**Gesture Recognition – Deep learning**

**Problem Statement:**

We need to develop a cool feature in the smart-TV that can recognize 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.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Experiment Number** | **Model** | **Parameters** | **Result** | **Decision + Explanation** |
| **1** | **Conv3D** | **Batch Size = 32,**  **dim\_x, dim\_y = 120,120,**  **seq\_idx = [0, 2, 4, 6, 8, 10, 12, 14, 16, 18]**  **epochs = 20** | **Throws Type Error when doing model.fit\_generator** | * **Use different libraries for read and resize as the scipy is deprecated.** * **Clean up the generator function with the updated libraries functions** * **Change the model.** **fit\_generator to model.fit function as the previous one is deprecated.** |
| **2** | **Conv3D** | **Batch Size = 128**  **2 Conv3D layers (32, 64 filters), 2 MaxPooling3D layers (pool size 2), 1 Dense layer (256 units), output Dense layer (5 units, softmax)** | **Accuracy: 1.00**  **Model is over fitting** | * **Add EarlyStopping, ReduceLROnPlateau as callbacks.** * **Add a method to pass less data to train the model.** |
| **3** | **Conv3D** |  | **Need a plot the training and validation metrics** | * **Put the models in a folder to reduce clutter.** * **Add a matplotlib to visualize the metrics.** |
| **4** | **Conv3D** | **learning\_rate=0.01** | **The Model is still overfitting.**  **Validation loss is not improving.**  **Accuracy: 0.9970** | * **Adding more Layers** |
| **5** | **Conv3D** | **Batch Size: 64**  **5 Conv3D layers (32, 64, 128, 256, 512 filters), 4 MaxPooling3D layers (varying pool sizes), 1 Dense layer (512 units), output Dense layer (5 units, softmax).** | **Accuracy fell down by a lot.**  **Validation Loss is not improving.**  **Accuracy: 0.2066**  **val\_loss: 1.6080** | * **Change Generator, add image enhancements** * **The Model fit is failing due to insufficient memory, changing the batch size.** * **Add data augmentation for improve generalization.** |
| **6** | **Conv3D** | **seq\_idx = range(7,26,2)** | **The model fitting is stopping after 11 epochs, because of the EarlyStopping as the val\_loss has stopped increasing** | * **Add Batch Normalization after every MaxPooling3D possibly get better stability in training.** |
| **7** | **Conv3D** |  | **Accuracy: 0.6705 val\_loss: 3.6772** | * **Add Dropout to improve model accuracy.** |
| **8** | **Conv3D** |  | **Training Accuracy is increasing but validation accuracy not.**  **Accuracy: 0.8750 val\_loss: 10.0498** | * **Try Conv2D + GRU as Conv3D not giving desired accuracy.** |
| **9** | **Conv2D + GRU** |  | **Validation Loss Decreased, but the accuracy also decreased as well.**  **accuracy: 0.6989**  **val\_loss: 1.9708** | * **Try Conv2D + Dense to try if simplifying the architecture helps.** |
| **10** | **Conv2D + Dense** |  | **Accuracy Increased, But Validation Loss is still high**  **accuracy: 0.8139 val\_loss: 1.3384** | * **Try using ConvLSTM2D.** |
| **11** | **ConvLSTM** |  | **No Improvement**  **accuracy: 0.5611 - val\_loss: 1.6358** | * **Try changing the architecture parameters.** |
| **12** | **ConvLSTM** |  | **Model stopped at 14th Epoch due to early stopping as the performance did not improve.**  **accuracy: 0.5739 val\_loss: 1.4696** | * **Try changing the architecture parameters.** |
| **13** | **ConvLSTM** |  | **The model training stopped at 11th Epoch due to early stopping as the validation loss did not improve.** | * **Try a lightweight architecture with least parameters.** |
| **14** | **ConvLSTM** |  | **Validation Accuracy did not increase.**  **accuracy: 0.5085 val\_loss: 2.0209** | * **Try ConvLSTM2D + Fewer Filters + BatchNormalization + Smaller Dense Layer + Global Pooling** |
| **Final Model** | **Conv3D + ConvLSTM2D** |  | **accuracy: 0.4062 val\_loss: 1.2165** | * **End** |