We have tried to implement two different kinds of model

1. Conv3D
2. Conv + RNN/GRU
3. Transfer Learning + GRU

|  |  |  |  |
| --- | --- | --- | --- |
| **Experiment Number** | **Model** | **Result** | **Decision + Explanation** |
| **1** | **Conv3D** | **Accuracy: 0.7156**  **Val Accuracy: 0.5143** | **Using smaller size of images for each video to see how the model fits.**  **This is a sample model**  **Model is overfitting as there are lots of parameters to train.** |
| **2** | **Conv3D** | **Accuracy: 0.71**  **Val Accuracy: 0.58** | **Increasing the resolution of the images used and keeping the model same as above**  **Dropped the no. of params to train.  Model is still overfitting** |
| **3** | **Conv3D** | **Accuracy: 0.96**  **Val Accuracy: 0.67** | **Changing the filter size to 3x3x3 and increasing the dense neurons to 256  Model is still overfitting** |
| **4** | **Conv3D** | **Accuracy: 0.90**  **Val Accuracy: 0.71** | **Keeping model same as above, but increasing the dropout rate to 0.3** |
| **5** | **CNN + LSTM** | **Accuracy: 0.7529**  **Val Accuracy: 0.63** | **Moved to Conv2D model will GRU Layers** |
| **6** | **Transfer Learning  imagenet + GRU** | **Accuracy: 0.9157**  **Val Accuracy: 0.89** | **Using transfer learning along with GRU** |

The model where pretrained model from imagenet is used gave us the accuracy of 91% and the validation accuracy of 89%. This shows this architecture can be improved further if more samples are given. One way to achieve this can be image augmentation that will improve no. of images for the training in different ways.