International Journal of Information Technology and Applications
Volume 1, Issue 2, June 2024 ISSN 3007-7397

Neural Sensor Network ST-YOLO for Plant Disease Detection and Growth Guidance

Linjuan Ma[1], Fuquan Zhang[2,*]

[1]*School of Computer Science and Technology, Beijing Institute of Technology, China*
[2]*College of Computer and Control Engineering, Minjiang University, China*

mlj@bit.edu.cn, zfq@mju.edu.cn

Abstract. In recent years, plant disease detection has grown in importance for biotechnology, and its uses include growth guidance for plants. This paper presents a novel deep neural method based on the sensor network Swin Transformer-YOLO (ST-YOLO) that performs plant disease detection and segmentation to prevent the plants and fruits from spreading disease and to guide biological growth. Several experiments were performed to validate the proposed method. The results showed that obtaining information on the disease detection and biological state based on the RoCoLe dataset using ST-YOLO was more accurate and efficient than other typical methods. In particular, the mAP50 of the proposed approach is 0.987, 0.7% higher than state-of-the-art approaches. Therefore, the proposed approach is practical.

Keywords: Deep learning, Sensor network, Plant detection, ST-YOLO.

1. Introduction

Biotechnology has recently attracted increasing academic attention, contributing to developing the agricultural sector. In particular, intelligent computing has been implemented in biotechnology to improve the survival rate of biological culture technology and reduce production costs (Trkolu et al., 2019; Gupta et al., 2019). In recent years, implementing intelligent computing has attracted considerable interest in the bioinformatics community (Haridasan et al.,2023). For example, one application branch of biotechnology focuses on plant and fruit disease detection to prevent the plants and fruits from sickness and spreading (Rashwan et al., 2022; Li et al., 2022). Furthermore, the approaches can be divided into two general classes: traditional model detection and deep neural network-based detection models.

2. Related Works

Numerous researchers have considered traditional modelling methods to implement plant and fruit disease detection (Hamuda et al., 2017; Yao et al., 2017; Islam et al., 2019; Zou et al., 2019; Thirthe et al., 2021; Bose et al., 2021). In particular, Hamuda et al. (2017) proposed an automatic crop detection using the HSV color space and morphological operations. Then, Yao et al. (2017) presented a three-layer detection method to detect and identify white-backed planthoppers based on AdaBoost, support vector machine (SVM), and histogram of oriented gradient (HOG). Moreover, an automatic plant detection method was developed using HOG and LBP features with SVM (Islam et al., 2019). Furthermore, Zou et al. (2019) combined HOG features with local binary patterns (LBP) features to detect tomato disease and diagnose pests. Additionally, a tobacco-plant detection method was proposed using scale-invariant HOG Descriptors (Thirthe et al., 2021). Similarly, Bose et al. (2021) presented a leaf disease of medicinal plants detection method based on the SVM classification algorithm.

Although substantial progress has been achieved in plant and fruit disease detection using traditional methods, no clear advancement has been obtained in the trade-off between the speed and efficiency of the detection method. To fill these gaps, several authors have focused on plant and fruit-disease detection methods relying on deep learning methods. For example, Kumar et al. (2019) proposed detecting fruit flowers using a refined semantic weight convolutional neural network (CNN). Moreover, Lim et al. (2020) introduced a deep neural network to perform real-time Kiwi fruit flower detection in an orchard environment. Boogaard et al. (2020) considered a robust node-detection algorithm based on a deep CNN to detect and track fruit-vegetable crops. Subsequently, a hybrid detection model for an automatic plant disease was proposed using a convolutional autoencoder (CAE) network and CNN (Bedi et al., 2021).

Most studies on deep neural networks in this field can be grouped into two broad areas: one-stage detection methods (Liu et al., 2016; Joseph et al., 2016; Joseph et al., 2017; Joseph et al., 2018) and two-stage detection methods (Girshick et al., 2014; Ren et al., 2015; Sandler et al., 2018). In one-stage methods, the features are extracted directly from the network to predict the classification and location of objects without region proposal. Thus, the task flow can be described as a feature extraction and classification or localization regression. In contrast, in two-stage methods, the region proposal is generated. Here, the tested object might be contained in a pre-selected box, and then the samples are classified using CNN. Moreover, the task flow can be described as a feature extraction, generation region proposal, and classification or localization regression. Based on the two-stage detection method, a multi-class apple detection method in dense-foliage fruiting-wall trees was presented using a faster region-CNN, and the fruit disease detection network was built upon Mask R-CNN (Gao et al., 2020; Chang et al., 2021). In contrast, based on the one-stage detection method, the architecture 'MangoYOLO' was proposed to perform real-time fruit detection (Koirala et al., 2019). In addition, Zheng et al. (2021) developed a one-stage neural network, namely YOLO BP, to detect green citrus in natural environments. Furthermore, Roy et al. (2021) presented a high-performance and real-time one-stage object detection framework based on YOLOv4 in plant disease detection.

As seen from the above paragraph, significant progress has been made in this field; however, the following limitations arise in actual practice: 1) It is still difficult to locate and detect the object area accurately due to sunlight and backlighting in the planting environment. 2) Noise overlap due to plant feature loss and occlusion in a complex natural environment. 3)Most studies focused on detection accuracy, and more effort needs to be paid to real-time performance. 2) Various studies have been conducted on detection methods, but only a limited amount of research has been conducted on both detection and segmentation. 3) The algorithms are subject to the model size, which can be challenging to deploy in the system, relying heavily on the quantity of data to have improved performance.

Therefore, to overcome the above limitations, this study aims to detect and segment plants and fruits using the lightweight neural network ST-YOLO to protect them from disease and other outside effects. There are several advantages in YOLO models, such as faster speed, high accuracy and it can provide a unified framework for training model. However, the detection of small objects is relatively poor (Joseph et al., 2016; Joseph et al., 2017; Joseph et al., 2018). Considering the scales on different images or on the same image can vary greatly and compared with text, the size of image is too large, and the computational complexity is higher. The swin transformer is adopted in our network to solve the multi-scale problem of visual image, greatly reduces the computational complexity of the network and huge improvements can be made in detection, segmentation and other downstream tasks (Liu et al., 2021). On the object detection dataset COCO, it was improved by about 2.7 points over the SOTA and the semantic segmentation dataset ADE20K, the mAP improved by about 3.2 points compared with SOTA at that time (Liu et al., 2021). Therefore, the proposed method can address the gaps in the literature and be more accurate, efficient, and easier to deploy in an existing system to realize real-time performance. Moreover, the main contributions of the study can be summarized as follows: 1) An improved lightweight module was added to the developed neural network to reduce the model size while considering more abundant information. 2) A novel backbone network was included in the model to reduce parameter calculation and network uptime. 3) The data augmentation can be automatically generated in the proposed model to improve the performance in both detection and segmentation tasks.

3. Proposed Method

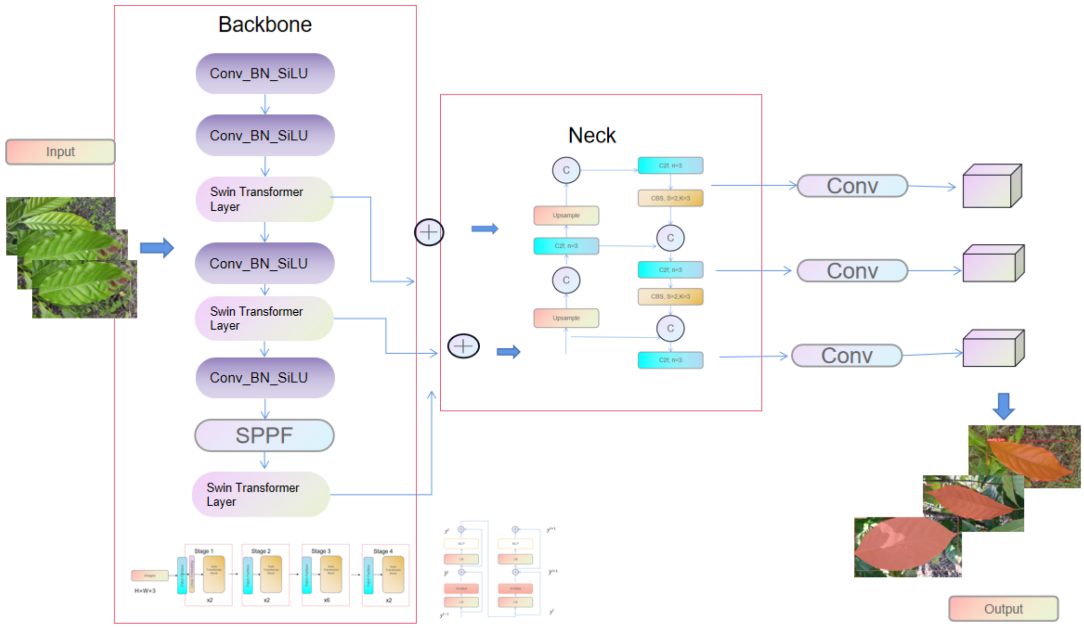

Figure 1. Architecture of the proposed ST-YOLO.

The architecture of the proposed object detection method ST-YOLO is shown in Figure 1. In traditional convolutional layer in deep neural network, it contains multiple convolution cores, whose function is to extract features from input data, then the output feature map will be input to the pooling layer for feature selection and information filtering. So the valuable information can be lost and the correlation between the local and the whole is ignored. What's more, it has low training efficiency due to much parameters. Compared with it, the Swin Transformer has better adaptability to plant data and better model effect also can learn dynamic parameters, which has stronger universality and does not depend on the data itself (Liu et al., 2021). However, its local ability to obtain information is weaker than that of YOLO algorithms. Therefore, ST-YOLO is proposed in our work to perform plant disease detection and segmentation with higher performance to prevent the plants and fruits from spreading disease and to guide biological growth. The image input first passes through the backbone network. The process begins by passing the image through the Conv layer. Subsequently, the image is sent to Swin Transformer Layer to Patch Partition for a block operation and Linear Embedding module to adjust the number of channels. The prediction results are obtained through feature extraction and down sampling in stages 1, 2, 3 and 4. Swin Transformer Blocks in each stage comprise two connected Transformer blocks based on W-MSA and shift window multi-head self attention (SW-MSA). When the image passes through each stage, its size is halved, and the channel is expanded to twice its original size. Then, the image passes through the Neck module to enrich the semantic feature and information. Finally, the model provides both detection and segmentation.

3.1. Data

The dataset used to test the model was RoCoLe (Parraga-Alava et al., 2019), a Robusta coffee leaf images dataset that can be used for both plant detection and segmentation research with segmentation annotations. In particular, the dataset contains 1560 coffee leaf images, with annotations regarding the healthy and unhealthy state and the according to the leaf area with spots to predict the severity of the disease. In addition, the dataset has six categories of labels, including healthy, red spider mite presence, rust level 1, rust level 2, rust level 3, and rust level 4, as shown in Figure 2. The rust level represents the severity scale spots of the affected leaf area.

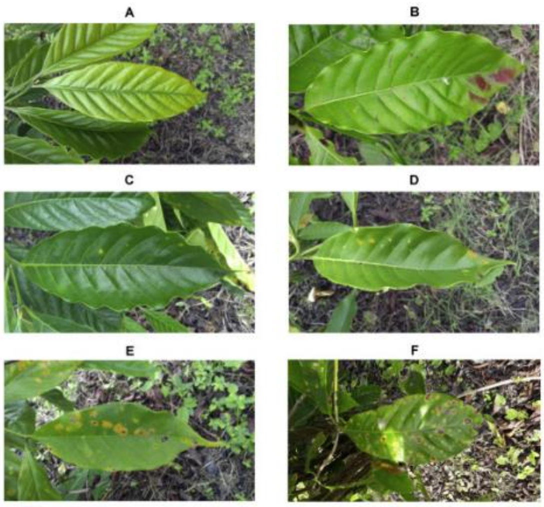

Figure 2. RoCole dataset in different states.

3.2. Backbone Network

The architecture of the backbone network is shown in Figure 3. First, the construction sequence of each Patch feature is obtained, and then the attention is hierarchically calculated. Next, the image is input to the Patch Partition for a block operation and subsequently sent to the Linear Embedding module to adjust the number of channels. Finally, the final prediction results are obtained through feature extraction and downsampling in stages 1, 2, 3, and 4. When the image passes through each stage, the size is halved, and the channel is expanded to twice its original size. In the Patch Partition, the image is divided into small chunks, differing from VIT. Then, the image undergoes 4 stages, each containing a Patch Merging and Swin Transformer Block. Patch Merging is similar to pooling; however, it cannot lose information. The Swin Transformer Blocks in each stage comprise two connected Transformer blocks based on W-MSA and SW-MSA. W-MSA computes the attention based on windows, whereas SW-MSA recalculates the attention when the windows slide.

4. Experimental Results

The experimental results are presented in this section. Figure 3 summarizes the training progress of the proposed model and presents its metrics. Moreover, the prediction results are shown in Figure 4, containing both detection and segmentation results of the Robusta coffee leaf images dataset.

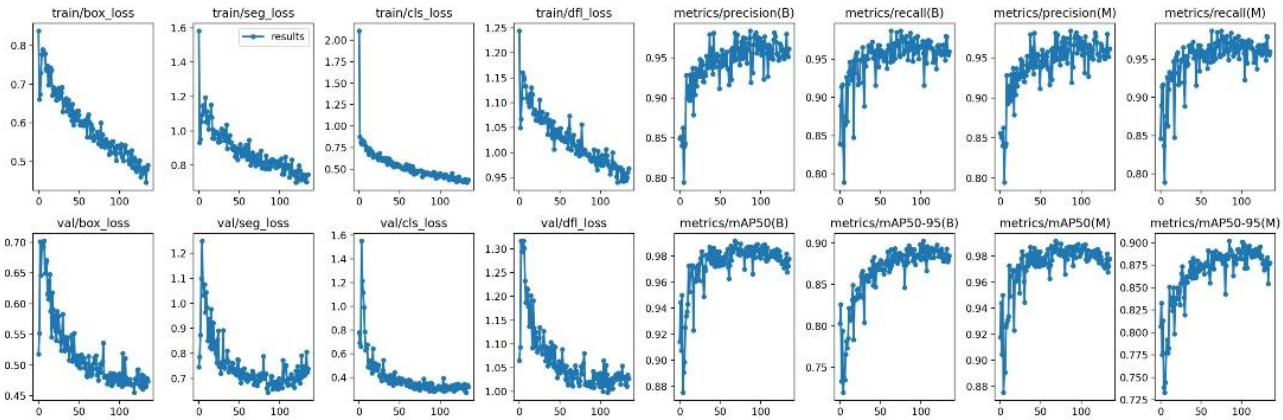

Figure 3. Metrics and loss of the ST-YOLO in training process.

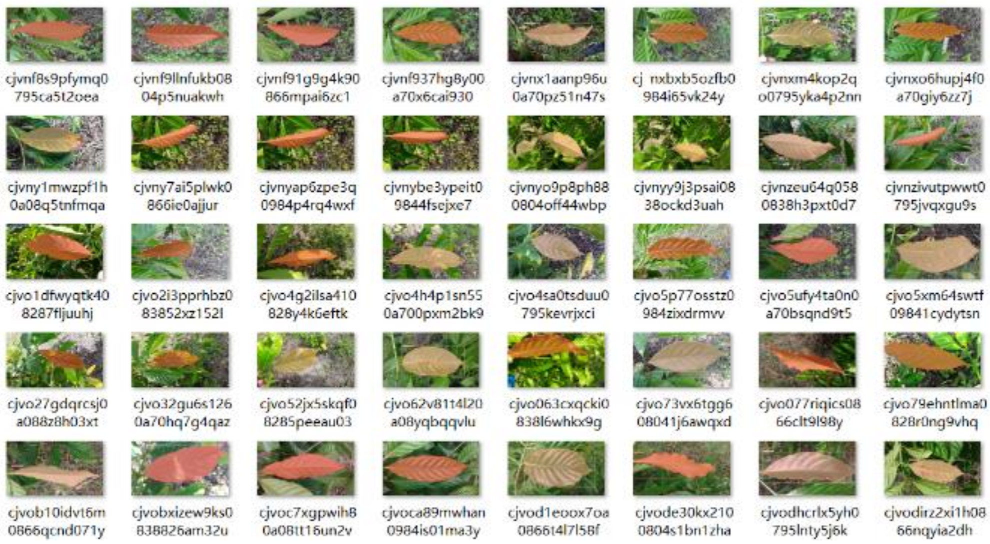

Figure 4. Detection and segmentation results of the Robusta coffee leaf images dataset.

5. Conclusion

In conclusion, this study showed that the proposed plant and fruit detection ST-YOLO method proposed in this study is meaningful and fruitful and can be applied for plant disease detection and growth guidance. Future work will focus on the lightweight of the algorithm to improve the more intelligence and convenience of the proposed network.

Acknowledgements

The authors acknowledge the helpful comments and suggestions of Prof. Zhang which improved the presentation and quality of the paper.

References

[1] Hamuda E, Ginley B M, Glavin M, et al. Automatic crop detection under field conditions using the HSV colour space and morphological operations[J]. Computers & Electronics in Agriculture, 2017, 133(Complete):97-107.

[2] Yao Q, Chen G T, Wang Z, et al. Automated detection and identification of white-backed planthoppers in paddy fields using image processing[J]. Journal of Integrative Agriculture, 2017.

[3] Islam M A, Yousuf M, Billah M M. Automatic Plant Detection Using HOG and LBP Features With SVM[J]. 2019(1).

[4] Zou Y, Zhang Y, Qin Y, et al. The Combination of HOG Features with LBP Features Applied to Tomato Disease and Pest Detection[J]. Journal of Nanjing Normal University (Engineering and Technology Edition), 2019.

[5] Thirthe G T, Chandrika J. Optimized Scale-Invariant Hog Descriptors for Tobacco Plant Detection[J]. WSEAS Transactions on Environment and Development, 2021(17-):17.

[6] Bose P, Dutta S, Goyal V, et al. Leaf Diseases Detection of Medicinal Plants Based on Support Vector Machine Classification Algorithm[J]. Journal of Pharmaceutical Research International, 2021:111-119.

[7] Lim J Y, Ahn H S, Nejati M, et al. Deep Neural Network Based Real-time Kiwi Fruit Flower Detection in an Orchard Environment[J]. 2020.

[8] Kumar M, Deolia V K. RSWECNN: Refined Semantic Weight based Convolutional Neural Network (WECNN) Segmentation for Fruit Flower Detection[J]. Journal of Advanced Research in Dynamical and Control Systems, 2019, 11(11-SPECIAL ISSUE):189-198.

[9] Gao F, Fu L, Zhang X, et al. Multi-class fruit-on-plant detection for apple in SNAP system using Faster R-CNN[J]. Computers and Electronics in Agriculture, 2020, 176:105634.

[10] Zheng Z, Xiong J, Lin H, et al. A Method of Green Citrus Detection in Natural Environments Using a Deep Convolutional Neural Network[J]. Frontiers in Plant Science, 2021, 12.

[11] Roy A M, Bose R, Bhaduri J. A fast accurate fine-grain object detection model based on YOLOv4 deep neural network[J]. arXiv e-prints, 2021.

[12] Chang H H, Kim E, Tan N, et al. Region-aggregated attention CNN for disease detection in fruit images[J]. PLoS ONE, 2021, 16(10):e0258880-.

[13] Bedi P, Gole P. Plant disease detection using hybrid model based on convolutional autoencoder and convolutional neural network[J]. Artificial Intelligence in Agriculture, 2021, 5:90-101.

[14] Boogaard F P, Rongen K, Kootstra G W. Robust node detection and tracking in fruit-vegetable crops using deep learning and multi-view imaging[J]. Biosystems Engineering, 2020, 192:117-132.

[15] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 779–788, 2016. 2

[16] Joseph Redmon and Ali Farhadi. YOLO9000: better, faster, stronger. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 7263– 7271, 2017. 2

[17] Joseph Redmon and Ali Farhadi. YOLOv3: An incremental improvement. arXiv preprint arXiv:1804.02767, 2018. 2, 4, 7, 11

[18] Liu W, Anguelov D, Erhan D, et al. SSD: Single Shot MultiBox Detector[J]. 2016.

[19] Ren S, He K, Girshick R, et al. “Faster R-CNN: Towards Real-Time Object Detection 404 with Region Proposal Networks”[J]. IEEE Transactions on Pattern Analysis and Machine 405 Intelligence, 2015, 39(6).

[20] Girshick, R.; Donahue, J.; Darrell, T.; and Malik, J. 2014. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition, 580–587.

[21] Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; and Chen, L.-C. 2018a. Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[22] Koirala A, Walsh K B, Wang Z, et al. Deep learning for real-time fruit detection and orchard fruit load estimation: benchmarking of ' MangoYOLO '[J]. Precision Agriculture, 2019.

[23] Parraga-Alava, Jorge; Cusme, Kevin; Loor, Angélica; Santander, Esneider (2019), “RoCoLe: A robusta coffee leaf images dataset”, Mendeley Data, V2, doi: 10.17632/c5yvn32dzg.2

[24] Liu Z, Lin Y, Cao Y, et al. Swin Transformer: Hierarchical Vision Transformer using Shifted Windows[J]. Computer Vision and Pattern Recognition, 2021.

[25] Trkolu M, Hanbay D. Apricot Disease Identification based on Attributes Obtained from Deep Learning Algorithms[C]// International Conference on Artificial Intelligence and Data Processing, 2019.

[26] Gupta A K, Gupta K, Jadhav J, et al. Plant Disease Prediction using Deep Learning and IoT[C]// International Conference on Computing for Sustainable Global Development. IEEE, 2019.

[27] Rashwan S A, Elteir M K. Plant leaf disease detection using deep learning on mobile devices[J]. International journal of computational vision and robotics, 2022(2):12.

[28] Haridasan A, Thomas J, Raj E D. Deep learning system for paddy plant disease detection and classification[J]. Environmental Monitoring and Assessment, 2023, 195(1):1-28.

[29] Li B, Tang J, Zhang Y, et al. Ensemble of the Deep Convolutional Network for Multiclass of Plant Disease Classification Using Leaf Images[J]. International journal of pattern recognition and artificial intelligence, 2022(4):36.

[30] Kingma D, Ba J. Adam: A Method for Stochastic Optimization[J]. Computer Science, 2014.

[31] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[J]. IEEE, 2016.

[32] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale[C]// International Conference on Learning Representations. 2021.

[33] Tan M, Le Q V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks[J]. 2019.

[34] Liu S, Qi L, Qin H, et al. Path Aggregation Network for Instance Segmentation[J]. IEEE, 2018.

[35] Du X, Lin T Y, Jin P, et al. SpineNet: Learning Scale-Permuted Backbone for Recognition and Localization[J]. 2019.