Research on Automatic Reading Recognition of Wheel Mechanical Water Meter Based on Improved U-Net and VGG16

LIUKUI CHEN[1], WEIYE SUN[1,*], LI TANG[1], HAIYANG JIANG[1], ZUOJIN LI[2]
[1]School of Intelligent Technology and Engineering, Chongqing University of Science & Technology
[2]Research Department, Chongqing University of Science & Technology
Chongqing, 401331, CHINA

Abstract: This paper proposes a deep learning scheme to automatically carry out reading recognition in wheel mechanical water meter images. Aiming at these early water meters deployed in old residential compounds, this method based on deep neural networks employs a coarse-to-fine reading recognition strategy, firstly, by means of an improved U-Net to locate the reading area of the dial on a large scale, and then the single character segmentation is performed according to the structural features of the dial, and finally carry out reading recognition through the improved VGG16. Experimental result shows that the proposed scheme can reduce the information interference of non-interested regions, effectively extract and identify reading results, and the recognition accuracy of 95.6% is achieved on the dataset in this paper. This paper proposes a new solution for the current situation of manual meter reading, which is time-consuming and labor-intensive, errors occur frequently; and the transformation cost is high and difficult to implement. It provides technical support for automatic reading recognition of wheel mechanical water meters.

Key-Words: Wheel Mechanical Water Meter, Reading Recognition, U-Net, VGG16
Received: September 23, 2021. Revised: June 21, 2022. Accepted: August 13, 2022. Published: September 1, 2022.

1 Introduction

With the increasing demand of water companies for faster collection of water consumption, many scholars at home and abroad have studied the automatic reading recognition of water meters [1-4]. According to the preliminary survey results, there are 7394 old residential compounds with backward public facilities built before 2000 in Chongqing. Chongqing Tap Water Company owns nearly 2 million users of water, and mechanical water meters account for nearly 60% of the installations of these users [5]. The "one household one meter" policy implemented in succession throughout the country represents the increasing attention of the state and society to saving water resources, but at the same time, it greatly highlights the cumbersome and inconvenient nature of the traditional meter reading method [6].

At present, the vast majority of water consumption audit work still adopts the method of naked eye identification and manual transcription by meter readers from door to door [7]. The disadvantages of this traditional manual meter reading method are becoming increasingly prominent. It is not only time-consuming and laborious, this also happens frequently that meter readers cannot enter narrow and rugged areas, which makes it difficult to collect data. Moreover, due to the long-time running, tiredness, dazzle, heavy workload of meter readers; the influence of light, silt, and other factors in the poor working environment; transcribing with the naked eye is very easy to lead to errors [8]. At the same time, with the intensive development of modern high-rise building construction, it is more and more difficult to check the reading of water meters only by manpower.

In order to keep up with the development of the times, it is necessary to get rid of the traditional meter reading method and carry out technological innovation. The traditional image classification and detection algorithms that have emerged in recent years are often difficult to deal with the contradiction between anti-noise performance and detection accuracy [9]. Under such circumstances, it has become a trend of the times to further reduce the consumption of human and material resources by means of deep learning.

2 Related Research

At present, the common application of automatic meter reading systems of digital display water meters mainly includes the following two forms:

The first is the wiring meter reading method based on the sensor. By installing a sensor in the water meter, the data collectors transmit the reading through the prearranged line in the form of electrical signal [10,11], but the work of laying the line from door to door is very heavy [12]. At present, there are also ways to transmit data and power at the same time based on existing wires. For example, Ben-shimol introduced an effective method of automatic meter reading from smart meters by using a power line communication network, and two intelligent polling algorithms based on application layer methods are used to ensure the effective transmission of data in the network [13]. This method has a simple structure, mature technology, and a relatively low price. However, it still faces the problem that the line is easy to be damaged in a humid and messy complex environment for a long time, and the maintenance work is not easy.

The second application mode is based on the smart card water meter. Users use water by purchasing a certain amount of water card in advance [14]. This method is simple to use and convenient to replace and install. Its main shortcomings are: on the one hand, due to the lack of information transmission devices to connect users and companies, timely water supply statistics and dispatching will be difficult to achieve [2]; On the other hand, the economic losses of users or water supply companies caused by the failure and damage of water cards or malicious modification of users' water cards by third parties also occur from time to time.

Under the background that the traditional meter reading methods are gradually difficult to meet people's requirements for accuracy and efficiency, the use of machine vision and deep learning to improve and innovate the traditional methods has gradually become the research focus of relevant scholars at home and abroad.

At present, the automatic recognition technology of water meter reading based on image processing has been theoretically studied, but it has not been applied on a large scale in the market. For example, Jing-wei Sun combined the color characteristics of the pointer, used the global threshold and local threshold to segment the water meter components, and then used the shape features to complete the reading location [15]. Shuai Shang extracted the water meter frame through the vertical projection method and region-based segmentation method; and then matched the template with the segmented image by using the template matching method to obtain the segmented single character matching result [16]. Tian-hua Liu transformed the water meter pictures into HSV color model, extracted the H-channel, removed the noise, and obtained the contour by using median filter and canny operator, and then calculated the center coordinates of the pointer by cluster circle fitting algorithm [17]. Ying Chen et al. Proposed an automatic recognition algorithm for water meter characters that can meet real-time requirements. The character image is cut into template size for image thinning, feature extraction, and character recognition, so as to achieve a high recognition rate [18]. Hao-lin Shi screened out most of the non-text regions according to the text region features, then extracted the HOG features of the training samples, trained the samples, and used SVM classifier to accurately locate the candidate regions [19]. Fan Zhang and others classify the character curve by calculating the gradient information of the image, obtain the edge features of the image, and then classify the characters to be detected according to the K-Nearest Neighbor classification algorithm (KNN) for character recognition. The test results show that the recognition rate of the edge gradient feature algorithm is 5.23% higher than that of the template matching algorithm [20]. Chen Yue carries out a series of image processing through OpenCV computer vision library, and the combination of image processing and neural network recognition was used to recognize the reading of water meter pictures [21]. Shuai-cheng Pan and others used a character recognition algorithm based on deep convolution neural network, by improving the classical CNN network structure, they constructed a convolution neural network model which can recognize characters and dial at the same time, and the test effect is good [22].

The methods proposed by the above researchers have achieved good results in dealing with their research objects, but there are still some deficiencies. For example, the common character recognition algorithms such as KNN algorithm have good effect but long running time, and the character recognition algorithm based on SVM has difficulties in solving multi-classification problems. In addition, in view of the "half character" display phenomenon of the wheel mechanical water meter, which is the research object of this paper, due to the structural characteristics of its gear transmission; Aging and blurring of the disk surface and interference of other digital characters; And in the complex environment, due to the random angle, random illumination and many other factors brought by the use of mobile phone shooting by nonprofessionals, the direct application of the above image processing methods will easily lead to the results of wrong feature extraction, wrong reading and so on. Therefore, according to the characteristics of the research object in this paper, we propose a method to segment the region of interest, locate the target character, and then carry out numeral recognition.

3 Holistic Design

Firstly, the water company outsources patrol personnel of the old residential compounds to collect the reading images of water meters in a certain area through mobile phones and other mobile devices. The patrol personnel does not recognize the image, but directly uploads it wirelessly to the server. On the server side, the processing strategy from coarse to fine designed in this paper is used. In the first place, through the positioning algorithm, the reading area in the dial image is extracted and the irrelevant area is removed. In the next place, segment each character in the target area, then recognize the obtained single character to get the reading result. After that, the results will be sent to the billing center. Finally, the water bill will be returned to each user after being summarized and sorted by the billing center. The basic logic block diagram of this remote meter reading system is shown in Fig.1.

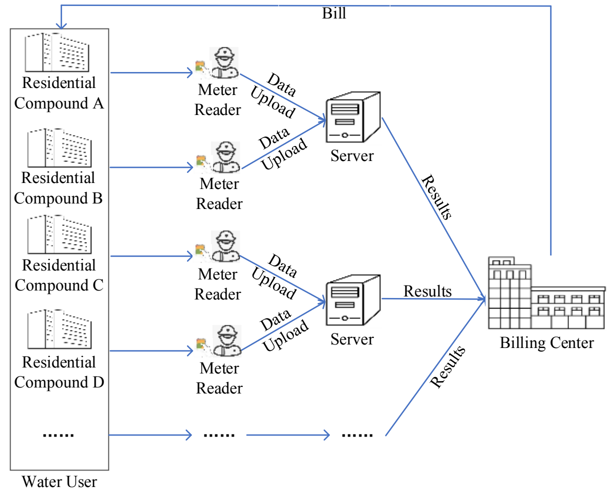

Fig.1 Logic block diagram of remote meter reading system

3.1 Technical route

In this remote meter reading system, the most important link is the automatic recognition of water meter reading. Based on the existing research at home and abroad, this paper further explores the application of automatic recognition technology for water meter reading. In view of the fact that it is difficult for patrol personnel to fix the angle of the captured image, which will cause image distortion. Moreover, there are many printed digital characters similar to the reading characters on the dial. These redundant interference information has not been removed, and the accuracy of direct disk recognition is not high. In order to solve this problem, the " coarse to fine" strategy of accurately locating the reading area at first and then identifying the reading is necessary. Therefore, the main research content of this paper is divided into two aspects: target region segmentation using segmentation network and character recognition using recognition network. Fig.2 shows the technical route of this paper.

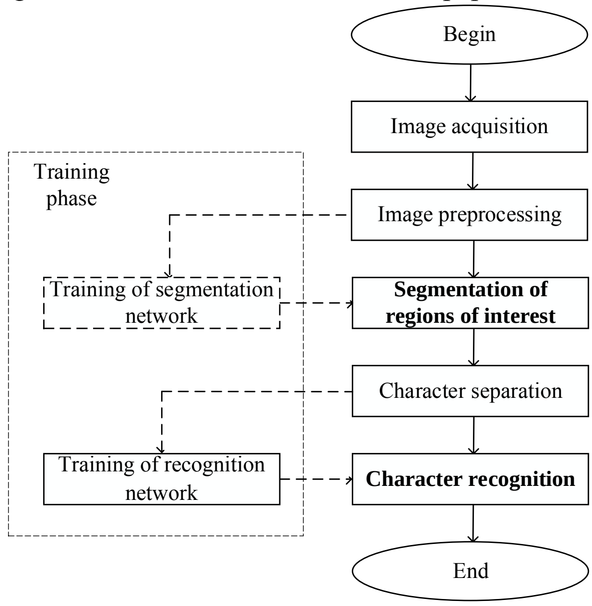

Fig.2 Technical roadmap for automatic recognition of water meter readings

3.2 Improved U-Net segmentation network

For the problem of target area segmentation, firstly, the network structure and parameter optimization of various semantic segmentation networks are studied, the U-Net network among them, which is simple and efficient, especially in learning a small amount of datasets, it can still achieve good recognition results, is determined to extract the binary classification of water meter reading area [23]. Semantic segmentation needs to judge the category of each pixel in the image, mark each pixel in the image as an object category for accurate segmentation, and finally, we get an image in which each pixel has a one-to-one corresponding type, so as to accurately extract the reading area of the dial. The network structure of the U-Net semantic segmentation model used in this paper is shown in Fig.3, In order to further improve the accuracy of the model while avoiding overfitting and degradation [24], the "shortcut connection" idea of ResNet is imported into the model.

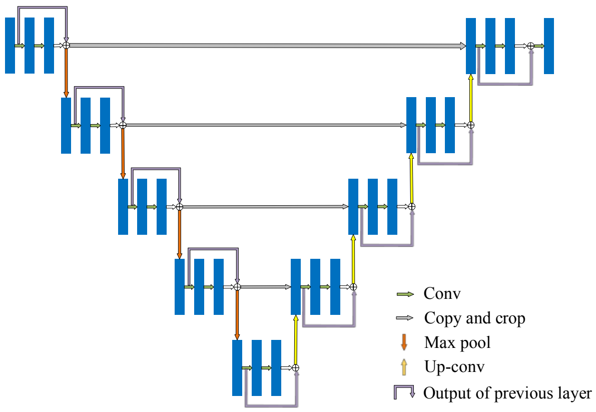

Fig.3 Structure diagram of improved U-Net segmentation network

U-Net network model is an Encoder-Decoder structure. The mechanical water meter dial image is down-sampled by the encoder to extract features, and then the image is restored to the original size by the decoder for pixel-by-pixel classification. The shortcut connection is introduced into the encoder and decoder to form residual blocks [25,26]. The formula of the residual structure is expressed as follows:

$$H(x) = x + F(x) \tag{1}$$

Where x is identity mapping, that is, the output of the previous layer of the network, $F(x)$ is residual mapping, that is, the output of the current layer of the network, and $H(x)$ is the network mapping from the input to the sum. Adding identity mapping to the network through the shortcut connection can ensure that the output of each layer can be in the optimal state. At the same time, it will not increase the parameters and computational complexity of the model.

The continuous convolution and pooling operations in the encoder and decoder are included in the residual blocks to perform feature extraction and step-by-step up-sampling of the target area of the dial respectively. The shallow convolution focuses on the texture features of printed characters in the reading area, and the deep convolution focuses on the essential features. In the meantime, multiscale feature fusion is performed between the encoder and decoder, such fusion connections run through the whole network, so that the final up-sampling feature maps have more shallow semantic information. Therefore, in the segmentation results of the dial, both micro features and macro features such as edges can be obtained, which enhances the segmentation accuracy.

Before network training, it is necessary to preprocess and expand the samples of water meter dataset by means of filtering and rotation transformation, so as to enhance the generalization ability and robustness of segmentation network, thus improve the effect of subsequent character recognition.

3.3 Improved VGG16 identification network

For the problem of dial reading recognition in this paper, in view of the shortcomings of the above recognition algorithms in dealing with the special research objects in this paper, for example, the common threading method is generally applied to the digital display instrument using the 7-segment nixie tube display scheme, but it is not applicable to the wheel type mechanical water meter; as a widely used classical algorithm, template matching algorithm has simple principle but poor flexibility. Due to installation error, looseness, gear wear and other reasons, the mechanical water meters with long service life may have the phenomenon of "half character" display. Obviously, the template matching algorithm cannot deal with such complex situations. The recognition method using convolutional neural network has better accuracy than other traditional algorithms in dealing with different backgrounds, different formats and different types of character recognition. Therefore, this paper explores the character recognition method based on the classical convolutional neural network model VGG16, and realizes the character recognition by extracting the features of the dial image.

As a classic and efficient classification and recognition convolution neural network, VGG16 shows high robustness in various recognition problems. The structure of VGG16 is very simple, and the original network structure is shown in Fig.4, it has 5 feature extraction modules composed of 13 convolution layers and 5 pooling layers, the 3 fully connected layers and the softmax output layer constitute the classification module and output the prediction results of ten-category classification. In the convolutional layer, since the entire network uses 3*3 small convolution kernels, the model has less parameters and better performance than using larger convolution kernels. However, in the fully connected layer, due to the large amount of parameters, the model has problems such as large amount of calculation, large memory consumption, and easy overfitting. For this reason, "convolutional layer + global average pooling (GAP)" can be used instead of "convolutional layer + fully connected layer" [27,28], as shown in Fig.5, the feature maps of the feature extraction module are directly associated with the categories of output, reducing the number of parameters originally located in the full connection layer. This replacement is equivalent to regularizing the network structure, so as to prevent overfitting problems in the model, and greatly reduce the memory occupation of VGG16 model [29].

The last convolution layer of the designed ten classification VGG16 network shall output 10 feature maps. Using the global average pooling instead of the full connection layer to directly associate the feature maps of the last convolution layer with the categories, respectively accumulate all pixel values of each feature map and calculate the average value, and then send the 10 average values to the softmax layer to obtain 10 probability values, that is, the probability value of the current picture belonging to a certain classification. The global average pooling operation integrates the global spatial information of the feature map and makes the network more robust to the spatial transformation of the input image.

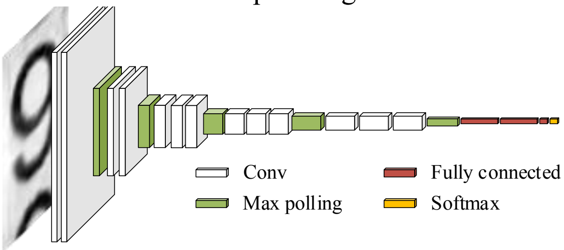

Fig.4 Structure diagram of VGG16 convolutional network

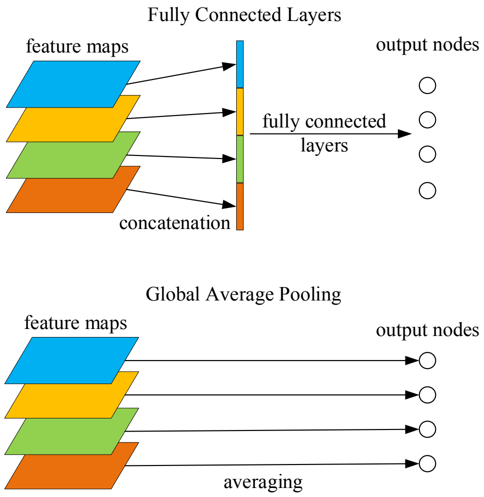

Fig.5 Replacing the full connection layer (up) with GAP (down) in VGG16

In the process of training, the "half character" display phenomenon that may exist in the wheel mechanical water meter that displays two characters at the same time is divided into two cases:

1. If one of the characters is completely displayed or more than half displayed, the label is set to the number corresponding to the character;

2. For the case that there are two characters both showing nearly half in the same character box, the label is set to the number corresponding to the smaller character according to the actual requirement that if the water consumption is less than a certain value, it should be rounded down.

4 Experiments and Results

In this project, 2924 images of wheel mechanical water meters are obtained from Chongqing Water Supply Company. Some image samples are shown in Fig.6. After the overall system is built, the experimental research is carried out according to the above technical route.

Fig.6 Some of the displayed data sample

4.1 Image preprocessing

In this paper, the image resolution of the original dataset has three different scales: 240 * 320, 540 * 960 and 960 * 1280. In order to carry out the follow-up network training smoothly, aiming at the problems of different sizes of the original water meter images and less samples, in this section, the size normalization processing and data expansion are carried out first.

For the original water meter images taken by hand, the camera focus is generally focused on the reading area. Therefore, by cutting the largest inscribed square of the original image from the middle, an image suitable for network training and removing some redundant information is obtained. For the processed image, if the dial reading area is missing in the sample, take the return operation to re-collect the sample. Then, through operations such as rotation, scaling, etc, the dataset is resized and expanded to four times the size of the original dataset, and finally, 11696 data samples with a unified size of 572 * 572 are obtained.

Then, the dataset is grayed based on the weighted average method [30], the weighted average method sums the gray values of the three channels of each pixel of the image according to a certain weight to obtain the average value, the calculation formula is shown in formula (2). Where $G_{ave}(x, y)$ represents the gray value result of any point of the image based on the weighted average method, $f(x, y)_R$, $f(x, y)_G$, and $f(x, y)_B$ are the gray values of three channels at any point of the image, and based on the sensitivity difference of human eyes to red, green and blue, W_R, W_G and W_B are taken as 0.299, 0.587 and 0.114 respectively.

$$G_{ave}(x, y) = \frac{(W_R * f(x, y)_R + W_G * f(x, y)_G + W_B * f(x, y)_B)}{3} \quad (2)$$

In the complex shooting environment, the following situations are inevitable in the original water meter pictures: uneven brightness of the dial caused by strong exposure or darkness in some areas caused by lighting factors, and blurred fonts caused by broken and aging of the plastic or glass disk and dust and sediment masking the dial. In addition, there are many printed words or patterns on the water meter pictures, which are very similar to the reading area in features, therefore, these bring great interference to the extraction of reading area for deep learning. In order to minimize the misidentification caused by external factors and dial itself, histogram equalization algorithm is used to make the gray distribution of the image uniform firstly, and the cumulative distribution function of the image is calculated:

$$s_k = T(r_k) = \frac{L}{M * N} \sum_{j=0}^{k} n_j, k = 1, 2, 3...L \quad (3)$$

Where s_k is the gray value mapped by the k-level gray value, r_k is the number of pixels of the k-level gray value, $M * N$ is the total number of pixels of the image, and L is the total gray level. According to the mapping relationship shown in the above formula, the original image is processed pixel by pixel to obtain the gray-scale transformed image.

After the contrast enhancement of the image, the mean filter is used to remove the noise in the water meter image. The mean filter replaces the gray value of the pixel with the average value of all pixels in the neighborhood of the pixel to be processed, which can effectively remove the noise and reduce the impact of these interference information on the subsequent processing. If M is the number of neighborhood pixels, s_{xy} represents the neighborhood pixel area with (x, y) as the center point, $g(x, y)$ represents the original image, and $f(x, y)$ represents the filtered image, then:

$$f(x, y) = \frac{1}{M} \sum_{(x,y)\in S_{xy}} g(x, y) \quad (4)$$

Mean filtering not only removes the noise of the image, but also makes the image smooth. Subsequent processing such as sharpening is required to improve the clarity of the image and enhance the boundary and detail information in the image. In conclusion, the influence factors and treatment methods in the pretreatment process are summarized in Table 1.

Table 1 Factors affecting reading segmentation and identification and treatment methods

Influence factors	Processing methods
Less samples	Data expansion
Image noise	Mean filtering
Uneven illumination	Contrast enhancement
Defocus blur	Image sharpening

4.2 Region of interest segmentation

After preprocessing the image, this paper uses the improved U-Net semantic segmentation network to perform binary segmentation on the water meter images to extract the reading area. There may be misidentified areas in the preliminary segmentation results, only the areas that meet the length-width ratio of the dial reading frame are retained through noise removal and contour detection. Fig.7 shows the comparison between the preprocessed data and the reading area segmentation results.

Fig.7 Pretreatment images of water meter (left) and reading area segmentation result (right)

After the location of the reading area is obtained through the segmentation model, firstly, the opening operation is used to eliminate the possible adhesion between the target area and the misidentified area, then other areas except the maximum contour are removed by contour detection. After that, fill the maximum contour with the smallest bounding rectangle area. Since the length-width ratio of the reading frame of the water meter is 4:1, if the length of the minimum circumscribed rectangle is greater than 4 times the width, delete the excess length from the right side of the rectangle to remove the possible non-target area on the right side. After restoring the shape and size of the rectangular reading box area, the target region extraction result is obtained by performing the "and" operation between this image and the original image. As shown in Fig.8:

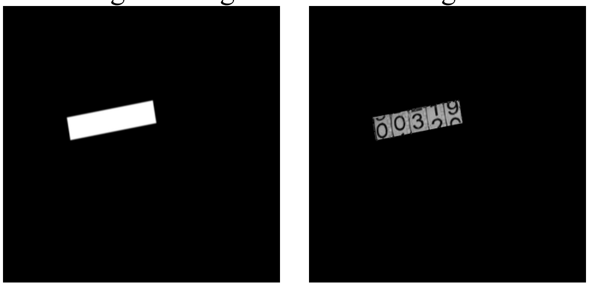

Fig.8 Rectangular filling image (left) and reading area extraction result image (right)

4.3 Reading identification

After obtaining the extraction result of the reading area, due to the inevitable tilt of the dial in the captured image, it is necessary to correct the tilt of the extraction result.. Generally, the selected object of tilt correction is usually the frame line of the object. In view of the fact that the actual processing results cannot guarantee the inclusion of the frame, and combined with the research object of this paper, there is an obvious black vertical line connecting the top and bottom between each reading character of the mechanical water meter. Firstly, this paper detects the tilt angle of these parallel interval vertical lines through Hough transform to correct the tilt of the reading area. In the standard parameterization mode, the straight line can be expressed as

$$\rho = x\cos\theta + y\sin\theta, \rho \geq 0, 0 \leq \theta < 2\pi \quad (5)$$

Where x, y are the coordinates of the straight line in the Cartesian coordinate system, ρ is the vertical distance from the straight line to the origin, and θ is the angle between the straight line and the x-axis. The straight line detection of Hough transform is realized by determining the intersection point of the curve transformed by the point on the straight line in the parameter space [31]. Then, the image is projected vertically and smoothed with Gaussian filter. And the parallel spaced vertical lines in the water meter image are determined by locating the peak points present in the projected curve. Relying on these parallel vertical lines and projected curves with the same spacing distance, the reading string can be divided into single characters. The results are shown in Fig.9:

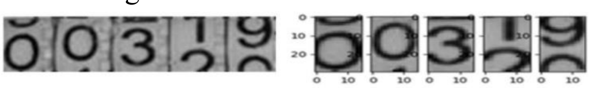

Fig.9 Tilt correction diagram (left) and single character segmentation diagram (right)

There are 8696 binary images in the training set and 3000 images in the testing set. In the process of sending the data to improved VGG16 network training, the learning rate is dynamically adjusted by gradient descent method. In order to further improve the performance of the model, this paper adds a certain network depth, generates recognition results and splices them in order, as shown in Fig.10:

Fig.10 The identification result after splicing

4.4 Analysis and comparison of experimental results

Table 2 shows the comparison of parameters between the improved VGG16 model and the original VGG16 model used in the reading recognition experiment. It can be seen from the table that the experimental scheme of "convolution layer + global average pooling" greatly reduces the parameters of the model, thus reducing the pressure for subsequent deployment of the model on the server.

Table 2 Comparison of parameters between improved VGG16 model and original VGG16 model

Model	Parameter quantity
Original VGG16 model	134.3M
Improved VGG16 model	14.7M

After analyzing the recognition results, it is found that the occurrence frequencies of characters in each reading are quite different from each other. Fig.11 shows the occurrence times of ten characters from "0" to "9" in the training set. According to the statistical chart and the actual water consumption of residents, the first two to three digits of the five-figure reading of the wheel mechanical water meter are highly likely to be "0", Thus, the sample's number of each character in the dataset is extremely unbalanced. The number of samples of the nine characters "1" to "9" is far less than the number of samples of the "0" character. The small number of samples of the nine characters makes it difficult for the recognition network to extract their sufficient features. Therefore, in order to further improve the recognition accuracy, after the single character segmentation operation, the second data expansion is carried out for characters other than "0" character in the training set. In the experiment of reading recognition on 3000 test images in this paper, the test results reach 95.6% recognition accuracy after secondary expansion of the data, as shown in Fig.12.

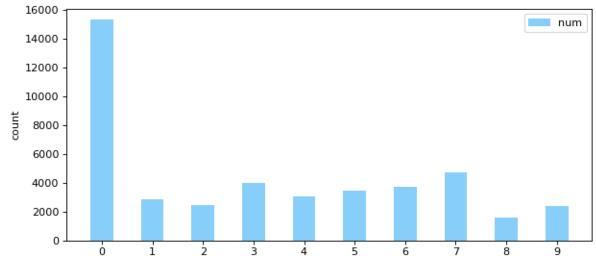

Fig.11 Statistics of the number of reading characters

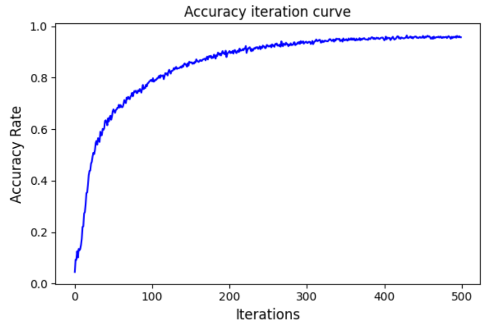

Fig.12 Recognition accuracy after data expansion

In the process of reading area extraction experiment, this paper also uses YOLOv5 target detection algorithm to compare with the improved U-Net network [32]. Fig.13 shows part of the results of detecting read regions using the YOLOv5 algorithm, where the pre-trained model YOLOv5s.pt is used, and the number of epochs for training is 150.

Fig.13 Positioning results based on YOLOv5 algorithm

As can be seen from the above figure that the YOLO algorithm can accurately locate the target area too, but it inevitably contains some background areas. It is not as precise as the pixel level segmentation and positioning of U-Net, and algorithms such as line detection still need to be carried out to extract exact edges of the reading area. At the same time, YOLO algorithm has higher requirements on dataset, model volume and complexity than U-Net [33]. The water meter dataset with small samples has high similarity, and the background similarity is also high, the type to be identified is single. Thus, it is more accurate and applicable to directly use U-Net, which is simple, efficient and suitable for small sample datasets, to obtain the mask of the reading area through instance segmentation than the rectangular box detection of YOLO algorithm.

Most of the existing instrument reading recognition algorithms and OCR (optical character recognition) technology have no good solutions to the character separation frame and character rotation display of wheel mechanical water meter, and the recognition accuracy is relatively low. Table 3 shows the comparison of the results of U-Net combined with VGG16 proposed in this paper from coarse to fine recognition method, BP neural network and OCR (optical character recognition) technology.

Table 3 Performance comparison of recognition algorithms

Algorithms	Correct recognition rate /%
BP neural network	82.9
OCR	88.7
U-Net+VGG16	95.6

Fig.14 Partial recognition results based on OCR technology

In the comparative experiment, the maximum number of iterations of BP neural network is 1000 and the learning rate is 0.01, but the learning efficiency and recognition accuracy are low, which means that the generalization ability of BP neural network is poor when dealing with the "half character" situation and the interference of redundant information. OCR recognition performance has certain requirements for the test sample itself. The printing lines and frames in the reading area of the dial and the light environment have brought some interference to the reading recognition of OCR technology. Fig.14 shows some results of calling the open OCR application program interface (API) provided by Baidu AI Cloud, it can be seen from the reading recognition results marked by the red box that even the OCR technology that has been put into commercial use at present has problems when it is directly applied to the reading of mechanical water meters, such as misreading the dividing line between the numbers as "1", and the number misreading caused by the inconsistency of the indication position, which is resulted from gear rotation. In addition, the response speed of calling the cloud API is greatly affected by the network, The applicability and robustness of this technology directly applied to the research of this topic are weak.

In the algorithm of this paper, the character region is accurately located and segmented to remove the influence of the background region and the separation line; Then combined with the "half character" processing scheme, the single character recognition is carried out, which solves the reading recognition problems in special cases such as the distortion of the number, the size difference and the deformity caused by the rotation of the gear, and achieves good experimental results. Compared with various character recognition algorithms, this strategy is more targeted and effective for the research object of wheel type mechanical water meter.

4.5 Discussion

In this study, the recognition strategy from coarse to fine is adopted. First, the dial reading area is accurately located at pixel level through the improved U-Net network, then the single character segmentation is carried out according to the structural characteristics of the reading area, and then the character recognition is carried out through the improved VGG16 network. The experimental results show that the technical scheme is very effective for the old wheel type mechanical water meter. The main shortcomings of this paper are:

1. The number of samples of the original dataset collected is still small, the type of dataset and the designed technical scheme are only for the wheel type water meter,

and do not include the pointer type water meter;

2. It is difficult to extract effective features by image processing in the case of too low brightness and serious obscuration of sundries;

3. The structure of the designed segmentation and recognition networks is still complex, and there are still some limitations in real-time processing, there is still room for improvement in accuracy.

The future development in the field of instrument identification requires higher resolution and picture quality for images, the development of high-definition cameras will inevitably lead to the algorithms requiring more computing resources, and the algorithms themself also require higher robustness and accuracy. In the follow-up of this study, it is necessary to further expand the scale of the dataset, and at the same time, the reading recognition of the pointer type water meter will be included in the scope of the study. On the basis of a larger scale of the dataset, more effective image preprocessing technology and more efficient and advanced target detection and recognition networks will be explored to achieve better detection accuracy.

5 Conclusion

This paper takes the time-consuming and laborious manual reading mode of the wheel mechanical water meter used by local residents as the starting point, and explores the application of the automatic recognition technology of water meter reading based on deep learning. By dividing the research problem into two sub-problems of "region of interest segmentation" and "character recognition", U-Net semantic segmentation network model and VGG16 convolution neural network model are built to solve the two key problems of "region of interest segmentation in complex shooting environment" and "accurate character recognition under the condition of missing dial digital information". It provides a perfect technical scheme for the research in the field of instrument reading recognition.

Acknowledgment:
The research of this paper is supported by four funds: (1) The National Science Foundation under Grant 61873043. (2) The Natural Science Foundation of Chongqing under Grant and cstc2020jcyj-msxmX0818 and cstc2020jcyj-msxmX0927. (3) The Science Technology Research Program of Chongqing Municipal Education Commission (Grant No. KJQN201901530). (4) the Campus Research Foundation of Chongqing University of Science and Technology under Grant CK2017zkyb024.References.

References:

[1] Yang, F (Yang, Fan) 1 ;Jin, LW (Jin, Lianwen) 1 ;Lai, SX (Lai, Songxuan) 1 ;Gao, X (Gao, Xue) 1 ;Li, ZH (Li, Zhaohai) 1 .Fully Convolutional Sequence Recognition Network for Water Meter Number Reading[J].*IEEE Access*,2019,Vol.7: 11679-11687.

[2] Zhu, Jiang 1,2 ;Li, Mingke 1,2 ;Jiang, Jin 1 ;Li, Jianqi 3 ;Wang, Zhaohong 1 ;Shen, Jinfei 3 . Automatic wheel-type water meter digit reading recognition based on deep learning.[J]. *Journal of Electronic Imaging*,2022,Vol.31(2): 23023.

[3] Hanson L . DEEP LEARNING APPLICATIO N IN INDUSTRY: WATER METER NUMBE RS RECOGNITION[J]. *European Journal of T echnical and Natural Sciences*, 2020:45-47.

[4] Chun-shan Li;Yu-kun Su;Rui Yuan;Dian-hui C hu;Jin-hui Zhu.Light-Weight Spliced Convoluti on Network-Based Automatic Water Meter Rea

ding in Smart City[J].*IEEE Access*,2019,Vol.7: 174359-174367.

[5] https://zixun.jia.com/article/341733.html.

[6] Tian-tian Li. Wheel Water Meter Reading Recognition Based on Improved Convolutional Neural Network [D]. Chongqing: Chongqing Normal University, 2019.

[7] http://www.xjishu.com/zhuanli/down/19220449.html.

[8] Hong Liu. Application status and development trend of intelligent water meter [J]. *Intelligent City*, 2019,5(7):30 31.

[9] Wen-xue Zheng, Zhi-min Yue, Xu-sheng Tang, Dan Chen. Calibration method of water meter based on machine vision [J]. *Journal of Mechanical & Electrical Engineering* 2019,36(03):271274.

[10] Wang, K (Wang, Kun) 1 .Application of Wireless Sensor Network based on LoRa in City Gas Meter Reading[J].*INTERNATIONAL JOURNAL OF ONLINE ENGINEERING*,2017, Vol.13 (12): 104-115.

[11] Mohd Zubairuddin;Pooja Thakre.Automatic Meter Reading using Wireless Sensor Module [J].*International Journal of Scientific Research in Science and Technology*,2018,Vol.4(8).

[12] Chen Zhangshao, Bi Sheng,Dong Min. Water Meter Reading Automatic Recognition System Based on Lightweight Convolutional Neural Network [J]. *Microcontrollers & Embedded Systems*,2021,21(12):12-15.

[13] Ben-Shimol Y , Greenberg S , Danilchenko K . Application-Layer Approach for Efficient Smart Meter Reading in Low-Voltage PLC Networks[J]. *IEEE Transactions on Communications*, 2018:1-1.

[14] Azerbaijan's Azersu opens tender to buy smart cards for water meters.[J].*Weekly Tenders Report*,2022.

[15] Jing-wei Sun. Research on Water Meter Reading Recognition [D]. Beijing: Beijing University of technology, 2016.

[16] Shuai Shang. Automatic Recognition of Water-Meter Reading [J]. *Computer engineering*, 2005, (5).

[17] Tian-hua Liu. Mechanical water meter reading recognition system based on machine vision [D]. Changsha: Hunan University, 2019.

[18] Ying C , Lei L I , Wen-yuan W , et al. Research on character recognition algorithm for domestic water meter[J]. *Modern Electronics Technique*, 2018.

[19] Hao-lin Shi. Method Research on Printed Character on Circuit Board based on Machine Learning [D]. Chengdu: University of Electronic Science and technology, 2019.

[20] Fan Zhang, Xiao-dong Wang, Xian-peng Hao. Intelligent Vehicle Character Recognition Based on Edge Features [J]. *Automation and Instrumentation*, 2020, No. 248 (06): 17-20 + 26.

[21] Chen Yue. Software Implementation of Intelligent Recognition System for Water Meter Reading Based on OpenCV [D]. Chengdu: University of Electronic Science and technology, 2019.

[22] Shuai-Cheng Pan, Lei Han, Yi Tao, et al. Research on character recognition technology for watermeter based on deep convolution neural network [J]. *Computer age*, 2020, No. 332 (02): 31-34.

[23] RONNEBERGER O,FISCHER P,BROX T. U-Net:Convolutional Networks for Biomedical Image Segmentation [C] // *International Conference on Medical Image Computing and Computer-assisted Intervention*. Munich: [s. n.] ,2015:234-241.

[24] HE K,ZHANG X,REN S,et al.Deepresidual learning for image recognition [C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*,2016:770-778.

[25] ZHANG Z, LIU Q, WANG Y. Road Extraction by Deep Residual U-Net[J] . *IEEE Geoscience and Remote Sensing Letters*,2018,15(5) :749-753.

[26] HE K,ZHANG X,REN S,et al. Identity mappings in deep residual networks [C]. *European Conference on Computer Vision*,Springer,Cham, 2016:630-645.

[27] Lin M, Chen Q, Yan S. Network in network[J]. arXiv preprint arXiv:1312.4400, 2013.

[28] Han, Jinyoung 1,2 ;Choi, Seong 1,2 ;Park, Ji In 3 ;Hwang, Joon Seo 4 ;Han, Jeong Mo 5 ;Lee, Hak Jun 6 ;Ko, Junseo 1,2 ;Yoon, Jeewoo 1,2 ; Hwang, Daniel Duck-Jin 6,7,8 .Classifying neovascular age-related macular degeneration with a deep convolutional neural network based on optical coherence tomography images.[J]. *Scientific Reports*,2022,Vol.12(1): 1-10.

[29] Wei Wang;Jinge Tian;Chengwen Zhang;Yanhong Luo;Xin Wang;Ji Li.An improved deep learning approach and its applications on colonic polyp images detection[J].*BMC Medical Imaging*,2020,Vol.20(1): 1-14.

[30] Yong-fei Hao, Xu-sheng Tang, Liang-li Cheng. Auto dashboard pointer detection based on machine vision [J]. *Journal of Mechanical & Electrical Engineering*,2022,39(1):134-140.

[31] Jie-xian Zeng, Gui-mei Zhang, Jun Chu, et al. Fit Line Using A Method Combined Hough Transform With Least Square [J]. *Journal of Nanchang Aviation University: Natural Science Edition*, 2003 (4): 6.
[32] https://github.com/ultralytics/yolov5
[33] Xue J , Zheng Y , Dong-Ye C , et al. Improved YOLOv5 network method for remote sensing image-based ground objects recognition[J]. *Soft Computing*, 2022:1-11.