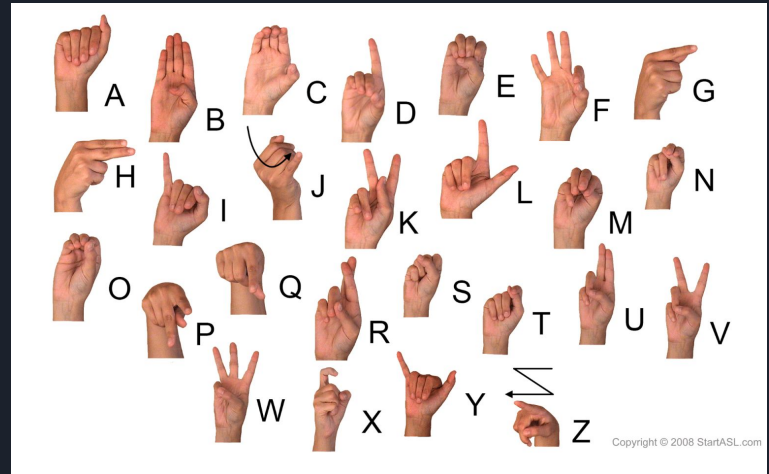
Rouzbeh Modarresi-Yazdi

Abdullateef Shodunke

Po-Kuan Wu (group mentor)

Overview

- American Sign Language (ASL) uses hand gestures to communicate.
 - Finger spelling for individual letters
- Much recent work using ML to recognize ASL
- **Goal:** Synthesize new techniques and innovations to produce a model that can recognize ASL fingerspelling sequences



startasl.com



Data

Several publicly available datasets, including:

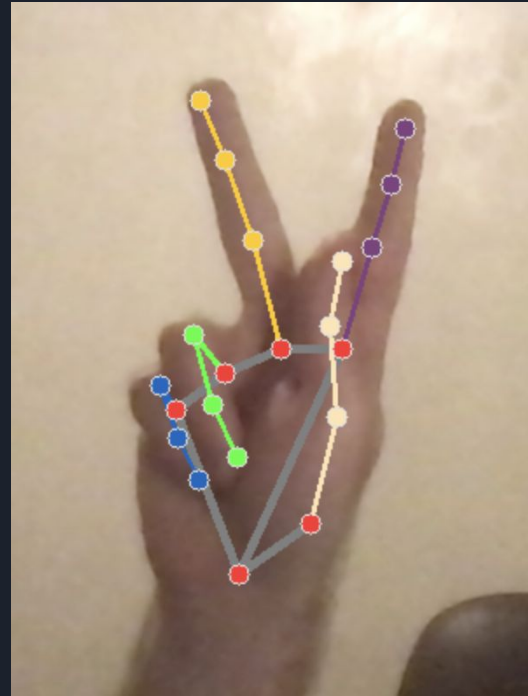
- Dis (<https://www.kaggle.com/datasets/ayuraj/asl-dataset>)
 - 2515 images, cropped
 - 400x400 in color
- Fuzzy-octo-guacamole (<https://github.com/good-soul/fuzzy-octo-guacamole/tree/main/Datasets/Dataset2>)
 - Hand images with bounding box
 - 28x28

Variations between datasets provided difficulty, but also diversity

- Conditions in which photo was taken
 - Controlled vs “natural”
- Blurriness
- Variations between signers

Data Processing

- Detect and model hands in images with off-the-shelf solutions
 - OpenCV
 - MediaPipe-Hands
- Crop and resize images to a standard format
- Augment data with randomly generated variations
 - Rotations
 - Reflections
 - Brightness/contrast
 - Random noise



The output of the MediaPipe-Hands module

Model Architecture

2D Convolutional Layer (with Relu)

x2

Max Pooling

2D Convolutional Layer (with Relu)

Batch Normalization

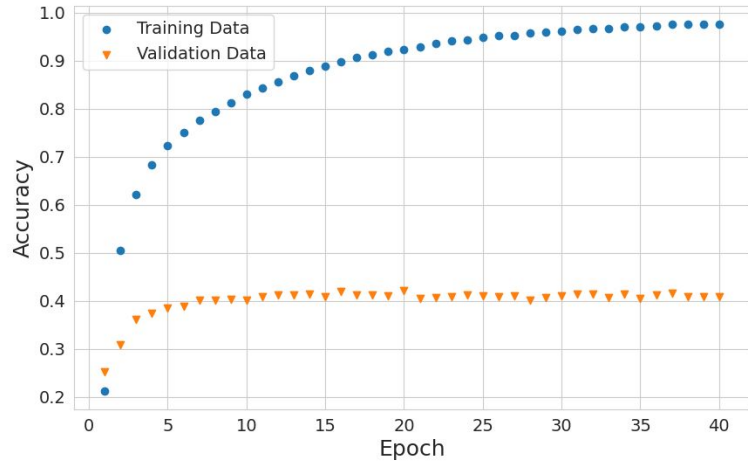
Max Pooling

Flatten

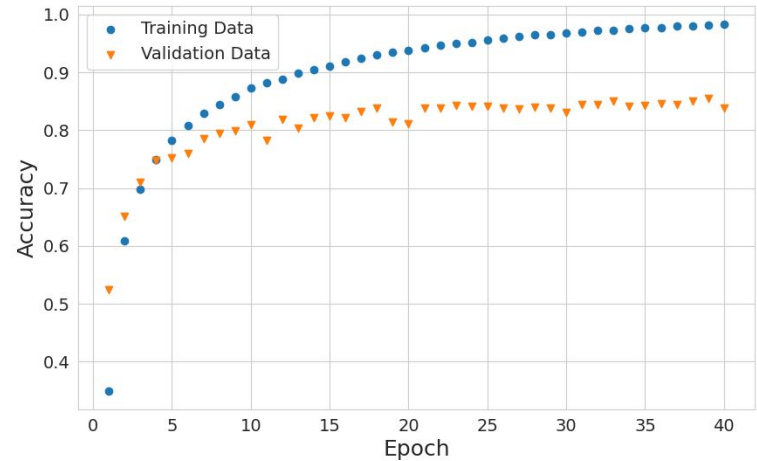
Fully Connected Linear Layers (x3 hidden) (with Dropout) (with Relu)

Classifier (linear layer with Softmax)

Accuracy Plots: with and without batch normalization



Without Batch Normalization



With Batch Normalization



Conclusion

- Results
 - 85% validation accuracy
- Future directions:
 - Model fine tuning
 - Alternative architectures
 - Video recognition