1. Introduction

Highly performance AI-OCR requires large amounts of data. Handwritten data is difficult to collect:
- Cost, Time, Manpower
- Many types of characters

English
- Alphabet (26 types)

Japanese
- Hiragana (46 types)
- Katakana (46 types)
- Kanji (Over 6000 types)

This paper proposes to generate handwritten character images and use them as training data.

- **Goal** of our research
  - Improvement of accuracy of character recognizers using generated images

- **Proposed approach**
  - Use a generative model that combines Y-Autoencoder (Y-AE) with Adaptive instance normalization (AdaIN)

2. Related Works

CycleGAN (font to handwrite Chinese Characters)

- Can only convert a single font image to only a single type of handwritten character image
- Few variations

cCGAN + cCVAE (generated handwritten images)

- The quality of the generated image was not stable

3. Proposed Model

- **Y-AE with AdaIN**
  - 46 different images can be generated from a single image.

4. Experiment

Comparison of F1-score of handwriting recognizers

- 46 characters types

Dataset: ETL
- Japanese handwriting character dataset
- Consists of 9 subsets

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Num. of images</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training</td>
<td>ETL7</td>
</tr>
<tr>
<td>Validation</td>
<td>ETL8G</td>
</tr>
<tr>
<td>Testing</td>
<td>ETL9G</td>
</tr>
</tbody>
</table>

Hyper-parameters of the handwriting recognizer

- Model : VGG16
- Optimizer : Adam
- Learning rate : 1e-4
- Iteration : Adjusted for the same number of data used per training

Results (F1-score [%])

<table>
<thead>
<tr>
<th>Training dataset</th>
<th>F1-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>ETL7 only</td>
<td>0.8664</td>
</tr>
<tr>
<td>Generated image only</td>
<td>0.9192</td>
</tr>
<tr>
<td>ETL7 + generated images</td>
<td>0.9281</td>
</tr>
</tbody>
</table>

5. Future Work

- Work on other characters, more complex character generation (Katakana, Kanji, etc ...)
- More detail analysis of generated images
  - Variations of generated images
  - Comparison with data augmentation, effect of combination
  - Selection of generated images for training