Motivation: For Multi-label classifier network, its hard to ensure its precision and recall for the following two cases.
   a. Rarely occurring classes
   b. Lesser amount of training data across classes

Data:
EMNIST – Extended MNIST
   a. Training Set – 240K Balanced Handwritten digits
   b. Test Set – 40k Balanced Handwritten digits

Features:
BAGAN: New harder input data set.
Normalized to (-1, 1).
Manual: Transformations
   a. Rotation
   b. Gaussian Noise
   c. Sharpen/Blur

Models:
Multi-label Classifier:

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

Datasets used:

Results:

<table>
<thead>
<tr>
<th>Dataset (5% classes of handwritten digits)</th>
<th>Fully Balanced Dataset (A)</th>
<th>Data set from B augmented with BAGAN to balance k (50% of 5% images from BAGAN)</th>
<th>Data set from B augmented with BAGAN to balance k (5% of 5% images from BAGAN)</th>
</tr>
</thead>
<tbody>
<tr>
<td>70% of EMNIST (total 65K images)</td>
<td>Acc@5 0.9825 precision recall f-score 0.9 1.0 0.97</td>
<td>Acc@5 0.9825 precision recall f-score 0.9 1.0 0.97</td>
<td>Acc@5 0.9825 precision recall f-score 0.9 1.0 0.97</td>
</tr>
<tr>
<td>50% of EMNIST (total 51K images)</td>
<td>Acc@5 0.9865 precision recall f-score 0.9 1.0 0.97</td>
<td>Acc@5 0.9865 precision recall f-score 0.9 1.0 0.97</td>
<td>Acc@5 0.9865 precision recall f-score 0.9 1.0 0.97</td>
</tr>
<tr>
<td>30% of EMNIST (total 31K images)</td>
<td>Acc@5 0.9805 precision recall f-score 0.9 1.0 0.97</td>
<td>Acc@5 0.9805 precision recall f-score 0.9 1.0 0.97</td>
<td>Acc@5 0.9805 precision recall f-score 0.9 1.0 0.97</td>
</tr>
<tr>
<td>15% of EMNIST (total 15K images)</td>
<td>Acc@5 0.9725 precision recall f-score 0.9 1.0 0.97</td>
<td>Acc@5 0.9725 precision recall f-score 0.9 1.0 0.97</td>
<td>Acc@5 0.9725 precision recall f-score 0.9 1.0 0.97</td>
</tr>
</tbody>
</table>

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-vmnnxw