Gesture Recognition – Deep learning

Problem Statement:

We need to develop a cool feature in the smart-TV that can recognise five different gestures performed by the user which will help users control the TV without using a remote.

The following table consists of the experiments done to build a model to predict the gestures from the given data set

Experiment Number	Model	Result	Decision + Explanation
1	Conv3D	Throws Generator error	Image shape was different. Due to different dimensions model was throwing error
2	Conv3D	Model is overfitting, val_losses were too high	Augmentation was introduced in generator method to overcome overfitting
3	Conv3D	Accuracy: 0.26, model stopped learning	Increase the amount of trainable data/ reduce the filter size
4	Conv3D	Training accuracy improved significantly, val_acc remains less	Removed few layers in the next model to simplify the model
5	Conv3D	Observed increase in val_acc, but model can still be improved	Check the model accuracy for Conv2D, LSTM
6	Time Distributed + Dense	Train Accuracy: 0.83, Validation Accuracy: 0.81	Model seems satisfactory, changing Batch Size
7	Time Distributed + Dense(Increase in Batch Size)	Train Accuracy: 0.8780, Validation Accuracy: 0.8750	This is good model with training and validation accuracies with number of params 128,517
Final Model	Time Distributed + ConvLSTM 2D	Observed max training accuracy: 0.8515 & val_accuracy: 0.8100	Best model observed in the complete experiment

Conclusion:

The Model built with Time distributed Conv2D and ConvLSTM2D (Final model) gave better results compared to all the other models.