**Case Study done on Jarvis Lab account**

Observations as experienced during the model creation and training process

* Started with small model having single Conv3D layer but results were not good. The accuracy level increased for first few
* Note that dropouts (usually it was set to 0.25 in dense connections) and batch normalization play an important role for regularization but not explicitly mentioned them in the model column.
* Started with batch size with 5 but it was working very fast so have increased batch size to 10, and it is doing reasonably good.
* Not only experimented with Batch Size, Epochs and Model layers. But also with no of frames. Started with lower number of frames 18,23,28
* Images are resized to 120x120.
* GRU and LSTM layer used in with MobileRNN network.

|  |  |  |  |
| --- | --- | --- | --- |
| **Experiment Number** | **Model** | **Result** | **Decision + Explanation** |
| **1** | **Conv 3D Model with 1 hidden Conv layer using 15 frames per video**  **3D layer followed by max pool 3D**  **+ image size 120 by 120** | **Training Loss: 0.9841**  **Training Accuracy: 62%**  **Validation Accuracy: 55%** | **Underfitted model as training and validation both accuracies are low. Even the results were not stable.** |
| **2** | **Conv 3D Model with 3 hidden conv layers using 18 frames per video**  **+ image size 120 by 120**  **+Batch size 10**  **+ Epochs 10** | **Training loss: 0.1843**  **Train accuracy : 94%**  **Val accuracy: 72%** | **Slightly better results** |
| **3** | **Conv 3D Model with 3 hidden conv layers using 23 frames per video**  **+ image size 120 by 120**  **+Batch size 10**  **+ 15 Epochs** | **Training loss: 0.1654**  **Training Accuracy: 94%**  **Validation Accuracy: 85%** | **Model still seems to be not doing good with test data set** |
| **4** | **Conv 3D Model with 3 hidden conv layers using 23 frames per video**  **+ image size 120 by 120**  **+Batch size 10**  **+ 20 Epochs** | **Training loss: 0.649**  **Training Accuracy: 98%**  **Validation Accuracy: 84%** | **20 Epochs and 23 frames gave good results. Still the valid accuracy is low.** |
| **5** | **Conv 3D Model with 3 hidden conv layers using 28 frames per video**  **+ image size 120 by 120**  **+Batch size 10**  **+ 20 Epochs** | **Training Accuracy: 93%**  **Validation Accuracy:88%** | **Not a good improvement still the model results fluctuating up and down in between Epochs** |
| **6** | **Mobilenet (retrain all weights)**  **+ GRU (128 cells)**  **+ Dense (128 nodes)**  **+ image size 120 by 120**  **+ 20 images per video** | **Training Accuracy: 96%**  **Validation Accuracy: 94%** | **Retrained all the weights of Mobilenet.**  **Batch size = 5**  **Epochs = 10** |
| **7** | **Mobilenet (retrain all weights)**  **+ GRU (128 cells)**  **+ Dense (128 nodes)**  **+ image size 120 by 120**  **+ 30 images per video** | **Training Accuracy: 98.8%**  **Validation Accuracy: 89%** | **Retrained all the weights of Mobilenet.**  **Batch size = 10**  **Epochs = 20** |
| **8** | **Mobilenet (fine tune after 50th layer)**  **+ LSTM (128 cells)**  **+ Dense (128 nodes)**  **+ image size 120 by 120**  **+ 20 images per video** | **Training Accuracy: 99.85%**  **Validation Accuracy: 91%** | **Batch size = 10**  **Epochs = 10**  **Model is overfitted** |
|  |  |  |  |
|  |  |  |  |