**Gesture Recognition – Mukul**

**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.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Exp. #** | **Model** | **Hyperparameters** | **Result** | **Decision + Explanation** |
| **1** | **Conv3D** | No of frames = 16  No. epochs = 20  Batch Size = 64 | Max. Training Accuracy 0.903743326663971  Max. Validation Accuracy 0.625 | The Model is learning but while Validating we are seeing the gaps trying to reduce the gap |
| **2** | **Conv3D** | No of frames = 30  No. epochs = 30  Batch Size = 64 | Max. Training Accuracy 0.9197860956192017  Max. Validation Accuracy 0.875 | There is slight in improvement in training accuracy and the validation is as is so still try to reduce the gap |
| **3** | **Conv3D** | No of frames = 30  No. epochs = 20  Batch Size = 64 | Max. Training Accuracy 0.8877005577087402  Max. Validation Accuracy 0.625 | Reduced parameters  There is dip in training accuracy and the validation is as is so still try to reduce the gap |
| **4** | **Conv3D** | No of frames = 20  No. epochs = 20  Batch Size = 64 | Max. Training Accuracy 0.8074866533279419  Max. Validation Accuracy 0.5 | Same pooling in Conv3D layer  There is dip in training accuracy and the validation is as is so still try to reduce the gap |
| **5** | **Conv3D** | No of frames = 20  No. epochs = 20  Batch Size = 64 | Max. Training Accuracy 0.9411764740943909  Max. Validation Accuracy 0.375 | Reduced kernel to (2,2,2), switching Batch Normalization before Max Pooling  There is dip in training accuracy and the validation is as is so still try to reduce the gap |
| **6** | **Conv2D+LSTM** | No of frames = 20  No. epochs = 30  Batch Size = 32 | Max. Training Accuracy 0.918767511844635  Max. Validation Accuracy 0.6875 | Switching Model architecture to Conv2D+LSTM  The Model is learning but while Validating we are seeing the gaps trying to reduce the gap |
| **7** | **Time**  **Distributed**  **+**  **ConvLSTM**  **2D** | No of frames = 20  No. epochs = 30  Batch Size = 64 | Max. Training Accuracy 1.0  Max. Validaiton Accuracy 1.0 | **Transfer Learning (MobileNet) with LSTM**  **This is the best model so far we can get. The validation accuracy is good .** |

**Conclusion:**

The Model built with Time distributed Conv2D and ConvLSTM2D (Experiment #7) gave better results compared to all the other models and also the model has very least number of parameters compared to other models.