Gesture Recogniton

**Model** #1:

Model : Con3D

Trainable params: 99,269

Non-trainable params: 576

Results :

**categorical\_accuracy: 0.9668**

**val\_categorical\_accuracy: 0.5700**

**Inferences:**

**On the training set the model perfoms well , but on the validation set shows HIGH Variance.**

**Model** #2:

Model : **Conv2D + TimeDistributed + GRU**

Trainable params: 99,269

Non-trainable params: 576

Results :

**categorical\_accuracy: 0.8160**

**val\_categorical\_accuracy: 0.3100**

**Inferences:**

**Performs poorer than the previous model both in the train and validation data.**

**Chaging the pooling to layer and removing the GRU to check another model.**

**Model** #3:

Model : **Conv2D + TimeDistributed + GlobalAveragePooling2D**

Trainable params: 13,589

Non-trainable params: 192

Results :

**categorical\_accuracy: 0.8160**

**val\_categorical\_accuracy: 0.3100**

**Inferences:**

**Performs poorer than the previous model both in the train and validation data.**

**Training with the LSTM model to see any change.**

**Model** #4:

Model : **Conv3D + LSTM**

Trainable params: 2,668,485

Non-trainable params: 448

Results :

**categorical\_accuracy: 0.9653**

**val\_categorical\_accuracy: 0.5400**

**Inferences:**

**With more traininig params, performance on train set is top again, but still the model suffers from high variance in val set**

**Model** #5:

Model : **Conv3D(4 layers)**

Trainable params: 910,965

Non-trainable params: 1,008

Results :

**categorical\_accuracy: 0.9759**

**val\_categorical\_accuracy: 0.5300**

**Inferences:**

**The training data performance has slightly increase but hte val data accuracy has decreased**

**Model** #Final:

Model : **Transfer Learning (mobilenet) + TimeDistributed + LSTM + GRU**

Trainable params: 3,669,317

Non-trainable params: 23,936

Results :

**categorical\_accuracy: 0.9991**

**val\_categorical\_accuracy: 0.9853**

**Inferences:**

**T**he architecture combines the efficiency of MobileNet, the temporal modeling capabilities of LSTM and GRU, and the sequential processing provided by TimeDistributed layers. This combination seems well-suited for the task of gesture recognition, leading to improved accuracy and generalization on the given problem statement.