CS7015 : Deep Learning (for Computer Vision)
Programming Assignment 1

Instructor: Dr. Anurag Mittal (BSB 368)

amittal@cse.iitm.ac.in

1 Part-1

Please train a network for image classification on the CIFAR-10 dataset (https://www.cs.toronto.edu/~kriz/cifar.html). You should test out different network configurations and 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 network 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.

We recommend to use PyTorch and a bolier plate code for PyTorch have been provided. 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).

Download the boiler-plate code from here.

PyTorch

- The given code is inspired from the PyTorch tutorial (https://pytorch.org/tutorials/beginner/blitz/cifar10_tutorial.html)
- The code automatically downloads the dataset locally for the first time.
- The given code contains only a basic model definition (line number 43 to 61)

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 learned 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 ($\sim$ 15 images) from CIFAR-10 test set, perform occlusion sensitivity experiment as follows: For each pixel position i(along x-direction), j (along y-direction):

1. consider a window ($3 \times 3, 5 \times 5, 7 \times 7, 9 \times 9$) around (i, j) and replace the content of the window with gray pixels. Refer [?] 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.

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 dates

- **Deadline :** 28/08/2019 11 : 59 PM IST
- **PDF & Code Upload:** The details will be shared in Moodle.

TAs:

- Gouthaman KV (gouthamankv@gmail.com)
- ArulKumar (arul.csecit@ymail.com)
- Vikram Singh (vik1409@gmail.com)