**Programming Problem**

**MINIST Features Visualization with CNN**

In this question, you will design and implement a simple Convolutional Neural Network (CNN) to classify and visualize the distribution of MNIST ten handwritten digits features in a 2D scatter plot.

The MNIST database of handwritten digits, has a training set of 60,000 examples, and a test set of 10,000 examples. The digits have been size-normalized and centered in a fixed-size image. The original source of MNIST is http://yann.lecun.com/exdb/mnist/, but you may take it wherever available.

**Train your network.** The network would be trained via Cross Entropy with mini-batch GD. Describe the information that associated to the network training:

 Data Preprocessing, data augmentation (If any), initialization

 Optimal hyperparameters such as learning rate schedule, momentum coefficient, L2 coefficient, dropout rate, batch number etc.

Present them systematically in the *table form*.

(a) Plot the following:

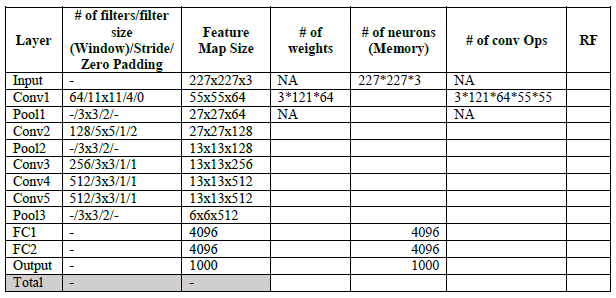
 Training and test loss (not classification accuracy) vs epoch.

 Classification accuracy on the training and test set vs epoch.

(b) Fill the table with the final accuracy that you obtained.

|  |  |
| --- | --- |
| **Training Accuracy (%)** | **Testing Accuracy (%)** |

1. A ConvNet with following specifications is given:



(a) Fill the table and calculate the *total weights*, *total memory per image* consumed (assuming 4 bytes floating points decision) and *total convolution operation*s during prediction.

PS: # of conv Ops includes FC layers as they are also considered as a special case of convolution.



(b) Does Receptive Field at Conv5 manage to cover the input image size?

Yes, 163is large enugh to cover 13\*13

(c) Calculate how many percent of weights in the two FC layers occupied in the network.

(18874368+16777216)/ 43809472 = 81 %

2. Suppose ***x****l*-1 and **x***l* are respectively the input and output of the residual block, F(***x***, **W**) is a function of the neuron and ⊕ is the elementwise addition. Given ***y****l*-1 = *h*(***x****l*-1) + F(***x****l*-1, **W***l-*1) and ***x****l* = ReLU(***y****l*-1) as shown in the figure 1 below.



Figure 1: Residual Block

(a) By making two assumptions and state the assumptions you make, show 𝒙𝐿=𝒙𝑙−1+Σ𝐹(𝒙𝑖,𝐖𝑖)𝐿−1𝑖=𝑙−1 where *L* is the total number of the blocks in the ResNet.

(b) Based on the equation obtained from (a), show 𝜕ℇ𝜕𝒙𝑙−1=𝜕ℇ𝜕𝒙𝐿(1+𝜕𝜕𝒙𝑙−1Σ𝐹(𝒙𝑖,𝐖𝑖)𝐿−1𝑖=𝑙−1). Reason why gradient vanishing problem would not occur in this case.

3. Pick one deep learning/computer vision paper that related to your background (job, interest etc). If your area has nothing to do with deep learning/computer vision, just pick any paper related to application of deep learning/computer vision. Write a brief note in **point form** (**not more** than half a A4 paper) with the following:

(a) Your background and why you choose this paper.

(b) The problem statement, i.e the problem that authors were trying to solve and what are the challenges of this problem.

(c) Describe their solution in brief. You can copy and paste the figure from the paper to assists your explanation with citation if necessary.