

## Article

# Pointer Meter Recognition Method Based on YOLOv7 and Hough Transform

Chuanlei Zhang \*, Lei Shi, Dandan Zhang, Ting Ke and Jianrong Li

College of Artificial Intelligence, Tianjin University of Science and Technology, Tianjin 300453, China; shil0406@163.com (L.S.); zdd15076311103@163.com (D.Z.); keting@tust.edu.cn (T.K.); lisa\_ljr@tust.edu.cn (J.L.)

\* Correspondence: 97313114@tust.edu.cn; Tel.: +86-139-1178-0653

**Abstract:** The current manual reading of substation pointer meters wastes human resources, and existing algorithms have limitations in accuracy and robustness for detecting various pointer meters. This paper proposes a method for recognizing pointer meters based on YOLOv7 and Hough transform to improve their automatic readability. The proposed method consists of three main contributions: (1) Using YOLOv7 object detection technology, which is the latest YOLO technology, to enhance instrument recognition accuracy. (2) Providing a formula for calculating the angle of a square pointer meter after Hough transformation. (3) Applying OCR recognition to the instrument dial to obtain the model and scale value. This information helps differentiate between meter models and determine the measuring range. Test results demonstrate that the proposed algorithm achieves high accuracy and robustness in detecting different types and ranges of instruments. The map of the YOLOv7 model on the instrument dataset is as high as 99.8%. Additionally, the accuracy of pointer readings obtained using this method exceeds 95%, indicating promising applications for a wide range of scenarios.

**Keywords:** pointer meter recognition; YOLOv7; Deeplabv3+; PGNet; Hough transform



**Citation:** Zhang, C.; Shi, L.; Zhang, D.; Ke, T.; Li, J. Pointer Meter Recognition Method Based on YOLOv7 and Hough Transform. *Appl. Sci.* **2023**, *13*, 8722. <https://doi.org/10.3390/app13158722>

Academic Editor: Christos Bouras

Received: 10 June 2023

Revised: 14 July 2023

Accepted: 25 July 2023

Published: 28 July 2023



**Copyright:** © 2023 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

The substation is in a very important position in the power system, and its main task is to convert and distribute electric energy. The normal operation and safety maintenance of substations are issues that need to be paid close attention to in daily inspections. The readings of the meters in the substation can reflect the operation of many pieces of equipment. However, due to the complex external environment of the meter, the large number, and the dependence of the original equipment of the substation on the meter, traditional substation inspection cannot effectively guarantee the real-time validity of the meter reading [1].

The main method of traditional substation inspection is manual inspection and recording. The operator manually judges to read the pointer reading, which requires the operator to have a high level of technology and experience, so it is inevitably interfered with by many human factors, such as the operation angle of the operator observing the dial, the error caused by the observation distance, the long and boring recognition process, and visual fatigue are easily affected by subjective factors, resulting in low work efficiency and inaccurate and wrong recognition results [2]. Traditional substation manual inspection requires personnel to inspect each meter one by one and record data manually. This method requires a lot of time and labor costs, especially in large substations where the inspection process is more cumbersome and time-consuming. The substation environment is complex and there are certain safety risks, such as high-voltage electricity and explosions. During manual inspections, staff need to frequently contact high-voltage equipment, which poses certain safety risks and may cause casualties. Manual inspection can perform only regular inspections on the meters and cannot realize real-time monitoring of the meters, so it is difficult to quickly find and deal with abnormalities in the meters.

There are many meters in a substation, such as an oil temperature meter, pressure meter, ammeter, and so on. There are two common types of gauges: digital and analog.

Figure 1 shows the digital and analog meters. These two types of meters have different characteristics. In the meter industry, the share of digital meters is increasing, because they have the advantages of easy reading and high precision; however, digital meters are not suitable in some special circumstances, such as harsh dust and oil environments. Compared to the digital meter, the pointer meter has greater advantages in this respect, such as being dustproof and waterproof, and having antifreeze, anti-interference, simple structure, low price, strong anti-interference ability, etc. [2]. Pointer meters are widely used in industrial production, automobile manufacturing, electric power industry, and other fields. They are mainly used to measure and display various physical quantities, such as current, voltage, temperature, and pressure, etc. Although digital meters are gradually replacing analog meters, there are still many analog meters in operation in remote areas and developing countries because of the high cost and time-consuming nature of replacing the meters [3].



**Figure 1.** (a) Digital meter; (b) pointer meter.

However, there are still some challenges in pointer meter recognition technology. The pointer shapes and colors of pointer meters are very different, which makes it difficult to accurately classify and locate all pointers in image recognition and processing. Analog gauges are typically used in a variety of lighting conditions, such that changes in ambient lighting may affect the recognition of the pointer's position. In addition, background noise, such as reflections, shadows, etc., will also affect pointer recognition, and these factors affect the recognition accuracy of meter readings. Analog gauges usually need to be photographed or scanned to obtain their image. However, if the image quality is poor, blurry, or low in pixel count, it can also become difficult to identify the pointer position. When the meter reading changes, it is necessary to identify and upload the data in real time, so the response speed of the identification system also needs to be further improved. Therefore, in future research, it is necessary to continuously improve and optimize the pointer meter recognition technology to improve its application effect.

Overall, the contributions of the work in this paper are as follows:

1. Yolov7 object detection technology is employed, which is the latest Yolo technology, to accurately and quickly locate instruments in complex backgrounds and enhance instrument recognition accuracy.
2. Relatively little research has been conducted on square pointer gauges to provide a formula for pointer readings. This formula fills the research gap in the field of square pointer instruments and provides an effective method for accurately reading the value indicated by the pointer.
3. For instruments of different models and ranges, the PGNet method is used to identify scale values and models, which exhibits high robustness.

By combining these technologies and methods, this approach achieves accurate, fast, and robust detection and localization in pointer-type instrument recognition, suitable for various complex scenarios and different types of instruments. A detailed description of the specific process will be provided in the third part.

## 2. Related Work

At present, a lot of research has been conducted on the automatic identification of pointer meters. The technology of pointer meter recognition can be divided into three types: technology based on template matching, technology based on OpenCV, and technology based on deep learning. (1) Techniques based on template matching [4]. This technology is a computer vision technology based on statistics, and its basic idea is to compare the image of the pointer meter with the established template image, so as to determine the value of the pointer meter. (2) Technology based on OpenCV [5]. The technology uses the OpenCV algorithm to process the image of the pointer meter to recognize the value on the pointer meter. (3) Techniques based on deep learning [6]. This technology can use the deep neural network to automatically extract the characteristics of the pointer of the meter, so as to identify the value on the pointer meter more accurately. In terms pointer meter recognition technology, there are many mature methods, such as deep learning, convolutional neural network (CNN) [7], recurrent neural network (RNN), and so on. These techniques have all made great progress in improving the accuracy and robustness of pointer gauge recognition.

Liu et al. [8] first used Faster R-CNN to determine the position of the object meter, then used the feature correspondence algorithm and perspective transformation to obtain high-quality images, and finally determined the pointer position and read the readings through Hough transform. Wu et al. [9] proposed a meter image skew correction algorithm based on binary mask and improved Mask-RCNN, which achieved high-precision ellipse fitting. Li et al. [10] introduced deep learning to detect and recognize the scale value text in the dial, then rectified the image and determined the center of the meter, and used polar coordinate transformation to convert the arc scale area into a linear scale area to obtain the scale marks and pointers. The position of the pointer was read by the distance method. Liu et al. [11] used a Gaussian scale space to enhance the scale invariance of the ORB algorithm. At the same time, the RANSAC was used to filter matching point pairs, which improves the accuracy of feature pointer matching. Finally, the pointer is fitted by Hough transformation and the meter reading is obtained. Cai et al. [12] proposed a new virtual sample generation technique to generate a large number of images from a small number of meter images to add to the training model. Zuo et al. [13] constructed a new deep learning algorithm, replacing RoiAlign in the existing mask RCNN with PrRoIPooling, classifying the meter types when fitting the pointer binary mask, and finally calculating the meter by the angle method readings. Wang et al. [2] used the Faster RCNN object detection method to locate the meter area and used the Poisson fusion method and the K-fold verification algorithm to expand the dataset and optimize the dataset quality. Finally, the pointer position was detected by Hough transform and the pointer reading was obtained. Laroca et al. [14] devised a two-stage approach to counter detection using the fast YOLO object detector and evaluated three different CNN-based methods for counter recognition. Salomon et al. [3] sorted out the UFPR-ADMR public real-world dialing table dataset, proposed a recognition baseline based on deep learning, and proposed the main problems and detailed error analysis in the process of recognizing the meter. Salomon et al. [15] introduced a new dataset UFPRADMR-v2 for ADMR based on [3], combining YOLOv4 with a new regression method (AngReg). Bishwokarma et al. [16] used a deep neural network YOLOv3 to detect and recognize meter counters and digits. Zhang et al. [1] proposed a digital meter method based on the connected domain analysis algorithm, using Faster R-CNN to locate the dial area, and then using YOLOv4 to detect the digital area. Combined with the connected area algorithm, it judged whether there was a decimal point after the number. Yan et al. [17] used Gaussian filtering and binarization methods to

identify pointer circular meters. Meng et al. [18] obtained the key point information of the pointer meter through the deep learning algorithm, determined the pointer rotation center and radius according to the key point information, and finally obtained the meter reading according to the pointer angle and the meter range. Zhang et al. [19] proposed a pointer meter recognition algorithm combining meter detection and localization, deep learning and image processing techniques. The algorithm used the Faster R-CNN algorithm to classify three types of meters (voltmeter, ammeter, and digital meter), followed by image processing for more accurate needle readings. Laroca et al. [20] introduced a new stage called corner detection and counter classification, which corrects meter regions and rejects illegible or wrong meters before the recognition stage. Dong et al. [21] regarded the pointer as a two-dimensional vector whose initial point coincides with the end of the pointer and whose direction is along the direction from the end to the top. Bayhan et al. [22] used Faster-RCNN and YoloV4 models to train military action images taken by drones. The results showed that the accuracy of Faster-RCNN reached 93%, while the accuracy of YoloV4 was 88%. Ozkan et al. [23] used Faster-RCNN and SSD MobileNet V2 models to train military action images taken by drones and transplanted the trained models to the Raspberry Pi 4 Model B electronic board. The results show that the accuracy rate of the Faster-RCNN model reached 91%, while the accuracy rate of the SSD MobileNet V2 model was 88%.

The basic principle of pointer meter recognition technology is to analyze and recognize the image of the pointer meter through image processing and machine learning technology, so as to read the meter data. The key steps of this technology include image preprocessing, feature extraction, pointer positioning, pointer recognition, and data reading. By analyzing the characteristics of the pointer in the image, such as pointer length, color, shape, etc., the position and direction of the pointer in the image are determined, and finally the meter reading is obtained by identifying the scale. This technology can be applied to different types of pointer meters, such as gauges, gas meters, etc.

The following is the structure of this article, Section 2 introduces the related work of pointer meter recognition; Section 3 introduces the recognition algorithm of pointer meters, and introduces the related technologies used, including the principles of YoloV7, DeepLabv3+, PGNet, Thining, Hough transform, and reading; finally, Section 4 provides a detailed analysis of a large number of experiments.

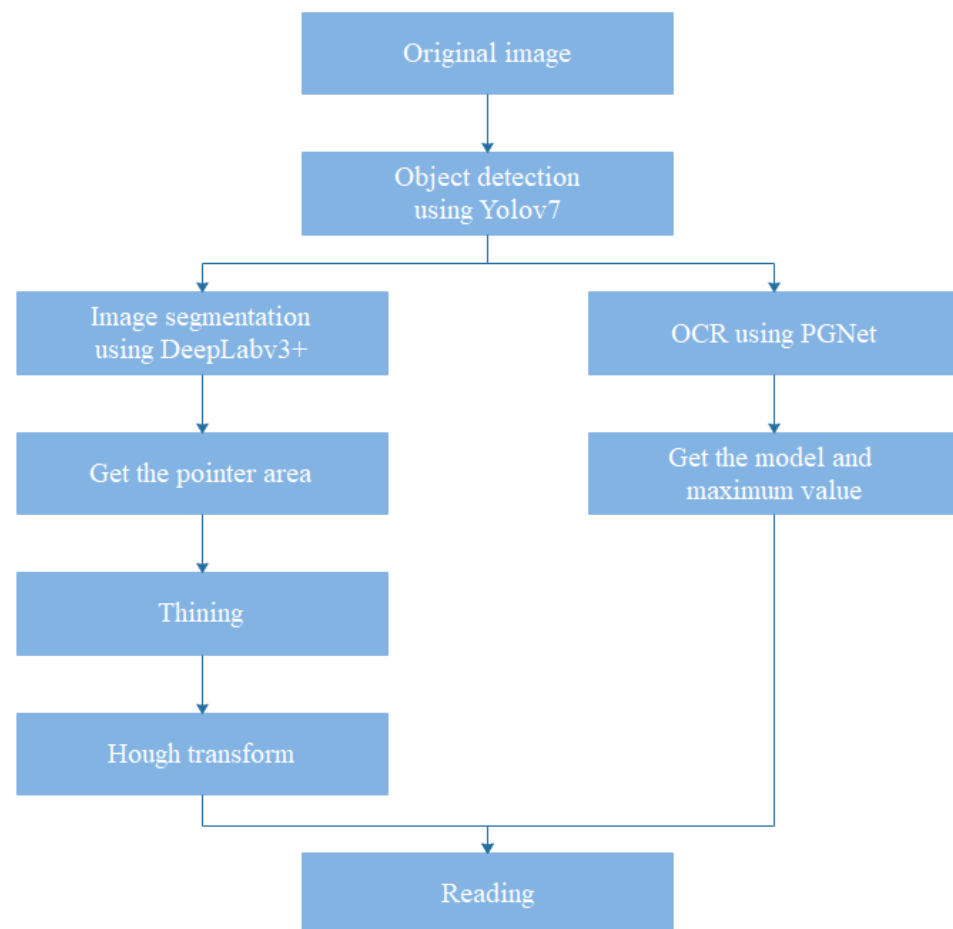
### 3. Pointer Meter Recognition Method

This paper proposes a pointer meter recognition method based on YoloV7 and Hough transform. First, use YoloV7 for object detection to locate the position of the meter and crop the image: input the cropped image into the DeepLabv3+ image segmentation model to extract the pointer area, use Thinning and Hough transform to determine the precise position of the pointer, at the same time, use PGNet in the OCR (Optical Character Recognition) method to identify the scale value and model of the pointer meter, and use the maximum value of the obtained scale value as the range of the pointer; finally, use the precise position and span of the pointer to determine the reading of the meter. The flow chart of the pointer meter identification method is shown in Figure 2. The following will introduce YoloV7, DeepLabv3+, Thining, Hough transform, PGNet, and reading.

#### 3.1. YoloV7

Object detection is an important task in computer vision, it refers to the identification of an object of interest from an image or video and determining its location. Some common object detection algorithms are Faster R-CNN [24], Single Shot Multi-Box Detector (SSD) [25], and YOLO [26].

YoloV7, the single-stage object detector, stands out as the swiftest and most precise real-time object detection system. The performance improvement of YoloV7 establishes a new and important benchmark [27].



**Figure 2.** Flowchart of Pointer Meter Identification Method.

The YOLO architecture is based on FCNN. The YOLO framework consists of Backbone, Neck and Head. Backbone is responsible for extracting image features, Neck is responsible for multi-scale feature fusion, and Head is responsible for target location and category prediction. The overall architecture of the Yolov7 model is shown in Figure 3.

First, preprocess the input image, align it into a  $640 \times 640$  RGB image, and input it into the backbone network. According to the three-layer output in the backbone network, the head layer continues to output three layers of feature maps of different sizes through the backbone network. After RepVGG block and conv, the three types of tasks of image detection (classification, background classification, border) are predicted, and the final result is output [28].

### 3.2. DeepLabv3+

Image segmentation is an important task in the field of image processing, it refers to separating different parts of an image, identifying their boundaries, and assigning them to different categories or regions. The purpose of image segmentation is to enable the image processing system to identify different objects in the image for further analysis and processing. Figure 4 illustrates the overall architecture of the DeepLabv3+ model. The primary component of its Decoder is a DCNN that incorporates dilated convolutions. This DCNN has the capability to utilize commonly employed classification networks like ResNet. Additionally, the Decoder incorporates a spatial pyramid pooling module, known as Atrous Spatial Pyramid Pooling (ASPP), which also employs dilated convolutions and pooling operations. The ASPP module primarily aims to introduce multi-scale information into the model.



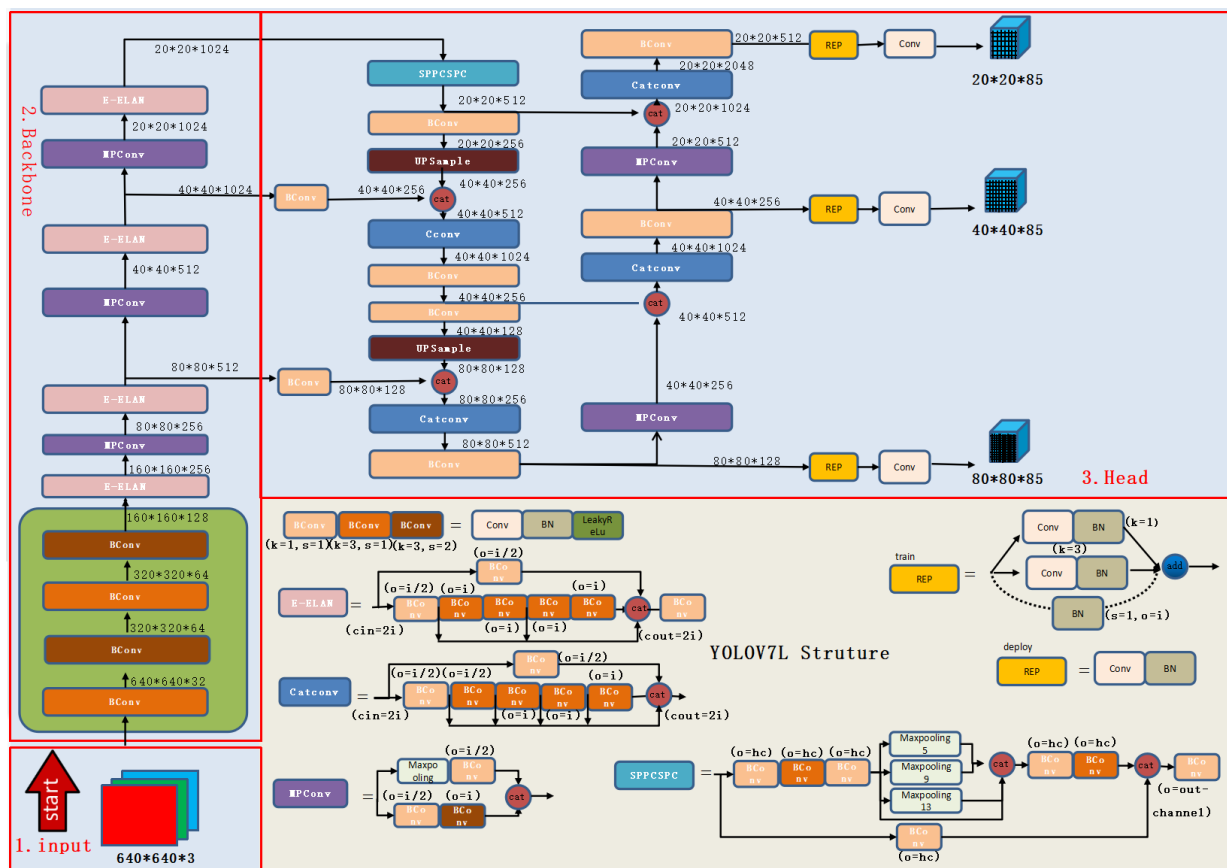


Figure 3. The overall architecture of the YOLOv7 model.

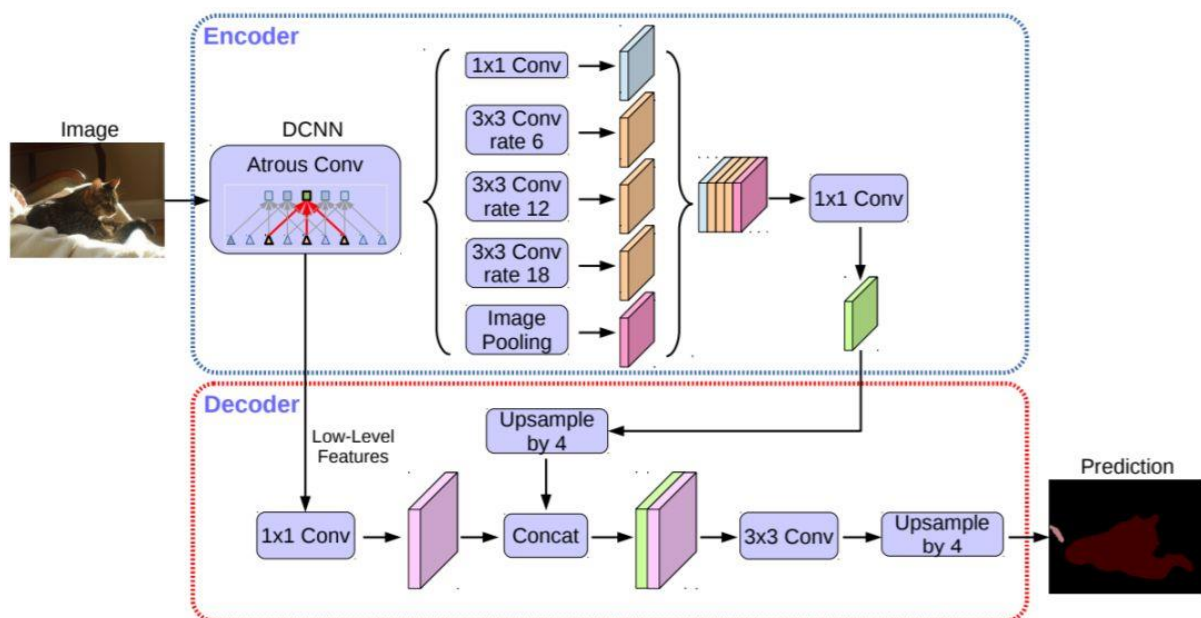


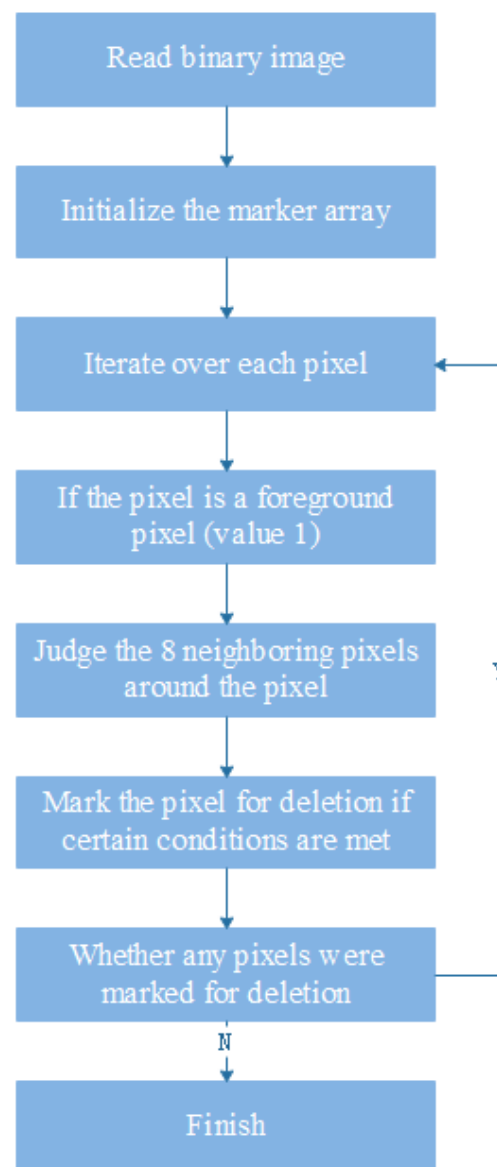
Figure 4. The overall architecture of the DeepLabv3+ model.

In comparison to DeepLabv3, DeepLabv3+ introduces the Decoder module, which further integrates both the underlying features and high-level features to enhance the accuracy of segmentation boundaries. This integration is achieved by leveraging the concept of an Encoder-Decoder architecture, building upon the foundation of Dilated-

FCN. Essentially, DeepLabv3+ incorporates the idea of an Encoder-Decoder to refine the segmentation process [29].

### 3.3. Thinning

The Thinning algorithm in OpenCV is a morphology-based image-processing algorithm for extracting skeletons or edges in binary images. The basic principle of the thinning algorithm is to gradually reduce the area of the object region through iteration until it reaches the final skeleton or edge. In each iteration, the algorithm performs Erosion on the object area and calculates the difference between the eroded result and the original area. If the difference is 0, it means that the region has reached the skeleton or edge; otherwise, the edge pixels in the difference value need to be removed, and then the erosion operation is performed again until the final result is reached. The refinement flow chart is shown in Figure 5. Common thinning algorithms include the Zhang-Suen algorithm, Guo-Hall algorithm, etc., which use different conditions and rules to achieve thinning operations.



**Figure 5.** The refinement flow chart.

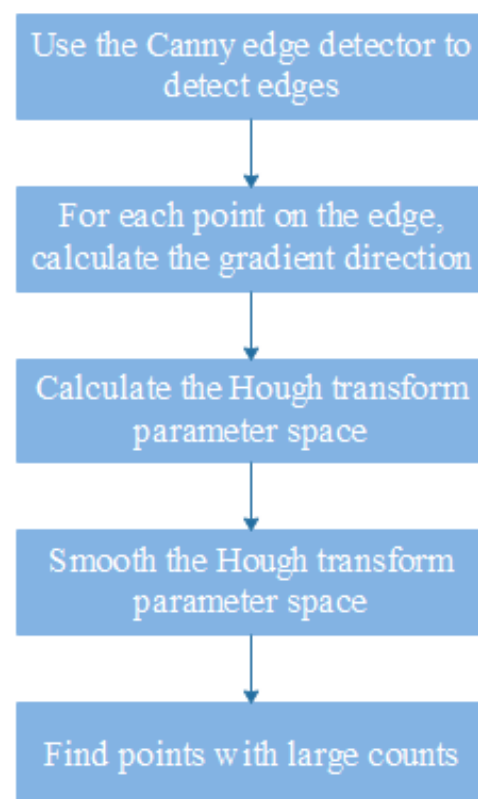
### 3.4. Hough Transform

Hough transform is a classic image processing algorithm based on parameter space, it is used to detect various shapes in images, including straight lines, circles, ellipses,

etc. Among them, line detection is one of the most commonly used applications of Hough transform.

The straight-line detection algorithm of Hough transform can be divided into the following steps: 1. Edge detection, for example canny edge detection. 2. Map the edge point to the Hough space and store it in the accumulator. 3. Find the most reasonable point in the accumulator through the threshold and other possible restrictions (for example, find the point with the largest value in the accumulator or the point greater than the threshold), and generate a straight line of infinite length. 4. Convert a straight line of infinite length to finite length and overlap it on the original image.

Hough transform can be well-applied to the detection of straight lines in the image, but due to the large amount of calculation and its sensitivity to noise, certain optimization and preprocessing are required. In practical applications, algorithm parameters can be adjusted according to specific scenarios and need to achieve the best detection effect. The flow chart of Hough transform is shown in Figure 6.



**Figure 6.** The flow chart of Hough transform.

### 3.5. PGNet

Optical Character Recognition (OCR) is the process of converting scanned or digital images of text into editable and searchable text data. The technology behind OCR enables computers to recognize and process written or printed text, making it possible to extract, store, and manipulate textual information from a variety of sources.

In recent years, end-to-end OCR algorithms have made good progress, including the MaskTextSpotter series, TextSnake, TextDragon, PGNet series, etc. Among these algorithms, PGNet has features that other algorithms do not have: PGNet loss is designed to guide training: no character-level annotation is required, NMS and ROI-related operations are avoided, and the prediction speed is improved. A module for predicting the reading order within a text line is proposed. A graph-based rectification module (GRM) is introduced to further improve the recognition performance. It makes the recognition accuracy higher and



the prediction speed faster [30]. The schematic diagram of the PGNet algorithm is shown in Figure 7.

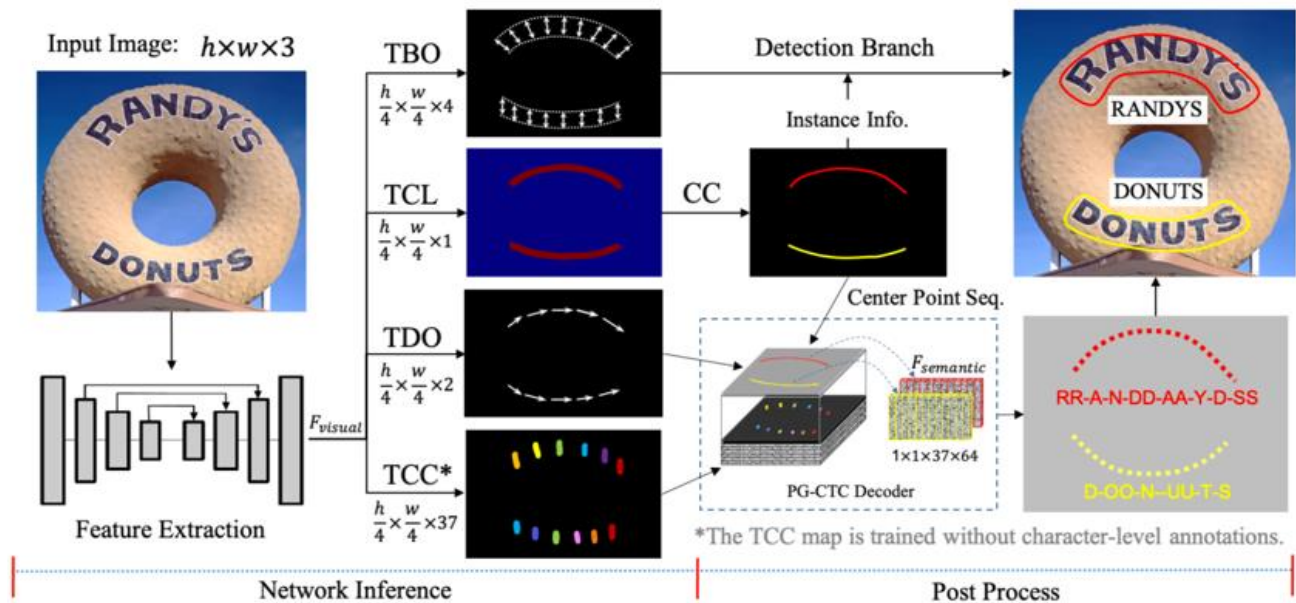


Figure 7. Schematic diagram of PGNet algorithm.

The input image will enter four different branches through feature extraction: the TBO (Text Edge Offset Prediction) module, the TCL (Text Centerline Prediction) module, the TDO (Text Orientation Offset Prediction) module, and the TCC (Text Character Classification) graph prediction) module. The output of TBO and TCL can get the text detection result after subsequent processing, while the TCL, TDO, and TCC are responsible for recognizing the text.

### 3.6. Reading

The following is the calculation formula for the reading:  $k$  represents the slope of the straight line, where  $k \geq 0$ .  $r$  is the radian value between the straight line where the pointer is located and the horizontal line, with the constraint  $0 \leq r \leq \pi/2$ .  $d$  is the angle value between the straight line where the pointer is located and the horizontal line, with the constraint  $0 \leq d \leq 90$ . Formula (3) is used to convert the radian value,  $r$ , from Formula (2) to an angle value. By using OCR, identify the model and scale value of the meter. Take the maximum scale value on the meter as the range of the meter and use this value as 'max' in Formula (4). 'num' represents the reading of the final pointer, and in Formula (4), 90 represents the rotatable range of the pointer, which is 90 degrees.

$$k = \frac{y_2 - y_1}{x_2 - x_1}, \quad (1)$$

$$r = \arctan(k), \quad (2)$$

$$d = \frac{r}{\pi} \times 180, \quad (3)$$

$$\text{num} = \frac{\text{max} \times d}{90}, \quad (4)$$

### 3.7. Comparison with Existing Technology

Previous studies have used a number of different methods and techniques to address the recognition and reading extraction problems of square pointer meters. These methods are shown in Table 1.

**Table 1.** Methods used in previous studies.

Previous Research	Type	Positioning Instrumentation	Identification Pointers and Readings
Wang et al. [2]	Round pointer instrument	Faster-RCNN	Hough transform
Salomon et al. [3]	water meter	Faster-RCNN, YOLOv3	Classification
Liu et al. [8]	Round pointer instrument	Faster R-CNN	Hough transform
Liu et al. [11]	Round pointer instrument	ORB algorithm	Hough transform
Zuo et al. [13]	Round pointer instrument	Mask RCNN	angle method
Meng et al. [18]	Round pointer instrument	/	key point detection

In past research, scholars mainly focused on exploring circular pointer meters and water meters, but research on the square pointer meters mentioned in this paper is relatively limited. There are obvious differences in shape and structure between the square pointer meter and the traditional circular pointer meter, so it is of great significance to study it deeply.

When locating instruments, commonly used technologies include Faster-RCNN, YOLOv3, ORB algorithm, and Mask RCNN. This paper uses the newly proposed YOLOv7 model, which significantly improves the accuracy of target detection.

In terms of identifying pointers and readings, previous studies such as Wang et al. [2], Liu et al. [8], and Liu et al. [11] directly used Hough transform to extract pointers, but this method is easily affected by complex backgrounds. However, this paper adopts the method of segmenting the pointer area by using image segmentation technology first, and then performing Hough transform, which greatly improves the robustness of instrument recognition and reduces the interference to complex backgrounds. Since the water meter has only 10 different states, Salomon et al. [3] use a classification method to distinguish different readings. However, this method is not suitable for complex ammeters or voltmeters. Conversely, Zuo et al. [13] only used the angle method to determine the reading of the pointer instrument, but in fact the pointer has a width, which may lead to errors in the reading. For this reason, this paper adopts the method of thinning first and then angle method, which can effectively extract the central axis of the pointer and improve the accuracy of reading. Meng et al. [18] used the key point detection method to determine the pointed center point and pointer endpoint of the meter, but due to the long-term use of the meter, the position of the key point may be blocked by stains, which seriously affects the accurate recognition rate of the meter center and pointer endpoint.

In addition, this paper also proposes the use of OCR technology to identify the instrument model and scale value, and take the maximum value of the scale value as the range of the instrument, which greatly improves the identification robustness in instruments of different models and ranges.

This paper proposes a series of new methods for square pointer instruments, including using the YOLOv7 model for target detection, combining image segmentation and Hough transform, extracting pointer readings by thinning first and then angle method, and using OCR technology to identify instruments' model and range. The application of these methods greatly enhances the accuracy and robustness of instrument recognition, making up for the shortcomings of previous studies.

## 4. Experiment

### 4.1. Experimental Environment

The experimental platform is the Ubuntu20.04 operating system, and the software used is CUDA11.6 and CUDNN8401. The experiment is performed under the deep learning

framework based on Pytorch1.12.1 + cu116 and paddle2.3.2 [31]. The computer configuration is shown in Table 2.

**Table 2.** Experimental hardware configuration.

Name	Type
CPU	i7-12700KF
Graphics card	R TX-3090
Memory	64G
Hard disk	6T

#### 4.2. Experimental Dataset

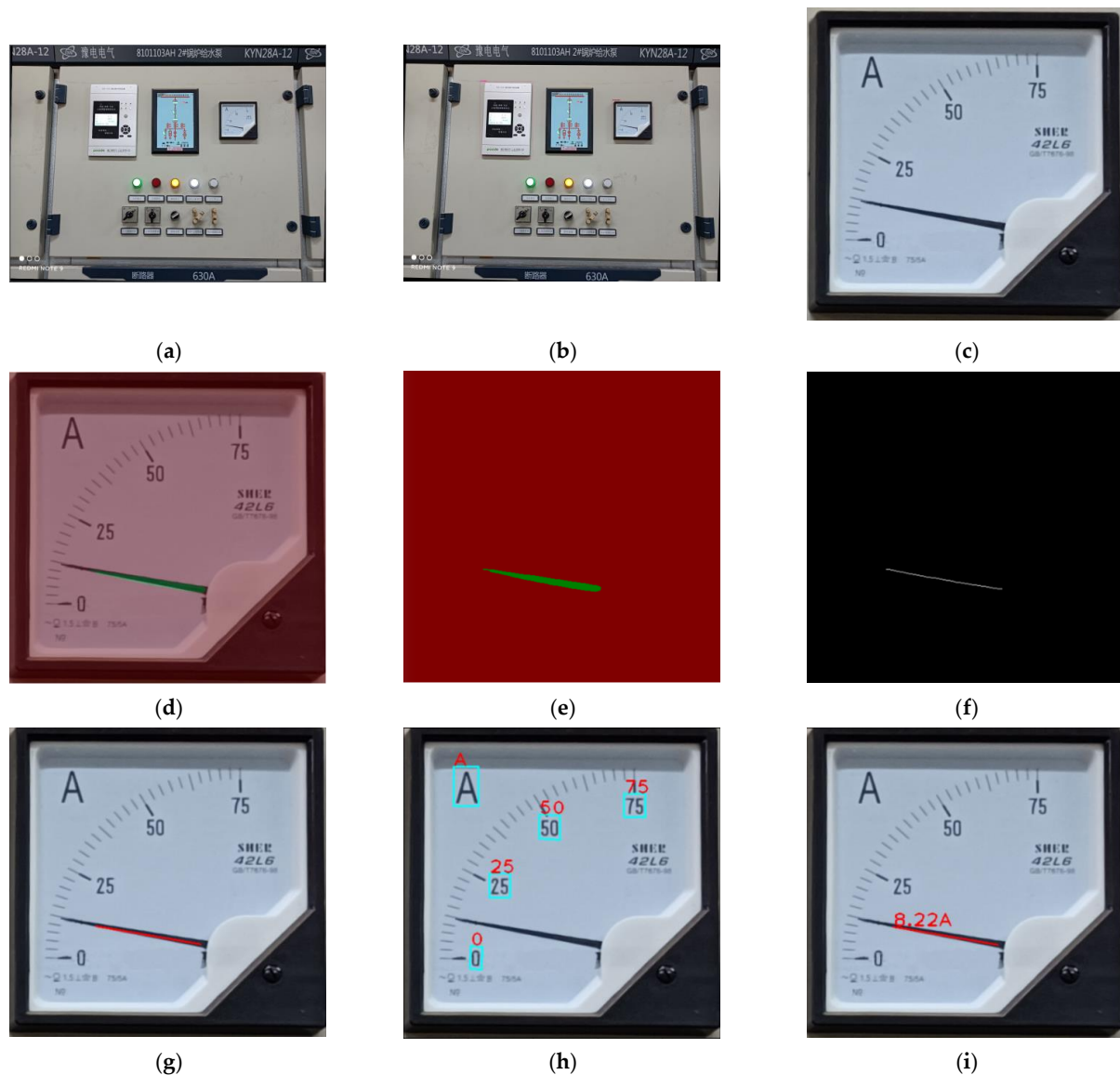
The experimental data set was taken by a mobile phone in the real scene of the enterprise, with a total of 110 pictures, and the number of pictures was enhanced to 5000 through data enhancement technology. The robustness and generalization ability of the object detection algorithm are improved by augmenting the training data. Common data augmentation methods include scaling, translation, mirror flipping, color perturbation, etc. In this experiment, the image enhancement method of mirror flip is not adopted, because the pointer of mirror flip does not exist and is meaningless in reality. The scaling ratio is 0.5–2, and gray bars are added to the excess space in the reduction process to ensure the consistency of the image size. This experimental data set belongs to the engineering application data set in a specific scene and has certain practical value. Some pictures of the experimental data set are shown in Figure 8. The dataset consists of an object detection dataset, image segmentation dataset, and ocr dataset. The object detection data set is labeled with labelme [32], which contains six categories, namely square table, ROC table, round table, complex table, white table, and digital table. This experiment focuses on the reading of the square watch, and other types of watches will be studied in follow-up research. The image segmentation data set is annotated with Baidu PaddlePaddle's interactive segmentation and annotation software EISeg [33], which is used to annotate the pointer part of the meter. The ocr data set is marked with PPOCRLabel [34], a semi-automatic marking tool of Baidu PaddlePaddle, to mark the model and scale value of the meter.



**Figure 8.** Pictures of the experimental data set.

#### 4.3. Experimental Results

The experimental results of the intermediate steps are shown in Figure 9. The nine pictures sequentially represent the original image, the object detection result, the cropped image of the square table, the overlay effect map of the image segmentation, the predicted mask result of the image segmentation, the thinned image, the Hough transform, the ocr result, and the reading.



**Figure 9.** Experimental results of the intermediate steps: (a) Original image; (b) Object detection result; (c) Cropped image of the square table; (d) Overlay effect map of the image segmentation; (e) Predicted mask result of the image segmentation; (f) Thinned image; (g) Hough transform; (h) OCR result; (i) Reading.

#### 4.4. Analysis of Experimental Results

##### 4.4.1. Object Detection Evaluation Index

mAP50, FPS, params, and Gflops are indicators to measure the performance of object detection models [35]. mAP50 (mean average precision at 50) is an indicator used to measure the performance of the object detection algorithm, indicating the average accuracy of the matching degree between the detected object and the real object, where 50 means the IoU (intersection-over-union ratio) threshold is 0.5 [36]. The higher the mAP50, the higher the detection accuracy of the algorithm [37]. FPS (frames per second) indicates the number of image frames that the object detection model can process per second, that is, the inference speed of the model. The higher the FPS, the faster the detection speed of the algorithm. params (parameters) refers to the number of parameters of the model, which is an important indicator for evaluating the complexity of the model. The fewer params, the smaller the model and lower the requirement of computing resources and storage space. Gflops (giga-floating-point-operations-per-second) indicates the number

of floating-point operations performed by the model per second and is an indicator to measure the computational efficiency of the model [38]. The higher the Gflops, the higher the computational efficiency of the model, which can complete the reasoning task faster. Their calculation formulas are in Formulas (5)–(9) [28].

$$AP = \frac{1}{m} \sum_i^m P_i = \frac{1}{m} \times P_1 + \frac{1}{m} \times P_2 + \cdots + \frac{1}{m} \times P_m = \int P(R) dR, \quad (5)$$

$$mAP = \frac{1}{C} (AP_1 + AP_2 + \cdots + AP_c), \quad (6)$$

$$FPS = \frac{1}{T}, \quad (7)$$

$$params = n1 \times m1 + n2 \times m2 + \cdots + nk \times mk, \quad (8)$$

$$Gflops = \frac{(2 \times n1 \times m1 \times h1 \times w1 \times d1 + n2 \times m2 \times h2 \times w2 \times d2 + \cdots + nk \times mk \times hk \times wk \times dk)}{10^9 \times T}, \quad (9)$$

In this experiment, we trained the model 300 times and compared the recognition effects of the five models. After analyzing the experimental results, we found that Yolov7 performed the best. The comparison of different models on the meter dataset is shown in Table 3.

**Table 3.** The comparison of different models on the meter dataset.

Model	Map50	FPS	Params	Gflops	Map5095
Yolov5-s	99.5	204.08	7.03 m	15.8	99.1
Yolov5-m	99.5	85.47	20.87 m	47.9	99.2
Yolov5-l	99.3	51.02	46.14 m	107.7	99.2
Yolox-s	96.7	25.63	8.94 m	26.77	94.51
Yolov7	99.8	92.59	36.51 m	103.2	99.5

According to the table, the maps of Yolov5 and Yolov7 are above 99%, far exceeding Yolox. The Yolov7 model has the highest Map50 score of 99.8, indicating that it performs best in object detection accuracy. Although it has the highest computational resource requirements, it is still the best choice, especially when accuracy is a key metric. Yolov5-s has the least number of parameters and Gflops, so it is the most lightweight model and requires the least computational resources. Additionally, Yolov5-s has the highest FPS, so it is suitable for tasks that need to process images quickly. The choice of which model to use depends on the specific application scenario and requirements. If you need high-precision object detection and have enough computing resources, you can choose the Yolov7 model. If you need to process images quickly or have limited computing resources, you can choose Yolov5-s.

#### 4.4.2. Image Segmentation Evaluation Index

In the field of image segmentation, the evaluation model quality is mainly judged by three indicators, accuracy rate (acc), average intersection over union (mean Intersection over Union, referred to as mIoU), and Kappa coefficient. The accuracy rate refers to the proportion of the pixels with correct category predictions to the total pixels. The higher the accuracy rate, the better the model quality [39]. The average intersection ratio is calculated separately for each data set category, the calculated intersection of the predicted area and the actual area is divided by the union of the predicted area and the actual area, and then the results obtained for all categories are averaged [40]. The Kappa coefficient, also known



as Cohen's kappa, is used to measure the accuracy of a classifier. The Kappa coefficient takes into account the agreement between the classifier and the random classifier, and its value ranges from  $-1$  to  $1$ , where  $1$  indicates complete agreement,  $0$  indicates agreement with the random classifier, and  $-1$  indicates no agreement at all. The higher the Kappa coefficient, the better the model quality.

The formulas of accuracy rate, average intersection ratio and Kappa coefficient are shown in Formulas (10)–(13) [33].

$$\text{Acc} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})}, \quad (10)$$

$$\text{mIOU} = \frac{1}{n} \times \sum \frac{\text{TP}}{(\text{TP} + \text{FP} + \text{FN})}, \quad (11)$$

$$e = \frac{(\text{TP} + \text{FP}) \times (\text{TP} + \text{FN}) + (\text{FN} + \text{TN}) \times (\text{FP} + \text{TN})}{(\text{TP} + \text{FP} + \text{TN} + \text{FN})^2}, \quad (12)$$

$$\text{kappa} = \frac{\text{acc} - e}{1 - e}, \quad (13)$$

Different models for image segmentation on the pointer dataset are shown in Table 4:

**Table 4.** Results of different models for image segmentation on the pointer dataset.

Model	MioU	Acc	Kappa	Dice	Class IoU	Class Precision	Class Recall
ann	0.8084	0.9967	0.7639	0.8819	[0.9967, 0.6202]	[0.9984, 0.7609]	[0.9983, 0.7703]
pspnet	0.807	0.9966	0.7618	0.8809	[0.9966, 0.6175]	[0.9985, 0.7431]	[0.9981, 0.7851]
deplabv3	0.8029	0.9966	0.7555	0.8777	[0.9966, 0.6093]	[0.9984, 0.7475]	[0.9982, 0.7672]
deplabv3+	0.8773	0.9981	0.8604	0.9302	[0.998, 0.7565]	[0.9991, 0.8577]	[0.999, 0.865]
danet	0.8059	0.9965	0.7601	0.88	[0.9965, 0.6153]	[0.9985, 0.7337]	[0.998, 0.7921]

In the image segmentation experiments, we used the models in Table 4 to train and evaluate on the pointer dataset. The results show that the deplabv3p model has excellent performance in all aspects. The mIoU (mean Intersection over Union ratio), acc (pixel accuracy), Kappa coefficient, and Dice coefficient of the model are 0.8773, 0.9981, 0.8604, and 0.9302, respectively. In addition, the category IoU index of the model is [0.998, 0.7565], indicating the performance difference of the model on different categories; the category precision is [0.9991, 0.8577], reflecting the classification accuracy of the model for each category, and the category recall rate is [0.999, 0.865], indicating the recognition ability of the model for each category. These evaluation indicators show that the deplabv3p model has high mIoU, acc, Kappa coefficient, and Dice coefficient when performing image segmentation tasks, and can also achieve high IoU, precision, and recall on different categories. This shows that the model has high accuracy and stability when dealing with image segmentation tasks and can play a role in a variety of practical application scenarios.

#### 4.4.3. OCR Evaluation Index

PaddleOCR calculates three OCR end-to-end related indicators, namely precision (precision rate), recall (recall rate), and Hmean (F1 value). Precision refers to the ratio of the number of correct characters in the OCR recognition result to the total number of characters. The recall rate refers to the ratio of the number of correct characters in the OCR recognition result to the total actual number of characters. The Hmean is the harmonic mean of precision and recall. Their calculation formulas are in Formulas (14)–(16) [41]:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad (14)$$



$$\text{Recall} = \frac{TP}{TP + FN}, \quad (15)$$

$$\text{Hmean} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \quad (16)$$

Among them, TP represents the number of characters correctly recognized by OCR, and FP represents the number of characters incorrectly recognized by OCR [42].

In this experiment, the accuracy rate of PGNet is 98.04%, the recall rate is 81.97%, and the Hmean rate is 89.29%. The reason why the Hmean value is not high may be that the pointer will cover the scale value in some cases, resulting in the inability to accurately identify the scale, which is a common problem. As shown in Figure 10, the pointer covers the “0” scale, making it impossible to accurately identify the “0” scale value. However, under normal circumstances, the pointer will not reach the maximum value of the meter, which has little effect on the operation of taking the maximum scale value.



**Figure 10.** The situation where the pointer covers the scale value.

#### 4.4.4. Meter Automatic Reading Evaluation Index

In this experiment, the absolute error, relative error, and accuracy are used as the measurement indicators of the meter reading, and their calculation formulas are in Formulas (17)–(19). Among them, the manual reading is performed by 20 workers, and their average value is taken as the manual reading. Results are rounded to two decimal places. The comparison between automatic identification and manual identification of pointer meters is shown in Table 5. The interpretation of each abbreviation is shown in Table 6.

$$\text{Absolute Error} = |\text{Automatic Reading} - \text{Manual Reading}|, \quad (17)$$

$$\text{Relative Error} = \frac{\text{Absolute Error}}{\text{Maximum Range}}, \quad (18)$$

$$\text{Accuracy} = 1 - \text{Relative Error}, \quad (19)$$

In this data table, we can see that there are certain errors between the automatic readings and the manual readings, but in most cases the relative errors are small, and the accuracy is high. The accuracy of the readings obtained by this scheme is not less than 95%. At the same time, we can find that the automatic identification meter type is consistent with the manual identification meter type, and the maximum range of automatic identification is basically the same as that of manual identification, which shows that the automatic identification algorithm of the meter proposed in this experiment is relatively reliable. Therefore, this solution can be used as an effective application of intelligent measurement technology in the industrial field, thereby improving the accuracy and efficiency of measurement and reducing the cost of manual operation.

**Table 5.** Comparison between automatic identification and manual identification of pointer meters.

AR	MR	AT	MT	AM	MM	WT	WM	AE	RE	ACC
8.22	11.90	A	A	75	75	Y	Y	3.68	0.05	95.09
9.18	8.76	kV	kV	12	12	Y	Y	0.42	0.04	96.50
10.55	10.40	kV	kV	12	12	Y	Y	0.15	0.01	98.75
71.32	73.10	A	A	150	150	Y	Y	1.78	0.01	98.81
21.75	23.32	A	A	50	50	Y	Y	1.57	0.03	96.86
120.78	115.25	V	V	300	300	Y	Y	5.53	0.02	98.16
10.68	10.05	A	A	15	15	Y	Y	0.63	0.04	95.80
161.35	155.88	V	V	250	250	Y	Y	5.47	0.02	97.81
3.47	3.05	A	A	20	20	Y	Y	0.42	0.02	97.90
123.56	128.92	V	V	150	150	Y	Y	5.36	0.04	96.43

**Table 6.** The interpretation of each abbreviation.

Name	Explanation
AR	automatic reading
MR	manual reading
AT	automatic identification of meter model
MT	manual recognition of meter model
AM	automatic recognition of the maximum range
MM	manual recognition of the maximum range
WT	whether the type is the same
WM	whether the maximum value is the same
AE	absolute error
RE	relative error
ACC	accuracy

## 5. Conclusions

In this paper, the deep learning method and computer vision technology are combined to propose a pointer meter recognition method based on Yolov7 and Hough transform, which realizes the automatic recognition and reading of pointer meters. Deep learning includes Yolov7, DeepLabv3+ and PGNet: use Yolov7 to determine the position of the meter; use DeepLabv3+ to extract the pointer part of the meter; and use PGNet to determine the model and range of the meter. Computer vision includes Thinning and Hough transforms, which can determine the precise position of needles and take meter readings. By harnessing the advantages of deep learning in object detection and image segmentation, it becomes possible to accurately locate the instrument and segment the pointer directly from the image, thereby avoiding the cumbersome process of traditional computer vision for pointer extraction. Therefore, the method effectively solves the problems of uneven illumination, complex background, image blur, and diversity of meter models in each image. The experimental results demonstrate that the utilization of the latest Yolov7 model achieved a remarkable map of 99.8% on the instrument dataset. The accuracy of the pointer readings obtained using this method is more than 95%. The method performs well in terms of anti-interference, accuracy, and robustness. When meter images are captured at highly standard shooting angles, they can be quickly and accurately recognized. However, when meter images are captured at larger angles, the recognition accuracy cannot meet the required standards. In the next phase of the work, efforts will be made to correct the meter images in order to improve the recognition accuracy.

**Author Contributions:** Writing—review & editing, L.S., D.Z., T.K. and J.L.; Supervision, C.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** <https://github.com/shilei0406/Pointer-instrument-identification> (accessed on 2 April 2023).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Zhang, Z.Y.; Hua, Z.X.; Tang, Y.C.; Zhang, Y.J.; Lu, W.J.; Dai, C.F. Recognition Method of Digital Meter Readings in Substation Based on Connected Domain Analysis Algorithm. *Actuators* **2021**, *10*, 170. [CrossRef]
2. Wang, L.; Wang, P.; Wu, L.H.; Xu, L.J.; Huang, P.; Kang, Z.L. Computer Vision Based Automatic Recognition of Pointer meters: Data Set Optimization and Reading. *Entropy* **2021**, *23*, 272. [CrossRef]
3. Salomon, G.; Laroca, R.; Menotti, D. Deep Learning for Image-based Automatic Dial Meter Reading: Dataset and Baselines. *arXiv* **2020**, arXiv:2005.03106. Available online: <https://ui.adsabs.harvard.edu/abs/2020arXiv200503106S> (accessed on 25 April 2023).
4. Brunelli, R. *Template Matching Techniques in Computer Vision: Theory and Practice*; John Wiley & Sons: Hoboken, NJ, USA, 2009.
5. Abiodun, O.I.; Jantan, A.; Omolara, A.E.; Dada, K.V.; Umar, A.M.; Linus, O.U.; Arshad, H.; Kazaure, A.A.; Gana, U.; Kiru, M.U. Comprehensive review of artificial neural network applications to pattern recognition. *IEEE Access* **2019**, *7*, 158820–158846. [CrossRef]
6. Zheng, Y.; Li, G.; Li, Y. Review of the Application of Deep Learning in Image Recognition. *Comput. Eng. Appl.* **2019**, *55*, 20–36.
7. Bhatt, D.; Patel, C.; Talsania, H.; Patel, J.; Vaghela, R.; Pandya, S.; Modi, K.; Ghayvat, H. CNN variants for computer vision: History, architecture, application, challenges and future scope. *Electronics* **2021**, *10*, 2470. [CrossRef]
8. Liu, Y.; Liu, J.; Ke, Y.C. A detection and recognition system of pointer meters in substations based on computer vision. *Measurement* **2020**, *152*, 107333. [CrossRef]
9. Wu, X.; Shi, X.; Jiang, Y.; Gong, J. A High-Precision Automatic Pointer Meter Reading System in Low-Light Environment. *Sensors* **2021**, *21*, 4891. [CrossRef]
10. Li, Z.; Zhou, Y.; Sheng, Q.; Chen, K.; Huang, J. A High-Robust Automatic Reading Algorithm of Pointer Meters Based on Text Detection. *Sensors* **2020**, *20*, 5946. [CrossRef]
11. Liu, Z.; Huang, H.; Wang, N.; Cao, Y.; Zeng, L.; Zhang, J.; Zhang, C. A pointer meter reading recognition method based on improved ORB algorithm for substitution inspection robot. *J. Phys. Conf. Ser.* **2022**, *2189*, 012027. [CrossRef]
12. Cai, W.D.; Ma, B.; Zhang, L.; Han, Y.M. A pointer meter recognition method based on virtual sample generation technology. *Measurement* **2020**, *163*, 107962. [CrossRef]
13. Zuo, L.; He, P.; Zhang, C.; Zhang, Z. A Robust Approach to Reading Recognition of Pointer Meters Based on Improved Mask-RCNN. *Neurocomputing* **2020**, *388*, 90–101. [CrossRef]
14. Laroca, R.; Barroso, V.; Diniz, M.A.; Gonçalves, G.R.; Schwartz, W.R.; Menotti, D. Convolutional neural networks for automatic meter reading. *J. Electron. Imaging* **2019**, *28*, 013023. [CrossRef]
15. Salomon, G.; Laroca, R.; Menotti, D. Image-based Automatic Dial Meter Reading in Unconstrained Scenarios. *arXiv* **2022**, arXiv:2201.02850. Available online: <https://ui.adsabs.harvard.edu/abs/2022arXiv220102850S> (accessed on 2 April 2023). [CrossRef]
16. Bishwokarma, R.; Paudyal, B.; Chapagain, P.; Bajgain, S.; Shakya, H.D. Deep Neural Network based Automatic System for Electricity Meter Reading in Nepal. In Proceedings of the International Conference on “Role of Energy for Sustainable Social Development in New Normal Era”, Kathmandu, Nepal, 28–29 December 2021.
17. Yan, X.; Wei, L.; Li, J.; Jia, G.; Yang, J. Research on Automatic Recognition of Pointer Meter Reading Based on Deep Learning Algorithm. *J. Phys. Conf. Ser.* **2021**, *1865*, 042017. [CrossRef]
18. Meng, X.; Cai, F.; Wang, J.; Lv, C.; Liu, H.; Liu, H.; Shuai, M. Research on Reading Recognition Method of Pointer Meters Based on Deep Learning Combined with Rotating Virtual Pointer. In Proceedings of the 5th International Conference on Information Science, Computer Technology and Transportation (ISCTT), Shenyang, China, 13–15 November 2020; pp. 115–118. [CrossRef]
19. Zhang, X.; Dang, X.; Lv, Q.; Liu, S. A pointer meter recognition algorithm based on deep learning. In Proceedings of the 2020 3rd International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE), Shenzhen, China, 24–26 April 2020; pp. 283–287.
20. Laroca, R.; Araujo, A.B.; Zanlorensi, L.A.; de Almeida, E.C.; Menotti, D. Towards Image-based Automatic Meter Reading in Unconstrained Scenarios: A Robust and Efficient Approach. *arXiv* **2020**, arXiv:2009.10181. Available online: <https://ui.adsabs.harvard.edu/abs/2020arXiv200910181L> (accessed on 5 April 2023). [CrossRef]
21. Dong, Z.; Gao, Y.; Yan, Y.; Chen, F. Vector Detection Network: An Application Study on Robots Reading Analog Meters in the Wild. *arXiv* **2021**, arXiv:2105.14522. Available online: <https://ui.adsabs.harvard.edu/abs/2021arXiv210514522D> (accessed on 15 April 2023). [CrossRef]
22. Bayhan, E.; Ozkan, Z.; Namdar, M.; Basgumus, A. Deep learning based object detection and recognition of unmanned aerial vehicles. In Proceedings of the IEEE 2021 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), Ankara, Turkey, 11–13 June 2021; pp. 1–5.
23. Ozkan, Z.; Bayhan, E.; Namdar, M.; Basgumus, A. Object detection and recognition of unmanned aerial vehicles using Raspberry Pi platform. In Proceedings of the 2021 IEEE 5th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), Ankara, Turkey, 21–23 October 2021; pp. 467–472.

24. Girshick, R. Fast r-cnn. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 1440–1448.
25. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. Ssd: Single shot multibox detector. In Proceedings of the Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, 11–14 October 2016; Part I 14; Springer: Berlin/Heidelberg, Germany; pp. 21–37.
26. Jiang, P.; Ergu, D.; Liu, F.; Cai, Y.; Ma, B. A Review of Yolo algorithm developments. *Procedia Comput. Sci.* **2022**, *199*, 1066–1073. [\[CrossRef\]](#)
27. Wang, C.-Y.; Bochkovskiy, A.; Liao, H.-Y.M. Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *arXiv* **2022**, arXiv:2207.02696.
28. Bian, H.; Liu, Y.; Shi, L.; Lin, Z.; Huang, M.; Zhang, J.; Weng, G.; Zhang, C.; Gao, M. Detection Method of Helmet Wearing Based on UAV Images and Yolov7. In Proceedings of the 2023 IEEE 6th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), Chongqing, China, 24–26 February 2023; Volume 6, pp. 1633–1640.
29. Ding, L.; Wang, J.; Wu, Y. Electric power line patrol operation based on vision and laser SLAM fusion perception. In Proceedings of the 2021 IEEE 4th International Conference on Automation, Electronics and Electrical Engineering (AUTEEE), Shenyang, China, 19–21 November 2021; pp. 125–129.
30. Wang, P.; Zhang, C.; Qi, F.; Liu, S.; Zhang, X.; Lyu, P.; Han, J.; Liu, J.; Ding, E.; Shi, G. Pgnet: Real-time arbitrarily-shaped text spotting with point gathering network. In Proceedings of the AAAI Conference on Artificial Intelligence, Virtual, 2–9 February 2021; Volume 35, pp. 2782–2790.
31. Ma, Y.; Yu, D.; Wu, T.; Wang, H. PaddlePaddle: An open-source deep learning platform from industrial practice. *Front. Data Computing* **2019**, *1*, 105–115.
32. Torralba, A.; Russell, B.C.; Yuen, J. Labelme: Online image annotation and applications. *Proc. IEEE* **2010**, *98*, 1467–1484. [\[CrossRef\]](#)
33. Liu, Y.; Chu, L.; Chen, G.; Wu, Z.; Chen, Z.; Lai, B.; Hao, Y. Paddleseg: A high-efficient development toolkit for image segmentation. *arXiv* **2021**, arXiv:2101.06175.
34. Du, Y.; Li, C.; Guo, R.; Cui, C.; Liu, W.; Zhou, J.; Lu, B.; Yang, Y.; Liu, Q.; Ma, Y.; et al. Pp-ocrv2: Bag of tricks for ultra lightweight ocr system. *arXiv* **2021**, arXiv:2109.03144.
35. Hu, Z.; Deng, Y.; Lan, J.; Wang, T.; Han, Z.; Huang, Y.; Zhang, H.; Wang, J.; Cheng, M.; Chen, G.; et al. A multi-task deep learning framework for perineural invasion recognition in gastric cancer whole slide images. *Biomed. Signal Process. Control* **2023**, *79*, 104261. [\[CrossRef\]](#)
36. Shuo, H.; Ximing, Y.; Donghang, L.; Shaoli, L.; Yu, P. Digital recognition of electric meter with deep learning. In Proceedings of the 2019 14th IEEE International Conference on Electronic Measurement & Instruments (ICEMI), Changsha, China, 1–3 November 2019; pp. 600–607.
37. Liang, Q.; Wang, W.; Liu, X.; Na, Z.; Jia, M.; Zhang, B. (Eds.) *Communications, Signal Processing, and Systems: Proceedings of the 8th International Conference on Communications, Signal Processing, and Systems*; Springer Nature: Berlin/Heidelberg, Germany, 2020.
38. Yan, X.; Jia, L.; Cao, H.; Yu, Y.; Wang, T.; Zhang, F.; Guan, Q. Multitargets joint training lightweight model for object detection of substation. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, 1–12. [\[CrossRef\]](#)
39. Liang, C.M.; Li, Y.W.; Liu, Y.H.; Wen, P.F.; Yang, H. Segmentation and weight prediction of grape ear based on SFNet-ResNet18. *Syst. Sci. Control Eng.* **2022**, *10*, 722–732. [\[CrossRef\]](#)
40. Cheng, Z.; Wang, Z.; Huang, H.; Liu, Y. Dense-acssd for end-to-end traffic scenes recognition. In Proceedings of the 2019 IEEE Intelligent Vehicles Symposium (IV), Paris, France, 9–12 June 2019; pp. 460–465.
41. Bashyam, V.; Taira, R.K. Identifying Anatomical Phrases in Clinical Reports by Shallow Semantic Parsing Methods. In Proceedings of the 2007 IEEE Symposium on Computational Intelligence and Data Mining, Honolulu, HI, USA, 1–5 April 2007; pp. 210–214.
42. Peng, H.; Yu, J.; Nie, Y. Efficient Neural Network for Text Recognition in Natural Scenes Based on End-to-End Multi-Scale Attention Mechanism. *Electronics* **2023**, *12*, 1395. [\[CrossRef\]](#)

**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.