

## Research Article

# Research on Taproots Identification Technology in *Panax notoginseng* Quality Intelligent Management System

Mingfang Chen <sup>1</sup>, Zhongping Chen,<sup>1</sup> Xiuming Cui,<sup>2</sup> Yongxia Zhang,<sup>1</sup> and Sen Wang <sup>1</sup>

<sup>1</sup>School of Mechanical and Electrical Engineering, Kunming University of Science and Technology, Kunming 650500, China

<sup>2</sup>Key Laboratory of Sustainable Utilization of *Panax Notoginseng* Resources in Yunnan Province, Kunming 650500, China

Correspondence should be addressed to Sen Wang; wangsen0401@126.com

Received 27 July 2021; Accepted 31 August 2021; Published 17 September 2021

Academic Editor: Yu-Ting Bai

Copyright © 2021 Mingfang Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the *Panax notoginseng* quality intelligent management system, the big roots and fibrous roots cannot be cut automatically because the machine cannot distinguish the taproot, big roots, and fibrous roots of *Panax notoginseng*, resulting in the automatic cutting mechanism unable to obtain the control trajectory coordinate reference of the tool feed. To solve this problem, this paper proposes a visual optimal network model detection method, which uses the image detection method of marking anchor frames to improve the detection accuracy. A variety of deep learning network models are modified by the TensorFlow framework, and the best training model is optimized by comparing the results of training, testing, and verification data. This model is used to automatically identify the taproots and provide the control trajectory coordinate reference for the actuator that cuts big roots and fibrous roots automatically. The experimental results show that the optimal network model studied in this paper is effective and accurate in identifying the taproots of *Panax notoginseng*.

## 1. Introduction

The *Panax notoginseng* industry is an important component of Yunnan province in China. The Yunnan *Panax notoginseng* processing plant needs to process a big amount of *Panax notoginseng* raw materials in harvest season. These raw materials will be powdered or sliced to prepare them for use in subsequent pharmaceuticals. However, there is an important and critical step in the process of making *Panax notoginseng* powder or slicing is to cut big roots and fibrous roots. The harvested *Panax notoginseng* is divided into three parts: the taproot, the big roots, and the fibrous roots; *Panax notoginseng* is shown in Figure 1. Although the effective ingredients contained in the three parts are the same, the saponin content of each part is quite different. Therefore, adding a big proportion of big roots and fibrous roots will reduce the effective content of saponin and affect the subsequent pharmaceuticals. At present, the big roots and fibrous roots of *Panax notoginseng* are cut by manual visual inspection. Due to the different shapes and sizes of *Panax*

*notoginseng*, human visual inspection will have different results because everyone has different criteria for judging the big roots and fibrous roots of *Panax notoginseng*, so the standards for cutting are different. Because long hours of manual work will cause visual fatigue, there will be different judgment standards at different times. These different judgment standards will cause differences in shearing, which will affect the subsequent pharmaceuticals made form, *Panax notoginseng*.

Target detection aims to find all objects of interest through the computer, including classifying the objects and determining the location of the objects [1]. It could not only effectively reduce the interference of human factors but also realize stable identification and high-precision detection. Therefore, target detection is widely used in industry and agriculture [2]. Common target detection is divided into traditional target detection and target detection based on deep learning [3].

Traditional target detection mainly uses artificial selection of features to detect objects [4]. Detection is usually divided into three stages. First, a certain part of the image is

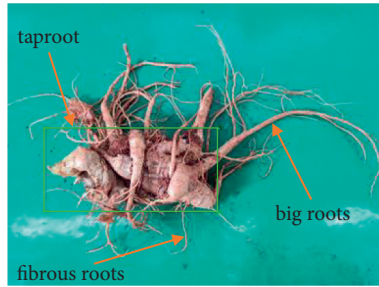


FIGURE 1: *Panax notoginseng*.

determined as a candidate area through a multiscale sliding window, then the features of the candidate area are extracted by the constructed features, and the classifier is used for classification [5–8]. Traditional target detection could better realize target recognition in which the image was in specific scenes and fewer categories [9]. However, if the external environment, such as the lighting changes, it is difficult to achieve a good detection effect, since it is easy to miss or incorrectly detect. At the same time, traditional target detection will also generate a large number of candidate boxes, which increases the computational overhead, and the accuracy and speed of detection cannot meet the requirements of practical industrial applications [10, 11].

Viola et al. [12, 13] used the Viola Jones detector to achieve real-time face detection for the first time by a direct detection method without any constraints at the time, but it required more computational ability than available at that time. Dalal and Triggs [14] used a directional gradient histogram feature descriptor that had a superior recognition ability for pedestrian detection. However, this algorithm is prone to missed detection when dealing with occlusion, large posture changes, and target angle change [15]. Felzenszwalb et al. [16] proposed a model based on variable parts to detect the target object through the ideas of an improved histogram of gradient (HOG) feature and support vector machine (SVM) classifier. This detection method has a certain robustness for deformed targets, but the method has poor stability for large-scale rotating targets and requires an artificial design of different excitation templates for detection.

Target detection based on deep learning has achieved great success in the field of detection [17]. Compared with traditional target detection, this algorithm based on deep learning could not only use fewer data sets to obtain high-precision detection results but also greatly improve the speed of target detection [18]. Therefore, it is widely used in industry and agriculture.

Aiming at the fact that currently the big roots and fibrous roots of *Panax notoginseng* are cut manually, this article proposed a visual inspection method for developing an optimal model of identifying *Panax notoginseng* taproots. This method could improve the accuracy and stably identify the taproots of *Panax notoginseng* and be prepared for the subsequent cutting of the big roots and fibrous roots of *Panax notoginseng* by machines. To improve the stability and reliability of target detection and to meet the need for rapid detection in industry, this paper uses the bounding box

method to mark the anchor frame to detect the target. To reduce the number of training samples, this article applies LabelImg mapping software to mark the target features of limited samples and rely on the TensorFlow framework to modify a variety of different deep learning network models. By comparing the training, testing, and verification performance of different network models, this article selects the optimal network model to detect the taproots of *Panax notoginseng*.

## 2. Detection of the Taproot Target of *Panax notoginseng*

Visual detection of the *Panax notoginseng* taproots involved five steps: image acquisition of *Panax notoginseng*, database construction, network model training, model performance comparison, and *Panax notoginseng* taproot detection. The specific implementation process is shown in Figure 2.

**2.1. Image Acquisition of *Panax notoginseng*.** The taproot detection of *Panax notoginseng* requires the collection of *Panax notoginseng* images. The acquisition tool used was CMOS industrial camera (Basler ace-acA2500-14gc, F1.4-F16), which has a higher acquisition frequency and acquisition speed. The use of a vision and motion module in LabVIEW to collect images of *Panax notoginseng* can fully meet the requirements of industrial collection [19]. The flowchart of image acquisition is shown in Figure 3. Through the collection results, this paper found that although the use of higher resolution provided a wider field of view in the image, it required a larger storage space and took more time for image collection and processing. The low image resolution reduced the field of view of the image, but it saved the storage space of the image and reduced the image collection and processing time. Because the detection of the taproot of *Panax notoginseng* was a large target detection, the recognition of *Panax notoginseng* taproots does not require higher acquisition speed and higher resolution in actual industrial applications, so this article sets the acquisition resolution of the industrial camera at  $1156 \times 868$  and sets the rate of image acquisition to 1 frame/sec.

**2.2. Dataset Construction.** It was necessary to construct a dataset of *Panax notoginseng* images before using detection on the taproot. In this paper, we selected *Panax notoginseng* images that contained only big root and fibrous root images from the collected images to construct the entire image dataset. The taproots of *Panax notoginseng* images have different directions, positions, shapes, and backgrounds. The big root and fibrous root images also had different shapes and backgrounds. This article used a random sampling method to divide the *Panax notoginseng* image dataset into a training set, test set, and validation set according to a certain proportion so that the samples were more random and extensive. This article also used the LabelImg plotting tool to anchor the taproot in the training set and test set of *Panax notoginseng*. We integrated the files generated after the anchor frame with the corresponding image files to

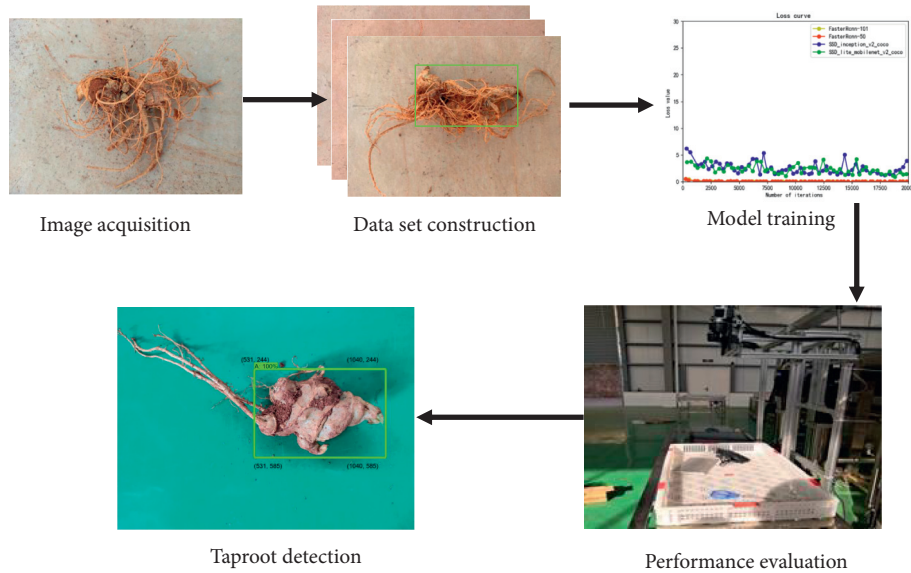


FIGURE 2: Determination of the taproot of *Panax notoginseng*.

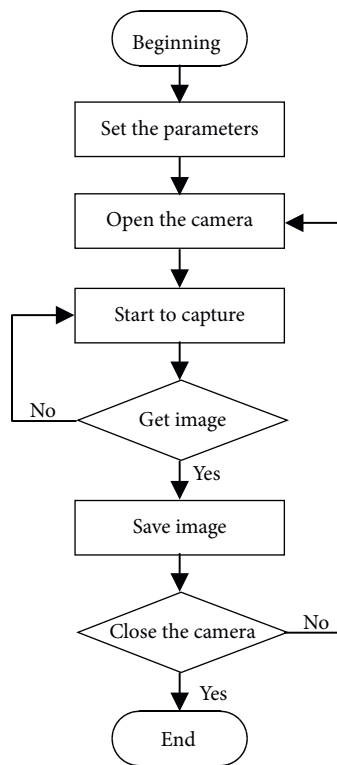


FIGURE 3: Image acquisition flowchart.

jointly construct the dataset. The process of constructing the dataset is shown in Figure 4.

2.3. *Network Model Training.* At present, the big roots and fibrous roots of *Panax notoginseng* are still mainly done by manual cutting. The shape of the roots are complicated, different workers have different standards for judging big roots during cutting, and the same workers also have different

standards for cutting in different time periods during the day. Target detection based on deep learning can solve the problem of inconsistent manual detection standards, and at the same time, it can avoid the situation of incorrect detection or missed detection of traditional target detection in a complex background with multiple features. Considering the practical applications of engineering, this article used the bounding box method to mark the anchor box to identify the taproot of *Panax notoginseng*, wrote the entire deep learning model

framework in Python, and used the constructed *Panax notoginseng* dataset. A public deep learning model downloaded from the Internet was used as a pretraining model, and the taproot recognition of *Panax notoginseng* was realized by modifying the parameters of the training model. The selected network model includes single-stage SSD [20] and two-stage faster-RCNN [21] as the network pretraining model. Under the TensorFlow framework, different network models were trained and tested, the average test accuracy between different optimal models was verified, and the optimal network that meets the taproot detection requirements was selected. The verification formula used in this article is shown in formulas (1) and (2), and the SSD model skeleton [22] used to train detection of the *Panax notoginseng* taproot is shown in Figure 5, where  $N$  is the total number of sheets in the training set and test set,  $N \in (1, 2, \dots, N)$ ,  $P_N$  is the detection accuracy of the model, AVPE is the average detection accuracy of the model,  $TP_N$  is the number of *Panax notoginseng* taproots correctly detected, and  $FP_N$  is the number of background detections as the taproot of *Panax notoginseng*:

$$P_N = \frac{TP_N}{TP_N + FP_N}, \quad (1)$$

$$AVPE = \frac{1}{N} \sum_1^N (P_1, P_2, \dots, P_N). \quad (2)$$

**2.4. Model Performance Comparison.** This paper used different deep learning network models to detect the taproot of *Panax notoginseng*. However, there are obvious differences in the detection effects of different network models, and the average detection accuracy could not effectively select the optimal network model. To better realize the comparison of various network models and select the optimal network model, this paper proposed a comprehensive performance verification formula that combines average accuracy, accuracy, missed detection rate, and false detection rate for deep learning model performance evaluation and comparison. The verification formula is shown in (6), where  $H_p$  is the number of the taproots of *Panax notoginseng* predicted by the deep learning model,  $H_t$  is the sample size of the real *Panax notoginseng* taproot,  $H_l$  is the number of missed detections of the taproot of *Panax notoginseng*, and  $H_f$  is the number of errors during the detection. In this paper, the determined deep learning model was verified, and the performance of different models was evaluated on the validation results until the optimal network model was selected:

$$P_a = \frac{H_p}{H_t}, \quad (3)$$

$$P_l = \frac{H_l}{H_t}, \quad (4)$$

$$P_f = \frac{H_f}{H_t}, \quad (5)$$

$$C_{PE} = \frac{1}{N} \sum_1^N (\text{mean}(AVPE + P_a) - P_l - P_f). \quad (6)$$

**2.5. The Taproot Detection of *Panax notoginseng*.** *Panax notoginseng* taproot detection was to perform image detection on the verification set according to the optimal model trained in deep learning. The hardware 3D diagram of detection is shown in Figure 6. The physical platform diagram is shown in Figure 7. The hardware mainly includes industrial cameras, industrial tablet computers, racks, and test platforms. Among them, the test platform was used, and the industrial tablet computer was used to embed the optimal network model to realize the detection and result display of the images collected by the industrial camera. The detection flowchart is shown in Figure 8, and the detection result is shown in Figure 9.

### 3. Recognition Results and Analysis of the Taproot of *Panax Notoginseng*

**3.1. Experimental Environment and Configuration.** For the effectiveness of the visual inspection of the taproot of *Panax notoginseng*, 175 images with multiple backgrounds and multiple shapes containing the taproot and 25 images with only big roots and fibrous roots were selected from the 1685 collected images. The 200 images were divided into a training set, test set, and validation set at a ratio of 7 : 2 : 1 by random sampling. The experimental platform environment is 2(Intel (R) Core(TM)i7-9700F CPU@3.00 GHz, GPU GeForce RTX 2070S, 32 GB memory, Windows7-64bit. The deep learning network model used two frameworks: SSD and Faster-RCNN under TensorFlow. The 4 network models included SSD\_inception\_v2\_coco, SSD\_Lite\_mobile-net\_v2\_coco, Faster-RCNN-Resnet50, and Faster-RCNN-Resnet101. We modified and trained the pretraining models of these 4 network models; the results of training used loss, AVPE, CPE, Pa, Pl, and Pf to compare the qualitative and quantitative analysis of the entire model.

**3.2. Model Training and Verification.** The results of training and verification of the four network models were batchSize = 4, learningRate = 0.0001, the number of iterations was 20,000, and the other parameters were defaulted. The generated iteration number and loss value curve are shown in Figure 10. It can be seen from the figure that the loss value of SSD\_inception\_v2\_coco was the largest at the beginning, and the loss values of RCNN-Resnet101 and Faster-RCNN-Resnet50 were the smallest. However, after 1000 iterations of SSD\_inception\_v2\_coco, the loss value decreased, but the model loss function curve was in a state of nonconvergence. At the same time, the loss function curve of the SSD\_Lite\_mobile-net\_v2\_coco model was also in a state of nonconvergence, and the loss functions of SSD\_Lite\_mobile-net\_v2\_coco and SSD\_inception\_v2\_coco were

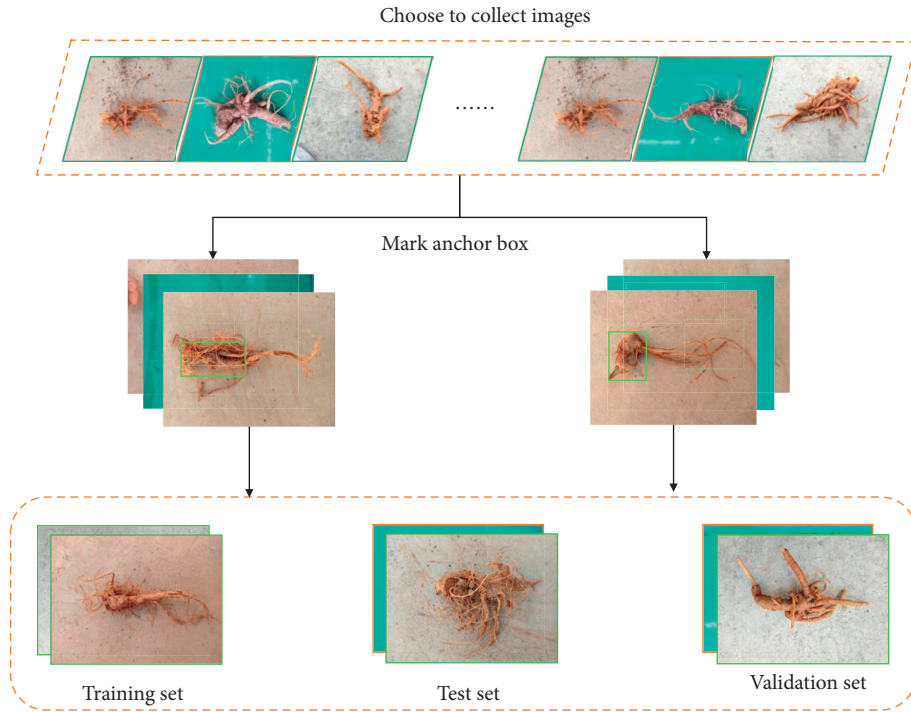


FIGURE 4: Dataset construction.

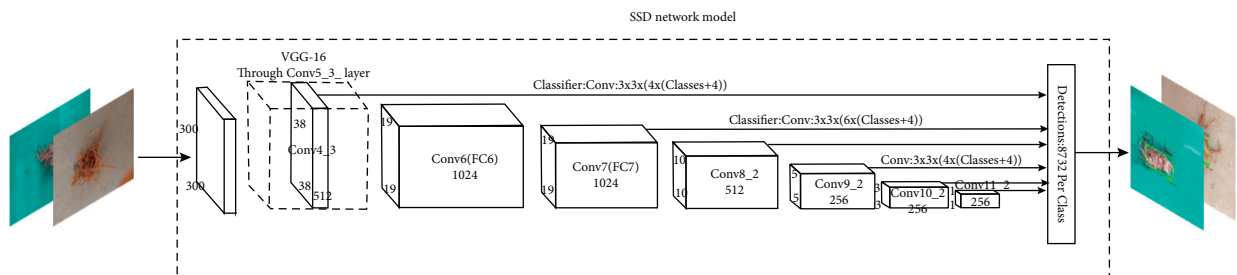


FIGURE 5: SSD network model skeleton.

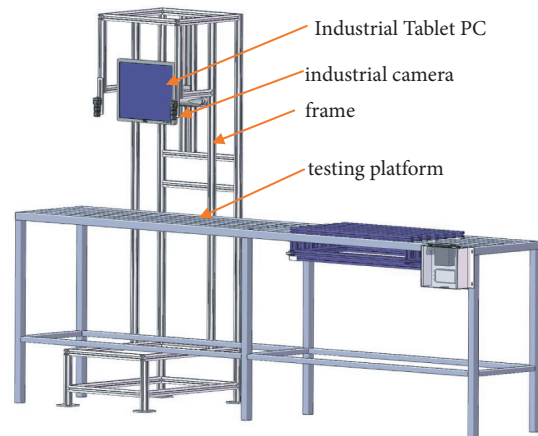


FIGURE 6: Hardware 3D diagram of taproot detection.

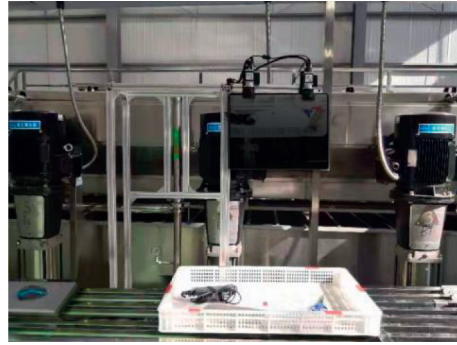


FIGURE 7: Physical detection platform.

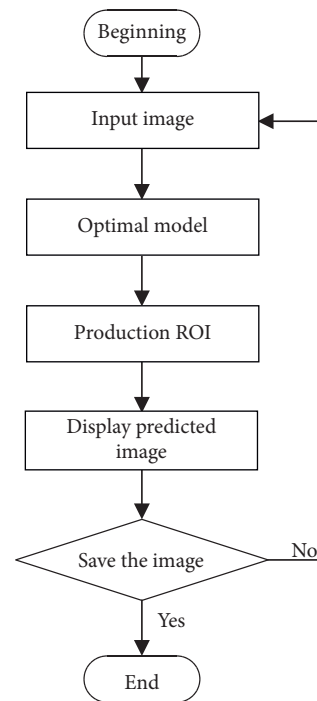


FIGURE 8: Flowchart of taproot detection of *Panax notoginseng*.

always higher than the function value of Faster-RCNN. The two model networks of Faster-RCNN had relatively small loss values at the beginning, as shown in Figure 11. However, compared with Faster-RCNN-Resnet50, the loss value of Faster-RCNN-Resnet101 was generally lower than that of Faster-RCNN-Resnet50. Although the loss value of Faster-RCNN-Resnet50 was lower than Faster-RCNN-Resnet101 when the number of iterations was 18000 times, the loss value of Faster-RCNN-Resnet101 was lower afterward. Therefore, compared with the four network models, Faster-RCNN-Resnet101 had a better modification effect. The comparison and verification of the optimal model among the four training models is shown in Figure 12. When the number of iterations was 19,811, the Faster-RCNN-Resnet101 model could detect the optimal verification result with an average detection rate of 99.99% for the taproot. The Faster-RCNN-Resnet50 model could detect the taproot with an average

detection accuracy of 99.90% when the number of iterations was 17,795. When the number of iterations of SSD\_Lite\_mobilenet\_v2\_coco was 17987, the average detection accuracy of the taproot of the optimal model was 98.36%. The optimal model of SSD\_inception\_v2\_coco had a poor average detection rate for the taproot, which was only 90.25%. The comprehensive results show that using the *Panax notoginseng* dataset, Faster-RCNN-Resnet101 was more suitable for target detection than the other models [23].

**3.3. Model Testing and Comparison.** To effectively verify the universality of the four deep learning models for the detection of the taproot of *Panax notoginseng*, in this paper, 18 images from the *Panax notoginseng* validation set were used to compare and analyze the performance of the four optimal performance models. Part of the

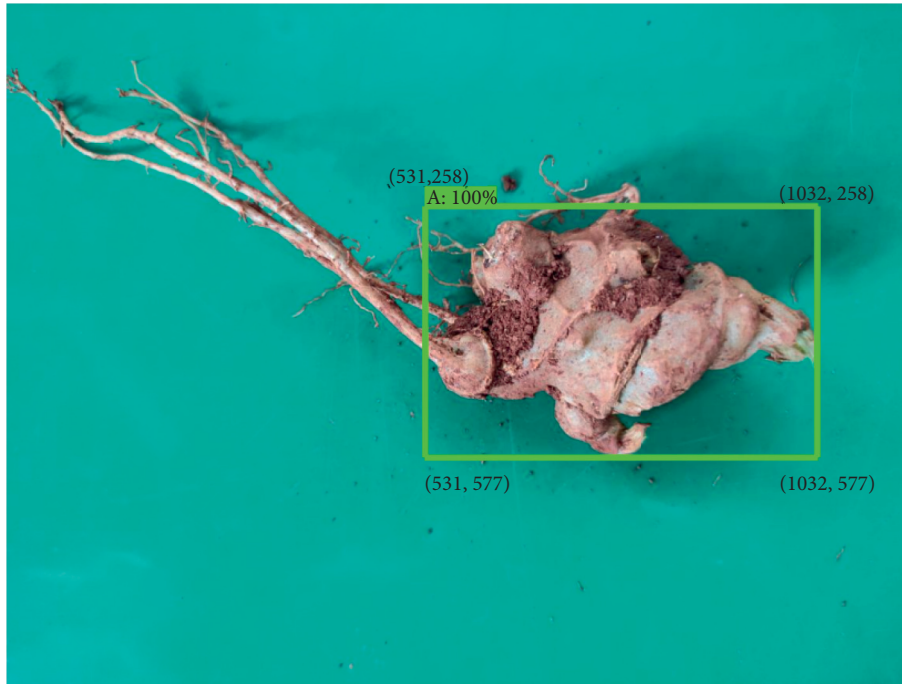


FIGURE 9: An example of taproot detection of *Panax notoginseng*.

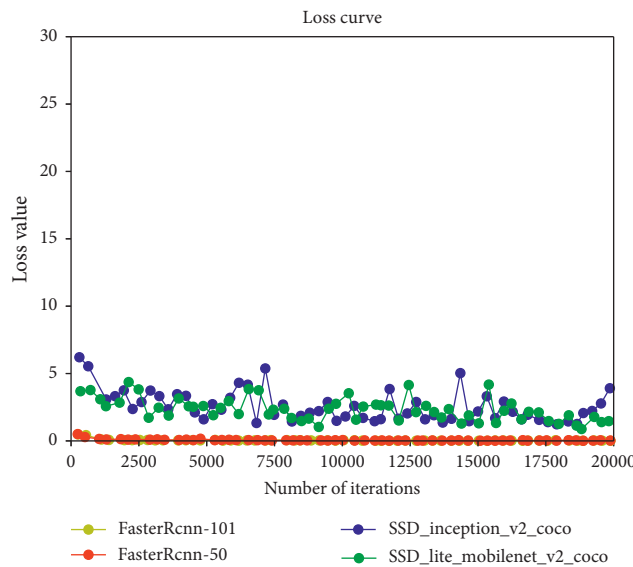


FIGURE 10: Four model loss function curves.

qualitative analysis is shown in Figure 13. By comparing and analyzing the results, it could be seen that the four network models of SSD\_inception\_v2\_coco, SSD\_lite\_mobilenet\_v2\_coco, Faster-RCNN-Resnet50, and Faster-RCNN-Resnet101 can detect the taproot of *Panax notoginseng*, but the detection of SSD\_inception\_v2\_coco was unstable, and the detection of some complex multi-root images would be missed. We listed all of the model detection results in the first row in Figure 13 and observed the categories and detection frame positions of the *Panax*

*notoginseng* taproots by the four models. The enumeration diagram is shown in Figure 14. Through observation, it was found that Faster-RCNN-Resnet101 can stably detect the taproot target. The SSD\_lite\_mobilenet\_v2\_coco and Faster-RCNN-Resnet50 models can also detect all taproot images. However, for complex *Panax notoginseng* images, there would be a problem of detecting multiple frames. Further qualitative comparison of some of the verification images in the *Panax notoginseng* verification set was carried out to achieve the accuracy comparison of the

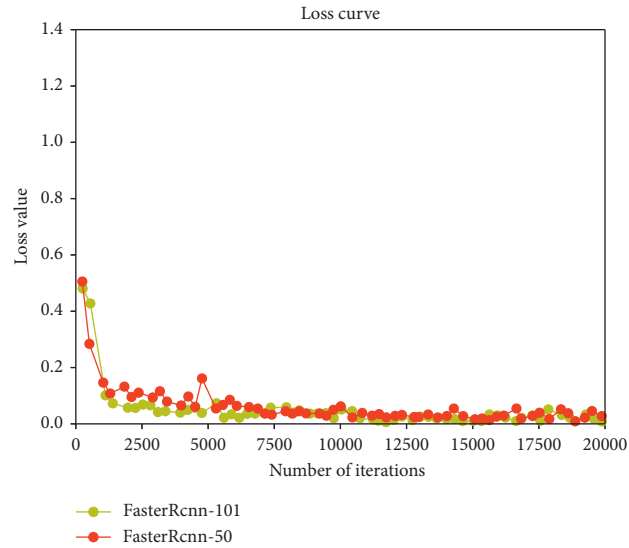


FIGURE 11: Faster-RCNN loss function curve.

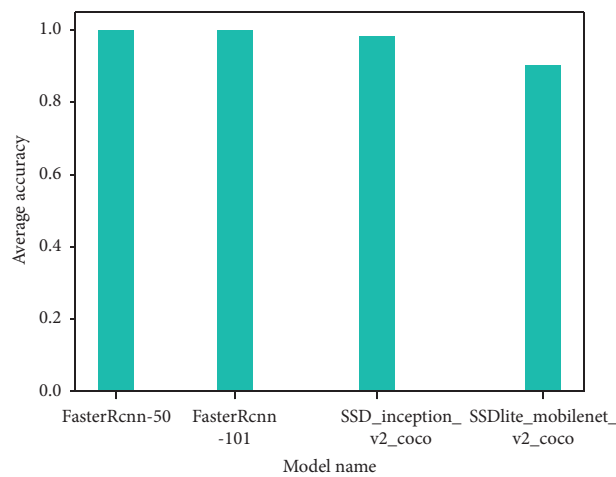


FIGURE 12: Average accuracy of the model.

bounding box. Figure 15 shows the comparison result of a certain area randomly selected in this paper. The frame selection accuracy of the four models would have certain differences. Taking the second row in Figure 13 as an example, Faster-RCNN-Resnet101 could detect surrounding targets other than the taproot of *Panax notoginseng*, and the detection confidence was 100%. Faster-RCNN-Resnet50 could also detect the taproot target, but it could not be detected correctly, and the detection accuracy rate was also lower than that of Faster-RCNN-

Resnet101. The two models SSD\_Lite\_mobilenet\_v2\_coco and SSD\_inception\_v2\_coco could be detected correctly, but the detection confidence was lower than that of Faster-RCNN-Resnet101. The quantitative comparison results in Table 1 also show that Faster-RCNN-Resnet101 had a stable and good detection effect on the taproot of *Panax notoginseng*. Therefore, by comparing the results, it is concluded that Faster-RCNN-Resnet101 was more suitable for the visual inspection of the taproot of *Panax notoginseng*.



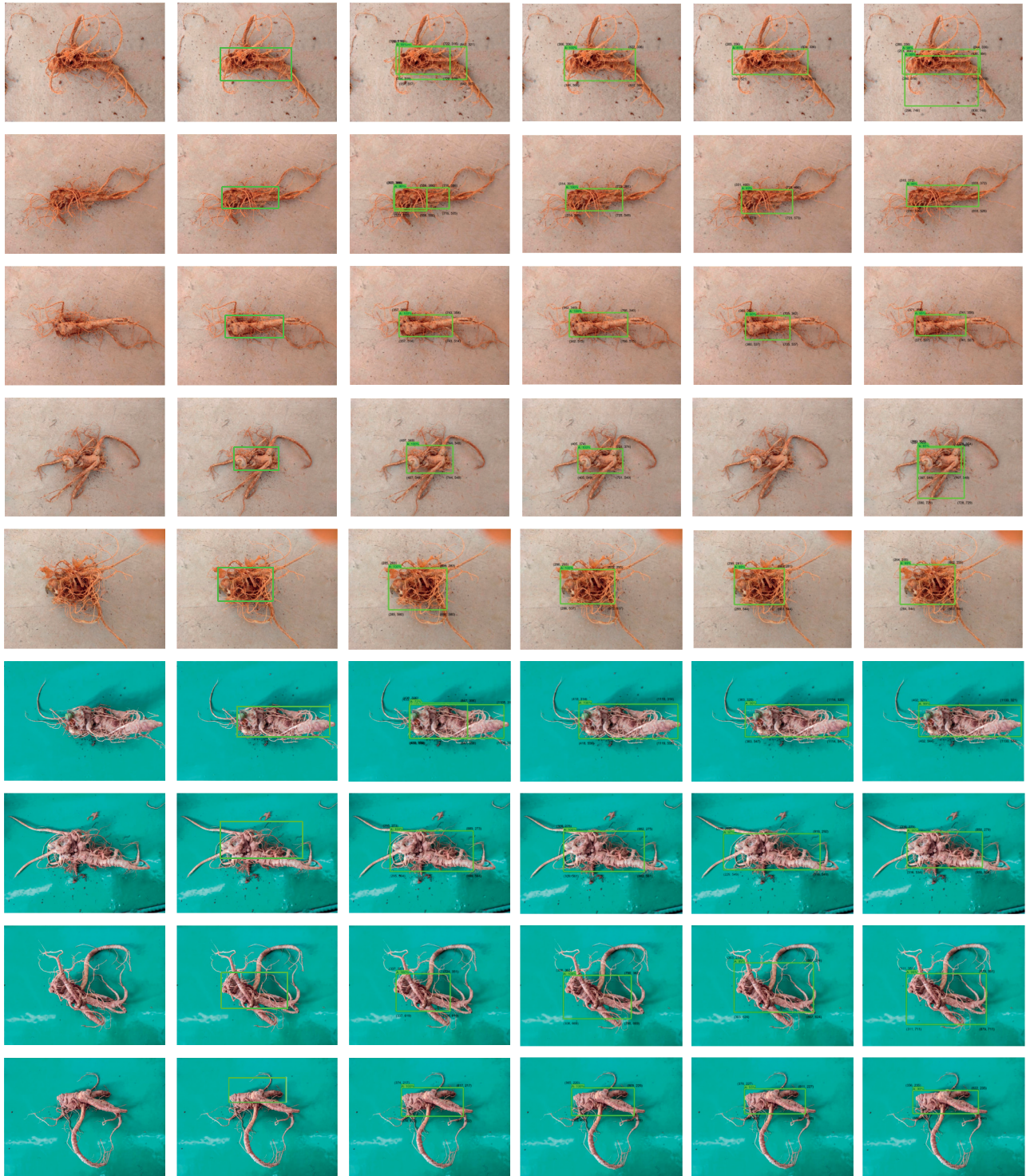


FIGURE 13: Comparison results of partial qualitative results of the validation of the four optimal models. From left to right, raw image, GT, Faster-RCNN-Resnet50, Faster-RCNN-Resnet101, SSD\_INCEPTION\_V2\_COCO, and SSD\_LITE\_MOBILENET\_V2\_COCO.

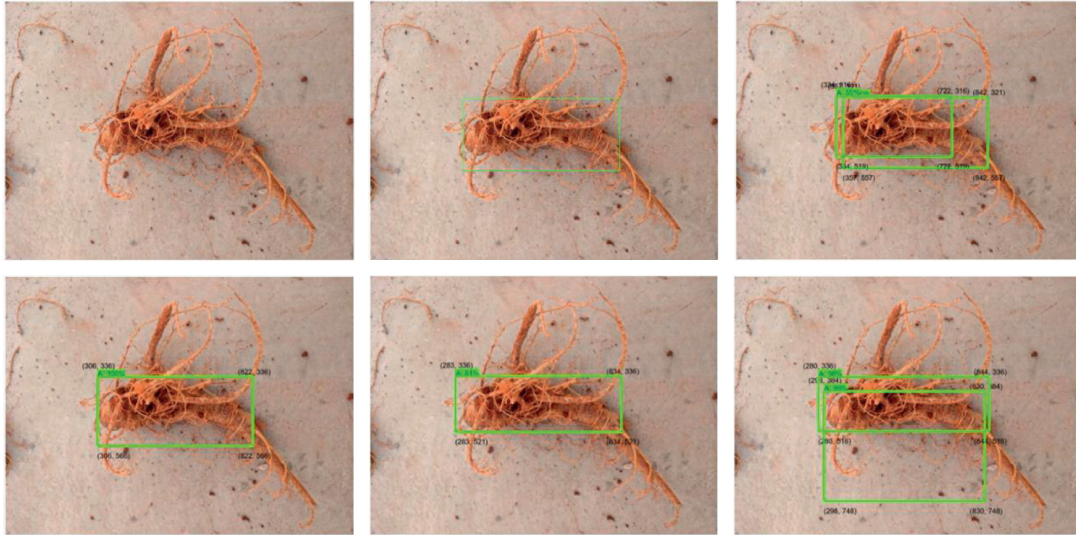


FIGURE 14: Comparison result of row 1 in Figure 13.

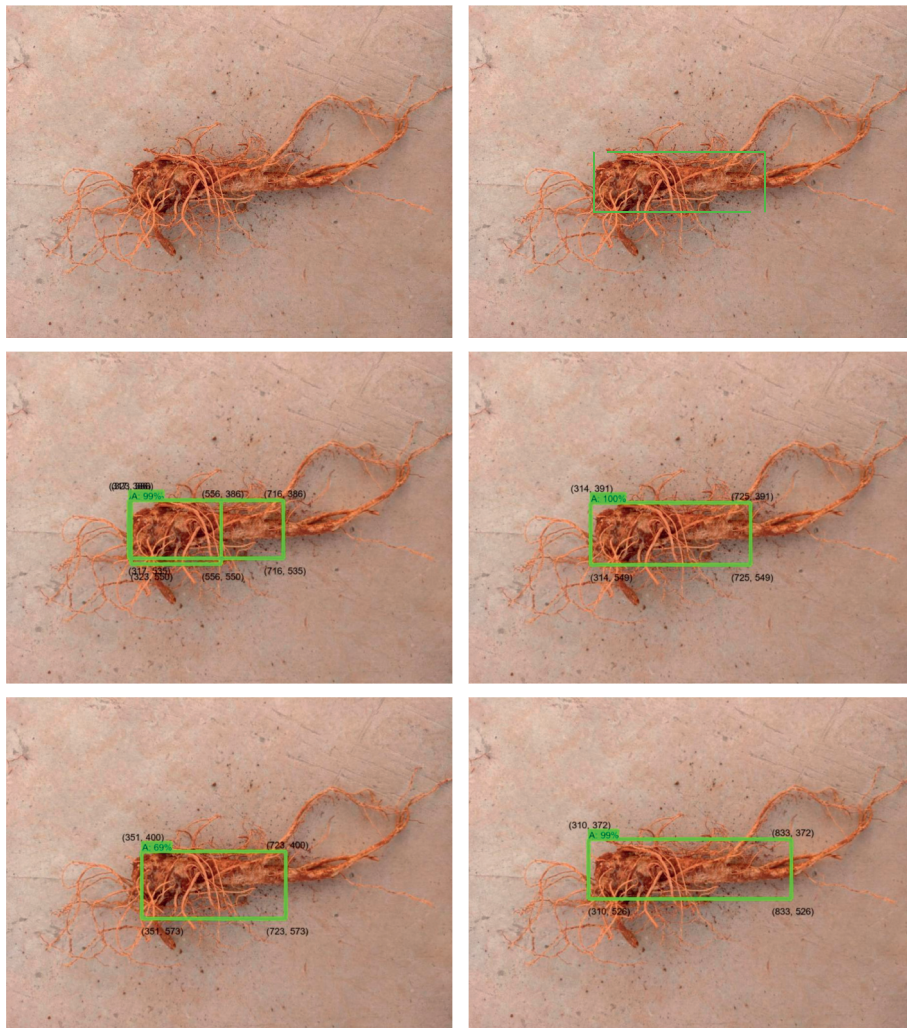


FIGURE 15: Confidence comparison of row 2 in Figure 13.

TABLE 1: CPE comparison results of the four optimal model validations.

| Model                      | $H_t$ | $H_p$ | $H_l$ | $H_f$ | $P_a$ (%) | $P_l$ (%) | $P_f$ (%) | AVPE (%) | $C_{PE}$ (%) |
|----------------------------|-------|-------|-------|-------|-----------|-----------|-----------|----------|--------------|
| SSD_inception_v2_coco      | 18    | 17    | 1     | 0     | 94.44     | 5.56      | 0         | 90.25    | 89.57        |
| SSD_Lite_mobilenet_v2_coco | 18    | 16    | 0     | 2     | 88.89     | 0         | 11.11     | 98.36    | 88.07        |
| Faster-RCNN-Resnet50       | 18    | 15    | 0     | 3     | 83.33     | 0         | 16.67     | 99.90    | 83.28        |
| Faster-RCNN-Resnet101      | 18    | 18    | 0     | 0     | 100       | 0         | 0         | 99.99    | 99.99        |

## 4. Conclusion

To avoid the defects of manual visual inspection and realize the taproot target recognition of *Panax notoginseng*. This article constructed an optimal model based on deep learning to detect the taproot target, relying on TensorFlow various deep learning network models for modification and training, and used the proposed optimization strategy to compare the training, testing, and verification of different network models to select the optimal network model, realizing the detection of the taproot target of *Panax notoginseng*. The experimental results showed that, by modifying and training different network models and selecting the optimal network model, the detection efficiency and accuracy of the target could be effectively improved, and the detection accuracy of the optimal model could reach 99.99%. However, this article does not give the identified taproot position parameters to the controller. Therefore, in future work, we can continue to modify the parameters of the optimal network model and transmit the taproot coordinates to the controller so that the actuator can accurately cut the big roots and fibrous roots of *Panax notoginseng*.

## Data Availability

The datasets used to support the findings of the study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported in part by the National Natural Science Foundation of China, under Grants 51965029 and 52065035, and the National Key Research and Development Plan Project, under 2017YFC1702503.

## References

- [1] K. Q. Chen and Z. L. Zhu, "A review of deep learning for multi-scale target detection," *Journal of Software*, vol. 32, no. 4, 2021.
- [2] D. Yang, X. Gao, L. Kong, Y. Pang, and B. Zhou, "An event-driven convolutional neural architecture for non-intrusive load monitoring of residential appliance," *IEEE Transactions on Consumer Electronics*, vol. 66, no. 2, pp. 173–182, 2020.
- [3] H. Y. Xu and Z. S. Long, "Multi-scale target detection algorithm for ship intelligent navigation," *Journal of Huazhong University of Science and Technology (Natural Science Edition)*, vol. 49, no. 3, pp. 50–55, 2021.
- [4] S. Li, Y. C. Zhang, and T. Lan, "A visual perception target detection algorithm for high resolution remote sensing images," *Journal of Xi'an Jiaotong University*, vol. 52, no. 6, pp. 9–16, 2018.
- [5] Z. Zhao, G. Xu, and Y. Qi, "Representation of binary feature pooling for detection of insulator strings in infrared images," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 23, no. 5, pp. 2858–2866, 2016.
- [6] C. A. Laurentys Almeida, A. P. Braga, S. Nascimento et al., "Intelligent thermographic diagnostic applied to surge arresters: a new approach," *IEEE Transactions on Power Delivery*, vol. 24, no. 2, pp. 751–757, 2009.
- [7] Q. Wu and J. An, "An active contour model based on texture distribution for extracting inhomogeneous insulators from aerial images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 6, pp. 3613–3626, 2014.
- [8] S. C. Zhu, Q. W. Gao, and Y. X. Lu, "Identification and location of insulator based on frequency tuning," *Transactions of China Electrotechnical Society*, vol. 33, no. 23, pp. 5573–5580, 2018.
- [9] X. L. Liao, *Research on Target Detection and Behavior Recognition Technology for Intelligent Video Surveillance*, University of Electronic Science and Technology of China, Chengdu, China, 2019.
- [10] L. L. Fan, H. W. Zhao, and H. Y. Zhao, "Review of target detection based on deep convolutional neural network [J]," *Optics and Precision Engineering*, vol. 28, no. 5, pp. 161–173, 2020.
- [11] H. L. Luo and H. K. Chen, "A review of target detection based on deep learning," *Acta Electronica Sinica*, no. 6, 2020.
- [12] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, Kauai, HI, USA, 2001.
- [13] P. Viola and M. J. Jones, "Robust real-time face detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137–154, 2004.
- [14] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proceedings of the IEEE Computer Society Conference on Computer Vision & Pattern Recognition*, IEEE, Puerto Rico, USA, June 2005.
- [15] J. Y. Li, B. Gao, and X. M. Tong, "Video pedestrian detection method based on YOLO and GMM," *Journal of China Academy of Electronic Sciences*, vol. 14, no. 3, pp. 265–271, 2019.
- [16] P. Felzenszwalb, D. McAllester, and D. Ramanan, "A discriminatively trained, multiscale, deformable part model[C]// 2008 IEEE conference on computer vision and pattern recognition," *IEEE*, pp. 1–8, 2008.
- [17] H. Liu, L. S. Zhang, and Y. Shen, "Real-time pedestrian detection method in orchard based on improved SSD [J]," *Transactions of the Chinese Society for Agricultural Machinery*, vol. 50, no. 4, pp. 29–35, 2019.
- [18] L. Jing, R. Wang, and H. Liu, "Pedestrian detection and location in orchard based on binocular camera and improved

- YOLOV3 algorithm,” *Transactions of the Chinese Society for Agricultural Machinery*, no. 9, 2020.
- [19] W. Miao, Q. Lin, W. Zhang, and N.-J. Wu, “A programmable SIMD vision chip for real-time vision applications,” *IEEE Journal of Solid-State Circuits*, vol. 43, no. 6, pp. 1470–1479, 2008.
- [20] L. Liu, W. Ouyang, X. Wang et al., “Deep learning for generic object detection: a survey,” *International Journal of Computer Vision*, vol. 128, no. 2, pp. 261–318, 2020.
- [21] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: towards real-time object detection with region proposal networks towards real-time object detection with region proposal networks,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2017.
- [22] W. Liu, D. Anguelov, and D. Erhan, “SSD: single shot multiBox detector,” in *Proceedings of the European Conference on Computer Vision*, Springer, Glasgow, UK, August 2016.
- [23] H. Ni, M. Wang, M. Wang, and L. Zhao, “An improved Faster R-CNN for defect recognition of key components of transmission line,” *Mathematical Biosciences and Engineering*, vol. 18, no. 4, pp. 4679–4695, 2021.