**Gesture Recognition model development experiments**

**Best Model: InceptionV3+LSTM – Train Accuracy: 0.984 and Val Accuracy 0.964**

Best model (26) is highlighted in green color

Other notable models are highlighted in blue color

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Exp# | Model | Layers | Model Parameters | Results (Accuracy) | Actions and Remarks |
| 1 | Conv3D | 4 CNN layers  (CONV+Maxpooling + BN) | num\_frames = 10 Batch\_Size = 32 Epochs = 1 filter\_size = (3,3,3) optimizer = ADAM | Model Compilation Error | Since num\_frames = 15, using (3,3,3) filter reduced the dimension (x) to negative, hence reduced the CNN layers to 3, increase num\_frames to 15 |
| 2 | Conv3D | 3 CNN layers  (CONV+Maxpooling + BN) | num\_frames = 15 Batch\_Size = 32 Epochs = 20 filter\_size = (3,3,3) Tot\_param = 3.758mn LRonPlatue(factor=0.1, patience=10) optimizer = ADAM | Train Acc = 0.957 Val Acc = 0.625 | Reduced the N/W depth Increased num\_frames Increased Epochs  Bias = Low Variance = high Overfitting |
| 3 | Conv3D | 3 CNN layers  (CONV+Maxpooling + BN) | Num\_frames = 15 Batch\_Size = 64 Epochs = 20 filter\_size = (5,5,5) Tot\_param = 4.041mn LRonPlatue(factor=0.1, patience=5) optimizer = ADAM | Train Acc = 0.818 Val Acc = 0.625 | Same the N/W depth Increased filter\_size Increased Batch\_Size Increased Epochs  Bias = moderatly low Variance = high Some overfitting |
| 4 | Conv3D | 3 CNN layers  (CONV+Maxpooling + BN) | num\_frames = 15 Batch\_Size = 64 Epochs = 20 filter\_size = (5,5,5) FC units = 512 Tot\_param = 7.702mn LRonPlatue(factor=0.2, patience=5) optimizer = SGD() | Train Acc = 1.0 Val Acc = 0.875 | Changed the optimizer function to SGD  Model is overfitting |
| 5 | Conv3D | 3 CNN layers  (CONV+Maxpooling + BN) | num\_frames = 15 Batch\_Size = 64 Epochs = 30 filter\_size = (5,5,5) FC units = 512 Tot\_param = 7.702mn LRonPlatue(factor=0.2, patience=5) optimizer = ADAM | Train Acc = 0.963 Val Acc = 0.875 | changed LRonPlatue parameters Increased Epochs Increased FC units Same filter\_size Same Batch\_Size  Bias = low Variance = some reduced overfitting |
| 6 | Conv3D | 3 CNN layers  (CONV+Maxpooling + BN) | num\_frames = 15 Batch\_Size = 64 Epochs = 50 filter\_size = (5,5,5) FC units = 512 Tot\_param = 7.702mn LRonPlatue(factor=0.2, patience=5) optimizer = ADAM | Train Acc = 0.973 Val Acc = 1. | Increased Epochs   Bias = low Variance = some reduced overfitting |
| 7 | Conv3D | 3 CNN layers  (CONV+Maxpooling + BN) | num\_frames = 15 Batch\_Size = 64 Epochs = 50 filter\_size = (5,5,5) FC units = 512, 256 Tot\_param = 8.374mn LRonPlatue(factor=0.2, patience=5) optimizer = ADAM | Train Acc = 0.856 Val Acc = 1. | Changed the atchitecture 2 CNN layers followed by Max Pooling and BN layer  2FC layers of 512 and 256 units  Something not right?? Val Acc is much higher than Train Acc |
| 8 | Conv3D | 3 CNN layers  (CONV+Maxpooling + BN) | num\_frames = 18 Batch\_Size = 64 Epochs = 50 filter\_size = (5,5,5) FC units = 256 Tot\_param = 7.7mn LRonPlatue(factor=0.2, patience=5) optimizer = ADAM | Train Acc = 0.824 Val Acc = 0.875 | 3x Conv+pooling+BN layers 1FC layers of 256 units  Bit of underfitting |
| 9 | Conv2D+GRU | 3 CNN layers  (CONV+Maxpooling + BN) + GRU(512) + FC(512) | num\_frames = 15 Batch\_Size = 64/32 Epochs = 1 filter\_size = (3,3) Tot\_param = 22.477mn LRonPlatue(factor=0.2, patience=5) optimizer = ADAM | OOM Error | Need to reduce the network depth and other parameters |
| 10 | Conv2D+GRU | 3 CNN layers  (CONV+Maxpooling + BN) + GRU(128) + FC(128) | num\_frames = 15 Batch\_Size = 16 Epochs = 30 filter\_size = (3,3) Tot\_param = 11.22mn LRonPlatue(factor=0.2, patience=5) optimizer = ADAM | Train Acc = 0.804 Val Acc = 0.75 | Batch\_size reduced to 16 Reduced GRU units Reduced FC units  There is opportunity to improve model performance |
| 11 | Conv2D+GRU | 3 blocks of CNN layers  (2CONV+Maxpooling + BN) + GRU(128) + FC(256) | num\_frames = 15 Batch\_Size = 32 Epochs = 50 filter\_size = (5, 5) Tot\_param = 11.915mn LRonPlatue(factor=0.2, patience=5) optimizer = ADAM | Train Acc = 0.91 Val Acc = 0.81 | Modified the network - 2Conv layers followed by BN and Maxpooling layer Increased filter size Increased Epochs to 50 Reduced dropout to 0.4 from 0.5 Both Train and Val performance improved |
| 12 | Conv2D+GRU | 3 blocks of CNN layers  (2CONV+Maxpooling + BN) + GRU(128) + FC(512) | num\_frames = 15 Batch\_Size = 32 Epochs = 50 filter\_size = (5, 5) Tot\_param = 11.949mn LRonPlatue(factor=0.2, patience=5) optimizer = ADAM | Train Acc = 0.92 Val Acc = 0.875 | Increased FC units to 512 Reduced dropout from 0.4 to 0.3 Training and validation accuracy increased |
| 13 | Conv2D+GRU | 3 blocks of CNN layers  (2CONV+Maxpooling + BN) + GRU(128) + FC(512) | num\_frames = 15 Batch\_Size = 32 Epochs = 50 filter\_size = (5, 5) Tot\_param = 11.949mn LRonPlatue(factor=0.2, patience=5) optimizer = ADAM(lr=0.0001) | Train Acc = 1. Val Acc = 0.857 | No droput Training and validation accuracy increased   Model Overfitting |
| 14 | Conv2D+GRU | 3 blocks of CNN layers  (2CONV+Maxpooling + BN) + GRU(128) + FC(512) | num\_frames = 15 Batch\_Size = 32 Epochs = 50 filter\_size = (5, 5) Tot\_param = 11.949mn LRonPlatue(factor=0.2, patience=5) optimizer = ADAM(lr=0.0001) | Train Acc = 1. Val Acc = 0.893 | Added dropout 0.2 Val Accuracy increased  Model still Overfits |
| 15 | Conv2D+GRU | 3 blocks of CNN layers  (2CONV+Maxpooling + BN) + GRU(128) + FC(512) | num\_frames = 15 Batch\_Size = 32 Epochs = 50 filter\_size = (5, 5) Tot\_param = 11.949mn LRonPlatue(factor=0.2, patience=5) optimizer = ADAM(lr=0.0001) | Train Acc = 1. Val Acc = 0.929 | Increased Dropout to 0.4 Val accuracy increased  Model overfits |
| 16 | Conv2D+GRU | 3 blocks of CNN layers  (2CONV+Maxpooling + BN) + GRU(128) + FC(512) | num\_frames = 15 Batch\_Size = 32 Epochs = 50 filter\_size = (5, 5) Tot\_param = 11.949mn LRonPlatue(factor=0.2, patience=5) optimizer = ADAM(lr=0.0001) | Train Acc = 0.937 Val Acc = 0.964 | Used 2 Droput layers (0.2 and 0.6) Reduced Overfitting |
| 17 | Conv2D+LSTM | 3 blocks of CNN layers  (2CONV+Maxpooling + BN) + LSTM(128) + FC(256) | num\_frames = 15 Batch\_Size = 32 Epochs = 50 filter\_size = (5, 5) Tot\_param =15.617mn LRonPlatue(factor=0.2, patience=5) optimizer = ADAM | Train Acc = 0.829 Val Acc = 0.875 | Changed the RNN algo from GRU to LSTM  Train accuracy droped |
| 18 | Conv2D+LSMT | 3 blocks of CNN layers  (2CONV+Maxpooling + BN) + LSTM(128) + FC(512) | num\_frames = 15 Batch\_Size = 16 Epochs = 50 filter\_size = (5, 5) Tot\_param = 15.651mn LRonPlatue(factor=0.2, patience=5) optimizer = ADAM(lr=0.0005) | Train Acc = 0.92 Val Acc = 0.875 | Increased the FC units to 512 reduced the dropout to 0.2 reduced the batch size to 16 to avoid OOM |
| 19 | Conv2D+LSMT | 3 blocks of CNN layers  (2CONV+Maxpooling + BN) + LSTM(128) + FC(256) | num\_frames = 15 Batch\_Size = 16 Epochs = 50 filter\_size = (5, 5) Tot\_param = 15.617mn LRonPlatue(factor=0.2, patience=5) optimizer = ADAM(lr=0.0001) | **Train Acc = 0.984 Val Acc = 0.929** | Reduce the Initial LR to 0.0001 No dropout used Training and Validation Accuracy Increased |
| 20 | InceptionV3+GRU | InceptionV3 + GRU(128) | num\_frames = 15 Batch\_Size = 32 Epochs = 1 filter\_size = (5, 5, 5) Tot\_param = 22.430mn LRonPlatue(factor=0.2, patience=5) optimizer = ADAM(lr=0.0005) | OMM | Need to reduce the Model/ Size and Batch\_size |
| 21 | InceptionV3+GRU | InceptionV3 + GRU(64) | num\_frames = 15 Batch\_Size = 16 Epochs = 30 filter\_size = (5, 5, 5) Tot\_param = 22.430mn LRonPlatue(factor=0.2, patience=5) optimizer = ADAM() | Train Acc = 0.67 Val Acc = 0.5 | Reduced the batch\_size to 16 Model underfits |
| 22 | InceptionV3+GRU | InceptionV3 + GRU(128) | num\_frames = 15 Batch\_Size = 16 Epochs = 30 filter\_size = (5, 5, 5) Tot\_param = 22.492mn LRonPlatue(factor=0.2, patience=5) optimizer = ADAM() | Train Acc = 0.865 Val Acc = 0.786 | Incresed GRU units to 128 No dropout Layer Both Train and Val accuracies improved |
| 23 | InceptionV3+GRU | InceptionV3 + GRU(128) | num\_frames = 15 Batch\_Size = 16 Epochs = 30 filter\_size = (5, 5, 5) Tot\_param = 22.492mn LRonPlatue(factor=0.2, patience=5) optimizer = ADAM() | Train Acc = 1. Val Acc = 0.964 | Reduced the initial LR to 0.0001 Val Accuracy increased  Mode overfits |
| 24 | InceptionV3+GRU | InceptionV3 + GRU(128) | num\_frames = 15 Batch\_Size = 16 Epochs = 30 filter\_size = (5, 5, 5) Tot\_param = 22.492mn LRonPlatue(factor=0.2, patience=5) optimizer = ADAM(0.0001) | **Train Acc = 0.984 Val Acc = 0.929** | Increased Epochs to 40 Reduced Initial LR to 0.0001 Introduced 2 Dropout layers (0.2, 0.5) |
| 25 | InceptionV3+LSTM | InceptionV3 + LSTM(128) | num\_frames = 15 Batch\_Size = 16 Epochs = 30 filter\_size = (5, 5, 5) Tot\_param = 22.524mn LRonPlatue(factor=0.2, patience=5) optimizer = ADAM(0.0001) | **Train Acc = 0.992 Val Acc = 0.929** | Changed the RNN model from GRU to LSTM No Dropout Layer  Train Accuracy increased Val Accuracy remained same  Model overfits |
| 26 | InceptionV3+LSTM | InceptionV3 + LSTM(128) | num\_frames = 15 Batch\_Size = 16 Epochs = 30 filter\_size = (5, 5, 5) Tot\_param = 22.524mn LRonPlatue(factor=0.2, patience=5) optimizer = ADAM(0.0001) | **Train Acc = 0.984 Val Acc = 0.964** | Added a dropout layer(0.4)  Reduced overfitting Both Train and val accuracies are good |