

Research on Instrument Digital Recognition Technology Based on Yolov8

Zhiyi Zhang

School of Computer Science and Engineering
Xi'an Technological University
Xi'an, China
E-mail: zhangzhiyi@st.xatu.edu.cn

Abstract—To improve the efficiency and reliability of industrial digital instrument reading, this paper proposes an automatic recognition method based on YOLOv8. Aiming at the low efficiency and poor robustness of traditional manual and rule-based methods, YOLOv8 is used to accurately detect the digital display area of instruments, benefiting from its fast inference speed and strong adaptability to complex environments. Subsequently, image preprocessing operations, including grayscale conversion, denoising, binarization, and morphological processing, are applied to enhance digital features, and individual digits are segmented and recognized using a ResNet34 classifier. Experimental results show that the proposed method achieves a detection accuracy of over 98.20% for digital display areas and a digit recognition accuracy of over 98.60%, demonstrating good robustness and practical applicability in complex industrial scenarios.

Keywords—Industrial Digital Instrument Recognition; Yolov8; Resnet34; Opencv

I. INTRODUCTION

As an important measuring instrument, the instrument is widely used in the industrial field. Common industrial instruments are mainly divided into pointer instruments and digital display instruments. This paper will focus on digital display instruments [1]. With the progress of artificial intelligence technology, many enterprises have turned to automated management in order to achieve cost savings and efficiency improvement. In this context, the reading of traditional instruments is gradually changing to intelligence. Using computer vision technology to realize the automatic identification of industrial instruments has become the trend of the times. This method can not only replace the complex manual operation, but also reduce human error, which provides a solid foundation for the automatic monitoring of industrial instruments.

Manual reading not only faces the risk of error, but also may lead to cheating. In the traditional workflow, the staff need to manually record or audit the instrument data, and then upload it to the system to generate reports. The whole process relies on human work as an intermediary, which makes the reading generation and report generation process vulnerable to human factors, resulting in unstable data reliability. Especially in industrial projects, instrument readings are directly related to the quality of the project and the company's economic interests. Manual auditors or inspectors have become a key factor affecting quality in the current model. Any improper operation or intentional tampering may bring serious consequences. Therefore, the automatic identification technology of industrial instruments based on artificial intelligence has become an effective alternative to solve such problems. The automatic identification technology of digital instruments proposed in this paper is designed to meet this demand. By eliminating the reading link in the quality inspection process of manual intervention projects, the accuracy and consistency of data collection are improved. In the process of instrument digital reading, the use of artificial intelligence technology not only reduces human error, but also reduces the risk of human intervention, which can effectively guarantee the stability and reliability of engineering quality.

At present, the mainstream methods of digital recognition algorithms can be roughly divided into three categories: threading method [2-4], pattern matching method [5] and image recognition algorithm based on deep learning [6-9]. Since digital meters are usually displayed by seven-segment digital tubes, the threading method plays an important role in the identification application

of digital meters. The threading method recognizes numbers by scanning the contour and structural characteristics of digital characters and using the continuity of the path. This method has the advantage of fast recognition and is suitable for scenes with clear and simple character contours. However, the threading method has a high requirement for the integrity of the character information in the display area, and may lead to a high error rate when the conditions such as character display angle, shooting angle and ambient light are not ideal, which limits its universal applicability in complex practical applications [10]. The pattern matching method compares the number to be identified with the preset template by establishing the template library of each number, and obtains the recognition result by calculating the difference between the two [11]. Although the pattern matching method is intuitive to operate, its accuracy is greatly affected by image quality and shooting angle. In the face of noise or blurred images, the recognition performance is poor and lacks strong robustness.

In recent years, the image recognition method based on deep learning has been widely used in digital instrument recognition tasks and has gradually become the mainstream. In [12], a method based on regional convolutional neural network is proposed. This method first locates the digital region, and then uses the YOLOv4 algorithm to accurately locate and identify the detected digital region, which greatly improves the efficiency and accuracy of digital detection. In [13], a digital binarization method based on color model and multi-threshold segmentation was proposed to extract the edge information of the digital region more accurately. Then, the YOLOv4 network was used to identify the numbers. Finally, the accuracy of 87.98 % mAP was obtained on the test set. Reference [14] proposed a digital meter reading method based on YOLOv8 target detection and digital recognition network, and gave two different digital recognition schemes. One is the combination of traditional digital image processing and deep learning. Firstly, the characters are segmented, and then the Le Net-5 network is used to classify and recognize the numbers. The other is based on the improved

CRNN model to identify the number directly, and the accuracy of the two is 96.63 % and 91.58 % respectively. Compared with the traditional methods, the image recognition method based on deep learning has strong adaptability in dealing with complex images, different illumination and angles, and can learn more features from the data, and has stronger generalization ability. Therefore, although the accuracy of deep learning methods may be slightly lower than that of traditional methods in some common scenarios, its excellent environmental adaptability and robustness make it a more ideal solution in complex and changeable environments.

Through the above algorithm comparison, combined with its advantages and disadvantages, this paper proposes an algorithm based on deep learning for instrument digital recognition in specific application scenarios. Firstly, YOLOv8 is used to detect the target of the instrument image and accurately locate the digital display area of the instrument. Then, the image is preprocessed to remove background noise and enhance the recognizability of the number. Finally, the preprocessed image is digitally segmented, and each segmented number is sent to the ResNet [15] network for recognition. The experimental results show that the proposed method has higher recognition accuracy, and shows higher accuracy and robustness than the traditional method. This method not only effectively improves the recognition rate of instrument numbers, but also maintains a good recognition effect in complex environments, and has strong practical application potential.

II. OVERALL ALGORITHM FRAMEWORK

The overall process of the instrument digital recognition algorithm is shown in Fig. 1. The algorithm is mainly composed of three core steps:

- Extraction of instrument display area based on YOLOv8 model: The YOLOv8 model is used to detect the target of the input image, accurately extract the digital display area of the instrument, and collect the color of the instrument display according to the color characteristics of the area;

- **Instrument character segmentation:** After obtaining the display area, a series of preprocessing is performed on the image, including grayscale, Gaussian filtering denoising [16], and binarization processing according to the color of the instrument digital display. Then, the morphological opening and closing operation is used to further remove the noise, while ensuring that the digital strokes of the seven-segment digital tube remain continuous. Finally, the contour extraction technology is used to identify the digital contour in the image, and the contour of each number is selected for character segmentation.
- **Instrument number recognition:** The segmented numbers are sent to the ResNet34 image classification network for recognition, and the recognition results are spliced to finally output the display number of the instrument.

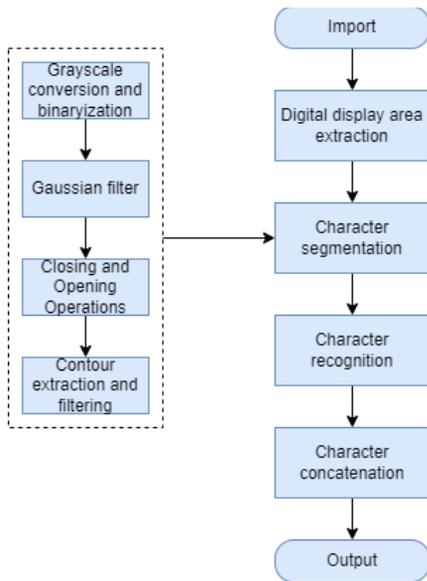


Figure 1. Algorithm flow chart

III. POSITIONING AND EXTRACTION OF INSTRUMENT DIGITAL

A. Digital display area extraction

Due to the complexity of the application scenario environment, this paper uses YOLOv8 target detection technology to identify and locate the instrument digital display area in the image to be detected. As a one-stage target detection

technology, YOLOv8 has the advantages of fast detection speed and high precision. It can efficiently and accurately locate the digital display area (ROI) of the instrument, effectively eliminate the interference of irrelevant features in the environment, and ensure the accuracy of digital recognition. While extracting ROI, YOLOv8 can also record the color information of the display characters. In the subsequent grayscale, by enhancing the characteristics of the corresponding RGB channel according to the actual character color, the effect of grayscale is guaranteed, and the recognition of the numbers in the image is effectively improved. In addition, based on the proportion of ROI in the original image, the algorithm can adaptively adjust the image preprocessing parameters, such as the size of the structural elements of the opening and closing operations, so as to ensure the continuity and integrity of the digital strokes.

The introduction of YOLOv8 not only effectively avoids the interference of external environment on digital recognition, improves the accuracy of instrument digital recognition, but also provides an adaptive strategy for image preprocessing. This strategy can dynamically adjust the image processing parameters according to different environmental conditions, thereby enhancing the adaptability of the algorithm and further optimizing the overall image processing and digital recognition effects.

B. Image preprocessing

In practical applications, the digital display colors of industrial instruments are diverse, and the collected images are usually RGB color images, which may contain different degrees of color noise. In order to effectively suppress noise interference and reduce the workload of data acquisition, the image is grayed and binarized to reduce the influence of color difference.

Take the green digital tube instrument as an example. The specific parameter configuration is shown in Table 1. The grayscale conversion adopts a weighted fusion method, with weights set as $\omega_R = 0.30$, $\omega_G = 0.59$, and $\omega_B = 0.11$. This conforms to the human eye's more sensitive perception of brightness in the green channel,

which can highlight the digital strokes and weaken the influence of background color differences. Subsequently, Gaussian filtering is used for noise reduction, with $ksize=(5,5)$ and $\sigma=0$. OpenCV automatically estimates the filtering strength to balance the smoothing effect and edge preservation. Finally, `cv2.THRESH_BINARY` is used for threshold binarization, with $thresh=120$ and $maxval=255$, to effectively separate the digital area from the background, providing clear input for subsequent character segmentation and recognition.

TABLE I. SUMMARY TABLE OF KEY PARAMETERS IN THE IMAGE PREPROCESSING STAGE

Step	Parameter	Value
Gray processing	ω_R	0.30
	ω_G	0.59
	ω_B	0.11
Gaussian filtering	$ksize$	(5, 5)
	σ	0
	$thresh$	120
binaryzation	$maxval$	255
	$type$	<code>cv2.THRESH_BINARY</code>

First, gray processing. This paper uses a custom weighted RGB-to-grayscale conversion, optimized for the actual digital display colors to achieve an ideal grayscale effect. This processing step significantly reduces the complexity of color information, and the processed results are shown in Fig. 2 (a). Then, in order to further remove the noise in the image, the Gaussian filtering technique is used to denoise the image. Gaussian filtering helps to smooth the details in the image and remove small-scale interference signals. Finally, the gray image is binarized by using the simple threshold method in OpenCV. By setting a custom threshold, the part of the image whose pixel value is greater than the threshold is converted to white, while the part less than the threshold is converted to black. The processed results are shown in Fig. 2 (c), which can distinguish the important features in the image more clearly and enhance the identifiability of the digital region.



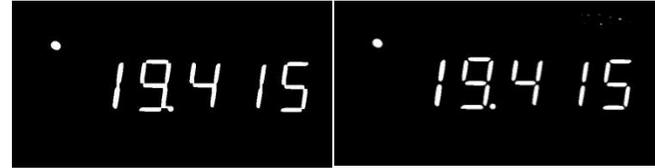
(a) Original drawing

(b) Graying



(c) Gaussian filter

(d) Binaryzation



(e) Closed operation

(f) Union operation

Figure 2. Instrument digital display area schematic diagram

C. Image denoising

In this paper, there will be noise interference in the image acquisition process, which makes the effect of image recognition worse. Common filtering noise reduction methods include Gaussian filtering, median filtering [17], mean filtering, and bilateral filtering [18]. In this study, the main concern is the irrelevant noise that may interfere with digital extraction in the digital display area of the instrument. Therefore, this paper uses a combination of Gaussian filtering and morphological opening and closing operations to remove noise [19].



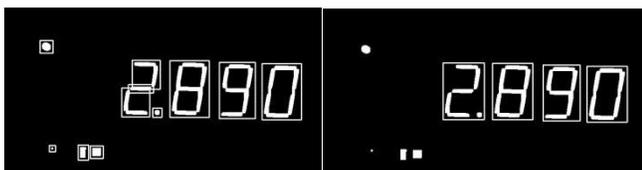
Figure 3. Gaussian filtering effect image

Firstly, between the grayscale and binarization processing in the image preprocessing stage, the Gaussian filter is used to blur the image. Through experimental verification, Gaussian filtering can effectively suppress irrelevant features in the image. Gaussian filtering reduces the intensity of noise by smoothing the image while retaining the main structural features in the image. The effect is shown in Fig. 3, where the right side is the binarized image after Gaussian filtering, which obviously removes some noise. Then, the morphological opening and closing operation is applied to further remove the remaining noise. The open operation can effectively remove the small

interference elements by performing the corrosion operation first and then the expansion operation. The closed operation first performs the expansion operation and then the corrosion operation, which helps to fill the small gap in the digital strokes and ensure the integrity and consistency of the strokes. The combination of opening and closing operations makes the irrelevant interference in the image effectively removed, and the seven-segment digital strokes in the image are further enhanced to facilitate the subsequent extraction and recognition of digital contours. The effect is shown in Fig. 2 (e) and Fig. 2 (f). The image after opening and closing operation is clearer and the boundary of the digital region is more coherent.

D. Contour extraction

Before the recognition of digital meter readings, each number needs to be segmented separately to reduce the burden on the data set. In order to accurately identify and segment the digital region in the image, this paper uses `cv2.findContours()` in the OpenCV library function to detect all contours in the image. The function can effectively extract the connected region from the binary image and return the point set of the contour. Thanks to the preceding morphological opening and closing operations, the digit strokes are more coherent, enabling more accurate contour extraction. Specifically, the `cv2.findContours()` function first identifies the edges in the image and extracts each contour. Then, in order to facilitate subsequent digital segmentation and recognition, each contour needs to be transformed into a regular rectangular box. Therefore, this paper uses `cv2.BoundingRect()` function, which can calculate a minimum rectangular box according to the coordinates of the contour and transform it into a matrix form, which can effectively eliminate the irregular area caused by noise or interference, so as to facilitate the extraction of each digital area in the image.



(a) Before treatment (b) After processing

Figure 4. Matrix box filtering effect image

In the processed binary image, although preprocessing such as grayscale, filtering, and opening and closing operations have been performed, there may still be noise, which will introduce unnecessary contours in the contour extraction process, as shown in Fig. 4 (a). These noise profiles not only increase the amount of calculation, but also may lead to false recognition in subsequent recognition steps. Therefore, the extracted contour must be screened and optimized to remove irrelevant noise and ensure the accuracy of digital recognition.

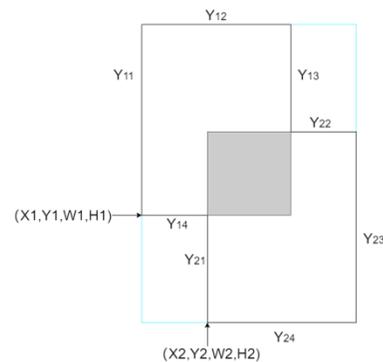


Figure 5. Diagram showing overlapping and intersecting of matrix boxes

In this paper, combined with the precise digital display area positioning of YOLOv8, the contour extraction process is further optimized. After YOLOv8 determines the digital display area, the number is usually located in the center of the vertical direction of the digital display area, and the height of the number is greater than half of the length of the entire area. Therefore, based on these features, we can first filter out the contours located in the central region of the image according to the spatial position of the contours, and focus on these contour regions. However, in some cases, even through location filtering, we may still encounter the problem of digital stroke discontinuity. For example, the number '2' in Fig. 4(a) is broken due to strokes, generating multiple contours. These contours may partially overlap, making it impossible for the numbers to form a complete connected region. In this case, we use the method of contour merging: by analyzing the boundary information of the contour, that is, (x, y, w, h) , we calculate the four boundary line equations of the

rectangular contour, and then compare them with the boundaries of other contours one by one.

If the two contour boundaries have a focus, then select the outermost boundary and regenerate a new rectangular contour, so as to ensure that a number only corresponds to a rectangular box. The schematic diagram of the specific process is shown in Fig. 5, where the gray area represents the overlapping part between the two contours. On this basis, we first remove the focus boundaries (Y13, Y14, Y21, Y22) in the overlapping area, and then extend the remaining borders until they intersect to obtain four new focuses to form a new border, that is, the blue border in the figure.

Finally, the contour that meets the requirements is selected according to the approximate digital height, and the noise contour below the specified height is eliminated. After the above series of processing, we obtained a clear and accurate digital contour. As shown in Fig. 4(b), all noise and invalid contours have been removed, which provides a guarantee for the stability of subsequent digital recognition.

IV. CHARACTER RECOGNITION

In this paper, ResNet network is used to recognize the extracted characters. Through a series of operations such as digital display area positioning, image preprocessing and character segmentation, the processed image is sent to the ResNet network for efficient and accurate digital recognition. Compared with the traditional threading method, ResNet not only shows higher recognition accuracy in practical applications, but also has stronger generalization ability. Through training for digital recognition tasks, ResNet can achieve efficient learning on relatively small data sets and avoid dependence on a large amount of labeled data.

ResNet is a deep convolutional neural network with residual learning at its core. With the increase of the number of layers, the traditional deep network is easy to encounter the problem of gradient disappearance or gradient explosion, which affects the training effect of the model. In order to solve this problem, ResNet introduces the idea of residual blocks, allowing signals to cross one or more layers and pass directly to the

following layers, avoiding the loss of information. In ResNet, the input signal is transmitted through two paths: one is through conventional operations such as convolution and activation functions, and the other is directly transmitted through jump connections, and finally the outputs of the two paths are added. This structure can effectively alleviate the degradation problem in deep network training. Through residual learning, the network can focus on learning residuals, that is, the difference between the current layer and the skipping layer, rather than directly learning complex mappings.

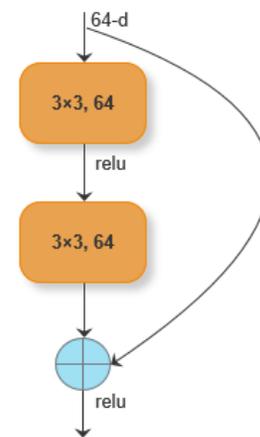


Figure 6. Schematic diagram of the ResNet34 residual module structure

The ResNet34 used in this paper is a variant of the ResNet series. It has a 34-layer network depth. Compared with ResNet50 or ResNet101, ResNet34 belongs to a shallower network structure. It stacks the convolutional layers through multiple residual modules, allowing the model to transmit signals through jump connections in each residual block, avoiding the training difficulty of traditional deep networks. Each residual module includes two or more convolutional layers, and the network's expression ability is enhanced by techniques such as Batch Normalization and ReLU activation function. The residual module structure of ResNet34 is shown in Fig. 6. Because of its relatively simple structure, ResNet34 has a small amount of parameters and calculation, which is suitable for tasks with limited resources, and can achieve good performance on relatively small data sets. The introduction of ResNet34 can efficiently classify the input digital images, and has stronger

anti-interference ability and adaptability than the traditional methods.

After the above image processing, including instrument digital display area positioning, image preprocessing, grayscale, denoising and character segmentation, the input RGB image has been converted into a clear binary digital image. The processed image removes unnecessary interference, and the boundary of each number is clearly visible. Then the segmented single digital image is input into the ResNet34 network for classification, and the final classification result is output. The specific digital recognition process is shown in Fig. 7. After the previous image preprocessing, the input image is first cropped and scaled to a size of 224×224 to meet the input requirements of the network. Then the image enters the first convolutional layer (the size of the convolution kernel is 7×7 , the number of channels is 64, the

step size is 2, and the filling is 3), and the input image is converted into a 56×56 feature map. After batch normalization and ReLU activation function processing, the feature map enters the maximum pooling layer, and the size is further reduced to 28×28 . Then, the feature map is passed through multiple BottleNeck modules of ResNet34 in turn. In this process, the size of the feature map gradually decreases, and the depth gradually increases, so that the network can extract more abundant feature information. Finally, the feature map is reduced to a size of 7×7 . Then, using the global average pooling operation, the output feature map of the last BottleNeck module is compressed into a vector, reducing the number of parameters and reducing the risk of overfitting. Finally, the vector is mapped to the output space through the full connection layer to obtain the final digital recognition result.

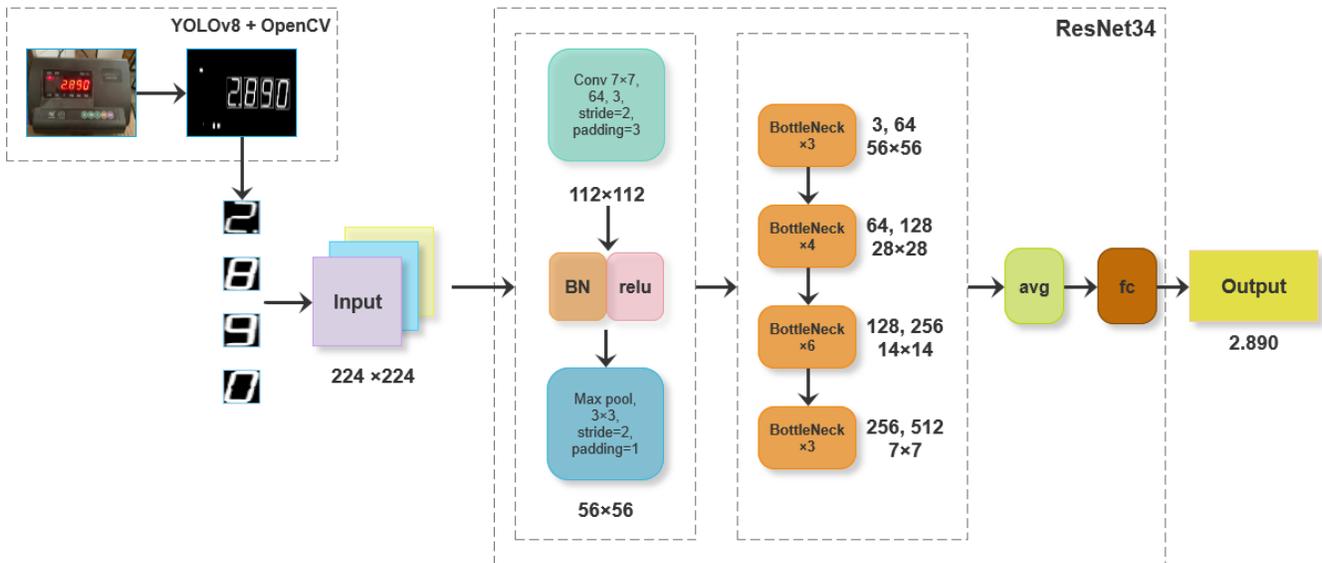


Figure 7. Overall architecture diagram of the algorithm model

V. EXPERIMENTAL RESULT ANALYSIS

A. Data set

The data set used in this paper is provided by the research group, including 3000 digital instrument images. Some of the instrument data are shown in Fig. 8.

These images are collected in the production environment, which meets the needs of practical application scenarios. The numbers in the image

are segmented separately, and the specific data set is distributed as shown in Table 2. Before the experiment, we divided the data set into training set and test set in a ratio of 3: 1. The data contained in the training set is used for the training of the algorithm model to fully learn the digital features. The test set is used to objectively evaluate the performance of the model to test its generalization ability on unseen data. To further ensure robustness, we also applied data

augmentation techniques such as rotation, scaling, and brightness adjustment during training to simulate real-world variations.



Figure 8. Instrument data diagram

TABLE II. TABLE SHOWING THE DISTRIBUTION OF DIGITAL INSTRUMENT DATA SETS

Category	Quantity / Number of Pieces
0	780
1	574
2	445
3	468
4	430
5	605
6	501
7	412
8	521
9	477

B. Test result

Aiming at the single-target positioning of the digital display area of the instrument, this experiment first uses the original data image to train the YOLOv8 model for accurately identifying the digital display area of the instrument. After several rounds of training and parameter adjustment, the final YOLOv8 model achieved a recall rate and accuracy rate of more than 98.20 % in this task.

Specifically, in complex environmental scenarios (Fig. 9), there may be frequent interference information such as pipelines and text labels around the instrument, which can easily lead to false detections or missed detections. YOLOv8 is still able to accurately locate the digital display area, indicating that the model has learned the relatively stable geometric layout and boundary patterns of the digital display area, thereby

reducing the influence of background noise on the detection decision. In the exposure scenario (Fig. 10), the excessively high brightness of the digital display area will cause local details to be lost or edges to become blurred. Traditional threshold or edge methods often have difficulty stably extracting the ROI; however, YOLOv8 is still able to provide a relatively complete detection box, indicating that it has certain robustness against the texture degradation caused by brightness saturation and is able to complete positioning using the remaining structural cues. Under different lighting conditions (Fig. 11), the overall brightness and contrast of the image change, and the gray difference between the digital display area and the background may weaken. The model still maintains effective detection, indicating that it has good adaptability to lighting changes.



Figure 9. Schematic diagram of the recognition result of the digital display area of instrument data in complex environments



Figure 10. Illustration of the recognition result of the digital display area for exposed instrument data



Figure 11. Schematic diagram of the recognition results of the digital display area of instrument data under different lighting conditions

After achieving precise positioning of the digital display area, this paper performs

preprocessing such as grayscale conversion, noise removal, and binaryzation on the ROI, and classifies and recognizes the individual digital images obtained through segmentation. This paper uses ResNet34 as the digital classifier to train and infer the extracted digital samples. After 10 rounds of iterative optimization, the overall accuracy of the classifier on the test set reaches over 98%, and the inference time for a single digit remains stable at 0.09 – 0.10 seconds, demonstrating good real-time performance. The specific experimental results are shown in Table 3 and Fig. 12.

TABLE III. PERFORMANCE SUMMARY OF THE RESNET34 DIGITAL CLASSIFIER ON THE TEST SET

Category	Quantity / Number of Pieces	Correct quantity / number of pieces	Reasoning speed / s	Accuracy rate / %
0	195	195	0.10	1.00
1	143	143	0.09	1.00
2	111	107	0.10	0.97
3	117	117	0.10	1.00
4	107	102	0.09	0.96
5	151	146	0.09	0.97
6	125	125	0.09	1.00
7	103	103	0.10	1.00
8	130	130	0.10	1.00
9	119	115	0.10	0.96

From the distribution of classification results by category, the recognition accuracy rates of the numbers 0, 1, 3, 6, 7, and 8 all reached 100%, indicating that the model has a stable feature representation ability for most seven-segment digital tube structures; while the accuracy rates of the numbers 2 and 5 were 97%, and those of the numbers 4 and 9 were 96%. Further analysis revealed that the minor misidentifications of the numbers 2 and 5 were mainly related to the stroke deformation caused by morphological opening and closing operations, especially for the seven-segment structure of the number 5, which was relatively special in some samples. Excessive expansion might lead to local stroke adhesion or rupture, thereby reducing the separability. This problem can be alleviated by introducing more refined morphological parameters for adaptive or structural constraints for the "5" category structure.

The misidentifications of the numbers 4 and 9 were more likely to occur during the character cropping stage. When the segmentation box was too tight, causing the strokes to be truncated, the key structural information of the number was incomplete, thus leading to category confusion. By optimizing the cropping boundary, the accuracy rate can be further improved.

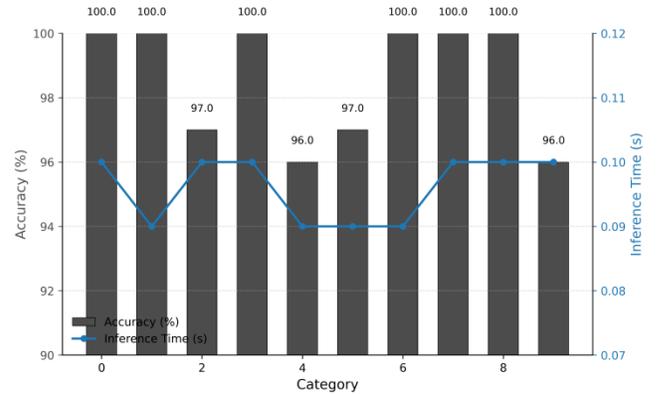


Figure 12. Performance of ResNet34 Digital Classifier on Test Set

From the comprehensive chart results, it can be concluded that ResNet34 maintains stable and fast inference speed while ensuring high accuracy, verifying the effectiveness and engineering usability of this classification module in the industrial instrument digital recognition task; and the current errors mainly concentrate on two controllable factors, "morphological deformation" and "cropping deviation", providing a clear direction for subsequent targeted optimization.

C. Contrast experiment

In the industrial instrument digital recognition task, to verify the effectiveness of the method proposed in this paper, a comparative experiment was conducted with traditional image processing methods, YOLO series detection algorithms, and general cloud-based recognition interfaces. The results are shown in Table 4. From the table, it can be seen that the recognition accuracy of the traditional threading method is 96.5%, and the test speed is 0.72 s, but its robustness is relatively low; OpenCV + threading method improves the recognition accuracy and speed based on the traditional method, with an accuracy rate of 98.7% and a test speed of 0.5 s, and its robustness is at a medium level. However, both of these traditional

methods highly rely on manual rules such as threshold segmentation and contour extraction, and are sensitive to changes in lighting, background interference, exposure, and digital deformation. Therefore, they have weak generalization ability and insufficient cross-scenario adaptability, making it difficult to meet the long-term stable application requirements in complex industrial environments.

TABLE IV. PERFORMANCE COMPARISON BETWEEN THE RESNET34 DIGITAL CLASSIFIER AND COMMON CHARACTER RECOGNITION METHODS

Method	Recognition accuracy rate	Test the speed	Robustness
The threading method	96.5%	0.72	lower
OpenCV+The threading method	98.7%	0.5s	medium
YOLOv4	87.63%	0.62	higher
YOLOv5	89.62%	0.59	higher
YOLOv8	92.23%	0.49	higher
Baidu Intelligent Cloud Digital Recognition Interface	93.00%	0.08s	higher
our	98.6%	0.32	higher

In contrast, the YOLO series of algorithms exhibit strong robustness in target detection under complex background conditions. For instance, both YOLOv4 and YOLOv5 demonstrated high robustness in this task. However, their recognition accuracy rates were only 87.63% and 89.62% respectively. The main reason for this is that the YOLO series methods are better at completing target detection and positioning tasks. However, for the recognition of digital tube numbers, simply relying on the target detection results often makes it difficult to precisely distinguish similar digital structures, especially when the digit strokes are broken, stuck together, or partially missing. Moreover, YOLO-like methods usually require larger-scale labeled data and more complex training parameter tuning processes, which have the drawbacks of high training difficulty and difficulty in accurately outputting digital

categories. Therefore, in the task of precise digital recognition, their advantages are not obvious.

In addition, general recognition interfaces, such as the Baidu Intelligent Cloud Digital Recognition Interface, achieved a recognition accuracy rate of 93.00% in this experiment, with a test speed of 0.08 seconds. They have certain universality and ease of use, but there are still issues with insufficient recognition accuracy in specific industrial instrument scenarios.

Through a comprehensive comparative analysis, it can be seen that the method proposed in this paper achieves a relatively ideal balance among recognition accuracy, processing speed, and system robustness: the overall recognition accuracy reaches 98.6%, the single test time is approximately 0.32 seconds, and it shows stable and reliable recognition results in the complex environment data as shown in Figures 9-11 in the previous text. The core advantage of this method lies in using YOLOv8 to precisely locate the digital display area, combined with targeted image preprocessing, character segmentation, and ResNet34 classification recognition, effectively reducing the reliance on fixed thresholds and manual rules in traditional methods, while compensating for the shortcomings of the YOLO series algorithms in the fine discrimination of digital categories. Therefore, the proposed method demonstrates stronger environmental adaptability and higher recognition reliability in complex industrial scenarios, has good engineering practicability, and is more suitable for application in automatic monitoring and intelligent management tasks of instrument digital data in actual production environments.

VI. CONCLUSIONS

With the rapid development of digital image processing and deep learning technologies, automatic recognition of industrial instrument readings has become increasingly important in modern industrial production, energy monitoring, and quality management. To address the limitations of traditional manual reading and rule-based methods in complex environments, this paper proposes an automatic digital instrument

positioning and recognition method based on YOLOv8 and ResNet34.

The proposed method first employs the YOLOv8 target detection algorithm to accurately locate the digital display area of instruments, effectively reducing the influence of environmental factors such as illumination variation, background clutter, and noise. Subsequently, a series of image preprocessing operations, including grayscale conversion, Gaussian filtering, binarization, morphological opening and closing, and contour extraction, are applied to enhance the digital features and achieve accurate segmentation of individual digits. Finally, a ResNet34 classifier is utilized to recognize the segmented digits, improving both recognition accuracy and robustness.

Experimental results demonstrate that the proposed method can achieve fast and accurate digital recognition under complex conditions, exhibiting strong anti-interference capability and good adaptability to variations in lighting, perspective, and noise. In practical applications, particularly in weighing systems, the method effectively reduces the risks associated with manual operation, such as misreading and intentional tampering, thereby improving the reliability, transparency, and automation level of quality monitoring.

Moreover, the proposed approach shows good real-time performance, high accuracy, and scalability, indicating its potential applicability to a wider range of industrial intelligent inspection and monitoring scenarios. Future work will focus on enhancing the model's generalization ability under extreme conditions, optimizing digit segmentation strategies, and exploring lightweight and end-to-end recognition frameworks for deployment on embedded and edge computing devices.

REFERENCES

- [1] Jiang Tao, Tao Jiancheng, Fan Xu. Research on reading of pointer instrument in gas gathering station based on deep learning [J]. *Automatic instrument*, 2022,43 (2) : 18-22.
- [2] Zhang Yin, Lv Yanming, Qian Yunjie. Research on character extraction and fast recognition of digital instrument based on OpenCV [J]. *Automation instrument*, 2023,44 (05): 46-50 + 55.
- [3] Lu Jingbin, Xu Li. Digital instrument recognition system based on adaptive feature extraction [J]. *Modern electronic technology*, 2017,40 (24) : 147-150.
- [4] Li Zhi, Xiong Jingxian, Yan Kun, et al. Character recognition of digital tube under complex illumination conditions [J]. *Instrument technology and sensors*, 2022 (4): 42-48,63.
- [5] Zheng Lin, Wang Fulong. The license plate recognition algorithm based on improved HOG feature [J]. *Software Guide*, 2022,21 (5): 193-197.
- [6] LIN D Z, LIN F, LV Y P, et al. Chinese Character CAPTCHA Recognition and Performace Estimation Via Deep Neural Network [J]. *Neurocomputing*, 2018,288:11-19.
- [7] ZHANG X S, WANG Y. Industrial Character Recognition Based on Improved CRNCRNN in Complex Environments[J]. *Computers in Industry*,2022,142:103732.
- [8] Hua Zexi, Shi Huibin, Luo Yan, et al. Substation digital instrument detection and identification based on lightweight YOLO-v4 model [J/OL]. *Journal of Southwest Jiaotong University* : 1-11.
- [9] Chen Kaifeng, Yu Weicong, Tang Yanwen, etc. Digital instrument recognition based on lightweight deep learning network [J]. *Science, Technology and Engineering*, 2023,23 (2): 674-680.
- [10] Chen Gang, Hu Zifeng, Zheng Chao. Fast digital recognition algorithm for digital meters based on feature detection [J]. *China Test*, 2019,45 (4): 146-150.
- [11] Yang Dian, Li Xiaoyan, Liu Peiyan, etc. Research on substation instrument recognition method based on OpenCV [J]. *Automation and instrument*, 2022,37 (4) : 75-80.
- [12] ZHANG Z Y, HUA Z X, TANG Y C, et al. R ecognition method of digital meter readings in substation based on connected domain analysis algorithm [J] . *Actuators*, 2021, 10(8):170-174.
- [13] Hou Zhuocheng, Ouyang Hua, Hu Xin, et al. Based on the improved YOLOv4 color digital meter reading recognition method [J]. *Electronic measurement technology*, 2022,45 (06) : 124-129.
- [14] Xu Kechao. Research and application of deep learning algorithm in industrial instrument reading recognition [D]. *University of Electronic Science and Technology of China*, 2024.
- [15] Li Bingbing, Zhu Ge, Cao Han, etc. Automatic reading method of multi-type digital instruments based on image classification [J]. *Software*, 2023,44 (12) : 70-75.
- [16] Hui Wanyu, Wu Yuxiu. Comparative study of image denoising and filtering methods [J]. *Journal of Luoyang Institute of Technology (Natural Science Edition)*, 2021,31 (01): 71-76.
- [17] Li Qiang, Chen Chuxia, Huang Tao, etc. A median filtering algorithm for removing high-density salt and pepper noise [J]. *Computer Knowledge and Technology*, 2021,17 (35) : 137-139.
- [18] Lin Xiezhao. An image denoising algorithm based on adaptive bilateral filtering [J]. *Electronic technology and software engineering*, 2022 (11) : 196-199.
- [19] Li Fenglin, Du Hongmei, Fan Yiwei, etc. A wheel polygon fault diagnosis method based on morphological filtering [nnJ]. *Machinery*, 2023,50 (08): 39-46.