

Deep Learning Course Project - Gesture Recognition

Write – up

Developing a cool feature in the smart-TV that can recognize five different gestures performed by the user which will help users control the TV without using a remote.

The gestures are continuously monitored by the webcam mounted on the TV. Each gesture corresponds to a specific command:

Thumbs up: Increase the volume Thumbs down: Decrease the volume Left swipe: 'Jump' backwards 10 seconds Right swipe: 'Jump' forward 10 seconds
Stop: Pause the movie

Each video is a sequence of 30 frames (or images).

And here we are using 2 architecture - 3D CNN and a CNN-RNN based model.

Basically totally 15 models are created, 7 using 3D CNN architecture, 4 models on using 2D CNN and GRU architecture and 3 models on 2D CNN and LSTM architecture.

In all the model I'm getting an accuracy of 90% using 3D CNN model.

Here the data is Augmented, Reshaped, cropped, Resized, Normalized before the data is sent to the model for better model preparation. Regarding the model and how the model has altered and why it's changed are mentioned in the below table. 3 .h5 files have been provided 1 for 3D CNN one for ConvGRU and another for ConvLSTM architectures.

Experiment Number	Model	Result	Decision + Explanation
1	Conv3D	Categorical accuracy = 87.7 % Validation accuracy = 81 %	Augmented, Resized, Normalized, taken batch size of 10, frame size of 30 and epoch as 20. The frame size is at its maximum, epochs are taken 20 so not to make the model too complex.
2	Conv3D	Categorical accuracy = 94.68 % Validation accuracy = 92.5 % (Good model)	Augmented, Resized, Normalized, taken batch size of 30, frame size of 30 and epoch as 30, number of epochs has been increased so as to make the model little bit complex and so as to increase the accuracy.

3	Conv3D	Categorical accuracy = 87.58 % Validation accuracy = 85 %	Data is Augmented to increase the quality, Resized it and also normalized. Taken a batch size of 20, epoch as 20, to check by decreasing the complexity can we increase the accuracy.
4	Conv3D	Categorical accuracy = 79.27 % Validation accuracy = 84 % (Getting more validation accuracy than categorical accuracy- not a good model)	Data is Augmented to increase the data quality, Resized it and also normalized, before giving the data to model, taken a batch size of 10(decreased further), epoch as 20, and frame size of 20(decreased from 30), to check by decreasing the complexity can we get the same accuracy as above.
5	Conv3D	Categorical accuracy = 52 % Validation accuracy = 60 % (Getting more validation accuracy than categorical accuracy- not a good model)	Added another 3D CNN and a dense layer added a dropout layer to increase the performance and also increased the number of epoch, for proper understanding of data
6	Conv3D	Categorical accuracy = 79 % Validation accuracy = 80 % (Getting more validation accuracy than categorical accuracy- not a good model)	Batch size reduced, number of epochs have taken 30, and frame size has been reduced to 20, and checking by increasing the number of CNN layers, dropouts and decreasing the frame, batch and epoch can we increase the model performance.
7	Conv3D	Categorical accuracy = 50 % Validation accuracy = 49 % (Not a good accuracy)	Decreased the frame size further to 16, taken batch size and number of epochs as 20, and removing the newly added layers as it didn't gave any significant result, building a model again.
8	ConvGRU	Categorical accuracy = 73.46 % Validation accuracy = 73 % (Good accuracy compared to other models and the model is not overfitting also.)	Using new architecture CNN-RNN model and we are using GRU model here as it is little bit lighter compared to LSTM layers. Frame size as 30, Batch size as 10, epoch as 30, so that we have taken smaller batch size (as it consumes less memory) and

			little bit high epoch (to increase the performance).
9	ConvGRU	Categorical accuracy = 96.73 % Validation accuracy = 76 % (Model is overfitting)	Trying 2D CNN and GRU architecture for model building Frame size as 30, batch size as 20, epoch as 20(to decreasing the overfitting of data), taking the learning rate as 0.001 as Learning rate is a hyper-parameter that controls how much we are adjusting the weights of our network with respect the loss gradient. And learning rate affects the model quickly.
10	ConvGRU	Categorical accuracy = 88.23 % Validation accuracy = 65.67 % (Model is overfitting)	Taking less frame size as 20, batch size as 10 so that the memory usage can be reduced, epoch size increased for better iterations on data.
11	ConvGRU	Categorical accuracy = 90.85 % Validation accuracy = 68 % (Model is overfitting)	In ConvGRU model decreasing the frame size, taking the batch size as 20 and decreasing the epochs as as to decrease the overfit of data. And a dense layer a dropout layer and an extra CNN layer is added for better model understanding of input data.
12	ConvLSTM	Categorical accuracy = 71.31 % Validation accuracy = 67.7 % (Good accuracy compared to other models and the model is not overfitting also.)	Using LSTM layer along with 2D CNN layers, and LSTM performs better with help of 3 gates in the LSTM cell, and we can prevent vanishing gradient issue, decreasing the batch size, increasing the epoch size and frame size.
13	ConvLSTM	Categorical accuracy = 74.18 % Validation accuracy = 59.33 % (Model is overfitting)	Using same frame size, increasing the batch size and decreasing the number of epochs checking the model may perform well to give good accuracy.
14	ConvLSTM	Categorical accuracy = 68.63 % Validation accuracy = 54.33 % (Model is overfitting)	Using batch size as 20 and increasing the number of epochs so that the number of iterations on the data is increased and to

			check whether the model performance improves.
15	ConvLSTM	Categorical accuracy = 50.86 % Validation accuracy = 53.81 % (Not a good Accuracy)	Using ConvLSTM layer with frame size 20, batch size as 15 and epoch as 35, we decreased the epochs to reduce overfit, as it was hugely overfitting in previous model.
Final Model	Conv3D	Categorical accuracy = 94.68 % Validation accuracy = 92.5 % (Good accuracy)	Augmented, Resized, Normalized, taken batch size of 30, frame size of 30 and epoch as 30, number of epochs has been increased so as to make the model little bit complex and so as to increase the accuracy. The model is giving good Categorical and Validation accuracy and the difference is also negligible and the model is not overfitting. Overall it's a good model.