CS6910 : Deep Learning (for Computer Vision)
Programming Assignment 1

Instructor: Prof. Anurag Mittal (BSB 368)

amittal@cse.iitm.ac.in

1 Part-A

Please train a network for image classification on your respective datasets (please ref Sec. 3). You should experiment with the following network parameters:

1. number of convolutional (conv) layers
2. fully connected (fc) layers
3. number of filters in different layers
4. maxpooling
5. training time (number of epochs)
6. stride

to come up a study of the effect of these parameters on the classification performance. Try to improve the performance as much as possible by modifying these parameters. Please present the results of such a study in the form of a table that shows the classification performance as a function of these parameters. Also look at some of the images that are mis-classified and see if there is an explanation for such mis-classifications.

2 Part-B

Reference paper: https://arxiv.org/pdf/1311.2901.pdf

From the models you have trained for Part - A, consider the one that gave you highest accuracy on test data and perform the following experiments on the same test data.

2.1 Occlusion sensitivity experiment

Given a fully-trained high-performance image classifier model, the question arises whether the model has really learnt the location of the object in the image or if the model just classifies the image based on surrounding or contextual cues. In this regard, to understand the behaviour of your model, for some of the selected images ($\sim$ 10 images) from test-set of your dataset, perform occlusion sensitivity experiment as follows: For each pixel position i(along x-direction), j (along y-direction):

1. consider a window ($N \times N$, choose an appropriate value of N) around (i, j) and replace the content of the window with gray pixels. Refer [1] figure 7, 8 for more information.
2. Pass the modified image (with respect to the position (i, j)) through the model and note down the probability for true class into an array. i.e., confidence(i, j). Plot the confidence array as an image and comment on the observations.

2.2 Filter Analysis

2.2.1 Filter Identification

Please choose some 10 random filters from your model, 2 in each layer. Then, find images where the response in this filter is the maximum across all images possible. Display some top 5 of the image patches corresponding to the maximum response position in the filter map.

2.2.2 Filter Modification

Now, switch off these 10 filters by making their weights 0. Find classes and images that start to mis-classify now, but were classified correctly earlier.

Please comment on the result obtained.

3 Dataset and boiler-plate code

3.1 Dataset

You must perform all the experiments only using the dataset-split alloted to you, refer this sheet for the same. These datasets are hand-crafted from Cifar10[2, 3] and Mini-ImageNet[4] datasets for the purpose of this assignment. A brief overview of the same is given below. Refer Mini-Imagenet and Cifar10.

Dataset #	Description	Total Images #		
		Train	Test	Val
1	5 classes from CIFAR 10	5000 images / class	500 images / class	500 images / class
2	5 classes from CIFAR 10	5000 images / class	500 images / class	500 images / class
3	33 classes from mini-ImageNet	400 images / class	100 images / class	100 images / class
4	33 classes from mini-ImageNet	400 images / class	100 images / class	100 images / class
5	33 classes from mini-ImageNet	400 images / class	100 images / class	100 images / class

Download your dataset from here.

3.2 Boilerplate code

We recommend to use PyTorch and the boilerplate code provided here. Please note that the testing should be done on a test set different from the training set, that should be separated into training and validation sets to check for over-fitting. This is already done in the provided codes and you need not fiddle around with it. (please check with the TAs in case of any doubts about this or if you write your own code, as this will be covered in later classes).

- The given code is inspired from the PyTorch tutorial (`https://pytorch.org/tutorials/beginner/blitz/cifar10_tutorial.html`)
- The given code contains only a basic model definition (lines 55 - 76)

Plagiarism

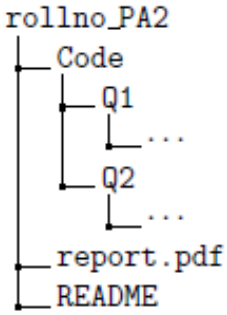
- You should do the assignment yourself. In case you take help from others, please mention in the pdf submitted.
- No sharing of code/experiments etc. will be allowed under any circumstances and may attract disciplinary action by the institute disciplinary committee.

Suggested Programming languages :

Python with PyTorch.

Submission Details

- **Deadline :** 10/15/2020 11 : 59 PM IST
- **What to submit :** You should prepare a report of the results obtained of your work. LaTeX is recommended for ease of work, but not essential. Submit a single tar/zip file containing the following files in the specified directory structure. Use the following naming convention: *rollno_PA1.tar.gz* or *rollno_PA1.zip*. A sample submission would look like this:



- **PDF & Code Upload:** On Moodle.

TAs:

- Gouthaman KV
- Arulkumar
- Asrar Ahmed
- Saikat Dutta
- Pawan Prasad

Please ask your doubts via the Moodle forum.

References

[1] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *European conference on computer vision*, pp. 818–833, Springer, 2014.

[2] A. Krizhevsky *et al.*, "Learning multiple layers of features from tiny images," 2009.

[3] "Cifar10, cifar100 datasets." https://www.cs.toronto.edu/~kriz/cifar.html. Accessed: 2020-09-25.

[4] "Miniimagenet dataset." https://github.com/yaoyao-liu/mini-imagenet-tools. Accessed: 2020-09-25.