

Gesture Recognition

Problem Statement

Imagine you are working as a data scientist at a home electronics company which manufactures state of the art smart televisions. You want to develop 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

Experiments

Below are the experiments conducted with various models and different set of hyperparameters for given data set

| Exp. | Model | Result | | Decision + Explanation |
|------|--------|----------------|---------------------|--|
| | | Train accuracy | Validation accuracy | |
| 1 | Conv3D | 0.20 | 0.17 | <i>All Images Batch Size = 10 Optimizer = SGD Number of Epochs = 5</i> Model accuracy is not improving |
| 2 | Conv3D | 0.20 | 0.20 | <i>All Images Batch Size = 16 Optimizer = SGD Number of Epochs = 5</i> Still no improvement with increase in batch size |
| 3 | Conv3D | 0.20 | 0.20 | <i>All Images Batch Size = 20 Optimizer = SGD Number of Epochs = 5</i> No improvement in result with increase in batch size |
| 4 | Conv3D | 0.20 | 0.18 | <i>All Images Batch Size = 32 Optimizer = SGD Number of Epochs = 5</i> Further increase in batch size reduced the validation accuracy |
| 5 | Conv3D | 0.20 | 0.20 | <i>All Images Batch Size = 32 Optimizer = Adam Number of Epochs = 10</i> No improvement in result with change in optimiser and epochs |
| 6 | Conv3D | 0.20 | 0.18 | <i>ODD Images Batch Size = 32 Optimizer = SGD Number of Epochs = 10</i> Odd image indexes reduced the validation accuracy |
| 7 | Conv3D | 0.20 | 0.24 | <i>EVEN Images Batch Size = 32 Optimizer = SGD Number of Epochs = 10</i> No improvement in result with change to even image indexes |

Gesture Recognition

| | | | | |
|----|-------------|------|------|--|
| 8 | Conv3D | 0.20 | 0.22 | <i>Image indexes from 5-27 Batch Size = 32 Optimizer = SGD Number of Epochs = 10</i> No improvement compared to previous results |
| 9 | Conv3D | 0.21 | 0.15 | <i>Image indexes from 5-27 Batch Size = 32 Optimizer = Adam Number of Epochs = 10</i> Change in optimiser further reduced the validation accuracy |
| 10 | Conv3D | 0.81 | 0.80 | With Dropouts=0.5 <i>All Images Batch Size = 10 Optimizer = SGD Number of Epochs = 10</i> Significant improvement in results by adding additional layers and including dropouts |
| 11 | Conv3D | 0.84 | 0.85 | <i>With Dropouts=0.5</i> <i>All Images Batch Size = 16 Optimizer = SGD Number of Epochs = 10</i> Further increase in accuracy with increase in batch size Third best model in our experiments |
| 12 | Conv3D | 0.80 | 0.79 | <i>With Dropouts=0.5</i> <i>All Images Batch Size = 20 Optimizer = SGD Number of Epochs = 10</i> No significant improvement in results with increase in batch size |
| 13 | Conv3D | 0.79 | 0.82 | <i>With Dropouts=0.5</i> <i>All Images Batch Size = 32 Optimizer = SGD Number of Epochs = 10</i> No significant improvement in results with increase in batch size |
| 14 | Conv3D | 0.84 | 0.69 | <i>With Dropouts=0.5</i> <i>All Images Batch Size = 32 Optimizer = Adam Number of Epochs = 10</i> Change in optimiser further reduced the validation accuracy |
| 15 | Conv3D | 0.93 | 0.77 | With Dropouts=0.25 <i>Image indexes from 5-27 Batch Size = 32 Optimizer = SGD Number of Epochs = 10</i> Observed overfitting of the model with change in image indexes and reduction of dropouts |
| 16 | Conv3D | 0.72 | 0.74 | With Dropouts=0.5 <i>Image indexes from 5-27 Batch Size = 32 Optimizer = SGD Number of Epochs = 10</i> Observed stable results with increase in dropouts |
| 17 | Conv3D | 0.44 | 0.48 | <i>With Dropouts=0.5</i> <i>Image indexes from 5-27 Batch Size = 32 Optimizer = Adam Number of Epochs = 10</i> Change in optimiser reduced the training and validation accuracy |
| 18 | Conv2D+LSTM | 0.49 | 0.48 | All Images Batch Size = 10 Optimizer = SGD Number of Epochs = 10 |

Gesture Recognition

| | | | | |
|----|--------------|------|------|--|
| | | | | Change in model from conv3d to conv2d+LSTM reduced the results |
| 19 | Conv2D+LSTM | 0.47 | 0.53 | All Images Batch Size = 16 Optimizer = SGD Number of Epochs = 10 Observed negligible improvement in results with increase in batch size |
| 20 | Conv2D+LSTM | 0.41 | 0.45 | All Images Batch Size = 20 Optimizer = SGD Number of Epochs = 10 Observed further reduction in accuracy results with increase in batch size |
| 21 | Conv2D+LSTM | 0.39 | 0.30 | All Images Batch Size = 32 Optimizer = SGD Number of Epochs = 10 Observed further reduction in accuracy results with increase in batch size |
| 22 | Conv2D+LSTM | 0.91 | 0.68 | All Images Batch Size = 32 Optimizer = Adam Number of Epochs = 10 Observed overfitting of the model with change optimiser |
| 23 | Conv2D+LSTM | 0.35 | 0.40 | ODD Images Batch Size = 32 Optimizer = SGD Number of Epochs = 10 Similar results as observed in experiment 20 |
| 24 | Conv2D+LSTM | 0.43 | 0.30 | EVEN Images Batch Size = 32 Optimizer = SGD Number of Epochs = 10 No improvement in result with change to even image indexes |
| 25 | Conv2D+LSTM | 0.38 | 0.32 | Image indexes from 5-27 Batch Size = 32 Optimizer = SGD Number of Epochs = 10 No improvement compared to previous results |
| 26 | Conv2D+LSTM | 0.87 | 0.65 | <i>Image indexes from 5-27</i> Batch Size = 32 Optimizer = Adam Number of Epochs = 10 Observed significant improvement in results with Adam optimiser but model seems to overfit |
| 27 | Conv2D+GRU | 0.52 | 0.23 | Image indexes from 5-27 Batch Size = 32 Optimizer = SGD Number of Epochs = 10 Change in model from conv2d+LSTM to conv2d+GRU reduced the results |
| 28 | Conv2D+Dense | 0.84 | 0.62 | Image indexes from 5-27 Batch Size = 32 Optimizer = SGD Number of Epochs = 10 Change in model from conv2d+GRU to conv2d+Dense improved the results but model seems to overfit |
| 29 | Conv3D | 0.95 | 0.90 | With Dropouts=0.25 All Images Batch Size = 20 Optimizer = SGD Number of Epochs = 10 Significant improvement in results by reducing the dropout and change in batch size BEST MODEL!!!! |
| 30 | Conv3D | 0.88 | 0.83 | With Dropouts=0.25 All Images Batch Size = 20 Optimizer = Adam Number of Epochs = 10 |

Gesture Recognition

| | | | | |
|----|---------------|------|------|---|
| | | | | Observed slight drop in results after changing optimiser but still giving good results |
| 31 | Conv2D+LSTM2D | 0.51 | 0.48 | All Images Batch Size = 32 Optimizer = SGD Number of Epochs = 50 Change in model from conv3D to conv2D+LSTM2D with increase in epochs reduced the results |
| 32 | Conv2D+LSTM2D | 0.79 | 0.74 | All Images Batch Size = 32 Optimizer = Adam Number of Epochs = 50 Observed significant improvement in results with Adam optimiser |
| 33 | Conv2D+LSTM2D | 0.90 | 0.88 | All Images Batch Size = 30 Optimizer = Adam Number of Epochs = 50 Observed significant improvement in results change in batch size Second best model in our experiments |

Conclusion

- As per the experiments conducted above following models gave us satisfactory results
 - Conv3D
 - ⇒ Experiment - 11
 - ⇒ Experiment - 29
 - ⇒ Experiment - 30
 - Conv2D+LSTM2D
 - ⇒ Experiment - 32
 - ⇒ Experiment - 33
- Out of these five experiments, experiment – 29 (Row is highlighted above) gave us very good results.
- However, this model consumes lot of memory due to huge number of parameters (Total params: 9,440,773).
- So, we are considering model having good accuracy and with less number of parameters (Total params: 13,781). Hence, we are going with second best fit model (**Conv2D+LSTM2D**) obtained in **experiment -33**

