


Article

Automatic Meter Reading from UAV Inspection Photos in the Substation by Combining YOLOv5s and DeeplabV3+

Guanghong Deng ¹, Tongbin Huang ¹, Baihao Lin ¹, Hongkai Liu ¹, Rui Yang ¹ and Wenlong Jing ^{2,*} ¹ Guangzhou iMapCloud Intelligent Technology Co., Ltd., Guangzhou 510095, China² Guangdong Province Engineering Laboratory for Geographic Spatiotemporal Big Data, Key Laboratory of Guangdong for Utilization of Remote Sensing and Geographical Information System, Guangdong Open Laboratory of Geospatial Information Technology and Application, Guangzhou Institute of Geography, Guangdong Academy of Sciences, Guangzhou 510070, China

* Correspondence: jingwl@lreis.ac.cn; Tel.: +86-18612649525

Abstract: The combination of unmanned aerial vehicles (UAVs) and artificial intelligence is significant and is a key topic in recent substation inspection applications; and meter reading is one of the challenging tasks. This paper proposes a method based on the combination of YOLOv5s object detection and Deeplabv3+ image segmentation to obtain meter readings through the post-processing of segmented images. Firstly, YOLOv5s was introduced to detect the meter dial area and the meter was classified. Following this, the detected and classified images were passed to the image segmentation algorithm. The backbone network of the Deeplabv3+ algorithm was improved by using the MobileNetv2 network, and the model size was reduced on the premise that the effective extraction of tick marks and pointers was ensured. To account for the inaccurate reading of the meter, the divided pointer and scale area were corroded first, and then the concentric circle sampling method was used to flatten the circular dial area into a rectangular area. Several analog meter readings were calculated by flattening the area scale distance. The experimental results show that the mean average precision of 50 (mAP50) of the YOLOv5s model with this method in this data set reached 99.58%, that the single detection speed reached 22.2 ms, and that the mean intersection over union (mIoU) of the image segmentation model reached 78.92%, 76.15%, 79.12%, 81.17%, and 75.73%, respectively. The single segmentation speed reached 35.1 ms. At the same time, the effects of various commonly used detection and segmentation algorithms on the recognition of meter readings were compared. The results show that the method in this paper significantly improved the accuracy and practicability of substation meter reading detection in complex situations.

Keywords: object detection; image segmentation; YOLOv5s; Deeplabv3+; meter reading

Citation: Deng, G.; Huang, T.; Lin, B.; Liu, H.; Yang, R.; Jing, W. Automatic Meter Reading from UAV Inspection Photos in the Substation by Combining YOLOv5s and DeeplabV3+. *Sensors* **2022**, *22*, 7090. <https://doi.org/10.3390/s22187090>

Academic Editors: Zhaoyang Wang, Hieu Nguyen and Minh P. Vo

Received: 15 August 2022

Accepted: 13 September 2022

Published: 19 September 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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

Meter reading is an extremely important task that is widely used in real life [1]. Many meter readings [2–4] require manual periodic inspection and recording of the meter readings to reflect whether the device is operating safely. Due to the various shapes, scales, pointers, and characters of the meter, and the existence of different equipment locations, such as high-voltage equipment in substations, it is difficult for regular manual inspections to be carried out.

At present, traditional meters with scales and pointers are still commonly used in the substation environment. In the actual application environment, the traditional manual counting not only involves many human factors, but also is easily disturbed by the environment and has the danger of an electric shock. With the gradual promotion of unattended substations, inspection robots or UAVs equipped with automatic meter identification technology have been widely used. Therefore, the realization of intelligent reading of meters by the computer vision method has become a research hotspot.

In recent years, the traditional meter reading methods have been mainly based on Hough transform [5–8] and image registration [9,10]. Firstly, the method based on Hough transform is to obtain the position of the pointer and the dial through Hough line detection and arc detection, and also to obtain the deflection angle of the pointer and calculate the reading of the meter. This method is susceptible to noise interference, resulting in inaccurate readings. Secondly, some researchers also use the method based on image template matching to identify the instrument. This method uses feature matching algorithms such as SIFT [11] and SURF [12] to register the image to be recognized as a standard image to read meter readings, and it is not conducive to dealing with multiple meter readings in complex backgrounds.

With the continuous development of deep learning and convolutional neural networks, many common object detection algorithms and semantic segmentation algorithms [13] are used in meter reading detection, such as YOLO [14,15], SSD [16], Faster RCNN [17], DeepLab [18], Unet [19], and Transformer [20], etc. In recent years, many researchers have used deep learning methods to conduct research in this field. Xing et al. [21] used a convolutional neural network to detect the area where the meter exists and post-process the meter image through a Hough transform. However, the robustness of the algorithm is not strong. Wan et al. [22] proposed a method based on a combination of object detection and image segmentation algorithms and a perspective transformation method based on image segmentation information to calibrate the image and finally obtain the meter reading. However, this method requires a two-stage detection of the dial and hands, which increases the calculation time. Tao et al. [23] used the SSD algorithm combined with the post-processing method to achieve meter readings, but this method has a large error, and the model size is also large and takes up a lot of memory.

Most of the above methods are not suitable for carrying on mobile terminals or UAVs, and there is the problem of large errors. In order to solve the above problems, this paper proposes a meter detection and recognition method based on UAV aerial images. Specifically, it is to deploy UAV slots in substations, plan routes, and collect meter images in different weather and time periods in substations. The meter area in the image is detected by using the YOLOv5 object detection algorithm [24], and the image of the meter area is intercepted. Then the intercepted meter area image is sent to the Deeplabv3+ [25] algorithm segmentation model for segmentation, and finally the meter reading is obtained after post-processing. The main contributions of this paper are as follows:

1. By combining UAV and deep learning vision technology, the problems of the low efficiency and the high cost of traditional manual inspection or robot inspection are solved;
2. The object detection algorithm YOLOv5s is introduced to improve the accuracy of detection of meter dial area and classification;
3. Deeplabv3+ is used for image segmentation and this method improves the detection accuracy of the pointer and the scale line;
4. Based on the image segmentation results, the concentric circle sampling method is proposed to flatten the dial to realize the reading of the dial image.

This study is outlined as follows: Section 2 shows the YOLOv5 algorithm structure, the Deeplabv3+ algorithm structure, and the post-processing method of the meter readings; Section 3 introduces various comparative experiments and experimental results from the same time and the experimental results are analyzed; Section 4 concludes the study; and Section 5 puts forward an outlook for the future in view of the shortcomings of the research.

2. Methods

2.1. Meter Reading Recognition Based on Object Detection and Image Segmentation

According to the characteristics of the UAV aerial photography substation meter image and the shortcomings of the traditional meter reading method, this paper adopted the object detection algorithm and semantic segmentation technology based on deep learning, which can accurately obtain the meter image and the meter corresponding to the

image. The scale and pointer area realized the reading of the substation equipment. Object detection technology was used to detect the meter area in the image, which generally refers to the minimum outer-enclosing rectangular area of the closely surrounded meter target; the image segmentation technology is used to further segment the meter pointer and scale area target pixels in the meter image.

The idea of this paper was firstly to use the YOLOv5s object detection [26,27] technology to detect the area where the metered target is located in the UAV aerial image and eliminate the interference of the non-target area in the image; secondly, to accurately segment the intercepted meter image into the scale and pointer positions in the meter image by Deeplabv3+ image segmentation technology; and finally, to obtain the meter reading by post-processing. The processing flow of this paper is shown in Figure 1.

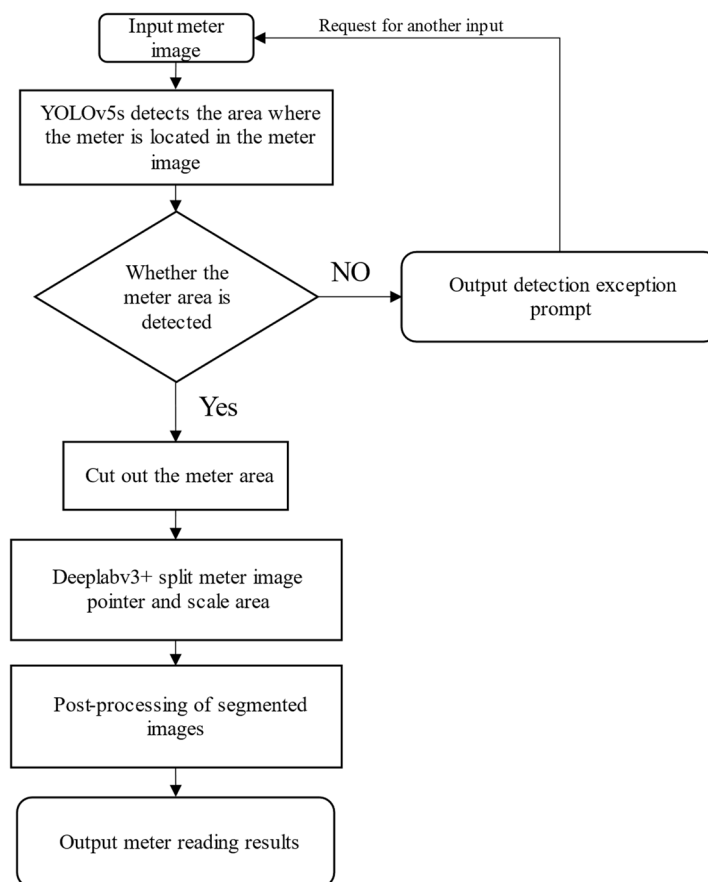


Figure 1. The process of identifying meter readings based on YOLOv5s and Deeplabv3+.

2.2. The YOLO Model

YOLOv5 [28] is a single-stage object detection algorithm. According to the depth and width of the network, YOLOv5 has four versions: YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x. The depth of the network directly affects the detection accuracy and speed of the detector. The detection object in this paper was the aerial meter data and the detection object was small. On the premise of ensuring the detection accuracy, the detection model was installed on the edge device, so the YOLOv5s version would be used.

The YOLOv5s network consists of three parts and its network structure is shown in Figure 2. The backbone used mosaic data enhancement for splicing images through random scaling, random cropping, and random arrangement. The input image was imported into the Focus module for slicing operation and the sample slices with a scale of $640 \times 640 \times 3$ were spliced into $320 \times 320 \times 12$. At the same time, a Cross Stage Partial Networks (CSP) [29] structure and Spatial Pyramid Pooling (SPP) [30] were introduced to realize convolution and pooling down-sampling in order to extract the features. The next part

was Neck, which was mainly the Feature Pyramid Networks (FPN) [31] and the Path Aggregation Network (PAN) [32]. The FPN layer transferred and fused the high-level strong semantic features through up-sampling from top to bottom. The PAN conveyed strong localization features from the bottom up and aggregated features from different backbone layers to different detection layers. The last part was the detection part, which took and output the feature maps of three scales, which are 17×17 , 20×20 , and 23×23 , respectively. In the post-processing process, the model generated multiple anchor boxes based on the object features and used non-maximum suppression (NMS) [33]. If the confidence of the object being predicted as the object category was greater than the set threshold, it was retained, thus completing the object detection process.

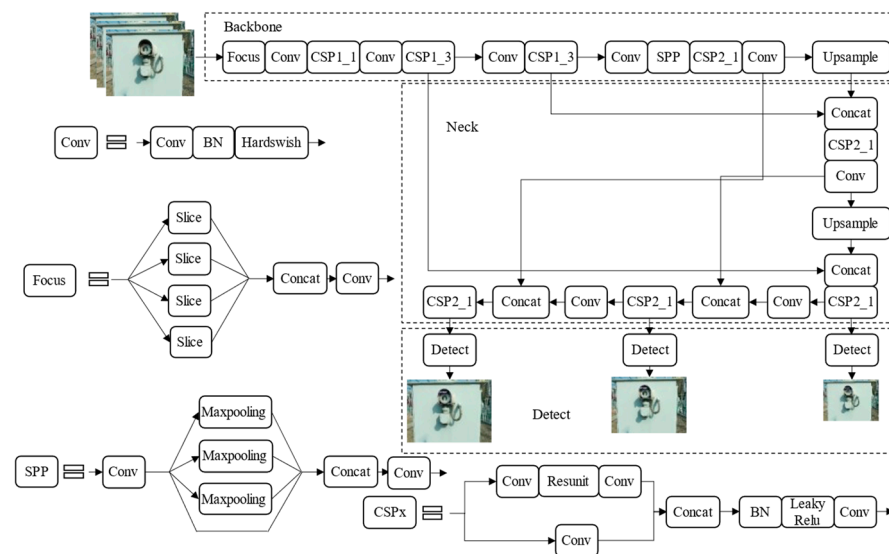


Figure 2. The YOLOv5s network structure diagram.

YOLOv5s was used to detect the area where the meter was located within the image. In order to improve the accuracy of the final meter reading, this paper firstly detected the dial area of the image through the YOLOv5s network model. Using the labeling software, Labellmg, to make meter datasets, five different kinds of labels (bj, bjA, bjB, bjH, and bjL) in the meter images were defined. The type of meter represented by each label is shown in Figure 3. The algorithm itself automatically imported the training set into the YOLOv5s network model for training and generated the corresponding detection model weights.

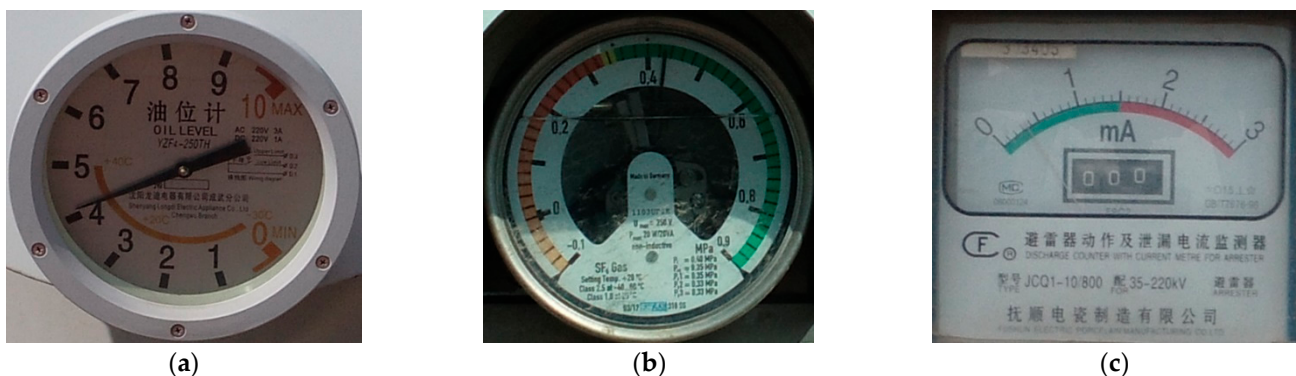


Figure 3. Cont.



Figure 3. Five kinds of meters. (a) bj, which refers to oil level, (b) bjA, which refers to Sulfur Hexafluoride Density Relay, (c) bjB, which refers to Discharge Counter With Current Meter For Arrester, (d) bjH, which refers to Discharge Counter and (e) bjL, which also is Sulfur Hexafluoride Density Relay.

2.3. Deeplabv3+ Split Tick Marks and Pointers

Deeplabv3+ is a well-polished state-of-the-art segmentation model, which has been widely used in many areas, such as the remote sensing area and medical image processing. This paper studied instrument segmentation detection. The proportion of the pointer and the scale was small, and the segmentation requirements were relatively fine. Deeplabv3+ has a better segmentation effect on fine objects, therefore, this paper chose Deeplabv3+.

Deeplabv3+ adopts a spatial pyramid pooling model and an encoding-decoding structure for semantic segmentation, it outputs feature map information in the upper backbone network, and it detects incoming features through a multi-rate and multi-effective field of view, which filters or pools operations to encode multi-scale features. The context information is enriched by encoding the semantic information and the decoder part gradually recovers the sharp target boundary information. In the meter detection images, because the pointer and the scale occupy a small proportion of the area where it is located, the segmentation of the image will be more difficult. The module structure of Deeplabv3+ is shown in Figure 4.

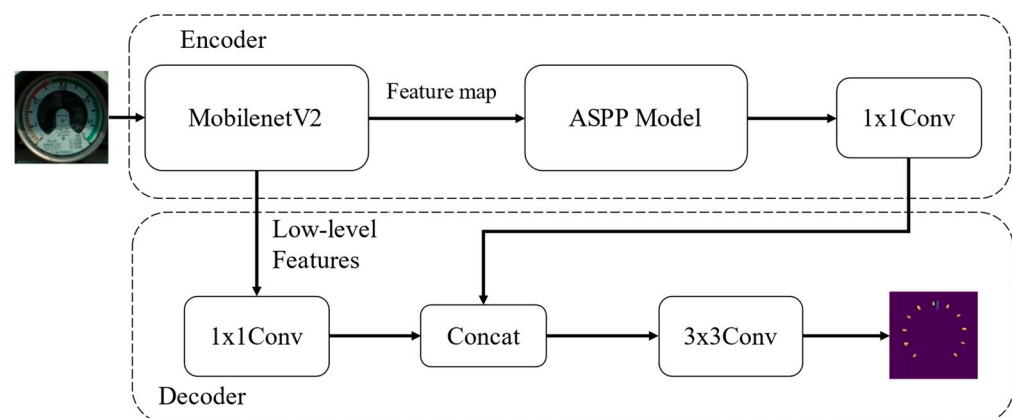


Figure 4. The Modules of Deeplabv3+.

Compared with the Xception series used in the Deeplabv3+ paper as the backbone feature extraction network, a different backbone network, MobileNetv2 [34], was used in this paper. It is more suitable for deployment on edge devices than Xception and has more ideal and more efficient parameters and speed. The extracted feature maps went through the upgraded ASPP module. The feature map was first reduced in dimension through a 1×1 convolution kernel, then through three depthwise separable convolutions, and finally it was output through Adaptive Pooling. The compressed feature layer was passed

into the decoder part through the backbone, and after being resized, it was concat with encode_data, and finally the output result was obtained through two convolution kernels and the Upsample.

In this paper, Deeplabv3+ was used to segment the tick marks and the pointer positions of the dial area. The training set of the segmentation network adopted the image of the dial area after being detected by the YOLOv5s detection network. Therefore, the cropped RGB image of the dial area was transmitted to the Deeplabv3+ network and it output the segmentation map as the same size as the input images. Figure 5 shows an example of the tick training set; the upper image is the original image, and the lower image is the label image.

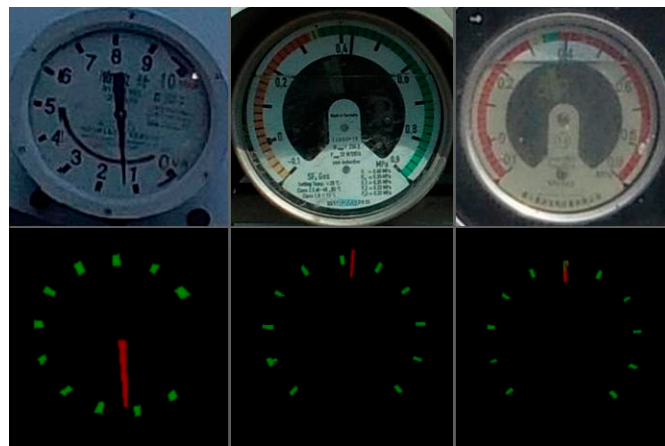


Figure 5. An example of segment datasets.

2.4. Post-Processing Methods

2.4.1. Erosion

The Deeplabv3+ segmentation removed most of the background and useless information, leaving only the scale and the pointer outline in the image; however, there were still many noises. In order to further eliminate interference, this paper performed erosion morphological processing on the segmented meter pointer and scale outline to remove the discrete point blocks. Assuming that the contour map point set is A , the convolution kernel is B , and B moves in order in A , the erosion image can be obtained. The segment image was obtained from the Formula (1):

$$A - B = \{x, y | (B)_{xy} \subseteq A\} \quad (1)$$

To remove the interference point blocks in the segmented image, the convolution kernel B , used in this paper, was designed as a 4×4 structure. The erosion operation eliminated the boundary points of the object, shrunk the boundary inward, and removed the objects that were smaller than the structural elements. The pointer and scale occupied fewer pixels in the segmentation map and were easily affected by noise. The effects of noise such as burrs and small bumps were removed by erosion. At the same time, two objects that are only connected by small blocks were disconnected to improve the reading accuracy of the meter.

2.4.2. The Flattening Method and Meter Readings

After semantic segmentation and erosion processing, the circular dial area of the pointer image was flattened into a rectangular area by the concentric circle sampling method. The length and geometric center of the two sides of the divided rectangular image were used as the diameter and the center of the initial concentric circle, respectively, and the initial rotation angle and the width and height of the flattened rectangular area were specified. The initial rotation angle was used to generate the initial sampling point of each

concentric circle. The width corresponded to the number of times of sampling for each concentric circle, and the height corresponded to the number of sampled concentric circles. Starting from the initial sampling point of the concentric circles, the pixel values were uniformly sampled on the circumference of the concentric circles in a clockwise direction. Taking the center of the concentric circles as the center, and shortening the radius by one pixel unit, a new concentric circle was generated. The steps for sampling were repeated several times, and finally the flattened rectangular area corresponding to the circular dial area was obtained.

The flattened rectangular area corresponded to the scale of the dial, and the midpoint coordinate of the line-segment was taken as the scale feature point for the scale, thereby forming a scale-center coordinate set that could represent the position of the scale. At the same time, the flattened image was scanned line by line from top to bottom, and the average value of the pointer pixel position was used as the coordinates of the pointer tip to indicate the pointer position.

Meter readings were accurately calculated by flattening the image, and the calculation formula is shown in (2):

$$R = \frac{\alpha}{\beta} \mu \quad (2)$$

where, R represents the meter reading, α represents the distance from the initial scale point of the flattened image to the pointer, β represents the distance from the initial scale of the flattened image to the end scale, and μ represents the total range of the meter.

2.5. Evaluation Indicators

Due to the complex environment of the substation, the aerial image of the UAV has a large receptive field, which will cause missed detection and false detection of the meter. Therefore, this paper used Precision and Recall to describe the meter detection model performance. The formulas of Precision and Recall are shown in (3) and (4):

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (4)$$

In the above formula, TP and FP represent the true and false positives, respectively, and FN represents the false negatives. In order to further evaluate the detection performance of the model, it was proposed to use the AP of a single category to represent the sum of the AP (average precision) values of each category, and to obtain the mAP (mean average precision) according to the AP value. The formulas for AP and mAP are shown in (5) and (6):

$$AP = \int_0^1 P(R) dR \quad (5)$$

$$mAP = \frac{\sum_{q=1}^Q AP(q)}{Q} \quad (6)$$

In Formula (6), Q is the number of categories.

The accuracy evaluation index of the image semantic segmentation model is expressed by mIoU (mean intersection over union), and the calculation formula is shown in (7):

$$mIoU = \frac{1}{k} \sum_{i=1}^k \frac{p_{ii}}{\sum_{j=1}^k p_{ij} + \sum_{j=1}^k p_{ji} - p_{ii}} \quad (7)$$

Among them, K represents the number of label categories in the data set; p_{ii} represents the number of the category label, i , in the data set, where the actual prediction is i ; p_{ij} represents the category label, j , where the actual predicted category is the number of i ; and p_{ji} represents the category label, i , where the actual predicted category is the number of j .

3. Experiment and Results

3.1. Experimental Conditions

3.1.1. Data Acquisition and Transmission

Datasets were acquired by using the company's self-produced drone nest, the developed platform scheduling software, and the DJI (Shenzhen DJI Sciences and Technologies Ltd., Shenzhen, China) Genie 4rtk (Real-Time Kinematic) UAV for data collection. The UAV itself integrates an HD video transmission system, a 360° rotating gimbal, and a 4K camera. The camera carried by the UAVs captured all the photos on an SD card. Transmission was carried out by a 4G/5G signal or, when it came back to the drone nest connected by LAN. All collected data was sent to the control center.

This paper collected the meter images in the Harbin substation and the filming hours were from 9:00 a.m to 6:00 p.m. The shooting environment included different weather patterns, lighting, and time periods, and the flight collection was specified according to the planned route. The flying height of the UAV was the same as the shooting point of the collecting meter, and the distance between the gimbal and the meter to be collected was about 1 to 1.5 m. Both the training set and the validation set are independent of each other. We collected a total of 1632 images in five different categories. There were 979 images in the training set and 653 images in the test set and the number of all the meter types is shown in Table 1.

Table 1. Dataset Statistics.

	bj	bjA	bjB	bjH	bjL
Training set	75	329	301	126	146
Test set	58	213	179	106	97

3.1.2. Experiment Platform

Server-side: Ubuntu 18.04, Intel® Silver 4210 CPU@2.20 GHz, NVIDIA GeForce RTX A100(80 GB) GPU. The model framework is Pytorch 1.7.0, and the related software is CUDA 11.1, CUDNN 8.0.5 and Python 3.8.

Ubuntu18.04, Intel® Xeon® Gold 5120 CPU@2.20 GHz, NVIDIA GeForce RTX 2080(8 G)*2 GPU. The model framework is PaddleX 1.3.3, and the related software is CUDA 10.2, CUDNN 8.1.0 and Python3.8.

3.2. Experimental Results

3.2.1. YOLOv5s Detection Results

After the training was completed, the YOLOv5s detection model needed to be tested before the semantic segmentation task, and the test results are shown in Figure 6. This shows that from different distances, different angles, and different weather conditions, the detection model could identify the position of the meter target in the image, and at the same time could correctly identify the meter type. This proves that the model used in this paper had a certain generalization ability and the training effect was ideal.



Figure 6. Dial detection results using YOLOv5s.

3.2.2. Deeplabv3+ Image Segmentation Results

After the Deeplabv3+ network was trained, the network model was tested. Figure 7 shows the test results of the tick marks and the pointers of the model after network segmentation. The upper layer is the original scene image and the lower layer is the corresponding segmented image.

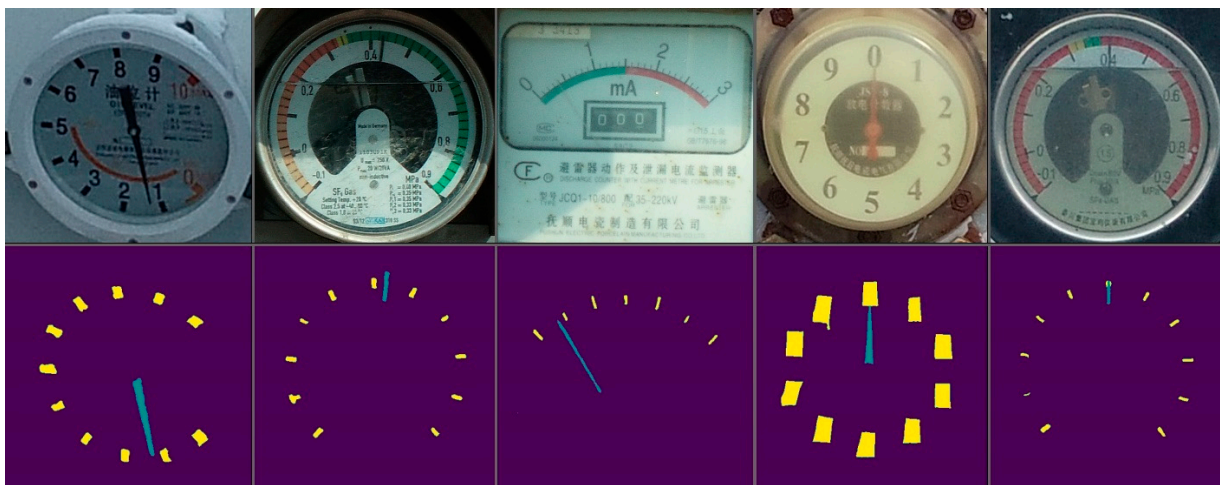


Figure 7. Image segmentation results of Deeplabv3+.

The test results show that the input image segmentation result was basically correct, and the position of the tick mark and the pointer could be accurately separated. The final image also needed to be corroded and flattened in order to correctly identify the dial reading.

3.2.3. Flattening Results

In the rectangular area, the scales were evenly arranged from left to right, the lower end of the pointer was close to the center of the dial, and the upper end was close to the scale. Figure 8 shows, the result of flattening the image. The first scale coordinate of the scale center coordinate corresponded to the start scale position, and the last scale coordinate corresponded to the end scale position. The first distance between the pointer tip coordinate and the start scale position was calculated and the second distance between the end scale position and the start scale position was calculated. The ratio of the first distance to the second distance was multiplied by the total range of the type of meter to obtain the readings of several pointer meters in the substation.

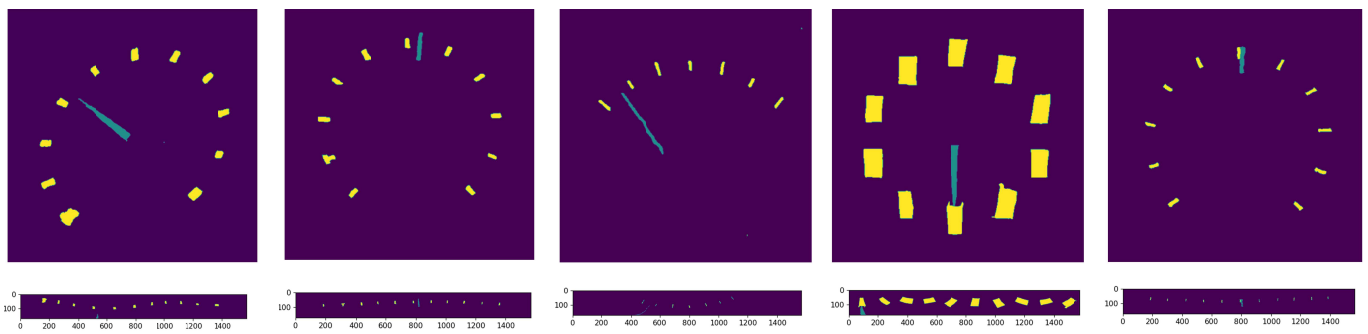


Figure 8. The result of flattening the image.

3.3. Comparative Requirements

In order to choose the algorithm version that was more suitable for the research in this paper, a comparative requirement between the parameters, FLOPS and running speed of each version of YOLOv5 was carried out. The experimental results are shown in Table 2.

Table 2. A comparison of yolov5 detection model results.

Model	Speed/ms	FPS	Params	FLOPS
YOLOv5s	22.2	45.0	7.5M	13.2B
YOLOv5m	27.0	37.0	21.8M	39.4B
YOLOv5l	29.2	34.2	47.8M	88.1B
YOLOv5x	30.8	32.5	89.0M	166.4B

It can be seen from Table 2 that the YOLOv5s model had the least number of parameters and the fastest speed.

In order to verify the feasibility of using the detection and segmentation algorithm in this paper, this study compared and tested a variety of commonly used algorithms on the premise of the same dataset. The results of this comparison are shown in Tables 3 and 4.

Table 3. The comparison of meter detection model results.

	bj/%	bjA/%	bjB/%	bjH/%	bjL/%	mAP50/%	Speed/ms	Model Size/MB
YOLOv3	99.536	99.623	99.592	99.575	99.567	99.579	27.4	123.4
YOLOv4	99.538	99.610	99.571	99.561	99.566	99.569	34.0	256.3
YOLOv5X	99.540	99.626	99.593	99.576	99.571	99.581	30.8	177.5
YOLOv5L	99.539	99.613	99.584	99.575	99.570	99.576	29.2	90.8
YOLOv5m	99.542	99.630	99.600	99.579	99.573	99.585	27.0	41.3
YOLOv5s	99.542	99.628	99.599	99.579	99.571	99.584	22.2	14.1

Table 4. A comparison of segmentation results.

	Backbone	bj _{mIoU} /%	bjA _{mIoU} /%	bjB _{mIoU} /%	bjH _{mIoU} /%	bjL _{mIoU} /%	Speed/ms	Model Size/MB
Deeplabv1	VGG16	61.69	52.35	44.54	67.01	37.33	15.8	82.0
Deeplabv2	Resnet101	57.74	47.12	33.55	77.09	45.06	56.0	176.9
Deeplabv3+	Xception65	85.62	76.95	82.14	82.93	80.03	66.8	165.1
Deeplabv3+	MobileNetV2	78.92	76.15	79.12	81.17	75.73	35.1	11.1

As can be seen in Table 3, this paper proposed the use of the YOLOv5s model with a faster detection speed for the detection of the meter dial of the aerial image. The Deeplabv3+ model with MobileNetV2 as the backbone network was used to segment the pointer and the scale of the meter image and the meter reading was realized through the meter post-processing technology. In this paper, the YOLOv5s model was used to detect a picture on the NVIDIA Geforce A100 GPU with an inference speed of 22 ms, which is significantly better than the YOLOv5m, YOLOv5L, YOLOv5X, YOLOv4 [35], and the YOLOv3 [36] models. The size of the model was only 14.1 MB; and mAP50 reached 99.584% in this dataset. The better detection accuracy and faster detection speed is able to meet the daily meter image inspection requirements.

In summary, compared with other commonly used detection algorithms, YOLOv5s has the smallest model and the fastest inference speed under the premise of ensuring model accuracy. It is more suitable for real-time detection on edge devices deployed on UAVs.

As can be seen from Table 4, for the five types of meters, the mIoU of the method used in this paper reached 78.92%, 76.15%, 79.12%, 81.17%, and 75.73%, respectively. It is obviously better than the Deeplabv1 model and the Deeplabv2 model, but not as good as than the original Deeplabv3+ model. However, the Deeplabv3+ model with MobileNetV2 as the backbone achieved a single image segmentation speed of 35.1 ms on NVIDIA Geforce 2080 GPU, which was significantly faster than the other three image segmentation models. The improved Deeplabv3+ model was only 11.1 MB and the parameters of the static model were only 2.8 MB. Furthermore, the model ran twice as fast, which greatly improved the model segmentation speed.

Table 4 shows that the Xception65 is the best backbone model, and this study did not need to achieve real-time detection. The method in this paper can be applied to substation meter reading whether using the Xception65 or the MobileNetv2 backbone network. Furthermore, the MobileNetv2 backbone network model detection speed was faster and we hope that the detection speed will be as fast as possible while still being of practical use.

Compared with the method of combining Faster R-CNN and U-Net in the literature [22], the method in this paper has the following advantages:

1. Compared with the Faster R-CNN algorithm, the YOLOv5 algorithm was used in this paper, and the detection speed was significantly faster;
2. The Deeplabv3+ image segmentation algorithm is mainly used in industrial applications, but the U-Net image segmentation method is mainly used for medical image segmentation, so it is better to use the Deeplabv3+ method for meter readings in industrial applications;
3. The post-processing methods such as concentric circle sampling in this paper were more robust than the industrial applications in paper [22].

3.4. Meter Reading Interface Display

Figure 9 shows the meter reading results of this method, which are 1.2544, 0.4285, 0.3073, 0.0000, and 0.3977, respectively. It can be seen from the above that the method in this paper could accurately read the value of the meter image, providing an effective and accurate reading method for the UAVs aerial photography of the meter image, which, in this case, used a pointer to indicate the value.

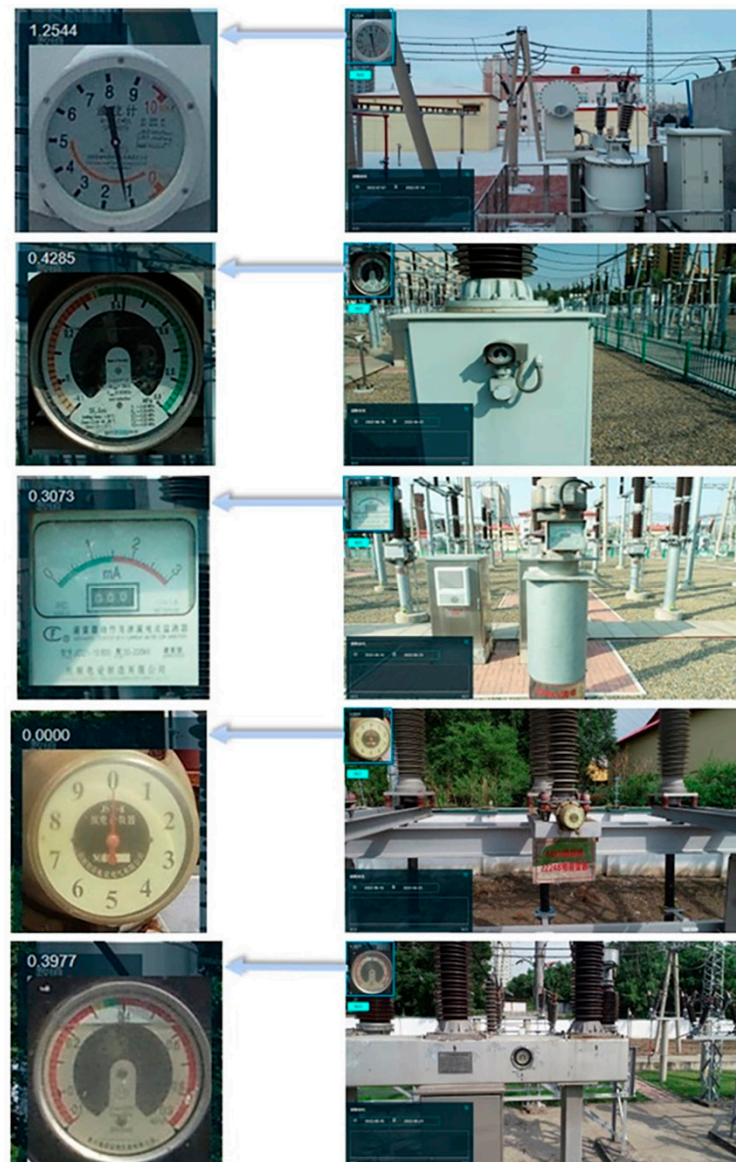


Figure 9. The results of the meter reading.

3.5. Comparing Readings

Figure 10 shows the images of the five meters and Table 5 shows the manually measured values compared to the recognized values. The error shows the absolute values of the manually measured values minus the recognized values, which are 0.0212, 0.0002, 0.0184, 0.0330, and 0.0017, respectively. The effectiveness of this method in automatic identification and reading of substation meter images in UAV aerial photography is, therefore, proven.



Figure 10. Meter images.

Table 5. Automated recognized results compared to manually read values.

	Manually Measured Values	Recognized Values	Error
bj	3.8500	3.8288	0.0212
bjA	0.4385	0.4383	0.0002
bjB	0.3900	0.3716	0.0184
bjH	0.0000	0.0330	0.0330
bjL	0.4055	0.4072	0.0017

The method proposed in this paper still has many aspects that need to be improved. We mainly realize that the reading detection of various pointer meters, and the provision of technical support for the inspection of substation meters, is needed. However, there are still many different types of meters that have not been studied, and when the area of the meter is not complete due to the external conditions, it will cause errors or lead to large reading errors.

4. Conclusions

This paper has designed a method that combined YOLOv5s and Deeplabv3+ and has implemented a substation meter detection and reading method through a series of post-processing methods. The test results have proven that the method proposed in this paper could accurately read various meter types at different angles and under different conditions. The main contributions of this paper are as follows

1. The use of UAVs to fly through designated routes at different times and different weather conditions and the collection of 1632 images, including five different types of meters for object detection model training;
2. The improvement of: the backbone network of the Deeplabv3+ semantic segmentation network; and the inference speed of the segmentation algorithm for a single image, which was twice the speed of the original model and a reduction in the size of the model weight;
3. The use of the erosion and concentric circle sampling method to flatten images to realize meter panel reading. The result has been to achieve an accurate reading of the meter readings while quickly detecting the meter area. In this paper, the inspection of substation instruments was combined with deep learning visual algorithms and mobile flying equipment. It is hoped that the work in this paper can provide some help for intelligent substation inspection.

5. Future Work

The main future work would be to continue to improve detection accuracy and speed, especially for different kinds of meters and more complex background conditions. At present, the intelligent inspection technology of meter readings has become an important development direction of the intelligent inspection of substations. In future research, we will increase the detection and segmentation of other components in the substation environment, and at the same time combine other algorithms such as object tracking and key point detection to achieve state estimation and prediction, which will further improve the intelligence level of substation inspection from all aspects.

Author Contributions: Conceptualization, G.D. and T.H.; validation, G.D.; formal analysis, B.L.; investigation, H.L.; resources, H.L.; data curation, R.Y. and B.L.; writing—original draft preparation, G.D.; writing—review and editing, B.L. and W.J.; visualization, T.H.; supervision, H.L. and W.J.; project administration, W.J.; funding acquisition, W.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the GDAS' Project of Science and Technology Development (2022GDASZH-2022010202) and the Science and Technology Program of Guangdong (2021B1212100006).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We thank the Geographical Science Data Center of the Greater Bay Area (<http://www.gbageodata.cn/>) (accessed on 1 May 2022) for providing the relevant data in this study.

Conflicts of Interest: Guanghong Deng, Baihao Lin, Hongkai Liu, and Rui Yang are employees of Guangzhou iMapCloud Intelligent Technology Co., Ltd. Tongbin Huang received a part-time salary for this work. Wenlong Jing has no conflict of interest to declare.

References

1. Li, C.; Su, Y.; Yuan, R.; Chu, D.; Zhu, J. Light-weight spliced convolution network-based automatic water meter reading in smart city. *IEEE Access* **2019**, *7*, 174359–174367. [[CrossRef](#)]
2. Wu, X.; Shi, X.; Jiang, Y.C.; Gong, J. A high-precision automatic pointer meter reading system in low-light environment. *Sensors* **2021**, *21*, 4891. [[CrossRef](#)] [[PubMed](#)]
3. Hong, Q.Q.; Ding, Y.W.; Lin, J.P.; Wang, M.H.; Wei, Q.Y.; Wang, X.W.; Zeng, M. Image-Based Automatic Watermeter Reading under Challenging Environments. *Sensors* **2021**, *21*, 434. [[CrossRef](#)]
4. Li, Z.; Zhou, Y.S.; Sheng, Q.H.; Chen, K.J.; Huang, J. A high-robust automatic reading algorithm of pointer meters based on text detection. *Sensors* **2020**, *20*, 5946. [[CrossRef](#)]
5. Fang, H.; Ming, Z.Q.; Zhou, Y.F.; Li, H.Y.; Li, J. Meter recognition algorithm for equipment inspection robot. *Autom. Instrum.* **2013**, *28*, 10–14.
6. Shi, J.; Zhang, D.; He, J.; Kang, C.; Yao, J.; Ma, X. Design of remote meter reading method for pointer type chemical instruments. *Process Autom. Instrum.* **2014**, *35*, 77–79.
7. Huang, Y.L.; Ye, Y.T.; Chen, Z.L.; Qiao, N. New method of fast Hough transform for circle detection. *J. Electron. Meas. Instrum.* **2010**, *24*, 837–841. [[CrossRef](#)]
8. Zhou, F.; Yang, C.; Wang, C.G.; Wang, B.; Liu, J. Circle detection and its number identification in complex condition based on random Hough transform. *Chin. J. Sci. Instrum.* **2013**, *34*, 622–628.
9. Zhang, W.J. Pointer Meter Recognition via Image Registration and Visual Saliency Detection. Ph.D. Thesis, Chongqing University, Chongqing, China, 2016.
10. Gao, J.W. Intelligent Recognition Method of Meter Reading for Substation Inspection Robot. Master's Thesis, University of Electronic Science and Technology of China, Chengdu, China, 2018.
11. Lowe, D.G. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110. [[CrossRef](#)]
12. Bay, H.; Tuytelaars, T.; Gool, L.V. SURF: Speeded up robust features. In Proceedings of the Ninth European Conference on Computer Vision, Graz, Austria, 7–13 May 2006; pp. 404–417.
13. Nanni, L.; Lumini, A.; Loreggia, A.; Formaggio, A.; Cuza, D. An Empirical Study on Ensemble of Segmentation Approaches. *Signals* **2022**, *3*, 22. [[CrossRef](#)]
14. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.
15. Redmon, J.; Farhadi, A. YOLO9000: Better, Faster, Stronger. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 7263–7271.
16. 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 European Conference on Computer Vision, Amsterdam, The Netherlands, 11–14 October 2016; Springer: Cham, Switzerland, 2016; pp. 21–37.
17. Ren, S.Q.; He, K.M.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. *Adv. Neural Inf. Process. Syst.* **2015**, *28*, 91–99. [[CrossRef](#)] [[PubMed](#)]
18. Chen, L.C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs. *arXiv* **2014**. [[CrossRef](#)]
19. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; Springer: Cham, Switzerland, 2015; pp. 234–241.
20. Wang, W.H.; Xie, E.; Li, X.; Fan, D.P.; Song, K.T.; Liang, D.; Lu, T.; Luo, P.; Shao, L. Pvt v2: Improved baselines with pyramid vision transformer. *Comput. Vis. Media* **2022**, *8*, 415–424. [[CrossRef](#)]
21. Xing, H.Q.; Du, Z.Q.; Su, B. Detection and recognition method for pointer-type meter in transformer substation. *Chin. J. Sci. Instrum.* **2017**, *38*, 2813–2821.
22. Wan, J.L.; Wang, H.F.; Guan, M.Y.; Shen, J.L.; Wu, G.Q.; Gao, A.; Yang, B. An automatic identification for reading of substation pointer-type meters using faster R-CNN and U-Net. *Power Syst. Technol.* **2020**, *44*, 3097–3105.

23. Ni, T.; Miao, H.F.; Wang, L.L.; Ni, S.; Huang, L.T. Multi-meter intelligent detection and recognition method under complex background. In Proceedings of the 2020 39th Chinese Control Conference (CCC), Shenyang, China, 27–30 July 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 7135–7141.
24. Huang, H.Q.; Huang, T.B.; Li, Z.; Lyu, S.L.; Hong, T. Design of Citrus Fruit Detection System Based on Mobile Platform and Edge Computer Device. *Sensors* **2021**, *22*, 59. [[CrossRef](#)]
25. Chen, L.C.; Zhu, Y.K.; Papandreou, G.; Schroff, F.; Adam, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In Proceedings of the European conference on computer vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 801–818.
26. Lv, Y.W.; Ai, Z.Q.; Chen, M.F.; Gong, X.R.; Wang, Y.X.; Lu, Z.H. High-Resolution Drone Detection Based on Background Difference and SAG-YOLOv5s. *Sensors* **2022**, *22*, 5825. [[CrossRef](#)]
27. Lyu, S.L.; Li, R.Y.; Zhao, Y.W.; Li, Z.; Fan, R.J.; Liu, S.Y. Green Citrus Detection and Counting in Orchards Based on YOLOv5-CS and AI Edge System. *Sensors* **2022**, *22*, 576. [[CrossRef](#)]
28. YOLOv5. Available online: <https://github.com/ultralytics/yolov5> (accessed on 10 June 2020).
29. Wang, C.Y.; Liao, H.Y.M.; Wu, Y.H.; Chen, P.Y.; Hsieh, J.W.; Yeh, I.H. CSPNet: A New Backbone that can Enhance Learning Capability of CNN. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA, 14–19 June 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1571–1580.
30. He, K.M.; Zhang, X.Y.; Ren, S.Q.; Sun, J. Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *37*, 1904–1920. [[CrossRef](#)]
31. Lin, T.Y.; Dollar, P.; Girshick, R.; He, K.; Hariharan, B.; Belongie, S. Feature Pyramid Networks for Object Detection. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE Computer Society, Honolulu, HI, USA, 21–26 July 2017.
32. Liu, S.; Qi, L.; Qin, H.; Shi, J.; Jia, J. Path aggregation network for instance segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 8759–8768.
33. Neubeck, A.; Gool, L. Efficient Non-Maximum Suppression. In Proceedings of the International Conference on Pattern Recognition, IEEE Computer Society, Hong Kong, China, 20–24 August 2006.
34. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.C. Mobilenetv2: Inverted residuals and linear bottle-necks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 4510–4520.
35. Bochkovskiy, A.; Wang, C.Y.; Liao, H.Y.M. YOLOv4 Optimal Speed and Accuracy of Object Detection. *arXiv* **2020**, arXiv:2004.10934.
36. Redmon, J.; Farhadi, A. YOLOv3: An Incremental Improvement. *arXiv* **2018**, arXiv:1804.02767.