Hand Gesture Detection

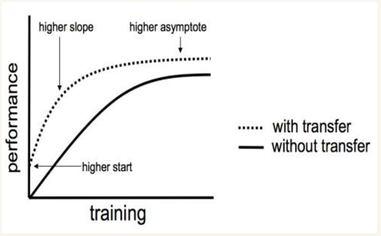
Tom Yang - Sabrina Chwalek - James Hu - Maxime Seknadje

BROWN

CSCI 1430

Motivation

Implement different hand gesture detection models using two different deep learning approaches.

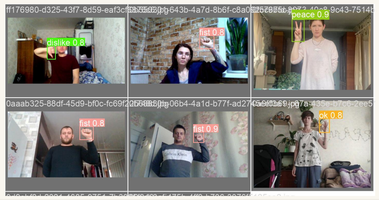


Specifically, we wanted to explore if transfer learning—which is faster than training a model from scratch— yields better results than training a MediaPipe hand landmark coordinate classifier.

Define the problem

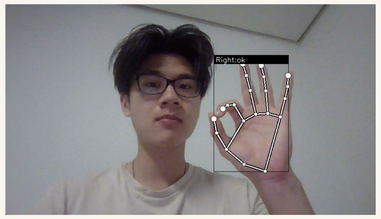
Transfer Learning

1. Label data for transfer learning using
2. Finetune yoloV5s object detection model



Landmark Classifier[1]

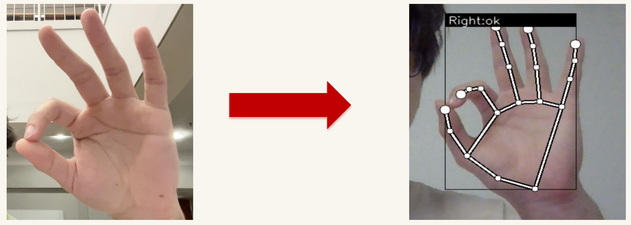
1. Built classification model using mediapipe



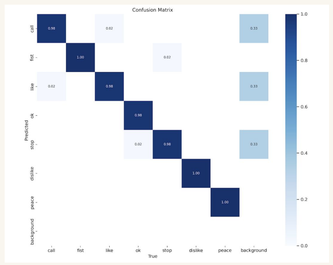
Goal

State Goals

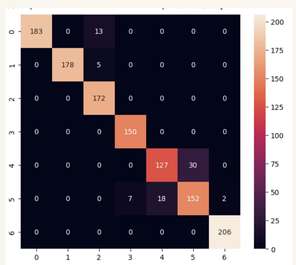
1. Detect hand-gesture in real-time from live video input
2. Train a classifier to determine type of hand gestures
3. Compare performance using two model architectures



Confusion Matrices

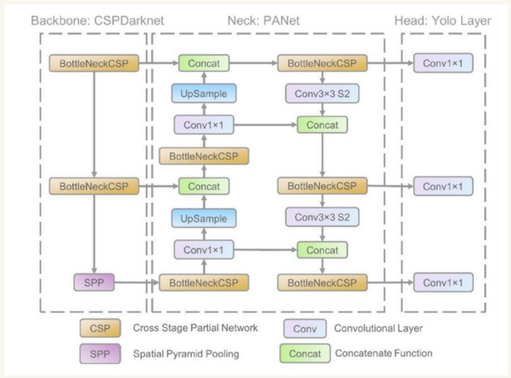


1. **(Left)** yoloV5s confusion Matrix: confuses call and like
2. **(Right)** Mediapipe confusion matrix: confuses dislike and stop

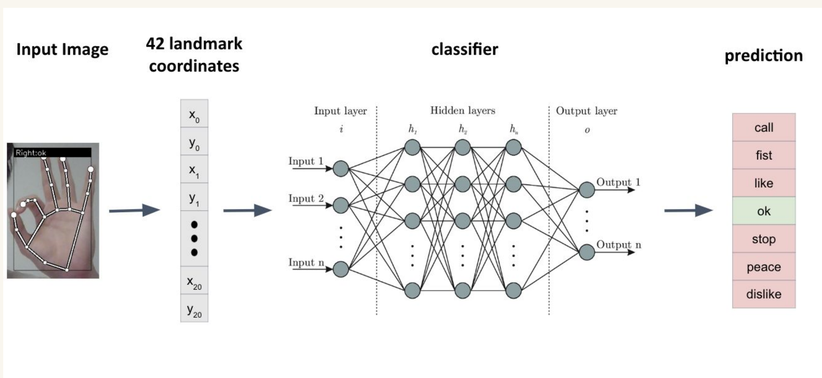


Architectures

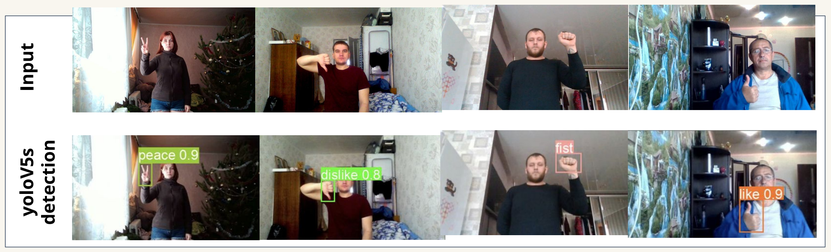
yolov5 architecture[3]

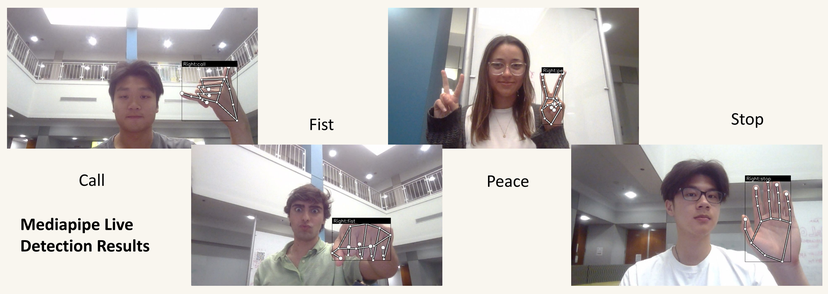


MediaPipe landmark classifier architecture



Results (images/figures)



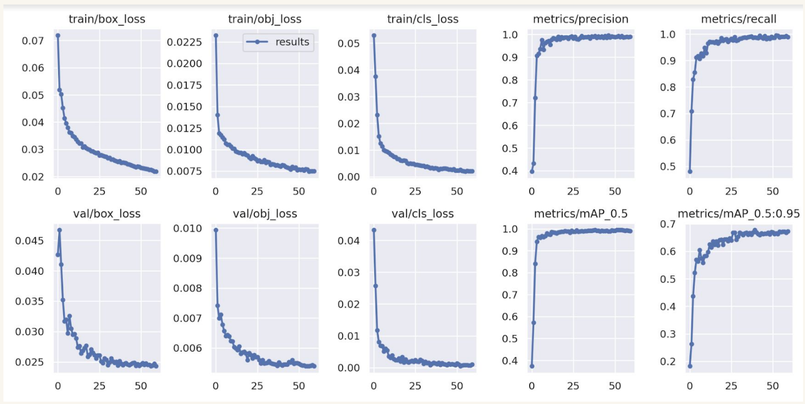


Failures

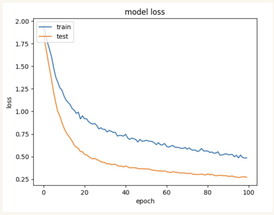
Transfer Learning: We fine-tuned yoloV5s on a dataset with fairly homogenous images. Therefore, our classifier does not perform as well when a user significantly rotates or angles their hand.

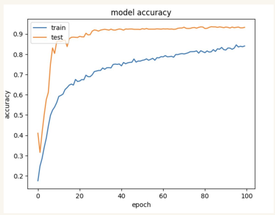
Landmark Classifier: We used MediaPipe to map 21 coordinates on the hand images we passed it, so MediaPipe did not correctly identify coordinates on some training gestures, which negatively affected the performance of our classifiers.

Model Loss & Accuracy



1. **(Left)** yoloV5s train/test loss, accuracy, precision & recall
2. **(Right)** Mediapipe classifier train/test loss and accuracy





References

[1] https://github.com/kinivi/hand-gesture-recognition-mediapipe
[2] https://www.kaggle.com/datasets/innominate817/hagrid-sample-120k-384p
[3] https://ig.opengenus.org/yolov5/

Acknowledgements

We would like to thank our TA, Qian Zheng for guidance, support and mentorship.