**Gesture recognition project** has different video frames containing gestures to detect certain predefined actions.

**Submitters -**

[**Gokulnath.nitt@gmail.com**](mailto:Gokulnath.nitt@gmail.com)

[**karthik.gadepalli@gmail.com**](mailto:karthik.gadepalli@gmail.com)

**Runtime -**

We have used GPU from JarvisLabs

**Summary -**

We have created 2 models -

**Model 1 -**

We have used Conv3D and MaxPooling3D

**Model 2 -**

We have applied Conv2D and LSTM

**Generator -**

Generator code has been added to handle the image preprocessing and provide in a batch for the CNN and RNN networks.

We have also added code to test the generator sample.

**Model Validation and Testing -**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model Name | Architecture | Number of Epochs | Train\_accuracy | Val\_accuracy |
| Model | Conv3D + MaxPooling3D | 20 | 80.51% | 35.94% |
| Model 2 | Conv2D + LSTM | 20 | 54.55% | 32.81% |

**Final Summary / Findings -**

1. Conv3D + MaxPooling3D has better validation accuracy when compared to CNN + RNN architecture.
2. Number of Epochs used on Conv3D used was only 20 as there are no significant improvements in validation accuracy.
3. Number of Epochs remained same as 20 for the RNN architecture.
4. Generators were used to feed the images in batches
5. Dropout layers are utilized to improve the overall accuracy
6. This model can further be enhanced by playing around the number of epochs, batch\_sizes, number of images used in the training, augmentation can also be used to increase the training data, image resolution can also be increased in the preprocessing step.