EXPERIMENTS

|  |  |  |  |
| --- | --- | --- | --- |
| Experiment Number | Model | Result | Decision + Explanation |
| 1 | Model 1:Conv\_3d1  (Conv3d) | Train\_accuracy: 0.86, Val. Acc= 0.72 Model is Overfit | Tried building the model with batch size greater than 32 but due to the GPU RAM limitations, OOM error raised up. Hence settled down with batch size 20 but ran with little higher batch size of 24 in the consequent runs.  IMG\_SIZE = 120\*120, batch size =20, images\_per\_video = 20, Epochs=20. Added dropouts in the dense layer as model was overfitting significantly. With this dropout, overfit is reduced. But accuracy as well dropped  Increasing Epochs above 10 started showing significant increase in Accuracy. Will add LSTM layer to see if this improves |
| 2 | Model 2 : rnn\_cnn1  (conv2d + LSTM) | Train\_accuracy: 0.86, Val. Acc= 0.70 Model is Overfit | Reduced the image size to 100\*100. Added LSTM layer with CONV2d + LSTM combination. No of trainable params also increased with LSTM addition.  2D CNN + RNN with Batch Size - 20, IMG SIZE 100x100, Images per Video - 20, Epochs – 20  Will add Transfer learning to see if this will help |
| 3 | Model 3: rnn\_cnn\_tl  (TL+2D cnn + LSTM) | Train\_accuracy: 0.99, Val. Acc= 0.81 Model is Overfit | Used Transfer learning by taking Mobilenet architecture on the top of conv2d + rnn(LSTM) architecture. Also Changing image size = 120\*120 Train accuracy is improved. Validation acc needs to be improved  Transfer Learning + 2D CNN + RNN(LSTMs) with Batch Size - 20, IMG SIZE 120x120, Images per Video - 20, Epochs - 20 |
| 4 | **Model 4: rnn\_cnn\_tl2**  **(TL + 2D CNN + GRU)** | Model is performing well.  **Train\_accuracy: 0.9925, Val. Acc= 0.97** | Used Transfer learning with Mobilenet architecture on the top of conv2d + rnn(GRU).  Transfer Learning + 2D CNN + RNN(GRUs) with Batch Size - 16, IMG SIZE 80x80, Images per Video - 20, Epochs - 17 |
| 5 | Model 5: conv\_3d1  (3d cnn) | Train\_accuracy: 0.96. Val. Acc= 0.90  No overfit. Improved accuracy compared to base model 1 | No overfitting. Increased no of epoch on base model conv3d. Increased image size and experimented with batchsize  3D CNN with Batch Size - 24, IMG SIZE 160x160, Images per Video - 20, Epochs - 30 |
| 6 | Model 6: rnn\_cnn1  Conv2d + rnn(LSTM) | Train\_accuracy: 0.89, Val. Acc= 0.80  Improved accuracy compared to model 2 | 2D CNN + RNN(LSTMs) with Batch Size - 24, IMG SIZE 160x160, Images per Video - 20, Epochs – 30. Increased epochs/image size compared to model 2 |
| 7 | Model 7: Rnn\_cnn\_tl2 model | Train\_accuracy: 0.99, Val. Acc= 0.70. Model is Overfit. | Transfer Learning + 2D CNN + RNN(LSTMs) with Batch Size - 24, IMG SIZE 120x120, Images per Video - 20, Epochs – 30.  Increased epochs/image size compared to model 3 |
| 8 | Model 8: nn\_cnn\_tl2 model(TL + 2D CNN + GRU) | Train\_accuracy: 0.997, Val. Acc= 0.95 | Transfer Learning + 2D CNN + RNN(GRUs) with Batch Size - 24, IMG SIZE 120x120, Images per Video - 20, Epochs – 17  Increased epochs/image size compared to model 4 |
| Final Model | Model 4: rnn\_cnn\_tl2  (TL + 2D CNN + GRU) | Train\_accuracy: 0.9925, Val. Acc= 0.97 | Choosing this model as this has both high train and val accuracies.  TL is able to extract most of the feature required for this gesture experiment compared to other models. |