A Few-Shot Learning Approach for Historical Ciphered Manuscript Recognition

Mohamed Ali Souibgui  Alicia Fornés  Yousri Kessentini  Crina Tudor
Motivation: Historical ciphered manuscript

Recognition process:
Transcription → Cryptanalysis

Difficulties:
1. Unknown alphabets (No dictionaries, or language models)
2. Changeable Symbols from a cipher to another.
3. Few resources (No labeled similar documents)
4. Hard to be segmented (To apply the classic unsupervised methods)
5. Etc ...

Document created using invented symbols to encrypt the content
1 System to Read All Alphabets
Read By Matching/Spotting

- **Input:**
  - Ciphered text line
  - Cropped alphabet symbols

- **Output:**
  - Similarity regions

- **Goal:**
  - Handling different alphabets (different training/ testing sets)
  - Segmentation free method
  - Use a few data

A few shot learning approach for text matching?
Few-shot text recognition

Few-shot Classification

Segmentation
Classification

Segmented symbols + classes

Few-shot Detection
Proposed Approach
Experiments: Data

Training

Omniglot dataset: few examples (20 per character) from 50 alphabets.

Testing

Copiale Lines

Synthetic Lines

Borg Lines
## Results: Without fine tuning

<table>
<thead>
<tr>
<th>Method</th>
<th>Description</th>
<th>Copiale SER</th>
<th>Borg SER</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SNN features + GMM</td>
<td>0.44</td>
<td>0.57</td>
</tr>
<tr>
<td>2</td>
<td>Manual Segmentation + 1</td>
<td>0.37</td>
<td>0.22</td>
</tr>
<tr>
<td>3</td>
<td>SIFT+ KNN + Label propagation</td>
<td>0.44</td>
<td>0.54</td>
</tr>
<tr>
<td>4</td>
<td>3 + Manual clusters cleaning</td>
<td>0.20</td>
<td>0.52</td>
</tr>
<tr>
<td>5</td>
<td>Ours (1 shot)</td>
<td>0.38</td>
<td>0.66</td>
</tr>
<tr>
<td>6</td>
<td>Ours (5 shots)</td>
<td>0.33</td>
<td>0.51</td>
</tr>
</tbody>
</table>

*Training (Synthetic)*

*Testing (Ciphers)*
Results: With fine tuning

<table>
<thead>
<tr>
<th>Method</th>
<th>Description</th>
<th>Copiale</th>
<th></th>
<th>Borg</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Training Pages</td>
<td>SER</td>
<td>Training Pages</td>
<td>SER</td>
</tr>
<tr>
<td>1</td>
<td>MDLSTM</td>
<td>25</td>
<td>0.13</td>
<td>7</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td></td>
<td>34</td>
<td>0.12</td>
<td>9</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td></td>
<td>42</td>
<td>0.08</td>
<td>11</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td></td>
<td>51</td>
<td>0.07</td>
<td>14</td>
<td>0.55</td>
</tr>
<tr>
<td>2</td>
<td>Ours (1 shot)</td>
<td>2</td>
<td>0.10</td>
<td>2</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>0.09</td>
<td>5</td>
<td>0.19</td>
</tr>
<tr>
<td>3</td>
<td>Ours (5 shots)</td>
<td>2</td>
<td>0.10</td>
<td>2</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>0.09</td>
<td>5</td>
<td>0.19</td>
</tr>
</tbody>
</table>
Conclusion and Future directions

Using the proposed approach we were able to:

- Recognize different alphabets using one model in a segmentation free fashion.
- Generalize to unseen data in term of alphabet and writing styles with, or without a few labeled lines.

As Future work:

- Adding language models for some ciphers to improve the decoding.
- Detection in full pages.
- Detecting all the symbol alphabet in one time.
Thank You!