CSE 591 - Introduction to Deep Learning
Mini Project 6

Bijan Fakhri
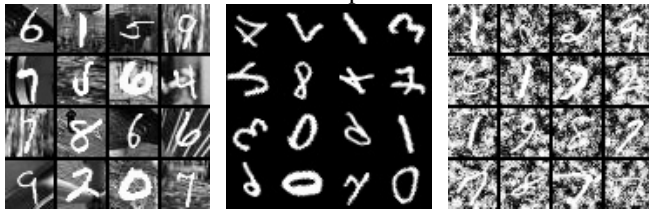*School of Computing, Informatics and*
Decision Systems Engineering
Arizona State University
Tempe, Arizona 85044
Email: bfakhri@asu.edu

Abstract—

1. Introduction

2. Autoencoders

Three datasets were tested in these experiments, all derivatives of the MNIST handwritten digit dataset. Below are YANN visualizer outputs of the three datasets.

From left to right, the datasets above are: the original MNIST dataset, a rotated MNIST dataset, and a noisy MNIST dataset. To measure the generalities, we prejudiced a network using one dataset, recorded its performance, then retrained the softmax layer of that network on another dataset. The generalization performance was then calculated using equation 2. This was done on all possible combinations of the three datasets.

3. Generative Adversarial Networks (GAN)

4. De-Noising with Autoencoders

5. Results

The results of the experiments are tabulated below. The dataset in the left column corresponding to a cell in the tabel represents the dataset on which the CNN was prejudiced with. The dataset in the top row corresponding to a cell represents the dataset the CNN was tested on (after training the softmax layer).

MNIST Dataset	Original	Rotated	Noisy
Original	1.00	0.27	0.39
Rotated	0.32	1.00	0.43
Noisy	0.39	0.28	1.00

From the results it is evident that the rotated version of the MNIST dataset is the most general, sporting an average generality of 0.58, slightly higher than the original and noisy datasets (both at 0.56). Interestingly, the generality of the rotated dataset with respect to the noisy dataset topped the charts at 0.43.

6. Conclusion and Discussion

The results unveiled interesting patterns in the generality of the datasets. The generalities with respect to the rotated dataset are interesting because they are the lowest (at 0.27 and 0.28). This could be because the atomic structures learned by the CNN are not rotation invariant. This could explain why the rotated dataset is the most general - prejudicing with it produces rotation invarient filters. It is also interesting to note that the noisy dataset is equally as general as the original. This is not initially expected because noise could act as a regularizer, but, because the CNN is using dropout, it is likely redundant, handing the rotated dataset the crown.

References

[1] R. Venkatesan, V. Gattupalli, and B. Li, "Neural dataset generality," *CoRR*, vol. abs/1605.04369, 2016. [Online]. Available: http://arxiv.org/abs/1605.04369