Week 2

11-777

Due: Thursday, Sept 14, by 4:30 pm

Instructions

A small yet significant part of the course is answering a few questions based on a paper you just read. By design, some of these questions are open-ended and the purpose is to start a thinking process. Some of those questions may already have been answered in recent literature (or in the given paper itself). We encourage you to, both, come up with your own ideas and survey recent papers before answering these questions.

A few key points to take care of:

- Answer **two out of the three questions** (you are welcome to answer all 3), unless stated otherwise.
- A lot of questions ask you to suggest changes or make alterations. It would be nice to support it with logical / mathematical arguments. Cite all sources used in the process of coming up with your answer. Figures and equations, if they support your arguments will be appreciated.

Questions

Visualizing and Understanding Convolutional Networks

1. Training CNNs is an arduous process which might take many days to give a reasonably trained model. Why and how do pre-trained weights help speed it up? What are the other ways to speed up this process of learning visual representations?

2. How can the architecture of a CNN be improved using the interpretable output obtained from the deconvnet? Is capturing different types of visual features more useful for different downstream tasks? Give some examples. Can you think of a general representation which will be effective across different types of inputs and targets?

3. Pooling seems to be a fairly important component of a convolutional neural network (CNN). What features of the pooling operation make it indispensable to CNNs? Or can CNNs learn representations without bringing pooling operation in the mix? Explain the process of unpooling (also based on the figure in the paper). Can you think of a better way to perform this operation?