Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Microscopic image recognition method of stomata in living leaves based on improved YOLO-X

Tianhong Dai
Northeast Forestry University
Shijie Cong
Northeast Forestry University
Jingzong Zhang
Northeast Forestry University
Kexin Li (dillonlkx@163.com)
Wuxi Vocational College of Science and Technology

Research Article

Keywords: Living plant stomata, Target detection, Deep learning, YOLO-X

Posted Date: April 1st, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1496525/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Abstract

Stomata are the main medium of water exchange in plants, regulating gas exchange and responsible for the processes of photosynthesis and transpiration. Stomata are surrounded by guard cells and the transpiration rate is controlled by opening and closing stomata. Stomatal state (open and close) plays an important role in describing the health of plants. In addition, counting the number of stomata is of great significance for scientists to study the number of opening and closeing stomata and to measure their density and distribution on the leaf surface through different sampling techniques. Although some techniques for calculating the number of stomata have been proposed, these methods are used to produce samples in isolation and then to identify and classify the states in the sample leaves. We improved YOLO-X and then implemented a transfer learning method to count the number of stomata and identify the stomatal opening and closing status of live black poplar leaves. In the end, the average accuracy and recall of the method were 98.3% and 95.9%, which helped researchers to obtain accurate information on leaf stomatal opening and closing status in an efficient and simple way.

1. Introduction

Stomata are epithelial openings in plant leaves and stems, which are structures unique to the plant epidermis. Stomata are generally regulated by movement through the opening and closing action of guard cells and play an important role in plant physiology. Stomata are the main outlet for water vapor to be discharged from the body to the outside during transpiration, and are the channels for photosynthesis and respiration to exchange gases with the outside world. The opening and closing of stomata in plants is controlled by guard cells. Each opening and closing of stomata affects the transpiration, photosynthesis and respiration of plants. When the stomata are opened, the transpiration of the plant becomes stronger, which in turn protects the plant from the sun. It also expels carbon dioxide, which facilitates the plant's carbon assimilation and photosynthesis.

With the continuous development of computer vision, target detection, segmentation and classification are gradually applied to the study of stomata of plants. Hiranya Jayakody et al. proposed a method to construct a cascaded object detectors using the histogram of oriented gradients feature to identify multiple stomata in large microscopic images. After testing on 1267 stomatal images of grapevines, the segmentation and skeletonization method was able to correctly identify 86.27% of stomatal opening areas[1]. A few years later, Alexandre Hild Aono et al. combined SVM with VGG16 and applied the dataset proposed by Hiranya Jayakody et al. for testing and achieved an accuracy of 97.1% for grapevine stomatal recognition[2]. It can be seen that using deep learning for stomatal recognition has better results compared to cascaded object detectors.

Karl C. Fetter et al. described a system for automatic calculation of the number of stomata. After treating the leaves with nail polish, they used a modified AlexNet to identify stomata in a variety of different microscopy images and achieved 98.1% accuracy on the ginkgo scanning electron microscopy entropy micrograph[3]. A year later, Luke Millstead et al. developed an optimal procedure for sample collection

and imaging by studying the suitability of using an automated microscope slide scanner for imaging nail polish blots. The stomata in the input images were detected by AlexNet and the average precision and recall of 79% and 85% were obtained for four plants: grape, apricot, sweet orange and periwinkle[4]. Meanwhile, Sofie Meeus et al. obtained 431 micrographs taken from specimen sheets of 19 herbarium specimens made by the nail polish method and achieved an average accuracy of 94% in calculating the number of stomata on images of unseen species used for training using the VGG19 architecture[5]. Angela Casado-García et al. proposed an automatic stomatal detection model based on YOLO. In the identification of a mixed dataset of three plants, soybean, navy bean and barley, test time enhancement was applied by using vertical flip, horizontal flip and color transformation to deal with the variability of stomatal shape, as well as the different textures and qualities of the images. To handle the variation in stomatal size, and using progressive resizing, the images were resized to three different sizes, resulting in accuracy and return rates of 92% and 93%[6]. From the work of the above authors, it can be seen that in the study of stomatal recognition of plant leaves, the use of isolated leaves treated with nail polish and then stomatal recognition using a network of deep learning can give good results in the recognition of leaf stomata.

With the continuous subdivision and development of deep learning networks, better target detection networks have been applied. Lucas Costa et al. used Mask-RCNN algorithm to identify stomata in citrus leaf epidermis microscopic images with 99% accuracy for citrus trees grafted on different rootstocks after nail polish coating of citrus leaf epidermis. They also identified stomata in the microscopic images of citrus leaf epidermis in the opening and closing states with 99.2% and 97.2% accuracy and 96.7% and 90.0% recall for identifying opening and closing stomata[7]. Abdul Razzaq et al. developed an automatic stomatal state identification and counting system for quinoa leaf images. Stomatal recognition was performed using SSD-MobileNet V2 and stomatal opening and closing classification was performed using SVM. Finally, the stomatal recognition and classification accuracy of the system was 98.6% and 97%[8]. Therefore, using a network with better target recognition for migration learning can lead to better results of the work.

Based on the above work, we found that researchers are using different models to identify the stomata of the treated leaves after nail polish coating treatment of the studied plant leaves, and the use of better target recognition networks is more helpful for the work. However, obtaining isolated plant leaves and applying nail polish smear treatment to the leaves makes the work more tedious in certain cases. With the continuous development of target recognition networks, the robustness of target feature extraction in target recognition networks is getting better and better. YOLO-X, a target recognition network proposed by Zheng Ge et al. in 2021, possessed excellent performance to accurately identify targets in relatively complex backgrounds. Therefore, we investigated the use of live black poplar leaves for direct stomatal acquisition and the identification of the collected stomatal data using the modified YOLO-X. The aim was to investigate whether good results could be obtained by using live plant leaves directly for stomatal data sets and using YOLO-X for stomatal identification.

This paper is organized as follows: Section I provides a general overview of the method, including the method of in vivo plant leaf stomatal collection, and the improved YOLO-X model; Section II presents the experimental results; Section III contains the discussion; and Section IV is conclude.

2. Materials And Methods

This section focuses on the work and methods needed to identify the stomata of the leaves of living plants. This section is divided into three parts: image acquisition and pre-processing, target detection network and evaluation, and the general block diagram is shown in Fig. 1. Image acquisition and pre-processing were mainly performed using a digital microscope of Keyence VHX-2000 to acquire images of the subepidermal stomata of living black poplar leaves. Target recognition network is based on YOLO-X, and the parameters and structure are optimized for the characteristics of leaf stomata. Evaluation mainly uses the evaluation indicators commonly used in target detection, namely accuracy, recall, and mean average precision.

2.1. Image acquisition and preprocessing

One-year-old black poplar trees were selected for live leaf subepidermal stomatal image acquisition using a Keyence VHX-2000 digital microscope. Using the real-time depth synthesis technology of the VHX-2000, a full-field clear image of the stomata of the living poplar leaves was acquired. The image resolution was 1600×1200, and stomatal images were acquired at this resolution with a 1000X lens. Stomatal images were also acquired from each location, including the upper, middle and lower leaves of different branches, as well as the upper, middle and base of a single leaf. An example of the collected images is shown in Fig. 2, and the collected stomatal images are labeled[9].

Finally, we obtained a total of 3160 data of living plant stomata. Based on the findings obtained by Sharada Prasanna Mohanty, we divided the training set, validation set and test set in the ratio of 7:1:2[10], as shown in Table 1.

Table 1
Dataset details.

Classes	Number of Training Data	Number of Testing Data
Open Stoma	1241	84
Close Stoma	1919	574
Total	3160	658

We manually annotated the dataset using Labelimage software. To obtain a larger dataset, we performed data enhancement on the obtained dataset at a scale of 1:10, using image processing methods such as upside down, mirror flip, Gaussian blur, brighten, panning and scaling, as shown in Fig. 3.

2.2. Target detection network

YOLO-X network as a new target recognition network, we optimized and improved YOLO-X on its basic structure, and the network structure is shown in Fig. 4. The improved YOLO-X network is divided into four parts, which are CSPDarknet, FPN, Yolo Head and EX-NMS.

CSPDarknet can be called the backbone feature extraction network of the YOLO-X network. The input images are first feature extracted in CSPDarknet, and the extracted features are used as feature layers. We acquire the features of the last three feature layers using FPN for feature enhancement to combine feature information at different scales. In the FPN, the same structure of Panet used in YOLO-V4 is used. The network up-samples the features to achieve feature fusion and down-samples the features afterwards to achieve feature fusion, and finally three feature-enhanced feature layers are obtained. Yolo Head is the classifier and regressor of YOLO-X network, and the structure is shown in Fig. 5. It determines the feature points in three aspects: determining the prediction frame, judging the presence of defects in the prediction frame, and identifying the types of defects in the prediction frame. Compared with the previous version of YOLO, which used a decoupling head where classification and regression were implemented in a 1X1 convolution, the Yolo Head was divided into two parts to implement classification and regression separately, and the results were integrated together for final prediction.

Although YOLO-X performs Non-Maximum Suppression(NMS) in the process of filtering and decoding the predicted results, it still appears that two states are identified out of one stomata in the output results, as shown in Fig. 6.

Therefore, we make a two-part adjustment to YOLO-X. On the one hand, in the decoding part of the prediction result, we use the Clou with better effect instead of Iou when calculating the IOU loss, see Eq. where $\rho^2\left(b, b^{gt}\right)$ is the Euclidean distance between the center point of the predicted frame and the real frame, c is the diagonal distance that contains both the smallest rectangle of the predicted frame and the real frame, α is the weight function, and v is used to measure the similarity of the aspect ratio, see Eq. On the other hand, since the stomata are all kept at a certain distance from each other and the prediction frames do not cross, we perform a separate NMS of the YOLO-X output to ensure that each stomata corresponds to only one prediction frame.

$$CIOU = IOU - \frac{\rho^2\left(b, b^{gt}\right)}{c^2} - \alpha v$$

$$\alpha = \frac{v}{1 - \mathrm{IOU} + v}$$

$$v = \frac{4}{\pi^2}\left(\arctan\frac{w^{gt}}{h^{gt}} - \arctan\frac{w}{h}\right)^2$$

The hyperparameters we set for the YOLO-X network are shown in Table 2.

Table 2
Hyper-parameters of the experiments.

Hyper-Parameters	Value
Optimization algorithm	SGD
learning rate	1.0×10 − 3
Epochs	300
Batch size	4

2.3. Evaluation

In this section, we used three evaluation metrics, including stomatal number recognition accuracy, stomatal opening and closing recognition accuracy, and mean average precision (MAP). These metrics are calculated from the base metrics, which are True Positive(TP), False Positive(FP), True Negative(TN), and False Negative(FN). TP is predicting positive classes as positive classes, FP is predicting negative classes as positive classes, TN is predicting negative classes as negative classes, and FN is predicts positive class as negative class.

Accuracy: The probability that the prediction is correct in the total sample, that is, the accuracy of stomatal opening and closing recognition.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Recall: It is the ratio of the number of relevant documents retrieved to the number of all relevant documents in the document library. Since this experiment only identifies stomata, the accuracy rate of stomata identification is the recall rate without distinguishing stomata status.

$$recall = \frac{TP}{TP + FN}$$

Precision: describes the percentage of correct predictions when the prediction result is in the positive category.

$$precision = \frac{TP}{TP + FP}$$

Mean Average Precision (MAP): a performance metric for this class of algorithms that predict target locations as well as categories.

3. Results

Because of the large amount of computation required when training CNN networks, all experiments were performed on a deep learning workstation with the workstation configuration shown in Table 3. The training process was performed in the Pytorch environment with the programming language python 3.6.

Table 3
Machine specifications.

Hardware and Software	Characteristics
Memory	64Gb
Processor	Intel Xeon E5-2650v4 CPU @ 2.20 GHz
Graphics	Titan Xp 12 Gb

In this experiment, we conducted experiments using FasterRCNN, YOLO-v4, YOLO-X, and improved YOLO-X. After five-fold cross-validation, the experimental results of each model are shown in Table 4, including accuracy, recall, and MAP.

Table 4
Performance measures (%) for every model.

Performance Measures	Accuracy	Recall	MAP
FasterRCNN	90.1	93.4	85.2
YOLO-v4	98.5	93.9	89.9
YOLO-X	98.6	95.7	91.5
Improved YOLO-X	98.7	95.9	92.0

The training curve of YOLO-X in the experiment is shown in Fig. 7. As can be seen from the figure, we first froze some of the network parameters, unfroze the parameters after the network reached convergence, and finally made the network converge again.

From the results, it can be seen that YOLO has more advantages in identifying stomata. The accuracy of the three YOLO networks is basically the same, but YOLO-X has a more obvious advantage in recall and MAP. Compared with YOLO-v4, YOLO-X can achieve an improvement of about 1.8% in recall, while MAP can achieve an improvement of about 1.6%. With the improvements made to YOLO-X, the improved YOLO-X achieved 0.1%, 0.2% and 0.5% improvement in accuracy, recall and MAP.

Moreover, Table 5 gives the confusion matrix of the improved YOLO-X model based on the performance metric with the best results. The performance of the network can be evaluated visually based on the results. Rows are associated with the output class, while columns are associated with the true class.

Diagonal units are associated with correctly classified observations, and non-diagonal units correspond to incorrectly classified observations.

Table 5
Confusion matrix based on YOLO-X model.

	CloseStoma	OpenStoma
CloseStoma	546	18
OpenStoma	9	77

In addition, we have calculated APs for both categories, as shown in Figure 8.

4. Discussion

With the development of deep learning, computer vision is applied to various fields. Target detection offers a possibility of extended research and applications for studying the stomatal state of plants. In the study, we used the digital microscope of Keyence VHX-2000 for the acquisition of the subepidermal stomatal image dataset of living black poplar leaves, and used the real-time depth synthesis technology of VHX-2000 to obtain a full-field clear image of the subepidermal stomata of living black poplar leaves. and finally obtained 3818 stomatal data. We divided the dataset into training set, validation set and test set according to the ratio of 7:1:2, and used the training set and validation set to train the improved YOLO-X.

In the identification of opening and closing stomata, it can be seen from the results that although we used data collection from living leaves of black poplar, we still obtained 98.3% accuracy and 95.9% recall in the identification of stomata, which indicates that data collection using living leaves on black poplar can achieve good results, which helps researchers to efficiently and simply obtain accurate information on leaf stomatal status.

In addition, YOLO-X, as one of several networks with the best results in the field of target recognition, also shows from the results that YOLO-X obtains better results in terms of accuracy, recall and MAP compared with FasterRCNN and YOLO-v4 because its structure is improved based on YOLO-v5. After we optimized and improved YOLO-X, the results of YOLO-X were further improved. In terms of stomatal recognition accuracy, the improved YOLO-X has 8.6% and 0.2% improvement over FasterRCNN and YOLO-V4. In the recognition of opening and closing stomata, we use Clou instead of Iou to calculate the IOU loss part for the problem that YOLO-X identifies two states with one stomata in the process of outputting results, and perform a separate NMS on the results outputted by YOLO-X, so that it has an accuracy of opening and closing stomata compared with FasterRCNN and YOLO-V4 2.5% and 2% improvement. Moreover, the improved YOLO-X also improves 6.8% and 2.1% over FasterRCNN and YOLO-V4 in MAP. This indicates

that it is possible to achieve the experimental requirements using a better performing target detection network.

However, it can be seen from Table 5 that YOLO-X misjudged a small percentage of both opening and closing stomata in the identification of stomata state. This is because stomata with small opening-closing degrees provide little photosynthesis and transpiration in practice, while stomata with large opening-closing degrees are the decisive factors in photosynthesis and transpiration. Therefore, when collecting stomatal data, there existed a small percentage of stomata whose opening-closing degrees were between opening and closing stomata, which would be misclassified by YOLO-X, as shown in Fig. 9. Thus, we try to avoid stomata with opening-closing degrees between opening and closing stomata as much as possible in the process of collecting data sets in the future.

5. Conclusions

In this paper, we propose a deep learning-based stomatal recognition method for microscopic images of living leaves. We use Clou instead of Iou to calculate the IOU loss part based on YOLO-X, and perform a separate NMS of the YOLO-X output. It is clear from the experimental results that although we did not perform picking and nail polish application treatment operations on the leaves. However, we were able to meet the requirements of the experiment with 98.7% accuracy and 95.9% recall when we used the improved YOLO-X for state identification against the stomata of living black poplar leaves. This indicates that stomatal identification using living black poplar leaves is feasible and has important implications for the study of stomatal states in living leaves of other plants.

Declarations

Authors' contribution

The first draft of the manuscript was written by Shijie Cong and all authors critically revised it for important intellectual contents. Conceptualization and design of the study: Kexin Li, Shijie Cong; Formal analysis and investigation: Shijie Cong, Jingzong Zhang; Writing -original draft preparation: Shijie Cong; Writing -review and editing: Tianhong Dai, Kexin Li. All authors read and approved the final manuscript.

Funding

Fundamental Research Funds for the Central Universities(2572019CP17); Natural Science Foundation of Heilongjiang Province of China(C201414).

Conflict of interest

The authors declared that they have no conflicts of interest to this work. We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted

References

1. H. Jayakody, S. Liu, M. Whitty, P. Petrie, Microscope image based fully automated stomata detection and pore measurement method for grapevines, Plant Methods, 13 (2017).
2. A.H. Aono, J.S. Nagai, G.D.M. Dickel, R.C. Marinho, P. de Oliveira, J.P. Papa, F.A. Faria, A stomata classification and detection system in microscope images of maize cultivars, Plos One, 16 (2021).
3. K.C. Fetter, S. Eberhardt, R.S. Barclay, S. Wing, S.R. Keller, StomataCounter: a neural network for automatic stomata identification and counting, New Phytologist, 223 (2019) 1671–1681.
4. L. Millstead, H. Jayakody, H. Patel, V. Kaura, P.R. Petrie, F. Tomasetig, M. Whitty, Accelerating Automated Stomata Analysis Through Simplified Sample Collection and Imaging Techniques, Frontiers In Plant Science, 11 (2020).
5. S. Meeus, J. Van den Bulcke, F. Wyffels, From leaf to label: A robust automated workflow for stomata detection, Ecology And Evolution, 10 (2020) 9178–9191.
6. A. Casado-Garcia, A. del-Canto, A. Sanz-Saez, U. Perez-Lopez, A. Bilbao-Kareaga, F.B. Fritschi, J. Miranda-Apodaca, A. Munoz-Rueda, A. Sillero-Martinez, A. Yoldi-Achalandabaso, M. Lacuesta, J. Heras, LabelStoma: A tool for stomata detection based on the YOLO algorithm, Computers And Electronics In Agriculture, 178 (2020).
7. L. Costa, L. Archer, Y. Ampatzidis, L. Casteluci, G.A.P. Caurin, U. Albrecht, Determining leaf stomatal properties in citrus trees utilizing machine vision and artificial intelligence, Precision Agriculture, 22 (2021) 1107–1119.
8. A. Razzaq, S. Shahid, M. Akram, M. Ashraf, S. Iqbal, A. Hussain, M.A. Zia, S. Qadri, N. Saher, F. Shahzad, A.N. Shah, A.U. Rehman, S.E. Jacobsen, Stomatal State Identification and Classification in Quinoa Microscopic Imprints through Deep Learning, Complexity, 2021 (2021).
9. K.X. Li, J.P. Huang, W.L. Song, J.T. Wang, S. Lv, X.W. Wang, Automatic segmentation and measurement methods of living stomata of plants based on the CV model, Plant Methods, 15 (2019).
10. S.P. Mohanty, D.P. Hughes, M. Salathe, Using Deep Learning for Image-Based Plant Disease Detection, Frontiers In Plant Science, 7 (2016).
11. S. Ren, K. He, R. Girshick, J. Sun, Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, IEEE transactions on pattern analysis and machine intelligence, 39 (2017) 1137–1149.

Figures

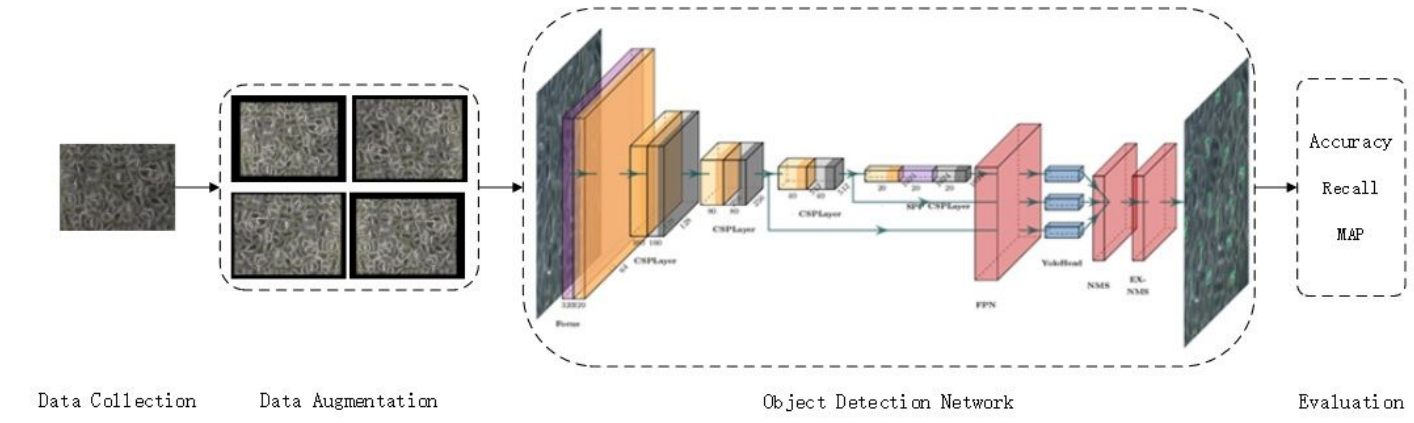


Figure 1

General Block Diagram.

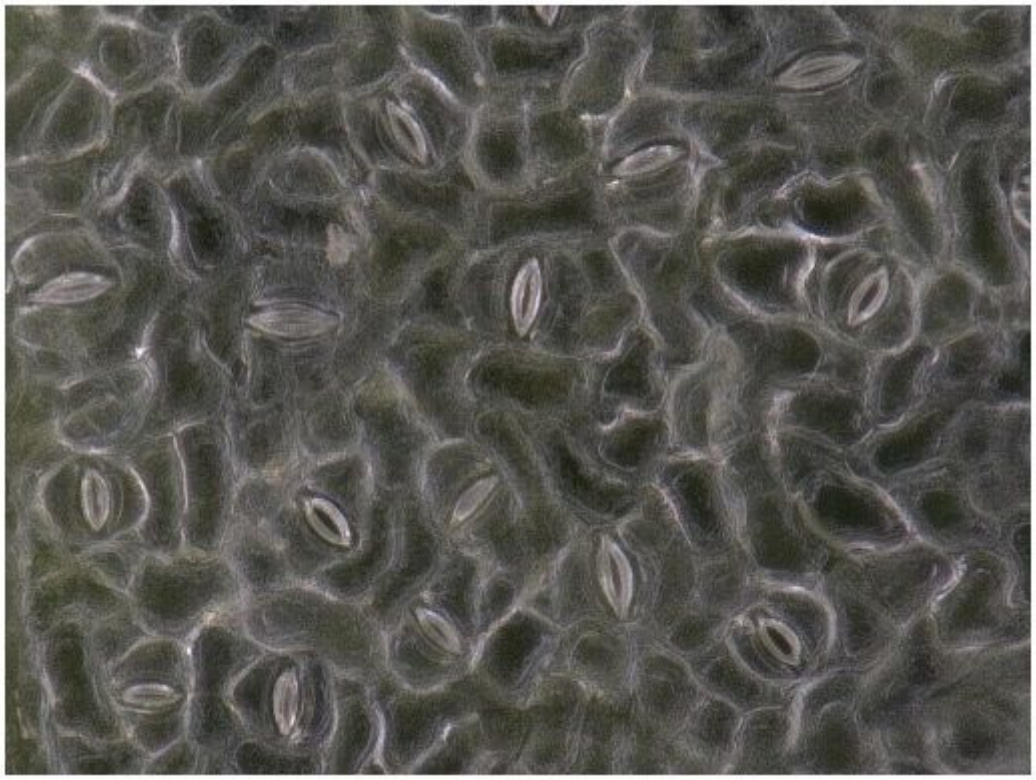


Figure 2

Microscopic images of stomata of black poplar leaves.

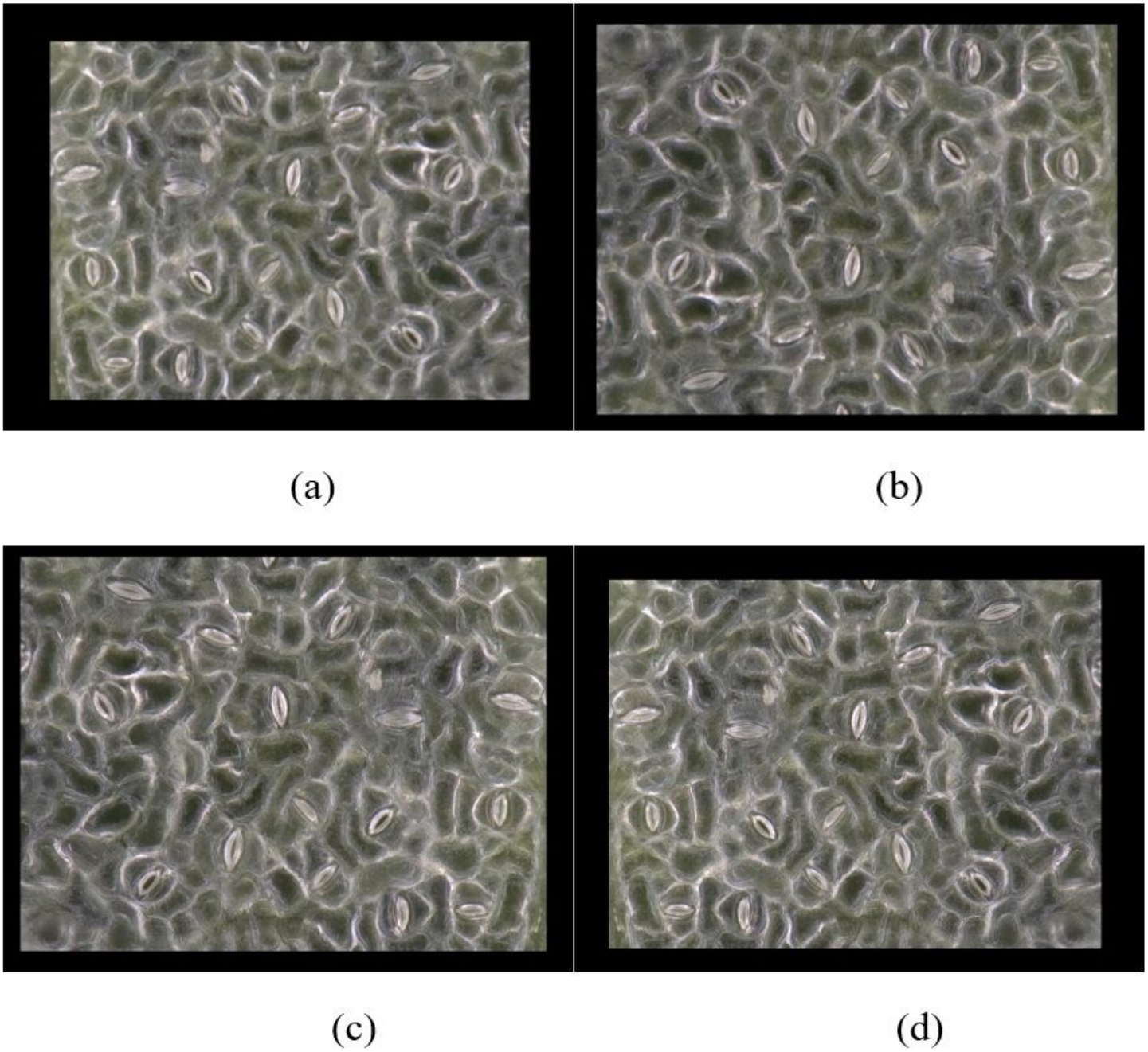


Figure 3

(a) Scaling, panning, brighten. (b) Scaling, panning, brighten, upside down, mirror flip, Gaussian blur. (c) Scaling, panning, brighten, mirror flip. (d) Scaling, panning, brighten, Gaussian blur.

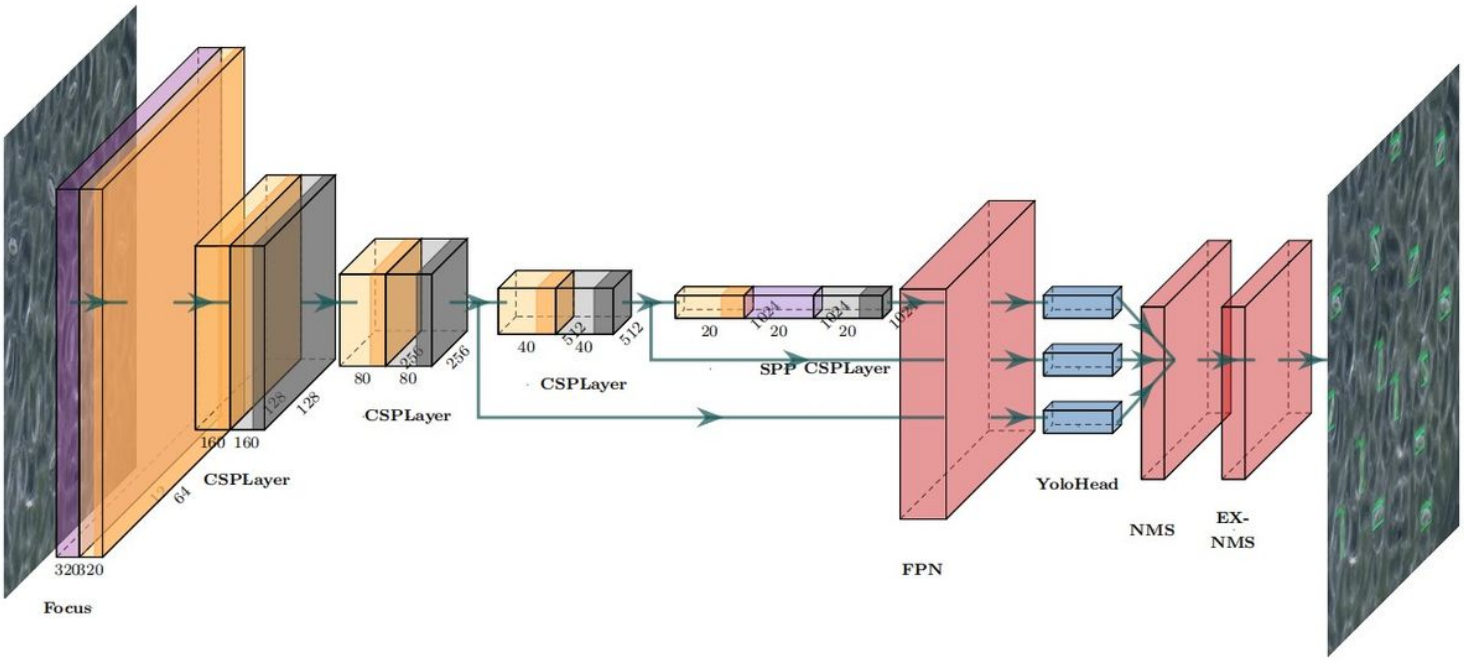


Figure 4

Improved YOLO-X network structure.

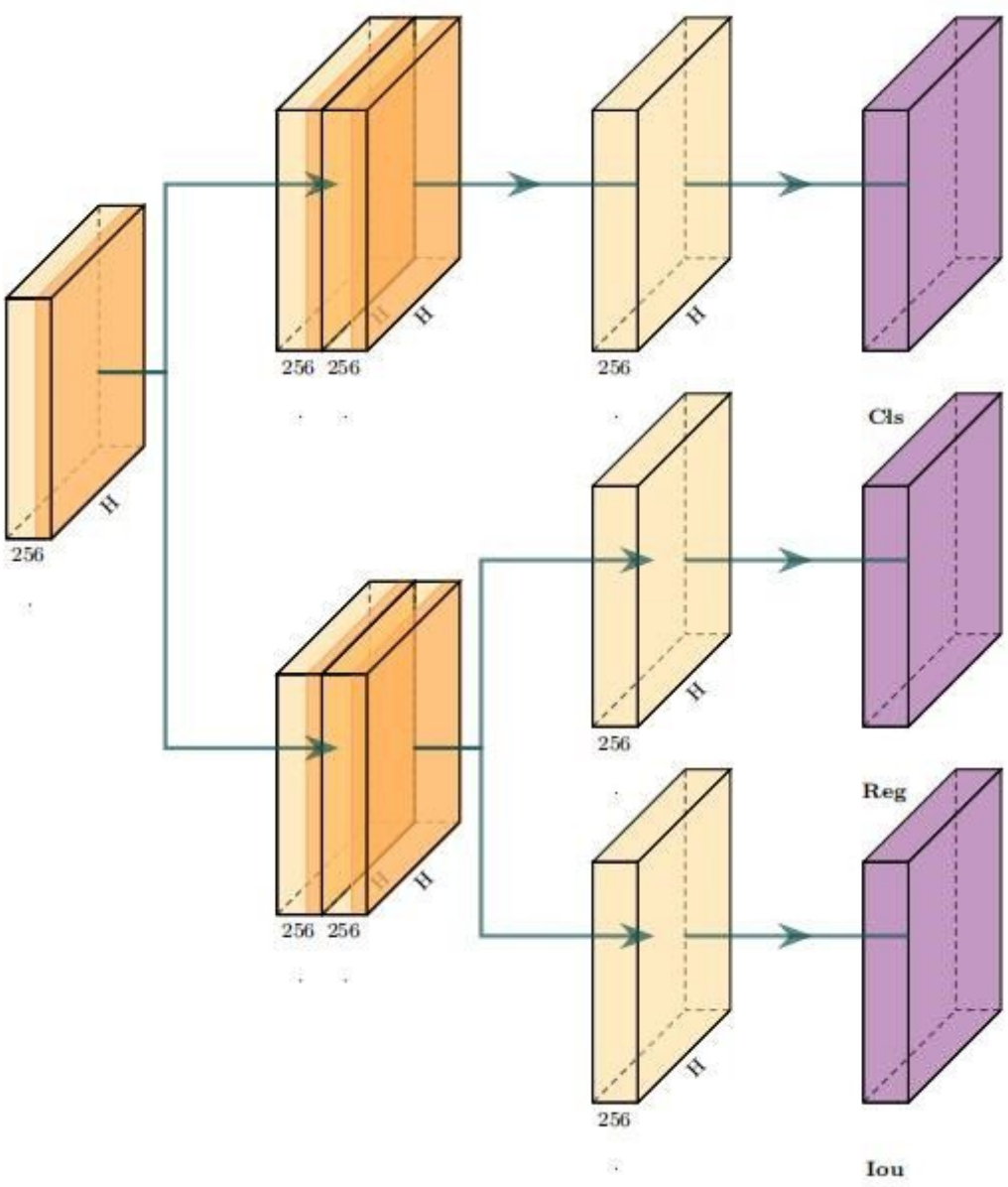


Figure 5

YOLO Head structure diagram.

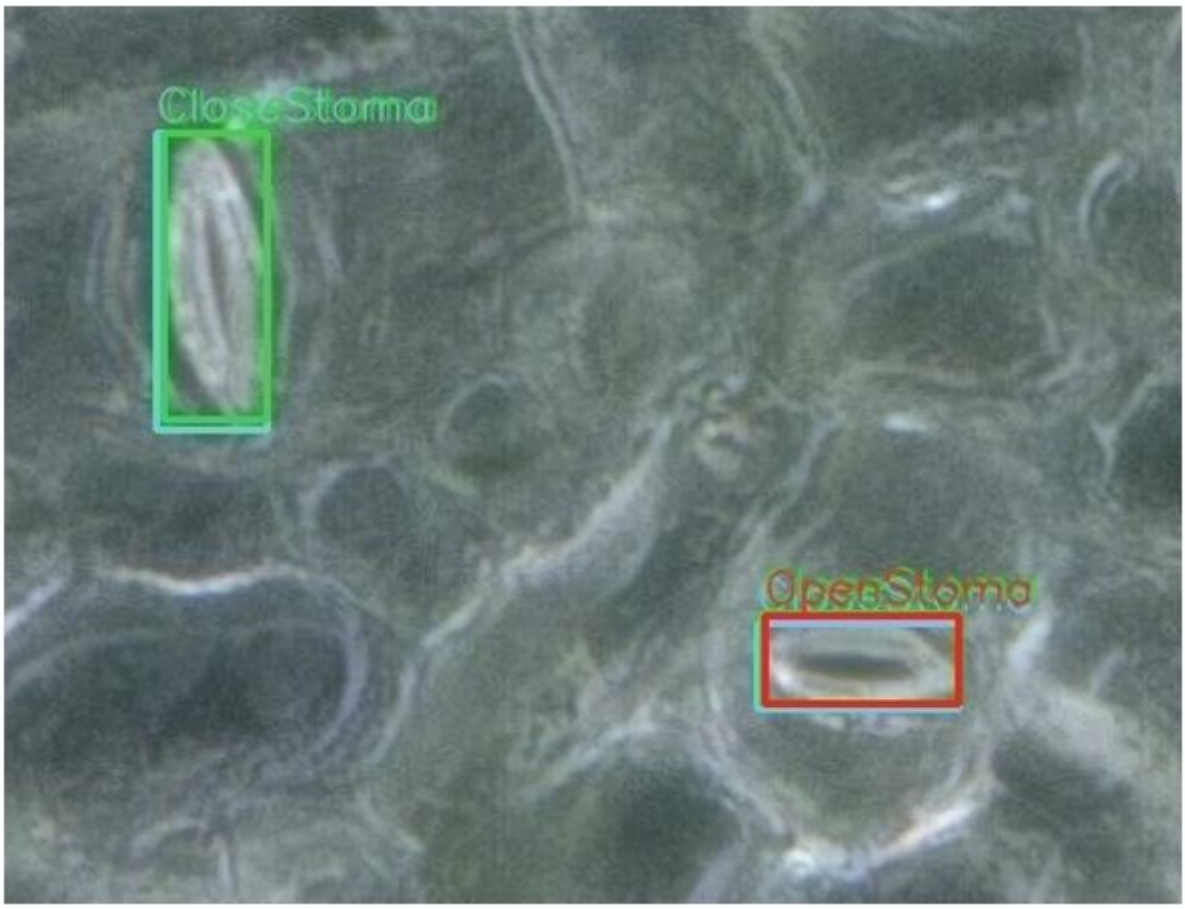


Figure 6

Results of YOLO-X output without EX-NMS.

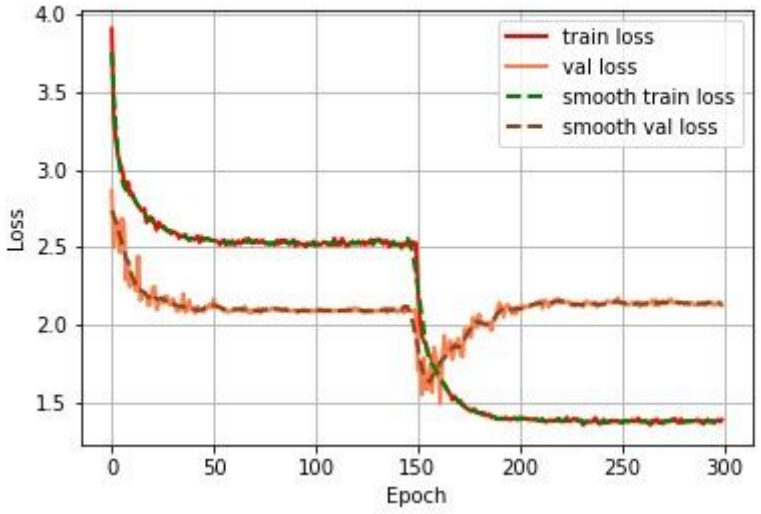


Figure 7

Improved YOLO-X training curve.

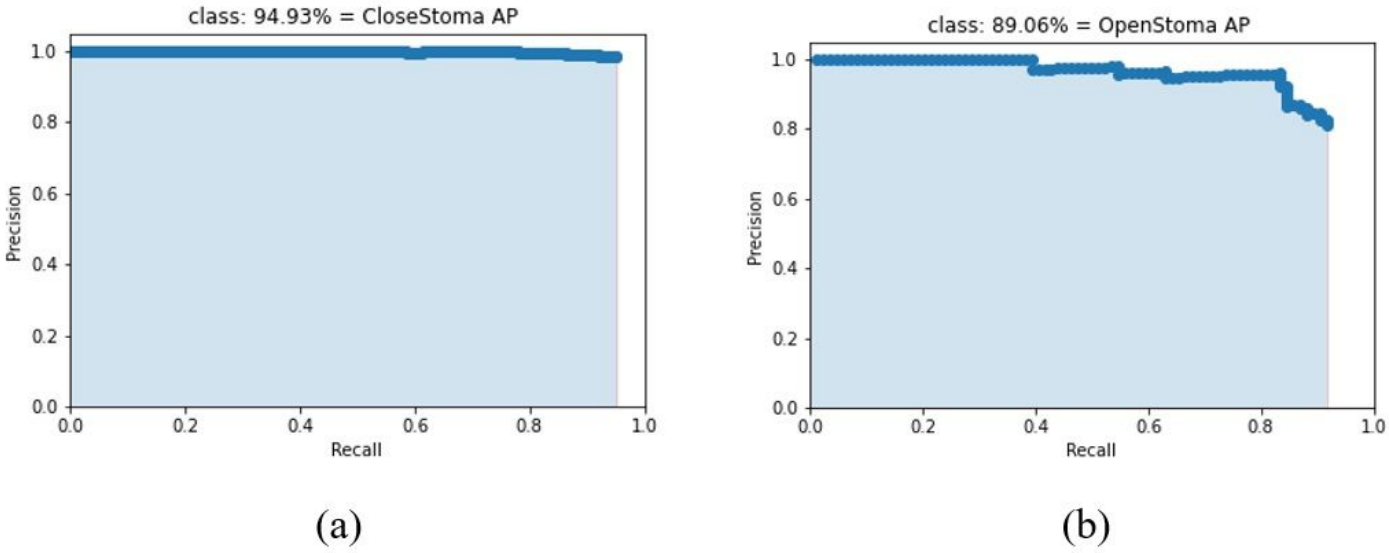


Figure 8

(a) Closed stomatal AP curve. (b) Opened stomatal AP curve.

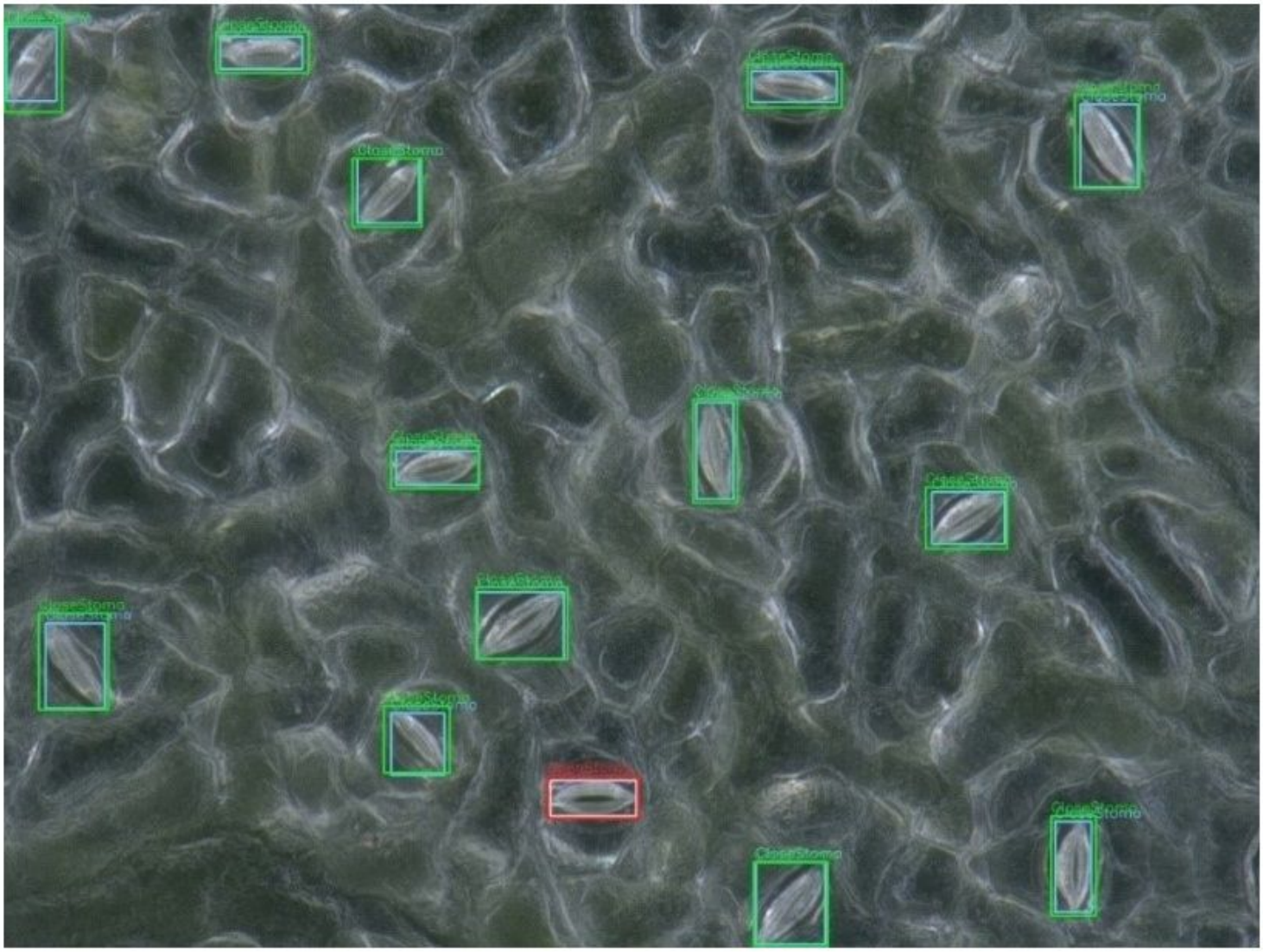


Figure 9

Output results of improved YOLO-X.