



Study on Indicator Recognition Method of Water Meter Based on Convolution Neural Network

Shuaicheng Pan, Lei Han^(✉), Yi Tao, and Qingyu Liu

Nanjing Institute of Technology, Nanjing 211167, China
psc351@foxmail.com, hanl@njit.edu.cn

Abstract. In order to improve the accuracy of water meter character recognition, this paper proposes a character recognition method based on deep convolutional neural network. Traditional identification methods need to build a large number of templates, which requires a lot of work, and are easy to be interfered by external light and sundries, so the identification accuracy is low. The object of the experiment is the water meter dial with the character of word wheel and the corresponding data set is established. A character recognition method based on deep convolutional neural network is proposed to solve the problem of half-character on water meter. First to pretreatment of data set, the main data set is rotating images and augmentation, and then according to the classical convolution neural network structure, construct a can identify characters at the same time and dial the convolutional neural network model, training on the data set tests, the experimental results show that the method effectively improves the water meter word wheel character recognition accuracy.

Keywords: Convolutional neural network · Intelligent table recognition · Image recognition · Character recognition · Halfword recognition

1 Introduction

Intelligent meter investment and operation and maintenance cost is low, fast, high efficiency, wide application prospects. The camera reading table mainly includes character positioning, character segmentation and character recognition. The method based on area aggregation runs faster, but is easily affected by low image resolution and noise [1]. The target recognition algorithm matches feature vectors through improved feature matching algorithm, which improves the accuracy but decreases the recognition speed. Template matching based on euler number groups the template images and matches the target image with the template with the same euler number. At the same time, the above methods are easily affected by environmental factors and are not robust.

To improve the performances of classification, several fusion methods have been developed by combining different feature extraction and classification methods. CNN-ELM classifier has been proposed by Zeng, Y. et al. [6] for traffic sign recognition and achieved human level accuracy. Guo, L. et al. [7] have proposed Hybrid CNN-ELM with improved accuracy and validate its performance on MNIST

dataset. Gurpinar, F. et al. [8] have replaced ELM to Kernel ELM for classification that minimizes the mean absolute error. Face features have been extracted from a pre-trained deep convolutional network. Kernel extreme learning machines are used for classification. Yoo, Y. et al. [9] proposed a novel fast learning architecture of CNN-ELM. Their core architecture is based on a local image (local receptive field) version of the ELM adopting random feature learning. Weng, Q. et al. [10] combines ELM classifier with the CNN-learned features instead of the fully connected layers of CNN to land-use classification.

The method proposed in this paper combines the characteristics of the classical network structure to build a new type of convolution neural network M_CNN (Modified_CNN) suitable for water meter image recognition. The experiment shows that it can not only recognize the word wheel figure of water meter but also recognize the digital image. And not easily affected by environmental factors. The rest of the paper is organized as follows: Sect. 2 describes the classical neural network used in this paper, Sect. 3 describes the network structure, Sect. 4 is the experiment, and finally the conclusion.

2 Related Work

The classical neural network based on CNN includes LeNet5, AlexNet, GoogLeNet, etc. As an early CNN, LeNet5 adopted ReLU activation function, added Dropout layer and expanded the amount of training data. However, AlexNet has only eight layers. The first five layers are convolution layer, and the last three layers are full connection layer. The network depth is not deep enough, and the discarding pool layer is prone to over-fitting.

GoogLeNet injected the modularization idea into CNN and used the convolution of 1×1 to carry out the elevating dimension. Its sparse connection structure improved the adaptability of the network to a variety of complex images, but it took too long to apply it to the calculation of character recognition.

The first four layers of M_CNN proposed in this paper can be regarded as a module. In the future, the structure of the module remains unchanged and only some parameters such as convolution kernel are changed. M_CNN is a convolutional neural network specially constructed for water meter characters, which USES the characteristics and ideas of classic CNN for reference, and more suitable for the use of camera meter reading scene.

M_CNN is improved on the basis of CNN_ON_MNIST [1, 6] and CNN_ON_CIFAR [3] to build a model that can recognize both the number and the small wheel. CNN_ON_CIFAR can effectively identify images in the cifar-105 dataset, which is built for 10 common objects in nature. This network can be used to identify the word wheel image of water meter. CNN_ON_MNIST can effectively recognize handwritten digits, this network can effectively recognize water meter character image.

By combining the two models, a new neural network M_CNN is constructed, which can recognize the word wheel image and character image of water meter at the same time. By using the self-built water meter training set and test set, the training, testing and generalization ability is better.

3 Proposed Model

As shown in Fig. 1, M_CNN draw lessons from the modular design idea, the two convolution layer, one of the biggest pooling layer, a dropout as a module, the network use a total of three such module, the first module number convolution kernels is 32, the second module has 64 convolution kernels, the third module has 128 convolution kernels, the convolution kernel module size is 3×3 , pooling nucleus is 2×2 , dropout ratio is 0.25, finally after a flattening all connections, regularization operation again suppress a layer of neurons activate prevent fitting, Finally complete connection layer can output a 128 - dimensional vector, softmax layer will this vector as input, and then calculate the test images respectively belong to the probability distribution of the 10 class, softmax using cross entropy loss function (cross-entropy) to calculate each forecast loss value size, commonly used loss value calculation function and hinge loss [4], but the hinge loss calculation value is without calibration, it is difficult to decide on all the classes. However, the value of cross-entropy can be used to determine all classes, which will give the prediction probability of each class and finally output the prediction result.

4 Experiment Results

4.1 Experimental Setup

This experimental platform is Intel i7-8700 processor, 16G of memory, GTX1080 graphics card, 4G video memory dell workstation, using Tensorflow as the back-end Keras platform to build M_CNN, and training and verification on the data set.

Character acquisition is completed through raspberry PI and macro camera, and character samples are shown in Fig. 2. USB HD macro camera is adopted with 8 million pixels and the maximum resolution is 640×480 . The word wheel sample is shown in Fig. 3. The image segmentation is processed by OpenCV, and five spinner code plates and four decimal dial are cut out.

The collected digital and pointer samples are divided into 10 categories from 0 to 9. The training set and test set are divided according to 4:1, and are independent of each other without intersection.

For the half-word problem of numeric characters, as shown in Fig. 3, the number shown in the upper part is 0, that is, the number that has been displayed completely recently is read according to the scroll direction of the word wheel, and the specific reading is read out in combination with the reading of the dial pointer. For the roulette half-word problem, according to the empirical table reading method, it is classified into the number of the last time, and then read more accurate Numbers according to the following X0.1, X0.01 and X0.001 dial. This is done for each dial. The last dial is almost always a pointer to an integer because it is the smallest unit indicator.

M_CNN model parameters include activation function, pooling method, pooled core size, discard ratio, classifier, loss function, optimizer, etc. See Table 2 for specific parameters.

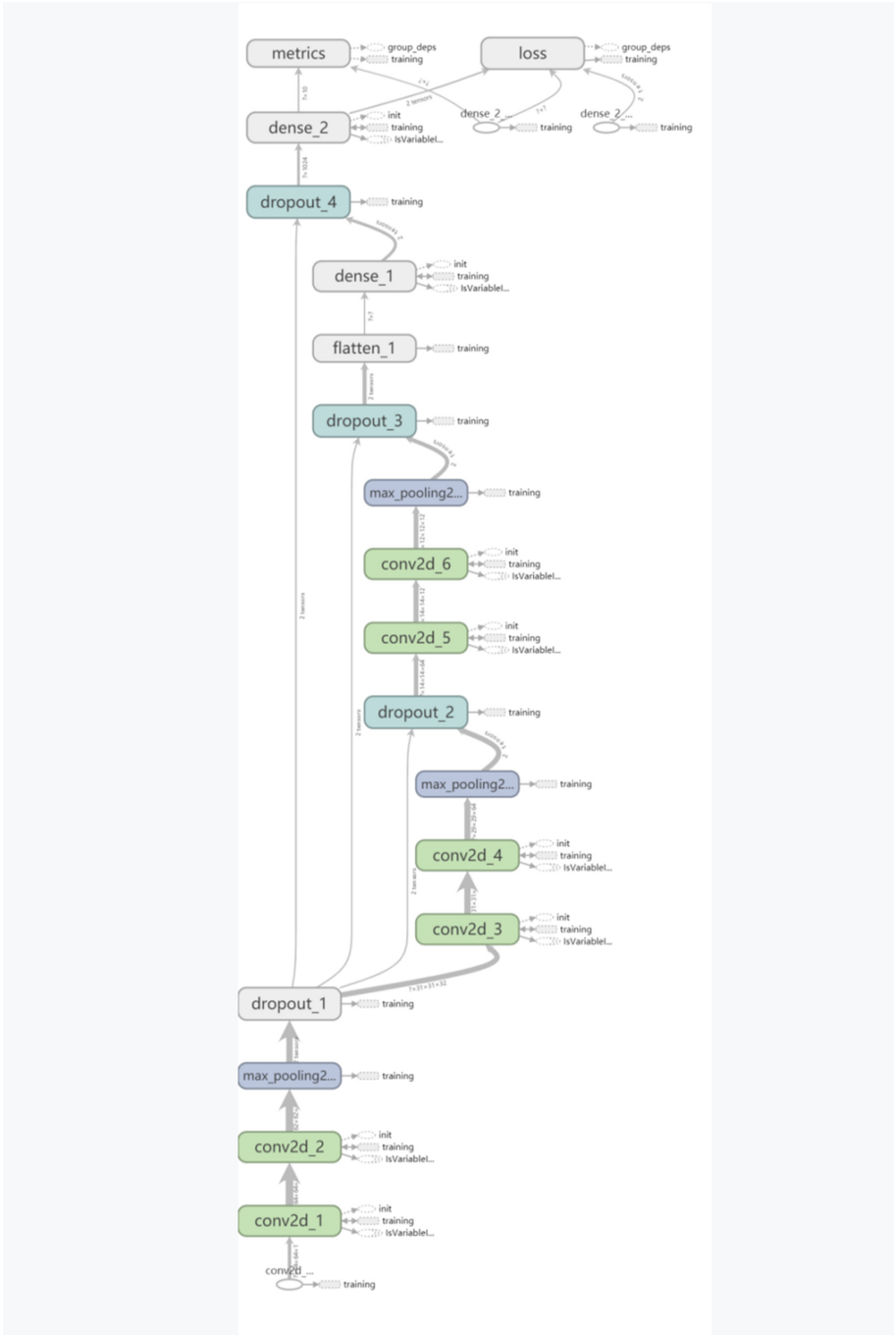


Fig. 1. Architecture of M_CNN.

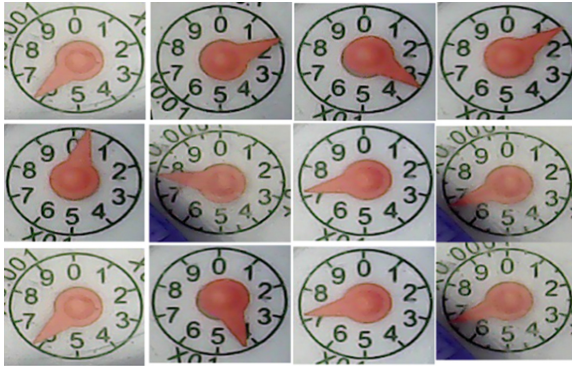


Fig. 2. Sample images.

Table 1. Data set specification.

The total number of sheets	9 200
Number of training sets	8 280
Number of test sets	920
Figure (height × width) (pixels)	136 × 136

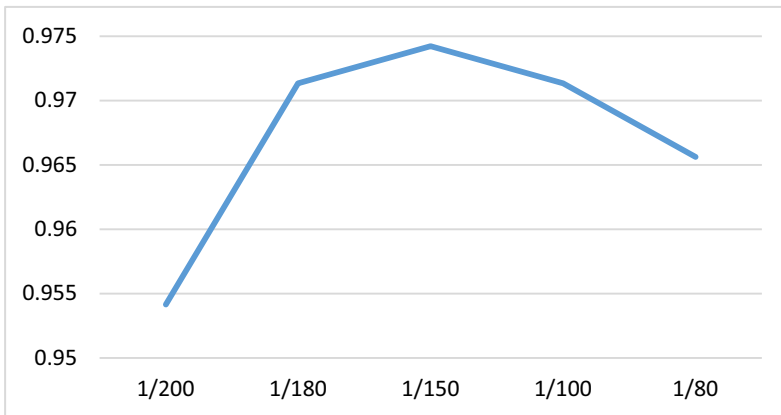


Fig. 3. Correlation between accuracy and rescaling factor.

4.2 Results and Analysis

Based on the self-built data set and training set, the traditional neural network and M_CNN were used for training test respectively, and the results were shown in Table 3 and Table 4. The loss changes and accuracy changes of CNN_ON_MNIST training are shown in Fig. 4. The loss value increases in the test set, indicating that the network

Table 2. Parameters specification.

Training parameters	Value
activation function	ReLu
Pooling method	Max pool
Pool Size	(2, 2)
Dropout	0.25
classifier	Softmax
Loss function	Categorical_crossentropy
optimizer	Adadelata
Rescaling factor	1/255

basically has no fitting and cannot be effectively identified. The change when CNN_ON_CIFAR is used is shown in Fig. 5. The loss fluctuation on the test set is large, indicating that the trained network cannot be well fitted. Overall, M_CNN showed better performance in terms of loss and accuracy.

Table 3. Comparison of accuracy between traditional neural network and M_CNN.

Network name	Recognition accuracy/%
CNN_ON_MNIST	85.67
CNN_ON_CIFAR	91.97
M_CNN	97.13
ResNet v1	94.84
ResNet v2	63.61

Table 4. Comparison of loss between traditional neural network and M_CNN.

Network name	Loss value
CNN_ON_MNIST	0.62
CNN_ON_CIFAR	0.21
M_CNN	0.19
ResNet v1	0.27
ResNet v2	1.76

Accuracy and loss values are compared as follows, respectively showing the accuracy and loss values of CNN_ON_MNIST, CNN_ON_CIFAR, ResNet and M_CNN. By comparison, M_CNN has better performance in accuracy and loss values. Although ResNet is a classic classification network, it does not perform well when both the word wheel and the character image are needed, and the training time is much longer than M_CNN.

4.3 Experimental Comparison

In order to solve the imbalance between sample data, the method of data expansion was proposed. It was found from the actual data collection that not every sample had the same number, which would interfere with the learning process. In order to balance the number of samples and avoid too few samples in the training process and poor fitting, re-scaling factor was added in the sample pretreatment. The experiment shows that when the rescaling factor is set as 1/150, the accuracy rate is the highest (See Fig. 4).

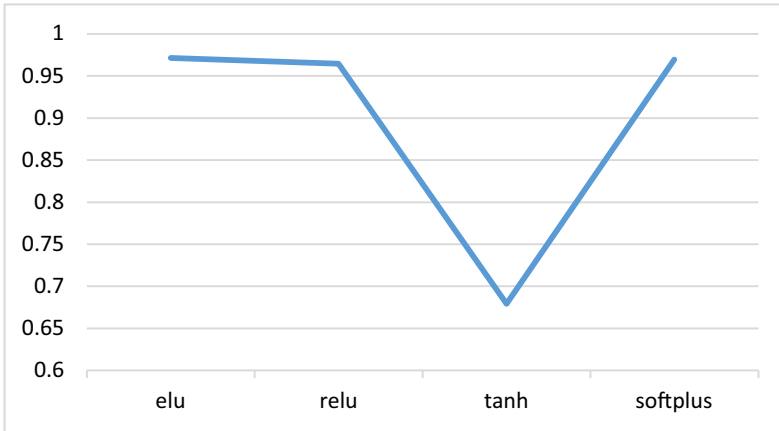


Fig. 4. Correlation between accuracy and activation function.

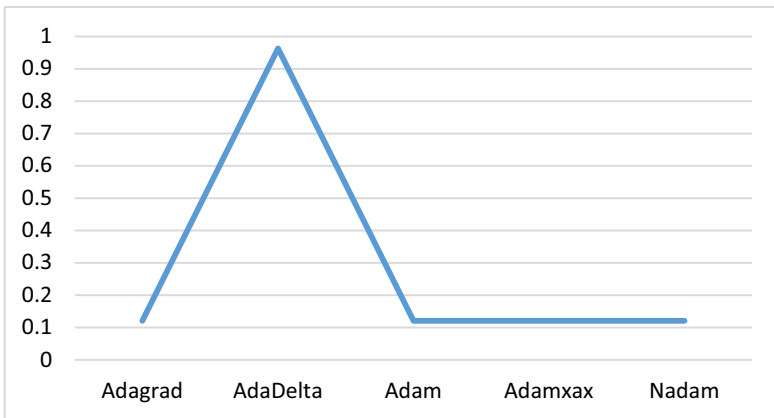


Fig. 5. Correlation between accuracy and optimizer.

Activation function is actually the function relation between the output of upper node and the input of lower node. The common activation function includes Sigmoid, tanh, ReLU, elu and softplus function. The Sigmoid function in deep neural network causes the gradient to disappear when the gradient reverse transmission, which has been less used. Use a different activation function in the penultimate full connection layer. The last full connection layer is a classifier, all other things being equal. In order to avoid chance, each activation function was repeated three times, and the accuracy rate was averaged. It is found that the activation function of elu is slightly better than that of relu.

The optimizer is used to minimize the loss function, for example, SGD, RMSprop, Adagrad, AdaDelta, Adam. When SGD randomly selects the gradient, noise will be introduced, so that the direction of weight updating is not necessarily correct, and the problem of local optimal solution is not solved. RMSprop is often used to train circular neural networks. The effects of Adagrad, AdaDelta and Adam on the model were compared. Experimental results show that the AdaDelta optimizer is more accurate than others. Adadelta is an extended version with greater robustness. Instead of accumulating all the past gradients, it adjusts the learning rate according to the updated movement window. So despite many updates, Adadelta continues to learn.

5 Conclusion

Based on the classic CNN_ON_MNIST and CNN_ON_CIFAR, this paper proposes MC_CNN which can identify water meters. Based on the raspberry PI platform, it can be read and identified by external hardware and software, without the need to transform the inside of the water meter. Compared with the traditional magnetic induction reading can effectively avoid the problem of magnetic disappearance. Compared with pattern matching, it can reduce the complexity of image preprocessing and the error of manual parameter selection. Compared with other neural networks, accuracy and loss values are also better. The optimal rescaling factor, optimizer and activation function were found by experiments. Mature development can be further extended to the identification of electricity meters, heating meters.

Acknowledgment. This work was supported by the Natural Science Foundation of Jiangsu Higher Education Institutions of China (No. 17KJB520010), and Research foundation of Nanjing Institute of Technology (No. CKJB201804), and Practice and innovation training program for college students in jiangsu province (No. 213345214301801).

References

1. Zhang, Y., Wang, S., Dong, Z.: Classification of Alzheimer disease based on structural magnetic resonance imaging by kernel support vector machine decision tree. *Prog. Electromagn. Res.* **144**, 171–184 (2014)
2. Chaplot, S., Patnaik, L.M., Jagannathan, N.R.: Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network. *Biomed. Sign. Process. Control* **1**, 86–92 (2006)

3. Maitra, M., Chatterjee, A.: A Slantlet transform based intelligent system for magnetic resonance brain image classification. *Biomed. Sign. Process. Control* **1**, 299–306 (2006)
4. El-Dahshan, E.S.A., Hosny, T., Salem, A.B.M.: Hybrid intelligent techniques for MRI brain images classification. *Digit. Sign. Process.* **1**, 299–306 (2006)
5. Zhang, Y., Wu, L., Wang, S.: Magnetic resonance brain image classification by an improved artificial bee colony algorithm. *Prog. Electromagn. Res.* **116**, 65–79 (2011)
6. Zeng, Y., Xu, X., Fang, Y., Zhao, K.: Traffic sign recognition using deep convolutional networks and extreme learning machine. In: He, X., Gao, X., Zhang, Y., Zhou, Z.-H., Liu, Z.-Y., Fu, B., Hu, F., Zhang, Z. (eds.) *IScIDE 2015. LNCS*, vol. 9242, pp. 272–280. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-23989-7_28
7. Guo, L., Ding, S.: A hybrid deep learning cnn-elm model and its application in handwritten numeral recognition. *J. Comput. Inf. Syst.* **11**(7), 2673–2680 (2015)
8. Gurpinar, F., Kaya, H., Dibeklioglu, H., Salah, A.: Kernel ELM and CNN based facial age estimation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 80–86 (2016)
9. Youngwoo, Y., Oh, S. Y.: Fast training of convolutional neural network classifiers through extreme learning machines. In: *2016 International Joint Conference on Neural Networks (IJCNN)*. IEEE (2016)
10. Weng, Q., Mao, Z., Lin, J., Guo, W.: Land-use classification via extreme learning classifier based on deep convolutional features. *IEEE Geosci. Remote Sens. Lett.* **14**, 704–708 (2017)