**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 | 10 | 98.72% | 22.66% |
| Model 2 | Conv2D + LSTM | 30 | 99.70% | 61.72% |

**Final Summary / Findings -**

1. CNN + RNN has better validation accuracy when compared to Conv3D architecture.
2. Number of Epochs used on Conv3D used was only 10 as there are no significant improvements in validation accuracy.
3. Number of Epochs increased to 30 for the RNN architecture as it improves the validation accuracy.
4. Generators were used to feed the images in batches
5. Dropout layers are utilized to improve the overall accuracy
6. We were getting errors while calling the callback functions to generate the .h5 files. So they are skipped when training the model.