Gesture-Recognition RNN Assignment

|  |  |  |  |
| --- | --- | --- | --- |
| **Experiment Number** | **Model** | **Result** | **Decision + Explanation** |
| **1** | **Conv3D (3 layers)**  Trainable params: 3,856,133 Non-trainable params: 320 | **categorical\_accuracy: 0.9668**  **val\_categorical\_accuracy: 0.5700** | **High variance in validation set**  **Try 2D instead of 3D with adding GRU** |
| **2** | **Conv2D + TimeDistributed + GRU**  Trainable params: 99,269 Non-trainable params: 576 | **categorical\_accuracy: 0.8160**  **val\_categorical\_accuracy: 0.3100** | **Low performance on both sets**  **Change pooling layer and remove GRU** |
| **3** | **Conv2D + TimeDistributed + GlobalAveragePooling2D**  Trainable params: 13,589 Non-trainable params: 192 | **categorical\_accuracy: 0.7572**  **val\_categorical\_accuracy: 0.3900** | **Performance on train set decreased, still val set at low performance compared with first trial**  **Try LSTM** |
| **4** | **Conv3D + LSTM**  Trainable params: 2,668,485 Non-trainable params: 448 | **categorical\_accuracy: 0.9653**  **val\_categorical\_accuracy: 0.5400** | **With higher trainable params, performance on train set is top again, but still the model suffers from high variance in val set**  **Add more layers in Conv3D** |
| **5** | **Conv3D (4 layers)**  Trainable params: 910,965 Non-trainable params: 1,008 | **categorical\_accuracy: 0.9759**  **val\_categorical\_accuracy: 0.5300** | **Slight improvement in train set, slight decrease in val set**  **Use transfer learning** |
| **Final Model** | **Transfer Learning (mobilenet) + TimeDistributed + LSTM + GRU**  Trainable params: 3,669,317 Non-trainable params: 23,936 | **categorical\_accuracy: 0.9991**  **val\_categorical\_accuracy: 0.9853** | **Peak performance in both train and val sets** |