Highlights

Efficient Recognition of Word-wheel Water Meters for Smart Urban Infrastructure

Shunyi Zhao,Qingxin Lu,Chengxi Zhang,Choon Ki Ahn,Kunming Chen

- The proposed word-wheel water meter image dataset comprises dedicated training and test sets for numeric area detection and water meter reading recognition. This dataset is sourced from images taken during on-site meter readings by meter readers in Hangzhou, China. It includes water meter images captured under various conditions, such as tilting, fouling, and small target areas, ensuring high representativeness, timeliness, and relevance.
- The proposed method for numeric area detection in water meter reading consists of two components: detection and correction. The detection component utilizes a deep model based on UNet, with a size of only 6.62MB, while maintaining a high level of accuracy. The correction step further refines the detected target area. Experimental results demonstrate that this algorithm significantly enhances the accuracy of the detection method, achieving an impressive 98.2% accuracy rate.
- The proposed Water Meter Reading Recognition Model (WMRRM) is employed for recognizing the readings within the numeric area. This model takes into full account the size characteristics of the water meter's numeric area and achieves reading recognition without the need for character segmentation, achieving an impressive 98.7% recognition accuracy with a mere 4.30MB in size.

Efficient Recognition of Word-wheel Water Meters for Smart Urban Infrastructure

Shunyi Zhao[a], Qingxin Lu[a], Chengxi Zhang[a,*], Choon Ki Ahn[b,*] and Kunming Chen[c]

[a] Key Laboratory of Advanced Control for Light Industry Processes, Ministry of Education, Jiangnan University, Wuxi, 214122, China
[b] School of Electrical Engineering, Korea University, 136-701, Seoul, South Korea
[c] Hangzhou Yuhang Environment (Water) Holding Group Co., Ltd., Hangzhou, China

ARTICLE INFO

Keywords:
Deep learning
water resource-saving
word-wheel water meter

ABSTRACT

Rapidly recognizing water meter readings is crucial for intelligent water management systems. Despite the widespread availability of smart water meters, the lower cost of word-wheel water meters means they continue to be used in most cases. As a result, manual reading and data review processes persist, hindering efficient management of water resources. Traditional recognition methods have been hampered by complex algorithms and insufficient robustness. This paper proposes a deep learning-based detection and recognition method for word wheel water meters, which involves dividing the reading process into three stages: detection, correction, and recognition. We have targeted algorithm design to suit the unique environment where the water meter is located. We then made specific refinements and improvements to the recognition method to improve the performance. The proposed method achieves high segmentation accuracy of 98.2% and recognition accuracy of 98.7% on a self-constructed data set collected in the whole of Hangzhou City, China. By streamlining manual meter reading and data review processes and increasing accuracy and efficiency, our approach holds great potential for facilitating effective water resource management.

1. Introduction

1.1. Background and Motivation

The automatic reading of traditional wheel-type water meters is a pressing need in developing smart infrastructure. As cities grow, they face numerous challenges, such as population expansion, overburdened infrastructure, traffic congestion, and increased governance difficulties. Fortunately, the emergence of smart cities provides a practical solution to these problems by addressing many issues that plague urban areas[1]. Smart infrastructure is essential in realizing smart cities [2, 3, 4] However, constructing infrastructure meeting the requirements of smart cities is technically challenging and economically costly. Therefore, it is crucial to improve existing urban infrastructure when building smart infrastructure.

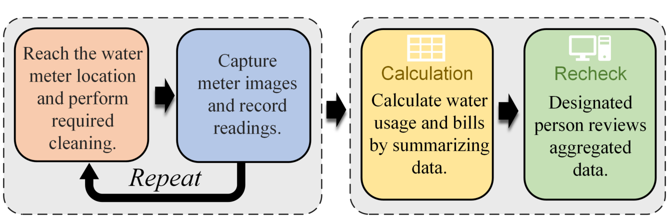


Figure 1: Standard operating procedure for traditional manual water meter reading process: high-cost and time-wasting.

*Corresponding author
shunyi@jiangnan.edu.cn (S. Zhao); lqx1229@stu.jiangnan.edu.cn (Q. Lu); dongfangxy@163.com (C. Zhang); hironaka@korea.ac.kr (C.K. Ahn); 365637286@qq.com (K. Chen)
ORCID(s): 0000-0001-5363-8305 (S. Zhao); 0000-0003-0993-9658 (C.K. Ahn)

Upgrading and retrofitting urban infrastructure can enable it to operate intelligently, meeting the demands of smart requests. Due to complex underground pipelines and other factors, the smartification of water supply systems in smart cities lags behind other types of infrastructure. The annual operations and maintenance costs for urban water systems worldwide reach approximately \$100 billion [5]. Making urban water supply systems intelligent can enhance their efficiency, reduce costs, and limit the waste of water resources. For instance, combining smart water meters with noise recorders can detect and locate leaks in water supply networks while automated equipment regulates pressure to minimize water loss [6, 7, 8, 9]. Despite the maturity of smart water meter technology and its ability to provide precise data in real time, most users still rely on traditional wheel-type water meters due to the high cost of equipment replacement. Replacing them entirely in the foreseeable future appears unlikely. Unfortunately, this challenges manual meter reading to the staff, which is a laborious and complicated process. As shown in Fig. 1, the process requires significant effort from one meter-reading worker alone. The most time-consuming and labor-intensive are the first and second steps. Usually, the distribution of water meters is relatively scattered, and the environment is very complex, such as buried underground and narrow corners, which significantly increases the workload of meter readers. Given that several million traditional water meters are still in use in most cities, not only the Hangzhou City of China, many meter-reading workers are required to read and record the readings on-site for most cities. The traditional method consumes substantial human and material resources, and each meter-reading worker may have to read hundreds of water meters daily, leading to inaccuracies and data review issues [10, 11].

To make matters worse, the traditional method's data timeliness is poor and unsuitable for efficient data use. The persistence of traditional water meters presents significant obstacles to the intelligent progress of urban water supply infrastructure. Therefore, developing a reliable automatic water meter reading system is crucial in addressing the above challenges.

1.2. Related Work

The recognition of meter readings on word wheel water meters has traditionally relied on a series of image-processing steps. These include preprocessing the input image to enhance quality and improve the definition of numeric edges through operations like grayscale conversion, binarization, filtering, and noise removal. Next, techniques such as edge detection, morphological transformations, or other methods are used to separate the digits on the meter from the image. Finally, recognition methods such as feature extraction and classification, template matching, or Support Vector Machines (SVM) [12, 13] are applied to recognize the segmented digits. For instance, Sablatnig et al. [14] employed Hough transform for pointer area detection and the rectangular box detection algorithm for target area segmentation. Subsequently, the Histogram of the Oriented Gradient (HOG) feature of the image was extracted for character recognition [15], with the SVM classifier utilized for classification. However, traditional feature extraction techniques often result in poor generalization ability of the algorithm. Template matching, another character recognition method, includes steps such as template selection, character segmentation, and matching. However, it is ineffective in finding a suitable template to match many character images.

Recently, a growing interest has been focused on utilizing deep learning for automated water meter readings. Liao et al. [16] introduced a deep learning-based object detector that accurately reads numbers from water meters. However, their dataset was not obtained from real scenes, which may limit its generalizability. Similarly, Yang et al. [17] proposed a fully convolutional sequence recognition network that still requires the integration of digit region detection algorithms to achieve optimal performance. In contrast, Peng and Chen [18] applied a deep residual neural network with impressive results, achieving over 80% accuracy in water meter reading recognition. In addition, Zhang et al.[19] proposed a method to automatically identify the position and reading of the water meter pointer by using target key point detection. Liang et al. [11] proposed two schemes of overall image recognition and image staged recognition for water meter reading recognition, and used Faster RCNN, YOLOv3 and other algorithms for comparative testing, and finally obtained a recognition accuracy of 90.61%. Jalel Ktari et al. [20] used YOLOv4 to detect the digital area of the water meter, used the Tesseract to recognize the readings, and deployed it in the mobile device. But the detection is less effective because there is no correction step for the detection. Together, these studies demonstrate the potential of deep learning-based approaches for water meter reading recognition and highlight the need for continued advancements in the accuracy and efficiency of these systems.

1.3. Contribution

The main contributions can be summarized as follows:

1) We present a new dataset of word-wheel water meter images with strong representativeness comprising specific training and testing sets for both numeric area detection and reading recognition networks. The dataset is derived from images captured by meter readers during the onsite reading process in Hangzhou City, China, ensuring timeliness and relevance. It encompasses water meter images under diverse conditions, such as tilted, defaced, and small target areas. Our experimental findings show that this dataset significantly enhances the robustness of meter reading networks in complex environments.

2) We present an effective approach for detecting and correcting water meter numeric areas. The proposed method employs the UNet model [21, 22] for numeric area detection, outperforming the Fully Convolutional Network-8s (FCN-8s) [23, 24] model with a segmentation accuracy 98.2%, surpassing it by 3.0%. Moreover, our UNet-based model is compact in size, with only 6.62MB. To further improve the performance, we introduce a rectification algorithm to correct and crop the predicted numeric areas. Experimental results show that the proposed algorithm is highly effective, leading to a remarkable enhancement of detection accuracy for both UNet and FCN-8s methods, increasing them by 2.4% and 2.2%, respectively.

3) Based on the convolutional neural network (CNN), we built a water meter reading recognition model (WM-RRM). This model considers the rectangular shape of the water meters' number area and the fact that the reading consists of 5 digits, enabling us to recognize the reading without character cutting. Our model outperforms VGG16 [25, 26] in terms of inference time and accuracy, achieving an impressive accuracy rate of 98.7%. Furthermore, our compact model size, less than 1% of the VGG model, makes it highly suitable for deployment on embedded devices.

2. Self-built Water Meter Dataset

We generated a professional dataset that allows for the automated detection of word-wheel water meter readings with high accuracy. This dataset was rigorously produced and thoroughly tested to ensure its maximum dependability, placing it as a significant resource for the study in this paper. The images in this dataset are entirely from the actual manual meter reading process of waterworks in Hangzhou City, China. Hangzhou is a megacity with a long history and the second-largest city in China's Yangtze River Delta region. The city's underground water pipes are complex and contain many traditional water meters of different ages. Therefore, the collected images of traditional water meters in Hangzhou are very representative. The model trained with this dataset can be easily extended to other cities. During the image

Figure 2: Various water meter pictures, which show the image of the water meter under the conditions of soil cover, tilted dial, blurred dial, uneven lighting, etc.

collection process of this dataset, we thoroughly considered the situations that are prone to occur in the real world based on the experience of front-line meter reading workers. For example, soil buried, dial defaced, dial tilted, small target areas, and different shooting equipment. In addition, due to various reasons, such as the time of the shooting and the weather at the time of the shooting, the image of the water meter will also have reflections and uneven lighting[27]. Fig. 2 shows examples of raw images from this dataset. The water meter status in the dataset is highly complicated, with varied degrees of coverage and visual distortion produced by different shooting angles. Traditional data processing methods face significant challenges as a result of these conditions.

2.1. Detection Dataset

The detection dataset is utilized for training and validating the detection network that identifies the numeric area of the word-wheel water meter. The label form of the water meter image in this dataset is the mask, as shown in Fig. 3, where the white area is the target area, and the black

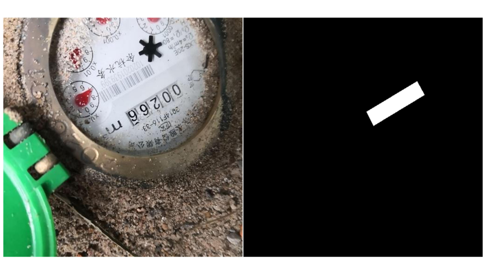
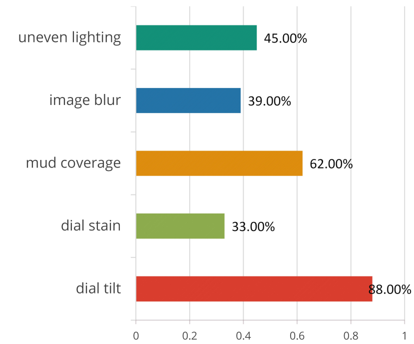
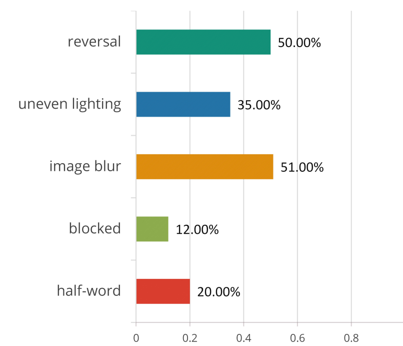


Figure 3: A example of detection dataset.

area is the irrelevant background. The detection data set contains 2100 images in 512×512 pixel color PNG format, 1600 images in the training set, and 500 images in the test set. Fig. 4(a) shows the proportion of water meter images in various situations in the data set. For example, the dial tilt image accounts for 88.00% of the total number of images in the data set. The composition of the data set is in line with the actual meter reading situation.



(a) The proportion of various image types in the detection dataset



(b) The proportion of various images in the recognition dataset

Figure 4: Data overview of self-built datasets.

2.2. Recognition Dataset

The recognition data set trains and verifies the digital recognition network. The data set contains 12,800 images of the numeric area of the water meter, all of which are extracted from the original image, and the images are in JPG format of 60×200 pixels. The number of pictures in the training and testing sets is 10,800 and 2,000, respectively. The format of the label is the file name; that is, the first five digits of each image file name are the readings, as shown in Fig. 5. This dataset contains common cases such as half-bit, blur, blocked by foreign objects, and uneven illumination. Its composition is shown in Fig. 4(b). Their levels of exposure, contrast, and other characteristics exhibit varying degrees of change.

3. Integral identification method of word-wheel water meter

Fig. 6 shows the basic framework of the integral identification method. The proposed method for word-wheel water meter recognition consists of three main parts.

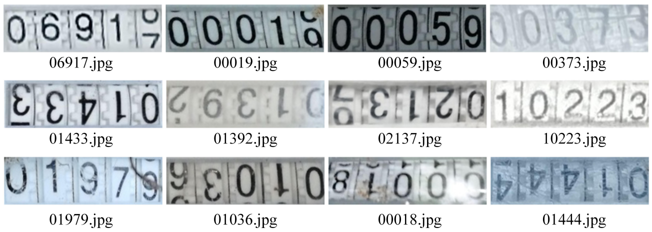


Figure 5: Examples of recognition dataset, which shows the situation where the reading area of the water meter is blocked by foreign objects, blurred, half-word, etc.

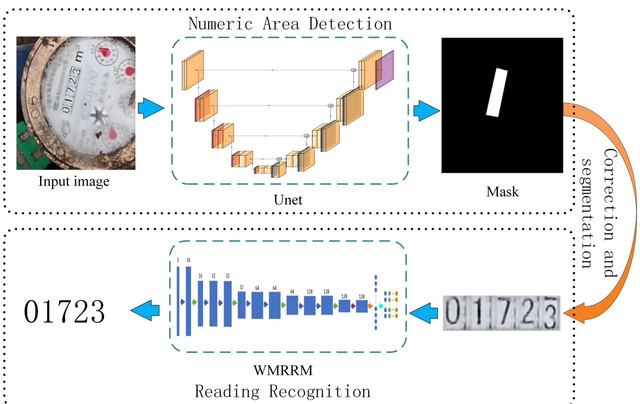


Figure 6: The basic framework, our method is able to directly use the original picture as input to obtain the recognition result of the water meter reading.

- The first part is the water meter numeric area detection. This part inputs the original image into the trained detection model to obtain the mask of the target area.
- The second part is the numeric area correction and segmentation. This part applies our correction algorithm to the target area and obtains the corrected image of the numeric area.
- The final part is the reading recognition. This part inputs the numeric area image of the water meter into the trained recognition model and generates the recognition result.

3.1. Numeric Area Detection

To identify the numbers on the water meter, the detection network aims to locate the numeric area of the water meter precisely from the original image. We present a digital area detection method that employs the UNet network. Olaf Ronneberger et al.[21] first proposed the UNet structure in 2015 to tackle the segmentation problem in biomedical images. The network can perform end-to-end training with limited images and has shown remarkable performance in the ISBI (International Symposium on Biomedical Imaging) challenge of segmenting neuronal structures in electron microscope stacks.

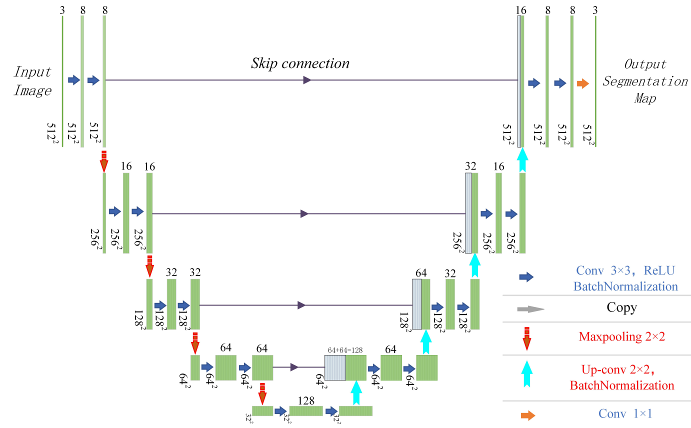


Figure 7: The detection network structure, the skip connection of the network, adopts splicing to fuse deep features and shallow features so that the deep network can obtain more feature information.

The detection network architecture employed in this study is illustrated in Fig. 7. UNet is an encoding-decoding framework. The left half of the network downsamples the input image to extract image features, while the right half upsamples the feature maps to restore the image size. We made some changes to the convolution block of UNet, as shown in Fig. 8, the convolutional layers use a 3×3 convolution kernel, with padding set to *'same'*, followed by the batch normalization layer and ReLU activation function. The downsampling operation uses maximum pooling with a 2×2 pooling window and a step size of 2. The upsampling method employs transposed convolution. UNet incorporates skip connections to fuse feature maps from corresponding positions of the encoder channels during the upsampling process at each level. By combining low-level and high-level features, the network can retain high-resolution detail information contained in the high-level feature maps, thereby enhancing the accuracy of image segmentation. The input and output image dimensions of the network are set to 512×512×3. MSE (Mean Squared Error) is used as the loss function when training the network and is given by:

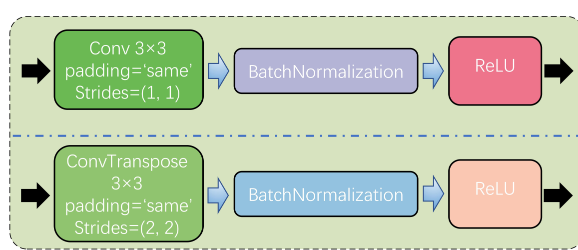


Figure 8: Structure of convolutional blocks and transposed convolutional blocks.

$$L = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{1}$$

where n is the number of output elements, y_i is the true value of the i-th output element, and $\hat{y}_i$ is the predicted value of the i-th output element. When the value of the loss function is minimal, the predicted value is closest to the true value.

We employ the widely used adaptive moment estimation (Adam) algorithm to optimize model weights during training. This algorithm dynamically tunes the learning rate and by using Adam, our model learned the underlying patterns in the data efficiently. Please see **Algorithm 1** for details,

Algorithm 1 Steps of Adam[28]

Require:

- α: step size
- $\beta_1, \beta_2 \in [0, 1)$: Exponential decay rates for the moment estimates.
- $f(\theta)$: Stochastic objective function with parameters θ, i.e. the loss function.
- θ_0: Initial parameter vector
  $m_0 \leftarrow 0$(Initialize 1^{st} moment vector)
  $v_0 \leftarrow 0$(Initialize 2^{nd} moment vector)
  $t \leftarrow 0$(Initialize timestep)

while θ_t not converged **do**

- $t \leftarrow t + 1$
- $g_t \leftarrow \bigtriangledown_\theta f_t(\theta_{t-1})$ (Get gradients w.r.t stochastic objective at timestep t)
- $m_t \leftarrow \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot g_t$ (Update biased first-moment estimate)
- $v_t \leftarrow \beta_2 \cdot v_{t-1} + (1 - \beta_2) \cdot g_t^2$ (Update biased second raw moment estimate)
- $\hat{m}_t \leftarrow m_t/(1 - \beta_1^t)$ (Compute bias-corrected first-moment estimate)
- $\hat{v}_t \leftarrow v_t/(1 - \beta_2^t)$ (Compute bias-corrected second-moment estimate)
- $\theta_t \leftarrow \theta_{t-1} - \alpha \cdot \hat{m}_t/(\sqrt{\hat{v}_t} + \varepsilon)$ (Update parameters)

end while
return θ_t (Resulting parameters)

where g_t^2 indicates the element-wise square $g_t \odot g_t$. Set the parameters $\alpha = 0.001$, $\beta_1 = 0.9$, $\beta_2 = 0.999$ and $\varepsilon = 10^{-8}$. All operations on vectors are element-wise. With β_1^t and β_2^t, we denote β_1 and β_2 to the power t.

3.2. Correction and Segmentation

To generate a precise digital display of a water meter image, we employ a detection network that produces a mask indicating the white region in the image. The white region of the mask represents the digital display area of the water meter. By performing contour detection on the white area, we can obtain a set of edge points that define the boundaries of the digital display area. From these edge points, we can determine the minimum circumscribed rectangle that encloses the white area. However, in cases where the digital area is significantly tilted, cropping the image based on the minimum circumscribed rectangle may not be appropriate. Such cropping would result in an image that includes irrelevant regions, as illustrated in Fig. 9. Hence, further refinement of the minimum circumscribed rectangle is necessary to achieve better results. As shown in Fig. 9, we correct

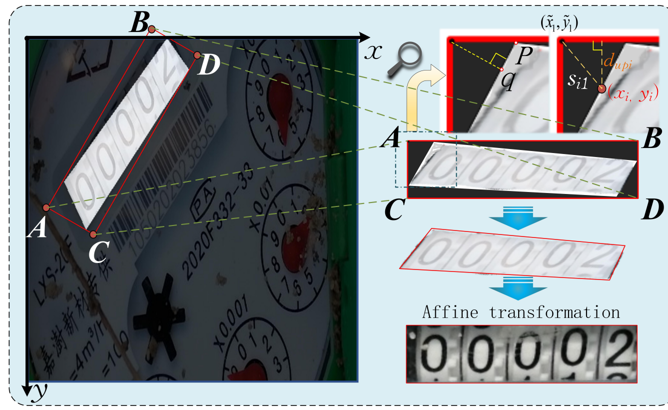


Figure 9: Minimum circumscribed rectangle correction, the left picture is the original picture covered by the predicted mask, in order to make it more intuitive to watch, the mask has been translucent.

the four vertices A, B, C, D of the rectangle respectively by introducing two variables: s_{ij} $(i = 1, 2, ...; j = 1, 2, 3, 4)$, the distance from the i-th edge point to the j-th rectangle frame vertex

$$s_{ij} = \sqrt{(x_i - \tilde{x}_j)^2 + (y_i - \tilde{y}_j)^2}, \tag{2}$$

C_{ij}, the weighted distance from the i-th edge point to the j-th vertex of the rectangle

$$C_{ij} = \begin{cases} \omega_1 d_{upi} + \omega_2 s_{ij} & (3) \\ \omega_1 d_{downi} + \omega_2 s_{ij} & (4) \end{cases}$$

where (x_i, y_i) $(i = 1, 2, ...)$ is the coordinates of the edge points of the white area (target area), $(\tilde{x}_j, \tilde{y}_j)$ $(j = 1, 2, 3, 4)$ is the coordinates of the four vertices A, B, C, D of the rectangular frame; d_{upi} and d_{downi} $(i = 1, 2, ...)$ are the distance from the i-th edge point to the upper and lower edges of the rectangle; ω_1 and ω_2 are the weighting coefficients. Here, choose $\omega_1 = 0.975, \omega_2 = (1 - \omega_1)$. If d_{upi} and d_{downi} are not introduced, and only s_{ij} is used to correct, then there will be a situation where the point q is mistakenly used as the upper left corner vertex of the new quadrilateral in Fig. 9. But we want to use p as a new vertex; this situation will result in an incomplete cropped target area.

- When correcting the upper two vertices A and B of the rectangle, use Eq. (3);

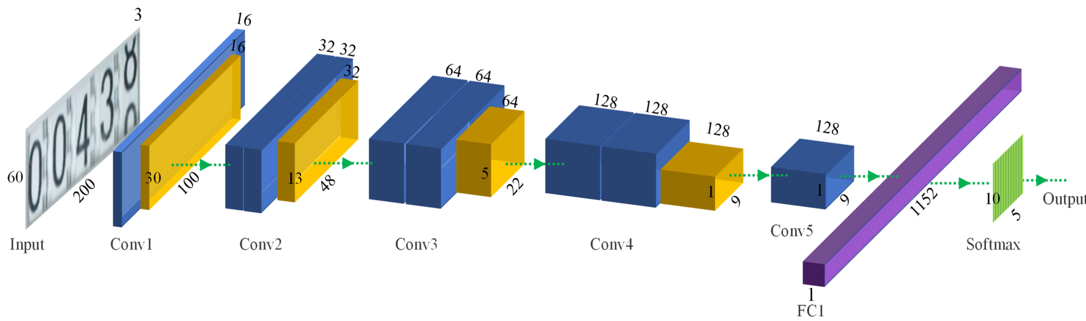


Figure 10: The structure of the Water Meter Reading Recognition Model, the model has a total of five convolutional blocks, the last layer of the first four convolutional blocks is the maximum pooling layer, the middle three convolutional blocks perform two convolution operations, the last convolution block is a convolution operation of size 1×1, the optimizer is Adam.

- When correcting the vertices C and D, use Eq. (4).

Consider the vertex A located at the upper left corner of Fig. 9, where $(\tilde{x_1}, \tilde{y_1})$ denotes its coordinate. First, we calculate the distance s_{i1} $(i = 1, 2, ...)$ from all edge points of the white area to A using Eq. (2), then calculate the distance d_{upi} $(i = 1, 2, ...)$ from all edge points to the upper side(AB) of the rectangle. Use Eq. (3) to calculate the minimum weighted distance C_{m1},

$$C_{m1} = \min_i \{\omega_1 d_{upi} + \omega_2 s_{i1}\} \tag{5}$$

then the point (x_m, y_m) $(i = m)$ with the smallest weighted distance to the vertex A replaces the vertex A as the upper left corner vertex of the new quadrilateral. Repeat this process for the remaining three vertex coordinates of the rectangle. Once the coordinates of the four new vertices are obtained, the original image can be cropped according to the new quadrilateral. A 60×200 water meter numeric area image can be obtained by performing an affine transformation on the cropped image.

3.3. Recognition Network

The primary objective of the recognition network is to accurately identify the reading on a water meter from a cropped image of its digital area. Conventionally, research in this field has concentrated on recognizing individual characters, necessitating character segmentation before recognition. Nonetheless, when dealing with datasets of high complexity, achieving precise character segmentation can prove to be an arduous task.

Despite recent advancements in OCR (Optical Character Recognition) research [29, 30, 31], applying existing networks directly to water meter recognition tasks has proven quite challenging [32]. Achieving satisfactory results requires a network specifically designed and trained to handle the unique complexities of water meter images.

Using CNN as the foundation, we have developed a WMRRM model that eliminates the need for character segmentation. The input and output configurations of the model have been customized to suit the specific dataset used in this study, and the network architecture has been depicted in Fig. 10. The core of the network consists of three identical blocks (Conv2, Conv3, and Conv4), each comprising two convolution operations leveraging 3×3 kernels, one maximum pooling operation, and an incremental increase in the number of convolution kernels. Additionally, Conv5 employs a 1×1 kernel convolution operation. Following feature extraction through the convolutional layer, the fully connected layer is responsible for classification, using the Softmax function as the activation function. Dropout layers have been incorporated post-Conv2, Conv3, Conv4, and Conv5 to mitigate overfitting. Specifically designed to process a 60×200 digital display image, the network produces a 10×5 matrix that corresponds to each reading column-wise, with each row representing one of the ten digits $(0, 1, ..., 9)$. Finally, the digit with the highest probability is selected as the recognized value.

Softmax is commonly utilized in multi-classification tasks to transform all input values into $(0, 1)$ while ensuring that the sum of all outputs equals 1, making it interpretable as a probability distribution. The definition of Softmax is:

$$S(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{k} e^{z_j}} \tag{6}$$

where k represents the number of outputs or categories; z is the output vector; z_j is the j-th output or category in z; i represents the category that currently needs to be evaluated. The category with the highest probability is the prediction category. The loss function of the recognition network is cross-entropy loss:

$$L = -\frac{1}{N} \sum_{i=1}^{N} \sum_{c=1}^{M} y_{ic} \log(p_{ic}) \tag{7}$$

where N represents the number of samples; M represents the number of categories; i represents the i-th sample; c represents the c-th category; p_{ic} is the prediction probability that the observed sample $i \in c$; and y_{ic} is a conditional function:

$$y_{ic} := \begin{cases} 1, & \text{if } i \in c, \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

Table 1
Detection and Segmentation results on test set.

Models		Train time	Total number of test images	Number of split images	Number of images segmented exactly	Test accuracy	Model size	Average inference time
UNet	**with Augmentation and Correction**	6160s	500	499	**491**	**98.20%**	**6.62MB**	**40ms**
	without Augmentation	1230s	500	484	466	93.20%	**6.62MB**	**40ms**
	without Correction	6160s	500	499	479	95.80%	**6.62MB**	**40ms**
FCN_8s	with Augmentation and Correction	5910s	500	**500**	477	95.20%	70.90MB	**40ms**
	without Correction	5910s	500	**500**	465	93.00%	70.90MB	**40ms**

4. Experimental process and analysis

We evaluated our method's performance by constructing the network using the TensorFlow deep learning framework.

4.1. Detection and Segmentation Experiments

In order to evaluate our proposed method, we conducted comparative experiments using the FCN-8s network, a well-established semantic segmentation network introduced by Jonathan Long et al. in 2015 for image segmentation tasks [23]. Like UNet, FCN-8s consists of an encoder and a decoder part, but its specific structure differs from UNet. FCN-8s also utilizes layer-skip connections, which are fused by summing feature maps. To improve the training data, we employed image rotation and shuffled the training data at the end of each round of training. The model was trained over 100 rounds, and we obtained the resulting model file in *.h5* format. We then assessed the segmentation effect on the test set by invoking various detection models and correction segmentation algorithms. The results of the water meter numeric area detection test set, consisting of 500 images, are presented in Table-1. The UNet-based detection model, combined with data enhancement and correction algorithms, achieved the highest accuracy of 98.20%, accurately segmenting 491 out of the 499 successfully detected target areas. Our ablation experiments and comparisons highlighted that data enhancement significantly improves the model's accuracy - without it, the accuracy of the detection model dropped to 93.20%. Furthermore, our experimental results demonstrate the effectiveness of the proposed correction algorithm, which enhances the accuracy of the UNet-based detection method by 2.4% and the FCN-8s-based detection method by 2.2%. The effectiveness of the correction algorithm, correcting the tilt of the target region and removing irrelevant regions, is illustrated in Fig. 11, thereby improving the recognition effect of subsequent recognition models. In comparison to FCN-8s, UNet demonstrated higher detection accuracy under the same conditions while also occupying less memory. This is particularly advantageous for deployment on embedded devices. After conducting multiple experiments, we found that the average inference time of both models was similar. Furthermore, training time for both UNet and FCN-8s was found to be nearly identical, given the same training set size and epoch.

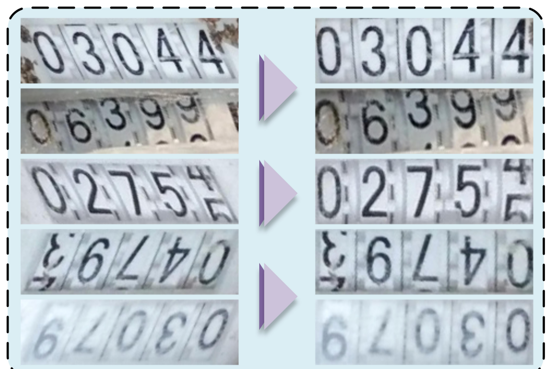


Figure 11: The effect display of the correction algorithm, the image on the left is cropped directly using the minimum circumscribing rectangle of the target area, and the image on the right is the cropping image obtained by correcting the minimum circumscribing rectangle using the rectification algorithm.

4.2. Reading Recognition Experiments

To evaluate the performance of our recognition network model, we conducted a comparison with the classic VGG16 network[25]. VGG16 is a deep convolutional neural network model proposed by Karen Simonyan and Andrew Zisserman, and it was named after the Visual Geometry Group. The objective of this model was to augment the depth and complexity of the network while reducing the number of parameters and mitigating the risk of over-fitting. However, since the size of the segmented image we used is 60×200, which is not compatible with VGG16's input size, we resized it to 224×224 before feeding it to the network. We trained the *.h5* files of the two models using the Adam optimizer

Table 2
Recognition result on testset.

Models	Train time	0-error accuracy	1-error accuracy	Model size	Average inference time
WMRRM	490s	**97.40%**	**98.70%**	**4.30MB**	**19.4ms**
VGG16	1390s	97.30%	98.30%	528.00MB	25.8ms

for WMRRM and the Stochastic Gradient Descent (SGD) algorithm using Nesterov momentum for VGG16. Finally, we tested both models on a set of 2000 images, including upside-down images, and obtained the results shown in Table-2.

In terms of practical application, it is important to consider the accuracy of identifying the entire sequence of water meter readings to determine a correct result. The evaluation index used is the recognition accuracy rate of the readings, which can be mathematically expressed as follows:

$$Acc = \frac{T_P}{M} \times 100\% \tag{9}$$

where T_P represents the number of samples with correct water meter reading recognition results, and M represents the total number of samples tested. The WMRRM surpasses VGG16 in accuracy on a dataset of traditional wheel-type water meter images. The 1-error accuracy metric indicates the rate of correctly identifying the last digit of half-digit number areas identified as another number, while the 0-error accuracy reflects overall accuracy. Our results show that WMRRM outperforms VGG16 on both metrics, including superior recognition of upside-down images. Furthermore, our model is over 100 times smaller than VGG16 and has a faster inference time, making it an excellent choice for deployment on embedded devices.

The proposed method achieved impressive results as shown in Fig. 12. It exhibits excellent robustness in complex environments and can accurately recognize water meter images even under challenging conditions such as mud coverage, blur, and tilted shooting angles. Overall, these findings demonstrate the superior performance of our recognition method, offering significant practical benefits for water meter reading automation.

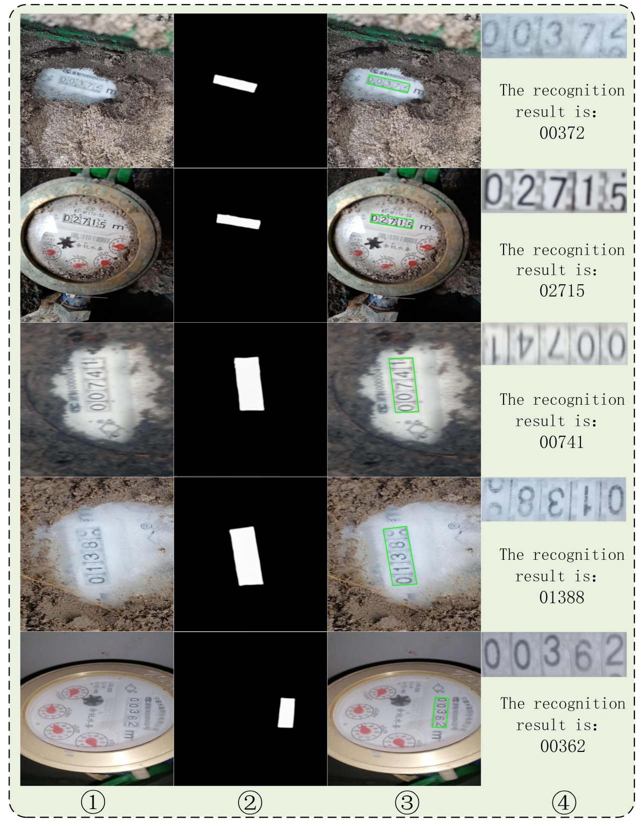


Figure 12: End-to-end reading recognition effect, in the figure ① is the original image, ② is the mask predicted by the detection model, ③ is the corrected digital area selected from the original image, and ④ is the cropped image and reading recognition results.

5. Conclusion

The complexities of water meter environments and diverse data collection methods often result in image disturbances such as tilted shooting angles, blurring, and dial contamination, posing significant challenges for existing character recognition methods. We propose a practical water meter reading recognition approach to address these issues. First, we repurposed the UNet network initially intended for medical image segmentation into a water meter numeric area detection network. We also proposed a correction algorithm to adjust the predicted numeric area, resulting in a 98.2% detection accuracy on our dataset using a model size of only 6.62MB. Next, we developed a CNN-based model for water meter reading recognition, better suited to the size characteristics of the numeric area of the water meter. Experimental results demonstrate a high recognition accuracy of 98.7%, with a small model size of only 4.30MB. Finally, we integrated the detection and recognition methods into a complete framework, offering a practical solution to traditional word wheel water meter reading recognition problems. This approach can be applied in embedded devices to enable traditional water meters' intelligence, installed on mobile devices for meter reading workers, or deployed in the cloud for data reading and review. The proposed method can potentially drive smart water management infrastructure development, offering significant benefits for water meter reading automation.

References

[1] X. Zhou, S. Li, Z. Li, W. Li, Information diffusion across cyber-physical-social systems in smart city: A survey, Neurocomputing 444 (2021) 203–213.

[2] F. G. Brundu, E. Patti, A. Osello, M. D. Giudice, N. Rapetti, A. Krylovskiy, M. Jahn, V. Verda, E. Guelpa, L. Rietto, A. Acquaviva, Iot software infrastructure for energy management and simulation in smart cities, IEEE Transactions on Industrial Informatics 13 (2017) 832–840.

[3] L. Zuo, P. He, C. Zhang, Z. Zhang, A robust approach to reading recognition of pointer meters based on improved mask-rcnn, Neurocomputing 388 (2020) 90–101.

[4] M. Shahidehpour, Z. Li, M. Ganji, Smart cities for a sustainable urbanization: Illuminating the need for establishing smart urban infrastructures, IEEE Electrification Magazine 6 (2018) 16–33.

[5] M. Moy de Vitry, M. Y. Schneider, O. F. Wani, L. Manny, J. P. Leitão, S. Eggimann, Smart urban water systems: what could possibly go wrong?, Environmental Research Letters 14 (2019) 081001.

[6] M. Quiñones-Grueiro, M. Ares Milián, M. Sánchez Rivero, A. J. Silva Neto, O. Llanes-Santiago, Robust leak localization in water distribution networks using computational intelligence, Neurocomputing 438 (2021) 195–208.

[7] Q. Sun, Y. Zhang, B. Lu, H. Liu, Flow measurement-based self-adaptive line segment clustering model for leakage detection in water distribution networks, IEEE Transactions on Instrumentation and Measurement 71 (2022) 1–13.

[8] S. C. Hsia, S.-H. Wang, S.-W. Hsu, Smart water-meter wireless transmission system for smart cities, IEEE Consumer Electronics Magazine 10 (2020) 83–88.

[9] S. El-Zahab, T. Zayed, Leak detection in water distribution networks: an introductory overview, Smart Water 4 (2019) 1–23.

[10] E. Farah, I. Shahrour, Smart water for leakage detection: Feedback about the use of automated meter reading technology, in: 2017 Sensors Networks Smart and Emerging Technologies (SENSET), IEEE, 2017, pp. 1–4.

[11] Y. Liang, Y. Liao, S. Li, W. Wu, T. Qiu, W. Zhang, Research on water meter reading recognition based on deep learning, Scientific Reports 12 (2022) 12861.

[12] V. N. Vapnik, An overview of statistical learning theory, IEEE Transactions on Neural Networks 10 (1999) 988–999.

[13] C. Cortes, V. Vapnik, Support-vector networks, Machine Learning 20 (1995) 273–297.

[14] R. Sablatnig, W. G. Kropatsch, Application constraints in the design of an automatic reading device for analog display instruments, in: Proceedings of 1994 IEEE Workshop on Applications of Computer Vision, IEEE, 1994, pp. 205–212.

[15] N. Dalal, B. Triggs, Histograms of oriented gradients for human detection, in: 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05), volume 1, IEEE, 2005, pp. 886–893.

[16] S. Liao, P. Zhou, L. Wang, S. Su, Reading digital numbers of water meter with deep learning based object detector, in: Pattern Recognition and Computer Vision: Second Chinese Conference, PRCV 2019, Xi'an, China, November 8–11, 2019, Proceedings, Part I 2, Springer, 2019, pp. 38–49.

[17] F. Yang, L. Jin, S. Lai, X. Gao, Z. Li, Fully convolutional sequence recognition network for water meter number reading, IEEE Access 7 (2019) 11679–11687.

[18] Y. Peng, Z. Chen, Application of deep residual neural network to water meter reading recognition, in: 2020 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA), IEEE, 2020, pp. 774–777.

[19] Q. Zhang, X. Bao, B. Wu, X. Tu, Y. Jin, Y. Luo, N. Zhang, Water meter pointer reading recognition method based on target-key point detection, Flow Measurement and Instrumentation 81 (2021) 102012.

[20] J. Ktari, T. Frikha, M. Hamdi, H. Elmannai, H. Hmam, Lightweight ai framework for industry 4.0 case study: water meter recognition, Big Data and Cognitive Computing 6 (2022) 72.

[21] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18, Springer, 2015, pp. 234–241.

[22] J. Wang, J. Hu, Y. Song, Q. Wang, X. Zhang, S. Bai, Z. Yi, Vmat dose prediction in radiotherapy by using progressive refinement unet, Neurocomputing 488 (2022) 528–539.

[23] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 3431–3440.

[24] O. Ghorbanzadeh, Y. Xu, P. Ghamisi, M. Kopp, D. Kreil, Landslide4sense: Reference benchmark data and deep learning models for landslide detection, IEEE Transactions on Geoscience and Remote Sensing 60 (2022) 1–17.

[25] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556 (2014).

[26] M. S. Hossain, M. Al-Hammadi, G. Muhammad, Automatic fruit classification using deep learning for industrial applications, IEEE Transactions on Industrial Informatics 15 (2019) 1027–1034.

[27] A. Salari, A. Djavadifar, X. Liu, H. Najjaran, Object recognition datasets and challenges: A review, Neurocomputing 495 (2022) 129–152.

[28] D. P. Kingma, J. Ba, Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980 (2014).

[29] X. Kong, K. Wang, M. Hou, X. Hao, G. Shen, X. Chen, F. Xia, A federated learning-based license plate recognition scheme for 5g-enabled internet of vehicles, IEEE Transactions on Industrial Informatics 17 (2021) 8523–8530.

[30] X.-F. Wang, Z.-H. He, K. Wang, Y.-F. Wang, L. Zou, Z.-Z. Wu, A survey of text detection and recognition algorithms based on deep learning technology, Neurocomputing 556 (2023) 126702.

[31] Z. Xiang, H. Wu, D. Zhou, Recognition of characters on curved metal workpiece surfaces based on multi-exposure image fusion and deep neural networks, Neurocomputing 512 (2022) 117–129.

[32] J. Zhang, W. Liu, S. Xu, X. Zhang, Key point localization and recurrent neural network based water meter reading recognition, Displays 74 (2022) 102222.