WHITE TEA BUD DETECTION BASED ON DEEP LEARNING RESEARCH /
基于深度学习的白茶嫩芽检测研究

Weiqiang PI[1], Rongyang WANG*[1], Qinliang SUN[1], Yingjie WANG[1], Bo LU[1], Guanyu LIU[1], Kaiqiang JIN[1]

[1] Huzhou Vocational and Technical College, College of Mechatronics and Automotive Engineering, Huzhou / China
Tel:15849125756; E-mail: piweiqiang20171@163.com
DOI: https://doi.org/10.35633/inmateh-70-45

Keywords: *White tea, Image detection, YOLOv5, Deep learning, Algorithm*

ABSTRACT
The quality of white tea buds is the basis of the quality of finished tea, and sorting white tea buds is a laborious, time-consuming, and key process in the tea-making process. For intelligent detection of white tea buds, this study established the YOLOv5+BiFPN model based on YOLOv5 by adding a Bidirectional Feature Pyramid Network (BiFPN) structure to the neck part. By comparing the YOLOv5 and YOLOv3 through the ablation experiment, it was found that the YOLOv5+BiFPN model could extract the fine features of white tea buds more effectively, and the detection average precision for one bud and one leaf was 98.7% and mAP@0.5 was 96.85%. This study provides a method and means for white tea bud detection based on deep learning image detection, and provides an efficient, accurate, and intelligent bud detection model for high-quality white tea sorting.

摘要
白茶嫩芽品质是成茶品质的基础，白茶嫩芽分选是制茶工序中费工、费时且关键的工序。为了实现白茶嫩芽智能化检测任务，本文通过建立白茶鲜叶数据集，基于 YOLOv5 模型，在 Neck 加入双向特征金字塔网络（BiFPN）结构，得到了 YOLOv5+BiFPN 模型。通过烧蚀实验对比 YOLOv5 模型和 YOLOv3 模型，发现 YOLOv5+BiFPN 模型可以更有效的提取白茶嫩芽中的细小特征，对一芽一叶的检测精度达 98.7%，mAP@0.5 达 96.85%。本研究为白茶嫩芽检测提供了一种基于深度学习图像检测的方法与手段，为名优白茶分选提供了一种高效、准确、智能化的嫩芽检测模型。

INTRODUCTION
In terms of the size of its tea plantations and annual production, China ranks first in the world and is one of the largest producers and consumers of tea. White Leaf No. 1, commonly known as Anji White Tea, is a rare kind of Chinese tea tree that only produces white tea with albino leaves. It was discovered in Anji, Zhejiang, China, in the 1980s. The Anji White Tea has a huge consumer market because of its high amino acid content, the finished tea with white leaves and green veins, fragrant and mellow taste. Anji white tea has a planting area of more than two hundred and ten thousand acres and an annual output value of more than 3.4 billion. With a national planting area of about two hundred and seventy thousand hectares, it has been widely planted by Guizhou, Jiangxi, Hubei, Sichuan, and other provinces. It also has a high nutritional and economic value *(Wang et al., 2023)*.

The quality of white tea buds is the basis of the quality of the tea, and white tea bud sorting is an important part of the tea production process. The white tea production process is mainly divided into three steps: picking, withering, and drying. Among them, the tea picking and sorting process accounts for more than 60% of the total workload of tea making, and the cost is more than 50%. The quality of tea leaves is determined by the experience and attitude of the tea picker, which is uneven and requires manual secondary sorting. With the development of agricultural machinery and the increase in labor costs, mechanized tea harvesting technology and equipment are now becoming more mature and gradually applied *(Yi et al., 2020)*.

However, mechanized tea picking can mix a large number of different types of tea leaves and impurities, and the quality varies, so it cannot directly enter the next step of the tea production process, and still needs tea fresh leaves sorting. Therefore, there is an urgent need for an efficient, intelligent, accurate, and low-cost tea fresh leaf sorting technology and method.

With the gradual improvement of the quality of life and the upgrading of consumption, consumers' requirements for the quality of Anji white tea have become higher and higher, and there is a huge price gap between the price of famous Anji white tea and the price of bulk Anji white tea. In the past two decades, scholars developed the sorting process according to the physical characteristics of tea leaves, such as rolling sieve *(Wang et al., 2016)*, round sieve *(Zhao et al., 2014)*, wind sieve *(Zhang et al., 2014)*, etc., which can separate the debris of tea fresh leaves (incomplete leaves, leaf stalks, etc.), but it is difficult to carry out the precise bud grade division of white tea buds so that one bud and one leaf, a bud and two leaves, a bud and three leaves and other types of tea fresh leaves are mixed together, reducing the quality of tea, so that only artificial sorting methods can be used to produce famous tea or directly producing low-priced bulk tea.

In recent years, with the development of machine vision and deep learning technology, precision agriculture is gradually moving towards the era of intelligence. Machine vision has been widely used in the fields of tea bud detection, quality identification, category recognition, and pest control due to its non-destructive and highly automated characteristics. Traditional machine vision methods were first applied in tea classification research, such as applying a support vector machine to Wuyi rock tea fresh leaf classification *(Lin et al., 2019)*; then using a support vector machine to extract texture features combined with distance matrix further improved tea fresh leaf classification accuracy *(Mao et al., 2020)*.

With the outstanding performance of deep learning models in image detection, the YOLOv3 model was gradually applied in tea detection, and scholars used the YOLOv3 model to achieve better detection results for single bud targets (one bud and one leaf) in tea fresh leaves *(Yang et al., 2019; Li et al., 2021)*, but due to the Feature in the YOLOv3 model YOLOv5 is rapidly coming into the view of scholars because of its excellent detection speed and accuracy for small targets, and is currently applied to the detection on Cucumber *(Zhao et al., 2022)*, Green asparagus *(Hong et al., 2023)*, green pepper *(Nan et al., 2023)*, and other crops, but few studies have been reported on the detection of white tea buds.

MATERIALS AND METHODS

Dataset Production

White tea No. 1 producing region in Anji County, Zhejiang Province (30°28′N, 119°42′E), is where the fresh white tea leaves were gathered. The fresh white tea leaves were returned to the lab in sealed preservation bags and placed on the experimental table the same day they were harvested. A dataset of fresh white tea leaves was created by manually selecting 1200 photos of the test bed's fresh white tea leaves taken with a Canon 6D II camera.

The white tea fresh leaf dataset has four categories: one bud and one leaf, one bud and two leaves, one bud and three leaves, and the other (incomplete leaves and other buds). Table 1 lists the number of samples for each of the four categories. The image size and the category labeling of white tea fresh leaves using labeling, as shown in Figure 1.

Table 1

The sample size for the fresh white tea leaves dataset

category	Train	Test	Total
One bud and one leaf	422	106	528
One bud and two leaves	157	39	196
One bud and two leaves	262	66	328
Other	119	19	148

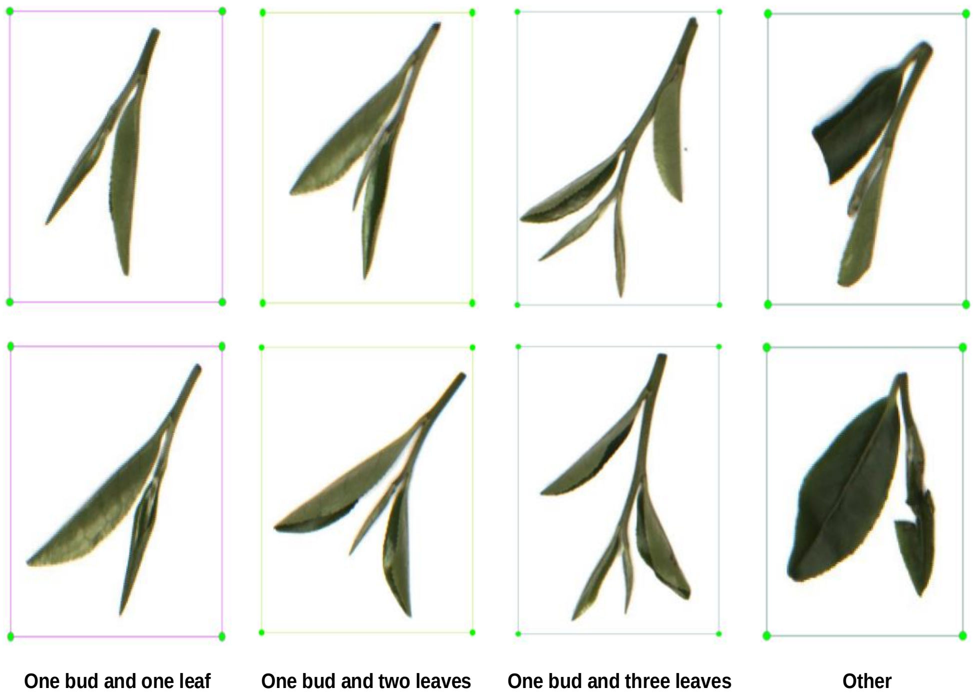

Fig. 1 - White tea fresh leaf category labelling

ALGORITHM PRINCIPLE

For the purpose of detecting white tea buds, this study will combine two original deep learning models—YOLOv3 and YOLOv5—and enhance the Neck layer on top of the YOLOv5 model. The core components of the original YOLOv3 and YOLOv5 models are Backbone, Neck, and Head, with Backbone retrieving picture data, the Neck doing feature fusion, and the Head performing prediction.

Darknet-53 serves as the YOLOv3 model's backbone. One DBL module, five Residual Blocks, and 53 convolutional layers make up Darknet-53. The DBL module is the tiniest one of them all in Darknet-53. In order to address the overfitting issue, the DBL module first convolves the image features before adding Batch Normalization processing and using the activation function Leaky Rectified Linear Unit; each Residual Block is made up of Zero Padding, DBL, and x Res Units. The DBL module uses convolution to complete the downsampling, and the Res Units follow suit to achieve jump connection using consecutive Conv (1*1) and Conv (3*3), which improves the gradient disappearance problem as the network depth increases. Darknet-53's final three Residual Block outputs are used to feed three different feature scales into the Neck Feature Pyramid Networks structure (FPN) for feature fusion and reconstruction, and the Head Convolution operation is then applied to the three feature maps that have been fused by the FPN structure (*Deng et al., 2021*). The structure of the YOLOv3 model is shown in Figure 2.

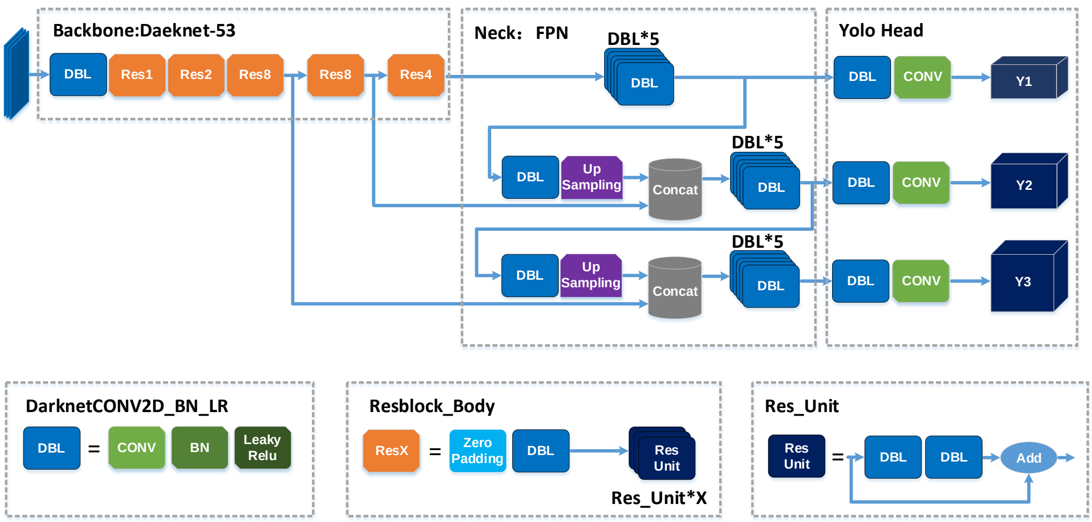

Fig. 2 - Model structure of YOLOv3

CSP Darknet-53, which includes the Focus module, DBL module, CSP module, and SPP module, serves as the skeleton of the YOLOv5 model. The Focus module divides the input picture into slices before entering the DBL module, the SCP1_X module splits the base layer mapping feature into two pieces, jump convolution is then used to fuse the features, and finally, the SPP module increases the ability of Backbone to extract features from MaxPool; to further enhance the feature fusion capabilities, the YOLOv5 model combines the SCP2_X module in Neck with the DBL module to construct the structure of Feature Pyramid Networks and Perceptual Adversarial Networks (FPN_PAN). Finally, the three feature maps from the fusion are applied to the Head. The structure of the YOLOv5 model is shown in Figure 3.

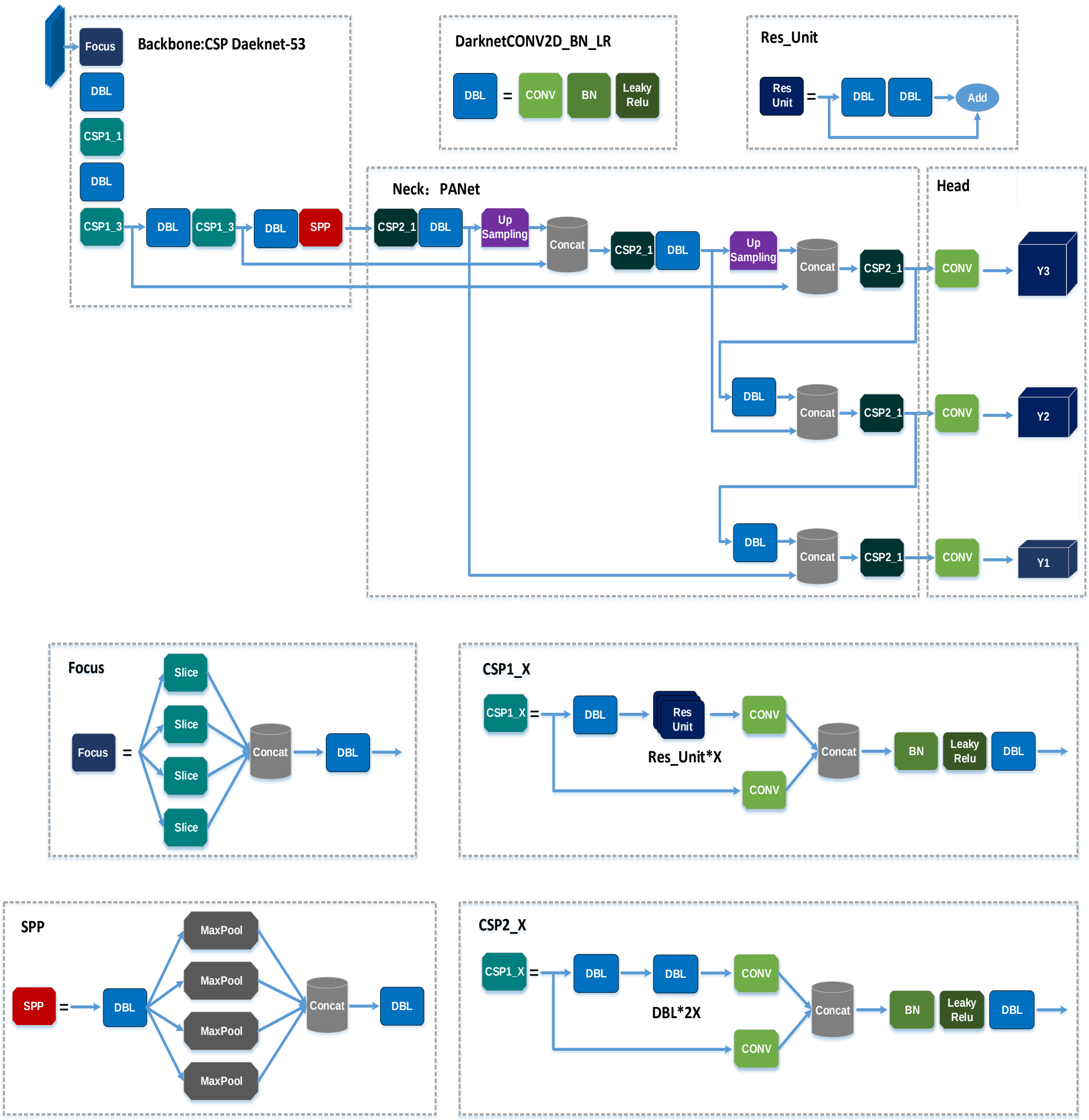

Fig. 3 – Model structure of YOLOv5

IMPROVED YOLOV5 MODEL

This study will improve the Neck portion of YOLOv5 and transform the FPN_PAN structure into a Bidirectional Feature Pyramid Network (BiFPN) structure (Qiu et al., 2023). The traditional FPN structure down-samples the image features before resampling them, which increases the feature level and makes it harder to detect small targets in the large feature layer. In contrast, the FPN_PAN structure adds a bottom-up channel to the FPN to fix the problem of one-way feature information transfer but does so at the expense of higher computational costs. By adding learning weights, the BiFPN structure outperforms the FPN_PAN structure by learning the relative weights of various input features and performing top-down and bottom-up feature fusion with fewer parameters. The schematic diagram of the three feature pyramid structures is shown in Figure 4, and the structure of the improved YOLOv5-BiFPN model structure is shown in Figure 5.

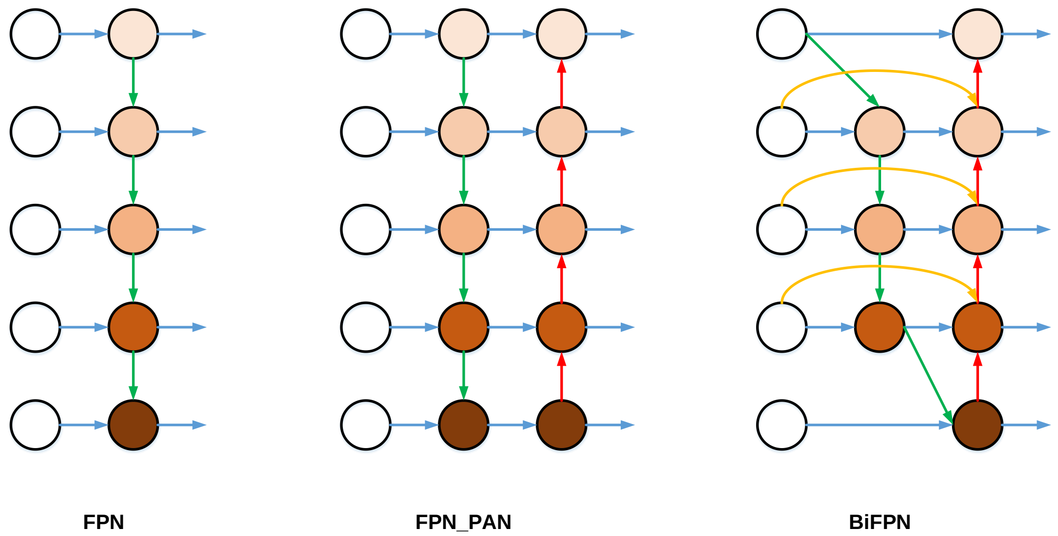

Fig. 4 – The structure of FPN, FPN_PAN and BiFPN module

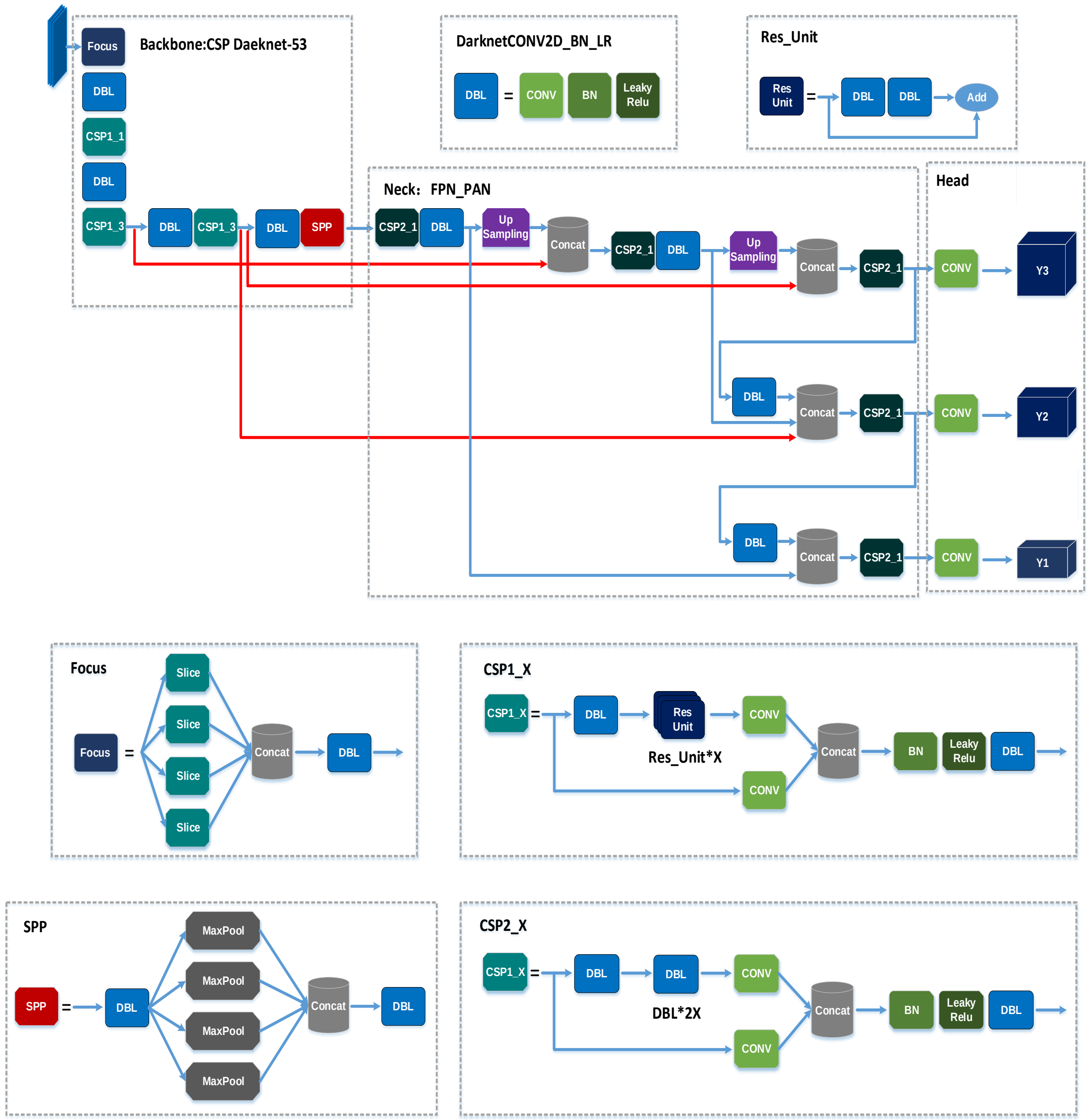

Fig. 5 – Model structure of improved YOLOv5-BiFPN

TEST ENVIRONMENT

Test environment configuration: Programming language Python3.7, deep learning framework Pytorch, graphics card Nvidia GeForce GTX1660 with 6GB video memory, processor Inter Core i7-9700K 3.60GHz octa-core, memory 32GB, and relevant libraries such as Torch, Torchvision, Opencv, Numpy, Scipy, Pillow, Matplotlib, Tqdm, H5py, etc.

NETWORK TRAINING

With 960 photos for the training dataset and 240 images for the validation dataset, three deep learning models—YOLOv3, YOLOv5, and YOLOv5-BiFPN—were trained on a graphics processing unit (GPU). The image sizes were then adjusted to 640 x 640. A total of 300 epochs were trained using Stochastic Gradient Descent (SGD), with an initial learning rate of 0.001 and a reduction to 0.0001 at the 200th epoch with a Batch Size of 4. The training loss rate curve of the three models is shown in Figure 6.

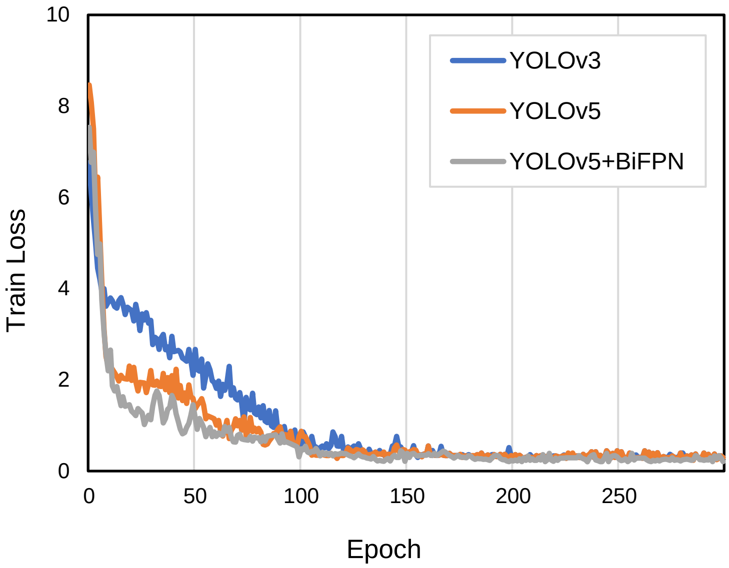

Fig. 6 - The training loss rate curve of the three models

Figure 6 shows that all three models produce positive training results. YOLOv3 has the lowest initial loss, followed by YOLOv5+BiFPN and YOLOv5 has the highest initial loss rate. As Epoch rises, the training loss of the three models eventually falls and stabilizes. Among them, YOLOv5+BiFPN converges the quickest and eventually converges between 0.20-0.24; YOLOv5 converges the second-fastest and eventually converges between 0.23-0.27; while YOLOv3 converges the slowest and ultimately converges between 0.26-0.29.

Evaluation Indicators

In this study, Recall (R), Average Precision (AP), and mean Average Precision (mAP) will be used as target detection evaluation indexes for white tea buds.

The precision and recall are respectively expressed by Equations (1) and (2):

$$Precision = \frac{TP}{TP+FP} \tag{1}$$

$$Recall = \frac{TP}{TP+FN} \tag{2}$$

where: TP (true positives) represents the number of positive samples correctly predicted by the model, FP (false positives) represents the number of positive samples incorrectly predicted by the model, and FN (false negatives) represents the number of negative samples incorrectly predicted by the model (*Li et al., 2022*).

AP and mAP are defined by Equations (3) and (4),

$$AP = \int_0^1 P(R)dR \tag{3}$$

$$mAP = \frac{\sum_{i=1}^{N} AP_i}{N} \tag{4}$$

where:

P represents Precision, AP_i represents the i-th category of target detection average precision, AP_i represents the average precision of the target detection for the i-th category, and N represents for the category of markers.

RESULTS

Comparison of Actual Detection Effects

In order to verify the actual detection effect of the YOLOv5+BiFPN model, the YOLOv5 model and the YOLOv3 model were used to test the images of the white tea fresh leaf dataset with four types of samples randomly selected, and the detection results are shown in Figure 7.

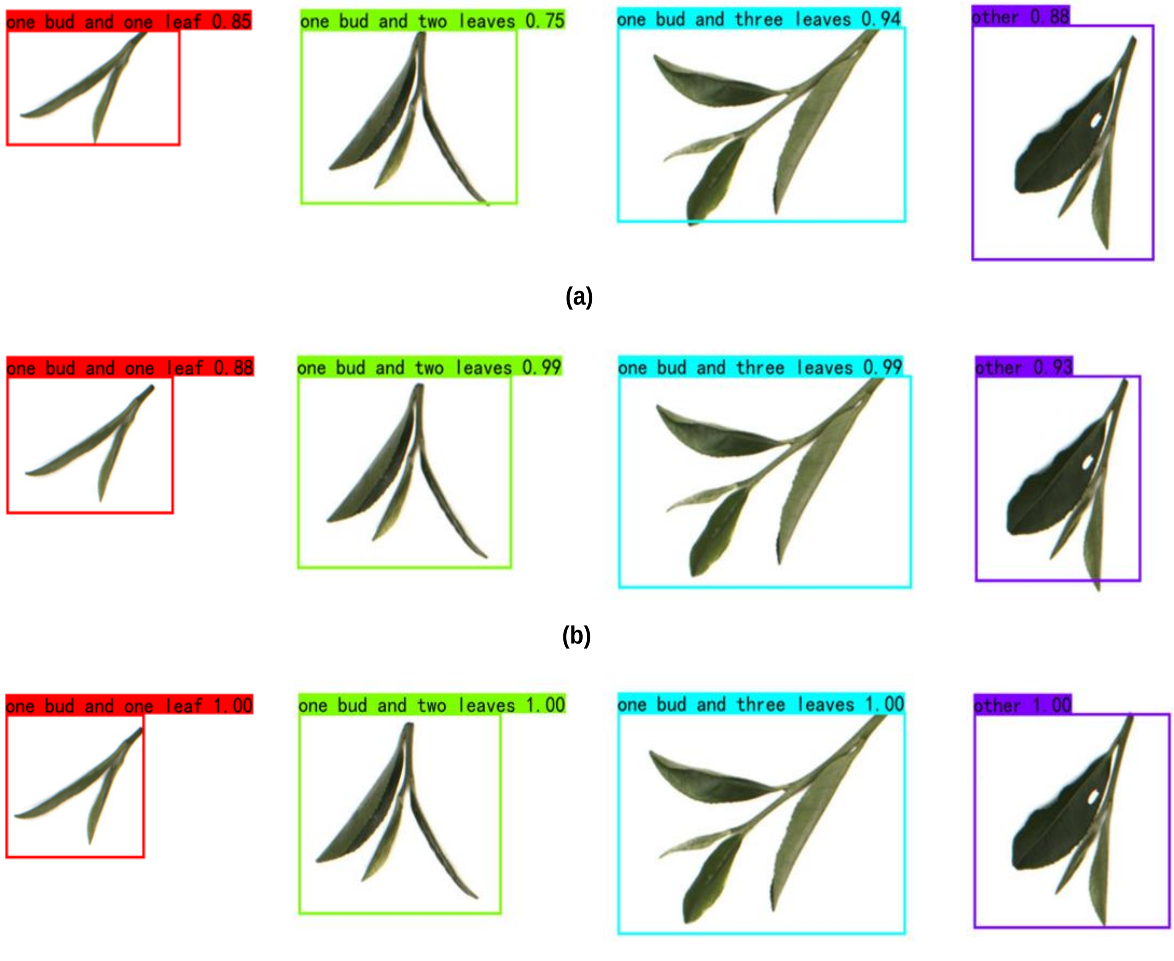

Fig. 7 - Comparison of detection effects of three models on white tea fresh leaves dataset.
(a) YOLOv3 model; (b) YOLOv5 model; (c) YOLOv5-BiFPN model

It can be seen from Figure 7 that the improved YOLOv5+BiFPN model has the highest confidence score and the best overall prediction box, but there is still room for optimizing the prediction box for others; The YOLOv5 model has the next best confidence score and the prediction box is better overall, but inaccurate for Other's prediction box; The YOLOv3 model had the lowest confidence score, the prediction box was not effective overall, the prediction box for one bud and two leaves, one bud and three leaves was inaccurate, and there was still room for optimization of the prediction box for other.

Comparison of Different Object Detection Models

In this study, the effectiveness of the improved model was verified by conducting ablation experiments on three models, and the results were compared on the same test data, and the experimental results are shown in Table 2.

Table 2

Comparison of detection performance for different algorithms

Model	**AP**				**mAP@0.5**
	one bud and one leaf	**one bud and two leaves**	**one bud and three leaves**	**other**	
YOLOv3	92.3%	89.6%	90.1%	86.4%	89.60%
YOLOv5	95.6%	92.3%	93.3%	89.9%	92.78%
YOLOv5+BiFPN	98.7%	96.4%	97.7%	94.6%	96.85%

Table 2 shows that the improved YOLOv5+BiFPN model has the highest average detection precision for all four types of buds in the white tea fresh leaf dataset, with the highest average detection precision of 98.7% for the one bud and one leaf category. mAP@0.5 was 96.85%. The average detection precision of all three models for the other class was the lowest among the four classes, which shows that there is still some difficulty for the model to detect the fine features in the bud incomplete leaves of the other class. The YOLOv5+BiFPN model improved the average detection precision of the other class by 4.7% compared to the YOLOv5 model and 8.2% compared to the YOLOv3 model, and combined with the YOLOv5+BiFPN model, the YOLOv5+BiFPN model improved the average detection precision of the other class by 4.7% compared to the YOLOv5 model. BiFPN model for one bud and one leaf class shows that the improved YOLOv5+BiFPN model has the advantage of extracting small features of white tea buds.

CONCLUSIONS

In this paper, the YOLOv5+BiFPN model was improved based on the YOLOv5 model, realized the high-precision detection of white tea buds in the established white tea fresh leaf dataset. The main results of this study are as follows:

1) The white tea fresh leaf dataset and YOLOv5+BiFPN white tea buds detection model were established, and the detection average precision of four categories of buds in the white tea fresh leaf dataset was 98.7%, 96.4%, 97.7%, and 94.6%, respectively. The mAP@0.5 was 96.85%.

2) The ablation experiment analysis showed that the YOLOv5+BiFPN white tea bud detection model improved mAP@0.5 with 4.07% higher than the YOLOv5 model and 7.25% higher than the YOLOv3 model.

3) The addition of the Bidirectional Feature Pyramid Network structure to the Neck in the YOLOv5 model can extract the small features in the fresh leaves of white tea more effectively, and its dual-channel and multi-scale feature fusion has more potential for the detection of buds such as one bud and one leaf and other.

In this study, the improved YOLOv5+BiFPN model is effectively combined with white tea fresh leaf images to provide a method and means for white tea bud detection based on deep learning image detection. The BiFPN structure can effectively improve the model's ability to extract the small features of white tea buds, providing an efficient, accurate, and intelligent buds detection model for the sorting of famous white tea.

ACKNOWLEDGEMENT

This work was supported by the National Natural Science Foundation of China (52105283), Huzhou Public Welfare Application Research Project (2020GZ01), Huzhou Natural Science Foundation Project (2022YZ54), Huzhou Vocational and Technical College Application of Special Public Welfare Technology for High-Level Talents (2022022GY08) and Teacher Professional Development Project for Domestic Visiting Scholars in Universities (FX2022113).

REFERENCES

[1] Dario L., Oscar P., Ruben U., et al., (2021), Cost-Performance Evaluation of a Recognition Service of Livestock Activity Using Aerial Images. *Remote Sensing*, 13(12), 2318. https://doi.org/10.3390/rs13122318

[2] Deng X., Yan X., Hou Y., et al., (2021), Detection of Behavior and Posture of Sheep Based on Yolov3. *INMATEH Agricultural Engineering*, 64(02), 457-466. https://doi.org/10.35633/inmateh-64-45

[3] Hong W., Ma Z., Ye B., et al., (2023), Detection of Green Asparagus in Complex Environments Based on the Improved YOLOv5 Algorithm. *Sensors*, 23, 1562. Sensors 2023, 23, 1562. https://doi.org/10.3390/s23031562

[4] Li X., Wang C., Ju H., et al., (2022), Surface Defect Detection Model for Aero-Engine Components Based on Improved YOLOv5.*Applied Sciences*, 12(14), 7235. https://doi.org/10.3390/app12147235

[5] Li Y., He L., Jia J., et al., (2021), In-field tea bud detection and 3D localization using an RGB-D camera. *Computers and Electronics in Agriculture*, 185: 106149. https://doi:10.1016/j.compag.2021.106149

[6] Lin L., Li C., Yang S., et al., (2019), Automated classification of Wuyi rock tealeaves based on support vector machine. *Concurrency and Computation practice & Experience*, 31(23):3.

[7] Mao YT., Huang Y., Wen XG., et al., (2020), Research on classification of fresh tea based on multiple features and multiple classifiers. *Journal of Chinese Agricultural Mechanization*, 41(12), 75-83. https://doi:10.13733/j.jcam.issn.2095-5553.2020.12.014

[8] Nan Y., Zhang H., Zeng Y., et al., (2023), Detection of Green Asparagus in Complex Environments Based on the Improved YOLOv5 Algorithm. *Computers and Electronics in Agriculture*, 23, 1562. Sensors 2023, 205, 107536. https://doi.org/10.1016/j.compag.2022.107563

[9] Qiu M., Huang L., Tang B., et al., (2023), Bridge detection method for HSRRSIs based on YOLOv5 with a decoupled head. *International Journal of Digital Earth*, 16(01), 113-129. https://doi.org/10.1080/17538947.2022.2163514

[10] Wang C., Zhang T., Zheng Y., et al., (2016), Simulation Study on the Screening of the Reciprocating Roller Screen with Two Degrees of Freedom for Tea. *Journal of Tea Science*, 36(06), 613-620. https://doi:10.13305/j.cnki.jts.2016.06.008

[11] Wang Z., Li S., Jin Z., et al., (2023), Accurate prediction technique of meteorological index relevant to frost damage for Anji white tea based on current planting zone. *Chinese Journal of Ecology*, 42(04), 1018-1024. https://doi:10.13292/j.1000-4890.202304.023

[12] Yang H., Chen L., Chen M., et al., (2019), Tender tea buds recognition and positioning for picking robot using improved YOLO-V3 model. *IEEE Access*, 23(11):180998-181011. https://doi:10.1109/ACCESS.2019.2958614

[13] Yi W., Cheng F., Qiu Y., et al., (2020), Research status and development trend of single tea-picking machine. *Journal of Chinese Agricultural Mechanization*, 41(11), 33-38. https://doi:10.13733/j.jcam.issn.2095-5553.2020.11.006

[14] Yuan H., Hua J., Deng Y., et al., (2016), Optimization of grading and classification technology for machine-picking leaves based on YJY-2 type classifier. *Transactions of the Chinese Society of Agricultural Engineering*, 32(06), 276-282. https://doi:10.11975/j.issn.1002-6819.2016.06.038

[15] Zhao F., Zhang J., Zhang N., et al., (2022), Detection of Cucurbits' Fruits Based on Deep Learning. *INMATEH Agricultural Engineering*, 66(01), 231-330. https://doi.org/10.35633/inmateh-66-32

[16] Zhang H., Xiao H., Mei S., et al., (2014), Status Quo and Future of Tea Sorting Technology. *Journal of Agricultural Engineering*, 04(05), 59-63. https://doi:10.13733/j.jcam.issn.2095-5553.2020.11.006

[17] Zhao Z., Zhang L., Wang X., et al., (2014), Condition Analysis and Numerical Simulation of Tea Plane Circular Sieve. *Journal of Tea Science*, 34(01), 1-8. https://doi:10.13305/j.cnki.jts.2014.01.001