Motor AI – Technical Challenge

Objective:

To control the Mountain-Car-v0 Gym Environment using webcam.

Dataset Collection:

There are 3 discrete deterministic actions for the mountain car:
1. Accelerate to the left
2. Don't accelerate
3. Accelerate to the right

So, we need 3 gestures for controlling the movement of mountain car.

The Dataset collected is as follows:
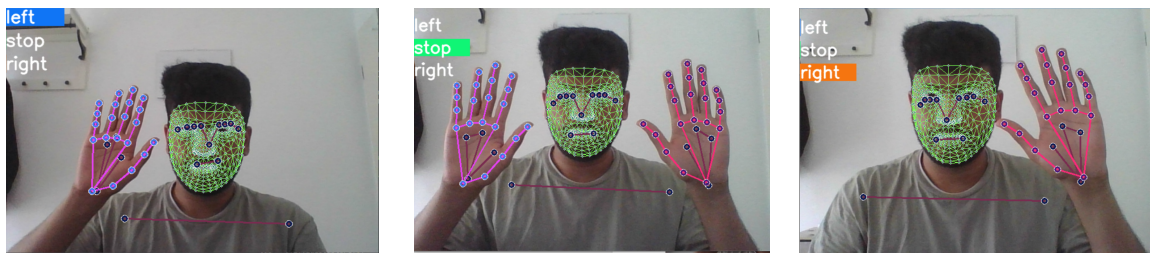
1.Left
2.Stop
3.Right

Fig 1: Gestures

- The key points for face, hand is estimated using mediapipe model and stored in NumPy array for data collection.
- The three gestures were chosen in such a way that key points are relatively distinguishable for good accuracy.
- The Training Dataset was constructed with 30 videos for each gesture.

Model Construction

Constructing a LSTM model with least parameters since the model inference time is kept in priority.

Optimal model is constructed using following methods:
- Bayesian Optimization
- Hyperband

This way model is automatically constructed after iterating over many trails and best model with configured with optimal number of layers, hidden units and learning rate.
Once the best model is found, it is further trained and evaluated with cross validation approach.

Model Evaluation

The model is evaluated with the test dataset and got the following metrics:
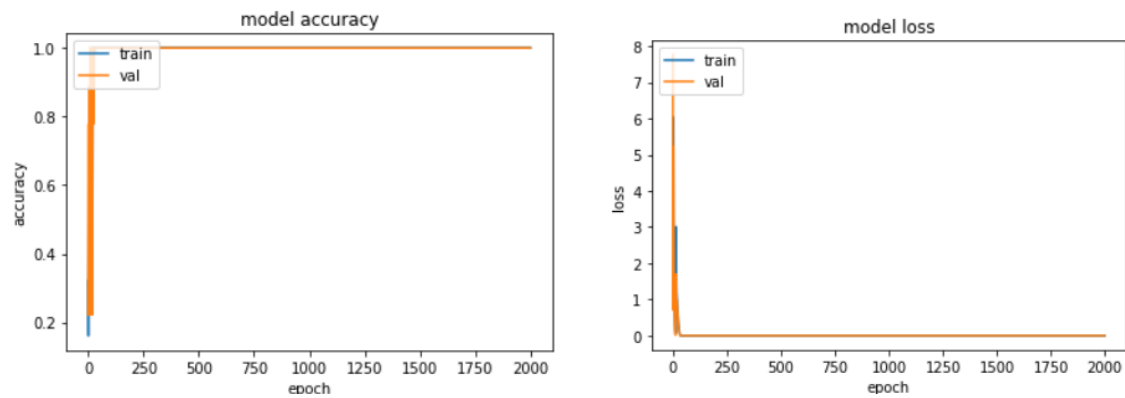
- Accuracy – 1.0
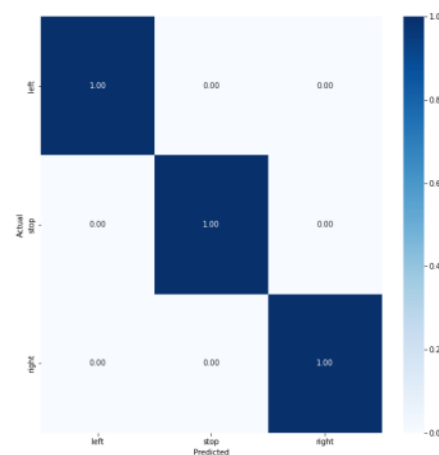- Recall score – 1.0

Fig 2: Accuracy and Loss Curves

Fig 3: Confusion Matrix

- Model can be further evaluated based on Cross validation and Leave one subject out (LOSO) methods.

Model Quantization

Quantization is applied to speed up the model at the inference stage by converting 32-bit floating point numbers to 8-bit integers.

It can be done using:
- Tensorflow Lite
- OpenVino

<u>Model Introspection</u>

We can visualize and debug the trained Keras neural net models using tf-keras-vis.
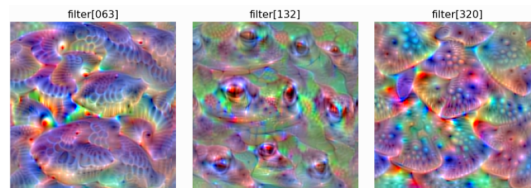Following are the few pictures showcasing the Activation Maximization of VGG16 model

Fig 4: Model Visualization

<u>Challenges Faced</u>

- Choosing Gestures which are easily differential from each other.
- Constructing the model with least parameter for faster inference results.