Detection Method of Damaged Camellia Oleifera Seeds Based on YOLOv5-CB

ZHANG XIAO, ERJIE SUN, FENGWEI YUAN, JIANG PENG, AND JUNJIE LIU
College of Mechanical Engineering, University of South China, Hengyang, Hunan 421001, China

Corresponding author: Fengwei Yuan (254551514@qq.com)

This work was supported in part by the Hunan Province Enterprise Science and Technology Ombudsman Program under Grant 2021GK5049, in part by the University of South China Research Initiation Fund Program under Grant 200XQD022, and in part by the Hunan Provincial Department of Education Program under Grant 19C1602.

ABSTRACT To solve the problems that the existing sorting equipment cannot effectively identify and sort damaged Camellia oleifera seeds and traditional manual sorting of damaged Camellia oleifera seeds is inefficient and slow, in this paper, a damaged Camellia oleifera seeds detection method based on YOLOv5, coordinate attention, and weighted bidirectional feature pyramid network was designed. In this study, according to the actual requirements, firstly, the Coordinate Attention module (CA) was added to the YOLOv5 algorithm to improve the detection precision of damaged Camellia oleifera seeds in stacked Camellia oleifera seeds. Secondly, the network structure was optimized and the weighted bi-directional feature pyramid network (BiFPN) was added. The module integrates multi-scale features from top to bottom to reduce the missed detection of slightly damaged Camellia seeds. The final experimental results show that compared with the original YOLOv5 model, the detection precision of the improved model YOLOV5-CB is improved by 6.1%, reaching 92.4%, and the mean Average Precision (mAP) is also improved from 87.7% to 93.4%, the average detection time of a single Camellia seeds image is 6.4ms, which meet the requirements of precision and real-time in practical application.

INDEX TERMS Camellia oleifera seed, YOLOv5, attention mechanism, feature pyramid, machine vision.

I. INTRODUCTION
Camellia oleifera Abel is a unique woody oil plant with high economic value in China, mainly growing in the south, including Jiangxi, Hunan, Guangxi, and other regions, the planting area of more than 12 million mu [1], its fruit, oil palm, olive, coconut and known as the world’s four largest woody edible oil [2]. Camellia oil is the main product of Camellia oleifera seed pressing and has high nutritional and medicinal value. The antioxidant Camellia oil protects against oxidative stress, which can lead to liver damage and gastrointestinal ulcers [3]. Camellia oil is also traditionally used to treat stomach pain and burns in China [4]. Long-term consumption of Camellia oil can also help prevent atherosclerosis, high blood pressure, and obesity, so it is known as the “Oriental olive oil” [5], [6]. China is a big country for Camellia oil plantations. The no.1 Central Document of 2022 supports the expansion of the Camellia oil plantation area [7], which will further increase China’s Camellia oil plantation area.

Camellia oleifera fruit includes two parts: husk and Camellia oleifera seeds. In the production and processing of Camellia oil, two ways are often used: pressing oil with the shell and removing the shell. Camellia oleifera husks do not contain oil, and the shell-crushing can effectively improve the oil yield and the quality of Camellia oil while reducing machine wear and improving the service life of machines and accessories [8]. Therefore, peeling oil is the most common oil-pressing method. The product of Camellia oleifera fruit after shelling treatment contains a large amount of peel and a small number of broken Camellia oleifera seeds caused by shelling. To ensure the quality of Camellia oil, separating the fruit dander and damaged Camellia oleifera seeds from the Camellia oleifera seeds used for oil extraction is necessary. For example, Li et al. designed a seed shell sorting machine composed of a rubber conveyor belt and vibration
support plate by taking advantage of the difference in the shape and friction coefficient of Camellia oleifera seed and fruit shell so that the oil Camellia oleifera seed shell moves upward while the Camellia oleifera seed moves downward, and Camellia oleifera seed and fruit shell are separated [9]. For example, Chen et al designed a cutting-type Camellia oleifera fruit depilator by taking advantage of the difference in the quality of Camellia oleifera seeds and husks and realized the separation of husks and seeds by adjusting the rotation speed of the rubber plate installed in a spiral manner [10]. For example, Lu et al. designed a kind of oil Camellia seed in practical application, and the existing algorithms often consider the accumulation and occlusion of Camellia oleifera more common for Camellia oleifera seeds to stack, block each other, and have a large base.

The above algorithms show that the application of machine vision in the field of Camellia oleifera seed sorting is completely feasible, but their deficiency is that they do not consider the accumulation and occlusion of Camellia oleifera seed in practical application, and the existing algorithms often cannot guarantee a high recognition rate in these situations.

The complete shell is conducive to the storage of camellia oleifera seeds, because the damaged shell of camellia oleifera seeds is vulnerable to microbial contamination during transportation and storage, and the mildew rate is much higher than that of camellia oleifera seeds with complete shell [15]. The mildew of camellia oleifera seeds is extremely dangerous, which not only seriously affects the quality of camellia oil, but also leads to the accumulation of toxic substances [16]. Using camellia oleifera seeds with damaged shells during camellia oil extraction has the risk of introducing mildewed camellia oleifera seeds [17]. Aspergillus is the most common mold in moldy camellia oleifera seeds, which will cause serious harm to the human body. For example, the most common aflatoxin in moldy agricultural products has strong hepatotoxicity, carcinogenicity, teratogenicity, and mutagenicity [18]. Therefore, it is of great significance to ensure the quality and safety of Camellia oil to accurately sort out the damaged camellia seeds.

Based on the YOLOv5m model in the YOLO series, this paper integrates the attention module and weighted bidirectional feature pyramid network and makes full use of the attention mechanism to important information. Weighted bidirectional feature towers can aggregate features with different resolutions. A new camellia oleifera seed detection network model is constructed to solve the problem that existing algorithms cannot effectively identify damaged camellia oleifera seeds in complex environments such as occlusion and stacking. Because coordinate attention (CA) and weighted bidirectional feature pyramid (BIFPN) are introduced at the same time, I named the network YOLOv5-CB.

II. MATERIALS AND EQUIPMENT
A. EXPERIMENTAL MATERIALS
The camellia oleifera seed sample used in this experiment was produced in Suzhuang Town, Kaihua County, Quzhou City, Zhejiang Province. The whole experiment process completely simulated the actual industrial assembly line. Use labeling software to manually mark the position of damaged camellia seeds on the collected data set, and place the marked image under the label file. The experiment collected 558 data images and removed 19 data sets without damaged Camellia oleifera seeds. Of the 539 pictures that meet the experimental requirements, 70% are randomly selected as the training set, 20% are randomly selected as the verification set, and 10% are randomly selected as the test set. Some data sets are shown in Figure 1 (a).

The cutout is a regularization technique of a convolutional neural network. It can improve the network’s robustness by randomly shielding some square areas during the training process [19]. Due to the insufficient number of samples, the overlapping environment cannot be fully simulated. This paper uses Cutout data enhancement technology to simulate overlapping samples during training. The characteristics of the shielding area randomly generated by Cutout technology can reasonably represent the overlap of camellia oleifera seeds in practical application. The training set processed by cutout technology is shown in Figure 1 (b).
B. EXPERIMENTAL EQUIPMENT
The equipment used in this experiment mainly includes a feeding device, an image acquisition device, a conveyor belt, and parallel manipulator four parts. The whole working process is as follows: the shell and the mixture of Camellia oleifera seeds fall onto the sieve plate from the feed port. The quality of the causal shell is different from that of the Camellia oleifera seeds. The sieve plate is used for the rough selection of Camellia oleifera seeds and fruit shells, and the fruit shells fall into the debris collection box, and the Camellia oleifera seeds fall onto the conveyor belt. Camellia oleifera seeds arrive at the image collection area with the conveyor belt. The camera takes pictures of the Camellia oleifera seeds on the conveyor belt and transmits the collected images to the computer. The computer determines the damaged Camellia seeds through the algorithm and controls the manipulator to put the damaged Camellia oleifera seeds into the corresponding collection box, and the intact Camellia oleifera seeds fall into the final collection box. The 3d assembly drawing of the experimental equipment is shown in Figure 2. The physical picture of the equipment is shown in Figure 3.

The image acquisition device is shown in Figure 4, including 1 camera and 2 fill lights. Two parallel fill lights are located on both sides of the top of the camera, which can effectively reduce the influence of shadows during image acquisition. The camera part is mainly composed of AI-230U150M industrial camera and an LB0814-5M lens.

III. METHOD
A. YOLOv5-CB NETWORK MODAL
The model used in this paper is the YOLOv5 network of the latest version of the YOLO series. YOLOv5 network can be divided into S, M, L, and X according to depth_multiple and width_multiple, as shown in Table 1.

To select a suitable network model, this paper carries on the experimental comparison to the commonly used YOLOv5 series algorithms. The performance pairs of YOLOv5 series models are shown in Table 2. It can be seen from the table that YOLOv5 series models have little difference in

| Model    | Precision/% | Recall/% | mAP/% | Size   | Time   |
|----------|-------------|----------|-------|--------|--------|
| YOLOv5n  | 82.1        | 73.3     | 84.4  | 3.73MB | 6ms    |
| YOLOv5s  | 69.6        | 85.3     | 84    | 13.7MB | 7.1ms  |
| YOLOv5m  | 86.3        | 84       | 87.7  | 40.2MB | 7.9ms  |
| YOLOv5l  | 84.3        | 78.7     | 83.9  | 88.5MB | 8.6ms  |
| YOLOv5x  | 87.3        | 73.3     | 82.3  | 165MB  | 16ms   |
Precision, Recall and mean Average Precision (mAP) of the data set collected in this experiment, but they cannot meet the detection requirements of practical application. Through comprehensive consideration of various parameters and model sizes, to meet the requirements of industrial application precision and real-time detection at the same time, this paper chooses YOLOv5m as the basic model for improvement.

It can be seen from Table 2 that the performance of YOLOv5l and YOLOv5x has decreased to a certain extent, which conflicts with the cognition that improving the depth can enhance the ability of network learning characteristics. Although increasing the depth can improve the ability to learn different features, it does not always improve the performance of the network [20]. In this data set, the characteristics of damaged camellia oleifera seeds are likely to be confused with those of normal camellia oleifera seeds after overlearning, leading to performance degradation.

The overall structure of the YOLOv5 network consists of four parts: the input end, the backbone network, the neck, and the output end. The input end performs preliminary processing such as data enhancement and adaptive image scaling. The backbone network mainly obtains the feature map through a series of operations such as convolution, pooling, cross-scale connection, and spatial pyramid. A series of mixing and combining are performed to complete the fusion of high and low-level image features to obtain more image feature information. Finally, the output terminal predicts the location information. Complete the whole target detection process. The improved model YOLOv5-CB is shown in Figure 5.

### B. COORDINATE ATTENTION MODUL

In the process of practical industrial application, the Camellia seeds on the production line are inevitably stacked and occluded. To make the model more focused on the feature information of damaged Camellia oleifera seeds, the attention module was introduced into the YOLOv5 model. Attention is regarded as an effective solution for general classification tasks [21] and has been widely used in various neural networks. At present, the Squeeze-and-Excitation attention (SE) mechanism is widely used [22]. This mechanism calculates channel attention through 2D pooling, which significantly improves the performance of current mainstream network models. But SE only encodes the information between channels, ignoring the equally important spatial relationship, which is used to capture objects in vision tasks. Spatial information is also crucial in capturing target structures in visual tasks [23]. Later, Woo et al. proposed a Convolutional Block Attention Module (CBAM) [24]. Unlike the SE module which only calculates channel attention, the CBAM module obtains position information by reducing the channel dimension of the input tensor and then uses convolution to calculate the space. Attention, which is sequentially inferred along two independent dimensions, channel, and space, is structured as shown in Figure. 6(a). The CBAM module is a lightweight general module, and it can improve the performance of various models to a certain extent. It is also a commonly used attention mechanism. However, the spatial attention calculated by convolution can only extract part of the information, and still cannot establish the long-range dependencies that are essential for vision tasks. In this case, Hou et al. proposed a Coordinate Attention mechanism (CA) [25], the structure is shown in Figure. 6(b).

The CA attention module embeds the position information into channel attention. To avoid the loss of position information caused by the 2D global pooling operation, the 2D global pooling is decomposed into two 1D global pooling, and then the two 1D global pooling is used. The global pooling operation aggregates the input features in the horizontal and vertical directions into two independent feature maps and encodes the two feature maps that have embedded directional features into an attention map, so the location information can be saved in the two attention maps, and finally, the two attention maps are applied to the input feature map by multiplication. Specifically, given input X, CA attention encodes each channel in two spatial ranges \((H,1)\) and \((1,W)\) along horizontal and vertical coordinate directions respectively, that is the output of channel \(c\) at height \(h\) is:

\[
x_c^h(h) = \frac{1}{W} \sum_{0\leq j<H} x_c(h, i)
\]

Similarly, the output of channel \(c\) with width \(w\) is expressed as:

\[
x_c^w(w) = \frac{1}{H} \sum_{0\leq j<H} x_c(j, w)
\]

This paper introduces the CA attention mechanism into the backbone network of the YOLOv5 model. One of the improvements is to add the CA module to the partial convolution operation at the beginning of the backbone network so that more coordinate information can be retained when the image is first processed. The second improvement is to add a layer of the CA attention module to the backbone network so that the feature information of the picture can be further processed before entering the neck.

### C. WEIGHTED BIDIRECTIONAL FRACTURE PYRAMID NETWORK

Recognition of objects at different scales is a fundamental challenge in visual tasks. The proposal of pyramids provides a flexible and convenient method for visual tasks of representing and manipulating image information [26], so feature pyramid becomes an indispensable component in the visual recognition task. Considering the intensive computation and memory of the pyramid module, Lin et al. developed a bottom-up pyramid structure with lateral connections, exploiting the inherent multi-scale structure of deep convolutional networks to build feature pyramid networks (FPN)
The importance of analyzing images at many scales stems from the nature of the image itself. Real-world scenes contain objects of various sizes, large and small, and these objects contain a lot of information such as size, position, color, etc. The FPN pyramid structure of the FPN may miss information on other scales. To solve this problem, Tan et al. proposed a simple and efficient weighted bi-directional feature pyramid network (BiFPN) [28], It repeatedly applies top-down and bottom-up multi-scale feature fusion while introducing learnable weights to learn the importance of different input features. As shown in Figure 7.

The difference between the weighted bidirectional feature pyramid and the previous works is that when integrating different input features, most of the previous feature pyramids are simply summed up, but the contributions of these different input features to the fusion output features are different. The multi-scale feature fusion of BiFPN aims to aggregate features of different resolutions. Because the resolutions of the input features are different, BiFPN uses a weighted feature fusion method. The expression is as follows:

\[ O = \sum_{i \in \epsilon} \frac{w_i}{\sum_j w_j} \cdot I_i \quad (3) \]

where \( \epsilon = 0.0001 \) is used to avoid numerical instability, and \( w \) is a learned parameter, similar to the attention mechanism.
used to distinguish the importance of different features in the feature fusion process. Let’s take the calculation of the output of layer 6 as an example:

The calculation expression is:

\[ P_{td}^6 = \text{Conv}(w_1 \cdot P_{in}^6 + w_2 \cdot \text{Resize}(P_{in}^7)) \] (4)

\[ P_{out}^6 = \text{Conv}(w'_1 \cdot P_{in}^6 + w'_2 \cdot P_{td}^6 + w'_3 \cdot \text{Resize}(P_{out}^5)) \] (5)

In the formula \( P_{td}^6 \) – the middle feature of the 6th layer from top to bottom, \( P_{in}^6 \) – the input feature of the 6th layer from left to right, \( P_{in}^7 \) – the input feature of the 7th layer from the top to the bottom, \( P_{out}^5 \) – the output of the 5th layer from the bottom to the top Features, \( P_{out}^6 \) – output features of layer 6, Resize – usually an upsampling or downsampling operation.

IV. RESULTS AND ANALYSIS

The hardware configuration of the experimental environment in this paper is Intel(R) Xeon(R) Gold 5218R CPU@ 2.10GHz, GPU RTX 3090, 24GB video memory. The software configuration is Windows10, and Cuda 11.1. Deep Learning Framework Platform: Pytorch 1.8.1, Python 3.8.5.

A. PERFORMANCE EVALUATION INDEX

Appropriate evaluation indicators should be selected when analyzing the experimental results. Since Precision and Recall have certain limitations [29], this paper conducted a comprehensive analysis based on three indexes: Precision (P), Recall (R), and mean Average Precision (mAP). The Precision refers to the proportion of the prediction targets of a certain category that are correctly predicted to be positive to all the predicted positive samples of the category. The Recall represents the proportion of positive correct predictions in all positive forecast samples [30]. Whose definition formulas are:

\[ P(\text{Precision}) = \frac{TP}{TP + FP} \] (6)

\[ R(\text{Recall}) = \frac{TP}{TP + FN} \] (7)

where TP(True Positive) – the correct Positive sample, which is Positive and judged to be Positive; FN(False Negative) – False Negative sample, a positive sample but judged to be Negative; FP(False Positive) – False Positive sample, which is negative but judged to be Positive; TN (True Negative) – the correct Negative sample, originally a Negative sample and judged to be Negative.

What we pursue is that both the Precision and the Recall can reach an ideal value. However, both influence and restrict each other in reality. The most common in this case is to use the Average Precision (AP) and the mean Average Precision (mAP) to measure the detection precision of an algorithm [31], and the detection effect of the model is reflected by the values of AP and mAP. Calculated as follows:

\[ AP = \int_0^1 P(R) dR \] (8)

\[ mAP = \frac{\sum_{i=1}^{N} AP_i}{N} \] (9)

In the formula, \( N \)–the number of detected categories. Since there is only one detection class in this experiment, mAP is equivalent to AP here.
B. EXPERIMENT

This paper also refers to the CBAM attention module commonly used today. As can be seen from Figure 8, the performance of the YOLOv5 algorithm which only introduces the CBAM attention module decreases instead of rising, indicating that this method is not suitable for Camellia oleifera seed detection. However, the performance of the model using the CBAM attention module and BiFPN pyramid network is significantly improved, but the mAP curve cannot reach the level of the YOLOV5-CA model. In comparison, the improved method adopted in this paper is more suitable for the detection of damaged Camellia oleifera seeds under the condition of the industrial conveyor belt. As can be seen from the figure, the performance of the algorithm after the introduction of CA attention and the BiFPN pyramid in this paper has been significantly improved. Although the precision and mAP are not as good as the original model at the beginning of training, the model curves after convergence are better than the original model.

It can be seen from Table 3 that YOLOV5-CB has better performance than other improved methods. Compared with YOLOV5m, the detection precision of YOLOV-CB is 6.1% higher than the original model, reaching 92.4%, and the mAP is 5.7% higher. Although the recall rate is reduced to a certain extent, the improved model can fully meet the requirements for detection precision in practical applications. In addition, the size of the improved model has also increased from 40.2MB to 44.3 MB. However, the average detection time of a single camellia seed image is 6.4 ms, which meets the requirements of the detection task of damaged camellia seeds in the industrial assembly line.

It is easy to see from Table 4 that the improved model in this paper is superior to several well-known target detection algorithms in all aspects of the dataset. On the one hand, the basic model of YOLOv5-CB uses the well-known YOLO
series of algorithms. On the other hand, the improved method is selected based on the characteristics of camellia oleifera seeds. First of all, based on YOLOv5, this paper integrates the CA focus mechanism module to make the algorithm pay more attention to the important feature information of the image, and save the coordinate information of the image feature. Secondly, a multi-scale feature fusion BiFPN weighted bi-directional feature pyramid network is introduced to further improve detection precision and reduce model loss and error detection. The final experimental results show that the YOLOv5-CB model has higher precision and average precision than the original model. The average detection time of a single camellia seed image is 6.4 ms, which meets the real-time requirements.

C. MODEL DETECTION RESULTS ANALYSIS

In practical application, we often face two difficulties. On the one hand, the number of Camellia oleifera seeds on the conveyor belt is large, and the proportion of damaged seeds is relatively small. At the same time, the damage of individual damaged Camellia oleifera seeds is small, and the proportion of damaged seeds on the whole picture is very low, which will lead to the occurrence of missing inspection. On the other hand, in the actual assembly line, Camellia oleifera seeds often slip onto the conveyor belt through the vibrating screen plate, which makes the damage in the dense Camellia oleifera seeds difficult to identify. In this paper, the optimization of the original YOLOv5 model can effectively improve the missed and false detection of the model in practical industrial applications.

Figure 9. shows the comparison of the detection effect of the original model and the YOLOv5-CB model on some samples. It can be seen from Figure 9. (a) that the improved YOLOv5 model is more accurate for the detection of slightly damaged Camellia oleifera seeds. As shown in Figure 9. (b), the YOLOv5-CB model has a higher detection precision for damaged Camellia seeds in dense Camellia oleifera seed groups. Through the detection results of evaluation indicators and some samples, the CA attention mechanism was introduced to strengthen the processing of damage feature information of the model, and effectively improve the detection precision of the model for slightly damaged Camellia oleifera seeds. Through the multi-scale feature fusion of BiFPN weighted bidirectional feature pyramid, the detection ability of damaged Camellia seeds in dense Camellia oleifera seeds was greatly improved.

It can be seen from the above that YOLOv5-CB performs well, but there are still some shortcomings. On the one hand, YOLOv5-CB has an increase of 4.1MB compared with the original model, which should be further lightweight to adapt to applications on mobile terminals and embedded devices. On the other hand, the detection time of YOLOv5-CB single image is 6.4 ms, but in fact, this is the detection speed in the laboratory environment. In practical application, the cost of using 3090 equipment is inappropriate. The model should be further simplified to improve the detection time under traditional equipment.

V. CONCLUSION

Improving the detection accuracy and efficiency of camellia oleifera seeds is of great significance for supporting the expansion of the camellia oleifera planting area. Aiming at the problems of low manual sorting efficiency and low machine selection accuracy of camellia oleifera seed damage, a detection method of camellia oleifera seed damage based on attention mechanism and feature pyramid was proposed. First of all, to simulate the real industrial environment, we designed a special experimental scheme to collect data sets and used Cutout technology to simulate overlapping and occluding environments. Secondly, through the experimental comparison of several models with different widths and depths of the YOLO series, we choose YOLOV5m, which has a relatively balanced effect and detection time, as the basic model. Finally, a coordinate attention module and weighted bi-directional feature pyramid network are introduced to strengthen the detection of slightly damaged camellia seeds in the model and improve the detection ability of the model for damaged camellia seeds in dense camellia seeds.

The final experimental results show that YOLOv5-CB meets the requirements of high precision and high-speed detection in practical applications, and its performance is superior to the general target detection model.

Although YOLOv5-CB meets the requirements of high precision and fast speed for industrial applications, we still need to further optimize it in the future. The YOLOv5-CB model is not superior to the original model in size, so we should lightweight it to adapt to mobile terminals and embedded devices with higher model requirements. Secondly, the detection time of YOLOv5-CB single image is 6.4 ms, but in fact, this is the detection speed in the laboratory environment. In practical application, the cost of using 3090 equipment is inappropriate. The model should be further simplified to improve the detection time under conventional equipment.

REFERENCES

[1] F. Luan, J. Zeng, Y. Yang, X. He, B. Wang, Y. Gao, and N. Zeng, “Recent advances in Camellia oleifera Abel: A review of nutritional constituents, biofunctional properties, and potential industrial applications,” J. Funct. Foods, vol. 75, Dec. 2020, Art. no. 104242, doi: 10.1016/J.JFF.2020.104242.
[2] G. Meng, J. Han, X. Zhan, Z. Tu, and J. Zhou, “Industry analysis and development strategy of oil? Tea Camllia in China,” China Oils Fats, vol. 46, no. 7, pp. 104–108, 2021, doi: 10.19902/J.CNKI.ZGYZ.1003-7969.2021.07.018.
[3] Y.-T. Cheng, C.-C. Lu, and G.-C. Yen, “Beneficial effects of camellia oil (Camellia oleifera Abel.) on hepatoprotective and gastroprotective activities,” J. Nutritional Sci. Vitaminol., vol. 61, pp. S100–S102, Jan. 2015, doi: 10.3177/JNSV61.1S100.
[4] C.-P. Lee and G.-C. Yen, “Antioxidant activity and bioactive compounds of tea seed (Camellia oleifera Abel.) oil,” J. Agricult. Food Chem., vol. 54, no. 3, pp. 779–784, Feb. 2006, doi: 10.1021/JF052325A.
[5] C.-P. Lee, P.-H. Shih, C.-L. Hsu, and G.-C. Yen, “Hepatoprotection of tea seed oil (Camellia oleifera Abel.) against CCl4-induced oxidative damage in rats,” Food Chem. Toxicol., vol. 45, no. 6, pp. 888–895, Jun. 2007, doi: 10.1016/J.FCT.2006.11.007.
[6] R. Wang, “Nutritional value of special oils and fats,” *Cereals Oils Process.*, vol. 1, pp. 9–10, Jan. 2004.

[7] Y. Gao and q. Chang, “Tang Renjian, director of the Central Agricultural Office and Minister of Agriculture and Rural Affairs, interprets the no.1 central document of this year—Striving to open up a new situation of rural revitalization,” *People’s Weekly*, no. 4, pp. 27–29, 2022.

[8] X. Luo and W. Zhu, “Processing technology of oil-tea camellia seed oil and comprehensive utilization of *Camellia oleifera* resources,” *China Oils Fats*, vol. 35, no. 9, pp. 13–17, 2010.

[9] Y. Li, Y. Wang, L. Deng, and Z. Chen, “Research on the effect of *Camellia oleifera* fruit shelter and sorting machine,” *Hunan Forestry Sci. Technol.*, vol. 42, no. 2, pp. 38–42, 2015.

[10] L. Chen, S. Hu, Z. Yao, J. Ren, F. Pan, and Y. Hong, “Design and experiment of cutting type *Camellia oleifera* fruit sheltering machine,” *Agricult. Eng.*, vol. 11, no. 5, pp. 80–85, 2021, doi: 10.3969/j.issn.2095-1795.2021.05.018.

[11] Y. Lu, W. Wang, B. Zhong, Q. Jin, and B. Wang, “Design and experiment of belt screen seed sorter for *Camellia oleifera* fused with photoelectric color sorting technology,” *Trans. Chin. Soc. Agricult. Mach.*, vol. 51, no. 51, pp. 429–439 2020, doi: 10.6040/j.issn.1000-1298.2020.S1.051.

[12] X. Lin, Z. Cheng, L. Li, J. Tan, and F. Wu, “A method of *Camellia* fruit sorting based on multi-features identification in preference aiNert,” *J. Chin. Agricult. Mechanization*, vol. 42, no. 9, pp. 187–194, 2021, doi: 10.13733/j.jcam.issn.2095-5553.2021.09.26.

[13] D. Yafei, H. Sisi, W. Yanqing, T. Yang, and H. Jun, “Sorting method of seeds and shells of the *Camellia oleifera* fruit based on machine vision,” *J. Chin. Agricult. Mechanization*, vol. 41, no. 6, pp. 171–178, 2020.

[14] Y. Xie, Y. Ding, F. Wang, S. Wei, and D. Yang, “Integrity recognition of *Camellia oleifera* seeds based on convolutional neural network,” *Trans. Chin. Soc. Agricult. Machinery*, vol. 51, no. 7, pp. 13–21, 2020.

[15] G. Zhu, H. Liu, Y. Xie, Q. Liao, Y. Lin, Y. Liu, Y. Liu, H. Xiao, Z. Gao, and S. Hu, “Postharvest processing and storage methods for *Camellia oleifera* seeds,” *Food Rev. Int.*, vol. 36, no. 4, pp. 319–339, May 2020, doi: 10.1080/08959674.2019.1649688.

[16] L. Feng, S. Zhu, F. Lin, Z. Su, K. Yuan, Y. Zhao, Y. He, and C. Zhang, “Detection of oil chestsnuts infected by blue mold using near-infrared hyperspectral imaging combined with artificial neural networks,” *Sensors*, vol. 18, no. 6, p. 1944, Jun. 2018, doi: 10.3390/s18061944.

[17] H. Jiang, X. Jiang, Y. Ru, Q. Chen, X. Li, L. Xu, H. Zhou, and M. Shi, “Rapid and non-destructive detection of natural mildew degree of postharvest *Camellia oleifera* fruit based on hyperspectral imaging,” *Infer. Phys. Technol.*, vol. 123, Jun. 2022, Art. no. 104169, doi: 10.1016/j.infrer.2022.104169.

[18] M. Abrar, F. M. Anjum, M. S. Butt, I. Pasha, M. A. Randhawa, F. Saeed, and K. Waqas, “Aflatoxins: Biosynthesis, occurrence, toxicity, and remedies,” *Crit. Rev. Food Sci. Nutrition*, vol. 53, no. 8, pp. 862–874, Jan. 2013, doi: 10.1080/10408398.2011.563154.

[19] T. DeVries and G. W. Taylor, “Improved regularization of convolutional neural networks with cutout,” 2017, arXiv:1708.04552.

[20] D. Bhatt, C. Patel, H. Talsania, J. Patel, R. Vaghela, S. Pandya, K. Modi, and H. Ghaayat, “CNN variants for computer vision: History, architecture, application, challenges and future scope,” *Electronics*, vol. 10, no. 20, p. 2470, Oct. 2021, doi: 10.3390/electronics10202470.

[21] J. Park, S. Woo, J.-Y. Lee, and I. So Kweon, “BAM: Bottleneck attention module,” 2018, arXiv:1807.06514.

[22] J. Hu, L. Shen, and G. Sun, “Squeeze-and-excitation networks,” in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7132–7141, doi: 10.1109/CVPR.2018.00745.

[23] C. Xie, H. Zhu, and Y. Fei, “Deep coordinate attention network for single image super-resolution,” *IET Image Process.*, vol. 16, no. 1, pp. 273–284, Jan. 2022, doi: 10.1049/ietipre.12364.

[24] S. Woo, J. Park, J. Y. Lee, and I. S. Kweon, “CBAM: Convolutional block attention module,” in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 3–19, doi: 10.1007/978-3-030-01234-2_1.

[25] Q. He, D. Zhou, and J. Feng, “Coordinated attention for efficient mobile network design,” 2021, arXiv:2103.02907.

[26] E. H. Adelson, C. H. Anderson, J. R. Bergen, P. J. Burt, and J. M. Ogden, “Pyramid methods in image processing,” *RCA Eng.*, vol. 29, no. 6, pp. 33–41, 1984.

[27] T.-Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, “Feature pyramid networks for object detection,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2117–2125, doi: 10.1109/CVPR.2017.106.

ZHUO JING received the B.S. and M.S. degrees in control science and engineering from Zhejiang University, Hangzhou, China, in 2017 and 2020, respectively. Currently, he is pursuing the Ph.D. degree with the University of South China. His research interests include image processing, computer vision, and pattern recognition.

ZHANG XIAO received the B.S., M.S., and Ph.D. degrees in mechatronic engineering from China Agricultural University, Beijing, China, in 2010, 2013, and 2020, respectively. In 2013, he joined with the College of Mechanical Engineering, University of South China, where he has been the Head of the Department of Intelligent manufacturing, since 2020. His research interests include image processing, computer vision, machine learning, and pattern recognition.

ERJIE SUN was born in Anhui, China, in 1996. He received the graduate degree in mechanical and electrical engineering from Hefei University, in 2016. He is currently with the School of Mechanical Engineering, University of South China. His research interests include image processing and computer vision.

FENGWEI YUAN was born in June 1977. He is an Associate Professor and the Director of the Department of Mechanical Design, Manufacturing and Automation, University of South China. He is the Executive Director of the Hunan Instrument Society. Mainly engaged in electromagnetic integration and measurement and control technology, embedded systems and intelligent instruments, and manufacturing information engineering.

JIANG PENG was born in Yueyang, Hunan, China, in 1997. He is currently pursuing the master’s degree with the University of South China. His research interests include object detection and image recognition based on deep learning.

JUNJIE LIU was born in Yongzhou, Hunan, China, in 1999. He is currently pursuing the master’s degree. His main research interests include deep learning and medical image processing.