Writer identification for offline Japanese handwritten character using convolutional neural network

Ryosuke Nasuno*, Shuichi Arai

Tokyo City University Graduate of Engineering
Tamazutsumi, Setagaya-ku, Tokyo, 158-0087, Japan

*Corresponding Author: nasuno15@ipl.cs.tcu.ac.jp

Abstract

In this paper, we propose to the some kind of features from Convolutional neural network (CNN) for writer identification. We use dataset of Japanese handwritten character, which consists 100 kind of words from each 100 writers. We evaluate two nature of handwritten words: the potential of writer identification for each word in Japanese and handwritten words contain the writer own unique identities. These nature cause a variation of classification accuracy about each handwritten character and about each writer for same word. The former, difference of accuracy is approximately 90% and the feature of each word from CNN have large influence on the accuracy. The latter, difference of accuracy is about 60% and unique writer style can be used to determine authorship of a handwritten document.

Keywords: offline writer identification, CNNs, Japanese handwritten character

1. Introduction

Writer identification is an important problem to understand written document and forensic application. In historical document analysis, identification of the writers is helpful to understand their historical context\(^1\). And, it is interested in detecting unique writer style of the authors for historical document dating\(^2\)\(^3\). In the area of scientific crime detection, they identify writer distinction for words already written a paper. They inspect shape, composition of strokes, and situation of calligraphy\(^4\) but they judge a diagnosis from their knowledge and experiment. So it will be helpful to learn the unique writer style for judge.

The task of writer identification can be categorized into online writer identification and offline writer identification. Online writer identification can use temporal information which is speed, direction or pen pressure. Offline writer identification can use only handwritten text. In addition, the latter can be categorized into allograph-based and textual-based methods\(^5\). Allograph-base methods rely on the local identifier computed from written words\(^6\)\(^7\). In contrast, textual-based methods rely on the global statistics which is the ink, width or angle\(^8\)\(^9\).

In this research, we propose an allograph-based method for offline Japanese handwritten character. We use activation features learned by CNN. In recent research, CNN become more attractive for many machine learning and computer vision classification problem such as semantic segmentation\(^10\)\(^11\), speech recognition\(^12\), musical chord recognition\(^13\)\(^14\). Moreover, CNN are among the top level on challenges like the Pascal-VOC or ImageNet\(^15\). In spite of this performance, CNN have not shown good performance on writer identification. A reason might be that some kind of handwritten character have a feature to classify easily and others do not have because of too simple structure or not unique writer style. To evaluate this, we show the difference of accuracy about same words written by different writer and about different words written by same writer.

2. Related work

CNN have been widely used in the field of image classification and object recognition. In the ImageNet Large Scale Visual Recognition Challenge, CNNs are among the top contenders\(^15\). In document analysis, CNNs have been used for word spotting by Jaderberg et al. \(^16\), and for handwritten character recognition by Bluche et al. \(^17\). In the \(^16\), they divide text spotting tasks into two sequential
tasks: detecting words region in the image, and recognizing the words with these regions. They use CNN for both tasks and character classification is 91% accuracy. In spite of these performance, CNN have not shown for offline writer identification[18]. J. Gall et al. [18] proposed the activation features from CNNs as local descriptors for writer identification. They evaluate on ICDAR13 dataset which contains English character written by 350 writers. They use the TOP-k scores to evaluate their experiment. TOP-k score are determined by calculating the percentage of words, where the k highest ranked documents were from the same writer. The score largely differ their TOP-1 from TOP-3. So, we evaluate a variation of classification accuracy about each handwritten character and about each writer for same word.

3. Convolutional neural network architecture

In our work, We employ AlexNet CNN[19] implemented in open source Caffe library to extract activation feature, and in this section, we describe about AlexNet. This architecture achieved significantly improved performance over the other non-deep learning methods for ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2012. This success has revived the interest in CNNs in computer vision. In fig.1, we show the architecture of AlexNet and this contains eight trained layers which are five convolutional layers and three fully-connected layers. For reducing overfitting, data augmentation and dropout are implemented in AlexNet. Data augmentation is the easiest method to reduce overfitting on train image data. This method is to artificially enlarge the dataset using label-preserving transformation. Dropout consists of setting to zero the output of each hidden neuron with probability 0.5. The neurons which are dropped out do not contribute to the forward pass and backward pass.

4. Experiments

4.1 Dataset

We use a dataset provided by Japanese Association of Forensic Science and Technology [20] for evaluation which contains 100 kinds of Japanese handwritten characters (96 of Chinese characters, 1 of Hiragana, 3 of Katakana) written 50 times by 100 writers. Fig.2 shows Hiragana, Katakana and Chinese character. Hiragana and Katakana are indigenous to Japan. We use 90 kind of words training and 10 kind of word testing which are grayscale images and 160 x 160 pixels.

4.2 Performance of AlexNet for writer identification

We now investigate the performance of AlexNet for writer identification. Table 1 shows how we split the dataset, 90 kind of word is train and 10 kind of word is test. Our purpose is to identify the owner of handwritten character, so we test on a kind of character which is not trained.

**Table 1 How we split the dataset in train and test**

```
<table>
<thead>
<tr>
<th>word</th>
<th>test</th>
<th>train</th>
</tr>
</thead>
<tbody>
<tr>
<td>町中木平地和場岩里塚田原寺入石瀬池郷稀上小沢三津口台浦部富大通内宮日路御古玉目西新ノル村岡之河倉道下谷島長八水一根出森山屋条前浜二橋保丘門東北井崎市坂見の堀白川本高神ケ吉生福子向</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

Fig.3 shows transition about each loss and identification accuracy of train and test. Each loss is presented by left vertical axis and each identification accuracy is presented by right vertical axis. AlexNet achieves Test identification accuracy of 85.24% on 100 writer and train identification accuracy is 91.40. This result shows that AlexNet do not overfit the train data due to data augmentation and dropout.

```
Fig.3 Transition of train and test
```

95
4.3 Identification accuracy for each word

Table 2 shows the result of identification accuracy for each kind of word. Complex Chinese character is better identification accuracy than simple Chinese character, Hiragana and Katakana. The highest identification accuracy of 91.90% is “富” and the lowest identification accuracy of 5.45% is “の”. This shows that activation features from Japanese handwritten words effect on identification writer hugely. It find out that writer unique style appears in complex character but in simple character does not.

Table 2 Identification accuracy of each word

<table>
<thead>
<tr>
<th>accuracy</th>
<th>Test word</th>
</tr>
</thead>
<tbody>
<tr>
<td>90% ~</td>
<td>富 石 池 木 海</td>
</tr>
<tr>
<td>80% ~</td>
<td>橋 木 村 原 木</td>
</tr>
<tr>
<td>70% ~</td>
<td>場 野和 神 向 道</td>
</tr>
<tr>
<td>60% ~</td>
<td>丘 城 崎 町 木</td>
</tr>
<tr>
<td>50% ~</td>
<td>上 水 丸 川 下</td>
</tr>
<tr>
<td>40% ~</td>
<td>戸 子 丸 小 外 下</td>
</tr>
<tr>
<td>30% ~</td>
<td>入 木 之</td>
</tr>
<tr>
<td>20% ~</td>
<td>ル</td>
</tr>
<tr>
<td>10% ~</td>
<td>ノ</td>
</tr>
<tr>
<td>0% ~</td>
<td>の</td>
</tr>
</tbody>
</table>

4.4 Identification accuracy of each writer

In this section, we evaluate how handwritten words have writer own identities which are able to classify the owner. Fig. 4 show identification accuracy for each writer in histogram. The best identification accuracy is 95.0%, but the worst identification accuracy is 31.5%. This shows that handwritten characters present writer’s unique identities. Fig. 5 and Fig. 6 show the handwritten words which are the best performance writer’s and worst performance writer’s. Comparison of these figures show that the highest identification accuracy writer is more unique writing style for Japanese. Moreover, both of “の”may not identify the owner for our eyes but “富” have much unique feature to identify the owner. So handwritten words contain writer own writing style which allow us to identify the owner of handwritten words so much.

5. Conclusions

In this paper, we proposed method of writer identification for activation feature learned from CNN. We evaluate difference about same documents written by different writer and about different documents written by same writer on the identification accuracy. Each writer have unique writing style and most unique writer can be identified 96.0% accuracy but worst accuracy writer can be identified only 38.5% accuracy. And even if the same writer draw words, each Hiragana and Katakana can be identified less than 30% because of its simple structure. For future work, we would like to research larger and more complex CNN architecture to identify a simple structure character such as Hiragana and Katakana. There is still room to improve identification accuracy to get more fine features.
Acknowledgment

We thank Japanese Association of Forensic Science and Technology for permission to use the dataset of Japanese handwritten words.

References


vation feature for generic visual recognition.” In ICML, 2014
(13) Filip Korzeniowski , Gerhard Widmer “A FULLY CONVOLUTIONAL DEEP AUDITORY MODEL FOR MUSICAL CHORD RECOGNITION” IEEE 2016
(19) Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton “Imagenet classification with Deep Convolutional Neural Networks” In NIPS 2012
(20) Kiichi Misaki, Daisuke Honjo, Umeda Michio “Creation of handwritten character databases and display analysis software for author recognition research (written in Japanese)” Japanese Association of Forensic Science and Technology, 2002