Research Article

Peng Luo*, Xuekai Hu, Yuhao Zhao, Yi Jiang, Fanglin Lu, Jifeng Liang, and Liang Xu

Smear character recognition method of side-end power meter based on PCA image enhancement

Abstract: Since it is difficult for manual recording to track the rapid change of indication of the power meter, the power meter images are collected by the camera and automatically recognized and recorded to effectively overcome the disadvantages of manual recording. However, the complex scene lighting environment and smearing character shadows make it difficult to transfer captured images directly to convolutional neural networks for character recognition. A smear character recognition method of side-end power meter under complex lighting conditions is proposed in this article. First, the uneven illumination image enhancement algorithm is studied. Through the estimation of the illumination component of the image, the fusion weight is calculated by the principal component analysis for multiscale fusion, and the up-sampling and down-sampling are adopted to reduce the calculation of the algorithm and achieve the rapid image enhancement. A convolution neural network framework based on deep learning is proposed to realize the segmentation of smear characters, and the final segmented individual characters are fed into a network to identify the meter readings. The experimental results show that the proposed smear character recognition method has fast recognition speed, and the recognition rate of samples with the smear character and complex illumination reaches 99.8%, which meets the requirements of power meter character recognition and is better than other algorithms.

Keywords: complex illumination, deep learning, power meter, smear character

1 Introduction

With the expansion of power consumption and the increase in the number of substations, more and more measuring instruments need to be monitored [1]. Accurate meter reading of measuring instruments is crucial to the operation of the substation [2]. In addition, the electric power inspection basically adopts the inspection robot to shoot and extract the data of the meter through the camera. However, there is a certain deviation in the recognition of the electric power meter under the complex illumination of the camera [3]. The character recognition accuracy of electric power meters needs to be improved under the complex illumination environment of the substations [4,5].

Currently, most of the instruments at the stations use LCD screens, but because the LCD screen has reflective characteristics, it is easy to cause uneven illumination of the collected images [6]. Moreover, the long black-and-white response time of some LCD screens causes serious shadows when the characters are quickly refreshed, that is, when one character does not completely disappear, another character has appeared, resulting in the superposition of two characters with different brightness in the collected image. In addition, the scene lighting environment is different, and sunlight is different indoor and outdoor, which makes the collected images vary greatly. Therefore, how to accurately identify instrument characters with serious shadow under the conditions of uneven illumination is a challenging problem [7].

Character recognition using neural networks has been a research focus in recent years. The algorithms [8,9] detect single characters before using the DCNN model to identify them. This approach often requires the implementation of an efficient character detector capable of precisely detecting and cutting each character from the original...
word picture. Other methods [10] treat scene character recognition as an image classification issue, requiring each English word to be assigned to a category label (90k words). Traditional character recognition approaches that are not based on neural networks have also contributed to new concepts and representations. Almazán et al. [11] and Rodriguez-Serrano et al. [12] proposed embedding word images and text strings in public-oriented quantum space and turning word recognition into a retrieval issue. Lee et al. [13] suggested a method for identifying feature pools that can learn the character information subregion. For character recognition, Bai et al. [14] and Gordo [15] used mid-level features. By exchanging features across two tasks, Yao et al. [16] unified character recognition and detection.

This article aims to design a fast image recognition method for smear characters under different illumination conditions, including image enhancement and trailing shadow character image segmentation to improve the accuracy of automatic recognition of power meter reading.

This article is divided into four sections. Section 1 presents the introduction. The status and development of power instrument character recognition are comprehensively expounded. In Section 2, a method for smear character recognition of edge power meter under complex illumination interference is proposed. First, the overall process of the algorithm is introduced. Then, the image enhancement algorithm for uneven illumination is studied to realize the rapid enhancement of instrument images. Finally, a convolutional neural network framework based on deep learning is designed to realize the perfect segmentation of instrument shadow characters. Section 3 presents the experimental results and analysis, including data collection and calibration, binary segmentation experiments, and character recognition experiments. Section 4 presents the conclusion, which summarizes the research content.

2 Method

2.1 Overall process

To achieve accurate and stable recognition of smear characters of power meters under complex lighting, improvements are mainly made in three aspects: image enhancement, binarization and character segmentation, and character recognition [17]. The overall process is shown in Figure 1. Among them, the fully convolutional network (FCN) [18] and the VGG [19] instrument character segmentation model were trained in advance, which are not marked in the flowchart. The major steps are as follows:

1. Camera images acquisition;
2. The fast adaptive image enhancement method based on PCA [20] fusion is used to correct uneven illumination, and then the binarization is carried out based on FCN and VGG instrument character segmentation method;
3. Remove small connected domains and noise by opening and closing operations;
4. Tilt correction;
5. Segmentation of single character;
6. Single characters are normalized and input to smear character recognition model for recognition.

2.2 Fast adaptive image enhancement based on PCA fusion

The original RGB image is converted into HSV color space, and the V component is extracted and downsampled. The illumination component of the image is
extracted by multiscale Gaussian filtering and up-sampled to the size before down-sampling. Then, the illumination component is refined by guided filtering, and the illumination component is adaptively adjusted by different correction coefficients to obtain two images. The weights of the two images are calculated by PCA and fused to enhance the V component. Finally, the image is converted from HSV space to RGB space to obtain the final enhanced image. The specific process is shown in Figure 2.

The calculation amount of multiscale Gaussian filtering increases with the increase of the input image size [21], which affects the real-time performance of the algorithm. Down-sampling and up-sampling can well reduce the calculation amount, and guided filtering can refine the up-sampling results and retain the local characteristics of the illumination component. To avoid the small down-sampling size and excessive loss of image details, the linear interpolation method is used for down-sampling, and the down-sampling size is a quarter of the input size.

To balance the global and local characteristics of the extracted illumination components, the multiscale two-dimensional discrete Gaussian function is used to extract the illumination components in the scene, and each component is weighted. Finally, the estimated illumination components are obtained. The calculation is as follows:

$$I_{r_{mg}}(x, y) = \sum_{i=1}^{N} \theta_i \cdot I_{r_{mg}}(x, y),$$

where $I_{r_{mg}}(x, y)$ denotes the illumination component extracted by the ith scale Gaussian function, $\theta_i$ represents the corresponding weight coefficient, and $N$ represents the number of scales. To balance the accuracy and calculation amount of illumination component extraction, $N$ is set to 3, $\theta_i$ is set to 1/3, and the parameters $c$ of the three scales of Gaussian filtering are 4, 20, and 63, respectively.

To reduce the amount of fusion computation, a simple source image weighted fusion is adopted without multiscale decomposition to avoid the bias of fusion results to a certain source image. PCA is used to determine the weight coefficient of each source image as the key to image fusion. The specific process and result are shown in Figure 3, and it can be seen that the image process by fast adaptive image enhancement algorithm is significantly improved in terms of color, brightness, and detail.

### 2.3 FCN- and VGG-based smear character segmentation

Combining FCN and VGG16 network, the input is the original RGB image and the output is a binary map without the smear. The network consists of two parts: feature extraction and image recombination. The feature extraction is a down-sampling composed of five convolutional blocks, each convolutional block includes a convolutional layer and a maximum pooling layer, and the Relu activation function is chosen behind the convolutional layer to alleviate overfitting. The number of convolutional kernels in the convolutional layer is 256, the size is 3 × 3, the step size is 1, and the pad is 1. The pooling layer has a pooling window size of 2 × 2 and a step size of 2.

The image reconstruction is performed by five deconvolution blocks to reduce the extracted information to images. The deconvolution block includes a deconvolution layer and a connection layer, where the convolution kernel
of the deconvolution layer is $4 \times 4$ with a step size of 2. The overall structure of the network is shown in Figure 4, and the mean square error (MSE) is used as the loss function, which is defined as follows:

$$\text{MSE}(\mathbf{y}, \mathbf{y}') = \frac{1}{n} \sum_{i=1}^{n} (y_i - y'_i)^2,$$  \hspace{1cm} (2)

where $n$ is the number of samples, $y_i$ is the label, and $y'_i$ is the network prediction result.

### 3 Experimental results and analysis

#### 3.1 Collection and calibration of data sets

The samples were collected in the instrument working environment. The light sources were placed in the left, right, and median positions of the instrument and combined with a 30 degree angle deflection between the camera and the instrument. A total of 7,600 color images were collected.

Label production is divided into two stages: preparation and production. In the preparation stage, the accurate position of the characters in the instrument image is mainly counted. The image with good illumination is selected, and the approximate region of the characters is manually calibrated. The minimum circumscribed rectangle is determined, and the angle $\theta$ between the long edge of the minimum circumscribed rectangle and the horizontal $x$-axis is calculated. The binary image is rotated and corrected to make all the characters on the same horizontal line. The center $P_1(x_1, y_1)$, $P_2(x_2, y_2)$, and $P_3(x_3, y_3)$ of the three calibration circles are determined by circle detection as the three points corresponding to the affine transformation in the label production stage. The position of each character is recorded by the vertical projection method, and the characters with complete segmentation
Figure 5: Label preparation phase. (a) Original image, (b) binary graph, (c) calculate rotation angle, and (d) location of each character.

Figure 6: The label production phase. (a) Original image, (b) binary graph, (c) affine transformation result, and (d) label after character replacement.
and no shadow are selected as the standard single character, as shown in Figure 5.

In the label production stage, Wolf algorithm is used to binarize the instrument image and detect the center of three calibration circles $P_1(x'_1, y'_1)$, $P_2(x'_2, y'_2)$, and $P_3(x'_3, y'_3)$, combined with the center $P_1(x_1, y_1)$, $P_2(x_2, y_2)$, and $P_3(x_3, y_3)$ detected in the preparation stage, the affine transformation matrix $A$ is calculated as shown in Eq. (3). Affine transformation of instrument image is obtained by affine transformation matrix, let $P_1$, $P_2$, and $P_3$ coincide with $P'_1$, $P'_2$, and $P'_3$, respectively. On each character position recorded in the preparation phase, the shadow character is replaced with its corresponding standard single character to complete the label production, as shown in Figure 6.

$$
\begin{bmatrix}
x_1 & x_2 & x_3 \\
y_1 & y_2 & y_3 \\
1 & 1 & 1
\end{bmatrix}
= A \cdot
\begin{bmatrix}
x'_1 & x'_2 & x'_3 \\
y'_1 & y'_2 & y'_3 \\
1 & 1 & 1
\end{bmatrix}
= \begin{bmatrix}
a & b & c \\
d & e & f \\
0 & 0 & 1
\end{bmatrix} \cdot
\begin{bmatrix}
x'_1 & x'_2 & x'_3 \\
y'_1 & y'_2 & y'_3 \\
1 & 1 & 1
\end{bmatrix}.
$$

To reduce the training time, the sample only represents the number of parts by the interceptor, and each sample is randomly rotated and shifted to expand three samples. Through this method, the total number of samples is 30,400, and 80% of the samples are randomly selected for training, and the remaining samples are used as testing samples.

### 3.2 Binary segmentation experiment

The experimental environment is Intel (R) Xeon (R) 2.4 GHz CPU E5-2640, GPU NVIDIA RTX 2080Ti, Python3.6 programming in TensorFlow framework. Without other pretraining models, batch size is 16, initial learning rate is $1 \times 10^{-4}$, period is 100. Compared with local threshold binarization algorithms such as Bernsen [22] black, Sauvola [23] Bradley [24], Wellner [25], and Wolf et al. [26], IoU and SNR are used as evaluation indexes.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Running time (s)</th>
<th>IoU</th>
<th>SNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bernsen</td>
<td>0.85</td>
<td>0.33</td>
<td>8.93</td>
</tr>
<tr>
<td>Niblack</td>
<td>1.06</td>
<td>0.51</td>
<td>6.74</td>
</tr>
<tr>
<td>Sauvola</td>
<td>0.28</td>
<td>0.66</td>
<td>10.17</td>
</tr>
<tr>
<td>Bradley</td>
<td>1.87</td>
<td>0.71</td>
<td>10.11</td>
</tr>
<tr>
<td>Wellner</td>
<td>0.07</td>
<td>0.70</td>
<td>10.97</td>
</tr>
<tr>
<td>Wolf</td>
<td>0.81</td>
<td>0.69</td>
<td>10.24</td>
</tr>
<tr>
<td>Ours</td>
<td>0.10</td>
<td>0.95</td>
<td>20.53</td>
</tr>
</tbody>
</table>

Figure 7: Processing results of different algorithms. (a) Original image, (b) ideal binary label, (c) Bernsen algorithm, (d) Niblack algorithm, (e) Sauvola algorithm, (f) Bradley algorithm, (g) Wellner algorithm, (h) Wolf algorithm, and (i) the method of this article.
cannot segment strong light region and smear character. Wolf algorithm overcomes the influence of uneven illumination to some extent, but cannot smear character. The network designed in this article has a clear and no smear character on the binarization effect of uneven illumination and shadow images. The IoU and SNR values are higher than those of other methods. It is suitable for images with different light intensities and rotations. It has a strong generalization ability and fast processing speed, only slightly higher than Welling.

3.3 Character recognition experiment

Smear characters are segmented for character recognition [27]. Character segmentation includes tilt correction, affine transformation, vertical projection, and horizontal projection. The specific effect is shown in Figure 8. Because the horizontal projection intervals of numbers, decimals and “VAC” characters are different, and “VAC” characters remain unchanged, decimals can be judged directly by the horizontal projection interval, so only numbers are reserved in segmentation. The single character is segmented by this algorithm, as shown in Figure 9, where the smear character is basically eliminated.

The individual characters after sample segmentation are categorized to make numerical labels, which are input into the CRNN model for training and testing, respectively.

To verify the effectiveness of the algorithms in this article, HOG + SVM [28], KNN [29], and CRNN [30] algorithms are used for character recognition. The CRNN algorithm parameters are as follows: input image size is $32 \times 32$, batch size is 64, training cycle is 3000, and learning rate is 0.0001. SVM uses linear kernel function, and HOG feature unit size is $3 \times 3$.

The experimental results are presented in Table 2. Compared with the direct recognition of characters, the recognition rate of all algorithms is higher after the processing of the proposed character segmentation algorithm, and the recognition rate of CRNN algorithm is improved by 21.71%, indicating that the proposed binarization algorithm provides a better basis for character recognition.

<table>
<thead>
<tr>
<th>Method</th>
<th>Training time (s)</th>
<th>Single-sample test time (ms)</th>
<th>Test recognition rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRNN (original)</td>
<td>—</td>
<td>—</td>
<td>78.09</td>
</tr>
<tr>
<td>HOG + SVM</td>
<td>32.18</td>
<td>5.7</td>
<td>98.95</td>
</tr>
<tr>
<td>KNN</td>
<td>—</td>
<td>12.6</td>
<td>98.75</td>
</tr>
<tr>
<td>LeNet-5</td>
<td>98.54</td>
<td>1.4</td>
<td>96.45</td>
</tr>
<tr>
<td>CRNN</td>
<td>90.16</td>
<td>2.16</td>
<td>99.80</td>
</tr>
</tbody>
</table>
4 Conclusions

In this article, we design a fast recognition method for power meter characters by combining adaptive image enhancement, smear character segmentation based on FCN and VGG for the problem of difficult recognition of meter smear characters under complex illumination. The experimental results show that the proposed smear character segmentation network can achieve adaptive binary segmentation of meter images, effectively remove the effects of digital trailing, offset or rotation under severe illumination uneven conditions, with high recognition rate and fast recognition speed of the whole method, meeting the requirements of fast meter character recognition.

Funding information: The work was supported by State Grid Science and Technology Project (kj2020-055): Research on Development Aligned Access Gateway and Networking Technology of Heterogeneous Sensor Network Based on Cloud-edge Computing Collaboration.

Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Conflict of interest: The authors state no conflict of interest.

References


