This explains the different architecture and their performance on the Gesture classification task. Here we tried mainly two broad architectures which are CNN-3D and CNN-2D+RNN. In each of the base model, we have tried adding different layers, neurons, and dropouts with GRU and Resnet. The result are as follows:

|  |  |  |  |
| --- | --- | --- | --- |
| **Exp. Number** | **Model** | **Result** | **Decision + Explanation** |
| **1** | **Conv3D** | **Train Accuracy: 68%**  **Val. Accuracy: 19%** | **For ablation=100 without any Augmentation.**  **Just to ensure the model working and learning.** |
| **2** | **Conv3D** | **Train Accuracy: 64%**  **Val. Accuracy: 20%** | **For ablation=100 with Augmentation.**  **Just to ensure the model working and learning.** |
| **3** | **Conv3D** | **Train Accuracy: 80%**  **Val. Accuracy: 35%** | **Same Architecture with full data with 32 batch size and no Augmentation.** |
| **4** | **Conv3D** | **Train Accuracy: 90%**  **Val. Accuracy: 35%** | **With added Dense layer and batch normalization.**  **Model is overfitting.** |
| **5** | **Conv3D** | **Train Accuracy: 98%**  **Val. Accuracy: 37%** | **With (2,2,2) filter. Val accuracy is very less and model is very over-fitting.** |
| **6** | **ResNet+ GRU** | **Train Accuracy: 18%**  **Val. Accuracy: 26%** | **Run for ablation 200. Model is working fine. No Augmentation.** |
| **7** | **ResNet+ GRU** | **Train Accuracy: 21%**  **Val. Accuracy: 24%** | **Run for ablation 200. Model is working fine with Augmentation.** |
| **8** | **ResNet+ GRU** | **Train Accuracy: 35%**  **Val. Accuracy: 50%** | **Run with Augmentation. No over-fitting but accuracy is very low.** |
| **9** | **ResNet+ GRU** | **Train Accuracy: 41%**  **Val. Accuracy: 51%** | **With added GRU neurons. Accuracy has not improved much with added Augmentation.** |
| **10** | **ResNet+ GRU** | **Train Accuracy: 80%**  **Val. Accuracy: 54%** | **Model is overfitting also accuracy is very low.** |
| **11** | **CNN2D+ GRU** | **Train Accuracy: 87%**  **Val. Accuracy: 76%** | **Model is a little overfitting, but accuracy has increased significantly.** |