Gesture Recognition – Deep learning

Case Study by: Prasad Maharana

Problem Statement:

We need to develop a cool feature in the smart-TV that can recognise five different gestures performed by the user which will help users control the TV without using a remote. The following table consists of the experiments done to build a model to predict the gestures from the given data set.

Results

Exp No.	Model	Result	Decision + Explanation	Parameters
1	Conv 3D	Train accuracy: 0.96 Val accuracy:0.16	The model is overfitting and we need to add drop out layers to reduce over fitting.	592,773
2	Conv 3D	Train accuracy 0.86 Val accuracy:0.27	The model is clearly overfitting. Reducing batch size and image size	1,541,829
3	Conv 3D	Train accuracy 0.70 Val accuracy:0.27	The model is again overfitting. Adding more layers to improve validation	1,336,629
4	Conv 3D	Train accuracy 0.92 Val accuracy:0.85	Huge improvement by reducing image size to 96 x 96 and reducing batch size. Adding drop out layers	2,130,549
5	Conv 3D	Train accuracy 0.92 Val accuracy:0.85	The model has good fit but reduced validation accuracy. Let's reduce parameters	2,130,549
6	Conv 3D	Train accuracy 0.92 Val accuracy:0.85	For the above low memory foot print model, we get a validation accuracy of 81% which a little lower than the previous models score but we are losing only 1% of the validation score with a reduction of (2130549-696645=1433904 parameters) or 60% parameter reduction which very good). Let's further reduce parameters	696,645
7	Conv 3D	Train accuracy 0.85 Val accuracy:0.79	There is drop in the overall performance but the model fits well	398,213
8	CNN + LSTM	Train accuracy 0.98 Val accuracy:0.85	For CNN - LSTM model we have a good performance but there is slight overfitting	1,657,445
9	Conv 3D + data augmentation	Train accuracy 0.75 Val accuracy:0.79	(3, 3, 3) Filter & 96 x 96 image resolution The model seems to overfit	1,541,829
10	Conv 3D + data augmentation	Train accuracy 0.75 Val accuracy:0.78	2, 2, 2) Filter & 96 x 96 image resolution. Increase epoch count to 20. Network is still overfitting	1,336,629
11	Conv 3D + data augmentation	Train accuracy 0.75 Val accuracy:0.79	Adding more layers. Network is still overfitting	2,130,549
12	Conv 3D + data augmentation	Train accuracy 0.78 Val accuracy:0.30	Adding dropouts has reduced the performance and it is overfitting again	2,130,549

Exp No.	Model	Result	Decision + Explanation	Parameters
13	Conv 3D + data augmentation	Train accuracy 0.85 Val accuracy:0.77	Very good performance the model fits well. Let's reduce the network parameters.	696,645
14	Conv 3D + data augmentation	Train accuracy 0.75 Val accuracy:0.78	After reducing network parameters, model's performance is quite weird and overfitting	398,213
15	CNN LSTM with GRU	Train accuracy 0.95 Val accuracy:0.81	The model fits well and augmentation has not improved any of the model's performance	1,934,565

Conclusion:

After doing all the experiments, we finalized Model 8– CNN+LSTM, which performed well.
Reason:
(Training Accuracy: 98%, Validation Accuracy: 85%)
Number of Parameters (1,657,445) less according to other models' performance
Learning rate gradually decreasing after some Epochs