Group 11: Review

October 19, 2019

1 Summary

In this report, the author runs experiments to compare the efficacy of feature extraction from deep CNNs and classification using Logistic Regression, Support Vector Machines and Linear Discriminant Analysis with feature extraction using scattering nets and using the aforementioned models for hand written digit recognition (on the MNIST dataset). The report is a concise description of the results.

2 Strengths

This report is pleasingly concise. Overall there is a decent amount of rigour in the experiments. The author is very clear on the apparent limitations of the report/experiments and clearly mentions those in section 1.1. The models run are clearly indicated ahead of all the analysis along with the size of their feature vectors.

3 Weaknesses

The weaknesses of this report are mostly technical and are covered in the section on technical quality. To summarize however, I am not exactly satisfied with the comparison of features extracted from different models (with different feature spaces) to rate the models as good or bad. While it is correct to analyze this retroactively once the prediction accuracies are available, without those metrics in place, I am not sure it is exactly rigorous to compare models like this.

No implementation details are provided.

4 Evaluation on Clarity and quality of writing

Overall: 4/5 (good)

The language used overall is very concise and contains no useless information. It would however be good to label the blobs in each plot in the visualization matrix to show which target output (digit) it represents. For example, it would be good to know

if the blue cluster in those plots represents the digits 2 or 3 (or whatever it actually represents). This would make it easy to intuitively argue why certain clusters are closer together or farther apart.

Typos and grammatical errors are to a minimum in the report and thus there is no hindrance to the quick reader. In Section 3.3, the heading erroneously says "Feature Visualization (VGG 19)" while the CNN in consideration is ResNet50.

5 Evaluation on Technical Quality

Overall: 4/5 (good)

From a technical point of view, the only objection that could possibly be raised in this otherwise rather rigorous report, is that it is not exactly ideal to compare the feature extraction capabilities of a pre-trained CNN by visualizing the output features in a low dimensional space. Especially when the features from different models live in completely different feature spaces. The resultant low dimensional mappings are a function of the dimensionality reduction methology used and the well formed clusters in the visualization cannot solely be attributed to the model only. As is evident from the visualizations, different dimensionality reduction methologies have very different outputs and the "clusters" have wildly different shapes. In order to be thorough one needs to define a certain type of distance metric between the original high dimensional representations and just the euclidean distance in a low dimensional setting might not be enough. While it is good to reduce dimensionality of the feature vectors (that aid humans to visualize the outputs), I am not certain that it should be considered gospel in model comparison.

6 Overall rating

Overall: 4/5 (good)

Overall, in my opinion, the experiments done as part of the report agree with the theme of the project and are adequately rigorous. The conclusions seem well thought out and are substantiated by experimental results. It would be nice to provide implementation details/package some code along with the report. Without that, I am not certain how the results mentioned in the report can be independently verified.

7 Confidence on assessment

Overall: 3

I have read the report in detail. However, I have not been able to double check the results as the code is not linked or packaged with the report.