mathematics

MDPI

Article

Utilizing Cross-Ratios for the Detection and Correction of Missing Digits in Instrument Digit Recognition

Jui-Hua Huang [1,*], Yong-Han Chen [2] and Yen-Lung Tsai [1]

[1] Department of Mathematical Sciences, National Chengchi University, Taipei 116011, Taiwan; yenlung@nccu.edu.tw
[2] Department of Applied Artificial Intelligence, Ming Chuan University, Taipei 111, Taiwan; 09160881@me.mcu.edu.tw
* Correspondence: 105751503@nccu.edu.tw

Abstract: This paper aims to enhance the existing Automatic Meter Reading (AMR) technologies for utilities in the public services sector, such as water, electricity, and gas, by allowing users to regularly upload images of their meters, which are then automatically processed by machines for digit recognition. We propose an end-to-end AMR approach designed explicitly for unconstrained environments, offering practical solutions to common failures encountered during the automatic recognition process, such as image blur, perspective distortion, partial reflection, poor lighting, missing digits, and intermediate digit states, to reduce the failure rate of automatic meter readings. The system's first stage involves checking the quality of the user-uploaded images through the SVM method and requesting re-uploads for images unsuitable for digit extraction and recognition. The second stage employs deep learning models for digit localization and recognition, automatically detecting and correcting issues such as missing and intermediate digits to enhance the accuracy of automatic meter readings. Our research established a gas meter training dataset comprising 52,000 images, extensively annotated across various degrees, to train the deep learning models for high-precision digit recognition. Experimental results demonstrate that, with the simple SVM model, an accuracy of 87.03% is achieved for the classification of blurry image types. In addition, meter digit recognition (including intermediate digit states) can reach 97.6% (mAP), and the detection and correction of missing digits can be as high as 63.64%, showcasing the practical application value of the system developed in this study.

Keywords: automatic meter reading; instrument degree interpretation; handling of missing digits; cross-ratio

MSC: 68T45

check for updates

Citation: Huang, J.-H.; Chen, Y.-H.; Tsai, Y.-L. Utilizing Cross-Ratios for the Detection and Correction of Missing Digits in Instrument Digit Recognition. *Mathematics* **2024**, *12*, 1669. https://doi.org/10.3390/math12111669

Academic Editors: Jesús García-Herrero and Johan Debayle

Received: 9 April 2024
Revised: 16 May 2024
Accepted: 24 May 2024
Published: 27 May 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Automatic Meter Reading technology represents an innovative approach to facilitate the automatic identification and recording of instrument digits through image processing techniques. This technology is primarily applicable to equipment monitoring and billing systems. With the ongoing advancement and widespread application of digital technology, intelligent instruments are gradually replacing traditional ones. However, in many regions and countries, especially within the realm of public utilities, the task of reading instrument digits still predominantly relies on manual operations by company staff on-site, complemented by photographing to preserve proof of the readings. This process is prone to errors and may necessitate another employee's second verification of the images in specific scenarios, such as when consumers dispute bill amounts or when there are significant discrepancies in consumption figures. This offline verification workflow is not only time-consuming but also inefficient.

This paper focuses on household gas meters as a methodological verification subject, addressing three main issues: the quality of input images, identification accuracy, and the repair of missing digit problems. It proposes corresponding solutions and optimization strategies:

(1) In the application scenario of automated meter readings, the installation locations of home gas meters vary greatly, leading to significant differences in the quality of images uploaded by users, substantially reducing the success rate of automated readings. Poor image quality is primarily caused by several factors: blurred images, localized reflections, poor lighting conditions, missing digits, digits in transitional states, and images containing elements unrelated to the meter, as shown in Figure 1. These issues are primarily due to the environment's variability and improper shooting handling (such as incorrect shooting angles or exposure settings). To address the variations in image clarity caused by different shooting environments, this study employs machine learning techniques to analyze a large dataset of images. Using a Support Vector Machine (SVM) as the classifier, this research has established a set of image quality assessment standards and categorized the images into clear and blurry classes. These classification results are used to determine whether it is necessary to request users to retake the images or whether the images are suitable for digit recognition, thereby enhancing the overall efficiency and accuracy of automated meter readings.

(2) In traditional machine learning algorithms, techniques used for meter digit recognition often fail to meet practical application demands due to slow recognition speeds and lower success rates (approximately 80% to 60%). Therefore, this study aims to explore the use of advanced deep learning models to enhance meter digits' detection and recognition efficiency. Our objective is to improve the consistency of recognition results with human-eye recognition to over 95% and ensure all visible digits are accurately identified. The choice to deploy the YOLO (You Only Look Once) series of models for image recognition tasks is based on their processing efficiency and feature identification advantages. Among the various variants, this research has chosen to implement YOLOv5, primarily considering its data augmentation techniques, lightweight and efficient backbone network architecture, and superior ability to extract and integrate target features. Furthermore, YOLOv5's multi-scale prediction strategy enhances the model's flexibility to adapt to targets of various sizes, further ensuring high efficiency and accurate detection performance. These characteristics make YOLOv5 an ideal choice for the meter digit recognition task in this study.

(3) When dealing with images captured in unrestricted environments, the phase of digit recognition often encounters the challenge of missing digits, a problem that persists even when employing advanced deep learning models, such as the YOLO series. Given the critical importance of thoroughly reading meter digits, the issue of missing digits becomes particularly significant. This is especially true when the visibility of digits is affected by reflections or lighting conditions that are too bright or too dark, rendering some digits unidentifiable. To address this challenge, this study proposes an innovative solution based on the geometric invariants between the centers of digit boxes. Specifically, the method utilizes the cross-ratio between the digit boxes, a value that remains constant even in images with perspective distortions, thereby significantly improving the precision of locating missing digits. This method not only accurately identifies and compensates for missing digits but also enhances the robustness and reliability of the overall digit recognition system. Through this approach, this research aims to substantially increase the accuracy of meter digit recognition under various environmental conditions.

This study explores the limitations of object-recognition deep learning models (such as the YOLO series) under the stringent requirements for high-precision recognition, particularly in applying Automated Reading Machines (ARMs), where complete digit recognition accuracy is imperative. Although traditional models achieve high accuracy, they fall short under the strict conditions of ARM tasks where no omission of target digits is allowed.

Therefore, this paper proposes an algorithm that integrates geometric invariants, specifically the cross-ratio, to detect and correct omissions in digit recognition, thus achieving a complete interpretation of missing digits.

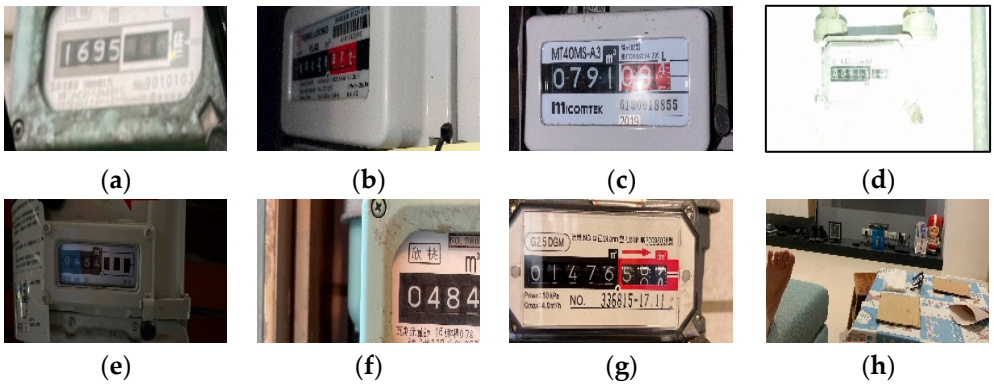


Figure 1. Various types of input images affecting the success of automatic meter readings. (**a**) image blur, (**b**) perspective distortion, (**c**) partial reflection, (**d**) over-exposure, (**e**) poor exposure, (**f**) missing digits, (**g**) intermediate state digits, and (**h**) non-instrument images.

The main contributions of this research include:

(1) Algorithm Innovation: We have developed an algorithm based on the cross-ratio that effectively detects and corrects image omissions caused by perspective distortions, size changes, or rotations. This method is pioneering in the existing literature on ARM and represents a significant advancement for practical applications.

(2) Architecture Optimization: Tailored to the diverse needs of ARM tasks, this study proposes a multi-tiered image processing architecture that considers interactivity, immediacy, and the processing capabilities of devices. This structure can select the most suitable machine learning or deep learning models to handle various types of image blurring, digit recognition, or detection of missing digits at the boundaries.

(3) Experimental Validation: Using datasets collected from real-world applications, the experiments demonstrate that our method surpasses existing technologies in terms of digit integrity and recognition accuracy, providing a more comprehensive and precise framework for applying ARM technology.

Through these innovations, this research not only enhances the accuracy and reliability of digit recognition but also paves the way for the practical use of deep learning in complex applications.

The remainder of this paper is organized as follows. Section 2 is a literature review to understand existing recognition technologies and functionalities. Section 3 describes the automated meter reading system, introducing the system's workflow and image quality classification. Section 4 presents the method of digit recognition using the deep learning model YOLOv5. Section 5 covers the identification and localization of missing digits. Section 6 demonstrates the experimental results, and Section 7 is a conclusion.

2. Related Work

Existing automatic meter reading technologies for instruments can be classified into three types: template matching, feature matching, and deep-learning-based image recognition. Below is a detailed literature review and analysis of these three methods.

(1) Template Matching Methods

The basic idea of this approach is to first establish a set of numeric templates from 0 to 9, before binarization and normalization of the image, and then compare the similarity between the local areas of the image and the preset digit templates. When the similarity is highest and exceeds a set threshold, the corresponding template digit is selected as the final identification result [1–5]. This method is simple to operate and easy to implement,

but its disadvantage lies in the need to design specific templates for different instruments and numeral fonts, and its robustness is poor in complex environments.

(2) Feature Matching Methods

The core of feature-matching methods is extracting features such as color, geometry, or frequency from the image and then using classifiers for digit recognition. Shu et al. proposed a method that identifies digits by analyzing areas in the image where brightness changes correspond to the characteristics of digits and extracting the brightness projection distribution of that area [6]. Oliveira et al. used homomorphic filters to reduce illumination effects and employed K-means clustering to identify the digits on electric meters [7]. Rodríguez and colleagues based their digit detection and recognition on the Hausdorff distance of segmented digits [8]. However, these methods based on manually defined features show unstable performance in practical applications, especially under insufficient lighting conditions.

(3) Deep-Learning-Based Methods

This category typically employs deep learning models, such as Convolutional Neural Networks (CNNs), to learn the recognition features of digits directly from images. This approach does not require manual feature design and has more robust generalization capabilities and robustness, especially when dealing with complex and variable environments. However, deep learning methods usually need many labeled data and significant computational resources. Compared to traditional methods of manually extracting features, CNNs can automatically learn more accurate feature representations, offering greater robustness to environmental changes.

(a) In the realm of digit recognition, researchers such as Li et al. have utilized lightweight CNNs to identify digits on odometers [9], while Gómez et al. applied CNN-based methods for the detection and recognition of digits on utility meters [10].
(b) Chen et al. introduced a deep-learning CNN model designed for processing entire rows of digits [11]. This model performed exceptionally well on synthetic datasets but its performance declined in real-world applications due to the discrepancies between real images and synthetic samples, leading to false detections, missed digits, and low recognition rates.
(c) Laroca et al. proposed a two-stage framework, combining a simplified YOLO structure, multi-task learning, and Convolutional Recurrent Neural Networks (CRNNs) for the detection and recognition of counters [12]. Son et al. also applied YOLO detectors to read gas meter digits [13].
(d) Yang et al. designed a method, combining Fully Convolutional Networks with attention mechanisms, capable of recognizing digits directly without segmentation [14], demonstrating good recognition performance.

Although these methods excel under certain conditions, in practical applications, images captured by users in unconstrained environments often fail to meet the preset requirements of the algorithms, leading to system malfunctions or unstable performance. Poor image quality is often not adequately considered in most of the literature due to blur, perspective distortion, partial reflection, insufficient lighting, missing digits, and intermediate digits. Table 1 lists some representative automatic meter reading studies in recent years and evaluates their consideration of the factors above. The data indicate that despite years of development in automatic meter reading technology, many shortcomings still need to be addressed with new solutions to enhance the system's stability and accuracy.

Table 1. A List of factors influencing recognition success in the AMR-related literature.

Reference	Blurred Images	Persp. Distortion	Partial Reflection	Poor Lighting	Missing Digits	Interm. Digits	Recognition Method
[15] (2013)				Y			SVM
[10] (2017)				Y			Image analysis
[13] (2019)							YOLO v3
[14] (2019)						Y	FCSRN network
[16] (2021)							Lightweight DNN
[17] (2021)							Image analysis
[12] (2021)		Y	Y	Y			YOLOv4
[18] (2021)				Y		Y	CNN
[19] (2022)	Y						Random forest
[20] (2022)						Y	BIF-MSP network
[21] (2022)							YOLOv5
[22] (2022)				Y			Lightweight CNN
[23] (2022)			Y	Y		Y	HRC-mCNN
[24] (2022)							Faster RCNN
[25] (2022)							CBAM network
[26] (2022)	Y			Y			MobileNet V2
[27] (2022)							CycleGAN
[28] (2023)		Y		Y			YOLOv3
[29] (2023)							MobileNet V2
[30] (2024)						Y	Faster-RCNN
[31] (2024)	Y			Y		Y	UNet + CNN
Our	Y	Y	Y	Y	Y	Y	YOLOv5

3. Image Quality Identification

3.1. System Overview

This study aims to meet the practical application needs of automatic digital instrument meter reading by executing the automatic identification of instrument digits. The focus of the research is on proposing a set of strategies to resolve digit recognition failures. The processing workflow is divided into two main stages, as illustrated in Figure 2.

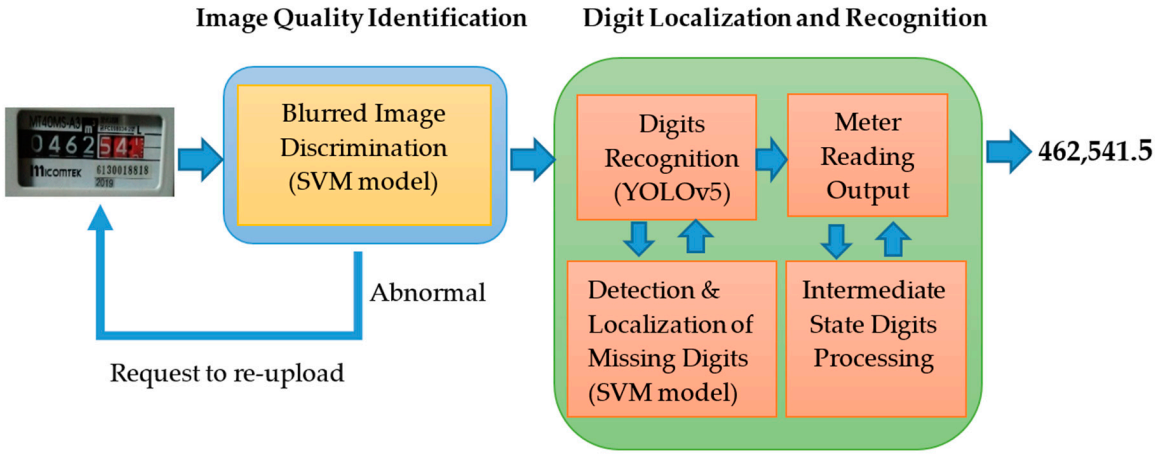


Figure 2. Workflow of the automatic meter reading system.

In the first stage, the system evaluates the quality of images that users upload. For those low-quality images that may hinder effective digit extraction and recognition, the system will prompt users to re-upload and provide suggestions for improving their photography.

Entering the second stage, the system employs the YOLOv5 algorithm to locate and recognize digits. The system will automatically perform error correction and digit restoration for missing and intermediate digits appearing in the recognition results.

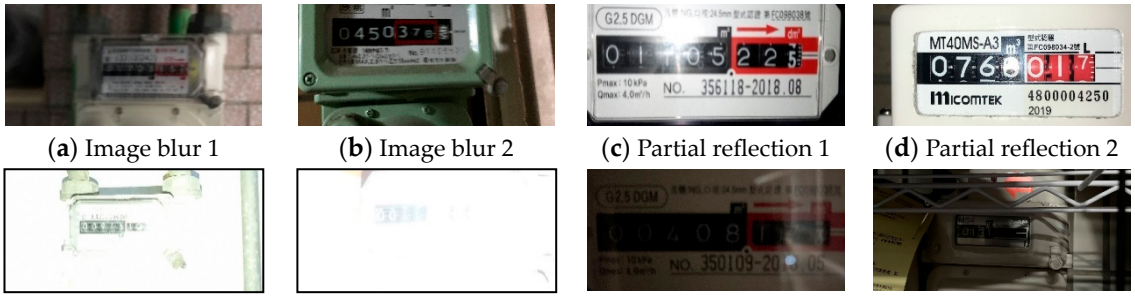
Overall, this research aims to enhance the accuracy and reliability of the automatic digital instrument meter reading system, ensuring high-quality digit recognition results under various conditions.

3.2. Types of Low Image Quality

Based on an in-depth analysis of previously uploaded images and cases of recognition failures, we have identified that the failure of image recognition can mainly be attributed to three categories: image blur, partial reflection, and poor exposure. The following will detail these three scenarios and provide corresponding illustrative examples.

(1) Image Blur

Image blur occurs when the camera shakes or fails to focus accurately during users' shooting of digital instruments. This blur typically manifests as unclear object edges, ghosting, and reduced overall contrast in the image. Figure 3a,b presents typical examples of image blur, from which the described characteristics can be observed.



(**a**) Image blur 1 (**b**) Image blur 2 (**c**) Partial reflection 1 (**d**) Partial reflection 2

(**e**) Over-exposure 1 (**f**) Over-exposure 2 (**g**) Poor exposure 1 (**h**) Poor exposure 2

Figure 3. Types of low-quality or blurred images.

To enhance the performance and reliability of the automatic meter reading system, sufficient attention must be given to such image blur issues. At the system design stage, an image quality assessment mechanism can be considered to evaluate the quality of uploaded images in real-time and prompt users to retake photos upon detecting blur. Moreover, exploring image enhancement technologies, such as deblurring algorithms, might improve the quality of blurred images, thereby enhancing digit recognition accuracy.

(2) Partial Reflection

Partial reflection often results from poor lighting conditions at the meter installation site or the automatic activation of the flash when using the camera's auto mode. This creates reflections on the glass cover of the instrument digits, preventing standard imaging of the digits and thereby affecting digit recognition. Figure 3c,d showcases examples of images under such conditions. Additionally, specific light sources may cause specular reflection in localized areas, further complicating recognition.

(3) Poor Exposure

The casing of digital instruments is usually made of metallic materials, and when users take photos at close range with the flash enabled, this can lead to overexposure, affecting the clarity and recognition of digits. Figure 3e,f provides examples of images in such scenarios. Conversely, underexposure occurs under insufficient lighting conditions without flash, similarly impacting the clarity and recognition of digits. Figure 3g,h showcases examples of images under such conditions.

In summary, poor exposure adversely affects the recognition of meter digits, whether through over- or under-exposure. Therefore, these image conditions must be thoroughly considered in the design and optimization process of the automatic meter reading system. Appropriate image processing techniques should be adopted to improve image quality, thereby increasing the accuracy and reliability of digit recognition.

3.3. Blurred Image Discrimination

To accurately determine the clarity of an image, this study adopts the method from reference [32], proposing a blurred image discrimination algorithm based on the local power spectrum slope. The core idea of this algorithm is to convert the input image into its Fast Fourier Transform (FFT) form, thereby obtaining the image's spectrum, which contains the amplitude (strength) of each frequency component.

It is important to note that since the input image is two-dimensional, the FFT converts it from the spatial domain to the frequency domain. In the spatial domain, the magnitude of color changes can be considered as frequency, and more minor color variations (i.e., lower frequencies) usually indicate a more blurred image. Based on this concept, we conduct spectral analysis on the FFT-transformed image and establish a classifier to distinguish between blurred and clear images.

3.3.1. Blurred Image Discrimination Algorithm

Given an input image $I(x, y)$, where $x = 0, \ldots, M-1$ and $y = 0, \ldots, N-1$, and the image is in color (three channels), the original image $I(x, y)$ is converted into a grayscale (single channel) image $Ig(x, y)$. The Fast Fourier Transform (FFT) is applied to $Ig(x, y)$ to obtain the image $f(u, v)$. The origin of $f(u, v)$ is then moved from the top left corner to the center, resulting in the image $f_s(u, v)$. Since the result after Fourier transformation is in complex form, containing amplitude and phase information of the image at different frequencies, the amplitude image is defined as:

$$f_m(u,v) = \sqrt{\mathrm{Re}(f_s(u,v))^2 + \mathrm{Im}(f_s(u,v))^2} \tag{1}$$

and the phase image as:

$$f_p(u,v) = \tan^{-1}\left(\frac{\mathrm{Im}(f_s(u,v))}{\mathrm{Re}(f_s(u,v))}\right) \tag{2}$$

After obtaining $f_m(u, v)$, taking its discrete square value makes the frequency response of blurriness and clarity in the image more pronounced:

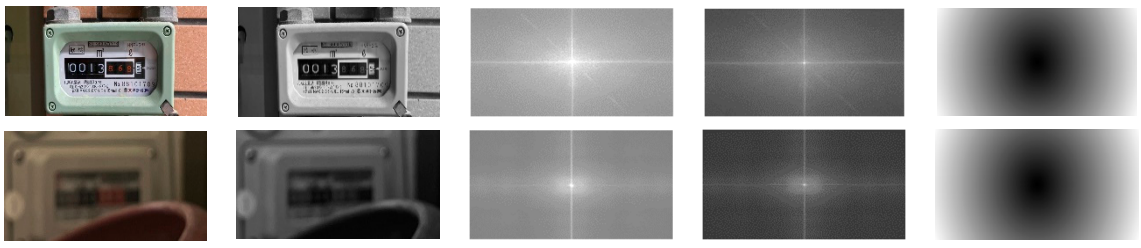
$$f_{ss}(u,v) = \frac{1}{\sqrt{MN}}\|f_m(u,v)\|^2 \tag{3}$$

Next, $f_{ss}(u, v)$ undergoes polar coordinate transformation:

$$f_{ss}(u,v) \equiv f_{sp}(r,\theta) = f_{sp}\left(\sqrt{u^2+v^2}, \tan^{-1}\frac{v}{u}\right) \tag{4}$$

Then, transform $f_{sp}(r, \theta)$ into a one-dimensional array $S(r)$, as shown in Figure 4:

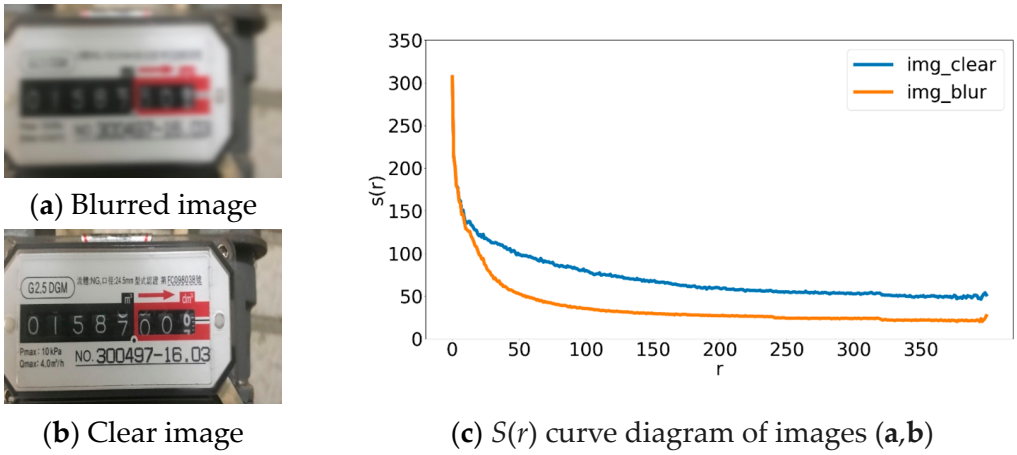
$$S(r) = \sum_{\theta} f_{sp}(r,\theta) \tag{5}$$



(**a**) Original image (**b**) Grey image (**c**) f_m (**d**) f_{ss} (**e**) $S(r)$ (polar coordinate)

Figure 4. The process of converting spatial images into spectrograms. The first row is the example of a clear image, and the second row is a blurred image.

Plotting $S(r)$ on the Y-axis and r on the X-axis creates a spectrum graph with frequency on the X-axis and amplitude on the Y-axis, showing the sum of amplitudes for each radius, as in Figure 5.



(**a**) Blurred image

(**b**) Clear image

(**c**) $S(r)$ curve diagram of images (**a**,**b**)

Figure 5. $S(r)$ Curve Diagram of Clear and Blurred Images.

To validate the effectiveness of the proposed blurred image discrimination algorithm, 50 blurred images and 50 clear images were meticulously selected from a database (as shown in Figure 6) and processed through the proposed discrimination procedure. These 100 images were divided into a training set (70%) and a test set (30%) for training and evaluating a machine learning model.



(**a**) (**b**)

Figure 6. Samples of clear and blurred images used for training and testing. (**a**) Training samples of clear images; (**b**) Training samples of blurred images.

In applying machine learning models, particularly classifiers such as Support Vector Machines (SVMs), selecting feature dimensions significantly impacts computational resource consumption and model computational efficiency. When feature dimensions are too high, it can lead to excessive consumption of computational resources and the potential problem of the curse of dimensionality, thereby affecting the model's performance and accuracy.

In this study, the specific context we faced was that for an image, when the quantity of the local power spectrum slope $S(r)$ reaches 400, meaning one image corresponds to 400 features, this creates a massive pressure on computational resources and operational efficiency. This research employed a method of fitting the $S(r)$ curve with a power function to reduce the feature dimensions. Through this approach, 400 features could be simplified to two parameters of the power function, thereby significantly reducing the feature dimensionality and enhancing the model's computational efficiency while maintaining sufficient recognition accuracy. This method improves model performance and lays a foundation for subsequent optimization and application.

Given data points on the $S(r)$ curve as (x_i, y_i), where $i = 1, \ldots, n$, let the power function be

$$y = a \cdot x^b \tag{6}$$

When fitting the $S(r)$ curve, set r as x and $S(r)$ as y. To simplify the calculation, first take the logarithm of both sides of Equation (6), resulting in

$$\log y = \log a + b \cdot \log x \tag{7}$$

For each data point, substituting into Equation (7) yields n equations. Thus, let

$$\theta = \begin{bmatrix} \log a \\ b \end{bmatrix}, \; Y = \begin{bmatrix} \log y_1 \\ \vdots \\ \log y_n \end{bmatrix}, \; X = \begin{bmatrix} 1 & \log x_1 \\ \vdots & \vdots \\ 1 & \log x_n \end{bmatrix} \tag{8}$$

and rewrite Equation (7) as

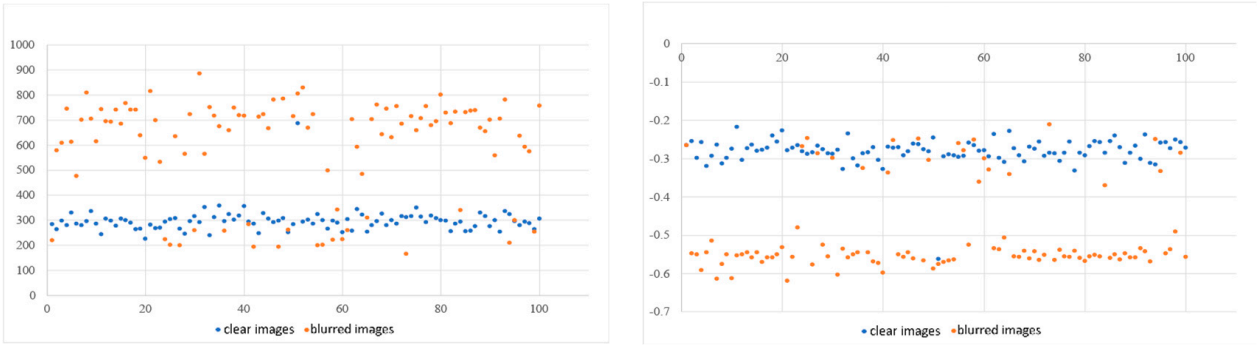
$$Y = X\theta. \tag{9}$$

Using the pseudo-inverse method, we obtain

$$\theta = (X^T X)^{-1} X^T Y \tag{10}$$

where $a = 10^{(\log a)}$.

Through the method described above, the features for each image can be reduced from 400 features of $S(r)$ to two features (a, b). However, as seen from Figure 7, the value of a is significantly greater than b. Therefore, using training data, the mean μ and standard deviation σ of the values a and b are calculated to perform a z-score computation, reducing the disparity between the values a and b. This approach effectively lowers computation and storage costs while also avoiding the problem of overfitting.



(**a**) Parameter a (**b**) Parameter b

Figure 7. Distribution of parameters (a, b) for clear and blurred images.

3.3.2. Non-Linear SVC Model

To establish an efficient and accurate model for image recognition, this study has selected the Support Vector Classifier (SVC) from Support Vector Machines (SVMs) as the classifier. During the model training process, parameters were carefully chosen to ensure the model could balance the proper classification of observations and maximize the classification margin.

In practice, the regularization parameter C was set to 10. This parameter controls the degree of penalty the model applies to misclassifications, with a more significant C value indicating that the model will strive harder to avoid misclassifications, thereby seeking a more significant classification margin. At the same time, the maximum number of iterations was set to 100,000,000 to ensure the model had sufficient time to converge to the optimal

solution. The image dataset was split into training and testing sets, with a ratio of 6:4. Following training, as demonstrated in Figure 8, the accuracy of the non-linear SVC model reached 99.285%.

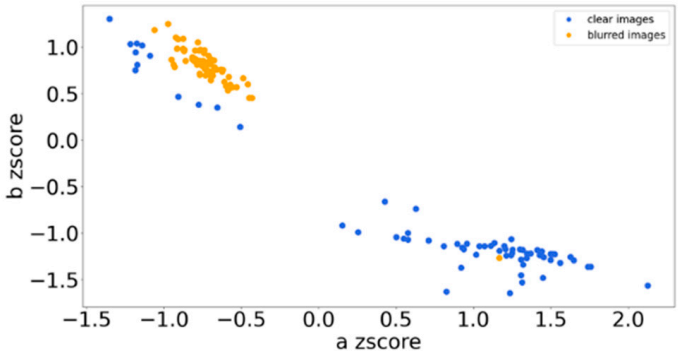


Figure 8. Distribution of training results for the non-linear SVC model.

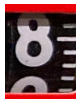
4. Digit Recognition

4.1. Digit Localization and Recognition

The readings of digital instruments are composed of the ten digits from 0 to 9, displayed through a dial mechanism. When the reading is between two consecutive integers, such as between 23456 and 23457, the unit digit displays parts 6 and 7, placing that digit in a transitional state. The appearance of digits in this transitional state significantly differs from the standard shapes of the digits 0 to 9.

To effectively address this issue, we propose a novel processing approach. In addition to the standard digit types from 0 to 9, transitional states between two consecutive integers are considered separate digit types. This means that in addition to the original 10 digit types, an additional 10 transitional-state digit types are introduced, bringing the total number of digit types that need to be recognized and processed to 20. These different types of digits and their corresponding representations are detailed and explained in Table 2. Through this approach, the system can more accurately identify and handle digits in transitional states, enhancing the overall system's recognition accuracy and reliability. This is important for improving automatic meter reading systems' performance and user experience.

Table 2. Digit labeling.

Labels	0	1	2	3	4	5	6	7	8	9
Examples										
Labels	10	11	12	13	14	15	16	17	18	19
Examples										

The YOLO (You Only Look Once) object detection model is highly influential in the current domain of object detection and has become an essential component of deep learning applications. Constructed on the foundation of CNN, YOLO is renowned for its efficiency and accuracy. The mechanism of YOLO involves dividing the input image into numerous small cells and analyzing each cell for potential objects and their locations. In this study, YOLOv5 was chosen as the primary model for executing digit localization and identification tasks after thorough testing and evaluation. This choice was based on a

comprehensive consideration of accuracy and speed to ensure the reliability and efficiency of experimental results.

The output of YOLOv5 includes bounding box information for digits in the image, denoted explicitly as $B(x, y, w, h, c)$. Here, (x, y) represents the coordinates of the center of the bounding box, normalized relative to the size of the input image; w and h indicate the width and height of the bounding box, respectively, again as ratios relative to the input image size; c is the confidence score, reflecting the Intersection over Union (IOU) value between the bounding box and the ground truth. Through this mechanism, YOLOv5 accurately locates digits within the image and provides relevant bounding box information and confidence scores. This lays a solid foundation for subsequent digit recognition and processing tasks, helping to enhance the overall system performance and identification accuracy. Moreover, the provision of confidence scores allows for a more refined assessment and filtering of the model's output, ensuring the reliability of the final results.

4.2. Training Dataset Construction

The dataset established for this research primarily originates from 51,992 images of meters uploaded by users over several years at a gas company. These images were taken in unpredictable field environments, thus encompassing a variety of anomalies such as blur (mainly due to camera movement), scale changes, rotation, reflection, shadows, and occlusions.

The image resolutions are mainly of two types: 3025×1324 and 320×140, depending on the specifications of the image storage systems over the years. Despite the variation in gas meter models installed and replaced over the years, the variation in digits displaying the readings is minimal, and therefore, the appearance and surface text distribution of the gas meters are not considered.

To ensure the dataset's quality, an initial screening was performed to exclude images that do not meet quality requirements, ensuring the digits on the instruments are clear and identifiable without obstructions. This step is crucial for subsequent digit recognition and analysis, ensuring the effectiveness of model training and application accuracy.

For each image in the dataset, detailed manual annotation was conducted to identify the upper-left corner coordinates (x, y) of the instrument digits' bounding boxes, along with their width (w) and height (h), and accurate labels were assigned. Given the sufficient quantity of data and the minor variations in size, tilt, and color of digits under normal shooting conditions, the decision was made not to use data augmentation techniques to expand the training dataset.

Digits usually displayed or in a transitional state were labeled according to the definitions in Table 2. Additionally, confirmed usage readings from the gas company's database were used as another form of labeling. This was to validate the recognition accuracy at both the digit level and the overall reading level, ensuring the accuracy and reliability of the model used. This step is essential for subsequent model training and evaluation, helping us to better understand and optimize the model's performance.

Upon detailed analysis of the dataset, it was observed that the digit "0" had a significantly higher instance count than other digits, which was anticipated. This is because new gas meters often start with a display of "00000000", and the digits on the far left increase more slowly than those in other positions, leading to an imbalanced distribution of sample sizes in the dataset.

A particular handling strategy for the "0" samples was needed during model training to address this issue. This could involve adjusting the weights of the "0" samples or employing techniques to balance the sample sizes of various categories within the dataset, ensuring the model learns the features of each digit category more fairly and accurately. This step is vital for enhancing the model's performance and stability in practical applications.

5. Missing Digit Detection and Correction

5.1. Identification and Localization of Missing Digits

In processing images from unconstrained environments, the digit recognition phase often encounters problems with missing digits, significantly impacting the accurate interpretation of instrument readings. The current literature lacks a comprehensive discussion on solving this issue; hence, in the face of applications in unconstrained environments, this study necessitates designing specific algorithms to overcome this challenge.

When digits in an image are affected by partial reflection or when the environmental lighting is too dark or too bright, even the use of advanced object detection models such as YOLO might not accurately detect digit positions, as shown in Figure 9. Additionally, the non-linear transformation of digit depth caused by perspective projection cannot be resolved merely by relying on the relative distances between digit bounding boxes to determine the presence of missing digits or to pinpoint their specific locations.

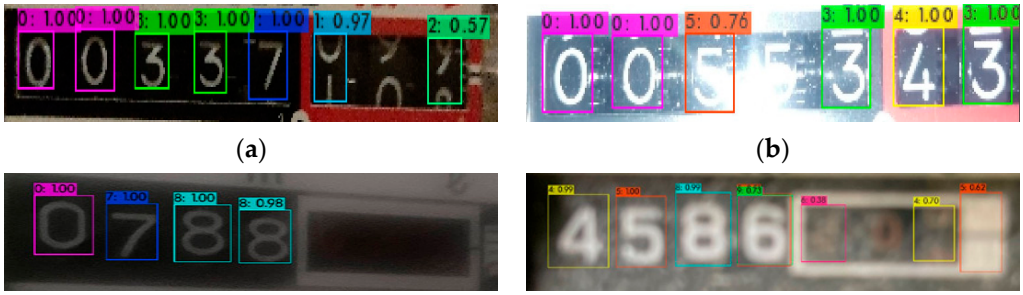


Figure 9. Examples of missing cases in digit recognition performed by YOLOv5. (**a**) Images under normal environmental lighting, (**b**) Images with localized reflection, (**c**) Images with Insufficient ambient light, and (**d**) Out-of-focus captured images.

This study proposes a method based on the distribution relationship of digit bounding boxes to address this issue. By utilizing the geometric invariants between the center points of digit boxes, it is possible to determine whether any digits are missing in the image, further identify which digit is missing, and predict its correct position. This method considers the spatial distribution characteristics of digits in the image, providing an effective solution for accurate digit recognition in complex environments.

5.2. Cross-Ratio

The cross-ratio is an invariant widely applied in geometry, characterized by its resistance to the effects of translation, rotation, and scale changes in an image due to perspective variation. Thus, it is considered a stable invariant in a three-dimensional space. Its basic concept can be described as follows: considering four collinear points A, B, C, and D, a proportional relationship based on the distances between these points can be established as:

$$CR(A,B,C,D) = \frac{\|\overline{AC}\| \cdot \|\overline{BD}\|}{\|\overline{BC}\| \cdot \|\overline{AD}\|} = \text{constant} \tag{11}$$

This proportional relationship remains unchanged under geometric transformations, meaning that even if the object's position, orientation, or size in the image changes, this ratio remains constant. This property makes the ratio a handy tool for analyzing and solving geometric problems, especially in computer vision and image processing.

5.3. Missing Digit Detection

In reading the degrees from digital instruments, the display typically consists of seven to eight digits arranged in a straight line and evenly distributed, forming a structural representation in space. However, in real-world applications, changes in the photography

stance often lead to perspective distortion of digit positions. A more challenging issue is that even with the use of current state-of-the-art deep learning models for digit recognition, it is difficult to ensure that every digit is identified entirely and accurately. We can preliminarily determine if any digits are missing by analyzing the number of identified digit bounding boxes. To precisely capture the degree information on digital instruments, the key lies in identifying which specific digits or digits are missing, thereby implementing targeted digit supplementation and correction measures. This process involves not only a deepening of recognition techniques but also challenges the accuracy and complexity of data processing, holding significant practical importance for enhancing the accuracy and reliability of digit recognition technology.

5.3.1. Filtering the Correct Bounding Boxes

Initially, the bounding boxes output by YOLO, amounting to m, are arranged along the horizontal X-axis of the image. This arrangement orders the bounding boxes from left to right as $B_i(x_i, y_i, w_i, h_i, c_i)$, where $i = 0, \ldots, m$, and $x_i < x_{i+1}$. However, in practice, these outputted frames still exhibit the following three types of errors, necessitating their elimination.

(1) Overlapping Bounding Boxes: More than one bounding box is included for the same digit. To address this, the bounding boxes overlapping with two or more other boxes are first eliminated, typically indicating that the boxes are oversized. For those overlapping with only one other bounding box, a new bounding box is defined by averaging the parameters of the two overlapping boxes, replacing the original boxes.
(2) Non-Meter Digits: Given that meters often display many non-meter digits and text, it is essential to pre-remove these elements. Since meter digits are arranged horizontally, bounding boxes of meter digits have similar y, w, and h values, whereas the parameters of non-meter digit boxes are considered outliers and are thus eliminated.
(3) Bounding Boxes with Low Confidence Scores: To ensure the recognized meter readings possess high robustness, bounding boxes with low confidence scores are eliminated. This is followed by processes to enhance their confidence scores by handling missing digits.

These three mechanisms for eliminating bounding boxes ensure that the remaining boxes are highly reliable. However, they also lead to the possibility of obtaining fewer than n boxes (for example, $n = 8$), resulting in the inability to capture all the meter's digits completely. This necessitates the importance of locating and rectifying missing digits.

5.3.2. Types of Missing Bounding Boxes

When the number of bounding boxes n is insufficient for the number of digits in the meter reading after filtering, initiating a missing bounding box handling procedure becomes necessary. For clarity in this dissertation, and without loss of generality, it is assumed that the meter readings consist of eight digits, and only up to two missing digits are permissible. Otherwise, the user is requested to resubmit the meter image.

Based on the assumptions above, the analysis of the possible scenarios for the appearance of meter digit bounding boxes totals 37 types, as illustrated in Figure 10. The diagonal lines in the figure represent the eight types where only one digit is missing (blue area); the yellow area denotes the 15 types, excluding the first, second, seventh, and eighth digits, as these do not involve the complexity of boundary digit handling and are thus observed separately. The cases involving boundary digit handling are divided into five major categories: missing the first and one of the third to seventh digits (five types, green area), missing the eighth and one of the second to sixth digits (five types, orange area), consecutively missing the first and second digits (red area), consecutively missing the seventh and eighth digits (brown area), and missing both the first and eighth digits (dark green area). Additionally, the type where all eight digits are present is not displayed in the figure.

Missing digits pair	1	2	3	4	5	6	7	8
1	1							
2	(1,2)	2						
3	(1,3)	(2,3)	3					
4	(1,4)	(2,4)	(3,4)	4				
5	(1,5)	(2,5)	(3,5)	(4,5)	5			
6	(1,6)	(2,6)	(3,6)	(4,6)	(5,6)	6		
7	(1,7)	(2,7)	(3,7)	(4,7)	(5,7)	(6,7)	7	
8	(1,8)	(2,8)	(3,8)	(4,8)	(5,8)	(6,8)	(7,8)	8

Figure 10. Types of missing bounding boxes.

5.3.3. Cross-Ratio Feature of Bounding Boxes

Given that on the plane of the instrument panel, the centers of four consecutive numerals (from K_i to K_{i+3}) are equidistantly aligned in a linear arrangement, the condition $||K_iK_{i+1}|| = ||K_{i+1}K_{i+2}|| = ||K_{i+2}K_{i+3}|| = D$ is adhered to. The perspective projection of these four digits' centers onto the image as $k_i(x_i, y_i)$ also lies collinearly; however, due to the nature of perspective projection, they do not remain equidistant. Let the distance between k_i and k_j on the image plane be $d_{ij} = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}$; according to the theory of cross-ratios, it can be derived that:

$$CR_i = \frac{d_{i,i+2} \cdot d_{i+1,i+3}}{d_{i+1,i+2} \cdot d_{i,i+3}} = \frac{\|K_iK_{i+2}\| \cdot \|K_{i+1}K_{i+3}\|}{\|K_{i+1}K_{i+2}\| \cdot \|K_iK_{i+3}\|} = \frac{2D \cdot 2D}{D \cdot 3D} = \frac{4}{3} \tag{12}$$

This implies that if the cross-ratio of the centers of four consecutive bounding boxes in the image approximates 4/3, it also signifies that these four digits are adjacent and none are missing.

In the image, some localization errors exist in the center coordinates of the digit bounding boxes detected by YOLO, resulting in these projection points not lying on a straight line. Thus, a linear fitting is first performed on these k_i points, and then these centers are projected onto the fitted line as p_i, hence expressed in terms of the line parameters as:

$$p_i = p_1 + t_i \cdot S,\ i = 1, \ldots, n, \tag{13}$$

where S denotes the unit direction vector of the fitted line.

Due to the randomness of the positions where digits are missing and the possibility of more than one digit being absent, the instrument digits corresponding to p_i are not necessarily K_i, and their quantity varies with the number of missing digits. In discussing the 37 missing types within this paper, p corresponds to numbers 8, 7, and 6, corresponding to zero, one, and two missing digits, respectively. When calculating the cross-ratio with four consecutive p values, it is possible to compute five, four, and three cross-ratios, respectively, thus denoting $CR_1 = CR(p_1, p_2, p_3, p_4)$, $CR_2 = CR(p_2, p_3, p_4, p_5)$, $CR_3 = CR(p_3, p_4, p_5, p_6)$, $CR_4 = CR(p_4, p_5, p_6, p_7)$, and $CR_5 = CR(p_5, p_6, p_7, p_8)$. This implies that with eight p values, CR_1 to CR_5 can be determined, and with six p values, CR_1 to CR_3 can be obtained.

Based on the 37 types analyzed in Section 5.3.2, the characteristics of their cross-ratios are explored as follows. Additionally, for clarity in explaining the subsequent processing for situation confirmation, this type's handling is denoted as L_{abc}, where 'a' represents the obtained quantity of p, 'b' is the detectable number of missing digits, and 'c' is the position involving boundaries. Here, 'S' denotes the starting digit and 'E' the ending digit, 'B' means that both 'S' and 'E' are included, and 0'' means that neither 'S' nor 'E' are included. In addition, it indicates the necessity for subsequent boundary digit localization and confirmation processing.

(A) No Missing Digits (find p_1 to p_8)

In this type, p_1 to p_8 are obtained, and five cross-ratios can be calculated by sequentially extracting the centers of four consecutive bounding boxes from the image. Since each combination of four points does not involve any missing digits, their cross-ratios are all approximately 4/3 (as per Equation (11)), and these five feature values, as shown in the first row of Table 3 (n = 8), do not require any additional procedures, denoted as L_{80B}.

Table 3. The 37 types of cross-ratio features for different numbers of missing digits and the corresponding positions of the missing digits.

n	Handling Procedure	CR_1	CR_2	CR_3	CR_4	CR_5	Missing Digits	Confirmation Procedure
8	L_{80B}	4/3	4/3	4/3	4/3	4/3	No	
7	L_{71S}	4/3	4/3	4/3	4/3		1	*S*
	L_{710}	3/4	4/3	4/3	4/3		2	
		9/8	3/4	4/3	4/3		3	
		3/2	9/8	3/4	4/3		4	
		4/3	3/2	9/8	3/4		5	
		4/3	4/3	3/2	9/8		6	
		4/3	4/3	4/3	3/2		7	
	L_{71E}	4/3	4/3	4/3	4/3		8	*E*
6	L_{620}	8/5	4/3	4/3			2, 3	
		12/10	6/4	4/3			2, 4	
		9/5	9/8	6/4			2, 5	
		6/4	6/4	9/8			2, 6	
		6/4	4/3	6/4			2, 7	
		16/15	8/5	4/3			3, 4	
		12/10	12/10	6/4			3, 5	
		9/8	9/5	9/8			3, 6	
		9/8	6/4	6/4			3, 7	
		8/5	16/15	8/5			4, 5	
		6/4	12/10	12/10			4, 6	
		6/4	9/8	9/5			4, 7	
		4/3	8/5	16/15			5, 6	
		4/3	6/4	12/10			5, 7	
		4/3	4/3	8/5			6, 7	
	L_{60B}	4/3	4/3	4/3			1, 8	*S*, *E*
	L_{60S}	4/3	4/3	4/3			1, 2	*S*
	L_{60E}	4/3	4/3	4/3			7, 8	*E*
	L_{61S}	6/4	4/3	4/3			1, 3	*S*
		9/8	6/4	4/3			1, 4	*S*
		6/4	9/8	6/4			1, 5	*S*
		4/3	6/4	9/8			1, 6	*S*
		4/3	4/3	6/4			1, 7	*S*
	L_{61E}	6/4	4/3	4/3			2, 8	*E*
		9/8	6/4	4/3			3, 8	*E*
		6/4	9/8	6/4			4, 8	*E*
		4/3	6/4	9/8			5, 8	*E*
		4/3	4/3	6/4			6, 8	*E*

(B) One Missing Digit (find p_1 to p_7)

Here, p_1 to p_7 are obtained and, by sequentially extracting the centers of four consecutive digit bounding boxes from the image, four cross-ratios can be calculated, namely CR_1, CR_2, CR_3, and CR_4. The cross-ratio of four consecutive digits (p_i to p_{i+3}) without any missing digit approximates 4/3; when one bounding box is missing, it spans the range of five actual digits on the meter panel, as shown in the three conditions of Figure 11, where the calculated cross-ratios will be:

$$CR_i^{(a)} = \frac{(t_2 - t_0) \cdot (t_3 - t_1)}{(t_3 - t_0) \cdot (t_2 - t_1)} \approx \frac{\|X_iX_{i+3}\| \cdot \|X_{i+2}X_{i+4}\|}{\|X_iX_{i+4}\| \cdot \|X_{i+2}X_{i+3}\|} = \frac{3D \cdot 2D}{4D \cdot 2D} = \frac{3}{4} \tag{14}$$

$$CR_i^{(b)} = \frac{(t_2 - t_0) \cdot (t_3 - t_1)}{(t_3 - t_0) \cdot (t_2 - t_1)} \approx \frac{\|X_iX_{i+3}\| \cdot \|X_{i+1}X_{i+4}\|}{\|X_iX_{i+4}\| \cdot \|X_{i+1}X_{i+3}\|} = \frac{3D \cdot 3D}{4D \cdot 2D} = \frac{9}{8} \tag{15}$$

$$CR_i^{(c)} = \frac{(t_2 - t_0) \cdot (t_3 - t_1)}{(t_3 - t_0) \cdot (t_2 - t_1)} \approx \frac{\|X_iX_{i+2}\| \cdot \|X_{i+1}X_{i+4}\|}{\|X_iX_{i+4}\| \cdot \|X_{i+1}X_{i+2}\|} = \frac{2D \cdot 3D}{4D \cdot D} = \frac{3}{2} \tag{16}$$

By integrating different missing digit types, the position of the missing digit can be determined from the four calculated cross-ratios, as detailed in the second to ninth rows of Table 3:

(a) The cross-ratio features from the third to the eighth rows are distinct, indicating that the types of digits missing between the second and seventh positions can be directly identified. Thus, L_{710} denotes that a missing digit can be determined from seven bounding boxes without involving boundary digits and does not require additional procedures.

(b) The feature values of the second and ninth rows are identical because the seven obtained bounding boxes are consecutive, with no missing digits in between. Therefore, the corresponding L_{71S} and L_{71E} require a subsequent procedure for locating and confirming the starting and ending digits, which will be elaborated on in Section 5.3.4.

(C) Two Missing Digits (find p_1 to p_6)

(a) Missing digits do not include the first or eighth digits: This indicates that the missing digits do not appear on the instrument's boundaries. The cross-ratio features for rows 10 to 24 in Table 3 are distinct, signifying that it can directly differentiate types with two digits missing between the second and seventh positions. Thus, L_{620} denotes that two missing digits can be determined from six bounding boxes without requiring subsequent boundary digit processing.

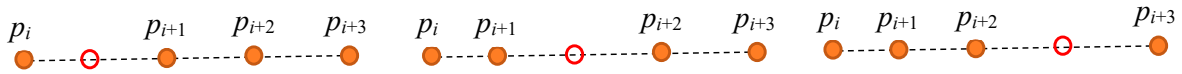
(b) Missing digits include the first digit but not the second or eighth: This implies that one missing digit is at the start, with the other missing digit being one of the third to seventh digits. The cross-ratio features for rows 28 to 32 in Table 3 are distinct, indicating that it is possible to directly identify the presence of a missing digit among the third to sixth digits. Hence, L_{61S} denotes that one missing digit can be determined from six bounding boxes, requiring a subsequent procedure for locating and confirming the starting boundary digit.

(c) Missing digits include the eighth digit but not the first or seventh: This suggests that one missing digit is at the end, with the other missing digit being one of the second to sixth digits. The cross-ratio features for rows 33 to 37 in Table 3 are distinct, indicating that it is possible to directly identify the presence of a missing digit among the second to sixth digits. Thus, L_{61E} denotes that one missing digit can be determined from six bounding boxes, necessitating a subsequent procedure for locating and confirming the ending boundary digit. Additionally, the cross-ratio features here are identical to those in (b), allowing for the identification of the two missing digits as either type (b) or (c) upon confirmation of the boundary digits.

(d) Missing digits include the first and eighth digits: This indicates that the missing digits are the starting and ending digits, presenting the cross-ratio feature of row 25 in Table 3. Therefore, L_{60B} denotes that six bounding boxes are consecutive without any missing digits, requiring further procedures for locating and confirming the starting and ending boundary digits.

(e) Missing digits include the first and second digits: This results in a situation where six consecutive bounding boxes do not have any missing digits, as presented by the cross-ratio feature of row 26 in Table 3. Thus, L_{60S} denotes that six bounding boxes are consecutive without any missing digits, similar to situation (d), requiring additional procedures for locating and confirming the first two consecutive starting digits.

(f) Missing digits include the seventh and eighth digits: Similar to (d) and (e), this results in six consecutive bounding boxes without any missing digits, as indicated by the cross-ratio feature of row 27 in Table 3. Therefore, L_{60E} denotes that six bounding boxes are consecutive without any missing digits, akin to situation (e), necessitating further procedures for locating and confirming the last two consecutive ending digits.



(**a**) Missing digit K_{i+} (**b**) Missing digit K_{i+2} (**c**) Missing digit K_{i+3}

Figure 11. Types of projected distributions for four consecutive digits with one missing digit.

5.3.4. Locating and Confirming the Starting and Ending Digits

For types where either the starting or ending digit is missing, relying solely on the method of cross-ratios does not provide sufficient information for accurate determination. Therefore, this study incorporates the technique of color histograms to assist in the decision-making process. When the missing digit is positioned on the outermost sides of the digit image, it implies that clues for the missing digit can be sought by analyzing the areas at the far left and right ends of the image.

(A) Locating the Starting Digit

In the process of addressing the absence of the first digit (as illustrated in Figure 12a), where the missing digit is K_1, the centers of the first three obtained digits, denoted as p_1, p_2, and p_3, are used to determine the expected position of K_1's projection (corresponding to the line parameter t_0). From Equation (17),

$$\frac{(t_2 - t_0) \cdot (t_3 - t_1)}{(t_3 - t_0) \cdot (t_2 - t_1)} = \frac{4}{3} \tag{17}$$

it follows that

$$t_0 = \frac{t_1(4t_3 - 3t_2) - t_3 t_2}{3t_3 - 4t_2 + t_1} \tag{18}$$



(**a**) Missing digit K_1 (**b**) Missing digit K_n

Figure 12. Distribution types of three consecutive digits with the first or last digit missing.

Inserting the above into Equation (13) allows for the determination of p_0's position.

(B) Locating the Ending Digit

For the procedure of checking for the missing last numeral (as shown in Figure 12b), where the missing digit is K_n (corresponding to the line parameter t_{n+1}), the centers of the

last three obtained digits, i.e., p_{n-2}, p_{n-1}, and p_n, are used to decide the expected position of p_{n+1}. From Equation (19),

$$\frac{(t_n - t_{n-2}) \cdot (t_{n+1} - t_{n-1})}{(t_{n+1} - t_{n-2}) \cdot (t_n - t_{n-1})} = \frac{4}{3} \tag{19}$$

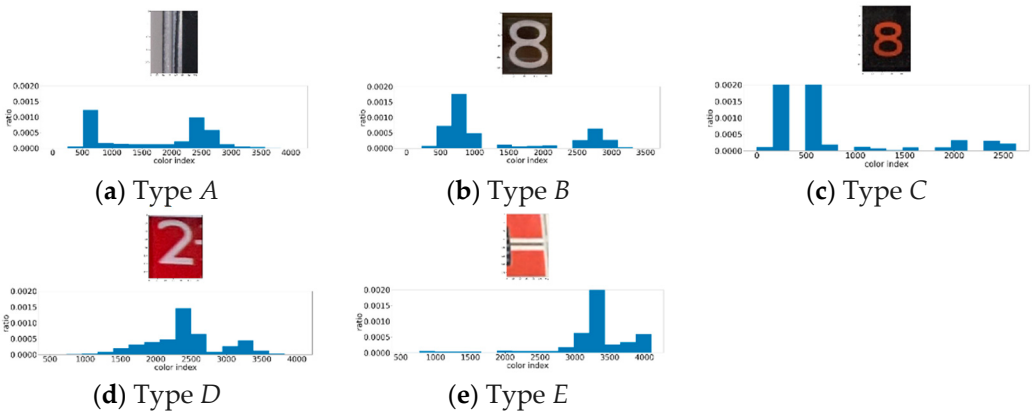
it follows that

$$t_{n+1} = \frac{t_n(4t_{n-2} - 3t_{n-1}) - t_{n-2}t_{n-1}}{t_n - 4t_{n-1} + 3t_{n-2}} \tag{20}$$

Inserting the above into Equation (13) facilitates obtaining p_{n+1}'s position.

(C) Confirmation of Meter Digits

Subsequently, with the anticipated boundary positions of the meter digits as centers, predefined boundary regions are selected for image category recognition. In this study, potential boundary areas of the meter are categorized into five types from *A* to *E*, representing the area of the meter panel to the left of the starting digit, between the starting digit and the second digit, the ending digit, the second digit to the left of the ending digit, and the meter panel area to the right of the ending digit, as illustrated in Figure 13. Each category possesses a specific color distribution, and by comparing these distribution patterns, we can accurately confirm and identify the missing digits.



(**a**) Type *A* (**b**) Type *B* (**c**) Type *C*

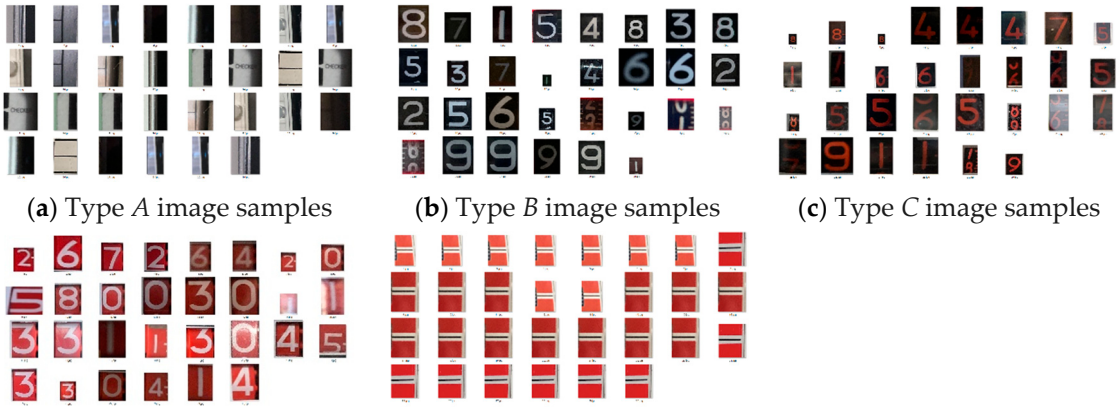
(**d**) Type *D* (**e**) Type *E*

Figure 13. Boundary image frames of the five types and their color histograms.

In this research, color histogram analysis of the boundary areas extracted from gas meter images was conducted to assist in identifying and addressing missing digits. The specific workflow is as follows. Initially, a simple quantization process was applied to the RGB channels of the image. By dividing the RGB values of each pixel by 16 and rounding off, the original 256-level pixel values were reduced to 16 levels, lowering computational complexity while retaining sufficient color information.

By combining the cross-ratio method with color histograms, the accuracy of digit recognition was enhanced, and a practical solution was also provided for addressing exceptional cases within images from the digit meters. The application of this comprehensive method demonstrates that combining multiple techniques and approaches can yield more reliable and robust performance in complex and variable real-world application settings.

The histogram data of these five types of samples (as shown in Figure 14) were used as features, and an SVM model was trained. During the training process, the Gaussian kernel (Radial Basis Function, RBF) was selected as the kernel function, and the regularization parameter (*C*) was set to 6 to achieve optimal model performance. This method improved the accuracy of missing digit identification, and a practical solution was provided for handling digit meter images in unconstrained environments.



(**a**) Type *A* image samples (**b**) Type *B* image samples (**c**) Type *C* image samples

(**d**) Type *D* image samples (**e**) Type *E* image samples

Figure 14. Five types of training samples used in the SVM classification model for recognizing digits at both ends of the meter.

5.3.5. Re-Processing of Missing Digits

The occurrence of missing digits often signifies that the image of the digit may be affected by factors such as blurring, shadows, partial reflections, etc., leading to the loss of recognition or low confidence scores in localization. After identifying the region of the missing digit using the previously described methods, image enhancement procedures are initially conducted to mitigate the effects of these interfering factors, followed by re-inputting into the YOLOv5 model for digit recognition through secondary detection.

This integrated approach, utilizing both machine learning models and deep learning detection techniques, effectively resolves issues related to missing end digits on gas meters, enhances the accuracy of digit identification, and ensures correct recognition of digits even in cases of missing elements. Consequently, this improves the overall system performance and reliability.

5.3.6. Meter Reading Output

After the determination of missing digits is concluded, the output of the meter reading will proceed. When the identified digit category L is one of 0 to 9, the designated number is L. If the digit category is one of 10 to 19 and the last digit, the designated number is L.5; otherwise, it is specified as L-10. Therefore, the final set of n numbers, L_i, where $i = 1, \ldots, n$, are consolidated on the meter to produce the output reading.

$$D = \sum_{i=1}^{n} L_i \cdot 10^{n-i} \tag{21}$$

6. Experiment Results and Discussion

6.1. Experiment 1: Image Blur Discrimination

This section focuses on an in-depth analysis of the system's capability in discriminating image blur. Starting with the identification effectiveness of images taken in different shooting environments, a qualitative analysis method combined with image examples is used to showcase the identification results.

(A) Dataset

We selected 1295 images (800 clear, 495 blurred) for testing to evaluate the system's performance in recognizing blurred images. After completing the tests, the results were carefully analyzed.

(B) Performance in Blurred Image Recognition

Using the blurred image recognition method described in Section 3.3, the trained SVM classifier was tested on 1295 images, yielding the following confusion matrix (Figure 15):

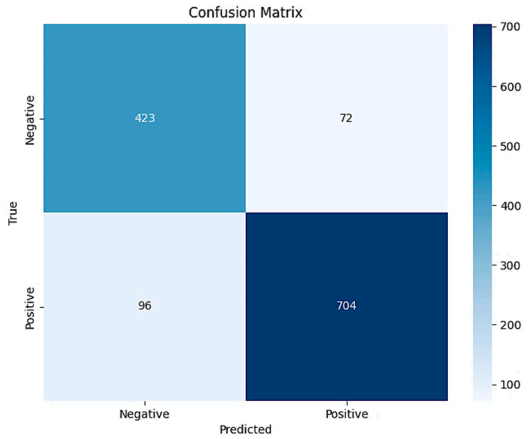


Figure 15. The confusion matrix of Experiment 1.

Here, the correctly identified images include 423 true positives (TP), where the original blurred images were correctly predicted as blurred, and 704 true negatives (TN), where the original clear images were correctly predicted as clear. The misidentified images include 72 false positives (FP), where the original blurred images were predicted as clear, and 96 false negatives (FN), where the original clear images were predicted as blurred.

Thus, the performance in blurred image identification is described as follows:

(1) Accuracy: 87.03 $\pm$ 1.83% (with confidence interval) of all test samples were correctly classified. This relatively high accuracy demonstrates the model's overall good recognition capability.

(2) Precision: Indicates that among the samples predicted as correct by the model, 85.45% $\pm$ 0.98% were indeed correct. This means the model has a high certainty in making correct predictions, with a relatively low misjudgment rate.

(3) Recall: The model predicted 81.50 $\pm$ 1.08% of all truly correct samples. This suggests that the model may have missed some samples that should have been correctly predicted.

Although the accuracy and precision are relatively high, a lower recall rate might indicate that the model could be overly conservative in some situations, leading to missing some correct samples. Further analysis is needed to determine in which scenarios the model is prone to errors and to adjust the model to improve recall.

6.2. Experiment 2: Meter Digit Recognition

This section presents the experimental results of the gas meter digit recognition model. It begins by showcasing the recognition results of various meter digits through qualitative analysis and image examples, illustrating the model's performance in recognizing different ones. Subsequently, accuracy is quantitatively evaluated, and results are organized in the corresponding tables.

(A) YOLOv5 Configuration

This study utilized YOLOv5 as the digit recognizer, with the architectural parameters set for the experiment outlined in Table 4. Key hyper-parameter settings include a learning rate of 0.01; momentum as the optimization method for deep learning training; mosaic augmentation, which randomly crops four training images and merges them into one new image; and mix-up augmentation, creating new training samples through the linear combination of two images and their labels; the training cycle was set for 100 epochs; batch size per training was 2.

(B) Results of Different Meter Digit Recognitions (Label 0–19)

This study validated the recognition performance of individual digits on gas meters. Recognition experiments were conducted using images from the image dataset, covering 9332 digits. Detailed data for the recognition results of different digits (0–19) can be found

in Table 5. The confusion matrix is shown in Figure 16. The results show an accuracy rate of 98.23%. This achieved the target accuracy rate of 95% and demonstrated that images processed for blur identification meet the image quality standards of YOLOv5 for digit recognition.

Table 4. YOLOv5 Architectural Parameters.

Setting	Value
Weights	yolov5x.pt
Config File (cfg)	models/yolov5x.yaml
Data File	data\mydata.yaml
Hyperparameters	
Learning Rate	lr0: 0.01
Momentum	momentum: 0.937
Mosaic Augmentation	mosaic: 1.0
Mixup Augmentation	mixup: 0.0
Number of Epochs	100
Batch Size	2
Image Size (imgsz)	640

Table 5. Results of different meter digit recognitions (Labels 0–19).

Label	0	1	2	3	4	5	6	7	8	9
Correct	1773	895	873	765	640	910	692	672	666	782
Samples	1800	907	879	780	656	921	701	681	672	788
Accuracy	98%	98%	99%	98%	97%	98%	98%	98%	99%	99%
AP	0.989	0.991	0.993	0.983	0.976	0.991	0.987	0.987	0.991	0.994
Label	10	11	12	13	14	15	16	17	18	19
Correct	50	43	37	52	39	87	33	47	45	66
Samples	63	46	37	54	45	98	34	49	53	68
Accuracy	79%	93%	100%	96%	86%	88%	97%	95%	84%	97%
AP	0.877	0.935	1	0.963	0.867	0.888	1	0.959	0.9	0.97
mAP	mAP ≈ 0.976									

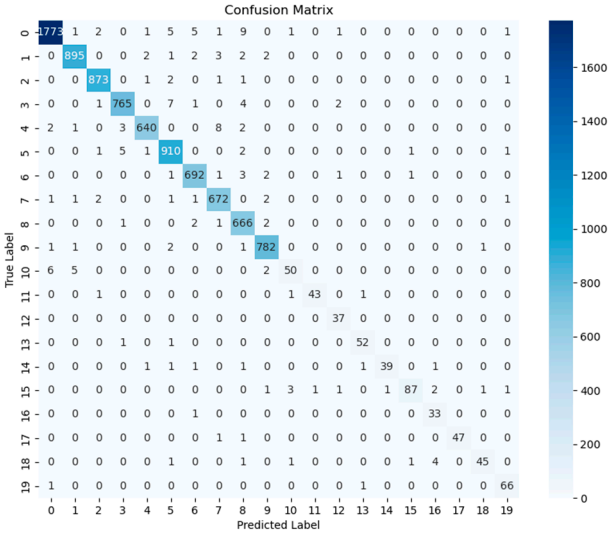


Figure 16. The confusion matrix of meter digit recognition in Experiment 2.

In this study, we conducted an in-depth analysis of digit recognition performance. For the complete range of digits, i.e., labels 0 to 9, the recognition accuracy reached an impressive 98.67%. However, when focusing on the mid-range digits, i.e., labels 10 to

19, the accuracy dropped to 91.22%. Notably, for the digit labeled "10", the recognition accuracy was significantly lower than the rest, at only 79%. This may be attributed to the sparser information within that digit frame, increasing the difficulty of its recognition. These results indicate that there may be significant differences in recognition performance across different digit ranges or specific digits in the field of digit recognition.

(C) Gas Meter Digit Recognition Results

This study also validates the performance of recognizing the position and type of digits in gas meter images, especially those processed after blur image identification. Given that some images mistakenly identified as clear still exist, it was necessary to examine whether these escaped blurred images would significantly impact the final digit recognition results. A recognition experiment was conducted using 1300 images from the image dataset, resulting in the following confusion matrix (refer to Figure 17):

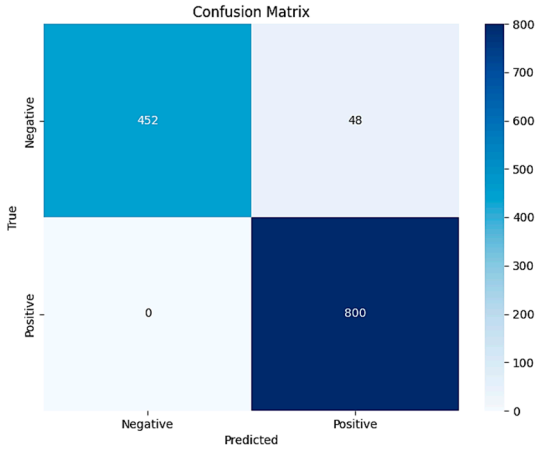


Figure 17. The confusion matrix of meter digit recognition for validation in Experiment 2.

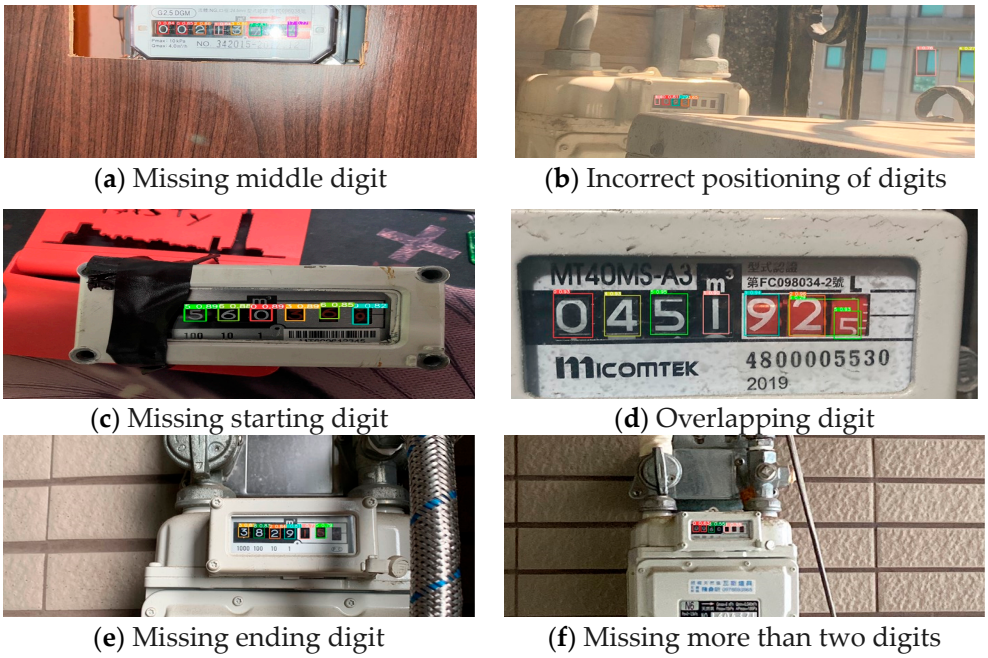
Here, 452 true positives (TP) represent the original blurred images correctly recognized, and 800 true negatives (TN) represent the original clear images predicted to be correctly recognized. There are 48 false positives (FP), original blurred images not recognized correctly, and 0 false negatives (FN), representing clear images that failed to be recognized correctly.

Thus, the final performance of digit recognition is described as follows:

(1) Accuracy: Indicates that 6.31 $\pm$ 0.52% of all test samples were correctly classified. This relatively high accuracy demonstrates the model's excellent overall recognition capability.
(2) Precision: Indicates that among the samples predicted as correct by the model, 90.40 $\pm$ 0.82% were indeed correct. Although this precision rate is relatively high, there is still a gap compared to the accuracy rate, suggesting that the model may have produced false-positive results in some cases.
(3) Recall: Indicates that the model successfully predicted all truly correct samples without missing any (100.00 $\pm$ 0.00%). This ideal recall rate shows the model's excellent ability to cover all correct samples.

6.3. Experiment 3: Digit Misses and Recognition Correction

This section discusses the experimental results regarding digit misses and recognition correction. Different types of missing digits are analyzed to showcase the system's capability to compensate for misses and recognition accuracy, culminating in the results being presented through qualitative analysis and image examples. Although YOLOv5 is recognized as a highly efficient object detection method, it cannot guarantee the complete detection of all digits in practice. Figure 18 demonstrates the six most common types of missed digit detections in applying gas meter digit recognition within this study.



(**a**) Missing middle digit (**b**) Incorrect positioning of digits

(**c**) Missing starting digit (**d**) Overlapping digit

(**e**) Missing ending digit (**f**) Missing more than two digits

Figure 18. Types of missing digits in digit detection using YOLOv5.

(1) Error Analysis of YOLOv5 Digit Detection

In this experiment, an algorithm analysis for blur judgment was conducted on 1300 test images. After the initial phase of blur image filtering, 776 images were deemed to be of sufficient quality. These 776 images were then subjected to YOLOv5 digit detection to examine the status of missing or overlapping digits, revealing 33 images with errors, resulting in an error rate of 4.25%. Figure 19 displays the distribution of numeral counts detected by the YOLOv5 algorithm across various images. Different types of gas meters primarily contain seven, eight, or nine digits. If the detection results in less than eight digits, it indicates a missing digit phenomenon; conversely, if the detection results in nine or more digits, it suggests a potential overlap of digits. This implies that the cross-ratio method proposed in this study must be employed for incorrect digit detection results to identify missing digits or eliminate overlapping digits, thereby reducing the error rate of YOLOv5 digit detection.

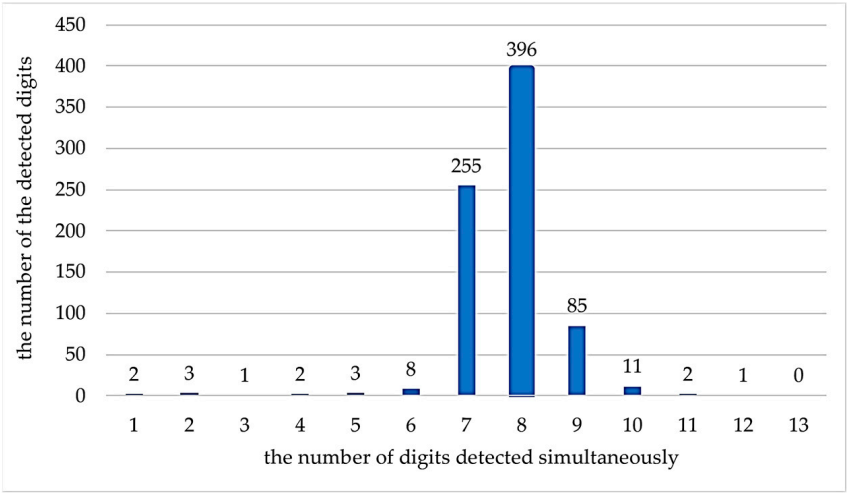
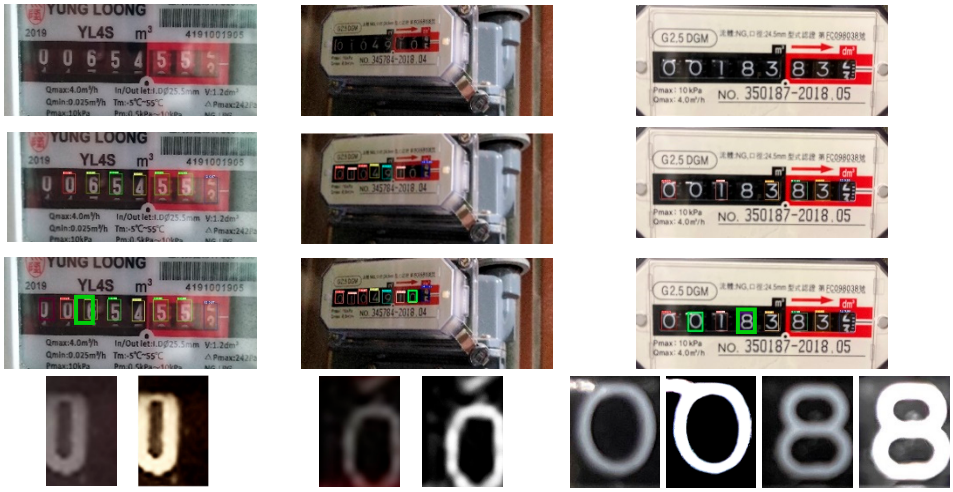


Figure 19. Distribution of detected digit counts.

(2) Locating and Repairing Missing Digits

Out of the 33 images identified by YOLOv5 as having missing digits, subsequent processing involving localization, confirmation, and image enhancement repaired 21 images successfully. The remaining 12 could not be detected, resulting in a 63.64% correction rate for digit recognition errors. Figure 20 shows the selected digit regions after the missing digits were localized, with the Intersection over Union (IoU) metric used to evaluate the accuracy between the model-predicted bounding boxes and the actual regions. The experimental results show an average IoU of 82.67%, indicating excellent capabilities in locating and repairing missing digits.



(**a**) Missing starting digit (**b**) Missing middle digit (**c**) Missing two digits

Figure 20. Examples of locating and confirming missing digits. Each column, from top to bottom, displays the original image, YOLOv5 digit recognition results (showing three types of missing types), locations of missing digits (green boxes), and the original image with predicted digit boxes and its enhancement processing.

7. Conclusions

This study focuses on optimizing the application of automatic meter reading technology in civil public utilities, specifically for reading residential gas meters. The method proposed aims to enhance the performance of AMR technology. We introduce an end-to-end AMR approach to address the identification challenges in unconstrained scenarios. This method specifically addresses common factors of identification failure, such as image blur, perspective distortion, and poor lighting, and has proven its efficacy in experiments.

In the initial phase, the system conducts a quality check on images uploaded by users, particularly distinguishing blurred images. When the image quality is poor, the system prompts the user to re-upload, ensuring the stability of the recognition accuracy rate. Through this strategy, the correct recognition rate for blurred images reached 87.03%.

The digits on the meters are located and identified using YOLOv5 technology. Since many blurred images are filtered out in the preliminary stage, the accuracy rate of digit recognition can reach 98.23%, fully meeting the requirements of practical applications. However, considering that YOLOv5 might still miss some digits, this study further introduces a method based on cross-ratios (projective invariants) specifically to address the issue of missed digits in digit recognition. Through experiments on different types of missing digits, this strategy has shown a missing rate of 4.2%, and the success rate after correction exceeds 60%, enhancing the method's flexibility and reliability.

In collaboration with a gas company, we have established a large dataset containing 52,000 images. After training with deep learning techniques, a highly efficient model was developed to achieve high precision in digit recognition. Experimental results confirm the

system's exceptional performance in automatic meter reading tasks and demonstrate its potential in practical applications.

For future work, we plan to conduct extensive testing to validate the applicability of our solution in relevant environments. A particular area for improvement is the text detection step, which prevents the capture of irrelevant information. Given the variety of digital meters from different eras and models commonly found in public utilities, we aim to enhance our ability to interpret various types of meters. Therefore, we will expand our dataset to include a broader range of meter types and identify the serial numbers of these meters to develop a more intelligent automatic billing system, which will record both the unit readings and user-specific serial numbers. This approach can be implemented on embedded devices, installed on mobile devices for use by meter reading personnel, or deployed in the cloud for data reading and review. The proposed solution can potentially advance the development of smart water or energy management infrastructure, offering significant benefits for the automation of meter readings.

Author Contributions: Conceptualization, J.-H.H. and Y.-H.C.; Methodology, J.-H.H., Y.-H.C. and Y.-L.T.; Validation, J.-H.H. and Y.-L.T.; Software: Y.-H.C.; Formal analysis, J.-H.H. and Y.-H.C.; Investigation, J.-H.H. and Y.-H.C.; Resources, J.-H.H. and Y.-L.T.; Data curation, Y.-H.C.; Writing—original draft preparation, J.-H.H. and Y.-L.T.; Writing—review and editing, J.-H.H. and Y.-L.T.; Supervision, Y.-L.T.; Project administration, Y.-L.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: All the data are available in the manuscript.

Acknowledgments: The authors thank Xintao Gas Company for providing photographs of gas meters with extensive numerical labels for this study.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zhao, S.; Li, B.; Yuan, J.; Cui, G. Research on remote meter automatic reading based on computer vision. In Proceedings of the 2005 IEEE/PES Transmission & Distribution Conference & Exposition: Asia and Pacific, Dalian, China, 15–18 August 2005; pp. 1–4.
2. Huang, X.; Nie, D.; Yong, H.E.; Wang, Q. Research on digital instrument automatic recognition technology of substation inspection robot. *Mech. Electron.* **2018**, *36*, 58–62.
3. Elrefaei, L.A.; Bajaber, A.; Natheir, S.; AbuSanab, N.; Bazi, M. Automatic electricity meter reading based on image processing. In Proceedings of the 2015 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT), Amman, Jordan, 3–5 November 2015; pp. 1–5.
4. He, Z.; He, Y.; Yang, Y.; Gao, M. A low-cost direct reading system for gas meter based on machine vision. In Proceedings of the 2017 12th IEEE Conference on Industrial Electronics and Applications (ICIEA), Siem Reap, Cambodia, 18–20 June 2017; pp. 1189–1194.
5. Kashid, S.G.; Pardeshi, S.A. Intelligent water metering system: An image processing approach. In Proceedings of the 2014 First International Conference on Networks & Soft Computing (ICNSC2014), Guntur, India, 19–20 August 2014; pp. 289–294.
6. Shu, D.; Ma, S.; Jing, C. Study of the automatic reading of watt meter based on image processing technology. In Proceedings of the 2nd IEEE Conference Industry Electron Applications, Harbin, China, 23–25 May 2007; pp. 2214–2217.
7. Oliveira, D.M.; Cruz, R.D.S.; Bensebaa, K. Automatic numeric characters' recognition of kilowatt-hour meter. In Proceedings of the 5th International Conference on Signal Image Technology Internet Based Systems, Marrakesh, Morocco, 29 November–3 December 2009; pp. 107–111.
8. Rodríguez, M.; Berdugo, G.; Jabba, D.; Calle, M.; Jimeno, M. HD-MR: A new algorithm for number recognition in electrical meters. *Turk. J. Electr. Eng. Comput. Sci.* **2014**, *22*, 87–96. [CrossRef]
9. Li, C.; Su, Y.; Yuan, R.; Chu, D.; Zhu, J. Light-weight spliced convolution network-based automatic water meter reading in smart city. *IEEE Access* **2019**, *7*, 174359–174367. [CrossRef]
10. G'omez, L.; Rusinol, M.; Karatzas, D. Cutting Sayre's knot: Reading scene text without segmentation. In Proceedings of the 13th IAPR International Workshop Document Analysis Systems (D.A.S.), Vienna, Austria, 24–27 April 2018; pp. 97–102.
11. Chen, G.; Zifeng, H.U.; Zheng, C. Fast recognition algorithm of digital instrument numbers based on feature detection. *China Test* **2019**, *45*, 146–150.
12. Laroca, R.; Barroso, V.; Diniz, M.A.; Gonçalves, G.R.; Schwartz, W.R.; Menotti, D. Convolutional neural network for automatic meter reading. *J. Electron. Imaging* **2019**, *28*, 013023. [CrossRef]

13. Son, C.; Park, S.; Lee, J.; Paik, J. Deep learning-based number detection and recognition for gas meter reading. *IEEE Trans. Smart Process. Comput.* **2019**, *8*, 367–372. [CrossRef]
14. Yang, F.; Jin, L.; Lai, S.; Gao, X.; Li, Z. Fully convolutional sequence recognition network for water meter number reading. *IEEE Access* **2019**, *7*, 11679–11687. [CrossRef]
15. Edward, V.C.P. Support vector machine based automatic electric meter reading system. In Proceedings of the IEEE International Conference on Computational Intelligence and Computing Research, Tamilnadu, India, 26–28 December 2013; pp. 1–5.
16. Sharma, A.K.; Kim, K.K. Lightweight CNN based meter digit recognition. *J. Sens. Sci. Technol.* **2021**, *30*, 15–19. [CrossRef]
17. Agrawal, A.; Jain, S.; Deep, V.; Sharma, P.; Mishra, S. Meter reading using O.C.R. for comfort and cost optimization. In Proceedings of the 2021 5th International Conference on Information Systems and Computer Networks (ISCON), Mathura, India, 22–23 October 2021; pp. 1–4.
18. Naim, A.; Aaroud, A.; Akodadi, K.; El Hachimi, C. A fully AI-based system to automate water meter data collection in Morocco country. *Array* **2021**, *10*, 1–15. [CrossRef]
19. Mahatya, K.J.B.; Jati, F.W.; Irawan, B.; Hasibuan, F.C. Random forest implementation in prepaid electric meter recognition. *J. Comput. Eng. Prog. Appl. Technol.* **2022**, *1*, 33–40. [CrossRef]
20. Liao, J.Y.; Hsieh, J.W.; Ma, C.M. Automatic meter reading based on bi-fusion M.S.P. network and carry-out rechecking. *IEEE Access* **2022**, *10*, 96710–96719. [CrossRef]
21. Nigar, N.; Muhammad Faisal, H.; Kashif Shahzad, M.; Islam, S.; Oki, O. An offline image auditing system for legacy meter reading systems in developing countries: A machine learning approach. *J. Electr. Comput. Eng.* **2022**, *2022*, 4543530. [CrossRef]
22. Hsu, T.C.; Tsai, Y.H.; Chang, D.M. The vision-based data reader in IoT system for smart factory. *Appl. Sci.* **2022**, *12*, 6586. [CrossRef]
23. Xiu, H.; He, J.; Zhang, X.; Wang, L.; Qi, Y. HRC-mCNNs: A hybrid regression and classification multibranch CNNs for automatic meter reading with smart shell. *IEEE Internet Things J.* **2022**, *9*, 25752–25766. [CrossRef]
24. Concio, M.L.W.; Bernardo, F.S.; Opulencia, J.M.; Ortiz, G.L.; Pedrasa, J.R.I. Automated water meter reading through image recognition. In Proceedings of the TENCON 2022—2022 IEEE Region 10 Conference (TENCON), Hong Kong, China, 1–4 November 2022; pp. 1–6.
25. Li, W.; Wang, S.; Ullah, I.; Zhang, X.; Duan, J. Multiple attention-based encoder-decoder networks for gas meter character recognition. *Sci. Rep.* **2022**, *12*, 10371. [CrossRef] [PubMed]
26. Chong, Y.J.; Chua, K.H.; Babrdel, M.; Hau, L.C.; Wang, L. Deep learning and optical character recognition for digitization of meter reading. In Proceedings of the 2022 IEEE 12th Symposium on Computer Applications & Industrial Electronics (ISCAIE), Virtual Conference, 21–22 May 2022; pp. 7–12.
27. Filinger, A. Improving Automatic Meter Reading Using Data Generated from Unpaired Image-to-Image Translation. Doctoral Dissertation, Hochschule für Angewandte Wissenschaften Landshut, Landshut, Germany, 2022.
28. Imran, M.; Anwar, H.; Tufail, M.; Khan, A.; Khan, M.; Ramli, D.A. Image-based automatic energy meter reading using deep learning. *Comput. Mater. Contin.* **2023**, *74*, 203–216. [CrossRef]
29. Carvalho, R.; Melo, J.; Graça, R.; Santos, G.; Vasconcelos, M.J.M. Deep learning-powered system for real-time digital meter reading on edge devices. *Appl. Sci.* **2023**, *13*, 2315. [CrossRef]
30. Sun, L.; Yuan, Y.; Qiao, S.; Qi, R. Detection of water meter digits based on improved faster R-CNN. *J. Comput. Commun.* **2024**, *12*, 1–13. [CrossRef]
31. Zhao, S.; Lu, Q.; Zhang, C.; Ahn, C.K.; Chen, K. Effective recognition of word-wheel water meter readings for smart urban infrastructure. *IEEE Internet Things J.* **2024**, *11*, 17283–17291. [CrossRef]
32. Liu, R.; Li, Z.; Jia, J. Image partial blur detection and classification. In Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, AK, USA, 23–28 June 2008; pp. 1–8.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.