Gesture Recognition model building approach:

We had total 663 video sequences for training and 100 for validation.

Video frames were of (360X360X3) and (160X120X3) dimensions.

All video frames were preprocessed to same dimensions and normalised.

Table 1-1

Experiment Number	Model	Result	Decision + Explanation	
1	**Conv3D**	**Training Acc: 92.7 Valid Acc: 85**	**Frames-20, Epoc-25, batch-40 - increasing layers to increase efficiency**	
2	**Conv3D**	**Training Acc: 92.7 Valid Acc: 80**	**Frames-20, Epoc-30, batch-40 - adding more layers and increasing number of epoc to increase effeciency**	
3	**Conv3D**	**Training Acc: 53 Valid Acc: 40**	**Reduce batch size to make model fast**	
4	**Conv3D**	**Training Acc: 91.3 Valid Acc: 86.6**	**Frames-20, batch-40, image size-100X100 and reduced parmeters**	Best Conv model
5	**Conv3D**	**Training Acc: 86.5 Valid Acc: 88.3**	**Frames:24, Epocs-20, BS-40, switching BatchNormalization before MaxPooling and reducing trainable parameters**	
6	**ConvLSTM**	**Training Acc: 66.6 Valid Acc: 34.9**	**CNN and LSTM architecture**	
7	**MobileNet with LSTM**	**Training Acc: 100.0 Valid Acc: 99.6**	**Transfer learning**	Best model with Transfer learn
Final Model	**Conv3D**	**Training Acc: 91.3 Valid Acc: 86.6**	**Frames-20, batch-40, image size-100X100 and reduced parmeters**	