Detection and Identification of Digital Display Meter of Distribution Cabinet Based on YOLOv5 Algorithm

Yanfei Zhou[1], Yunchu Zhang[1,2(✉)], Chao Wang[1], Shaohan Sun[1], and Jimin Wang[1]

[1] Shandong Jianzhu University, Jinan 250101, China
yczhang@sdjzu.edu.cn
[2] Shandong Key Laboratory of Intelligent Buildings Technology, Jinan 250101, China

Abstract. Aiming at the problem of low recognition accuracy of digital display meter readings when the inspection robot performs inspection tasks, a YOLOv5-based digital display meter detection and recognition algorithm for distribution cabinets is proposed. For the digital display meter image captured by the inspection robot spherical camera, the YOLOv5 model is used to locate the target character area. After scale normalization, image correction, filtering and noise removal and other image pre-processing operations, the character recognition is completed in combination with the traditional machine vision algorithm, and the reading results are automatically output. For the characteristics of digital tube characters, the threading method, support vector machine algorithm and PaddleOCR algorithm are used for comparison, and a suitable algorithm model is selected to recognize numeric, alphabetic and decimal point characters. The experimental results show that the accuracy of detecting and identifying digital display meters using the YOLOv5 model and PaddleOCR algorithm is 95.3%.

Keywords: Digital display meter · YOLOv5 · Machine vision · Character recognition · Deep learning

1 Introduction

The power distribution cabinet is the final stage of electric power transmission and distribution system equipment, which is widely used in factories, building parks, urban infrastructure and other energy-using places. Digital display meter has the advantages of convenient meter reading, fast recording and accurate reading, which is widely used in power distribution cabinet. Due to technical limitations and economic cost constraints, manual observation of meter data has become the mainstream way of distribution cabinet inspection work. However, the manual observation method has the disadvantages of low efficiency, high leakage rate, and large workload to a certain extent. Therefore, accurate real-time detection of digital meter readings is important to ensure power supply safety, improve power supply efficiency, reduce line loss and save energy.

With the development of China's intelligent information era, manual inspection can no longer meet the needs of industrial sites. Real-time monitoring and accurate identification of digital meter data are the current research's main content. Traditional digital

meter detection and identification algorithms include template matching, support vector machines (SVM), threading, etc. Ju Gao [1] et al. proposed a method to add character feature matching based on the template matching method, which solves the problem of misjudgment of similar characters to a certain extent, but the ability to recognize different tilted characters of the same font is very poor. Yanling Zhang [2] et al. proposed using SVM to recognize the instrument panel parameter symbols, which improved the accuracy level of the instrument panel and the recognition rate of special parameter symbols. However, for the instrument image data with different scenes and environmental noise, there are still problems such as misjudgment, which does not apply to the processing of large sample data. Wenliang Liu [3] proposed to use the improved multi-threshold localization segmentation method and threading method to recognize the digital characters of seven-segment digital tube type digital display meters, and the correct rate was greatly improved. However, under the influence of lighting conditions and shooting angles, the phenomenon of digital misjudgment and failure to judge the results may occur.

In recent years, the rapid development of deep learning algorithms has resulted in the emergence of many target detection algorithms based on deep learning that solve some bottlenecks that traditional machine vision methods would encounter [4]. Digital instrument detection and recognition algorithms based on deep learning mainly address two problems: digital area localization and character recognition. Xun Xiong [5] et al. proposed to use contour extraction algorithm to locate dial character regions and use the improved convolutional memory neural network model (CLSTM) for character recognition, which improved the accuracy by 4.2% compared with the traditional LSTM network. Longyu Zhang [6] proposed an improved Tiny-EAST algorithm based on scene text detection algorithm to detect and locate and identify target characters of digital meters with 99.7% character detection accuracy. Peng Tang [7] et al. proposed to use Mask-RCNN algorithm to detect and identify digital meters, which has worse detection rate than YOLOv3 algorithm, but has higher accuracy. Chaoran Qu [8] et al. proposed to use DB segmentation algorithm for character region detection, and by introducing the attention mechanism improved CRNN algorithm for character recognition of the detected data, and the results show that the character recognition rate can reach 96%, but the detection rate on the test set is low.

This paper proposes a digital display meter detection and recognition algorithm based on YOLOv5. Using the images of various devices in the distribution cabinet collected by the inspection robot as the data set, YOLOv5s deep learning framework is selected to determine the target character area of the digital display meter by training the model. Then the localized character image is segmented into single characters and input to the character recognition model in turn to complete the detection and recognition of the digital display meter. The experimental results show that digital display meter reading recognition accuracy reaches 95.3%.

2 Method

Industrial meter recognition is different from text recognition in general-purpose situations, which can be disturbed by complex environments and more factors, causing reading recognition errors. Using a traditional machine vision algorithm or OCR algorithm

for instrument character detection and recognition does not solve the above problems, so this paper proposes a deep learning-based instrument character region detection and recognition algorithm for complex backgrounds. Digital display meter reading recognition mainly includes character region detection and positioning, character segmentation and character recognition. This section will introduce the algorithm model and method selection for each part in three parts.

2.1 Instrument Character Area Detection and Positioning

There are many different power devices on the distribution cabinet, such as pointer meters, digital display meters and indicators. If the reading recognition of the digital display meter is to be completed, the primary key is to detect the character area of the digital display meter from different devices.

The mainstream deep learning algorithms for target detection are the YOLO [9–12] series of algorithms. YOLO is a single-stage target detection algorithm that transforms the original target detection problem into a regression problem by directly performing classification probability regression and envelope coordinate regression on the input image to achieve target detection with fast detection capability. The YOLOv5 algorithm has comparable performance with YOLOv4 [12], but the YOLOv5 model uses the Pytorch framework, which makes the algorithm inference faster and more suitable for deployment in engineering.

The network structure of YOLOv5 consists of four parts: Input, Backbone, Neck and Head (see Fig. 1). After inputting the images, mosaic [13] data enhancement, adaptive anchor calculation, and image scaling are used sequentially to expand the dataset to improve the generalization of the target detection model. The backbone network uses Focus structure and C3Darknet-53 structure. Neck consists of FPN [14] and PANet [15] structures. Head uses CIoU_loss [16] as the loss function for bounding box regression, while redundant prediction boxes are filtered using DIoU_NMS [17].
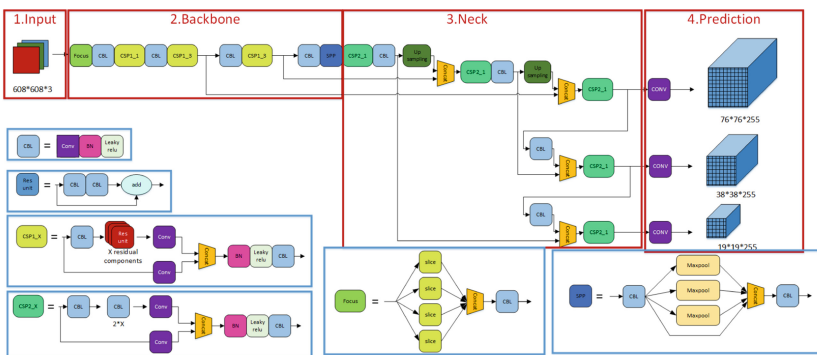

Fig. 1. YOLOv5s network diagram

According to the depth and width of the backbone network, YOLOv5 is divided into four magnitudes of models, which are noted as YOLOv5s, YOLOv5m, YOLOv5l and YOLOv5x. With the increase of model parameters and the model frame sequentially, the target detection capability gradually improves. To achieve the accurate localization of the digital display meter's character area and meet industrial sites' deployment requirements, this paper selects the YOLOv5s model to classify and locate the digital display meter on the distribution cabinet panel. YOLOv5s is the network with the smallest depth and width of the feature map in the YOLOv5 series, and its model parameter number is only 7.5M, which is suitable for deployment in the actual industrial sites. The results of the meter character area detection tested by the trained YOLOv5 model are shown in Fig. 2.
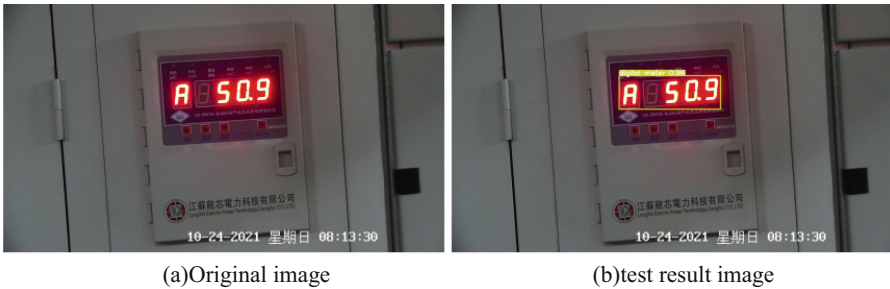

(a)Original image   (b)test result image

Fig. 2. Instrument character area detection results chart

2.2   Instrument Character Segmentation

In the actual inspection, the light source divergence of the digital tube will cause the character connection problem in the digital display meter image, and the incorrect camera shooting position will also cause the character skew. Therefore, this paper uses the position information of the character region of the digital display meter obtained from the YOLOv5 model to perform image pre-processing of the image of the character region, including image graying, skew correction, Gaussian denoising and morphological processing. The maximum interclass variance method (Otsu) [18] is used to extract the binarized images. The principle of the maximum interclass variance method is to use the idea of the maximum variance between the target region and the background region for the purpose of segmenting the image.

The character region of the digital display instrument is composed of several characters combined, so it is necessary to segment the individual characters sequentially before character recognition, so this paper adopts a combination of the horizontal projection method and vertical projection method for character segmentation of the binarized instrument character image [19]. The basic principle of the projection method is to perform the pixel statistics in horizontal and vertical directions on the binary image, and the peak and trough positions of the instrument character information in horizontal and vertical

directions can be obtained. The continuous peaks are character regions, and the constant troughs are character intervals according to which individual characters can be segmented. The result of character segmentation by the projection method is shown in Fig. 3.
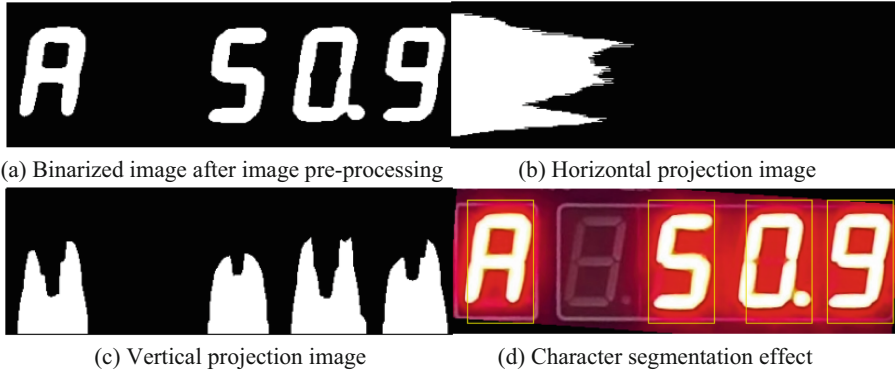

(a) Binarized image after image pre-processing   (b) Horizontal projection image

(c) Vertical projection image   (d) Character segmentation effect

Fig. 3. Projection method to split the character effect

2.3   Instrument Character Recognition

Traditional character recognition algorithms include template matching, SVM and threading method. Template matching requires many templates to be prepared in advance, and the character characteristics in different environments are different, so character recognition using a template matching algorithm is not very applicable. This section introduces the principles of the threading method and SVM to recognize characters. In addition, the PaddleOCR [20] algorithm introduced by Baidu Flying Pulp is also referred to as recognizing meter characters.

Threading Method. According to the characteristics of the digital tube-type characters, character recognition can be realized by the threading method. The principle of the threading method is to use the basic features of the 7 display segments in the 7-segment digital tube to complete the accurate judgment of the characters by extracting the feature information [21]. The seven display segments of the digital tube are labeled as ABCDEFG in the clockwise direction, as shown in Fig. 4., and the horizontal and vertical segmentation lines L1 to L7 are made in turn, corresponding to the seven display segments.
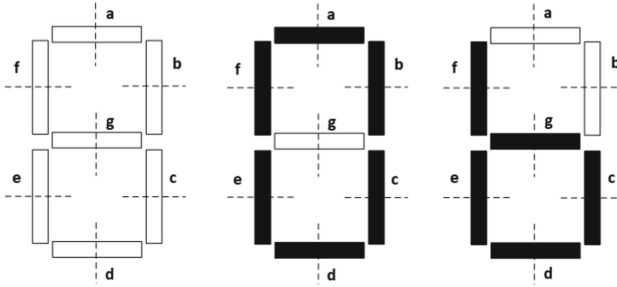
Fig. 4. Seven-segment digital tube

Scan the display segments, count the number of white points, and set the appropriate threshold value. If it is greater than or equal to the threshold value, that is, the field exists in the pen segment, it will be recorded as "1", otherwise it will be recorded as "0". This results in a series of binary codes. Each number is then noted as the corresponding numeric result according to its characteristics on the display segment. For example, the binary code of the number "0" in the seven display segments is "0111111", which is converted into decimal as "63" and recorded as "0". The binary code of the letter "b" in the seven display segments is "1111100", which is converted to "124" in decimal and is recorded as "B". The results of 0 to 9 and the corresponding codes of ABC are shown in Table 1.

Table 1. Digital tube character code correspondence table.

	a	b	c	d	e	f	g	Binary	Decimal
0	✓	✓	✓	✓	✓	✓		0111111	63
1		✓	✓					0000110	6
2	✓	✓		✓	✓		✓	1011011	91
3	✓	✓	✓	✓			✓	1001111	79
4		✓	✓			✓	✓	1100110	102
5	✓		✓	✓		✓	✓	1101101	109
6	✓		✓	✓	✓	✓	✓	1111101	125
7	✓	✓	✓					0000111	7
8	✓	✓	✓	✓	✓	✓	✓	1111111	127
9	✓	✓	✓	✓		✓	✓	1101111	103
A	✓	✓	✓		✓	✓	✓	1110111	119
B			✓	✓	✓	✓	✓	1111100	124
C	✓			✓	✓	✓		0111001	57

Due to the special position of the decimal point, it is easy to stick together with other characters after image pre-processing, which makes the threading method unable to realize the decimal point recognition. According to the location characteristics of the decimal point located in the lower right corner of the character, the process of lower right corner counting pixels is used for decimal point recognition. The specific method is to select a fixed region in the lower right corner of a single character picture, iterate the pixel value of the part, set a threshold value, and determine whether the decimal point exists by counting the number of white points.

SVM. SVM is a practical data classifier that is easier to apply than neural networks. The goal of SVM is to generate a model that can predict the target value [22]. Given a set of points $\{(x_1, y_1), (x_2, y_2), \ldots \ldots, (x_m, y_m)\}$, where $x_i \in R^n$ denotes the sample points, and $y_i \in \{-1, 1\}$ represents the class to which the corresponding sample point x_i belongs, the SVM requires solving this optimization problem as follows.

$$\min_{w,b,\xi} \frac{1}{2} w^T w + C \sum_{i=1}^{i} \xi_i \tag{1}$$

$$y_i(w^T \phi(x_i) + b) \geq 1 - \xi_i, \xi_i \geq 0 \tag{2}$$

The training vector x_i is projected to a high-dimensional space defined by ϕ. The SVM is to find a linearly differentiable hyperplane in this high-dimensional space. $C > 0$ is the penalty parameter of the training error. $K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$ is the kernel function, and the radial basis kernel function (RBF) is chosen in this paper, where $K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$ and $\gamma = 0.01$ are the kernel parameters.

All images obtained by character segmentation, including numbers, letters and other characters, are uniformly normalized to a standard size of 28×28 and named with the value of each character as a classification folder as the sample images for SVM training in this paper and some of the instrumentation character samples are shown in Fig. 5.

Fig. 5. Part of the instrument character samples

This paper combines the Principal Component Analysis (PCA) algorithm [23] for extracting character sample image features and then uses SVM for classification. The basic principle of PCA algorithm is to expand the n × n images into $1 \times n^2$ one-dimensional vectors. Suppose there are m samples, which are processed by the PCA algorithm to form an m × n2 array, with each row representing one sample. Each row represents one sample. In this paper, the image size is 28×28, the sample category is

14, and the training sample for each type is 240, so each image is expanded into a 1×784 one-dimensional vector, which eventually forms a 3360×784 array file as the data set for SVM training.

PaddleOCR. PaddleOCR is an ultra-lightweight OCR (Optical Character Recognition) system open-sourced by Baidu, which mainly consists of three parts: DB text detection [24], detection frame correction [25] and CRNN text recognition [26]. In this paper, we use PaddleOCR source code to recognize the meter character area detected by YOLOv5 directly, and the detection effect is shown in Fig. 6. To improve the recognition accuracy, this paper crops and saves the character region images based on the location information of the meter character regions detected by YOLOv5, construct the dataset, and retrains the model based on the PaddleOCR source code to achieve the character region recognition of meters.

3   Experimental Results and Analysis

3.1   Experimental Platform

In this paper, an intelligent inspection robot equipped with a visual inspection system is used as the experimental platform, and the camera equipment is a Hikvision thermal imaging dual-spectrum network intelligent dome camera. The experimental hardware is a PC, the operating system is Windows 10 64-bit system, the PC processor model is Intel(R) Core(TM) i7-7700, the graphics card model is NVIDIA GeForce GTX 1050Ti. The algorithm uses the deep learning Pythorch framework, and the programming language is Python.

3.2   Acquisition and Labeling of Experimental Data Sets

In this paper, the experimental data is taken by using the intelligent spherical camera of the hanging rail intelligent inspection robot to shoot the equipment operation status of the distribution cabinet in the distribution room, and 2432 data images of different types and different lighting conditions are selected as the experimental data set to ensure the reliability of the experiment. Use the LabelImg labelling tool to rectangular label boxes with a total of 3 device types, namely "pointer instrument", "digital meter", and "colour indicator", and save the text file in YOLO format. The labelled images are shown in Fig. 6. To distinguish different models of the same type of equipment, five other equipment categories are derived, namely "cos instrument", "pointer instrument A (pointer instrument with indicator)", "digital meter A (circuit breaker digital meter)", "white indicator", and "meter light (indicator on pointer meter)", a total of eight categories.

Fig. 6. Labelimg labeled images

3.3 Model Training

The YOLOv5 digital meter character region localization model uses the deep learning framework Pytorch, with SGD stochastic gradient descent and learning rate decay strategies selected as hyper-parameters to train the network. The initial learning rate is 0.01, the batch size is 8, the number of iterative rounds (epochs) is 150, and the input image resolution is 640×640. The learning rate momentum factor (SGD momentum) is 0.937, and the weights of the loss function are $\lambda_{box} = 0.05$, $\lambda_{cls} = 0.5$ and $\lambda_{obj} = 1.0$.

Considering the small sample size of characters, this paper uses the SVC (classification algorithm of support vector machine) of the SVM algorithm in the Scikit-learn machine learning library to complete the digital classification training. All single-character images obtained by projection segmentation are normalized to 28×28 pictures and then used as the data set for digital classification training. The kernel function (kernel) is selected as "RBF" (Gaussian kernel function), the kernel function coefficient gamma is 0.01, and the penalty coefficient C is 15.

The PaddleOCR character recognition model uses the deep learning framework Pytorch. The backbone network is selected as the CRNN recognition model of Resnet34_vd, and the hyperparameters are chosen as the stochastic gradient descent method SGD with the restart. The learning rate descent method is selected as the cosine annealing function Cosine. The initial learning rate is 0.001, and the number of iterations (epoch_num) is 200. The configuration file also has default data enhancement, including colour space transformation (cvtColor), blur, jitter, gauss noise, random crop, perspective, colour inversion (reverse), etc.

3.4   Results Analysis

Analysis of Instrument Character Area Detection Results. In this paper, a total of 2432 inspection images were selected as the dataset for YOLOv5 model training, of which 1946 were in the training set, accounting for 80%; 243 were in the test set, accounting for 10%; and 243 were in the validation set, accounting for 10%. The training was completed in 9 h using the YOLOv5s model on PC, and the training results are shown in Fig. 7.
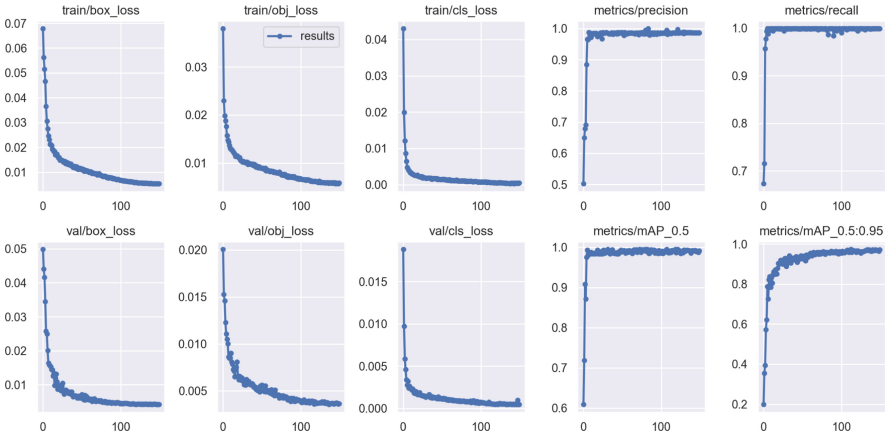

Fig. 7. YOLOv5s model training results image

The horizontal coordinates of YOLOv5s model training results represent the number of training rounds epoch, and the main indicators are the target and envelope loss curves loss, Precision, Recall, and mAP@0.5, mAP@0.5:0.95 for the training and validation sets in order. 0.5:0.95 value approaches 1 as the number of training rounds increases, and the training effect is good.

In this study, mAP (mean Average Precision) is used as the evaluation metric of the model. P(Precision) represents the proportion of correctly identified samples to all identified samples in the dataset, and R(Recall) represents the proportion of correctly identified samples to all models and is calculated as follows.

$$P = \frac{TP}{TP + FP} \times 100\% \tag{3}$$

$$R = \frac{TP}{TP + FN} \times 100\% \tag{4}$$

TP indicates the number of correctly identified distribution cabinet devices, FP shows the number of incorrectly identified distribution cabinet devices, and FN shows the number of missed identified distribution cabinet devices.
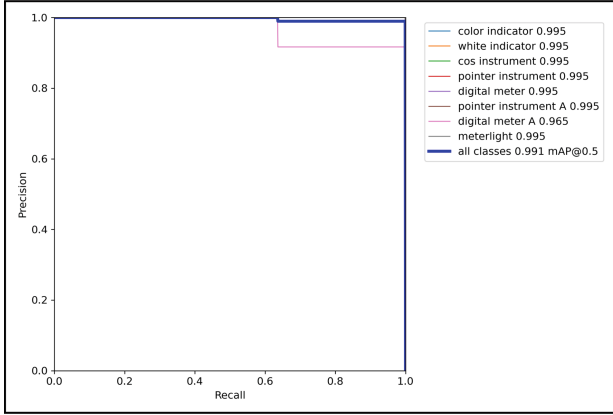

Fig. 8. P_R curve of YOLOv5 model training

The P-R (Precision-Recall) curve can be plotted according to Eq. (3), as shown in Fig. 8. The average precision AP (Average Precision) of a single category can be calculated by averaging the Precision values of the P-R curves. The mAP is the average value of each classification accuracy AP, and the relevant formula of mAP is as follows:

$$AP_i = \int_0^1 p(r)dr, i = 1, 2, ..., n \tag{5}$$

$$mAP = \frac{\sum AP_i}{n}, i = 1, 2, ..., n \tag{6}$$

where p is the precision rate, r is the recall rate, and n is the number of categories. The effect of using the YOLOv5s model to detect the character region of the distribution cabinet meter is shown in Fig. 9.

Fig. 9. Distribution cabinet instrumentation character area detection effect

Analysis of Instrument Character Recognition Results. In this paper, after completing the detection and positioning of the character area of the meter using YOLOv5, the character recognition is performed by character area cropping, image pre-processing and character segmentation operations, using the threading method, SVM and PaddleOCR, respectively. According to the character characteristics of the digital display meter in the data set of this paper, the decimal point is connected with the numeric characters in the segmented single character picture, which is difficult to recognize. The threading method and SVM algorithm determine whether the decimal point exists by scanning the lower right corner of the single character binarization picture to count the number of white points and output the character recognition results sequentially according to the coordinate order. The instrument character recognition results are shown in Fig. 10.

Fig. 10. Character recognition results

Character recognition was tested using the threading method, SVM and PaddleOCR for comparison, using a verification set of 240 digital display meter pictures, and the test results are shown in Table 2.

Table 2. Comparison of character recognition results of different algorithms

Algorithm	Accuracy (%)	Relative error (%)	FPS
Threading method	77.9	22.1	62
SVM	93.8	6.2	71
PaddleOCR	95.3	4.6	257

As shown in Table 2, the relative errors of the SVM algorithm model and PaddleOCR detection on the test set are significantly reduced, and the correct rates are improved by 15.9% and 17.4%, respectively, relative to the threading method. The recognition speed of the PaddleOCR algorithm is faster, and it only takes 4 milliseconds (ms) to process one frame of the image, which can ensure the real-time performance of the algorithm. The experimental results show that using the YOLOv5 algorithm to detect the meter character region, combined with the PaddleOCR algorithm to recognize characters, can improve the accuracy of digital display meter detection and recognition in complex scenes, and the final meter detection and recognition accuracy is 95.3%.

4 Conclusion and Prospect

For the actual needs of digital display meter detection and recognition in the process of robot inspection, this paper proposes a YOLOv5-based digital display meter detection and recognition algorithm. The method uses the YOLOv5s model to locate the character region of the digital display meter image, obtains the target character image, and uses the PaddleOCR algorithm model to complete the character recognition, improving the accuracy and speed of digital display meter recognition in complex scenes. Future work will consider improving the network structure to increase character recognition accuracy while maintaining speed.

References

1. Gao, J., Ye, H.: An effective algorithm for secondary recognition of digital images of water meters. J. Southeast Univ. (Nat. Sci. Ed.) **43**(S1), 153–157 (2016)
2. Zhang, Y.L.: Research on instrument dial parameter symbol recognition technology. Guangdong University of Technology (2008)
3. Liu, W.L.: Research and implementation of digital recognition in seven-segment digital display instrumentation. Dalian University of Technology (2013)
4. Hua, Z.X., Shi, H.B., Luo, Y., et al.: Identification of digital instrumentation detection in substations based on lightweight YOLO-v4 model. J. Southwest Jiaotong Univ., 1–11 (2021)
5. Xiong, X., Chen, X.D., Wu, L., Lin, X.L.: Digital dial reading recognition based on convolutional memory neural network. Comb. Mach. Tools Autom. Mach. Technol. **07**, 72–75 (2019)
6. Zhang, L.Y.: Research on detection and identification algorithm of substation instrumentation based on deep learning. Tianjin University of Technology (2021)
7. Tang, P., Liu, Y., Wei, H.G., Dong, X.Y., et al.: Automatic identification algorithm of digital meter readings based on mask-RCNN offshore booster station. Infrared Laser Eng. **50**(S2), 163–170 (2021)
8. Qu, C.R., Chen, L.W., Wang, J.S., Wang, S.G.: Research on industrial digital meter recognition algorithm based on deep learning. Appl. Sci. Technol., 1–7 (2022)
9. Redmon, J., Divvala, S., Girshick, R., et al.: You only look once: unified, real-time object detection. In: Computer Vision & Pattern Recognition (2016)
10. Redmon, J., Farhadi, A.: YOLO9000: better, faster, stronger. In: IEEE Conference on Computer Vision & Pattern Recognition, pp. 6517–6525 (2017)
11. Redmon, J., Farhadi, A.: Yolov3: an incremental improvement. arXiv preprint arXiv:1804.02767 (2018)
12. Bochkovskiy, A., Wang, C.Y., Liao, H.Y.M.: Yolov4: optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934 (2020)
13. Yun, S., Han, D., Oh, S.J., et al.: CutMix: regularization strategy to train strong classifiers with localizable features. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 6023–6032 (2019)
14. Lin, T.Y., Dollár, P., Girshick, R., et al.: Feature pyramid networks for object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2117–2125 (2017)
15. Liu, S., Qi, L., Qin, H., et al.: Path aggregation network for instance segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 8759–8768 (2018)

16. Rezatofighi, H., Tsoi, N., Gwak, J.Y., et al.: Generalized intersection over union: a metric and a loss for bounding box regression. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 658–666 (2019)
17. Zheng, Z., Wang, P., Liu, W., et al.: Distance-IoU loss: faster and better learning for bounding box regression. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, no. 07, pp. 12993–13000 (2020)
18. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Trans. Syst. Man Cybern. **9**(1), 62–66 (1979)
19. Fang, X.W., Fu, X.W.: ID card segmentation algorithm based on projection method and Caffe framework. Comput. Eng. Appl. **53**(23), 113–117 (2017)
20. Qiu, D., Weng, M., Yang, H.-T.: A fast lane line detection method based on improved probabilistic Hough transform. Comput. Technol. Dev. **30**(5), 43–48 (2020)
21. Song, Y.Y., Tang, D.L., Wu, X.L., et al.: Study on digital tube image reading by combining improved threading method and HOG+ SVM method. Comput. Sci. **48**(S2), 396–399+440 (2021)
22. Wang, X.D., Wei, C.B., Feng, H.R., et al.: Digital identification method for digital display instruments in substation protection rooms based on SVM. Electron. Meas. Technol. **42**(2), 92–95 (2019)
23. Chen, X., Xu, S.Y., et al.: Identification of braided river reservoir inclusions based on support vector machine and principal component analysis. J. China Univ. Petr. (Nat. Sci. Edn) **45**(04), 22–31 (2021)
24. Liao, M., Wan, Z., Yao, C., et al.: Real-time scene text detection with differentiable binarization. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, no. 07, pp. 11474–11481 (2020)
25. Yu, D., Li, X., Zhang, C., et al.: Towards accurate scene text recognition with semantic reasoning networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 12113–12122 (2020)
26. Li, W., Cao, L., Zhao, D., et al.: CRNN: Integrating classification rules into neural network. In: The 2013 International Joint Conference on Neural Networks (IJCNN), pp. 1–8 IEEE (2013)