



# Translating ASL Using Object Detection

—

Kyle Lindstrom



Kyle Lindstrom

[Github](#)

[LinkedIn](#)

# Agenda

Business Problem

Data Overview

Modeling

Conclusions



# Understanding the Problem

## The Problem

- Around 1 million ASL users in the United States
- Over 99% of the population can't understand

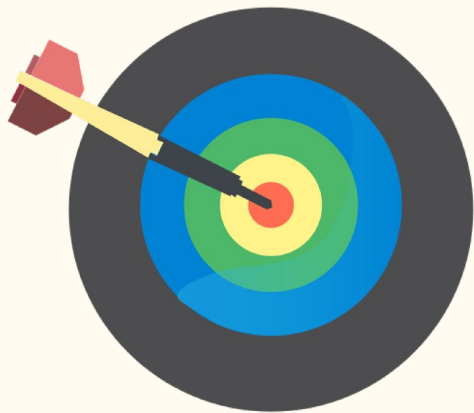
## The Stakeholder

- Google translate wants to incorporate sign language
- They'd like to identify live action video phrases

## The Solution

- Develop a neural network capable of the required predictions
- Identify and track still hand signals as well as complex motions

# Key Performance Indicator



## Accuracy

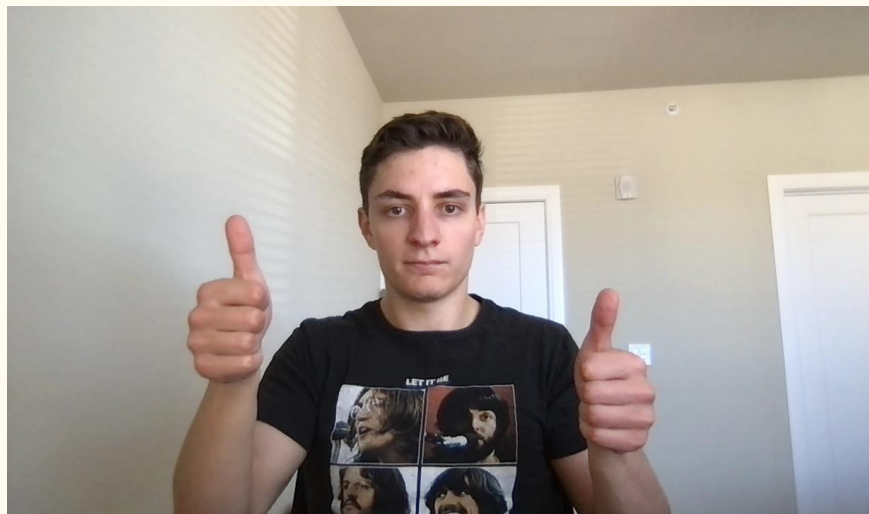
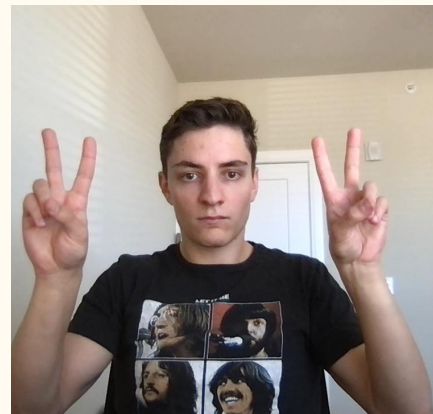
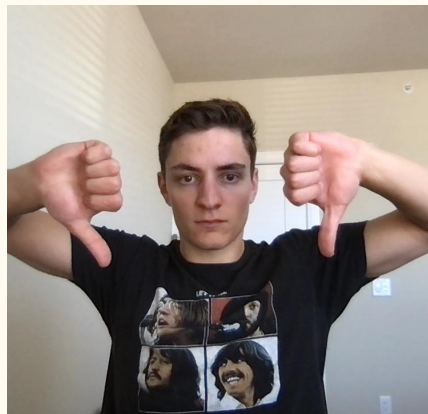
What percentage of  
the model's predictions  
are correct?

# The Data

—

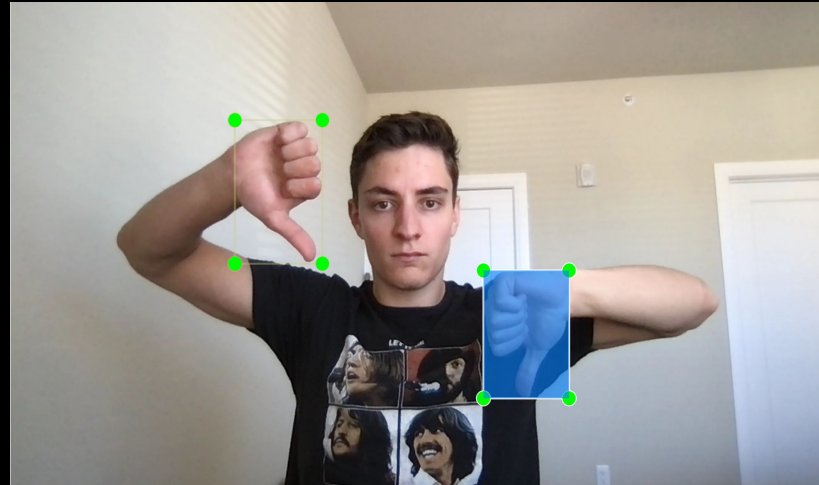
# Procedure

1. Collect many images of the gestures and symbols to be translated
2. Annotate the data accurately using LabelImg and MediaPipe
3. Split into training, validation, and test sets



# Single Frame Hand Signals

Collected using OpenCV  
Annotated using LabelImg





# Complex Gestures

Collected using OpenCV  
Annotated using MediaPipe



# Limitations



1. Only pictures of me
2. Same lighting/background
3. Single person in image

# Final Dataset

Still Images

**120 total examples split into 3 categories:**  
Thumbs Up, Thumbs Down, Peace

Action Gestures

**240 total examples split into 4 categories:**  
Hello, Nice To Meet You, Thanks, Goodbye

# Results

—

# Hand Signals

## Model Architecture:

SSD MobileNet V2 FPNLite

## Testing Accuracy (24 Test Instances):

100%



# Action Gestures

## Model Architecture:

Sequential with LSTM and Dense Layers

## Testing Accuracy (24 Test Instances):

100%



# Future Insights



1. More variety in training data
2. Add a full suite of ASL gestures
3. Explore possibility of combining models

# Thank You!



Kyle Lindstrom

[Github](#)

[LinkedIn](#)

