Research on Image Recognition Algorithm Technology for Power Line Business Audit

Qianqian Mao*, Zheqiong Pan, Dongmei Si, Xufeng Kong
State Grid Ningbo Power Supply Company, Ningbo 315000, China

*Corresponding author e-mail: magic316@qq.com

Abstract. The development of electric power business is not only limited to the on-site office of the business hall, but also can be carried out on the mobile Internet. At present, the power supply company's review of business such as name change, transfer, reclassification-peak shaving, etc., because customers use different methods to handle business, Resulting in a wide variety of certificate pictures, which caused great trouble to the staff. This paper studies an image recognition algorithm that can be applied to power business audits, can accurately identify various types of certificate images, and automates the overall business process. Implementation provides preparation. Users can use the mobile phone APP and computer network to handle all the business processes of electricity business online, know the main items of each link, the materials that should be provided, and the various procedures that should be handled, etc., electricity customers do not have to go out, just log on to the Internet All business can be handled well, which greatly facilitates the efficiency of users in handling various power businesses and saves time very convenient.

1. Introduction
With the transformation of the power business, the company's daily response center received a large number of business work orders. The main piece of business is the name change, transfer, reorganization-peak-shaving audit, and usually receive the work order to be reviewed every hour, . The task of the staff becomes more and more heavy. Once the audit is overtime or wrong, there will be a risk of receiving complaints. Taking the State Grid Ningbo Yinzhou District Power Supply Company as an example, according to data statistics, the company processed more than 80,000 transactions in 2017, and the change business accounted for 91.23%. The processing time of a single change business is about 8 to 10 minutes, and 4 people are needed every day. Responsible for the approval of this business, and the results are not good, the efficiency is low, and it is easy to cause customer complaints [1]. At present, the renaming, transfer, and reclassification-peaking valley table audit module in the State Grid and State System has a complicated operation process, which mainly has the following shortcomings: 1) Due to the large amount of loading content, the staff often needs some Necessary waiting time; 2) The website system is affected by the browser and the system resolution. The page display is incomplete, and the operation efficiency is affected; 3) There are many types of uploaded pictures, and the sources include self-service device uploads in the business hall and user phone photo uploads. The clarity is also very different; 4) The operation steps are cumbersome, the audit content is more, the human vision frequently stares at the numbers in the picture, it is easy to
produce fatigue, and the error rate increases significantly. These shortcomings greatly aggravated the workload of the staff and the processing time of a single transaction [2].

Therefore, it is necessary to study an image recognition algorithm that can be applied to power business audits, and can prepare to recognize various types of document images, laying a solid foundation for the automation of business processes.

2. Analysis on the difficulty of online recognition technology

Considering that the reviewer needs to verify whether the ID information of a single business is correct, and the ID in most pictures only occupies a small part of the picture, as shown in the above picture, it needs to be enlarged and viewed, and manual identification is extremely laborious. As shown in Figure 1, the ID card only occupies a very small part of the entire shooting area, causing the staff to zoom in on the photo to see clearly, which greatly reduces the working speed [3].

![Figure 1. Customer information photo area](image)

Secondly, considering that the reviewer still needs to verify whether the certificate number and address on the real estate certificate, application form, room ticket, purchase contract and other materials are consistent with the system, but the photocopy of the real estate certificate, application form, room ticket, purchase contract and other photos are uploaded Later, the size is inconsistent, the angle is inconsistent, the contrast is low, and the color scale is not obvious, which is not conducive to human eye recognition. And densely packed words are more prone to fatigue and errors.
In general, the main difficulties are: 1) The sources of images are diverse, and the differences in images are very large, which poses a huge challenge to the accuracy of recognition. 2) The proportion of the area to be identified in the image is inconsistent, which poses a challenge to the identification area. Secondly, the image is inverted, twisted, tilted, glare, highlight, etc. The difference is too large, which poses a huge challenge to the accuracy of the classifier algorithm identified. 3) The website system is complex and the amount of data is large. Separating and extracting key information consumes manpower. In order to make the simplified audit module stable and reliable, it takes time to debug.

3. Design of image recognition algorithm

The text content recognition of a picture is to recognize the text information contained in the picture from an ordinary picture [4]. Due to the limitations of the shooting environment, external light, shooting angle, picture compression algorithm, shooting equipment and other factors, the quality of the picture, Sharpness and resolution are very different. In order to make the recognition text and the background color show a large difference, the R component is selected to grayscale the color image, and then the image is binarized by acquiring the global threshold and the local threshold of the image.

1) Binarize the image: First use the Otsu algorithm to obtain the global threshold T of the entire image, then use the Beresen method to calculate the gray scale mean Tbn of the current window, and finally use the maximum grayscale value and minimum gray scale value of the entire image to calculate A correction factor b, as shown in the following formula:

Here, C is an empirical coefficient, usually taken as 0.12, g2 is the maximum value of grayscale in the image, and g1 is the minimum value of grayscale.

2) Black and white inverse color processing, the black and white inverse color processing is performed on the image obtained above, that is, black background, white text, and then use the findContours technology to detect the outer contour of the white pixel block in the binary image, and extract to meet the aspect ratio and Area required profile.

3) Number segmentation: Observation found that the cropped character image was obtained. In this image, it is obvious that the character and background color are highly differentiated, so the color is reversed to segment the character.

4) Feature extraction: Extract feature vectors of characters, that is, extract gradient distribution features + gray distribution features + horizontal projection histogram + vertical projection histogram, and finally each character can correspond to a 1 * 72 feature vector.

5) Neural network training: the training pictures used are all obtained by segmentation from multiple pictures, and then feature extraction, and the training image is generated by segmenting the image, and then the features are extracted according to the above method to obtain the training matrix.
and label matrix for Model training, mainly through BP multi-layer neural network and deep learning model for character recognition training.

6) Classifier classification: The classifier trained through the neural network framework is used to classify the extracted character feature vectors to achieve character recognition [5].

4. Research on Algorithm Design of Image Recognition System

The criterion of the identification algorithm is the minimum mean square error, that is, the expected value of the square of the difference \( e(n) \) between the ideal signal \( d(n) \) and the filter output \( y(n) \) is the smallest, and the weight coefficient \( w_i(n) \) is modified according to this criterion. The resulting algorithm is called the minimum mean square Algorithm (identification). Most research on adaptive filters is based on the identification algorithm proposed by Windrow. This is because the design and implementation of the identification algorithm are relatively simple and are very suitable for many applications. Let the tap coefficient of the order FIR filter be \( w_i(n) \) and the input and output of the filter be \( x(n) \) and \( y(n) \) respectively, then the FIR transversal filter equation can be expressed as:

\[
y(n) = \sum_{i=1}^{N} w_i(n)x(n-i)
\]  

(1)

Let \( d(n) \) stand for "expected response" and define the error signal:

\[
e(n) = d(n) - y(n) = d(n) - \sum_{i=1}^{N} w_i(n)x(n-i)
\]  

(2)

Using vector form to express weight coefficients and inputs \( w \) and \( X(n) \), the error signal \( e(n) \) can be written

\[
e(n) = d(n) - W^T X(n) = d(n) - X^T(n)W
\]  

(3)

The square of the error is:

\[
e^2(n) = d^2(n) - 2d(n)X^T(n)W + W^T X(n)X^T(n)W
\]  

(4)

After taking the mathematical expectations on both sides of the above formula, the mean square error is obtained:

\[
E\{e^2(n)\} = E\{d^2(n)\} - 2E\{d(n)X^T(n)\}W + W^T E\{X(n)X^T(n)\}W
\]  

(5)

Define the cross-correlation function vector:

\[
R_{xd}^T = \{E d(n)X^T(n)\}
\]  

(6)

And autocorrelation function matrix:

\[
R_{xx} = E\{X(n)X^T(n)\}
\]  

(7)
So, the mean square error can be expressed as:

$$E\{e^2(n)\} = E\{d^2(n)\} - 2R_{xd}^T W + W^T R_{xx} W$$  \hspace{1cm} (8)

Derivate the weight coefficient $W$ by equation (8) to obtain the gradient of the mean square error function:

$$\nabla(n) = \nabla E\{e^2(n)\} = \left[ \frac{\partial E\{e^2(n)\}}{\partial W_1^T}, ..., \frac{\partial E\{e^2(n)\}}{\partial W_N^T} \right]^T = -2R_{xd} + 2R_{xx} W$$  \hspace{1cm} (9)

Let $\nabla(n) = 0$, you can find the best weight coefficient vector:

$$W_{opt} = R_{xx}^{-1} R_{xd}$$  \hspace{1cm} (10)

Substituting $W_{opt}$ into equation (8), the minimum mean square error is obtained:

$$E\{e^2(n)\}_{min} = E\{d^2(n)\} - R_{xd}^T W_{opt}$$  \hspace{1cm} (11)

The basis of this algorithm is the steepest descent method in the optimization method. According to this steepest descent method, the "next moment" weight coefficient vector $W(n+1)$ should be equal to the "current moment" weight coefficient vector $W(n)$ plus a negative mean square error gradient $-\nabla(n)$ proportional term, ie

$$W(n+1) = W(n) - \mu \nabla(n)$$  \hspace{1cm} (12)

$\mu$ in the formula is a constant that controls the convergence rate and stability, and is called the convergence factor. It is not difficult to see that the identification algorithm has two keys: the calculation of gradient $\nabla(n)$ and the choice of convergence factor $\mu$. It is very difficult to calculate the gradient $\nabla(n)$ accurately. A rough but very effective approximate method for calculating $\nabla(n)$ is: directly take $e^2(n)$ as the estimated value of the mean square error $E\{e^2(n)\}$, that is

$$\hat{\nabla} = \nabla [e^2(n)] = 2e(n)\nabla [e(n)]$$  \hspace{1cm} (13)

$\nabla [e(n)]$ in the formula is:

$$\nabla [e(n)] = \nabla [d(n) - W^T(n)X(n)] = -X(n)$$  \hspace{1cm} (14)

Substituting (14) into equation (13), the gradient estimate is obtained:

$$\hat{\nabla}(n) = -2e(n)X(n)$$  \hspace{1cm} (15)

Therefore, the Windrow-Hoff identification algorithm is finally:
Design of image recognition algorithm

The text content recognition of a picture is to recognize the text information contained in the picture from an ordinary picture. Due to the limitations of the shooting environment, external light, shooting angle, picture compression algorithm, shooting equipment and other factors, the quality of the picture, Sharpness and resolution are very different. In order to make the recognition text and the background color show a large difference, the R component is selected to grayscale the color image, and then the image is binarized by acquiring the global threshold and the local threshold of the image [6].

1) Binarize the image: First use the Otsu algorithm to obtain the global threshold $T$ of the entire image, then use the Beresen method to calculate the grayscale mean $T_{bn}$ of the current window, and finally use the maximum grayscale value and minimum grayscale value of the entire image to calculate $A$ correction factor $b$, as shown in the following formula:

Here, $C$ is an empirical coefficient, usually taken as 0.12, $g_2$ is the maximum value of grayscale in the image, and $g_1$ is the minimum value of grayscale.

2) Black and white inverse color processing, the black and white inverse color processing is performed on the image obtained above, that is, black background, white text, and then use the findContours technology to detect the outer contour of the white pixel block in the binary image, and extract to meet the aspect ratio and Area required profile.

3) Number segmentation: Observation found that the cropped character image was obtained. In this image, it is obvious that the character and background color are highly differentiated, so the color is reversed to segment the character.

4) Feature extraction: Extract feature vectors of characters, that is, extract gradient distribution features + gray distribution features + horizontal projection histogram + vertical projection histogram, and finally each character can correspond to a $1 \times 72$ feature vector.

5) Neural network training: the training pictures used are all obtained by segmentation from multiple pictures, and then feature extraction, and the training image is generated by segmenting the image, and then the features are extracted according to the above method to obtain the training matrix and label matrix for Model training, mainly through BP multi-layer neural network and deep learning model for character recognition training.

6) Classifier classification: The classifier trained through the neural network framework is used to classify the extracted character feature vectors to achieve character recognition. [4].

The designed adaptive filter adjusts the filter parameter $W_i$ to make the observation signal $\hat{s}(n)$ based on the past observation samples most approximate the original signal $s(n)$ in a certain sense. At this time, on the one hand, the recovery error:

$$\eta(n) = s(n) - \hat{W}^T X(n)$$  \hspace{1cm} (17)

On the other hand, $\hat{W}^T X(n)$ can be regarded as the prediction of $x(n)$. Therefore, the prediction error can be defined:

$$e(n) = x(n) - \hat{W}^T X(n)$$  \hspace{1cm} (18)

The purpose of designing the adaptive filter is naturally to minimize the recovery error $\eta(n)$. But because the real signal $s(n)$ is unknown, $\eta(n)$ is unobservable or incalculable. On the contrary, the prediction error $e(n)$ is observable, and its relationship with the recovery error is:

$$W(n+1) = W(n) + 2 \mu e(n)X(n)$$  \hspace{1cm} (16)
The noise sequence $\eta(n)$ is independent, so the minimization of the unobservable recovery error $\eta(n)$ is equivalent to the minimization of the observable prediction error $e(n)$. Specifically, consider the following formula to minimize.

$$e(n, W) = \sum_{i=1}^{n} \lambda^{n-i} |e(i)|^2$$

In the formula, $\lambda$ is the forgetting factor, usually $0 \leq \lambda \leq 1$. By

$$\frac{\partial e(n, W)}{\partial W} = \sum_{i=1}^{n} \lambda^{n-i} [x(i) - W^T X(i)]^T = -2 \sum_{i=1}^{n} \lambda^{n-i} [x(i) - W^T X(i)] X(i) = 0$$

The equivalent relationship can be obtained:

$$\sum_{i=1}^{n} \lambda^{n-i} X(i) X^T(i) W = \sum_{i=1}^{n} \lambda^{n-i} x(i) X(i)$$

Make:

$$R(n) = \sum_{i=1}^{n} \lambda^{n-i} X(i) X^T(i)$$

$$U(n) = \sum_{i=1}^{n} \lambda^{n-i} x(i) X(i)$$

Then formula (24) can be abbreviated as:

$$R(n) W(n) = U(n)$$

Assuming that $R(n)$ is non-singular, then:

$$W(n) = R^{-1}(n) U(n)$$

This is the formula for the filter parameters of the filter. The reason for this is $W(n)$ changes with time.

5. Advantages of online image review system
The system adds online declaration and status inquiry. Applicants can log in to the electric company's online service platform to apply without leaving the house. On the online service platform, the service guide and the handling process are clear at a glance, and applicants can check the progress of the handling at any time through the "business status" column of the individual declaration on the online service platform. The approval service has increased service outlets, making it easier for power users
to work nearby. Reduce the number of submissions in the window. It is very convenient to complete the online pre-review on the next day [7].

6. Conclusion
In general, the main difficulties are: 1) The sources of images are diverse, and the differences in images are very large, which poses a huge challenge to the accuracy of recognition. 2) The proportion of the area to be identified in the image is inconsistent, which poses a challenge to the identification area. Secondly, the image is inverted, twisted, tilted, glare, highlight, etc. The difference is too large, which poses a huge challenge to the accuracy of the classifier algorithm identified. 3) The website system is complex and the amount of data is large. Separating and extracting key information consumes manpower. In order to make the simplified audit module stable and reliable, it takes time to debug.

Acknowledgments
This work is supported by the State Grid Zhejiang Electric Power Co., Ltd. Science and Technology Project “Three Types and One Type business hall self-service terminal experience interactive technology”

References
[1] Eduardo, N. D. S. Rigoberto, E. M. M., & Marco, J. D. S. Advanced image processing of wire-mesh sensor data for two-phase flow investigation. Latin America Transactions, IEEE (Revista IEEE America Latina), 13(7) (2015) 2269-2277.
[2] Huang, Y. F. Jen, Y. J., Chen, L. C., Chen, K. H., & Chattopadhyay, S. Design for approaching cicada-wing reflectance in low- and high-index biomimetic nanostructures. Acs Nano, 9(1) (2015) 301-311.
[3] Dufour, S. & De Koninck, Y. Optrodes for combined optogenetics and electrophysiology in live animals. Neurophotonics, 2(3) (2015) 031205.
[4] Zhang, H. Corkhill, C. L., Heath, P. G., Hand, R. J., Stennett, M. C., & Hyatt, N. C. Effect of zn- and ca-oxides on the structure and chemical durability of simulant alkali borosilicate glasses for immobilisation of uk high level wastes. Journal of Nuclear Materials, 462(1) (2015) 321-328.
[5] Sasikumar, Y. Kumar, A. M., Gasem, Z. M., & Ebenso, E. E. Hybrid nanocomposite from aniline and ceo2 nanoparticles: surface protective performance on mild steel in acidic environment. Applied Surface Science, 330(5) (2015) 207-215.
[6] Maris, E. Fries, P., & Ede, F. V. Diverse phase relations among neuronal rhythms and their potential function. Trends in Neuroences, 39(2) (2016) 86-99.
[7] Zhang, Y. Yang, S., Su, X., Shi, E., & Zhang, H. Automatic reading of domestic electric meter: an intelligent device based on image processing and ZigBee/ethernet communication. Journal of Real-Time Image Processing, 12(1) (2016) 133-143.