## **Deep Learning Course Project- Gesture Recognition**

# Group Members

* **Prateek Bhattu**
* **Rishi Garhwal**

# Problem Statement

As a data scientist at a home electronics company which manufactures state of the art smart televisions. We want 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.

* Thumbs up :  Increase the volume.
* Thumbs down : Decrease the volume.
* Left swipe : 'Jump' backwards 10 seconds.
* Right swipe : 'Jump' forward 10 seconds.
* Stop : Pause the movie.

# Objective

Our task is to train different models on the 'train' folder to predict the action performed in each sequence or video and which performs well on the 'val' folder as well. The final test folder for evaluation is withheld - final model's performance will be tested on the 'test' set.

# Data Generator

This is one of the most important part of the code. In the generator, we are going to pre-process the images as we have images of 2 different dimensions (*360 x 360* and *120 x 160*) as well as create a batch of video frames. The generator should be able to take a batch of videos as input without any error. Steps like cropping, resizing and normalization should be performed successfully.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Experiment Number** | **Model** | **Result** | **Decision + Explanation** | **Parameters** |
|  | **Conv3D** | **Error** | **Reduce the batch size and reduce the number of neurons in Dense layer** | **1,736,389** |
|  | **Conv3D** | **Training Accuracy :0.96**  **Validation Accuracy: 0.93** | **Overfitting**  **Reduce the number of epochs** | **504,709** |
|  | **Conv3D** | **Training Accuracy :0.89**  **Validation Accuracy: 0.39** | **Overfitting** | **1,736,389** |
|  | **Conv3D** | **Training Accuracy :0.78**  **Validation Accuracy: 0.41** | **Early stopping** | **2556533** |
|  | **Conv3D** | **Training Accuracy :0.84**  **Validation Accuracy: 0.69** | **Validation accuracy is low** | **1657445** |
|  | **Conv-LSTM** | **Training Accuracy :0.99**  **Validation Accuracy: 0.79** | **Validation accuracy is low** | **3,840,453** |
| **Final Model** | **Time Distributed-LSTM** | **Training Accuracy :0.98**  **Validation Accuracy: 0.95** | **Model Selected. Best Validation accuracy achieved in the complete experiment.** | **3,693,253** |

Conclusion:

The Model built with Time distributed Conv2D and ConvLSTM gave better results compared to all the other models and also the model has very least number of parameters compared to other models.