Introduction to Machine Learning - 10601
Homework 7

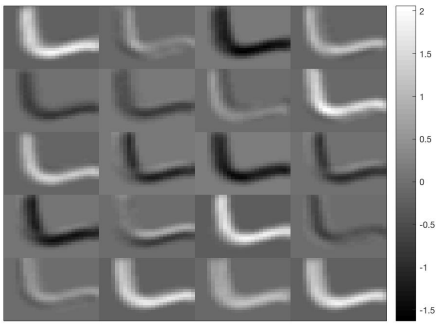
[redacted]
Andrew id: [redacted]

Start date: Apr 5th
Due date: Apr 17rd

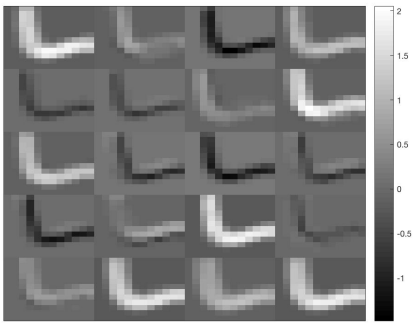
Feature Visualization

Q 8: Visualize the output of the second and third layers. Show 20 images from each layer on a single figure file.

I used MATLAB function *imaqmontage* to visualize 20 images in one figure. The visualization of output of second layer $(24 * 24 * 20)$ and the third layer $(12 * 12 * 20)$ are as below.



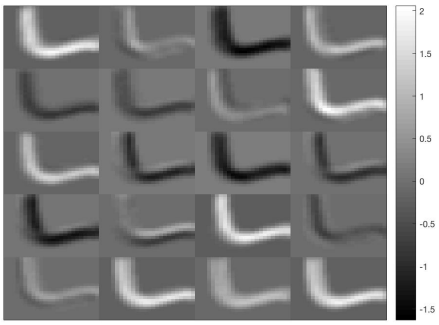
(a) Output of second layer



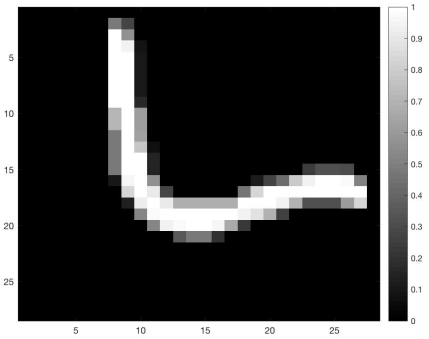
(b) Output of third layer

Q9 (a): Compare the output of the second layer and the original image (output of the first layer), what changes do you find?

The output of the first layer and the second layer are as below. Each image of the 20 images is an output of a convolution filter and extracted various features of the input image. From these two images we could see the 20 channels extracted different features from the original image, like some are horizontal edges and some are vertical edges, etc.



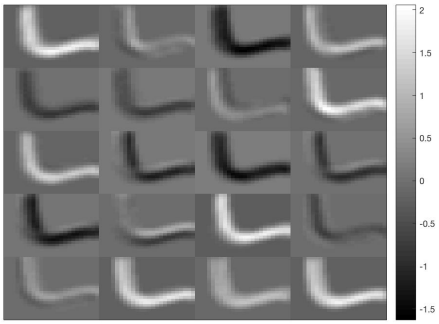
(a) Output of second layer



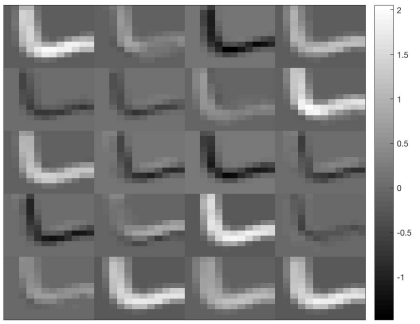
(b) Output of first layer

Q9 (b): Compare the output of the third layer and the output of the second layer, what changes do you find?

The output of the third layer and the second layer are as below. The output of the second layer (convolution) is of size $24*24*20$ and the output of the third layer is of size $12*12*20$, so the total number of channels are remain the same but with a smaller dimension in each channel. From these two images we could see even though the output of the third layer is relatively coarser than the output of the second layer however still represented the same feature of the original image.



(a) Output of second layer



(b) Output of third layer

Q9 (c): Explain your observation.

From the input image to the output of the convolution layer. We know that by convolution we could handle an image in many ways like blur it or detect the edges. With the filter (convolution kernel) applied to a part of the image, the output was computed by the pixels' value in this area and the parameter of the filter. With different parameters, we could detect different interested information of the input image. From the output of the second layer we could see they detected different features from the input image, they all have the same shape with the input but each of them emphasize a specific characteristic of the input like intensity, curve, edge, etc.

In the third layer (max pooling layer), a $K * K$ window was applied to each channel of the input and extract the maximum of them, and this window passing through the input with a known stride value. The reason we can do this and keep the feature unchanged at the same time is that neighbor pixels in an image has similar values. By max pooling in a small area of input, we basically picked out the most "representative" value among all pixels in this area. Besides, we diminished the amount of total pixels which in turn diminished the computation in the Network, it could also prevent out model from over fitting the training dataset. From the two images we could see each of the 20 channels still represent the same feature which the convolution layer detected but with a smaller dimension in each channel (The total number of channels didn't change), which means we kept the information we want and diminish the computation at the same time.

Q10: Run your final code on the entire MNIST dataset by setting fullset to true in testLeNet. Report the final test accuracy and time taken by testLeNet to run the training procedure for 10,000 iterations.

Below is a plot of the test accuracy over the whole dataset. I did not change the set of the test code so the test accuracy was print every 500 iterations. I marked some coordinates with data cursor. From the plot we could see the final accuracy was 98.93%, and at about 9500 iteration the accuracy was 99.09%, which near to 99.1%.

The total time of this Network is : 8174s (173 min).

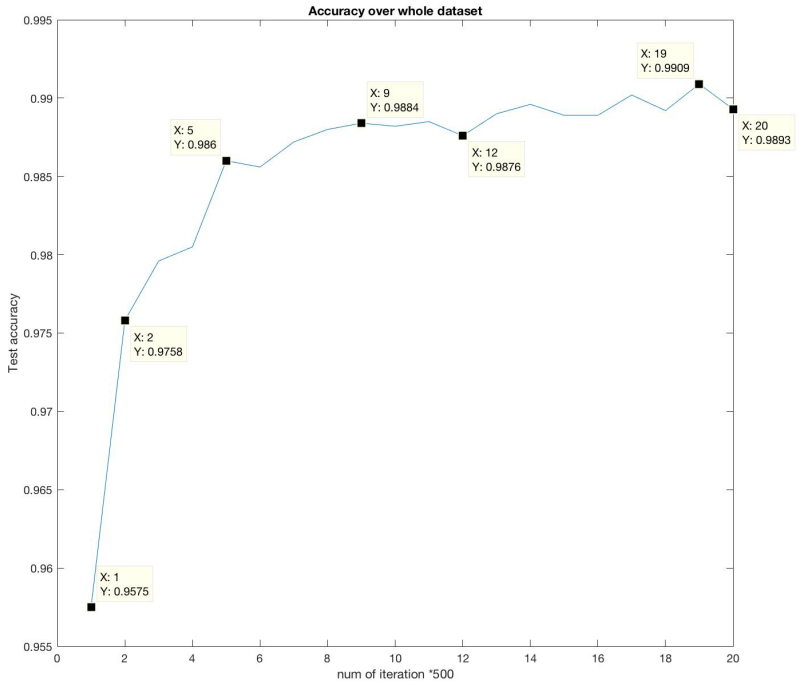


Figure 4: Test accuracy of whole dataset

Collaboration Policy

– Did you receive any help whatsoever from anyone in solving this assignment? **No.**
– Did you give any help whatsoever to anyone in solving this assignment? **No.**
– How many hours did this assignment take? **About 30 hours.**