An Automatic Electricity Meter Reading System Based on Neural Vision Location

Jin Yin¹, Pengfei Xia¹, Jingsong He¹*, Li Gu¹ and Kejun Yang²
¹University of Science and Technology of China, Hefei, China
²Anhui Nari Jiyuan Electric Power System Tech CO., LTD, Hefei, China

*Email: hjss@ustc.edu.cn

Abstract. Remote meter automatic reading and smart meter reading is a vital part of smart grid and has been widely concerned. Traditional manual method is manpower consuming. Developing an automatic electricity ether reading system is necessary in practice. Two particular aspects in this task should be concerned. One is different-perspective problem, which causes by the variations of camera position and leads to pool display localization. Another is information diversity problem. Namely, the electrical meter used in this paper needs to recognize effective information from diversity information. This paper describes a novel automatic electricity meter reading system, which can automatically locate and recognize effective information. For the camera perspective variability problem, we utilize a new localization method based on neural network. Via actual meter monitoring experiment, the result shows our system can efficient recognition and achieve 97% accuracy.

1. Introduction
With the rapid development of industrial automation and intelligent instruments, the power industry is facing following challenges, including lack of automated analysis, huge resources consuming, etc. Therefore, smart grid gradually become a research hotspot [1, 2]. Remote meter automatic reading and smart meter reading, as a vital part of smart grid, are widely concerned. Traditional method of manual reading is manpower consuming, especially in practical application, where meter-data need to be all-round monitored. So, it is necessary to develop an automatic electricity meter reading system which can apply in real-world.

There are two particular aspects in the process of electricity meter reading. One is different-angel problem, which means pool display localization effect owing to the variations of camera position(level), shot lengths and angles, as shown in Figure 1.a. It is apparent that every dial in images has its own unique angle and dimension. Especially, the localization task is important in the process of recognition because small deviation will affect final recognition performance. Another challenge in electricity meter reading tasks is information-diversity problem. That is to say, the electrical meter used in this paper is not like house's electrical meter mentioned in [3,4,5]. House's electrical meter only focus on energy consumed, however current electrical meter contains variable information and not all diversity information is useful, so the major aspect of digit recognition system lies in recognizing effective information from diversity information, as shown in Figure 1.b.

For above aspects, there are several works for the location and recognition. Previously, information extraction from images and video usually make full use of image features, such as edges, strokes and corners feature [6,7,8]. In the past few years, due to the effective application of neural networks, some
work extract information via neural networks [9]. Similarly, we use semantic segmentation to solve the camera visual variability problem in this paper.

The semantic segmentation is a core computer vision task, which predicted the category of per-pixel that it belongs to. On the basis of encoder-decoder structure, U-net appends a useful supplementary structure to increase segmentation accuracy. It has been demonstrated an excellent capability in various image analysis tasks, such as remote sensing image analysis [10] and medical image analysis [11].

![Figure 1](image1.jpg)

**Figure 1.** Different dial in real-world.

The first row is different picture captured from different video. This is different-angel problem. The second row is different picture captured from same video but different time. This is information-diversity problem.

In this paper, we proposed a novel automatic electricity meter reading system based on neural vision location, which consists of two major stages: automatic location system and digit recognition system. The proposed automatic location system utilizes semantic segmentation to accurately determine display coordinates. Comparing with other methods, semantic segmentation has high precision and efficiency, this is mainly because semantic segmentation is less affected by other interference information, such as dial information. The digit recognition method is to read the corresponding digit displayed in reading area. The key features of the developed system are:

- It is effective to solve the problem having risen from the camera orientation.
- In the process of digit recognition, our system can automatically select effective information from diversity information collected.

2. Automatic localization system

The important aspect in the process of automatic electricity meter reading is accurate localization of reading area. Considering the outstanding performance of U-net in various semantic tasks, including biomedical image [12] and remote sensing [13] image we choose U-net structure as our segmentation network. Semantic segmentation and determined coordinates are mainly units in automatic localization system.

2.1. Semantic segmentation

2.1.1. Network. U-net structure is a symmetric structure with a contracting path(left path) and an expansive path(right path). Following basic U-net, contracting path is composed of continuous two $3 \times 3$
convolutions module, each convolution followed by activation function and a 2×2 max-pooling operation with stride 2. After two 3×3 convolution module, the number of feature channel is doubled. Similarly, every 3×3 convolutions modules in expensive path, each convolution followed by activation function and a 2×2 up-sampling. After 3×3 convolution module, the number of feature channel is halved which concatenate with corresponding feature map from contracting path, which can propagate context information from lo-level features to high-level features. This is important to edge-preserving. In our system, we shrink channel numbers and use parametric rectified linear unit (PReLU) instead of rectified linear unit (Relu) owing to “dying Relu” problem.

Figure 2. The framework of entire system.

2.1.2. Training process. For our training process, we set the “batchsize” to 16 and use batch normalization [14] to accelerate training. Since there are only few training samples for our system, it is essential to use appropriate data pre-processing and augmentation method. Considering very little training data available, we primarily partition entire image (1280×720) into multiple parts, whose resolution are 128×128. New images are generated by synthesizing three new rotations (90, 180, 270), random flipping, random adding noise and random blurring. The augmented images are used to train the network with the stochastic gradient descent.

Due to segmentation image produced by semantic segmentation presents lack of precision near boundaries of the reading area, the separation border is enhanced using weighted cross entropy. The basic cross-entropy for binary classification can be calculated by [15]:

\[ C = -\frac{1}{n} \sum_y \left[ y \ln p + (1 - y) \ln (1 - p) \right] \]  

(1)

Where \( y \in \{0, 1\}, y = 1 \) means the ground-truth class is consistent with predict class. \( p \in [0, 1] \) specifies estimated probability for the ground-truth class. The notable property of cross-entropy loss is same weight to all pixels. We consider address boundary problems by means of up-weights the loss assigned to boundary pixels. The proposed enhanced cross-entropy loss as:

\[ EC = -\frac{1}{n} \sum_x \alpha \left[ y \ln p + (1 - y) \ln (1 - p) \right] \]  

(2)

In practice \( \alpha \) may be treated as a parameter. In our system, we set \( \alpha = 5 \).

2.2. Determine coordinates
As in the previous section, The reading area is indicated using U-net structure and edge enhancement method. In order to segment display area from entire dial image, we consider determining coordinates by minimum bounding rectangle (MBR).
3. Digital recognition system
Digital recognition system is to identify different information displayed on different occasions. The system mainly consists of three different units: pre-processing, title digit recognition and content digit recognition. An overview of digital recognition system is depicted in Figure 3.

3.1. Pre-processing
In order to make further processing more conveniently, we used the following strategy to optimize the segmentation results: reading area belonging to diverse camera perspective have different shapes, therefore we performed region-correction step to rectify reading area to a fixed size. Then execute with image segmentation to separate content and title, because only the area where the content and title are the regions of interesting. The title digit represents the information categories, which is up to 19 classes but only 11 classes are effective. A curial step, during digital recognition system, is to distinguish title information. This detail is important because from the result of title information, effective content information can be judged.

3.2. Title digit recognition
Through the image pre-processing, the content region and title region have been extracted. Then title digit in the regions of interesting is identified firstly. The numeric are in the range of 001 and 019. Obviously in this situation, classification using neural works but recognition after digit segmentation can improve speed. Especially, The scale of the rectangle should be converted to fixed scale before classification. Figure 3 shows the process of title digit recognition.

3.3. Content digit recognition
The content digit, stands in contrast to title digit recognition to contain a small number, is more challenging. Focus on this problem, these steps are applied in order to eliminate noise and improve accuracy.

Firstly, the image process of filtering and binarization are performed on content image to get the binary image of the content, then shrink area where the digits are to reduce interference.

Secondly, evaluate if the information displayed on content area is valid. In order to ensure the validity of the information, we consider the consistency of title digit and corresponding content digit format. For example, the valid digit format (see Figure 3) contains nine digits and a decimal point, of which six...
numbers exist before the decimal point. To judge the validity of digit format, we construct a column histogram which counter the total value of pixels for each column in binary image. By this way, can we obtain the number of black pixel blocks and validity of digit format.

Thirdly, segment each digit and train a neural network to classify. Knowing the title digit and validity of digit format, the automatic positioning and segmentation can be utilized to obtain each digits. Similarly, the position is determined by a row histogram and a column histogram. By analyzing the histogram, the exact location of the digits is obtained. Based on above result, each digit can be segmented and converted to uniform scale. Afterwards, a neural network model is trained to classify each digit.

**Figure 4.** The process of digit recognition.

4. Experiments
To demonstrate the performance of proposed system, we test our system on five video sequences with a 1280×720 resolution camera collected from real-world, the first image frame from video is used to automatically locate. The details of our experiments are specified in this section, including each part of the test and overall result.

The U-net is trained by stochastic gradient descent (SGD) on Keras. The first step is to remark edge pixels only the area where the digits are where the digits are. After training 50 epochs, we determine coordinates after obtaining segmentation map from U-net model. To exhibit our contribution, we show the process of semantic segmentation as follows.

For following digit recognition problem, we tested the digit detection strategy presented in Section 3. In order to provide a practical system, we create a human-computer interaction. Figure 5 shows the finally experiment result, yellow rectangles means information class, green rectangles represents valid information. The obtained overall accuracy is 97% and the speed is approximately 10frame/s. The devices used for our experiments as follows, GPU:GTX 1060 (3GB); CPU: Intel Core i7-7700.

**Figure 5.** The final experiment result.
5. Conclusion
The current development trend in electric power industry is towards intelligent direction. Aim to provide a practical tool, we proposed a novel electricity meter reading system. The proposed system, including automatic localization system and digital recognition system, is able to automatically detect the area of interest and recognize effective information. Experimental results show that our system has ability to automatically recognise and its accuracy can achieve 97%.

Acknowledgments
This study was funded and supported by the National Natural Science Foundation of China through Grant No.61273315.

Reference
[1] Gungor V C, Sahin D, Kocak T, Ergut S, Buccella C, Cecati C and Hancke G P 2011 IEEE transactions on Industrial informatics 7 529-539
[2] Farhangi H 2010 IEEE power and energy magazine 8.
[3] Zhao S, Li B, Yuan J and Cui G 2005 Research on remote meter automatic reading based on computer vision Transmission and Distribution Conference and Exhibition: Asia and Pacific, 2005 IEEE/PES (IEEE) pp 1-4.
[4] Elrefaei L A, Bajaber A, Natheir S, AbuSanab N and Bazi M 2015 Automatic electricity meter reading based on image processing Applied Electrical Engineering and Computing Technologies (AEECT), 2015 IEEE Jordan Conference on (IEEE) pp 1-5
[5] Gupta N and Shukla D 2016 Design of embedded based automated meter reading system for real time processing Electrical, Electronics and Computer Science (SCEECS), 2016 IEEE Students’ Conference on (IEEE) pp 1-6.
[6] Jung K, Kim K I and Jain A K 2004 Pattern recognition 37 977-997.
[7] Shen H and Coughlan J 2006 Finding text in natural scenes by figure-ground segmentation Pattern Recognition, 2006. ICPR 2006. 18th International Conference on vol 4 (IEEE) pp 113-118
[8] Kim K, Byun H, Song Y, Choi Y W, Chi S, Kim K K and Chung Y 2004 Scene text extraction in natural scene images using hierarchical feature combining and verification Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on vol 2 (IEEE) pp 679-682.
[9] Vanetti M, Gallo I and Nodari A 2013 Pattern Recognition Letters 34 519-526.
[10] Lu K, Sun Y and Ong S H 2018 Dual-resolution u-net: Building extraction from aerial images 2018 24th International Conference on Pattern Recognition (ICPR) (IEEE) pp 489-494
[11] Oktay O, Schlemper J, Folgoc L L, Lee M, Heinrich M, Misawa K, Mori K, McDonagh S, Hammerla N Y, Kainz B et al. 2018 arXiv preprint arXiv:1804.03999
[12] Alom M Z, Hasan M, Yakopcic C, Taha T M and Asari V K 2018 arXiv preprint arXiv: 1802.06955.
[13] Zhang Z, Liu Q and Wang Y 2018 IEEE Geoscience and Remote Sensing Letters
[14] Ioffe S and Szegedy C 2015 arXiv preprint arXiv:1502.03167
[15] Lin T Y, Goyal P, Girshick R, He K and Dollár P 2018 IEEE transactions on pattern analysis and machine intelligence