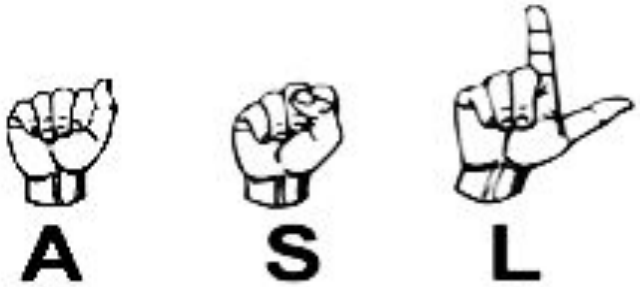
Signing Off on



Using Deep Learning to recognize hand signs from the ASL alphabet

Image source: https://silentword.org/wp-content/uploads/2020/06/ABC-ASL-scaled.jpg

Objective: Create a model th
can classify hand gestures as
ASL Alphabet hand signs.

Data

Kaggle Dataset:

- 77,518 images
- 26 image classes
  - A-Z, plus space, minus J
- ~3,000 per image class

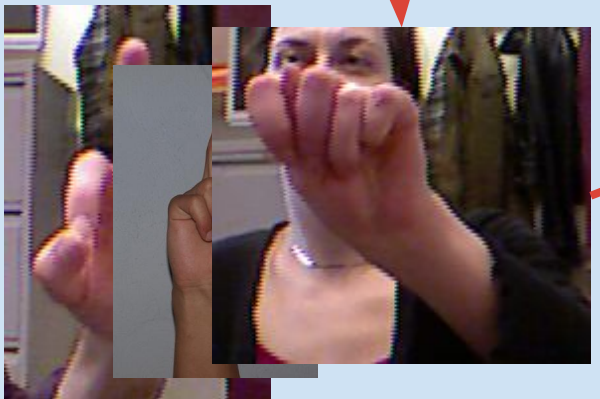
Tools utilized

- Pandas/NumPy
- Tensorflow/Keras
- OpenCV
- MediaPipe
- Matplotlib

Methodology - Preprocessing

kaggle

64x64 px



OpenCV

Methodology - CNN

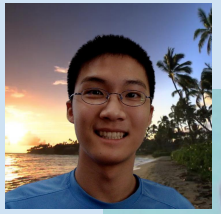
(64,64,3)

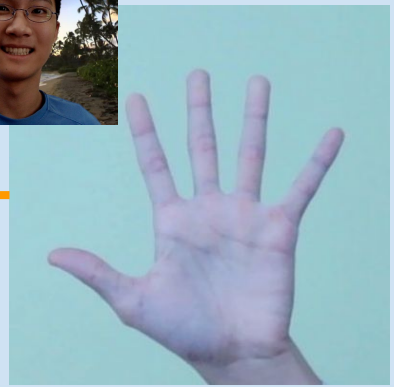
80-20 Train/Test split

Training time: 2 hours

CNN

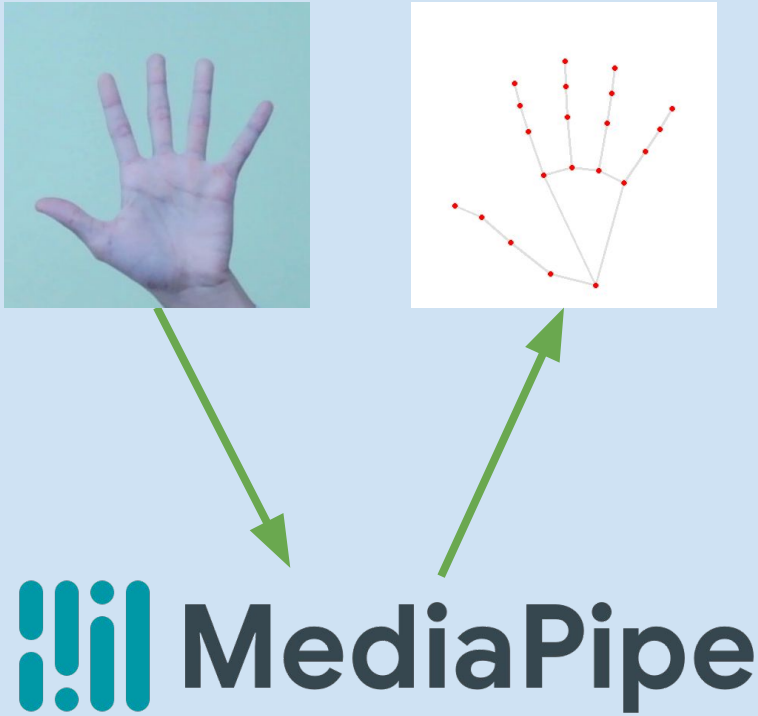
???





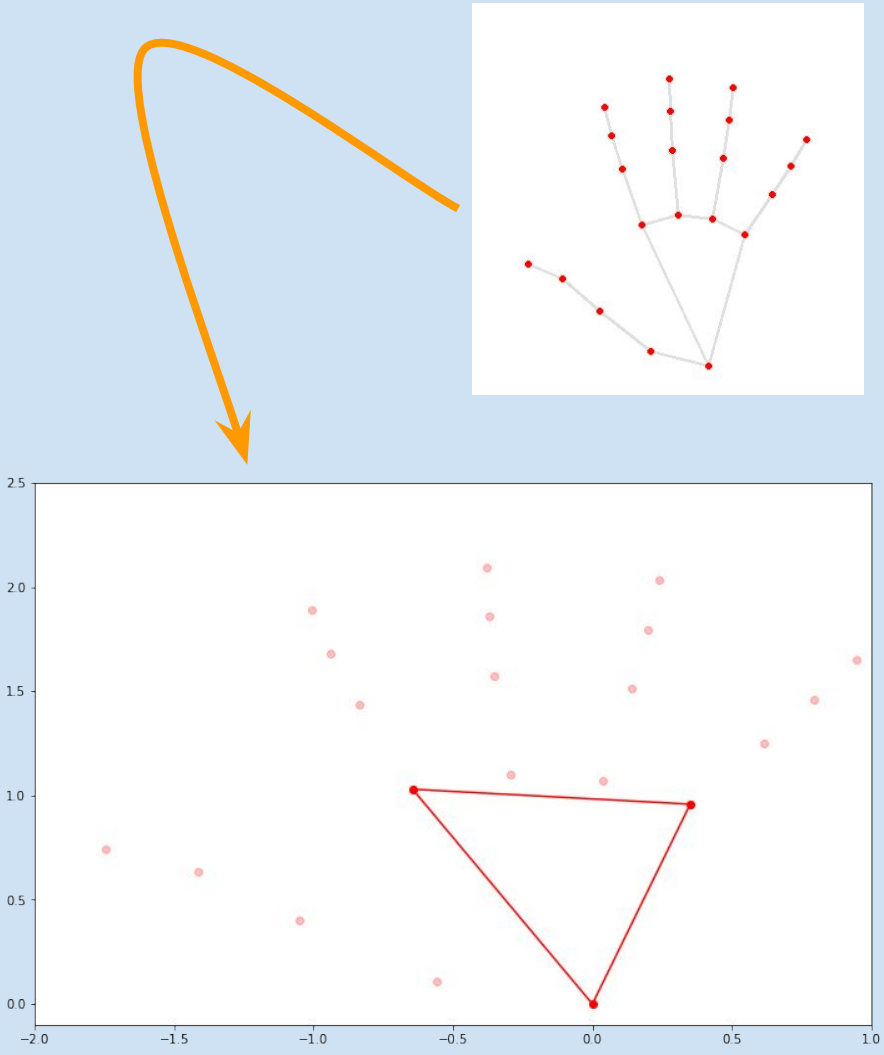
MediaPipe

Google’s MediaPipe library can highlight key points in a photo of a hand, allowing for simpler networks and faster modeling

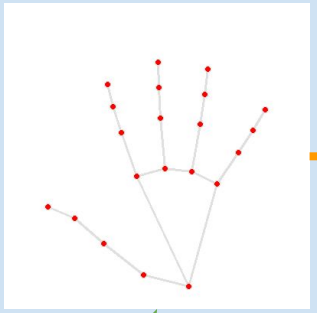


Regularization

In order to prevent overfitting, create a representation of the hand, scaled to the size of the palm.

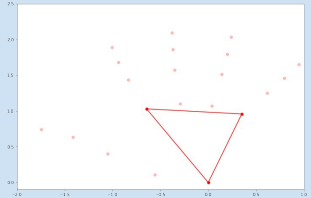


Methodology – MLP



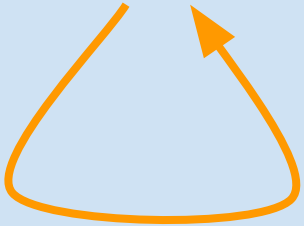
(21,2)

(42,1)



MLP

Training time: 2 minutes



Regularization

Model

Input Layer: 42

Hidden Layers: 96, 256, 128, 64 (relu)

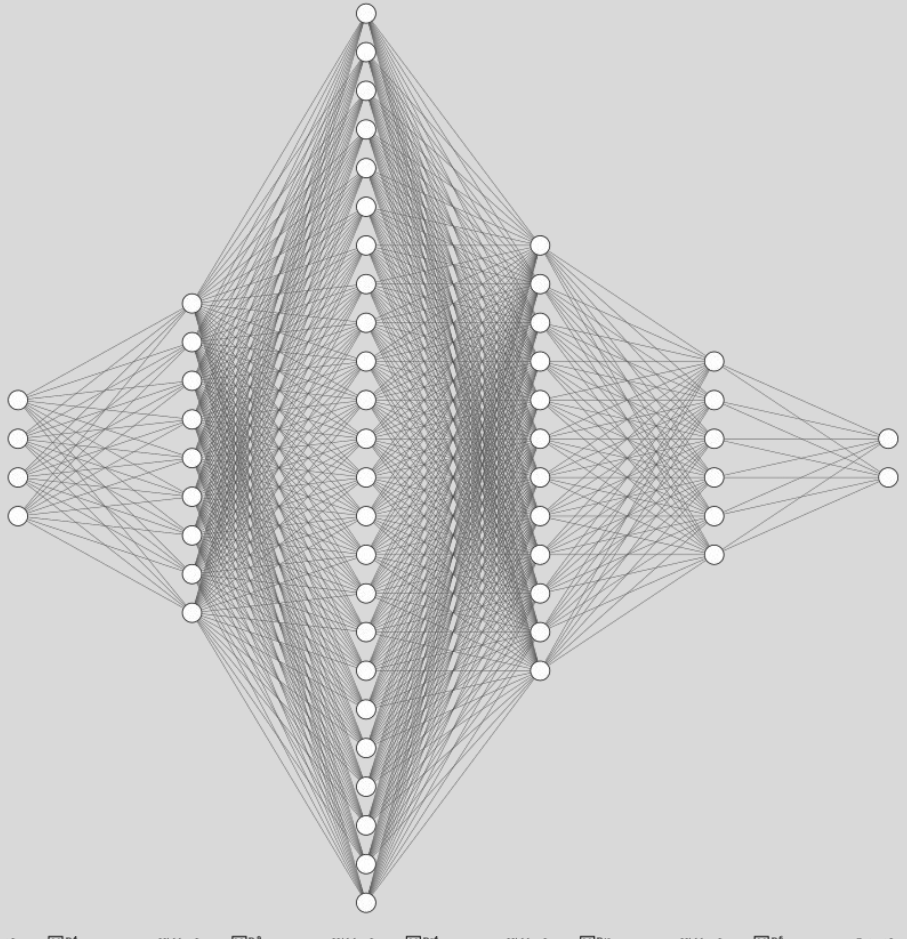
Output: 26 categories (softmax)

Loss metric: Categorical Crossentropy

Optimizer: ADAM

Metric: Accuracy

A highly simplified view of the model



Made using http://alexlenail.me/NN-SVG/

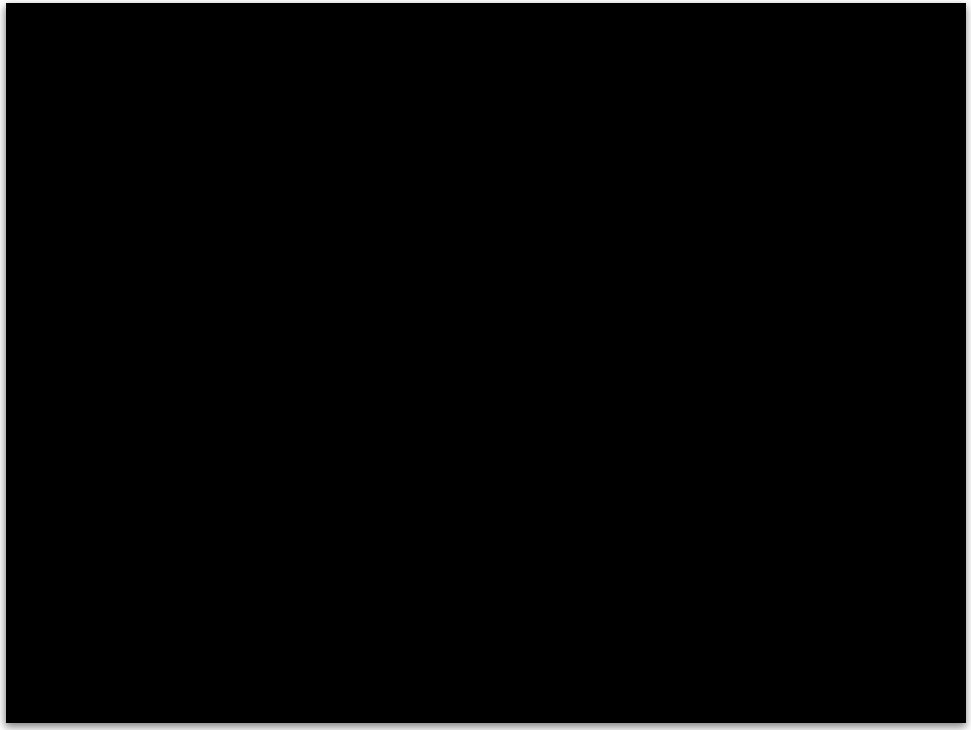
Useful Applications

This model can be used to translate hand signs in a video.

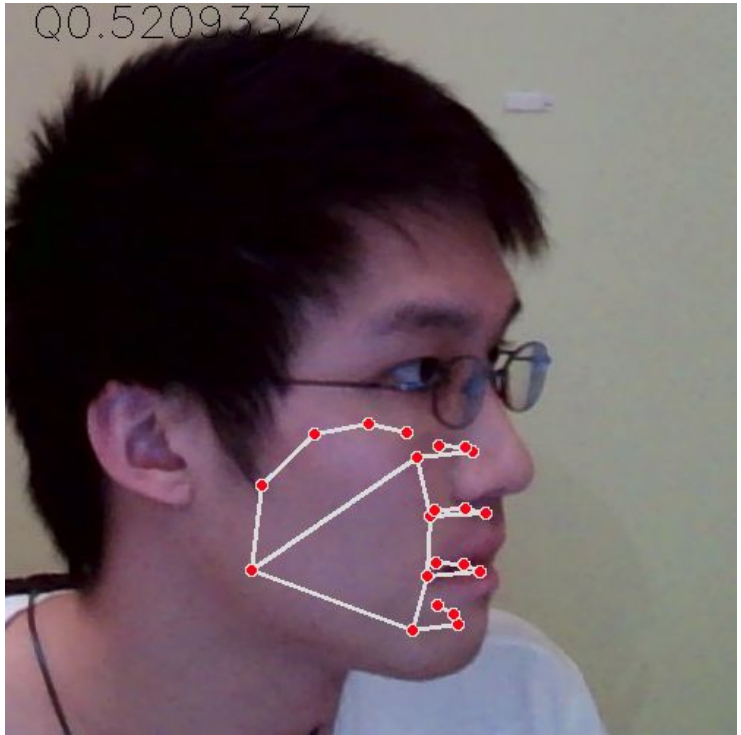
Right: Matthew making a valiant attempt to sign a complete sentence with ASL, and failing miserably

Translation:

I- -L-I-K-E- -T-U-R-T-L-E-S



Less-useful applications



Data scientists making completely useless features for their models:

('F', 0.9997948)

('X', 0.49064082)

Future Work

- Get more data; lack of varied backgrounds limits CNN effectiveness
- Gather videos of more complex gestures, and use RNN to interpret motion
- Code a feature that builds a sentence from hand signs, using NLP data for auto-correct.

Thank you!