|  |  |
| --- | --- |
| **Model 1** | **Decision + Explanation** |
| **Conv3D** | **Chose batch size 64 , it gave Out of memory error** |
|  | **Chose batch size 32, It took a long time to respond** |
|  | **Chose batch size 16, executed successfully** |
| **Conv3D** | **SGD Optimizer used with learning rate of 0.1, got the model summary as :**  **Total params: 9,440,773**  **Trainable params: 9,439,365**  **Non-trainable params: 1,408** |
| **Conv3D** | **By the 14th epoch(total 30 epochs) got the max accuracy of 76%** |
| **Model 2** | **Decision + Explanation** |
| **CNN LSTM** | **Got the model Summary as :**  **Total params: 2,573,541**  **Trainable params: 2,573,061**  **Non-trainable params: 480** |
| **ConvLSTM** | **We get the traing accuracy of app. 80 %** |
| **Final Model** | **We choose Model 1 using Conv3D as :**  Even though the parameters are more in Conv3D the model look much lesser time than CNN RNN models to finish with a comparable accuracy |

We have decided to select Conv 3D model with batch size 16 and accuracy 76% percent as the final model. This is because the model took comaparatively very less time to run with the give GPU specifications. Also since its basic gesture recognition Conv 3d can provide fairly good results with a large dataset. As you can see the best accuracy which we got for CNN LSTN model is 80% which is better than Con3D but with more epochs we can further increase the accuracy of it and make it close to the CNN-RNN network with lesser training time.