Key Points -

* Set the number of images used for each video to 16
* Image\_height and image\_width set to 160 x 160
* Batch\_size set to 64
* Using P100 GPU on Kaggle with 16GB VRAM
* Each model trained for 50 epochs except ResNet
* Model checkpoint at every epoch

|  |  |  |  |
| --- | --- | --- | --- |
| **Experiment Number** | **Model** | **Accuracy** | **Insights** |
| **1** | **Conv3D** | **80% (Best Model)** | Training loss steadily decreases while validation loss fluctuates, picked the model with highest validation accuracy |
| **2** | **Conv2D + GRU** | **57.50%** | Both training and validation loss gradually decrease, but the model is not able to predict well on validation data. |
| **3** | **TimeDistributed + ConvLSTM2D** | **28.75%** | Model reached the maximum validation accuracy of 62.50%, validation loss increased and then decreased |
| **4** | **Transfer Learning using ResNet** | **62.50%** | ResNet152 used after freezing weights; used ‘imagenet’ weights; added TimeDistributed Dense and GRU layers to be fine-tunable |

