REPORT

Task 1: CNN with Tensorflow

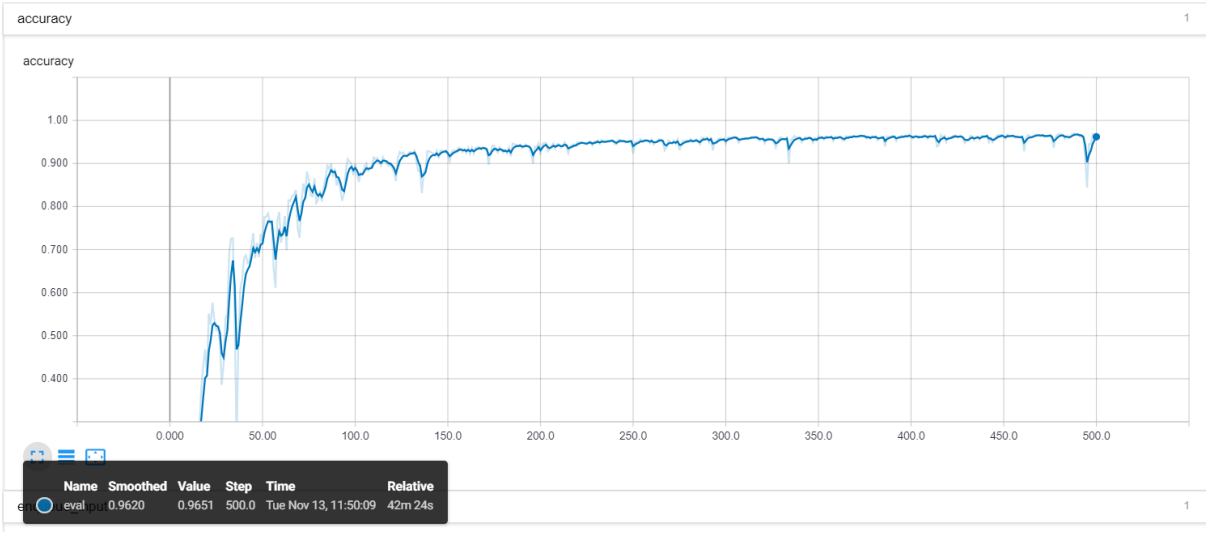
The network graph for this task looks like
I/P → CONV1 → RELU → POOL1 → CONV2 → RELU → POOL2 → FC → RELU → SOFTMAX → O/P

Since I was able to create the tensorflow environment on my own system, I executed the code in a Jupyter Notebook. CNN is built using Tensorflow Estimators. For testing, the Filter Size varied between {3, 5, 7}, and LR varied between {0.1, 0.01, 0.001, 0.0001}, number of epochs varied between {500, 1000} whereas number of filters is kept constant at 16.

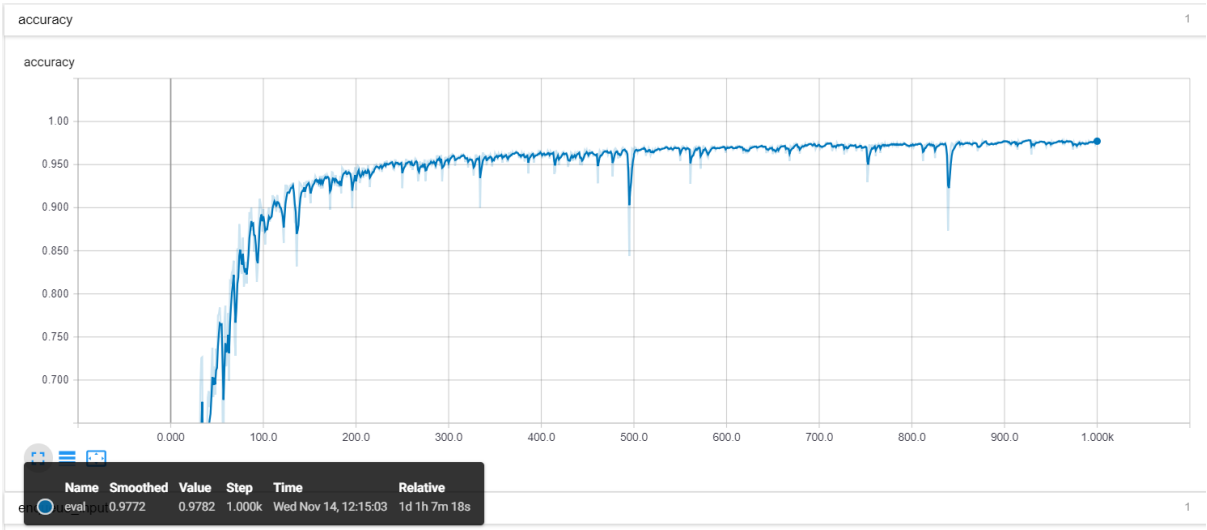
Task 2: CNN and Learning Rates

Comparing the CNN for the following LRs – {0.1, 0.01, 0.001, 0.0001}, with the filter size being 3x3.

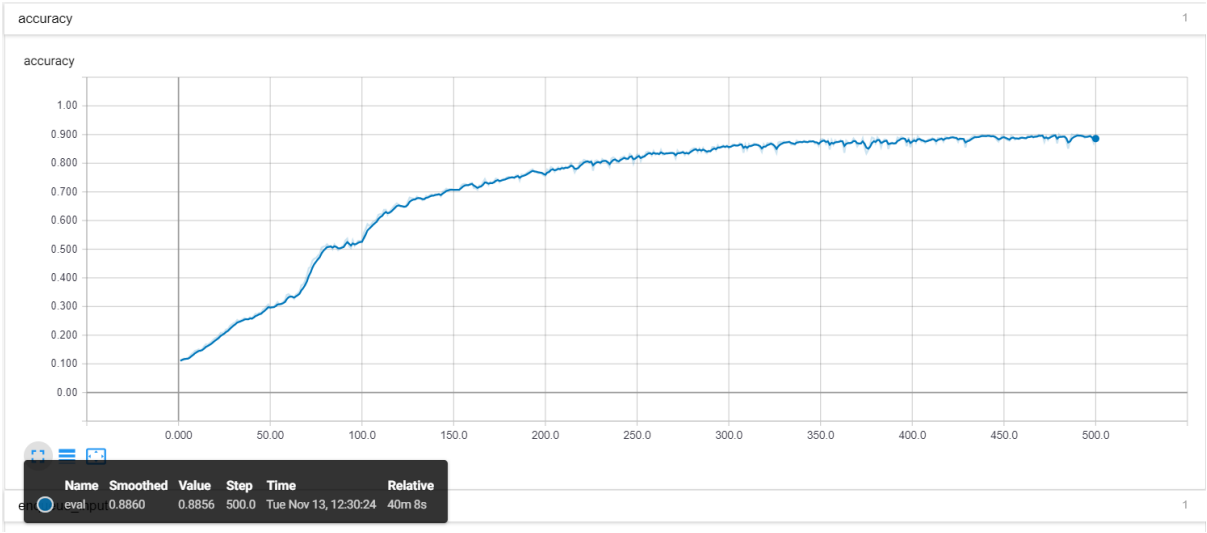
CNN at LR 0.1, Filter Size 3x3, Epochs 0-500 : ***VALIDATION ACCURACY 96.51%***



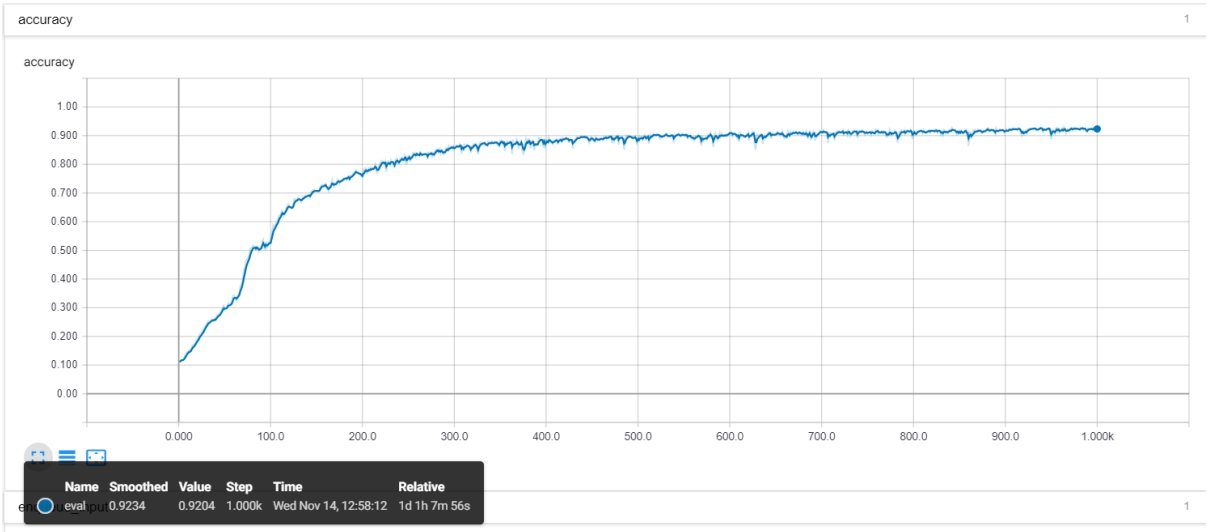
CNN at LR 0.1, Filter Size 3x3, Epochs 0-1000 : ***VALIDATION ACCURACY 97.82%***



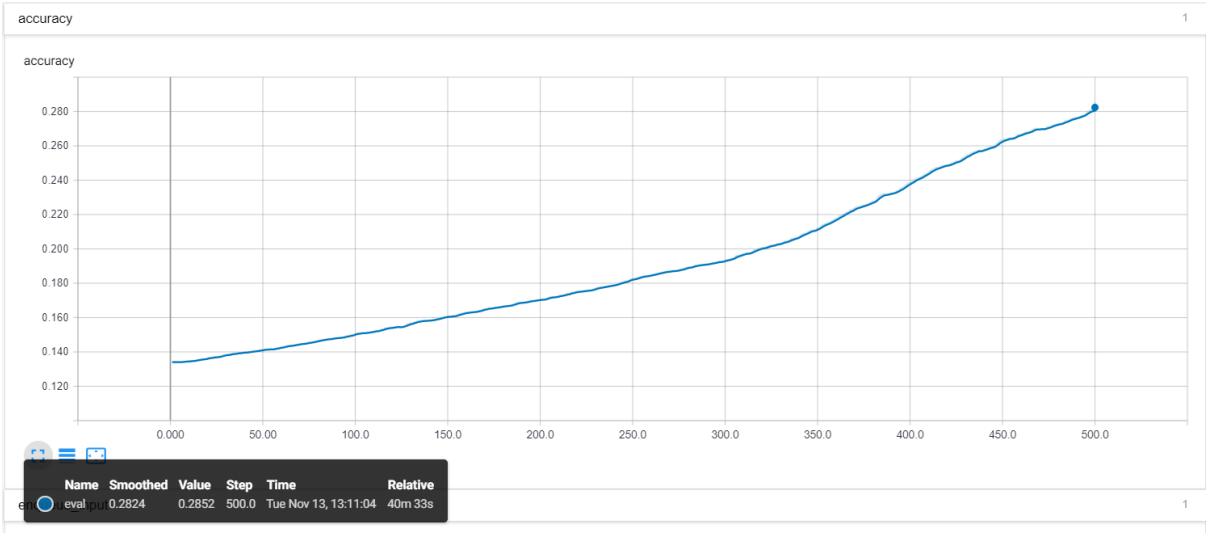
CNN at LR 0.01, Filter Size 3x3, Epochs 0-500 : ***VALIDATION ACCURACY 88.56%***



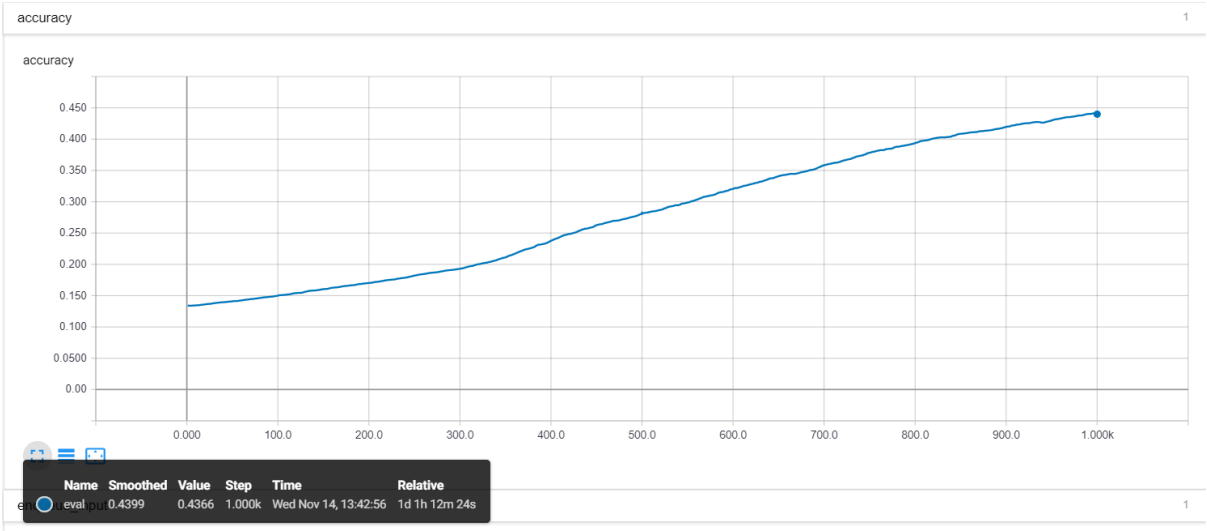
CNN at LR 0.01, Filter Size 3x3, Epochs 0-1000 : ***VALIDATION ACCURACY 92.04%***



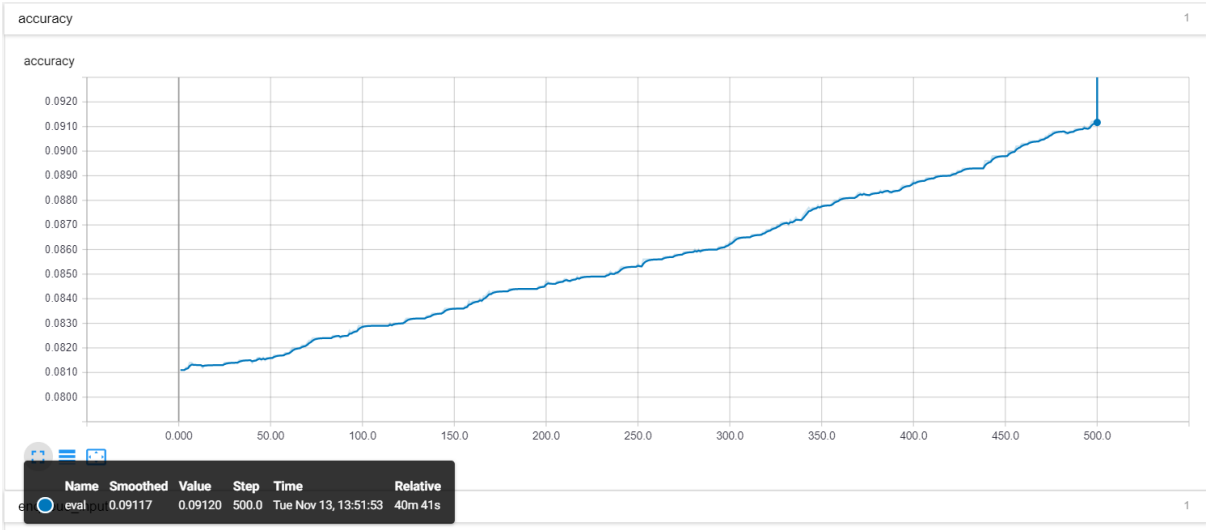
CNN at LR 0.001, Filter Size 3x3, Epochs 0-500 : ***VALIDATION ACCURACY 28.52%***



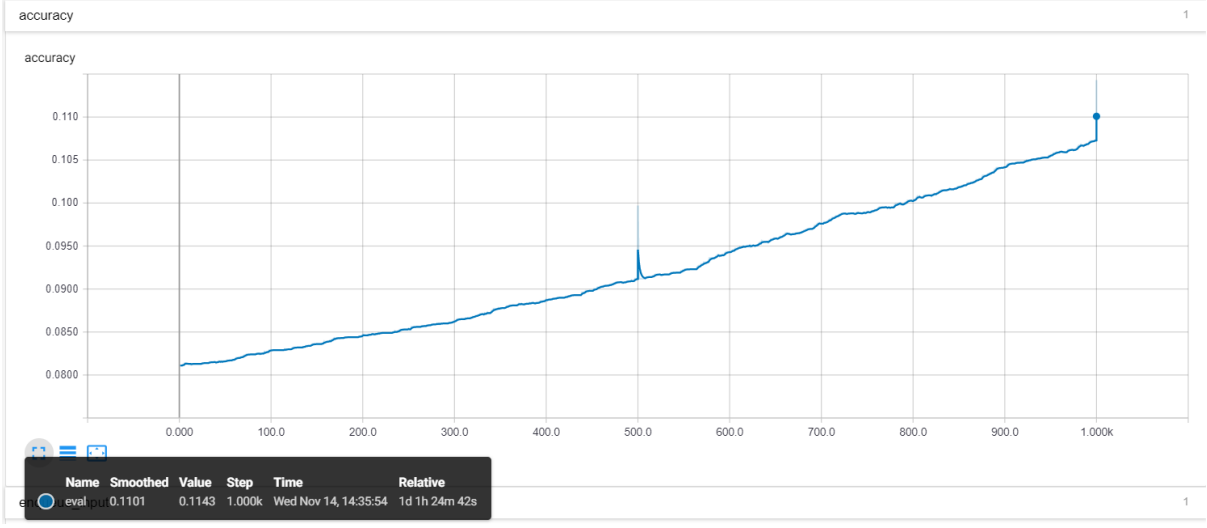
CNN at LR 0.001, Filter Size 3x3, Epochs 0-1000 : ***VALIDATION ACCURACY 43.66%***



CNN at LR 0.0001, Filter Size 3x3, Epochs 0-500 : ***VALIDATION ACCURACY 09.12%***



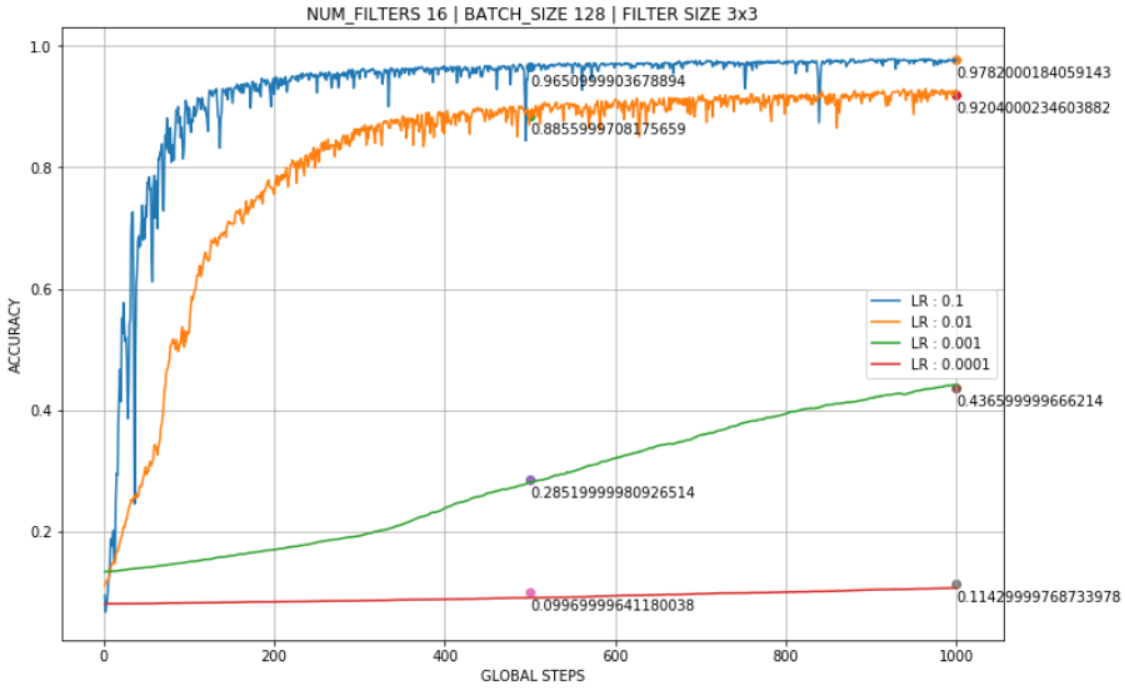
CNN at LR 0.0001, Filter Size 3x3, Epochs 0-1000: ***VALIDATION ACCURACY 11.43%***



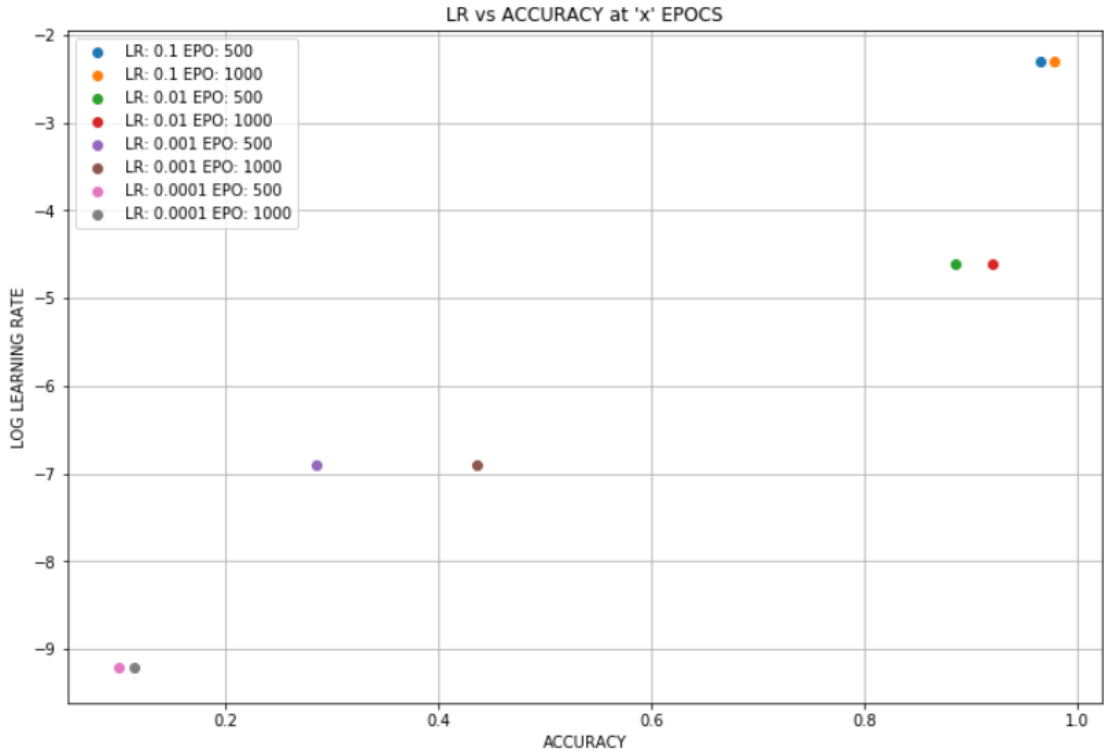
Observations – Constant Filter Size 3x3, Varying LR, Varying Epochs

Learning Rate	Number of Epochs	Validation Accuracy (%)	Comments
0.1	0 – 500	96.51	The network learns the quickest for a learning rate of 0.1, achieving the highest accuracy for minimum number of epochs. (Best accuracy to epoch ratio)
	0 – 1000	97.82	With an increase in the number of epochs, the accuracy improves marginally.
0.01	0 – 500	88.56	The next best learning rate after 0.1. High accuracy achieved on this LR as well.
	0 – 1000	92.04	We can see a marginal improvement in the accuracy with an increase in number of epochs. We can probably achieve similar accuracy as 0.1 in the next 500 epochs.
0.001	0 – 500	28.52	The learning is slow on LR 0.001.
	0 – 1000	43.66	With an increase in the number of epochs, the accuracy improves by almost 55%. This LR may achieve high accuracy given enough time, say about 2500+ epochs.
0.0001	0 – 500	9.12	Little to no learning is observed here. This could be because of ReLU gradients dying off or vanishing.
	0 – 1000	11.43	Since there is not much improvement when the epochs are doubled, it can be inferred that the network will not learn at such a slow learning rate.

Filter Size 3x3, Accuracy vs Global Steps for all 4 LRs.



Filter Size 3x3, Accuracy vs LR for 500 & 1000 Epochs

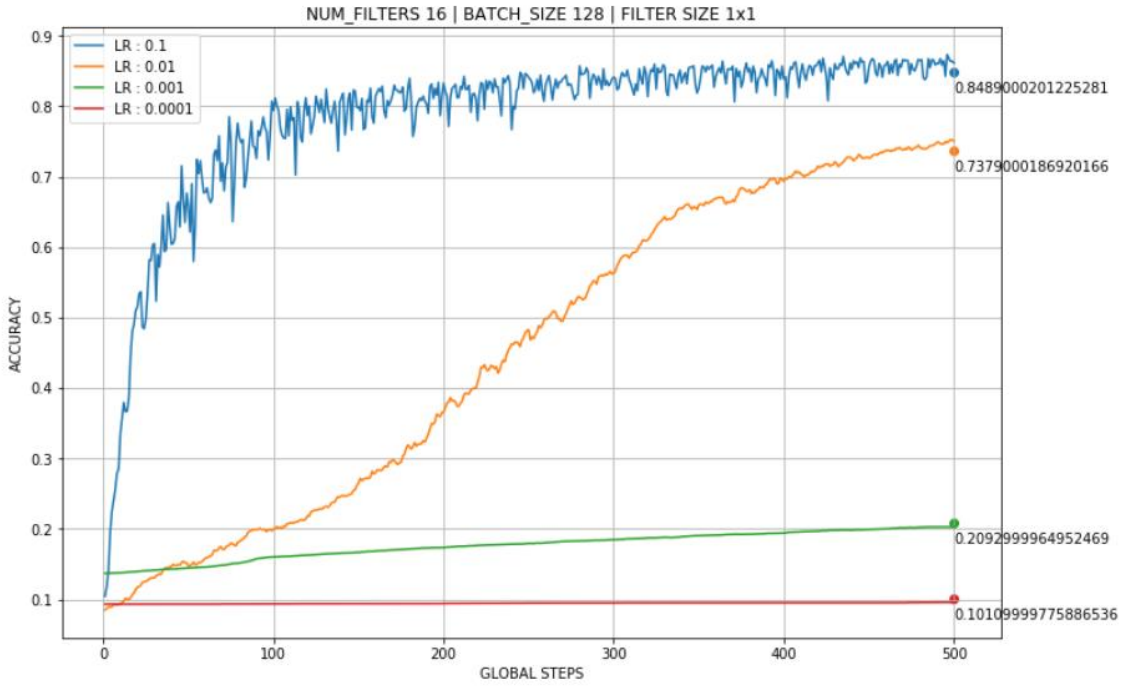


For the same LR, accuracy improves with an increase in the number of epochs.

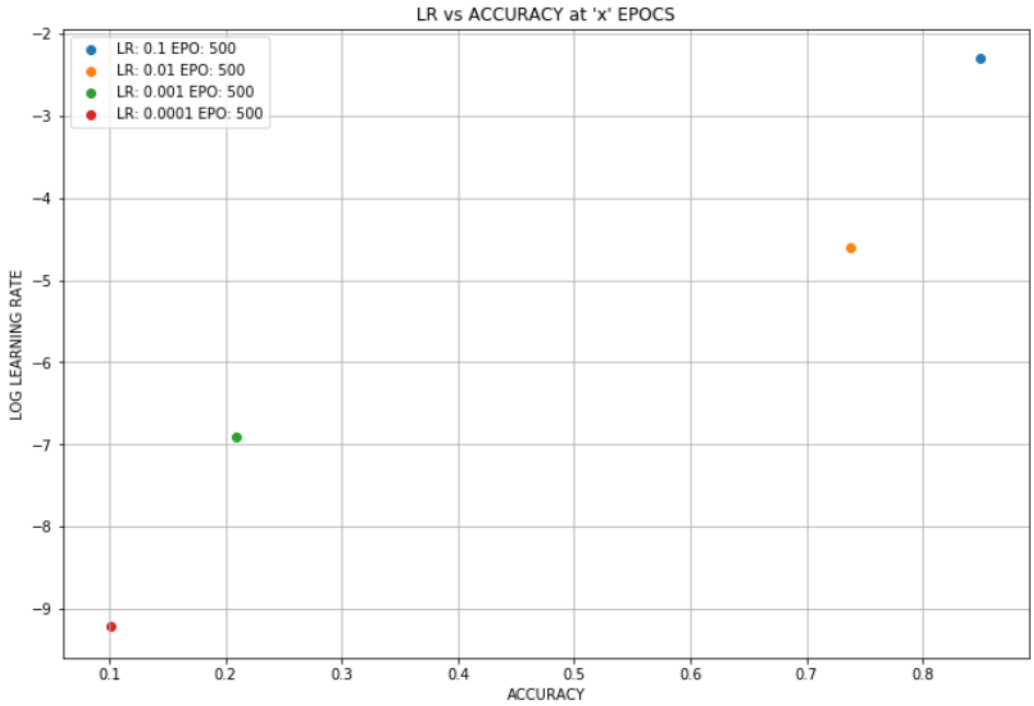
Task 3: CNN and Filter Size

Comparing the CNN for the following LRs – {0.1, 0.01, 0.001, 0.0001}, with the filter size being 1x1, 3x3 and 5x5, at constant epochs 500. Graph for Filter Size 3x3 provided above.

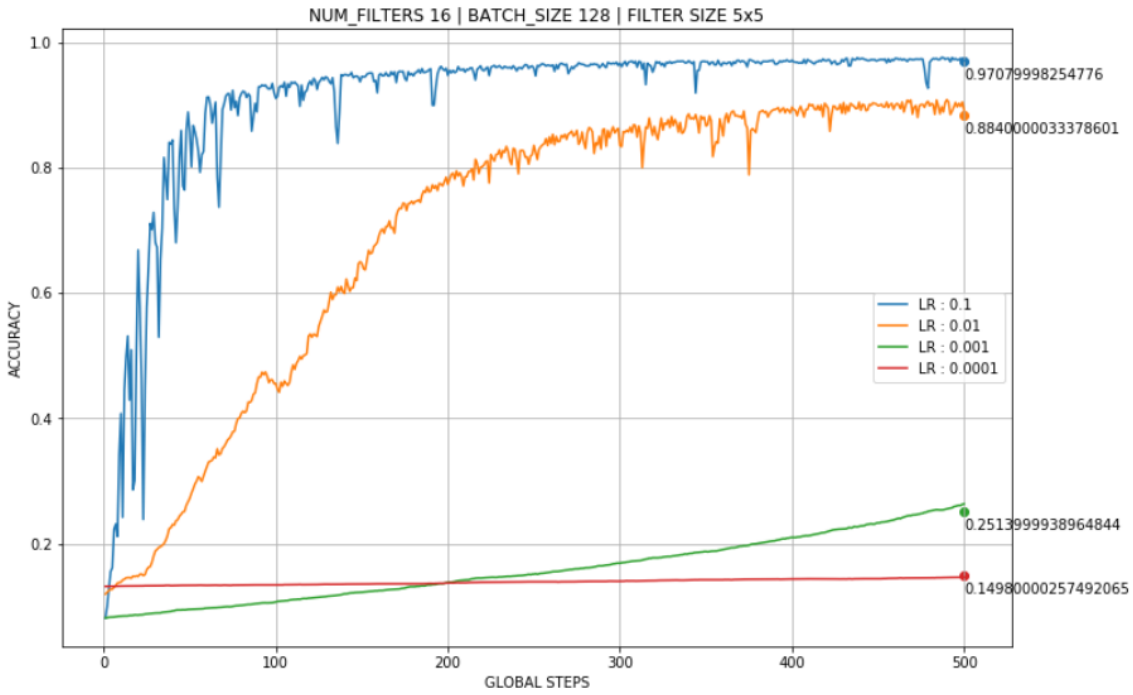
Filter Size 1x1, Accuracy vs Global Steps for all 4 LRs



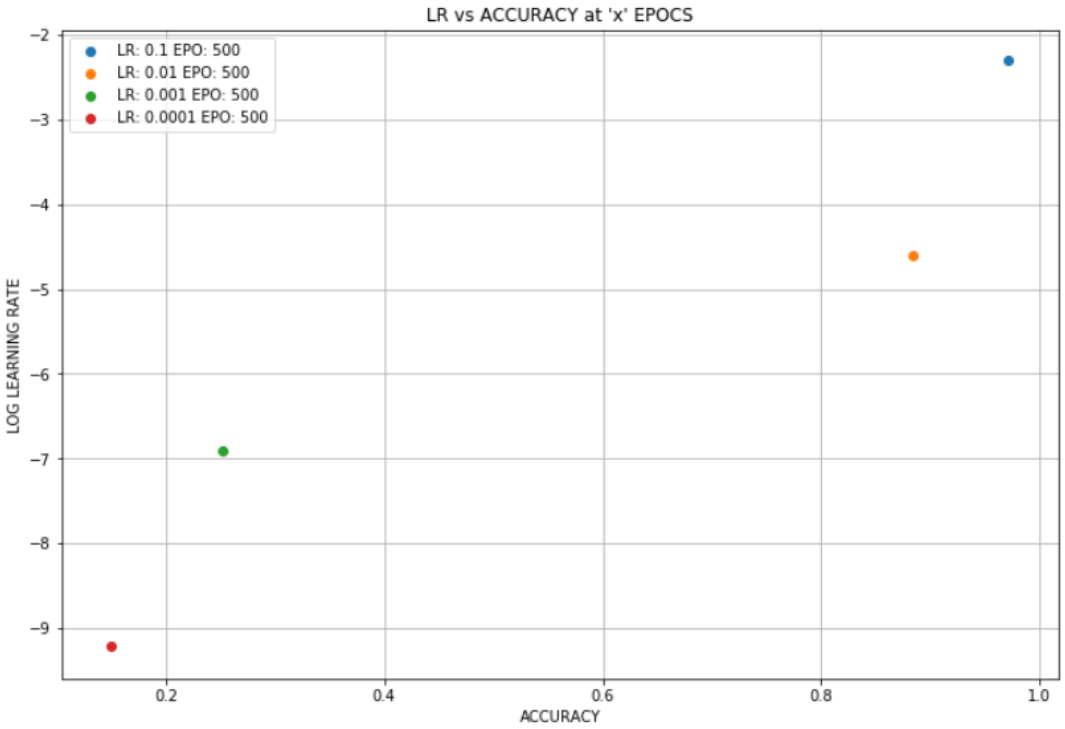
Filter Size 1x1, Accuracy vs LR for 500 Epochs



Filter Size 5x5, Accuracy vs Global Steps for all 4 LRs



Filter Size 5x5, Accuracy vs LR for 500 Epochs



The computation time increases with an increase in the filter size. This is expected as the number of operations increase per convolution for a larger filter. Larger filters have more expressive power and can detect more complex features, so they may be useful in cases where the data/image is complex or composed of a lot of objects.

Observations – Constant Epochs 500, Varying LR, Varying Filter Size

Learning Rate	Filter Size	Validation Accuracy (%)	Comments
0.1	1x1	84.89	Max accuracy and early convergence is observed for LR 0.1. Increase in Filter Size has a positive effect on the accuracy.
	3x3	96.51	
	5x5	97.08	
0.01	1x1	74.79	The accuracy and convergence rates initially increase with change in Filter Size and then stay about the same.
	3x3	88.56	
	5x5	88.40	
0.001	1x1	20.93	A smaller filter size works better in this case. This leads to evaluating the trade-off between changing the LR vs changing the Filter Size.
	3x3	28.52	
	5x5	25.14	
0.0001	1x1	10.11	There is little to no learning observed for 0.0001 due to dying ReLU or vanishing gradients. Change in Filter Size causes slight improvement but it is not substantial.
	3x3	9.12	
	5x5	14.98	

Task 4: Random Search

CNN for Random Search:
I/P → CL1 → RELU → POOL1 → CL2 → RELU → POOL2 → DROP → FC → RELU → DROP → SOFTMAX

```
Best found configuration: {'batch_size': 40, 'filter_size': 5, 'learnin
g_rate': 0.08695726656130962, 'num_filters': 21}
Test Accuracy:
0.9957
Test Error:
0.0042999999999999705
```