다양한 컨볼루션 신경망을 이용한 태국어 숫자 인식

Athita Onuean[1,*] · 정한민[1] · 김태홍[2]

[1]한국과학기술정보연구원, 과학기술연합대학원대학교 · [2]한국한의학연구원

Handwriting Thai Digit Recognition Using Convolution Neural Networks

Athita Onuean[1,*] · Hanmin Jung[1] · Taehong Kim[2]

[1]KISTI, UST · [2]KIOM

E-mail : athita@kisti.re.kr / jhm@kisti.re.kr / thkim@kiom.re.kr

요 약

필기체 인식 연구는 주로 딥러닝 기술에 초점이 맞추어져 있으며, 최근 몇 년 동안 많은 발전을 이루었다. 특히, 필기체 태국어 숫자 인식은 태국 공식 문서와 영수증과 같은 숫자 정보를 포함한 많은 분야에서 중요한 연구 분야지만, 동시에 도전적인 분야이기도 하다. 대규모 태국어 숫자 데이터 집합의 부재를 해결하기 위해, 본 연구는 자체적인 데이터 집합을 구축하고 이를 다양한 컨볼루션 신경망으로 학습시켰다. 정확도 메트릭을 이용하여 평가한 결과, 배치 정규화 기반 VGG 13이 98.29%의 가장 높은 성능을 보였다.

ABSTRACT

Handwriting recognition research is mainly focused on deep learning techniques and has achieved a great performance in the last few years. Especially, handwritten Thai digit recognition has been an important research area including generic digital numerical information, such as Thai official government documents and receipts. However, it becomes also a challenging task for a long time. For resolving the unavailability of a large Thai digit dataset, this paper constructs our dataset and learns them with some variants of the CNN model; Decision tree, K-nearest neighbors, Alexnet, LaNet-5, and VGG (11,13,16,19). The experimental results using the accuracy metric show the maximum accuracy of 98.29% when using VGG 13 with batch normalization.

키워드

Handwriting Thai Digit Recognition, Convolution Neural Network, Thai Digits Dataset

I. Introduction

The handwriting recognition challenges are the several handwriting styles of a person, low quality of the data due to the degradation of document/image or low standard capturing device over time, and cursive handwriting [1][2]. The lack of handwriting Thai digit dataset is incredibly challenging and problematic for a researcher on a deep convolution neural network.

* speaker

Deep learning architecture such as convolution neural network (CNN) has multi-layer feed-forward neural networks that automatically extract information from training data. In many applications such as computer vision, character, and handwriting recognition [3-5], CNN can learn complex information from a large number of training image data [6]. The advantage of CNN is that it automatically extracts features and classifies them to be unnecessary for handcraft feature extraction and selection. The purposes of this paper are to propose a new dataset of Thai handwriting digits and explore our initial ideas of using CNN to create a deep learning recognition system.

II. Data Description

The 1st inscription of King Ramkamhaeng is historical evidence showing that Thai scripts were invented and used from 1826. From time to time, the scripts have been changed in both form and orthography. The Thai language has unique digit characters that are different from Arabic digit characters. These digit characters are often used in official writings.

In this study, we constructed our new Thai digit dataset. We created a standard data collection sheet and collected handwriting from 1,072 university students. Each student wrote the digits in specified squares block on the sheet, and they wrote digits twice. Each data collection sheet was capture by using a true-color level document scanner with 300 DPI. We prepared our dataset by only segmentation each image. So our dataset is original, JPG file format, true-color, without de-noising or cleansing processing. After the segmentation process, the Thai digits dataset has 8,763 images assigned to a training set and another 2,000 images as a testing set. Figure 1 shows a sample of training the Thai digits dataset that is written in a sequence from 0 to 9.

Figure 1. The examples of handwriting styles on the proposed dataset

Table 1. Comparison of Performance Results

Models	Test accuracy
Decision Tree	40.27%
KNN	66.79%
AlexNet	96.02%
LeNet-5	87.55%
VGG_11_BN	97.83%
VGG_13_BN	**98.29%**
VGG_16_BN	97.99%
VGG_19_BN	98.16%

III. Experimental Results

In our experiment, we train our Thai handwriting digits data set with machine learning, and various CNN models consist of the Decision tree, K-nearest neighbors, Alexnet, LaNet-5, VGG (11,13,16,19) with batch normalization. All classifiers are repeated five times by shuffling the training set and average accuracy on the test set. The hyper-parameter used to train all the models are as follows: 1) batch size 32, 2) dropout 0.5, 3) epoch 100, and 4) optimizer Adam. The results of testing accuracy of eight models show that VGG-13 with BN show outperform others. Thus, VGG_13 is the best fit for handwriting Thai digit recognition.

We plot a confusion matrix as shown in Figure 2. The dark diagonal cells present the correct classification according to the true label and predicted label. We found a slight confusion of two groups: 1) digit four and five, 2) digit four and eight. Example of confused images shown in Figure 3 divided into two groups of confusion; (a) depicts true label of digit four and predicted of digit five, while (b) show predicted of digit eight. The red label indicates false prediction, and the green one is an accurate prediction. The label located in front of a bracket is the predicted label, and the label inside is a true label.

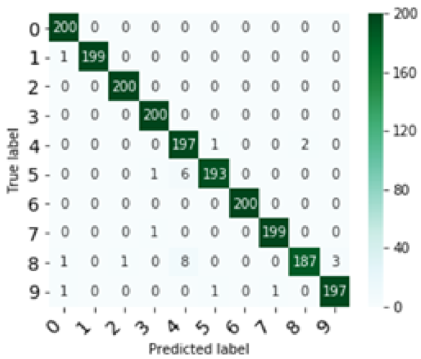
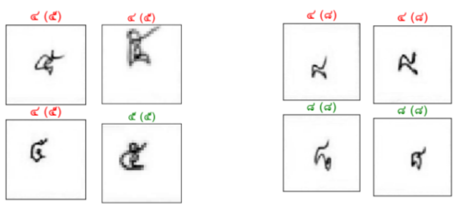


Figure 2. Confusion matrix



(a) Predicted: 4; True: 5 (b) Predicted: 4; True: 8

Figure 3. Example of confused digits

IV. Conclusion

This study constructs a handwriting Thai digit dataset written by 1,072 participants. Our proposed dataset available at Burapha University**. The variations of the Machine learning and CNN model's accuracy were observed. Among all the observations, the result shows maximum accuracy of 98.29% by using VGG 13 with batch normalization that best fits our dataset. As a result, we found two groups of confusion between similar labels. Therefore, in further studies, we will be focusing on improving the performance of handwriting Thai digit recognition and minimize the confusion between similar digits.

References

[1] N. Arica, and F. T. Yarman-Vural, "An overview of character recognition focused on off-line handwriting," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, Vol. 31, No. 2, pp. 216-233, May 2001.

[2] M. Maheshwari, D. Namdev, and S. Maheshwari, "A systematic review of automation in handwritten character recognition," *International Journal of Applied Engineering Research*, Vol. 13, No. 10, pp. 8090-8099, November 2018.

[3] A. El-Sawy, E. B. Hazem, and M. Loey, "CNN for handwritten arabic digits recognition based on LeNet-5," in *Proceeding of International conference on advanced intelligent systems and informatics,* Springer, Cham, Cairo: Egypt, pp. 566-575. October 2016.

[4] A. Baldominos, Y. Saez, and P. Isasi, "A survey of handwritten character recognition with MNIST and EMNIST," *Applied Sciences*, Vol. 9, No. 15, pp. 3169, August 2019.

[5] S. Pratt, A. Ochoa, M. Yadav, A. Sheta, and M. Eldefrawy, "Handwritten digits recognition using convolution neural networks," *Journal of Computing Sciences in Colleges*, Vol. 34, No. 5, pp.40-46, 2019.

[6] H. Khosravi, and E. Kabir, "Introducing a very large dataset of handwritten Farsi digits and a study on their varieties," *Pattern recognition letters*, Vol. 28, No. 10, pp. 1133-1141, 2007.

** http://services.informatics.buu.ac.th/datasets