|  |  |  |  |
| --- | --- | --- | --- |
| **Experiment Number** | **Model** | **Result** | **Decision + Explanation** |
| 1 | **Conv3D Basic model with filter size (2, 2, 2)** Frame Resize = (128x128) | - Accuracy: 73% - Loss is continuously decreasing on each Epoch | Both Model with even and old filter size accuracy is same. Hence next step to add Batch Normalisation step to check if accuracy is increasing in either cases. |
| **Conv3D Basic model with filter size (3, 3, 3)** Frame Resize = (128x128) | - Accuracy: 73% - Loss is continuously decreasing on each Epoch |
| 2 | **Conv3D Model with filter size (2, 2, 2) and Batch Normalization** Frame Resize = (128x128) | - Accuracy: 77% - Loss is continuously decreasing on each Epoch - Model has overfitted. | Model with filter size (3, 3, 3) has better accuracy. It is always recommend to use old size filter. Hence going forward we are using (3, 3, 3) filter size Reference: <https://datascience.stackexchange.com/questions/23183/why-convolutions-always-use-odd-numbers-as-filter-size>  Next step to see if dropout helps in improving accuracy with filter size (3, 3, 3) |
| **Conv3D Model with filter size (3, 3, 3) and Batch Normalization** Frame Resize = (128x128) | - Accuracy: 82% - Loss is continuously decreasing on each Epoch - Model has overfitted. |
| 3 | **Conv3D: Model with Batch Normalization and Dropout** Frame Resize = (128x128) Filter size = (3, 3, 3) | - Accuracy: 61% - Loss is struck in local minima. - Model has not overfitted. | This convey clearly that dropout are not helping. By adding Dropouts, there is actual loss of information and model not able to classify properly. Hence we will not be using dropout for further models. |
| 4 | **Conv3D Model with More Layers/Less Training params** Frame Resize: (128x128) Filter size = (3, 3, 3) + Batch Normalization | - Accuracy: 72% - Loss is struck in local minima. - Model has not overfitted. | Adding Convolution Extra layer has not help in achieve better performance compare to experiment No. 2. Hence we will not be using Extra layer for further experiments. |
| 5 | **Conv3D Model with Less Layers** Frame Resize: (128x128) Filter size = (3, 3, 3) + Batch Normalization | - Accuracy: 80% - Loss is continuously decreasing on each Epoch. - Model has overfitted. | Generated Model size is around 470 mb which is not feasible to run within a webcam memory. Hence we will not be reducing layers for further experiments. |
| 6 | **Conv3D Model with Data Augmentation (Crop)** Frame Resize: (120x120) Filter size = (3, 3, 3) + Batch Normalization | - Accuracy: 78% - Loss is continuously decreasing on each Epoch. - Model has overfitted. | With Cropped Frame size (120x120), there is some information of loss around the Edges. Hence accuracy is down compare to experiment No 2. |
| 7 | **Conv3D Model with Data Augmentation (15 Frames)** Frame Resize: (128x128) Filter size = (3, 3, 3) + Batch Normalization | - Accuracy: 67% - Loss is continuously decreasing on each Epoch. - Model has overfitted. | By reducing frame from 30 to 15, there is some information of loss and sequence of information gesture is not continous which might be confusing the model. Hence accuracy is low for validation dataset. |
| 8 | **Conv3D Model with Data Augmentation (Crop and 15 Frames)** Frame Resize: (120x120) Filter size = (3, 3, 3) + Batch Normalization | - Accuracy: 65% - Loss is continuously decreasing on each Epoch. - Model has overfitted. | This is same case as experiment no 7. |
| 9 | **GRU Model with Transfer Learning (Trainable=False)** Frame Resize: (128x128) + Batch Normalization | - Accuracy: 35% - Loss is struck in local minima. - Model has not overfitted. | This was just an experiment and expected as we are not allowing weights of the transfer model to be updated. |
| 10 | **GRU Model with Transfer Learning (Trainable=True)** Frame Resize: (128x128) + Batch Normalization | - Accuracy: 89% - Loss is continuously decreasing on each Epoch. - Model has overfitted. | The model has generalised to most extend.  GRU model has learn the gesture pattern to more extend compare to experiment no 2.  Next step would be to check GRU with 15 Frames. |
| 11 | **GRU Model with Transfer Learning (Trainable=True) and Data Augmentation (15 Frames)** Frame Resize: (128x128) + Batch Normalization | - Accuracy: 76% - Loss is continuously decreasing on each Epoch. - Model has overfitted. | This indicates that for the model to correctly predict the human gesture it requires all 30 Frames. Removing frames are actual loss of data and model is not able to classify correctly. This we have observed in case of experiment no 7 as well. |