**Gesture Recognition:**

Target class: 5

Data point: Each video has 30 frames

Experimented models:

1. 3D Convolutional Neural Networks
2. CNN + RNN

Hyperparameters:

Batch size, number of frames , image resolution

Observations:

Image size most suited (120,120,3)

Number of frames in range(15-30)

|  |  |  |  |
| --- | --- | --- | --- |
| **Experiment Number** | **Model** | **Result** | **Observations + Explainations** |
| **1** | **Conv3D** | **Complex model with with multiple layers and**  **Train\_accuracy : 0.8**  **Val\_accuracy:0.2** | **Classic case of overfitting** |
|  |  | **Varied frame count(reduced)**  **Train\_accuracy : 0.88**  **Val\_accuracy:0.2** | **Not much impact** |
|  |  | **Train\_accuracy : 0.81**  **Val\_accuracy:0.2** | **Decreased batch size**  **Testing with uneven batch size so number of batches will be not a factor of dataset count** |
| **2** | **Conv3D** | **Added extra layers**  **Train\_accuracy : 0.75**  **Val\_accuracy:0.15** | **Reduction in loss**  **Loss performance is better** |
| **3** | **Conv3D** | **Added regularization to model 2**  **Dropout =0.5**  **Accuracy: 0.4** | **Extra regularized**  **Under performing model** |
|  |  | **Less dropout**  **Dropout = 0.25**  **Accuracy : 0.62** | **Slight performance improvement over last iteration** |
| **4** | **Conv3D** | **Train\_accuracy : 0.75**  **Val\_accuracy:0.15** | **Different activation function for faster convergence**  **Increased images count**  **Increase in training loss** |
| **5** | **Conv3D** | **Train\_accuracy : 0.89**  **Val\_accuracy:0.80** | **Fresh model without batch normalization and less number of layers**  **For accuracy improvement**  **Model handled overfitting well** |
|  | **Conv3D** | **Train\_accuracy : 0.99**  **Val\_accuracy:0.80** | **Added more kernels as well as epoch size for model to learn dataset better** |
|  | **Conv3D** | **Train\_accuracy : 089**  **Val\_accuracy:0.86** | **Increased dropout to regularize**  **Optimal batch size** |
| **6** | **Conv3D** | **Train\_accuracy : 0.92**  **Val\_accuracy:0.80** | **Added single dropout layer**  **To model 6**  **Same loss performance as model 6** |
| **7** | **CNN+ LSTM** | **Train\_accuracy : 0.74**  **Val\_accuracy:0.15** | **Base model to check**  **Need to check with increased epochs and more number of dense and LSTM cells** |
|  | **CNN + LSTM** | **Train\_accuracy : 0.82**  **Val\_accuracy:0.23** | **Epochs =10**  **Loss has stabilized**  **Training performance improved** |
| **8** | **CNN + GRU** | **Train\_accuracy : 0.82**  **Val\_accuracy:0.26** | **Model has performed better than LSTM with less number of epochs** |
| **9** | **CNN + Simple RNN** | **Train\_accuracy : 0.76**  **Val\_accuracy:0.20** | **Longer time taken for model compilation** |
| **Final Model** | **CNN3D**  **Model 5\_3** | **Train\_accuracy : 0.89**  **Val\_accuracy:0.86** | **Best train and validation performance** |