Individual Final Report

Introduction

We took a while to decide what dataset we want to explore for the project. We first chose a small Hand Signs dataset with black and white images produced with infrared technology. Then we think this would be too simple and we wondered in tackle some object recognition problem. Due to the time constraints and knowledge limitation we decided to stick with image processing but using a more challenging data. Finally, we get back to the hand signs problem but this time using a much bigger dataset with colored images.

Individual work

My duties in the project were:

- Solve the second exam.
- Research previous work about hand gesture recognition and the approaches that have been used.
- Writing the bibliographic review of the report and assembling the whole report using LaTex.

I have spent a good amount of time selecting pertinent material related to hand gesture recognition in general and, more specifically, the sign language. Then I read and summarized some of the more relevant papers. In this process, I understood better what was the convolutional models used in these previous works, and started to study about them. More specifically, I studied about LeNet5, AlexNet, VGG Net, and GoogleLeNet, and also read the original papers about LeNet, Network In Network, and Google inception modules. With this better understanding, I could give some advice of how we should proceed to devise our solution proposal.

I also helped Junior and Rahul doing the implementation. In the very beginning, I had some complications to load the data. The idea was to do something similar as in the code of the first exam but there the data was already embedded in the PyTorch package. So, took a while to figure it out that I should use ImageFolder, what exactly type of transformations should be done, and how to use the class *SubsetRandomSampler* to split the data. These first tests can be seen in the file *HandGesture.py*. My final contribution in this regard is the function *getTrainTest* that can be found in the file *ASLAlexNet_II.py*.

I also helped Rahul in coding the ROC Curve.

Results

My contribution this time was more theoretical. I released a draft of the report with much antecedence and I think the summary I have written about well known CNN architectures was useful for the team during the implementation task.

Conclusion

During my research and writing I learned much more about convolution and its role as feature extractor. Became clearer why it works with images (highly spatially correlated) and how the whole field of representation learning is interesting and important. I also particularly enjoyed having understood better about Network In Network and its idea of enhancing the abstract power of the convolutions operations by replacing linear transformations for a more powerful function approximator like MLP. During the paper, the authors explain about maxout layers, and how they form a piecewise linear function which is capable of modeling any convex function. I learned the concept of piecewise linear functions and how they can be used to approximate convex functions in the Linear Optimization course, what helped me to understand better the NIN paper. I am glad that happened as I decided to take the optimization course to help me to go deeper in the machine learning algorithms.