The below is the model execution write up for Gesture Recognition Experiment on various iterations to determine the best model.

|  |  |  |  |
| --- | --- | --- | --- |
| Experiment Number | Model | Result | Decision + Explanation |
| 1 | **Conv3D** | **Training Acc: 20.81%**  **Val Acc: 29%** | Model is not performing well because of huge image width and height, many dropouts which is causing the model to discard things and learning rate is 0.01 |
| 2 | **Conv3D** | **Training Acc: 76.17%**  **Val Acc: 27%** | Heavy Overfitting obtained. Batch size = 128 is not optimal for validation set. Using more Dropouts to reduce overfitting and also reducing batch size |
| 3 | **Conv3D** | **Train Acc: 88.23%**  **Val Acc: 40.99%** | Reducing number of frames and adding dropout and batchnormalization() and modifying learning rate to 0.002 |
| 4 | **Conv3D** | **Train Acc: 89.59**  **Val Acc: 55%** | Increasing the number of epochs = 30. Changing Max Pooling Size to 1,2,2. Removing the learning rate for the model to decide the best value. Reducing the batch size to 32 |
| 5 | **Conv3D** | **Train Acc: 96.22%**  **Val Acc: 76.99%** | Though after the hyper parameter turning the accuracy has not increased so, tweaked the image size due to which training acc and val acc are getting better and closer with a simple Conv3D model |
| 6 | **Conv3D** | **Train Acc: 94.72%**  **Val Acc: 72%** | Added Pool Size and Strides to MaxPooling and increased number of frames to 30 and epochs = 30 |
| 7 | **Conv3D** | **Train Acc: 94.72%**  **Val Acc: 73%** | Modifying the model kernel size and adding dropouts to the existing model |
| 8 | **ConvLSTM** | **Train Acc: 97.43%**  **Val Acc: 69.99%** | Building a CNN Model with Mobilenet and LSTM integration |
| 9 | **ConvLSTM** | **Train Acc: 98.03%**  **Val Acc: 79.00%** | Building a CNN Model with Conv2D with 16,32,64,128,256 and LSTM 256 with batch size of 20, n\_frames = 18 and epochs 30 |
| 10 | **ConvLSTM** | **Train Acc: 98.19%**  **Val Acc: 80.00%** | Using the above model with 18 frames, 84\*84 size images and LSTM 128 with only one batchnormalization have better efficiency |
| 11 | **ConvGRU** | **Train Acc: 96.38%**  **Val Acc: 81.99%** | Using the above model with 18 frames, 84\*84 size images and GRU 128 with only one batchnormalization have better efficiency |
| Final Model | **ConvGRU** | **Train Acc: 96.38%**  **Val Acc: 81.99%** | **Based on various parameters this model performance is more effective** |