R Noah Padilla

11/24/2020

Dealing with Small Datasets

Introduction

In this exercise out goal was to use a small set of labeled data and are supposed to implement several techniques to improve the accuracy. These experiments were conducted in the Google Colab environment. The methods performed are shown below.

Implement and test the following techniques (you may combine them):

1. Data augmentation for training. For each training image x, create 8 additional ones by shifting x by one row and/or one column in all 8 directions.
2. Data augmentation for training. Generate additional images by averaging two images of the same class using random weights, as described in the slides.
3. Data augmentation for testing. For each testing image x, create 8 additional ones by shifting x by one row and/or one column in all 8 directions. Average the prediction for the 9 images and return the maximum as the final prediction for x.
4. Semi-supervised learning. Use the unlabeled set (without using the y values) to try to improve your results
5. Adversarial learning.
6. Any other technique you want, as long as it doesn't involve accessing additional data.

Results

I used the base code as a reference for each experiment and only considered that last 5/20 epochs. The base results and methods (1-6) are shown below.

Base Results

EPOCH	LOSS	ACCURACY
15	.5417	.8067
16	.5511	.8025
17	.5580	.7914
18	.5837	.7862
19	.5830	.7881
20	.5533	.8080

 # 1-SHIFTING THE IMAGES

EPOCH	LOSS	ACCURACY
15	.5475	.7968
16	.5530	.7895
17	.5294	.8065
18	.5591	.7894
19	.5285	.8065
20	.5319	.8107

RESULTS

With the results shown above, a little to no increase in accuracy was obtained using the shift technique.