**Problem Statement**

Imagine you are working as a data scientist at a home electronics company which manufactures state of the art **smart televisions**. You 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.

The gestures are continuously monitored by the webcam mounted on the TV. Each gesture corresponds to a specific command:

* 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

Each video is a sequence of 30 frames (or images). In the next couple of lectures, our subject matter expert Snehansu will walk you through the structure of the dataset.

**Understanding the Dataset**

The training data consists of a few hundred videos categorised into one of the five classes. Each video (typically 2-3 seconds long) is divided into a **sequence of 30 frames(images)**. These videos have been recorded by various people performing one of the five gestures in front of a webcam - similar to what the smart TV will use.

Below is the summary of 4 differen

SUMMARY:

|  |  |  |  |
| --- | --- | --- | --- |
| **Experiment Number** | **Model** | **Result** | **Comment** |
| **1** | **Conv2D+GRU**   * **Batch\_Size: 32** * **Activation: Relu** * **Kernel: (3,3,3)** * **Using 18 alternate frames.** | **Categorical\_Acc: 0.46**  **Val\_Acc: 0.51** | **The model is underfitting. It is reaching peak accuracy by the 12th Epoch and then the accuracy increases marginally but does not fit well.** |
| **2** | **Conv2D+LSTM**   * **Batch\_Size: 32** * **Activation: Relu** * **Kernel: (3,3,3)** * **Using 18 alternate frames.** | **Categorical\_Acc: 0.46**  **Val\_Acc: 0.57** | **The model is underfitting. It is reaching peak accuracy by the 12th Epoch and then the accuracy increases marginally but does not fit well.** |
| **3** | **Conv3D**   * **Batch\_Size: 32** * **Activation: elu** * **Kernel: (3,3,3)** * **Using 18 alternate frames.** | **Categorical\_Acc: 0.85**  **Val\_Acc: 0.77** | **The model is having a higher fit but it is still underfitting.** |
|  | **Conv3D**   * **Batch\_Size: 32** * **Activation: relu** * **Kernel: (3,3,3)** * **Using 18 alternate frames.** | **Categorical\_Acc: 0.83**  **Val\_Acc: 0.86** | **There is no over/under fitting for this model. Model as high accuracy.** |

We have considered the Conv3D with Activation: Relu as the final model.