



Optical Character Recognition of Power Equipment Nameplate for Energy Systems Based on Recurrent Neural Network

Xun Zhang^{1*}, Wanrong Bai¹ and Haoyang Cui^{2*}

¹Gansu Power Grid Co., Ltd., Electric Power Research Institute, Lanzhou, China, ²College of Electronic and Information Engineering, Shanghai University of Electric Power, Shanghai, China

OPEN ACCESS

Edited by:

Bin Zhou,
Hunan University, China

Reviewed by:

Xiaoshun Zhang,
Shantou University, China
Du Gang,
North China Electric Power University,
China

*Correspondence:

Haoyang Cui
cuihy@shiep.edu.cn

Specialty section:

This article was submitted to
Process and Energy Systems
Engineering,
a section of the journal
Frontiers in Energy Research

Received: 13 December 2021

Accepted: 28 December 2021

Published: 10 August 2022

Citation:

Zhang X, Bai W and Cui H (2022)
Optical Character Recognition of
Power Equipment Nameplate for
Energy Systems Based on Recurrent
Neural Network.
Front. Energy Res. 9:834283.
doi: 10.3389/fenrg.2021.834283

To address the problems of poor accuracy and response time of optical character recognition of power equipment nameplates for energy systems, which are ascribed to exposure to natural light and rainy weather, this paper proposes an optical character recognition algorithm for nameplates of power equipment that integrates recurrent neural network theory and algorithms with complex environments. The collected image power equipment nameplates are preprocessed via graying and binarization in order to enhance the contrast among features of the power equipment nameplates and thus reduce the difficulty of positioning. This innovation facilitates the application of image recognition processing algorithms in power equipment nameplate positioning, character segmentation, and character recognition operations. Following segmentation of the power equipment nameplate and normalization thereof, the characters obtained are unified according to size, and then used as the input of the recurrent neural network (RNN); meanwhile, corresponding Chinese characters, numbers and alphabetic characters are used as the output. The text data recognition system model is realized via the trained RNN network, and is verified by inputting a large dataset into training. Compared with existing text data recognition systems, the algorithm proposed in this paper achieves a Chinese character recognition accuracy of 99.90%, an alphabetic and numeric character recognition accuracy of 99.30%, and a single image recognition speed of 2.15 ms.

Keywords: energy systems, optical character recognition, artificial intelligence, power equipment nameplate, recurrent neural network (RNN), deep learning

1 INTRODUCTION

In the process of operating and inspecting power grid equipment, the operations and inspection personnel will record both manually (in text form) and electronically unstructured information such as power equipment failure phenomena, defects and corresponding elimination methods, and operations and maintenance measures (Li et al., 2021a; Li and Yu, 2021a; Li et al., 2021b; Li and Yu, 2021b). Such data not only indicate the working history and operational status of equipment, but also bear a wealth of latent fault information (Krizhevsky et al., 2012). Where traditional methods of record-keeping are applied, experts must rely on experience when assessing the operational status of the equipment. Due to subjective factors such as personnel experience, familiarity with the

equipment and the text description method, false detections, missed inspections or mis-operations often occur, and those in turn can lead to major economic losses. Rapid increases in the scale and number of substations have resulted in a surging accumulation of unstructured text data, the misinterpretation of which poses great safety hazards as regards the daily operation and inspection of power grid equipment. Further, utilities providers face the challenge of replacing traditional methods with those compatible with smart grid systems (Nair and Hinton, 2010).

The existing nameplates of power equipment are very complex (Li and Yu, 2022). They contain asset configuration information, equipment operation information, work tickets, operation tickets, operation and inspection logs, long-document reports, authoritative standards, etc (Li et al., 2022). The work tickets and operation tickets are in the form of short texts, and include mainly details of routine visits, repair and defect elimination, and existing defects (Li and Yu, 2022). The formats of above-mentioned different types of text data are varied. These include handwritten unstructured texts, paper and electronic documents, digital electronic archives, etc (Li et al., 2022). Therefore, it is difficult to directly analyze text data consistently or efficiently. In addition, text data also have the following characteristics: 1) There may be numbers, units, and even formulae in the text. The quantitative information of above-mentioned data plays a decisive role in the classification of defects, but these text characters are easily lost in the text mining process (Kumar and Singh, 2005). 2) The content of the text records varies in terms of details and length. Manual records are easily affected by subjective factors. Differences in description produce different expressions of the same phenomenon, and complicate the reader's understanding of the semantics of text (Hinton and Salakhutdinov, 2006). 3) There are many technical terms relating to power equipment operations and inspection, and there are different ways of describing the same parts or processes. An artificial intelligence algorithm can recognize a large number of images via training, so as to realize quick and accurate recognition in practice. Scholars in the field have conducted much research into the recognition of power character and symbols in complex environments (Poon and Domingos, 2011).

At present, machine learning is used widely in OCR technology. Cun et al. first proposed the use of convolutional neural networks in the field of optical character recognition (Lecun et al., 1998). In their study, the main recognition target was handwritten characters, that is, handwritten numbers with a size of 32*32 pixels on a bank check. First, target images are inputted into the convolutional neural network. Following multi-layer feature extraction, a variety of different types of local features could be obtained. Through analysis of the final local features and the relative position relationship between local features, the desired results could be obtained (Lecun et al., 1998).

Wang et al. (2012) proposed a multi-layer convolutional neural network for the positioning and recognition of characters in real-life pictures. They first created a detection sliding window for obtaining a series of areas where characters are likely to appear, and analyzed the spaces in the characters.

After obtaining the text area, they employed an algorithm similar to Viterbi for segmenting the character area, and then inputted the segmented single character into the next convolutional neural network for recognition. Finally, following splicing, they compared the single character recognition results against entries in the dictionary in their system to obtain a word with the highest probability, that is the final output, or the character content contained in the character area of the image. In the end, they achieved an accuracy of 90% using the ICDAR2003 dataset and an accuracy of 70% using the Street View Text dataset.

He et al. (2016) proposed non-segmenting recognition methods for complete character images by using a combination of Convolutional Neural Network (CNN) and RNN.

Zhang et al. (2021) first used CNN to perform sliding window operation on the original image, and transmitted the output result of each sliding window as input to the RNN network for processing, so as to obtain the final result. Shi et al. first inputted the character picture to the CNN network, and then inputted the pixels of each column of the feature to the RNN network in the order of left to right before outputting the final result. Johnson and Bird, 1990 proposed an image recognition technology that integrates image processing technology and pattern recognition technology. Kim et al. (2000) proposed an image recognition method based on an artificial neural network and support vector machine model. In earlier times, picture pixilation was low, and the pictures taken were affected highly by the outside world due to the poor performance of image extraction equipment (cameras, computers, etc.). The computer operating speed was limited, and algorithms lacked the necessary robustness for complex environments and could not function as universal image recognition algorithms. In addition, early image recognition algorithms mainly processed images based on the basic texture characteristics of images, and then used suitable template or traditional artificial neural networks to achieve final character recognition. Since most foreign images only contain letters and numbers, the recognition rate can exceed 90%. In recent years developers have achieved considerable advances in computer hardware and semiconductors, and have arrived at more creative solutions for machine learning and deep learning for image recognition technology. Eskandarpour and Khodaei, 2018 proposed an image recognition method based on fuzzy support vector machines, which uses the particle swarm optimization algorithm to adjust the parameters of the support vector machine. They performed recognition experiments using a collection of Malaysian images, yielding relatively ideal experimental results. Li et al. (2022) proposed a new method of optical character recognition, nearest neighbor algorithm, which became the first classification model due to its efficiency and simplicity in processing noisy data sets and large data sets. Multi-class support vector machines have been devised to resolve confusion between similar characters and symbols, and significantly improve the recognition rate of optical symbols by combining the k-nearest neighbor algorithm with a multi-class support vector machine. Rafique et al. (2018) proposed a neural network method and applied it to public data sets for optical character recognition. In a comparative study, they demonstrated that their method has a higher recognition rate

than that of traditional methods. Epshtein et al. (2010) proposed an optical character recognition method that first uses the winnows classifier of the weak sparse network to extract the candidate region, and then uses the convolutional neural network classifier to filter it. When the recognition system was deployed in the United States, they demonstrated that this method could attain more accurate image recognition in the presence of differences in character width and spacing and noise interference. Zaheer and Shaziya, 2018 proposed an optical character recognition method with regional convolutional neural network based on deep learning technology, which can recognize image and video sequences more accurately than traditional recognition methods. At the same time, Austria company launched an optical character recognition system based on machine learning technology that can be directly applied to the mobile phone platform, whereby optical character recognition can be automatically performed on a stored image.

Nevertheless, the above algorithms have the following problems:

- (1) SVM performs well with small-scale samples as its generalization ability is outstanding by classification algorithms standards; however, it is less effective with large-scale high-dimension data.
- (2) They cannot be optimized easily as existing methods regard OCR as two independent tasks - text detection, and recognition.
- (3) Their recognition speed is too slow to meet real-time requirements as it is restricted by the complicated underlying processes to a certain extent.

In order to solve the shortcomings of low recognition rate or limited range in terms of character recognition in the aforementioned methods, the method proposed in this paper first performs preprocessing (including grayscale and binarization) on the target original image, and then performs multiple operations such as positioning and character segmentation for power grid characters, prior to image recognition processing. The algorithm is designed to enhance the contrast between colors in and around the power equipment nameplate and thus reduce the difficulty of positioning. Secondly, the image recognition processing algorithm is used to perform operations such as power equipment nameplate positioning, character segmentation, and character recognition. The characters obtained after segmentation of the power equipment nameplate are normalized, unified according to size, and then used as the input of the recurrent neural network (RNN), while corresponding Chinese characters, numbers and alphabetic characters are used as the output. The text data recognition system model is realized via the trained RNN network. It is verified using a large data set and its training is compared with an existing text data recognition system. The algorithm proposed in this paper achieves a Chinese character recognition accuracy of 99.90%, an alphabetic and numeric character recognition accuracy of 99.30%, and a single image recognition speed of 2.15 ms.

The remainder of this paper is structured as follows:

Section 2 describes in detail the RNN-based OCR algorithm; **Section 3** provides and analyzes the case study results; and, **Section 4** draws conclusions on the aims, arguments and findings presented in this paper.

2 RNN-BASED OCR ALGORITHM

2.1 Optical Character Recognition Process Based on RNN Neural Network

The structural diagram of the image recognition method based on Long short term memory (LSTM) is shown in **Figure 1**. The specific process is as follows:

- (1) Image preprocessing. Given that the quality of the power equipment nameplate is affected by the natural environment and lighting conditions during the process of obtaining the original image of power equipment nameplate, it is necessary to perform image preprocessing operations such as graying, edge detection, and binarization on the image of power equipment nameplate in order to enhance the degree of contrast in the image area of the power equipment nameplate, and improve the accuracy of positioning.
- (2) Power equipment nameplate positioning. First, the color features are extracted from the RGB color space, and the rough positioning is completed on the basis of the color feature information. Then, the gray level jump is used to roughly determine the area of the power equipment nameplate. Finally, the area is filtered in the vertical and horizontal directions. The area that meets the aspect ratio of the power equipment nameplate is the target area.
- (3) Character segmentation. The vertical projection method is used to analyze the pixel distribution of the binarized power equipment nameplate image in the determined target area, and then the appropriate threshold is selected in order to complete the character segmentation.
- (4) Training network. The characters obtained from segmentation of the power equipment nameplate are normalized, unified according to size, and then used as the input of the recurrent neural network (RNN).
- (5) Character recognition. After the obtained nameplate pictures of power equipment are processed via steps (1) to (3), they are inputted into the trained RNN network via step (4) to obtain the recognition results.

2.2 RNN Algorithm

The data processed using the fully connected network or CNN has a common feature: each two data are independent of each other and are not correlated, so these networks will expand in space to deeply analyze the information contained in each single datum. However, if the data are based on spoken conversation (vocal communication), in which information content is reflected by the correlation between each data point, then the two networks mentioned above are powerless. It is for this reason that RNN has been developed. RNN is used mainly to process time-series data, and has become a popular research method in disciplines such as natural language processing.

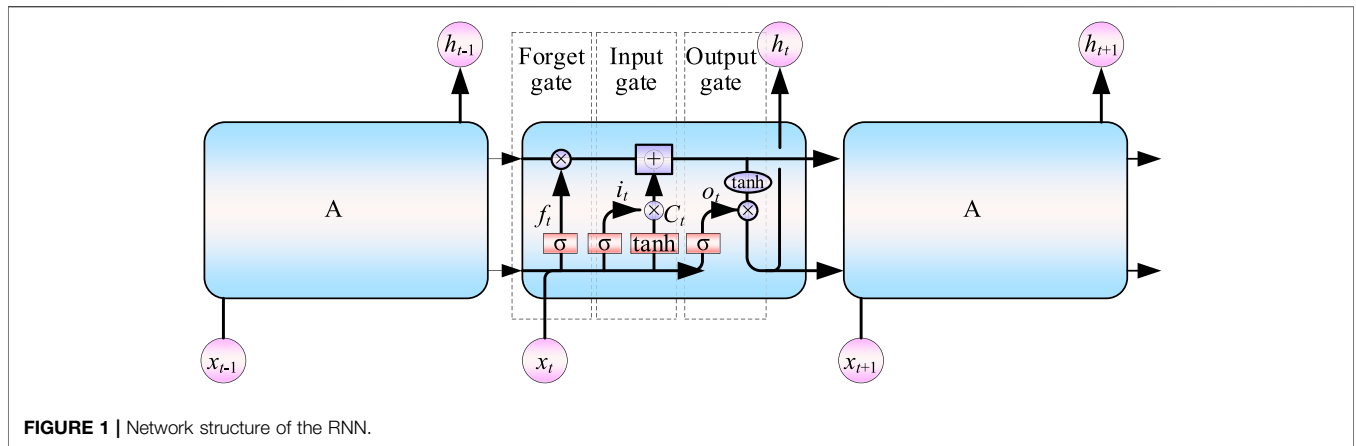


FIGURE 1 | Network structure of the RNN.

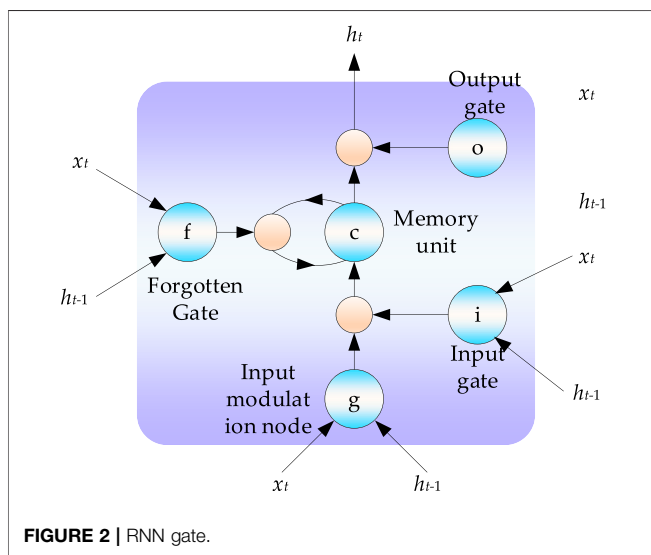


FIGURE 2 | RNN gate.

In **Figure 1**, the left side shows a simple RNN structure. At each moment, x_t is input. From the neural operation of W , the output y_t is obtained, and a recessive state h_t is generated at the same time. The recessive state is combined with the input datum x_{t+1} at the next moment as the joint input at the next moment, and then through the operation of the neuron W the output y_{t+1} and the recessive state h_{t+1} are obtained. This process is repeated until all the data are used. When the process is expanded according to the timeline, the structure shown on the right side of **Figure 1** can be obtained. The operation of the neuron W can be expressed by **Eq. 1**, wherein ω is the coefficient, b is the bias, h_{t-1} is the network output at the previous moment, and $f(\cdot)$ is the activation function.

$$y_t = f(\omega \cdot [h_{t-1}, x_t] + b) \tag{1}$$

After the RNN receives the input x_t at moment t , the value of the hidden layer is s_t , and the output value is s_{t+1} . However, unlike traditional neural networks, the value of s_t depends not only on x_t , but also on the hidden value s_{t-1} at the previous moment. The mathematical formula of the RNN is as follows:

$$o_t = g(V \times s_t)V \tag{2}$$

$$s_t = f(Ux_t + Ws_{t-1}) \tag{3}$$

Equation 2 shows how the output layer of the RNN is calculated. The output layer can be regarded as a fully connected network, wherein V represents the weight matrix of the output layer and g represents the activation function. **Eq. 3** expresses how the hidden layer is calculated, wherein U represents the weight matrix of the input x , W represents the weight matrix of s_{t-1} at the previous moment as the input at this moment, and f is the activation function of the cyclic layer. By repeatedly substituting **Eq. 3** into **Eq. 2**, one obtains the final expression of the RNN:

$$o_t = Vf(Ux_t + Wf(Ux_{t-1} + Wf(Ux_{t-2} + Wf(Ux_{t-3} + \dots)))) \tag{4}$$

As shown in **Figure 2**, the detailed steps of the three gates in the RNN can be summarized as follows:

- (1) Forget Gate f : This decides what kind of information will be eliminated, that is, how much cell state c_{t-1} from the previous moment is retained to current c_t . The value in the interval (0,1) is the input in each state, wherein, 0 stands for “completely forgotten”, and 1 is the opposite. The forget gate is calculated as follows:

$$f_t = \sigma(W_{fx}x_t + W_{fh}h_{t-1} + b_f) \tag{5}$$

Among them, f_t is the activation of forget gate f_t , W_{fx} is the weight vector from the input layer to f_t , W_{fh} is the weight vector from the hidden layer to f_t , and b_f is the offset of f_t . $\sigma(\cdot)$ is the sigmoid activation function.

- (2) Input Gate i : This determines how much information current x_t retains for current state c_t , and comprises two parts. First, the sigmoid layer (Input gate layer) determines what value to be updated at the next moment. Then, a new candidate value vector is created and added to the state in the tanh layer. This is calculated as follows:

$$i_t = \sigma(W_{ix}x_t + W_{ih}h_{t-1} + b_i) \quad (6)$$

Among them, i_t is the activation of input gate i_t , W_{ix} is the weight vector from the input layer to i_t , W_{ih} is the weight vector from the hidden layer to i_t , and b_i is the offset of i_t .

- (3) Output Gate o : This determines how much C_t transmits to output h_t of the current state. The output is based on the cell state. First, the sigmoid layer is run to determine the output of the cell state. Then, the cell state is processed through the tanh layer to get the genetic information, which is a value in the interval (0,1). After the value is multiplied by the output of sigmoid gate, only the target part will be outputted.

$$o_t = \sigma(W_{ox}x_t + W_{oh}h_{t-1} + b_o) \quad (7)$$

Among them, o_t is the activation of output gate o_t , W_{ox} is the weight vector from the input layer to o_t , W_{oh} is the weight vector from the hidden layer to o_t , and b_o is the offset of the Output gate.

The calculation results of input gate, forget gate, and Output gate differ from one another, but all three sets are obtained by multiplying the current input sequence x_t and the previous state output h_{t-1} by the corresponding weight plus the corresponding offset, and then solving via the sigmoid activation function. The immediate state is activated by using the tanh activation function (hyperbolic tangent activation function), for which the formula is as follows:

$$\bar{C}_t = \tanh(W_{cx}x_t + W_{ch}h_{t-1} + b_c) \quad (8)$$

Equation 8 is combined with formulae (1) (2) (3) to update the old cell state, and then the old state is multiplied by f_t to forget the expiration information. Then, $i_t * \bar{C}_t$ is added to obtain the new candidate value, that is:

$$C_t = f_t * C_{t-1} + i_t * \bar{C}_t \quad (9)$$

Thereafter, the formula of the output h_t of the RNN unit is:

$$h_t = o_t * \tanh(C_t) \quad (10)$$

3 CASE STUDIES

3.1 Experimental Data

The experimental setup is as follows:

Both the simulation model and programs described above have been developed using a server consisting of 48 CPUs. Each CPU is a 2.10 GHz Intel Xeon Platinum processor, and the RAM of the server is 192 GB. The simulation software package used here is MATLAB/Simulink version 9.8.0 (R2020a).

The standard power equipment nameplates in China are composed of Chinese characters, numbers, and Arabic letters, with a total of 65 different characters; however, due to the particularity of China's power equipment nameplates, it is more difficult to identify the nameplates of power equipment in China (Huang et al., 2021).

Matlab is an internationally recognized, excellent numerical calculation and simulation analysis software, the use of which

substitutes complicated calculations with large data sets. In this study, Matlab is used to realize the recognition module of power equipment nameplates (Gao et al., 2017).

Since the training of the RNN network is based on a large number of labeled training sets, the data set used should reflect the many different power equipment nameplates in substations all over the country. In reality, it is too costly and impractical to capture pictures of all existing power equipment nameplates, and so there is no standard data set available for training the model proposed in this study; the method in this paper divides the training data collecting process into shooting and collection of power equipment nameplates. In addition, the data set contains the pictures of power equipment nameplates of substations obtained in a wide variety of environmental conditions including, for example, 300 pictures of power equipment nameplates of substations during exposure to natural light, rainy weather, etc. These are high-definition power equipment nameplates of substations taken by mobile phones, unmanned aerial vehicles, etc. In addition, 9000 power equipment nameplate images substations were obtained via an online search using Google. In order to increase the diversity of the samples, the collected plate images were stretched, zoomed, and cut; then, all image of power equipment nameplates of the substations were preprocessed to obtain the training samples which is shown below (Ren et al., 2017).

The RNN network trained more than 9300 pictures as the training samples. **Figure 3** shows the power equipment nameplates samples of some substations used in the experiment.

3.2 Experimental Results

In order to illustrate the superiority of the RNN algorithm-based image recognition system, this paper evaluates the recognition capability of each algorithm from three aspects, namely the recognition accuracy of Chinese characters, the recognition accuracy of letters and numbers, and the image recognition time (Li et al., 2021d). The test set of this paper includes 30 images of power equipment nameplates from all over the country. The experimental results are shown in **Table 1**.

The results in **Table 1** show that the accuracy in terms of Chinese character recognition, letters, and numbers recognition signifies an improvement on previous results in the experiment; at the same time, the image recognition time is relatively short. DBN (Epshtein et al., 2010), BP (Eskandarpour and Khodaei, 2018), gate feedback neural network (GRNN) (Raghunandan et al., 2018), ANNs, and the feedback neural network can only recognize characters and numbers, but cannot recognize Chinese characters. SVM (Kim et al., 2000), CNN (He et al., 2016), DNN (Zaheer and Shaziya, 2018) and RNN can recognize Chinese characters, alphabetic characters and numbers, indicating that the deep-learning algorithm is a more suitable candidate for recognition of power equipment nameplates.

Regarding the evaluation criteria, the results in **Table 1** also show that the proposed power equipment nameplates recognition algorithm based on the RNN network attain an accuracy of 99.90% for Chinese character recognition, an



FIGURE 3 | Training samples.

TABLE 1 | Comparison of experimental results.

Algorithm	Chinese character recognition accuracy	Letter and number recognition accuracy rate (%)	Time(ms)
SVM Kim et al. (2000)	96.3%	96.91	5.97
CNN He et al. (2016)	98.68%	98.12	3.11
DNN Zaheer and Shaziya, 2018	98.23%	99.15	3.32
DBN Epshtein et al. (2010)	-	93.27	5.22
BP Eskandarpour and Khodaei, 2018	-	91.93	5.92
GRNN Raghunandan et al. (2018)	-	91.91	5.26
ANNs Rafique et al. (2018)	-	90.92	5.91
FNN Cho et al. (2014)	-	95.13	5.17
RNN	99.90%	99.30	2.15

accuracy of 99.30% for letter and numeric character recognition, and a single image recognition speed of 2.15 ms. Compared with DBN, BP, gate feedback neural network, ANNs, and feedback neural network, the algorithm proposed in this paper not only is more capable of recognizing letters and numbers, but also can more accurately read Chinese characters at an improved speed; specifically, compared with SVM (Kim et al., 2000), CNN (He et al., 2016), DNN (Zaheer and Shaziya, 2018), the accuracy of Chinese characters recognition is increased by 3.60%, the accuracy of letter and number recognition is increased by 2.41%, and the operating efficiency of the algorithm is increased by at least 67%. Compared with other methods, the operational efficiency of the algorithm in this paper is substantially higher. In addition, it not only recognizes Chinese characters, but does so more accurately, attaining a level of accuracy that is up to 2.41% above that or rival methods. In summary, the above comparative experiments show that the method proposed in this paper has higher recognition accuracy, and better algorithm operational efficiency.

4 CONCLUSION

From the information presented in the previous sections, the following conclusions are drawn:

This paper proposes an optical character recognition algorithm for nameplates of power equipment that integrates recurrent neural network theory and algorithms into complex environments. The collected image power equipment nameplates have been preprocessed via graying and binarization in order to enhance the degree of contrast in the images of power equipment nameplates and thus reduce the difficulty of positioning. Then, image recognition processing algorithms were used for power equipment nameplate positioning, character segmentation, and character recognition operations. The characters obtained from segmentation of the power equipment nameplate were normalized, unified according to size, and then used as the input of the recurrent neural network (RNN), while corresponding Chinese characters, numbers and alphabetic characters were used as the output. The text data recognition system model is realized via the trained RNN network, and its capabilities have been verified using a large data set and training

process. Compared with existing text data recognition systems, the algorithm proposed in this paper achieves a Chinese character recognition accuracy of 99.90%, an alphabetical and numeric character recognition accuracy of 99.30%, and a single image recognition speed of 2.15 ms.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

REFERENCES

- Cho, K., Van Merriënboer, B., Bahdanau, D., and Bengio, Y. (2014). *On the Properties of Neural Machine Translation: Encoder-Decoder Approaches*. Available at: <https://arxiv.org/abs/1409.1259> (Accessed December 10, 2021).
- Epshtein, B., Ofek, E., and Wexler, Y. (2010). "Detecting Text in Natural Scenes with Stroke Width Transform," in 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, 13-18 June 2010.
- Eskandarpour, R., and Khodaei, A. (2018). Leveraging Accuracy-Uncertainty Tradeoff in SVM to Achieve Highly Accurate Outage Predictions. *IEEE Trans. Power Syst.* 33, 1139–1141. doi:10.1109/tpwrs.2017.2759061
- Gao, Y., Chen, Y., Wang, J., and Lu, H. (2017). *Reading Scene Text with Attention Convolutional Sequence Modeling*. Available at: <https://arxiv.org/abs/1709.04303v1> (Accessed December 10, 2021).
- He, P., Huang, W., Qiao, Y., Loy, C. C., and Tang, X. (2016). "Reading Scene Text in Deep Convolutional Sequences," in 30th AAAI Conference on Artificial Intelligence, Phoenix, Arizona, USA, February 2016.
- Hinton, G. E., and Salakhutdinov, R. R. (2006). Reducing the Dimensionality of Data with Neural Networks. *Science* 313, 504–507. doi:10.1126/science.1127647
- Huang, S., Wu, Q., Liao, W., Wu, G., Li, X., and Wei, J. (2021). Adaptive Droop-Based Hierarchical Optimal Voltage Control Scheme for VSC-HVdc Connected Offshore Wind Farm. *IEEE Trans. Ind. Inf.* 17, 8165–8176. doi:10.1109/tii.2021.3065375
- Johnson, A. S., and Bird, B. M. (1990). "Number-Plate Matching for Automatic Vehicle Identification," in IEE Colloquium on Electronic Images and Image Processing in Security and Forensic Science, London, UK, 22-22 May 1990.
- Kim, K. K., Kim, K. I., Kim, J. B., and Kim, H. J. (2000). "Learning-Based Approach for License Plate Recognition," in IEEE Signal Processing Society Workshop, Sydney, NSW, Australia, 11-13 Dec. 2000.
- Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). "ImageNet Classification with Deep Convolutional Neural Networks," in 25th International Conference on Neural Information Processing Systems, Lake Tahoe, Nevada, USA, December 2012.
- Kumar, S., and Singh, C. (2005). "A Study of Zernike Moments and its Use in Devnagari Handwritten Character Recognition," in International Conference on Cognition and Recognition, Mandya, Karnataka, India, December 2005.
- Lecun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-Based Learning Applied to Document Recognition. *Proc. IEEE* 86, 2278–2324. doi:10.1109/5.726791
- Li, J., Yao, J., Yu, T., and Zhang, X. (2021d). Distributed Deep Reinforcement Learning for Integrated Generation-Control and Power-Dispatch of Interconnected Power Grid with Various Renewable Units. *IET Renewable Power Generation*. doi:10.1049/rpg2.12310
- Li, J., and Yu, T. (2021a). A New Adaptive Controller Based on Distributed Deep Reinforcement Learning for PEMFC Air Supply System. *Energ. Rep.* 7, 1267–1279. doi:10.1016/j.egy.2021.02.043

AUTHOR CONTRIBUTIONS

XZ: Conceptualization, methodology, software, data curation, writing-original draft preparation, visualization, investigation, software, validation. WB: Writing- Reviewing and editing. HC: Supervision.

FUNDING

This work was jointly supported by science and technology projects of Gansu State Grid Corporation of China (52272220002U).

- Li, J., and Yu, T. (2021b). A Novel Data-Driven Controller for Solid Oxide Fuel Cell via Deep Reinforcement Learning. *J. Clean. Prod.* 321, 128929. doi:10.1016/j.jclepro.2021.128929
- Li, J., Yu, T., and Yang, B. (2021b). A Data-Driven Output Voltage Control of Solid Oxide Fuel Cell Using Multi-Agent Deep Reinforcement Learning. *Appl. Energ.* 304, 117541. doi:10.1016/j.apenergy.2021.117541
- Li, J., Yu, T., and Zhang, X. (2022). Coordinated Automatic Generation Control of Interconnected Power System with Imitation Guided Exploration Multi-Agent Deep Reinforcement Learning. *Int. J. Electr. Power Energ. Syst.* 136, 107471. doi:10.1016/j.ijepes.2021.107471
- Li, J. (2022). Large-Scale Multi-Agent Reinforcement Learning-Based Method for Coordinated Output Voltage Control of Solid Oxide Fuel Cell. *Case Studies in Thermal Engineering* 4, 101752. doi:10.1016/j.csite.2021.101752
- Li, J., Yu, T., and Zhang, X. (2022). Coordinated Load Frequency Control of Multi-Area Integrated Energy System Using Multi-Agent Deep Reinforcement Learning. *Appl. Energ.* 306, 117900. doi:10.1016/j.apenergy.2021.117900
- Li, J., Yu, T., and Zhang, X. (2021c). Emergency Fault Affected Wide-Area Automatic Generation Control via Large-Scale Deep Reinforcement Learning. *Eng. Appl. Artif. Intelligence* 106, 104500. doi:10.1016/j.engappai.2021.104500
- Li, J., Yu, T., Zhang, X., Li, F., Lin, D., and Zhu, H. (2021a). Efficient Experience Replay Based Deep Deterministic Policy Gradient for AGC Dispatch in Integrated Energy System. *Appl. Energ.* 285, 116386. doi:10.1016/j.apenergy.2020.116386
- Nair, V., and Hinton, G. E. (2010). "Rectified Linear Units Improve Restricted Boltzmann Machines," in 27th International Conference on International Conference on Machine Learning, Haifa, Israel, June 2010.
- Poon, H., and Domingos, P. (2011). "Sum-Product Networks: A New Deep Architecture," in 2011 IEEE International Conference on Computer Vision Workshops, Barcelona, Spain, 6-13 Nov. 2011.
- Rafique, M. A., Pedrycz, W., and Jeon, M. (2018). Vehicle License Plate Detection Using Region-Based Convolutional Neural Networks. *Soft Comput.* 22, 6429–6440. doi:10.1007/s00500-017-2696-2
- Ragunandan, K. S., Shivakumara, P., Jalab, H. A., Ibrahim, R. W., Kumar, G. H., Pal, U., et al. (2018). Riesz Fractional Based Model for Enhancing License Plate Detection and Recognition. *IEEE Trans. Circuits Syst. Video Technol.* 28, 2276–2288. doi:10.1109/tcsvt.2017.2713806
- Ren, S., He, K., Girshick, R., and Sun, J. (2017). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* 39, 1137–1149. doi:10.1109/tpami.2016.2577031
- Wang, T., Wu, D. J., Coates, A., and Ng, A. Y. (2012). "End-to-End Text Recognition with Convolutional Neural Networks," in 21st International Conference on Pattern Recognition, Tsukuba, Japan, 11-15 Nov. 2012.
- Zaheer, R., and Shaziya, H. (2018). "GPU-based Empirical Evaluation of Activation Functions in Convolutional Neural Networks," in 2nd International Conference on Invention Systems and Control, Coimbatore, India, 19-20 Jan. 2018.
- Zhang, K., Zhou, B., Or, S. W., Li, C., Chung, C. Y., and Voropai, N. I. (2021). "Optimal Coordinated Control of Multi-Renewable-To-Hydrogen Production System for Hydrogen Fueling Stations," in IEEE Transactions on Industry Applications, 1. doi:10.1109/TIA.2021.3093841

Conflict of Interest: Author HC is employed by Shanghai University of Electric Power. Author XZ and WB are employed by Gansu Power Grid Co., Ltd. Electric Power Research Institute Electric.

The remaining author declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of

the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Zhang, Bai and Cui. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.