

Reading Group: Deep Convolutional Neural Networks

600.482/682 Deep Learning

Spring 2020

March 2, 2020

Reading groups are a good way to keep track of the most relevant research papers (which are now produced at an astronomical rate!). In reading group discussion we can collectively figure out the key points/findings of a paper, e.g parts that we could not understand might be trivial for a peer. Moreover, trying to explain what we know to someone else can reinforce our knowledge or reveal some holes in our understanding. We can also leverage the ‘wisdom of the crowds’ to gain a wider perspective on the paper by considering points raised by others that we missed.

In this session, we will discuss three groundbreaking papers on convolutional neural networks for computer vision tasks. These papers will show you that even though in theory a sufficiently large network can learn any decision boundary, this is far from truth in practice. Designing effective neural networks requires: 1) choosing the optimal architecture on a virtually infinite configuration space (you can take any network and just add another neuron to get a new architecture); 2) finding appropriate/sufficient data to train the model; 3) applying a gradient descent based algorithm to estimate the parameters; 4) additional tricks of the trade.

Papers

Our intuitions about what is going on in the high-dimensional, non-linear spaces induced by deep neural networks are often wrong. These papers show how successful models were built and what was reasoning behind some of the design choices. You will notice that all of them had the same underlying goal: scale-up convolutional neural networks in hopes of learning better models. However, some innovations aimed to improve architecture design, while others tried to tackle some of the thorny optimization issues that we encounter when training such models. Why are deeper models harder to design and optimize? After reading these papers you should be able to answer these questions and have better intuitions about how to go about designing your own groundbreaking architectures. Moreover, the models proposed in these papers are publicly available so they are good baselines to consider when you start working on your projects.

1. ImageNet Classification with Deep Convolutional Neural Networks [<https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>]
2. Very Deep Convolutional Networks for Large-Scale Image Recognition [<https://arxiv.org/pdf/1409.1556.pdf>]
3. Deep Residual Learning for Image Recognition [http://openaccess.thecvf.com/content_cvpr_2016/papers/He_Deep_Residual_Learning_CVPR_2016_paper.pdf]

Questions

Please try to answer the following questions, for each of the papers. Notice that more recent papers build on the successes of the previous ones. Try to compare and relate the findings on this set of papers as a whole. What trends can you observe?

1. Architecture

- (a) What are the output sizes of the first two conv layers?
 - (b) How many parameters in each layer?
 - (c) How was subsampling performed?
 - (d) What new components, techniques or tricks were proposed?
2. Training
- (a) How were the inputs pre-processed?
 - (b) What kinds of regularization were used?
 - (c) How were the parameters initialized?
 - (d) How was training parallelized?
 - (e) How was the data augmented?
3. Inference
- (a) How were the inputs pre-processed?
 - (b) How were predictions made?

Rules of Engagement

Reading groups are normally held by smaller groups but this will not be an obstacle — we will just have to make some adjustments to the format. We divided the class into three groups (one per paper) and each of you was already assigned to a group (see Piazza, if you are not in any group send an email to samir@jhu.edu). The discussion will be lead by the instructor and TAs and everyone is expected to participate. Here is a summary of the rules of engagement.

1. Each group will pick one champion to summarize the paper in 3-5 mins (ideally a volunteer). The champions join the panel and receives extra points for participation.
2. The other group members keep track of the discussion. In the following class, the group compares notes and puts together a presentation in 10-15 mins (5-10 slides) to be posted on Piazza. In the meantime, the members of the panel including the Champions can just chill.