**DL TRACK- Gesture Recognition Case Study**

* **Abhisek De**
* **Vipin Panthri**

**Note:**

* Since we are running our project in our respective systems, we have limited amount of resources. Hence, we did try to get the best model in those resources.

**Inferences made:**

1. The time taken to train the model depends upon the following factors:

* No of frames
* The size of the images.

Higher the number of these factors, greater is the training time.

1. In Conv3D and Conv2D with LSTM models, the size of the frames and the batch size is limited to 75x75 and 18 frames respectively. With the current model, any number higher than this in these parameters will throw an out of memory error.
2. Out of memory error is also persistent because of the batch size. The GPU memory is directly dependent on the batch size. Higher the batch size, lesser is the time taken to train the model but the memory allocated from the GPU Memory is high. Hence there is a tradeoff here.
3. Conv2D with LSTM model has performed better than Conv3D model.
4. Conv3D model with smaller filter(2x2x2) and dropout layers have performed a bit better than the Conv3D model with a 3x3x3 filter.
5. Finally, the transfer learning model has performed the best out of all these and will be our final model. We have used Mobilenet due to the light weighted architecture than others. VGG16 was also having memory issues hence discarded that.

**SUMMARY TABLE**

|  |  |  |  |
| --- | --- | --- | --- |
| **Experiment Number** | **Model** | **Result** | **Decision + Explanation** |
| **1** | **Conv3D(100x100 image size and batch size>18)** | **Throws out of memory error** | **Reduce the size of the image/Reduce the number of layers and set the batch size according to the GPU Memory.** |
| **2** | **Conv3D(75x75 image size,batch size=18,filter size=2x2x2)(Base Model)** | **Training Accuracy: 0.57**  **Validation Accuracy: 0.58** | **Low validation accuracy because of too many trainable parameters. Hence we will decrease the parameters and try again.** |
| **3** | **Conv3D(75x75 image size,batch size=18,filter size=2x2x2,reduced parameters)** | **Training Accuracy: 0.76**  **Validation Accuracy:0.67** | **The validation and training accuracies have improved after reducing the parameters.** |
| **4** | **Conv3D(75x75 image size,batch size=18,filter size=3x3x3)** | **Training Accuracy: 0.39**  **Validation Accuracy:0.49** | **The model has performed poorly as compared to others. This might be because of the big filter size. Reducing the parameters.** |
| **5** | **Conv3D(75x75 image size,batch size=18,filter size=3x3x3, reduced parameters)** | **Training Accuracy: 0.52**  **Validation Accuracy:0.72** | **Reducing the parameters increased the accuracies but still the model performs poorly. Hence changing the architecture.** |
| **6** | **Conv2D with LSTM** | **Training Accuracy: 0.72**  **Validation Accuracy: 0.67** | **Model performs good but the validation score is low. Hence will try after reducing parameters.** |
| **7** | **Conv2D with LSTM with reduced parameters** | **Training Accuracy: 0.70**  **Validation Accuracy: 0.65** | **Reducing the parameters doesn’t help with the accuracy hence going for transfer learning.** |
| **8** | **Transfer Learning with LSTM** | **Training Accuracy: 1.00**  **Validation Accuracy:0.79** | **This model performs good but is a bit overfitting. This can be eliminated by training all the weights.** |
|  |  |  |  |
| **Final Model** | **Transfer Learning with LSTM** | **Training Accuracy: 1.00**  **Validation Accuracy:0.79** | **This model performs good but is a bit overfitting. This can be eliminated by training all the weights** |

In one of the epochs while running Mobilenet model, the training accuracy is almost 93% and the validation accuracy is 81%. We will be submitting that .h5 file for the test data.