

Improved Reliability Using Data Augmentation

Mike Murugan, Amit Kulkarni

(mimuruga@Stanford.edu, kulkarn1@Stanford.edu)

Motivation: For Multi-label classifier network, its hard to ensure its precision and recall for the following two cases.

- Rarely occurring classes
- Lesser amount of training data across classes

Data:

EMNIST – Extended MNIST

- Training Set – 240K Balanced Handwritten digits
- Test Set – 40k Balanced Handwritten digits

Features:

BAGAN: New harder input data set. Normalized to (-1, 1).

Manual: Transformations

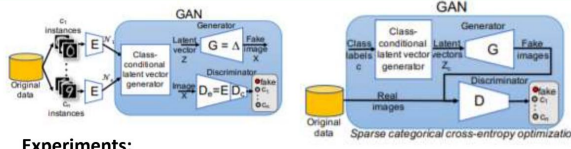
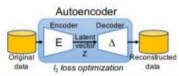
- Rotation
- Gaussian Noise
- Sharpen/Blur

Models:

Multi-label Classifier:



BAGAN:



Experiments:

Classification:

Induce population trimming and imbalance in the data to exaggerate the effect of misclassification error.

Datasets used:



Results:

Dataset (10 classes of handwritten digits)	Fully Balanced Dataset (A)	Dataset with 95% images removed from '0' digit class (B)	Data set from B augmented with BAGAN to balance it (95% of '0' images from BAGAN)	Data set from B augmented with BAGAN to balance it (95% of '0' images from BAGAN)
1% of EMNIST (total images 2400)	Accuracy Score: 0.955 precision recall f1-score '0' 0.98 1.00 0.99	Accuracy Score: 0.9125 precision recall f1-score '0' 0.96 0.60 0.74	Accuracy Score: 0.945 precision recall f1-score '0' 0.97 0.90 0.94	Accuracy Score: 0.9425 precision recall f1-score '0' 0.95 0.97 0.96
20% of EMNIST (Total 48000 images)	Accuracy Score: 0.991875 precision recall f1-score '0' 1.00 0.99 0.99	Accuracy Score: 0.98775 precision recall f1-score '0' 1.00 0.94 0.97	Accuracy Score: 0.99075 precision recall f1-score '0' 1.00 0.95 0.98	Accuracy Score: 0.990375 precision recall f1-score '0' 0.99 1.00 0.99
Full EMNIST dataset (240000 images)	Accuracy Score: 0.9951 precision recall f1-score '0' 1.00 1.00 1.00	Accuracy Score: 0.99255 precision recall f1-score '0' 1.00 0.97 0.98	Accuracy Score: 0.9917 precision recall f1-score '0' 1.00 0.96 0.98	Accuracy Score: 0.994725 precision recall f1-score '0' 0.99 1.00 0.99

Discussion:

For smaller population of data (1% with unbalanced classes (5% in digit "0"), BAGAN based data augmentation helps to improve Multi-label classifier accuracy. For Medium or Larger population of data (20% and more), Manual data augmentation surpasses BAGAN based augmented data in improving Multi-label classifier accuracy. BAGAN is a generic GAN (could be used for other datasets such as photos). Training and computation involved in BAGAN are very high compared manual data augmentation.

Future Work:

BAGAN is good in data augmentation for low data regime with unbalanced classes. But in high data regime generated data lacks diversity. Variable auto encoders comes handy in this case, but would generate blurrier image, hence CVAE GAN is promising area to investigate.

<https://youtu.be/w2Tft-vmnxw>