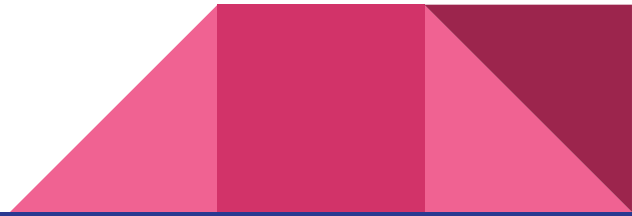
Sign Language Recognition with standard Laptop Webcams: A Study across Machine Learning models

A project made by:
Victor Winkelmann & Hannes Östergren

Supervisor:
Alexander Durr

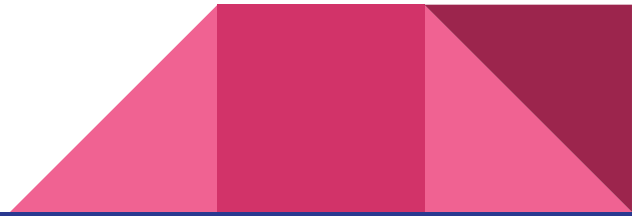
Introduction

- Goal: Translate gestures and static postures into spoken language using ordinary webcams
- Low hardware requirements
- Scalable and modular

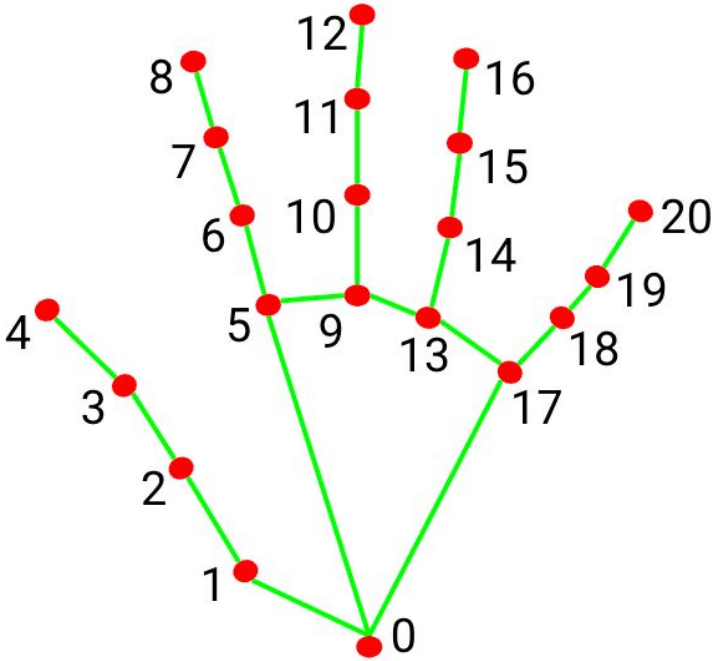


Background

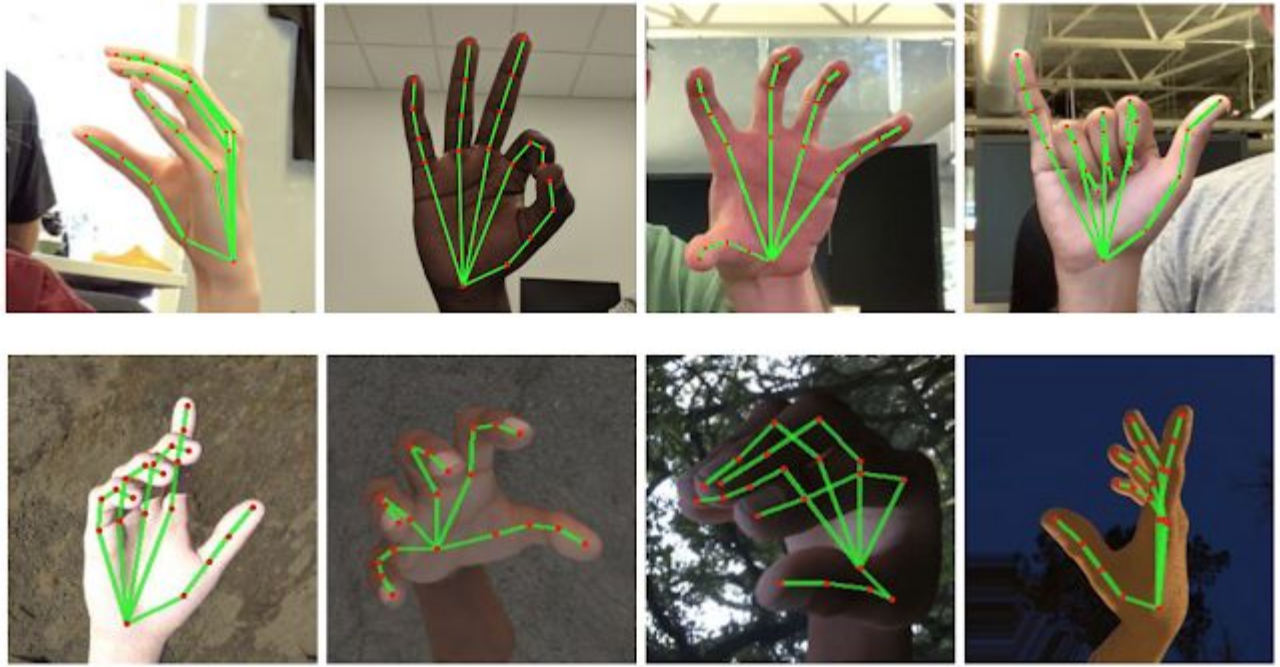
- Open pose
  - Drawbacks
- Mediapipe
  - High performing
  - Modular
- Image recognition vs landmarks
  - Landmarks result in higher performance
  - Filters unnecessary information

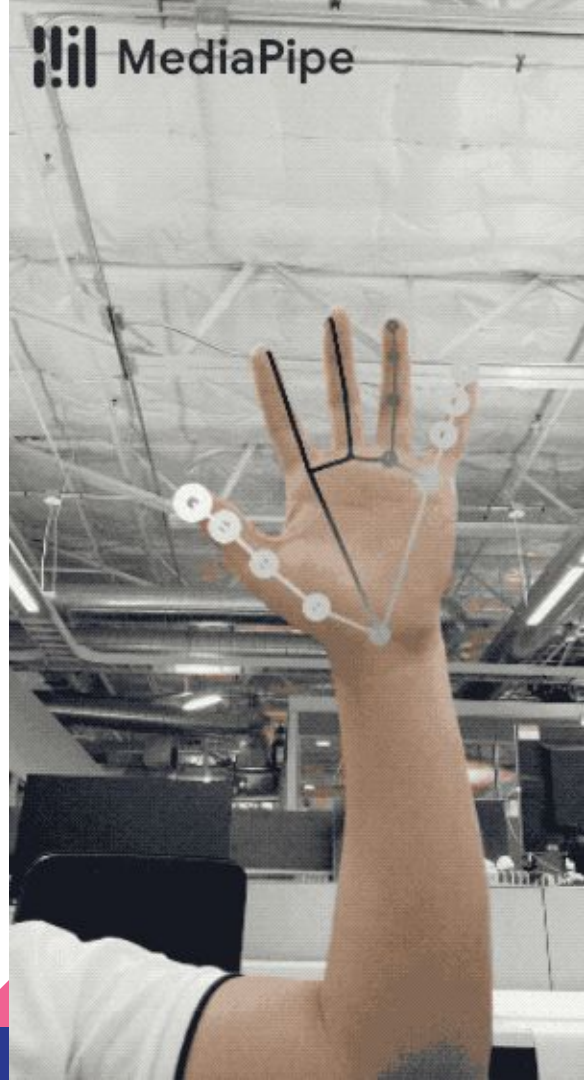


Mediapipe

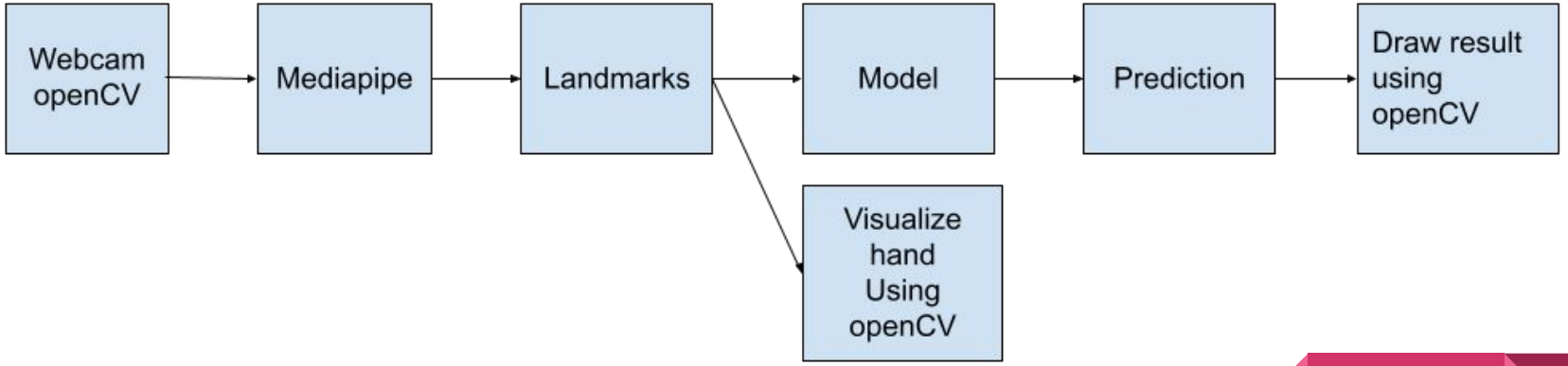


0. WRIST
1. THUMB_CMC
2. THUMB_MCP
3. THUMB_IP
4. THUMB_TIP
5. INDEX_FINGER_MCP
6. INDEX_FINGER_PIP
7. INDEX_FINGER_DIP
8. INDEX_FINGER_TIP
9. MIDDLE_FINGER_MCP
10. MIDDLE_FINGER_PIP
11. MIDDLE_FINGER_DIP
12. MIDDLE_FINGER_TIP
13. RING_FINGER_MCP
14. RING_FINGER_PIP
15. RING_FINGER_DIP
16. RING_FINGER_TIP
17. PINKY_MCP
18. PINKY_PIP
19. PINKY_DIP
20. PINKY_TIP



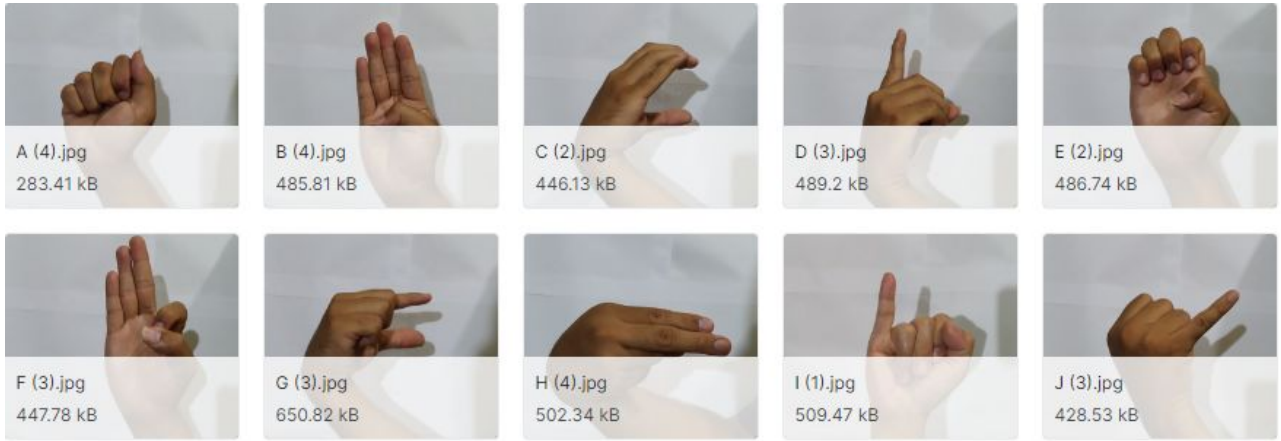

MediaPipe

Pipeline


Webcam openCV
Mediapipe
Landmarks
Model
Prediction
Draw result using openCV
Visualize hand Using openCV

Dataset

- Kaggle dataset



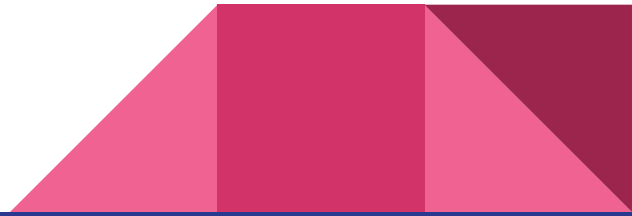
- 1237 datapoints

Model selection

- Scikit-learn (Sklearn)
  - Provides tools for machine learning and statistical modeling like classification.
- Keras
  - Is a high level neural network library that runs on top of tensorflow. It’s a high-level API used for easily building and training models.
- Using scikit-learn vs keras
  - Building an optimal model is a very hard task and has therefore taken up most of our time.
  - Using Sklearn has been a lot easier since you do not have to build the network
  - Keras brings more customizability but are more complicated

Optimization using Keras tuner

- What is Keras tuner
  - Finds optimizations by testing different configurations
  - Defined search space
- Using keras tuner to find an optimal model
  - Sklearn tuner
  - Keras tuner



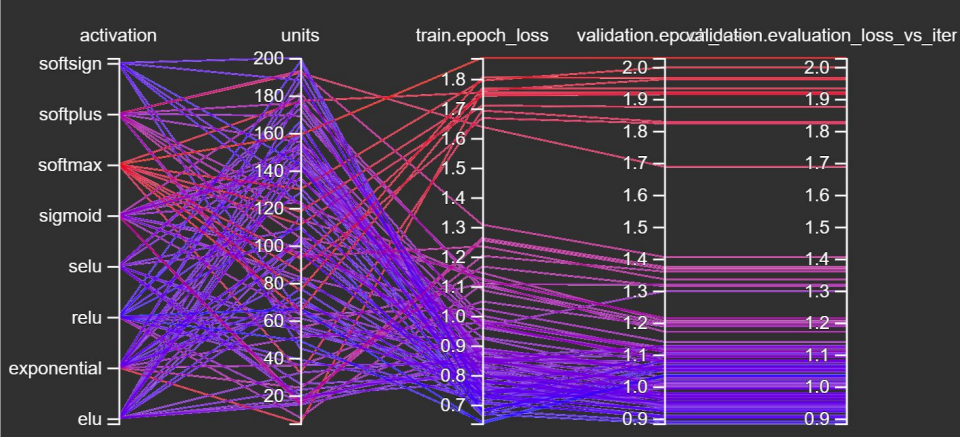
Performance

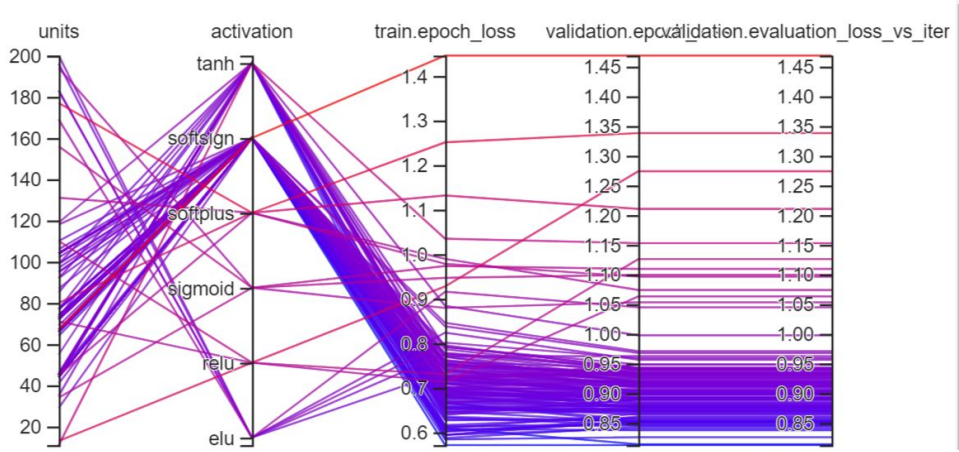
- Different models:
  - Nearest neighbours
  - SVM
  - Gaussian_process
  - Decision tree
  - Random_forest
  - Neural network model
  - Keras model
- Best performance according to the sklearn tuner/keras tuner:
  - Support vector classification, C = 8.2, gamma = 2 -> 87.5 accuracy
  - Keras model

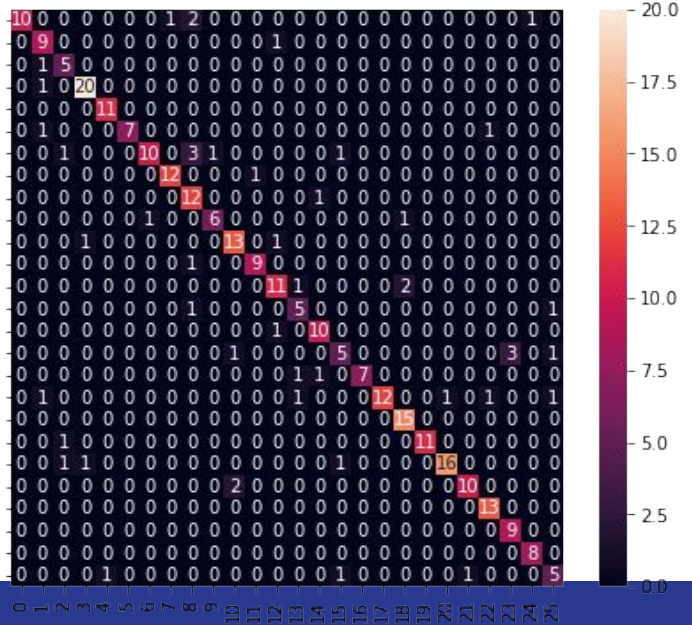
```
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense (Dense)               (None, 66)                2838

 dense_1 (Dense)             (None, 26)                1742

=================================================================
Total params: 4,580
Trainable params: 4,580
Non-trainable params: 0
_________________________________________________________________
<keras.engine.sequential.Sequential object at 0x000001F26F32E1C0>
29/29 [==============================] - 0s 1ms/step - loss: 0.5748
10/10 [==============================] - 0s 1ms/step - loss: 0.8150
0.8150477409362793
```

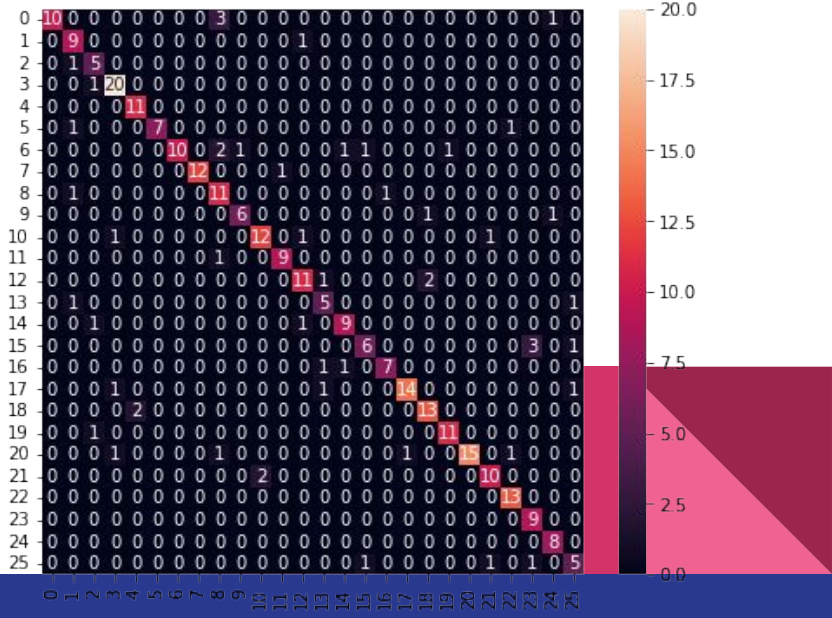
```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense (Dense)               (None, 114)               4902

 dense_1 (Dense)             (None, 26)                2990

=================================================================
Total params: 7,892
Trainable params: 7,892
Non-trainable params: 0
_________________________________________________________________
<keras.engine.sequential.Sequential object at 0x00000227E0CB5870>
29/29 [==============================] - 0s 1ms/step - loss: 0.6335
10/10 [==============================] - 0s 1ms/step - loss: 0.8586
0.8585554361343384
```

Both uses softsign as activation fcn


activation
units
train.epoch_loss
validation.epoch_loss
validation.evaluation_loss_vs_iter
softsign
softplus
softmax
sigmoid
selu
relu
exponential
elu

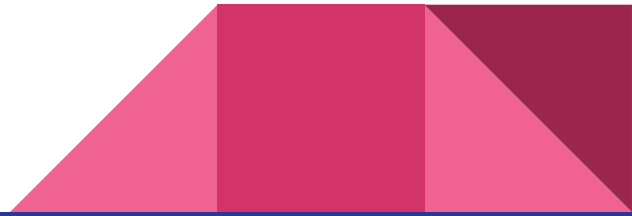

units
activation
train.epoch_loss
validation.epoch_loss
validation.evaluation_loss_vs_iter
tanh
softsign
softplus
sigmoid
relu
elu





Future work

- Deploy the code on android
- Object recognition
- Dynamic classification
- Remove false classifications



Questions?