Assignment 3
CSCI 6515 – Fall 2021
Posting date: Nov 22, 2021
Due Date: Dec 7, 2021, 11.30 PM (Halifax Time)

Background Information

A vision system should be able to identify what are the main objects in the scene while navigating in an environment. Suppose that the client wants to automate the delivery of mails by using drones that will carry the package and delivery to the correspondent address. For this system, one task that the drone should be able to accomplish is the recognition of the houses' numbers. Motivated by this task, we try to find a model that can recognize the house numbers and match them to the number on the delivery address of the package. We mainly focus on number recognition. Therefore, we will do experiments on the Street View House Numbers (SVHN) dataset, which is one of the most popular object recognition datasets that contain houses' numbers. We attempt to find a Convolution Neural Network model that is simple and still able to provide a good precision in number recognition. Finally, we decided to do a few experiments with different CNN architectures, parameters, and optimizers in order to detect the number in the image.

To access and collect the SVHN dataset:
http://ufldl.stanford.edu/housenumbers/

Note that this dataset is provided in two different formats: 1) original format, 2) cropped format. We suggest doing the assignment with format 2 as it seems simpler.
Also, note that you may use open-source libraries such as SK-learn, NumPy, and Pandas. However, feel free to use other open-source libraries.

Your Tasks:

[1] Your first task is doing some research on the dataset and how they solve this task on literature. You may pre-process the images. In this step, you do a descriptive analysis of your data to better understand it. This step creates a dataset to you be able to work on. You also have to include one summary visualization of the data.
i. Describe the dataset, the format selected and why you selected that format.
ii. Brief summary of approaches used in literature.
iii. Describe any preprocessing used and why.

[2] In this task, you have to apply CNN model on the dataset in order to recognize the numbers. In this task, you have to apply layers of CNN, max-pooling, Fully-connected layers. Do experiments varying the number and the order of the layers. For this task use the same optimizer and same optimizer parameters. Do diagrams to show the architectures you executed.

i. Brief description how CNN works.

ii. Describe how you split the dataset.

iii. Select an evaluation metric and explain why you select that.

iv. Describe the obtained results.

v. Analysis on the results, such as overfitting/underfitting, variance bias, statistical test. Which model provide the best performance?

[3] In this task, you have to apply the best CNN model found in task 2. In this model, change max-pooling to average-pooling and analyze the results. Next, in the best model obtained, check if the addition of batch-normalization layers will improve the results. For this task use the same optimizer and same optimizer parameters. Do diagrams to show the architectures you executed.

i. Describe the obtained results when changing the pooling operation.

ii. Do an analysis and describe conclusions and insights in terms of pooling operation. Which model provide the best performance?

iii. Describe the obtained results when including the batch-normalization layer.

iv. Do an analysis and describe conclusions and insights in terms of batch-normalization layer. Which model provide the best performance?

[4] In this task, you have to do experiments changing the optimizer and its parameters using the best model you found.

i. Describe the obtained results.

ii. Do an analysis and describe conclusions and insights in terms of the optimization.

[5] In this task, you have to do experiments in your CNN models. Based on your analysis, which modification you can do to improve your model? Feel free to do modifications in the architecture or in the optimizer.

i. Describe the ideas or approaches you believe will improve your model and why?

ii. Describe the results.

iii. Describe conclusions and insights.

How to Submit

Assignment 3 will be done in Jupyter Notebook. After you finished your coding and make sure it is working well, you run your code step by step in a Python notebook. As you can run many models, select in maximum 3 of them for each task or just the main ones and print the notebook in a PDF file called code.PDF. Please, keep your code organized and commented. Then you prepare a document and answer the questions in the document. You must add one page describing what you have done before your answers. You also add one page on the summary of your results at the end of the document. In the end, you add all the references you have used in this assignment. Therefore, this document includes the following:

1. Dataset description and preprocessing (Task 1)
2. Task description followed by the questions (Task 2)
3. Task description followed by the questions (Task 3)
4. Task description followed by the questions (Task 4)
5. Task description followed by the questions (Task 5)
6. Summary of your results and conclusions
7. References

We call this file Report.PDF. Now you merge Report.PDF and code.PDF and create your final document to submit. Therefore, the final document (FinalReport_YourBannerID.PDF) includes the content from both report.pdf and code.pdf files.

Now you submit the FinalReport_YourBannerID.PDF to Brightspace.