Homework 2

TAs

September 2015

1 Instructions

For this lab, you are allowed to use any of the following toolboxes: Theano, Torch and Caffe.

- Send your handin in one zip file named **yourandrewid.tar.gz (or .zip)**
- Submit your code; please use logical file names relating to the assignments.
- Please provide a README file with instructions on how to execute your code.
- Submit one pdf file with all your answers, figures etc.
- Submit your homework to the homework directory in AWS using the following command or FTP tools if you are windows users

```
scp -i 11785_homework.pem yourandrewid.tar.gz
student@ec2-54-209-64-139.compute-1.amazonaws.com:/home/ubuntu/HW2/
```

- For late submission, each day, your score will be one grade lower, for example, for A to A-.
- please avoid submitting large files.

Released: Fri, 2015/09/25
Due: Wed, 2015/10/07

2 Part1

Download PASCAL VOC2007 dataset from VOC2007 website. This dataset includes training, validation and testing sets. You can find the images inside JPEGImages folder and the labels inside the Imagets/Main folder. Together there are 2501 training images, 2510 validation images and 4952 testing images.

Train a Denosing Autoencoder with different noise levels using the training set and visualize the learned filters as in here. Here you can use the whole image (re-size to a fixed size, e.g., 256x256) or sub-sampled small patches with a fixed size (e.g., 30x30). Comparing both of them will give you extra credits.

3 Part 2

Given the PASCAL dataset from Part 1, train a Stacked Denosing Autoencoder as in here using the training set and use it as:

- a fixed network to extract features. Train RBF SVMs using the extracted features from the training set. Again, here you can use a small network trained with sub-sampled inputs and average the output of the network to get image-level features when you train the SVMs.
- a pre-trained network to initialize a feed-forward network and fine-tune the feed-forward network using the training set. Here you will have to use the whole re-sized images as the inputs. Given the fact that an image may have multiple labels in this dataset, you may not want to use a softmax layer as your output layer.

Explore the network structure and SVM parameters to achieve the best validation accuracy (mAP for SVM) you can get. Report the best network structure, SVM parameters the corresponding testing mAP.