Proceedings of the 3rd International Conference on Signal Processing and Machine Learning
DOI: 10.54254/2755-2721/4/20230450

Handwritten English character recognition using convolutional neural network

Liu Hanqi

Beijing National Day school, Yuquan road haidian district Beijing 100039 China

liuhanqimark@outlook.com

Abstract. Converting paper material into electronic material is still a necessary work nowadays, however, recognition of handwritten characters still has limitations in their recognition rate, owing to the presence of various shapes, scales, and formats in different people's handwritten characters. Machine learning has significant value in reducing human power. A Convolutional Neural Network model that is revised from LeNet-5, is used for handwritten letter recognition. This study uses the EMINST dataset to train the model, and the final recognition rate is about 93.44%.

Keywords: convolutional neural networks, handwritten character recognition, LeNet-5, EMINST

1. Introduction

With the development of technology and society, paper materials gradually get out of public sight. More and more paper materials need to be turned into the electronic version, to bring convenience for people and to decrease the cost of preserving paper materials. However, taping characters into a system can take lots of time and human power. Handwritten character recognition provides a way to reduce the human cost of transformation. By extracting images, the computer can convert Handwritten characters into data stored in the system. Convolutional Neural Network (CNN) has shown its superiority in the field of Handwritten recognition. Therefore, this study focuses on Handwritten English character recognition based on CNN. Incited to finding the underlying ability of CNN, the present study attempts to investigate the effect of LeNet-5.

2. Literature review

Since handwritten characters have an infinite number of variations, handwriting character recognition has been a heated topic, so there have been multiple studies and efforts on different algorithms. For example, the group of Zhao Bin has investigated the SVM (support vector machine) for numeral recognition based on an experiment on the NIST numeral database and shows that it possesses good generalization ability [1]. SVM has the most proper hyperplane which can be attained under certain conditions, thus it has great generalization ability. However, there are still some disadvantages in resolving the multi-class problem, training speed, and testing speed. In the research of Nurul Ilmi's group [2], they used KNN (K Nearest Neighbors) for the classification algorithm and Local Binary Pattern Variance (LBPV) for the feature extraction method to realize handwriting recognition. KNN

© 2023 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/4/20230450

algorithm is suitable for big volume data, and LBPV is good at recognizing texture images, but the accuracy is only about 89.81% on the digital Mnist dataset. JianXie researched the RF (random forest) algorithm based on a decision tree is applied to handwritten numeral recognition [3], to provide solutions to the field of handwritten numeral recognition. After optimizing and training each decision tree and setting multiple decision trees, the recognition accuracy of the algorithm has been greatly improved. The Mnist data set have been used for his experiment, and the results show that the random forest algorithm can reach a high recognition accuracy of about 97%. Pradeepta K Sarangi and P Ahmed have studied the suitability of the Artificial Neural Network (ANN) model with back-propagation learning which is based on recognizing the handwritten numerals of Odia language (an Indian regional language) and using LU decomposition of the matrix for feature extraction [4]. However, its relatively low percentage of accuracy 85.30% could be due to the local minima problem of backpropagation. Suthasinee Iamsa-at and Punyaphol Horata investigate and compare the recognition ability of Deep-Learning Feedforward-Backpropagation Neural Network (DFBNN) and Extreme Learning Machine (ELM) [5], and the result shows that recognition rates of DFBNN are slightly higher than ELM. This paper presents a method for applying HOG for feature extraction that can enhance the recognition rates for both classifiers. Li Wenzheng and Cui Wanxin compare and contrast three different ANN models: Perceptron Neural Network [6], Back Propagation Neural Network, and Hopfield neural network. The results show that Perceptron Neural Network has simple computation and short training time, but the anti-noise performance of the trained network is poor. Compared to that, Back Propagation Neural Network needs a longer training time, but Back Propagation Neural Network performs better than the Perceptron Neural Network whether it is trained with noise or without noise. Hopfield network can be generated directly according to the set equilibrium without training. However, the classification problems that the Hopfield network can handle have some limitations, and it may generate spurious equilibrium and perform wrong convergence, thus affecting the classification results.

With further development, researchers gradually found the priority of CNN. There are different architectures of CNN on account of characteristics and properties, LeNet-5 designed by Yan LeCun in 1998 has been a popular model in handwritten character recognition [7], but the results of LeNet-5 for complex problems are not ideal. Gao Can investigated improving LeNet-5 by adjusting activate function and sampling process [8]. The recognition efficiency and convergence speed of handwritten character recognition are effectively improved to a certain extent. Compared to that, Xiang Benhua and Wang hui have researched another architecture AlexNet [9], and they use it to recognize Greek letters, the accuracy of case letters is about 98.27% and 96.07%. In their study, their dataset is collected from 60 different students, so more diverse data can be collected to train better models in the future. Moreover, preprocessing method of datasets can be used to improve the recognition effect. In the paper of Ghazanfar Latif's group [10], they propose a deep learning architecture for the recognition of handwritten Multilanguage. It shows that the recognition rate is increasing with the number of hidden layers, however, more hidden layers need more training time. The average precision is about 99.29 and it helps people to explore deep learning for universal handwritten character recognition. Zhu Shiwen[11] used CRNN (Convolutional Recurrent Neural Network) as a basic structural framework which is composed of RNN and CNN with EnCTC (Connectionist Temporal Classification with Maximum Entropy Regularization) as a loss function. But before the CNN network, the STN (spatial transformation network layer) is inserted, followed by CNN, followed by a BiLSTM(stacked bidirectional long and short memory network), and finally the transcription layer.

In order to increase the accuracy of character recognition, many studies have different possible solutions. Fu Pengbin's group has focused on the problem of confusable characters in handwritten letter recognition, instead of changing the architectures of CNN [12], they improve the preprocessing process for multi-dimensional feature extraction. In addition, they remain different candidate words for confusable characters, considering the corpus and the relationship in words, the confusable characters are finally recognized. Compared to that, Aanchal Jain and Brij Kishore Sharma analyzed Activation Functions for Convolutional Neural Network, and the classification rate of sigmoid activation function is about 99.65%, which is higher than Tanh and Linear [13].

DOI: 10.54254/2755-2721/4/20230450

3. Method

3.1. dataset

In this study, the EMNIST dataset of handwritten letters. It consists of 88800 and 14800 samples for training and testing. The train and test sets are totally different, each letter has images of a resolution of 28 × 28 pixels. More detailed information can be found in the following [14].

3.2. model

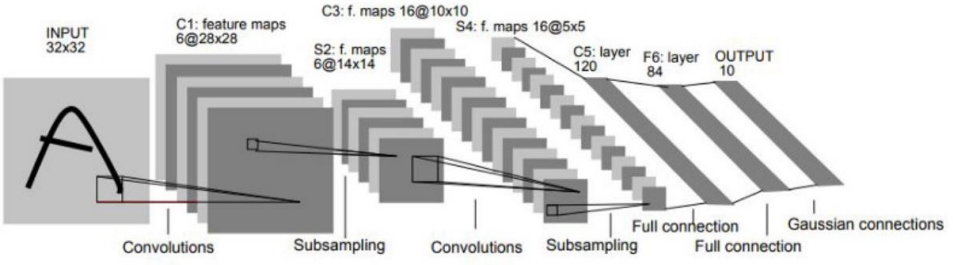


Figure 1. the steucture of LeNet-5.

The convolutional neural network model used in the experiment is improved on the basis of the LeNet-5 model. LeNet-5 is a classic CNN model which was proposed by Yann LeCun, Leon Bottou, Yosuha Bengio, and Patrick Hafner in 1998 [7]. It consists of two sets of convolutional, each followed by average pooling layers. And then, the fifth layer is a flattening convolutional layer, followed by two fully-connected layers, and finally a softmax classifier. Because this model is used for numeral recognition, outputs are 10 different values. Some more details about LeNet-5 are mentioned in [7]. Specific improvement content and analysis such as:

- Select the ReLU activation function. The ReLU activation function makes the output of some neurons equal to 0, which makes the convolutional neural network possess sparse properties to a certain extent, reduces the number of parameters, reduces the network computation, and makes the training of the network model easier.
- In the pooling layer, mean-pooling is changed to max-pooling for the pooling layer (or call subsampling layer). Mean pooling can better retain the background information of the image, while maximum pooling can better extract the texture information of the image.
- The convolution kernel size and its number, the number of neurons, and other parameters were adjusted. Too large or too many convolution kernels in the convolutional layer will affect the training speed of the network model. Otherwise, no effective feature information can be extracted.

4. Result

Each sample from the training set was fed to a CNN for training. The proposed CNN architecture was tested on the mentioned above datasets. Adjust quantities of epoch and parameters of the model, in order to improve the efficiency. When batch size is defined as 64, the accuracy and loss rate are shown in the figure. A categorical cross-entropy loss function is used in the model.

$$L = \frac{1}{N}\sum_i L_i = -\frac{1}{N}\sum_i L_i \sum_{c=1}^{M} y_{ic}\, log(p_{ic}) \quad (1)$$

DOI: 10.54254/2755-2721/4/20230450

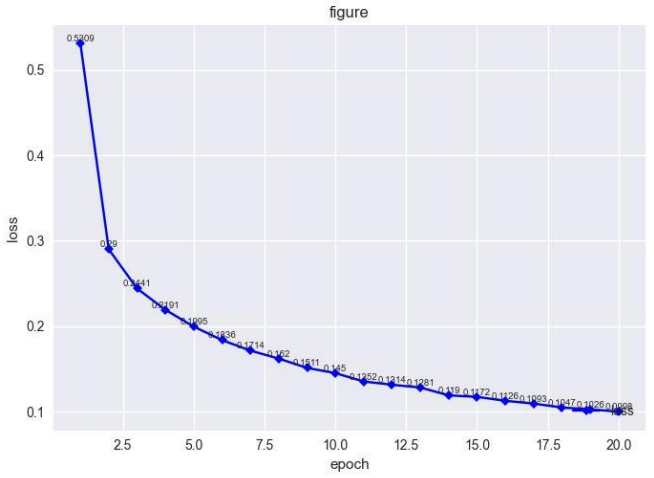


Figure 2.the loss by different epochs.

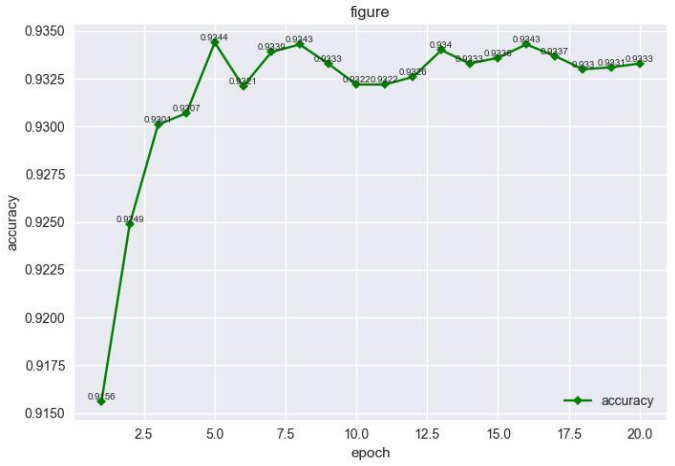


Figure 3. accuracy of model by different epochs.

The final accuracy is about 93.44%. The results indicate that better recognition rates were achieved with 5 epochs. However, with more epochs, the more complex the network and more computational time are needed to produce results. Excessive training can also lead to overfitting. For this study, the recognition rate of 93.44% is adequate and extra epochs will increase the time to produce results will not be justified.

5. Discussion

The model used in this article is revised from the model of numeral recognition. Although the network structure of the two is not very different, the accuracy is lower than its performance in numeral recognition for some reasons. Handwritten letters have totally different 26 classes, not to mention uppercase and lowercase. Because more types needed to be classified, the effect of the network structure will be decreased.

Moreover, with more letters characters are more likely to have similar structures, some confusable characters like “i” and “r” have much lower accuracy than other letters.

To these problems, there are several possible solutions. First, it is adopting an advanced structure of the model, finding a better network structure to increase the recognition accuracy, although complex

DOI: 10.54254/2755-2721/4/20230450

network structures will require higher computational power. The computational cost and accuracy are worth discussing, but the relatively optimal solution to the corresponding problem can be found through further research. Second, data preprocessing cannot be ignored. For example, in [9] and [12], feature extraction is applied before training, which is used to extract more image characteristics of letters ranging from outline to shape, and the results show better performances.

6. Conclusion

In conclusion, more and more paper material needs to be turned into electronic material, so handwritten character recognition is an effective way to resolve this problem. In this field, CNN is a heated topic, so this paper intends to research a revised model based on LeNet-5. By the optimization and adjustment of parameters, a CNN model is trained by the EMNIST dataset, and this model provides an accuracy of 93.44%. The research provides some insights into handwritten character recognition and provides some possible solutions for the preservation of paper materials. By studying on LeNet-5 structure, this paper proves the superiority of CNN.

However, there are some limitations to this study, which can provide suggestions and recommendations for future studies. LeNet-5 is still limited in letter recognition and cannot achieve extremely high accuracy. Some more complex CNN architectures can be adopted. At the same time, the data set has not been preprocessed more, and operations such as feature extraction may be used to improve the accuracy.

References

[1] Zhao Bin, Liu Yong, & Xia Shao-Wei. (n.d.). Support vector machine and its application in handwritten numeral recognition. Proceedings 15th International Conference on Pattern Recognition. ICPR-2000.

[2] Zhang, Chen, Chang, & Zhang. (2021). Research on handwritten digit Recognition technology based on KNN algorithm. Computer Programming Skills & Maintenance.

[3] Xie. (2021). Handwritten numeral recognition based on random forest algorithm. China Hi-Technology.

[4] K Sarangi, & Ahmed. (2012). Recognition of Handwritten Odia Numerals Using Artificial Intelligence Techniques. The International Journal of Computer Science & Applications, 2(2), 41–48.

[5] Iamsa-at, S., & Horata, P. (2013, December). Handwritten Character Recognition Using Histograms of Oriented Gradient Features in Deep Learning of Artificial Neural Network. 2013 International Conference on IT Convergence and Security (ICITCS).

[6] Li, & Cui. (2021). Comparative study on the performance of different artificial neural networks in letter recognition. Internet of Things Technologies.

[7] Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278–2324.

[8] Gao. (2017). An improved method of handwritten character recognition based on CNN. Innovation in Science and Technology.

[9] Xiang, & Wang. (2021). Research on Handwritten Greek letter Recognition based on AlexNet. Communication Technology.

[10] Latif, G., Alghazo, J., Alzubaidi, L., Naseer, M. M., & Alghazo, Y. (2018, March). Deep Convolutional Neural Network for Recognition of Unified Multi-Language Handwritten Numerals. 2018 IEEE 2nd International Workshop on Arabic and Derived Script Analysis and Recognition (ASAR).

[11] Zhu. (2022, May). Research on English Handwriting Recognition Based on CRNN and EnCTC. Computer & Digital Engineering, 50(5), 1093–1097.

[12] Fu, Song, & Yang. (2021, November). Confusable handwritten English recognition based on multi-dimension al features and candidates. Computer Engineering and Design, 42(11).

DOI: 10.54254/2755-2721/4/20230450

[13] Aiquan Yuan, Gang Bai, Lijing Jiao, & Yajie Liu. (2012, March). Offline handwritten English character recognition based on convolutional neural network. 2012 10th IAPR International Workshop on Document Analysis Systems.

[14] Cohen, G., Afshar, S., Tapson, J., & van Schaik, A. (2017, May). EMNIST: Extending MNIST to handwritten letters. 2017 International Joint Conference on Neural Networks (IJCNN).