Research on Water Meter Reading Recognition Based on Deep Learning

Yue Liang
Nanchang University
Wenjuan Wu
Nanchang University
Shaobo Li
Nanchang University
Yiqi Liao
Nanchang University
Taorong Qiu (qiutaorong@ncu.edu.cn)
Nanchang University

Research Article

Keywords: Water Meter, table picture, labor costs

Posted Date: January 19th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1234493/v1

Abstract

Most of the current water meters are manually carried with a palm-integrated machine to read the meters on the spot, which has low labor efficiency and high labor costs, which leads to problems such as low efficiency for the entire water department. In response to this problem, this article aims to in-depth study and analyze a suitable water meter image recognition model to improve labor efficiency, reduce labor costs, and thereby improve the overall efficiency of the water department. The main content of this paper is to select a model with high recognition accuracy for the numbers in the black rectangle. This paper first introduces the existing in-depth learning models, such as Faster RCNN, SSD, and YOLOv3. Then two datasets are sorted out, one is the original water table picture dataset, the other is the data set cut out of a black rectangle box. Then two plans are proposed, one is to train the whole picture directly and recognize the water meter digits through the in-depth learning network, the other is to train the black rectangle picture and recognize the water meter digits through the in-depth learning network. Finally, by comparing the three models from different angles, it is determined that YOLOv3 in plan B has the best recognition effect, and the accuracy rate reaches 90.61%, which can greatly improve the labor efficiency and save labor costs.

Introduction

Current water meters are mostly read on the spot by hand, and the management personnel of the business office is responsible for the arrangement of the meter-reading personnel, the adjustment of the meter-reading circuit and the formulation of the meter-reading plan. The meter reader downloads the meter reading task to the water supply management department through the spot by hand, and then carries the meter reading machine to the spot to read the meter. After completing the meter reading plan for the day, the meter reading result is uploaded to the meter reading system of the water supply management department. In this working mode, there are many problems: meter reading personnel reading efficiency is low, heavy task; It is difficult to arrange the meter reading plan. The meter reading results fed back by meter reading personnel cannot be judged accurate or not; The cost of meter reading is high and so on, so there is an urgent need for effective means to improve work efficiency and reduce personnel expenses.

In recent years, the development of artificial intelligence technology has become more and more mature, in which deep learning learns a large amount of data and automatically extracts features, and then trains specific algorithm models to obtain output results[1]. Deep learning is widely used in various aspects, and target detection is one of them[2]. The role of target detection is to detect specific categories of targets on pictures, which can be widely used in various fields of production and life, such as military field[3], biological field[4] and industrial manufacturing field[5] and Object detection field[6]. With the maturity of image recognition technology in computer vision, more and more target detection algorithms have been proposed. From RCNN[7][8] to MASKRCNN[9], RCNN series (regions with CNN features), YOLO series[10][11][12](You only look once), SSD[13] (Single Shot Multi Box Detector) and so on, the recognition accuracy in

improving step by step, both the speed of recognition. Therefore, the application of Internet and deep learning technology to meter reading companies can greatly improve labor efficiency and save labor costs.

Related Studies

Faster RCNN algorithm for target detection based on candidate regions

In Fasterr CNN[14], a candidate region detection box is generated by introducing Region Proposal Networks (RPN network[15]), and Anchor Box is introduced to integrate feature extraction, candidate region recommendation, bounding box regression and classification into a network. And share the features extracted by the convolutional layer. The backbone network for Fasterr CNN is VGG16.The algorithm consists of two steps: first, the RPN network determines whether the candidate box is the target or not, and then the target type is determined by the multi-task loss objective function. The network structure is shown in Figure 1.

The specific algorithm steps are as follows:

(1) The convolutional layer extracts feature maps from the input images for subsequent networks and RPN networks.

(2) After the feature map is input into the RPN network, the candidate region is obtained, and the candidate box matrix and its score are output.

(3) The feature map and candidate regions are pooled to output the proposal Feature Map through the Roi pooling layer, and then sent to the subsequent full-connection layer.

(4) The full-connection layer conducts classification and border regression according to proposal Feature Map to get the position and score information of the final detection box.

Target detection SSD method based on regression

The SSD model takes VGG16 as the main network structure, and changes the last two full connection layers of VGG16 into the convolution layer, and adds four convolution layers to construct the SSD network structure. SSD uses regression mode to quickly detect the category and location of objects. Also, feature maps of different scales are extracted based on regions to detect target regions. A total of 6 feature maps were extracted from the SSD model for detection, which presented an inverted pyramid structure from large to small. The model structure is shown in Figure 2.

Yolov3 algorithm for target detection based on regression

The YOLOV3 model is composed of two structures, one is the feature extraction layer (Darknet53)[16] and the other is the YOLO detection layer. Darknet53 network consists of convolutional layer and residual unit. The convolutional layers are made up of 1*1 and 3*3 convolutional kernels, using batch normalization and LeakvReLU activation functions at each convolutional layer[17].

Yolov3 generates the position of the bounding box and category information of the entire image through the neural network in one step. The feature map of a specific size is obtained through the feature extraction network, and the image is divided into $S*S$ networks. The grid is responsible for detecting the target and predicting N bounding boxes, including position information (x,y,w,h) and a confidence level[18]. This format uses $Pr\ (Object)*IOU$ to determine the accuracy information of the subject and the bounding box. When $Pr\ (Object)$ equals 1, it means there is an Object, and when Pr (Object) is 0, it means there is no Object. IOU is the value of the union over the intersection ratio of the predicted border and the real border, and the maximum value of IOU can be used to predict the object category. If there are K classes, YOLOV3 will output $S*S*(N*(4+1+K))$. $S*S$ is the size of feature graph, N is the number of boundary boxes, K is the number of categories, 4 represents (x,y,w,h), and 1 represents confidence.

Materials And Method

This data is obtained by meter reading personnel from 5 water supply branches and subordinate offices of Jiaxing Water Supply and Drainage Co., Ltd., who take pictures of water meters in real scenes with mobile phone cameras within the specified scope. Some images of water meters are shown in Figure 3.

Preprocessing of water meter image data set

Build a black rectangular box for target detection network recognition

Due to the obvious features of the black rectangular box, it is easy to identify. After comparative analysis of Fasterr CNN, SSD and YOLOV3, this paper selects the YOLOV3 target detection network with good recognition accuracy and speed to identify and cut the black rectangular box, and obtains the black rectangular box water meter image data set. The steps to build a Yolov3 network are as follows:

First, Calculate the width and height of the mark. All the black rectangular boxes are manually marked as the same class (this class is named screen) through the labelImg widget, generate xml files one by one, read out the coordinates in the xml file, calculate the width and height of the mark box, namely w and h, by calculating the values of *xmax - xmin* and *ymax - ymin*.The image labeling effect is shown in Figure 4.

Second,Clustering algorithm is used to design the target detection box which conforms to the standard. Calculate all w and h of the annotation box and gather 9 center points to design 9 target detection boxes.

(1) Nine target detection boxes are gathered by K-means algorithm. Build an array object with each group w and h. There are 3600 array objects in total, form a two-dimensional array, initialize 9 data points as the center of mass (w_1, h_1), (w_2, h_2), (w_3, h_3), ..., (w_9, h_9), that is, nine classes. Calculate the distance from the center of mass for each array object in a two-dimensional array, such as (w_a, w_b). If (w_a, w_b) is closer to (w_1, h_1), then (w_a, w_b) is divided into the class (w_1, h_1). After the first cycle, for each class, the average value of w and the aaverage value of h are calculated to obtain (w_1', h_1') as the new center of mass, and then the distance and new center of mass are recalculated until no new center of mass position is generated. The classification effect figure 5 is as follows.

(2) Hierarchical clustering method was used to divide 9 target detection boxes.3600 array objects equals 3600 clusters. The Euclidean distance between clusters[19] is calculated to determine whether it is the same class. Define a distance threshold. If the distance between clusters is closer and does not exceed the threshold (the threshold is set as 130 here), the two clusters are classified into one cluster; if the threshold is exceeded, the two clusters are divided into two categories. The center of mass of the clusters is represented by the mean value of the pair-distance of their respective data points.

Part of the effect of hierarchical clustering is shown in Figure 6.

The classification comparison between K-means and hierarchical clustering algorithm is shown in Table 1, and its distribution is shown in Figure 7.

Table 1
Comparison of clustering algorithms

Clustering algorithm	Total width difference	Total height difference	Average width	Average height
K-means	593.23	113.39	210.02	60.42
Hierarchical clustering	67.58	14.48	310.51	84.98

It can be seen from the figure that, compared with the classes divided by hierarchical clustering, the categories divided by K-means are more obvious in width and height, and can better classify the target box of identifying large target, medium target and small target. Combined with the water meter image data set, K-means can be used for better classification. The effect of some image target boxes is shown in Figure 8.

Third, The training model of migration[20]. In this experiment, model parameters pre-trained by COCO were selected for fine-tuning.

The specific training steps are as follows:

(1) Preprocess the water meter image, unify the image size, for example: 416*416, shrink the long side to 416, shrink the short side and fill it with pixel RGB (128,128,128). The processed image is shown in Figure

9.

(2) Select the pre-training model. In this experiment, model parameters pre-trained by COCO were selected for fine-tuning.

(3) Process the data set and set the data set to VOC format.

(4) Set the parameters of the experiment. Some parameter Settings are shown in Table 2.

Table 2
identifies part of the network parameters of the black rectangular box

Parameter Names	Parameter Values
BATCH_SIZE	2
IOU_LOSS_THRESH	0.5
ANCHOR_PER_SCALE	3
LEARN_RATE_INIT	1e-4
LEARN_RATE_END	1e-6
EPOCH	30×1800

Epoch represents the number of training sessions, 30 of 1,800 each. Batch size represents the two images used at each time. After repeated experiments, the batch size of 2 can meet the environmental requirements of this experiment. The learning rate is set between 1e-6 and 1e-4.

(5) Training and testing.

The data of the black rectangular box is processed

When marking the image, considering that there are numbers in other areas of the dial, in order to further eliminate the interference of other numbers, the black edge above and below the number of the water meter is also put into the box when marking. The effect is shown in Figure 10.

a b

Figure 10 Annotations of some water meter images

Resize the image according to the network, such as 416*416, as shown in Figure 11.

Split the Training Dataset and the Test Dataset

The initial water meter images in this paper are 4000 pieces, and the water meter images after the segmentation of the black rectangle are 4000 pieces in total. In this paper, the two data sets are randomly

divided into 10 pieces, 9 pieces as training data sets and 1 piece as test data sets. Therefore, the two data sets respectively have 3600 pieces of training data sets and 400 pieces of test data sets.

Data enhancement of water meter image

Data enhancement is to expand more data by changing the picture data of the training and synchronously changing the position information of the label, but not changing the category of the label, so as to improve the quality of the picture. Meanwhile, this paper only expands the training set. In computer vision, many image transformation methods are used to expand data sets. In this paper, the following data expansion methods are used in the image preprocessing stag

(1)The original image was flipped left and right to expand the data set and increase the diversity of the data.

(2)The original image is cropped randomly. Set the random function and the clipping threshold (for example, 0.5). If the value of the random function is less than the clipping threshold, it will be clipped randomly; otherwise, it will not be clipped randomly.

(3)Translate the original image. The fruit images in the data set were randomly shifted horizontally within the range of [0, width*0.1] or vertically within the range of [0, height*0.1].

After the above geometric transformation, the images were amplified by about 3 times, and the whole training set was expanded to 14400 images. An example diagram of a water meter turning left and right is shown in Figure 12.

a b

Figure 12 Water meter image after geometric transformation

Construction and selection of reading recognition model

Deep learning model building environment

Fasterr CNN uses the candidate boxes generated by RPN algorithm, and the candidate boxes are the same network as the CNN network for target detection, which reduces the number of candidate boxes to 300 and improves the quality of the selection boxes at the same time. To solve the problem of slow speed in Fasterr CNN, SSD adopts the idea of feature stratifying extraction, border regression and scoring simultaneously, and integrates the Anchors idea in Fasterr CNN, which is suitable for multi-scale target detection and faster at the same time.YOLOV3 is a target detection network with a very balanced speed and accuracy. Combined with this experimental environment, this paper uses Fasterr CNN in the two-stage target detection algorithm and SSD and YOLOV3 in the one-stage target detection algorithm to train the classification model of water meter graphics. The introduction of the experimental environment is shown in Table 3.

Table 3
Introduction to the experimental environment

CPU	AMD Ryzen 7 1700 Eight-Core Processor 3.00GHz
Memory	16GB
Operating System	windows64
GPU	NVIDAGeForce GTX 1060 6GB
Programming Language	Python3.7
Computer image vision library	OpenCV
Deep Learning Framework	TensorFlow2
CUDA Version	10.0

Selection of model plans

*First, construct Faster RCNN to identify water meter numbers.*The specific training steps are as follows:

(1) Make the data set in VOC format. There are 14400 trainval files (training set) and 400 test files (test set).

(2) Debug the environment. The .pyx files of cython_bbox and bbox are compiled to generate .c files so that the files can be run on windows to generate bounding box.

(3) Determine the pre-training model. Select the public VGG16 as the pre-training model and fine-tune its parameters.

(4) Debugging parameters are shown in Table 4.

(5) Training and testing.

Table 4
Faster RCNN partial parameters

Parameter Names	Parameter Values
BATCH_SIZE	32
LEARNING_RATE	0.000001
EPOCH	12000
RPN_POSITIVE_OVERLAP	0.7
RPN_NEGATIVE_OVERLAP	0.3

Second, construct SSD to identify water meter image. The specific training steps are as follows:

(1) Preprocessing on the water meter image. Scale the size of the image to 300*300.Make VOC data sets and convert images and labels to TF Records format.

(2) Determine the pre-training model and fine-tune it using the public VGG16 model parameters.

(3) Set experimental parameters as shown in Table 5.

(4) Training and testing

Table 5
Partial experimental parameters of SSD

Parameter Names	Parameter Values
BATCH_SIZE	3
END_LEARNING_RATE	1e-6
EPOCH	12000
MATCH_THRESHOLD	0.5
NUM_CLASS	11
TRAIN_IMAGE_SIZE	300

Since the image size is changed to 300*300, the batch size setting of 3 can best meet the requirements of the experiment. When the learning rate is greater than 0.001, the loss function of the model does not converge. Through many experiments, 1E-6 was selected as the learning rate of this experiment.

Third, construction of YOLOV3 recognition water meter image.

(1) Nine Anchorbox were selected by clustering algorithm. By calculating w and h of each annotation box, an array object is constructed with each group of w and h. There are a total of 16,402 array objects. The K-means algorithm and hierarchical clustering algorithm described in the second step of Section 3.2.1 are used respectively to gather 9 AnchorBox classes. The experimental results are shown in Table 6, and the category distribution is shown in Figure 13.

Table 6
Experimental results of K-means and hierarchical clustering

Clustering algorithm	Total width difference	Total height difference	Average width	Average height
K-means	93.8	106.5	30.62	34.86
Hierarchical clustering	13.88	16.56	40.54	47.84

It can be seen from the experimental results that the 9 AnchorBox classes gathered by the K-means clustering algorithm are more average compared with the hierarchical clustering algorithm, and can be used for multi-scale recognition with sizes ranging from [20, 23] to [114.03, 129.53], which is more in line with the design of YOLOv3. So select the 9 AnchorBox separated by K-means.

(2) Use transfer learning. The specific training steps are as follows:

 Resize the image to 416 by 416.

 Pre-training model selection. In this experiment, model parameters pre-trained by COCO were selected for fine-tuning.

 Processing data sets. Put the data in VOC format and generate voc_train and voc_test texts.

 Set experimental parameters. Parameter Settings are shown in Table 7.

 Training and testing.

Table 7
Some parameters of YOLOV3

Parameter Names	Parameter Values
BATCH_SIZE	2
IOU_LOSS_THRESH	0.5
ANCHOR_PER_SCALE	**3**
LEARN_RATE_INIT	1e-4
LEARN_RATE_END	1e-6
EPOCH	30×7200

Experimental comparison and analysis

Program design and improvement

In this paper, there are two plans for the outline design of water meter digital recognition model.

Plan A: Overall picture recognition. Image preprocessing stage, the image processing into a uniform format size, and the use of transfer learning. In the training stage, the water meter pictures were input into the deep learning model to learn ten categories. Categorize each goal. According to the predicted result, sort by the position size of *x*. Finally output the number on the water meter. The main process is shown in Figure 14

Although the plan A can identify the number in one step, the target will detect other numbers on the dial due to the existence of some external factors such as illumination, dust, interferences and some digital

interference of the dial itself. Therefore, plan A is improved on this basis and plan B is proposed, as shown in Figure 15.

Plan B: The image is recognized step by step. The first step of image preprocessing, adjust the picture size, and then carried out on the water meter image target detection for the first time, select YOLOv3 target detection depth learning network, and then will deal with good image input to the network, and learn a category, namely, with black rectangle category(screen), and used to predict the (*x, y, w, h*) to cutting of pictures to process the dataset further. The second step is to retrain the model and learn from 10 categories(0~9). The two-character criteria are used to determine the number based on the larger proportion of numbers in a box. According to the predicted results, the *x* position is sorted by the predicted size. Finally output the number on the water meter.

Comparison of model and plans results

mAP (mean average precision) is a commonly used evaluation index in multi-task target detection. Firstly, whether the prediction box or the prediction box is TP or FP is determined according to whether the IOU of the prediction box and the real box of each figure is greater than 0.5. Then, the prediction box is ranked from high to low according to the confidence of each prediction box, and the Precision and Recall under different confidence threshold values are obtained. PR curve is drawn and the area is obtained. AP is an indicator of the detection quality of a class, and mAP is an indicator of the detection quality of multiple classes, namely the average value of multiple categories of AP. Since the model needs to be used in the project, there are certain requirements for the accuracy. The higher the accuracy, the more efficient the user will be. In addition, for the prediction speed and model size of the model, the better the better on the basis of not affecting the accuracy requirements. To sum up, the evaluation indexes of the model are mAP, accuracy, recognition speed and model size.

First, Experimental results of identifying black rectangular boxes are shown in Table 8. The histogram is shown in Figure 16.

Table 8
Experimental results of identifying black rectangular boxes

Models	mAP(%)	Recognition speed(ms/piece)	Model Size(MB)
Faster RCNN	87.48	382.74	108.1
SSD	89.26	**60.15**	**95.84**
YOLOv3	**91.30**	66.87	231.1

It can be concluded that YOLOV3 has a significant effect on the recognition of the black rectangular box, can accurately identify the target area, and eliminate the interference of lighting, stains and black background. On this basis, this paper cuts the predicted target region and makes a data set. Some images are shown in Figure 17:

a b

Figure 17 Display of detection images of some water meters with black rectangular boxes

The image is cut according to the predicted (*x,y,w,h*), and the cut part of the image is shown in Figure 18:

a b

Figure 18 shows a picture with a black rectangular box cut out

At this point, this paper obtains a cut data set. According to the regulations of the company, this paper divides the numbers of water meters into ten categories. For single characters, each number represents a category; for double characters, the number with a large proportion represents the whole category of double characters.

Second, The result of identifying the water meter picture after cutting.

1. The experimental results of the Faster RCNN are shown in Figure 19.

(2) The experimental results of SSD are shown in Figure 20.

(3) The results of YOLOv3 experiment are shown in Figure 21.

- *Third, Comparison of experimental results of the plans.*

The experimental results of Plan A are shown in Table 9. The histogram is shown in Figure 22.

Table 9
Experimental results of Plan A

Models	mAP(%)	Accuracy	Recognition speed(ms/piece)	Model Size(MB)
Faster RCNN	77.07	82.96	397.6	111.5
SSD	75.67	78.63	**61.62**	**98.27**
YOLOv3	**78.10**	**87.30**	68.54	234.8

The experimental results of Plan B are shown in Table 10. The histogram is shown in Figure 23.

Table 10
Experimental results of Plan B

Models	mAP(%)	Accuracy	Recognition speed(ms/piece)	Model Size(MB)
Faster RCNN	75.46	81.53	395.1	109.4
SSD	78.61	80.42	**60.42**	**98.58**
YOLOv3	**82.02**	**90.61**	69.84	232.5

AP values for individual characters in Plan Bare compared as shown in Table 11. The histogram is shown in Figure 24.

Table 11
Single character AP values for Plam B

Single character AP	Faster RCNN	SSD	YOLOv3
zero	79.85	83.11	84.29
one	74.64	77.46	78.97
two	78.42	81.31	83.77
three	80.31	82.99	86.70
four	76.23	80.24	82.94
five	79.59	81.87	85.33
six	70.44	75.22	80.45
seven	69.54	72.87	81.33
eight	69.34	72.66	77.11
nine	76.28	78.34	79.36

In combination with the results in the table it can be seen that for the water meter image data in the recognition, the best is YOLOv3 network, can accurately identify the water meter black rectangular box in the digital images, and basic would not identify the black rectangle beyond the Numbers, if there is a character by two box predicted that if the two boxes of IOU value is greater than 0.8,By comparing the confidence, the greater one was selected as the predicted value, and the accuracy reached 90.61%.Compared with YOLOv3, the accuracy of SSD network is not very good at identifying the numbers in the black rectangle, and the proportion of SSD network identifying the numbers outside the black rectangle is higher than that of YOLOv3.The influence later is sorted according to the size of the predicted *x*. So it's not as effective as YOLOv3.The accuracy of Faster CNN network is also not as high as that of Yolov3, and compared with YOLOv3, Faster CNN cannot effectively identify every data in the black rectangular box. In addition, Faster CNN does not carry out good non-maximum suppression for the selection of candidate box, which leads to the recognition of multiple numbers into one number and the failure of candidate box to match the number region well, which does not meet the practical requirements.

Fourth, Testing.

Deploy the model to the server, and through the simple test on the page, the water meter number can be output effectively, as shown in Figure 25.

Conclusion

According to the actual demand of water meter reading recognition, this paper deeply studies and analyzes the existing deep learning model, and proposes two plans. Compared with Plan A, Plan B's mAP of YOLOv3 improved by nearly 4 percentage points, and its accuracy rate increased by about 3 percentage points, reaching 90.61%. It can accurately identify the numbers in the black rectangular box and output the sorted numbers, which can greatly improve labor efficiency and save labor costs.

Declarations

Acknowledgements

We are truly grateful to the people who helped with this research.

Author contributions statement

Y.L., W.W. conceived research; S.L., Y. L. performed research; T.Q. reviewed the manuscript. And all the participants completed the manuscript together.

Competing interests

The authors declare no competing interests.

References

1. Xu, D., Wang, L. & Li, F. Review of typical object detection algorithms for deep learning. Comput. Eng. Appl. 57, 10–25(2021).
2. Lu, F. et al. Overview on deep learning-based object detection. Comput.Syst.Appl. 30,1–13(2021).
3. Zhang, B., Jia, J. & Wang, W. Improvement of military target detection algorithm based on yolov3. Netw. Secur. Technol.Appl. 43–45 (2021).
4. Zhang, M. et al. Method for moving object detection of underwater fish using dynamic video sequence. J. Graph. 42,52–58 (2021).
5. Zhuang,F. Comparative study and application of three target detection algorithms in recognition of welding spot position of automobile door panel. Ph.D. thesis, South China University of Technology (2021).
6. Zhang, L. et al. Vehicle object detection method based on candidate region aggregation. Pattern Analysis Appl. 24,1635—-1647, DOI: 10.1007/s10044-021-01009-4 (2021).
7. Zhou, X., Wang, K. & Li, L. Review of object detection based on deep learning. Electron. Meas. Technol. 40, 89–93(2017).
8. Girshick, R., Donahue, J., Darrell, T. & Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation (2014). In: Proceedings of the IEEE conference on computer vision and

pattern recognition,pp 580-587.

9. He, k., Gkioxari, G., Dollár, P. & Girshick, R. Mask r-cnn (2017). In: Proceddings of IEEE International Conference on Computer Vision,pp 2980-2988.
10. Redmon, J., Divvala, S. & Girshick, R. Yolo9000: better, faster, stronger (2017). In: Proceddings of IEEE Conference on Computer Vision and Pattern Recognition, pp 6517-6525.
11. Redmon, J. & Farhadi, A. You only look once: unified, real-time object detection (2016). In: Proceddings of IEEE Conference on Computer Vision and Pattern Recognition, pp 779-788.
12. Redmon, J. & Farhadi, A. Yolov3: An incremental improvement (2018). Preprint at https://arxiv.org/abs/1804.02767.
13. Liu, W. et al. Ssd: single shot multibox detector (2016). In: Proceddings of European Conference on Computer Vision, pp 27-31.
14. Ren, S. , He, K. , Girshick, R. & Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis Mach. Intell. 39, 1137–1149 (2017).
15. Shang, X., Wen, Y., Xi, X. & Hu, F. Target tracking system based on the siamese guided anchor region proposal network. J. Image Graph. 26, 415–424 (2021).
16. Wang, J., Wang, Z., Lei, B. & Yan, X. Darknet framework based fine classification of vehicles in high altitude view. Mod. Electron. Tech. 44, 124–129 (2021).
17. Lawal, M.O. Tomato detection based on modified yolov3 framework. Sci Rep 11, 1447, DOI: 10.1038/s41598-021-81216-5(2021).
18. Cheng. Xand Qiu. G., Jiang. Y & Zhu. Z. An improved small object detection method based on yolo v3. Pattern Analysis Appl. 24, 1347—-1355, DOI: 10.1007/s10044-021-00989-7 (2021).
19. Gao,S. Study on Detection Method of Germinated Potato Based on SVM and Weighted Euclidean Distance. Ph.D. thesis, ilin Agricultural University (2020).
20. Yang, Y., Fan, D., Yin, H., Han, J. & Miao, S. Transmission line fault phase selection model based on deep-transfer learning and its transferability. Electr. Power Autom. Equip. 40, 165–172 (2020).

Figures

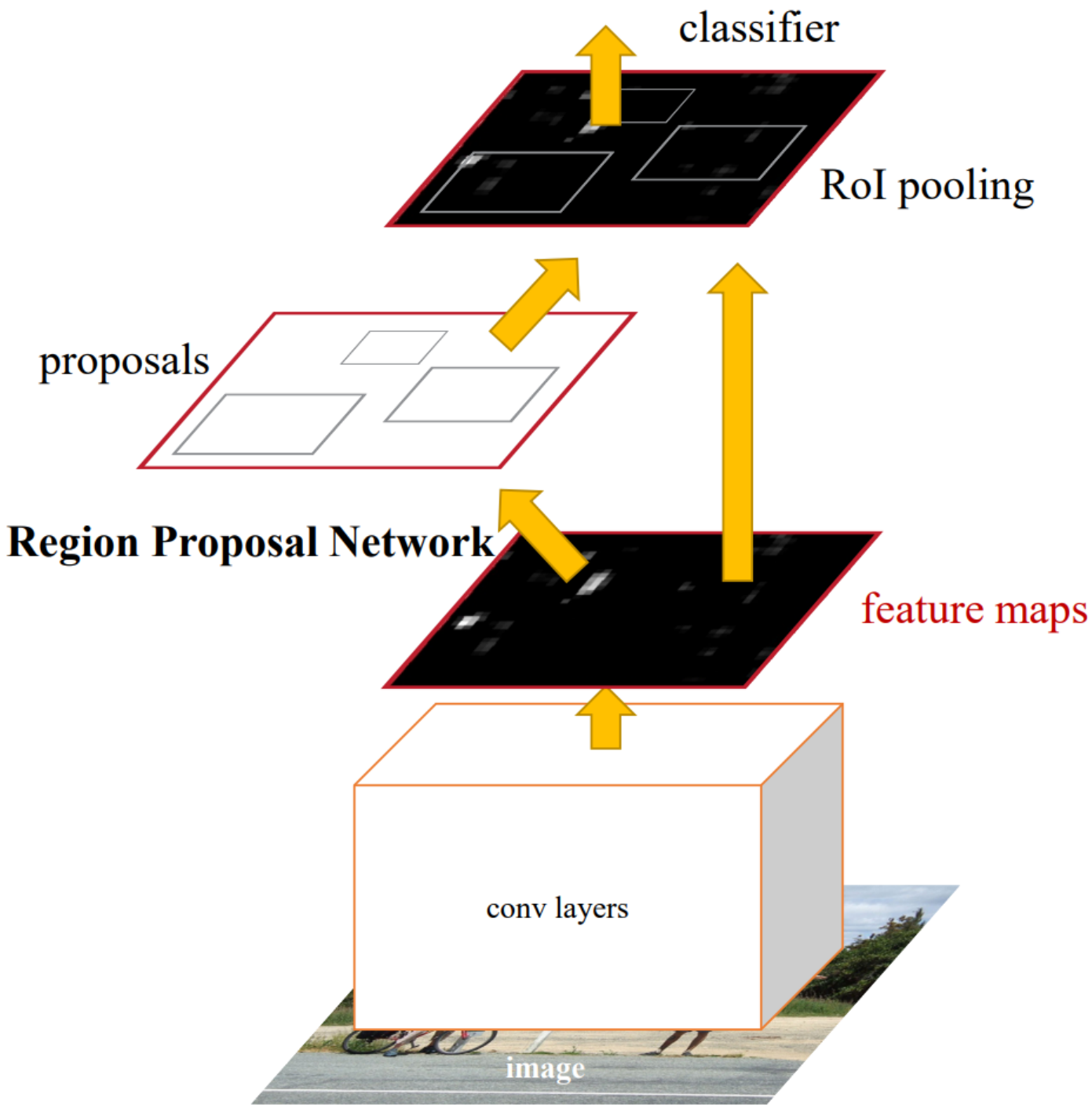

Figure 1

Faster RCNN network structure diagram[14]

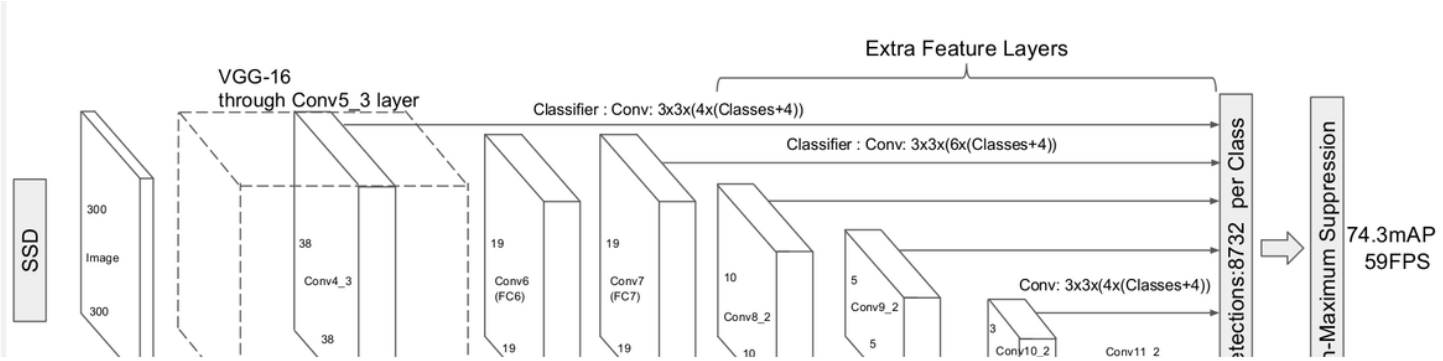

Figure 2

Model structure diagram of SSD[13]

Figure 3

Part of water meter image display

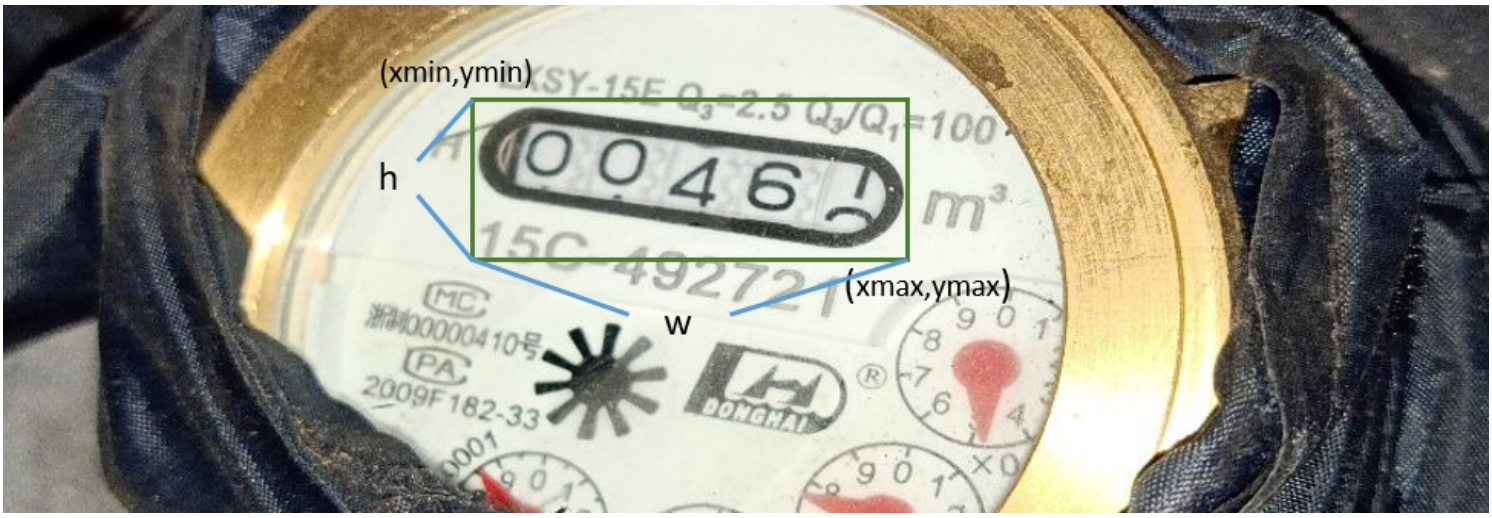

Figure 4

Picture label effect drawing

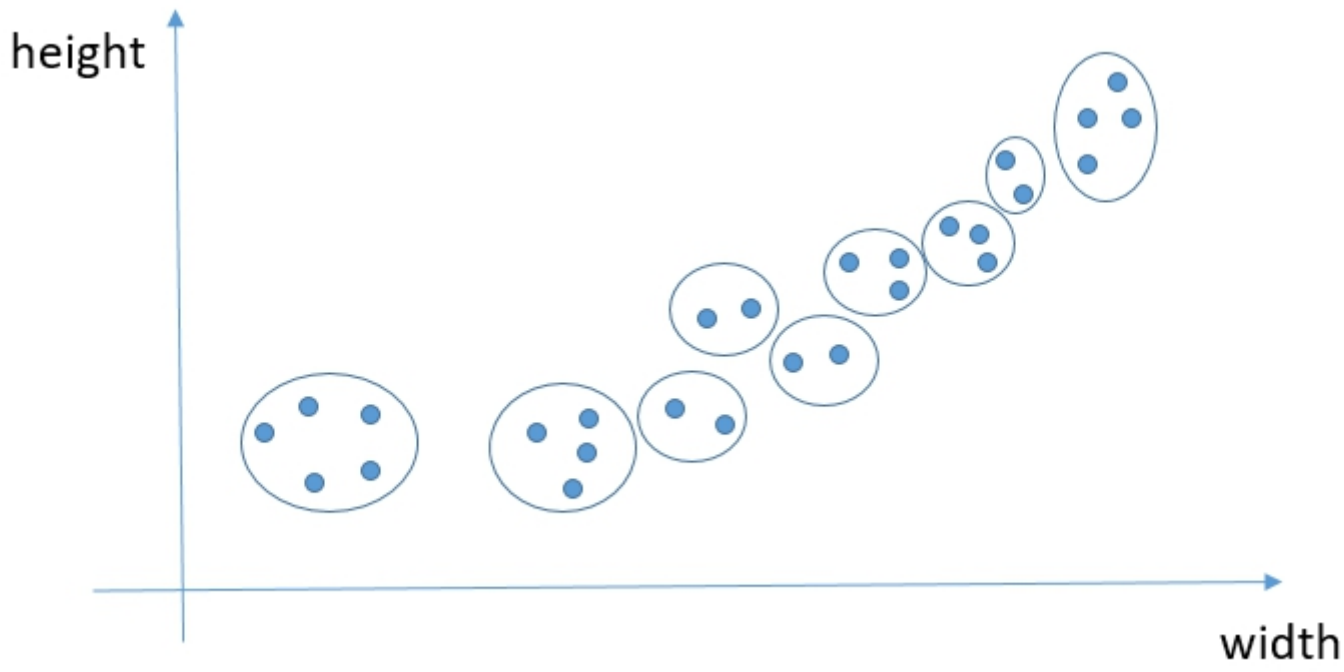

Figure 5

K-means classification effect diagram

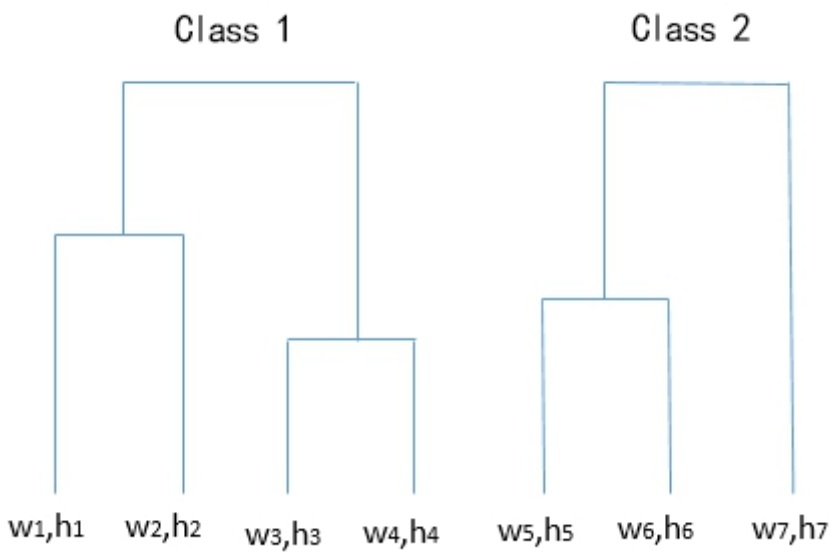

Figure 6

Partial renderings of hierarchical clustering

Figure 7

Distribution of K-means and hierarchical clustering classification

Figure 8

Effect of water meter image target box

Figure 9

Image of water meter with uniform size

Figure 10

Annotations of some water meter images

Figure 11

Water meter image after resizing

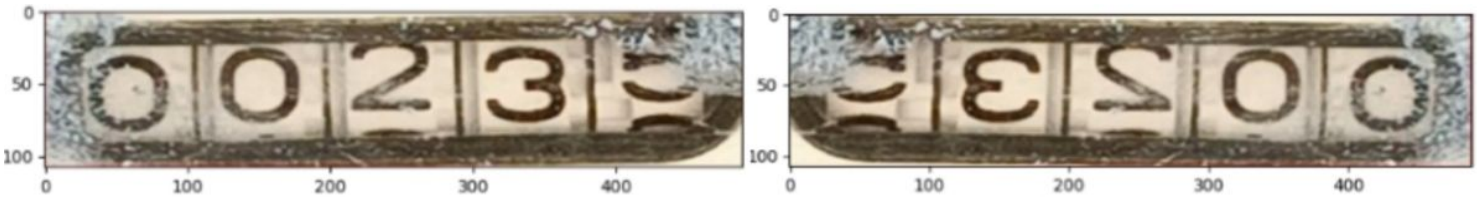

Figure 12

Water meter image after geometric transformation

Figure 13

Category distribution diagram of K-means and hierarchical clustering

Figure 14

Flow chart of Plan A

Figure 15

Flowchart of Plan B

Figure 16

Experimental results of identifying black rectangular boxes

Figure 17

Display of detection images of some water meters with black rectangular boxes

Figure 18

shows a picture with a black rectangular box cut out

Figure 19

A partial rendering of the Faster RCNN test

Figure 20

Partial renderings of SSD tests

Figure 21

A partial rendering of the YOLOv3 test

Figure 22

Histogram of experimental results of Plan A

Figure 23

histogram of experimental results of Plan B

Figure 24

Histogram of AP values for a single character in Plan B

Figure 25

Test result diagram