A State-of-the-art Deep Learning Approach for Handwriting-based Recognition on Gender and Handedness

Yanbang Wang, Jiangshan Li, Tiancheng Cai
{ywangdr, jiangsli, caitch}@stanford.edu, Stanford University

Problem & Task

Motivation
- Handwriting-based identification of gender and handedness has a variety of applications in forensic biometrics and archaeology, while most existing methods rely heavily on handcrafted features combined with classic statistical methods.

Results
- We report a deep-learning based method that achieved state-of-the-art performance on both English and Chinese handwriting datasets.
- Our model outperforms humans by a large margin and achieves an error rate at least 30% lower than the existing best models in two different test settings.
- We collect and publish a new handwriting dataset for gender classification consisting of 350 handwriting samples from 32 writers.
- We build a website that allows viewers to predict gender from their own handwriting images.

Features

We use preprocessed handwriting images as input to our CNNs. We consider data augmentation methods including rotation, shifting, rescaling and morphological operations (binary erosion/dilation) to apply to training data.

Models

Binary Classification loss:
\[ J = -\frac{1}{m} \sum_{i=1}^{m} (\alpha y_i \log \hat{y}_i + \beta (1 - y_i) \log (1 - \hat{y}_i)) \]
\[ \alpha, \beta: \text{weights for (potentially skewed categories)} \]

IAM-OnDB
- CNN as shown in FIGURE 3, with BatchNorm, dropout = 0.5, Adam Optimizer

HIT-MW
- CNN as shown in FIGURE 4, with BatchNorm, dropout = 0.3, Adam Optimizer

Results

Table 1: Accuracy results on IAM-OnDB

<table>
<thead>
<tr>
<th>Labels</th>
<th>Train Accuracy</th>
<th>Test Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>0.969</td>
<td>0.958</td>
</tr>
<tr>
<td>Female</td>
<td>0.969</td>
<td>0.739</td>
</tr>
</tbody>
</table>

Table 2: Accuracy results on HIT-MW

<table>
<thead>
<tr>
<th>Labels</th>
<th>Train Accuracy</th>
<th>Test Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>0.782</td>
<td>0.755</td>
</tr>
<tr>
<td>Female</td>
<td>0.782</td>
<td>0.755</td>
</tr>
</tbody>
</table>

Discussion

- We achieve state-of-the-art accuracy on the task of gender classification for both Chinese and English Datasets.
- We didn’t follow up on the classification of handedness because of the skewness of dataset (roughly 9:1).
- The task of gender classification from handwriting is inherently difficult, and we are amazed that we are able to beat human performance (around 64.88% on English) by a large margin.

Future

- We would study the transferability of our model further by integrating a third languages (for example, having train, dev and test sets coming from three different languages).
- We would work on the visualizing what the network has learned.
- We would like to study the fairness issue involved in such model.

Reference

