X-Disciplinarity

DOI: https://doi.org/10.63174/xdi.PARX2270

Original Article

# C3-Light Lightweight Algorithm Optimization under YOLOv5 Framework for Apple-Picking Recognition

https://doi.org/10.63174/xdi.PARX2270

Received: 19 Feb 2025

Accepted: 28 Feb 2025

Published: 7 Mar 2025

Open Access

Kecheng SHAN[1,†], Quanhong FENG[2,†], Xiaowei LI[1], Xianglong MENG[1], Hongkuan LYU (LV)[1], Chenfeng WANG[1], Liyang MU[1], Xin LIU[1, 2,*]

**Abstract**: As the fruit-picking process is a labour-intensive and time-consuming task, the accurate and efficient recognition of apples during picking is of great significance for improving the overall efficiency of apple harvesting, reducing labour costs, and enhancing the quality of fruit picking. Although YOLOV5 algorithm can effectively detect apple status, its use on portable mobile devices still faces problems such as running lag. This paper is dedicated to the optimization of the C3-Light lightweight algorithm based on the widely used YOLOv5 framework for apple-picking recognition. The network structure of the C3-Light algorithm is redesigned. By introducing novel convolutional block arrangements and fine-tuning the hyperparameters related to the network layers, the model's feature extraction ability is enhanced while maintaining a relatively simple architecture. Through these improvements, the calls for hardware resources are remarkably reduced. Experimental results clearly demonstrate that the lightweight C3-Light model can maintain the original high-level accuracy. Specifically, it reduces GFLOPs by approximately 17% compared to the original model, which means a significant decrease in the computational complexity. Moreover, the GPU memory usage is decreased by 11%, indicating better memory utilization efficiency.

## 1. Introduction

In recent years, with the development of artificial intelligence, intelligence has been realized in many scenes, and intelligent machinery has been used to replace manual labour. Among them, machine vision is widely used in scenes such as defect detection, picking and recognition. The YOLO algorithm has developed rapidly in recent years and occupies an important position in machine vision.

The YOLO series of algorithms plays an important role in the field of computer vision, and related studies are constantly emerging. Muhammad Hussain[1] conducted a comprehensive review of YOLOv1 to v8, analyzing its architectural evolution, training strategies, and practical applications, pointing out that YOLO has innovative advantages in object detection but also faces some challenges, such as dealing with occlusion and fine-grained object detection. Junyan Tian et al. [2] proposed an object detection method based on YOLOv5 for intelligent sweeping robots, which improved the detection performance and model efficiency by introducing multiple modules . Jinsu An et al. [3] improved YOLOv5 for small object detection on satellite images by adding new feature fusion layers and prediction heads, etc., improving the detection accuracy. Jianting Shi et al. [4] improved YOLOv5 for steel surface defect detection by adding the attention mechanism and clustering the anchor boxes, improving the detection effect. Jianrong Cao et al. [5] proposed a hybrid underwater target detection algorithm based on YOLOv5 combined with CBAM and CloU, improving the feature extraction and fusion ability . Lijuan Sun et al. [6] proposed a lightweight apple detection method, YOLOv5-PRE, based on YOLOv5. By introducing lightweight structures and attention mechanisms, the detection performance was improved. Meiyan Li et al. [7] improved the YOLOv5 algorithm for remote sensing image target detection by clustering the anchor box and optimizing the model, improving the detection performance . Zexuan Guo et al. [8] proposed the MSFT-YOLO model based on Transformer for detecting steel surface defects. By adding modules and using multi-step training methods, the detection accuracy was improved. Yu Zhang et al. [9] proposed a vehicle detection method based on the improved YOLO v5, using the Flip-Mosaic algorithm to improve the detection accuracy. These studies provide valuable references and improvement directions for object detection in different scenarios.

The visual recognition algorithm for fruits has always been a crucial subject in the research on rapid detection of fruit yield. Several local and international researchers have developed various detection algorithms for different types of fruits, such as tomatoes [10–13], grapes [14–18], lychee [19–21], and apples [22–24]. Early target fruit detection methods mainly relied on traditional image processing or machine learning methods to identify fruits based on their colors, shapes, textures, and other features [25]. For example, Yu et al. [19] employed color and texture features to train the random forest binary classification model for identifying litchi fruits. The proposed method achieved a recognition accuracy rate of 89.92% for green litchi and 94.50% for red litchi. Additionally, Syazwani et al. [26] demonstrated excellent fruit counting performance by using ANN-

[1] School of Mechanical and Electronic Engineering, Shandong Agriculture and Engineering University, Jinan 250100, Shandong Province, PR China

[2] School of physics and electronics, Shandong Normal University, Jinan 250358, Shandong Province, PR China

[†] Made the same contribution

[*] Corresponding Author: liux951127@163.com

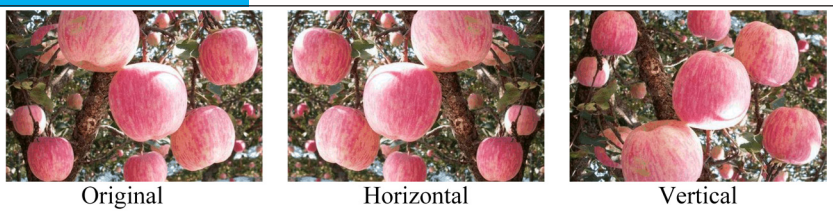


**Figure 1** Data rotation result

GDX (Artificial Neural Network, Variable Learning Rate Backpropagation) classification to detect pineapple crown images, with an accuracy of 94.4%. As for Fu et al. [27], they focused on the combined HOG and LBP (Oriented Gradient and Local Binary Pattern) texture features while using the SVM (Support Vector Machine) classification algorithm. They achieved an average scale detection rate of 89.63% for bananas and an average detection time of 1.325 s, with the shortest detection time being 0.343 s. To sum up and considering different studies, the feature-based image processing method has a slow detection speed, low accuracy, and poor adaptability to the extreme lighting environment in the orchard.

## 2. Materials of data

### 2.1. Data Source

The apple image data set used in this research is collected with different apple orchards from the web. The images cover various growth stages of apples, different lighting conditions, and different angles and distances of shooting. In this study, 87 images have been collected and sifted from the Internet. In order to ensure the diversity and representability of the data set, after labelling using labelling tool, horizontal inversion and vertical inversion can be used to expand the data set, and 261 pictures can finally be obtained. Horizontal and vertical rotation of the image results are shown in the **Figure. 1**. Horizontal inversion is the mirror symmetry of the original image, and turn the original image upside down and the vertical inversion is gotten.

### 2.2. Data Annotation

The collected apple images are annotated using professional annotation tools - labelling. Each apple in the image is marked with a bounding box. Taking into account the limitations of the picking robot arm, the labels are divided into three categories (**Figure. 2**), U-cover (uncovered), A-cover (covered by apple), and T-cover (covered by tree). When it is applied to the picking of the robot arm, the unshielded fruit is picked first. If the shielded fruit of the apple becomes unshielded after the previous apple is picked, it is picked directly; otherwise, the robot arm changes the picking position through path planning. Tree branch occlusion has the lowest priority, and when only this state exists in the camera field of view, apples will move to the next picking position. The annotation process strictly follows the YOLO format to ensure the consistency and standardization of the annotation information.

### 2.3. Data Set Division

The data set is randomly divided into a training set and a validation set according to the ratio of 8.5:1.5. The training set is used to train the model, and the validation set is used to adjust the model parameters during the training process to prevent overfitting. The test set is randomly selected from all images for validation.

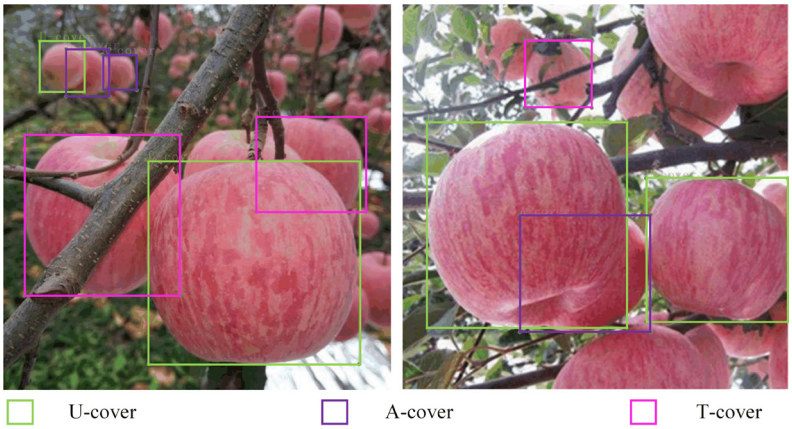


**Figure 2** Annotated examples

## 3. Method and Experimental setting

### 3.1. Method

You Only Look Once (YOLO), a real-time object detection technique, divides an image into grid cells and simultaneously predicts the bounding boxes and classes of objects within each cell. This unique approach endows YOLO with remarkable speed and precision in object detection. Over time, the YOLO algorithm has continuously evolved, culminating in the advanced YOLOv11. YOLO offers a diverse range of models with varying sizes, namely the n, s, m, l, and x scales, which are determined by specific scaling factors.

Lower versions of YOLO run faster but less accurate, and higher versions have improved accuracy, but have higher performance requirements for the hardware, increasing the cost of the hardware. Therefore, considering hardware cost and accuracy, the YOLOv5 version is the most widely used version. Among them, YOLOv5 6.1 is improved on the basis of 5.0, replacing the Focus module with the Conv module, which ensures accuracy and speeds up the detection speed. This study is optimized based on YOLOv5 6.1. The **Figure 3** shows the network architecture of YOLOv5. The annotation of Attention is the insertion position of the attention mechanism in this study.

### 3.2. Proposed method

The environment for apple picking is complicated, and ordinary YOLOv5 will recognize items other than apples as apples in the recognition of complex scenes. Therefore, CBAM (Convolution block attention module) and LSE (Squeeze and Excitation) attention mechanism is added in the backbone network or feature fusion part.

The LSE attention mechanism aims to enhance the model's attention to important features by creating a bottleneck that weights the feature graph channels. The principle is that the input feature graph x is transformed by 1x1 and 3x3 convolution to get feature graph x1; At the same time, the local

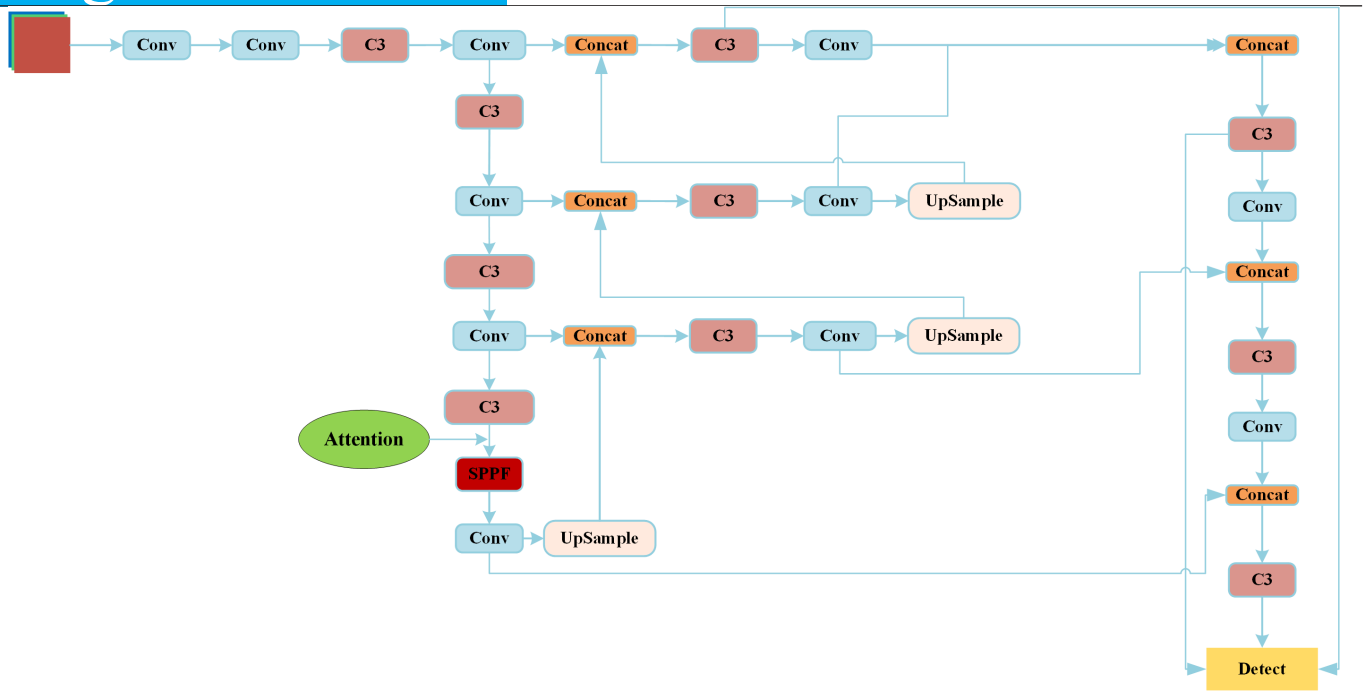


**Figure 3** Network structure diagram of YOLOv5

information is obtained by adaptive average pooling of x. After two 1x1 convolution and ReLU transformations, the attention diagram Att is obtained after Sigmoid activation, and its size is adjusted to be consistent with x1. Then Att and x1 are multiplied element by element to get the weighted feature map, and residual connections can be made according to the conditions. The mechanism has many advantages, such as enhanced feature representation to improve detection accuracy, local and global information fusion, high flexibility and adaptability with adaptive average pooling and adjustable parameters, and a good balance between computational efficiency and performance through 1x1 convolution, reduction of BatchNorm2d layer and residual connection.

The CBAM attention can weigh features in channel dimension and spatial dimension, so that the network pays more attention to the key feature areas of apple, and suppress the interference of background and irrelevant information. In the channel attention module, through global average pooling and multi-layer perceptron processing of features of different channels, the importance weight of each channel is learned, so as to highlight the channel where the apple features are located. In the spatial attention module, the feature map is pooled in the horizontal and vertical directions to obtain the spatial attention weight and further focus on the location region of the apple.

Due to the mobile factors of apple picking devices, the size of the hardware is limited. At present, most devices for processing visual information use small intelligent microcontrollers such as Raspberry PI, but in actual use, YOLOv5 is still difficult to use on this hardware and often occurs. In view of this situation, the lightweight detection network is proposed to reduce the stackage

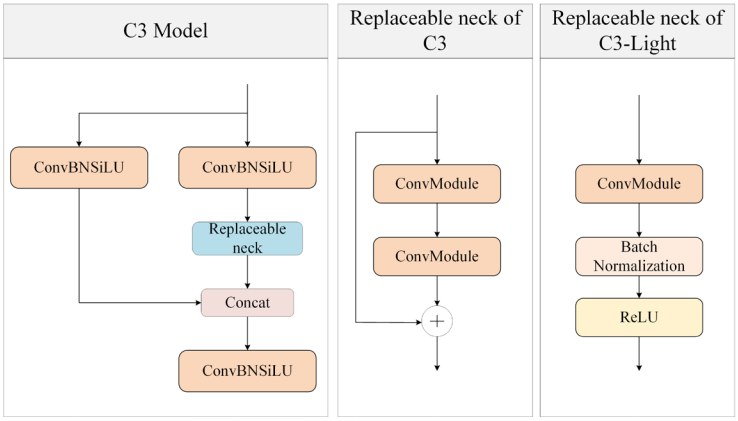


**Figure 4** Network structure diagram of C3-Light

during operation, and the C3 module is improved (C3-Light) to reduce the network pressure and adapt to the microcontroller, the corresponding C3-Light principle is given in the **Figure 4**. In this study, the addition of the attention model and replacement of the C3 model are in the backbone part of the model. The **Figure 5** shows the modified YOLOv5 network structure. C3 in the head is replaced with C3-light, which reduces the number of network layers in the early stage of processing after image input and speeds up the early stage of image processing. In the Attention part of the figure, attention mechanisms can be added.

### 3.3. Experimental Equipment and Experiment Setting

The experiment is carried out on a server equipped with an NVIDIA GeForce RTX 4080 Super GPU (memory of GPU is 16 GB), Intel Core i9-13900K CPU, and 32 GB memory of DDR5 version. The GPU provides strong computing power support for the training and inference of the deep learning mod-

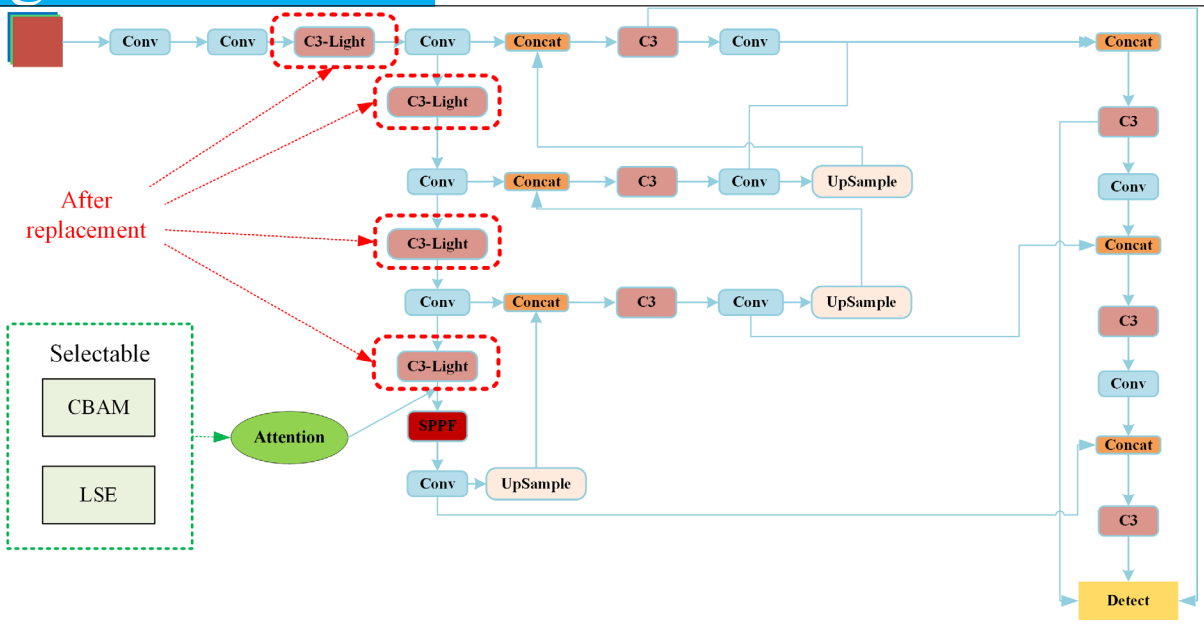


**Figure 5** Revised Network structure diagram of YOLOv5

el, ensuring the efficiency of the experiment.

The deep learning framework used in this research is PyTorch. The Python programming language is used for code implementation, and related libraries such as NumPy and OpenCV are also used for data processing and visualization. The details of the software and experiment are shown in **Table 1**.

Three basic groups of comparison experiments are set up. The lightweight C3-Light model proposed by combining the 5s model in YOLOv5 with the mainstream attention mechanism LSE and CBAM is used to compare the lightweight results.

**Table 1**. Experimental parameter setting.

| Parameter | Value |
|---|---|
| Python version | 3.8.1.8 |
| PyTorch version | 2.1.0 |
| Wandb version | 0.16.3 |
| CUDA version of NVIDIA | 12.0 |
| CUDA version of PyTorch | 11.7 |
| Epochs | 500 |
| Input image size | 640 |
| Batch size | 8 |

over Union) thresholds.

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

$$nAP = \frac{1}{n}\sum_{i=1}^{n} AP_i \quad (3)$$

In the **Equation 1, 2 and 3**, TP represents the true cases, FN represents the false negative cases, FP represents the false positive cases, and n represents the total number of cases.

## 4. Experimental Results Analysis

### 4.1. Comparative Experiments

To visually showcase the detection effectiveness of the algorithm proposed in this paper, a comparison experiment was conducted between the YOLOv5s model and the proposed model in terms of the confusion matrix. The results of this comparison are presented in **Figure 6**. In the confusion matrix, the rows denote the true categories, and the columns represent the predicted categories. The proportions of successfully predicted categories are indicated by the values within the diagonal regions, whereas the proportions of incorrectly predicted categories are reflected by the values in the off-diagonal regions.

The comparison results in Figure. 6 show that the diagonal region of the model confusion matrix of CBAM is darker than that of the YOLOv5s matrix, indicating improved accuracy. The diagonal region of the confusion matrix of the proposed CBAM-C3L model has little change in color than that of the diagonal region of the YOLOv5s matrix, which indicates that the model proposed in this paper maintains the original accuracy in accurately predicting object categories.

The experimental results of **Figure 7**. show that the optimized C3light algorithm based on YOLOv5 has significantly improved performance compared with the original algorithm. In terms of mAP, the optimized algorithm has increased by 0.4% compared with the original algorithm. In comparison with other mainstream algorithms, the optimized algorithm has increased by 0.5%, the improvement is more obvious in complex algorithms.

### 4.2. Visualization of the Results

In order to more intuitively compare the effectiveness of the proposed algorithm and the comparison algorithm in detecting apple picking recognition, the same test set is used for apple detection, and part of the detection effects are

### 3.4. Evaluation Index Selection

The evaluation indexes include accuracy (Precision), recall rate (Recall), and mean average precision (mAP). Accuracy refers to the proportion of true positive examples among all the samples predicted as positive examples by the model., Recall refers to the proportion of positive examples that are correctly predicted as positive examples by the model among all actual positive example samples. These indexes can comprehensively evaluate the performance of the model in apple-picking recognition, where mAP is the main evaluation index, which reflects the average precision of the model in different IoU (Intersection

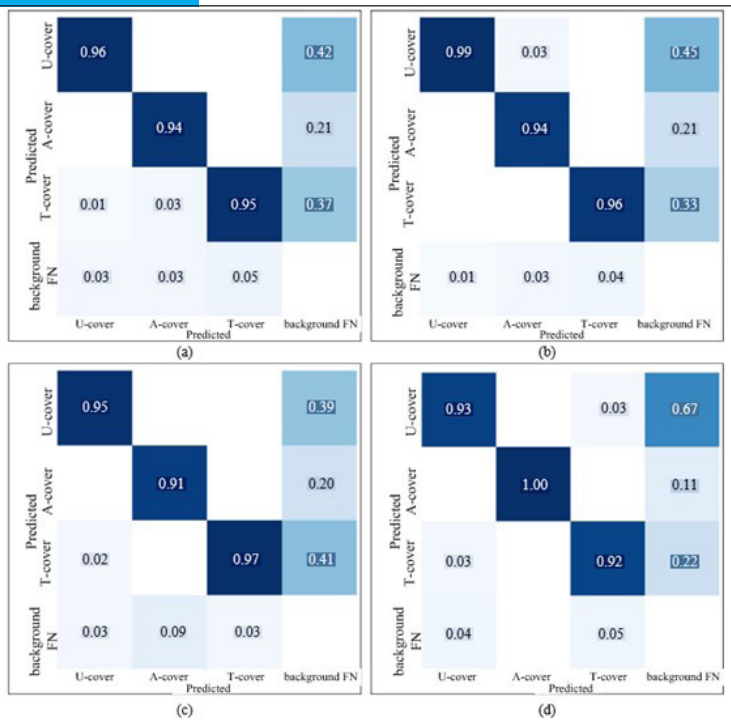


**Figure 6** Confusion matrix (a) 5s (b) 5s-C3L (c) 5s-CBAM (d) 5s-CBAM-C3L

shown in Figure. 8. As can be seen from the **Figure 8**, the original YOLOv5s model has the problem of missing detection, and the detection performance of some categories is particularly poor, while the C3-Light model combined with CBAM attention mechanism is superior to the original YOLOv5s model in missing detection effect. The application of C3-Light model in the two models can maintain a good state of detection accuracy, and the accuracy is reduced by 0.1% on average. When the apples are partially occluded, although the recognition accuracy is slightly affected, it still keeps great accuracy. This shows that the algorithm has a certain degree of robustness in dealing with complex orchard environments.

**4.3. Comparison of Microcontroller Adaptation with Different Models**

The parameters of the model in the running process can show the smoothness of the algorithm transplanted to the single-chip computer. Furthermore, the evaluation metrics also involve the model computational complexity (GFLOPs), the number of model parameters (Parameters), the memory using GPU (G-memory) and layers (Layers). The model's computational complexity serves as a crucial indicator for evaluating the efficiency of the model, while the number of model parameters is the key metric for assessing the capacity of the model. The detection speed, indicated by the frames per second, is utilized to appraise the model's ability to process image frames per second.

In order to verify the proposed C3-Light effect, another LSE attention mechanism is added in this part, which is improved by SE. As can be seen from the **Table 2**, the convolution layer is fixed and reduced by 55 layers. After replacing C3 model in 5s, CBAM and LSE models, GFLOPs decreased by 17.5%, 16.9% and 16.2%. Reduced by 10.6%, 12%, 11% on GPU memory calls. The reduction of these parameters can improve the smooth running of the model in the single-chip computer and reduce the delay.

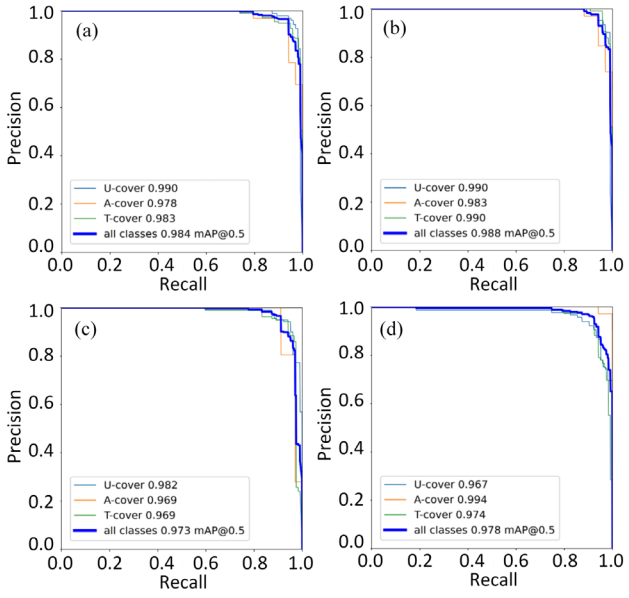


**Figure 7** P-R curve (a) 5s (b) 5s-C3L (c) 5s-CBAM (d) 5s-CBAM-C3L

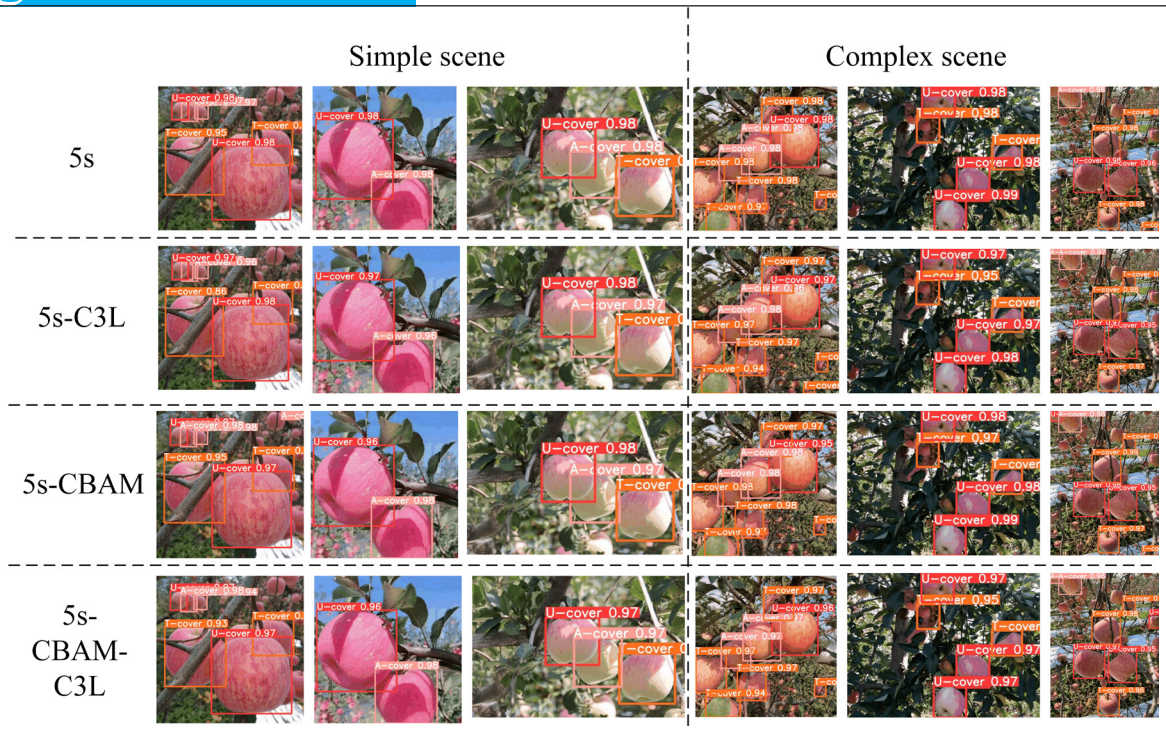


**Figure 8** Comparison of detection results

## 5. Conclusion

The innovation of this research lies in the comprehensive optimization of the C3light algorithm based on YOLOv5. By introducing the CBAM attention mechanism and improving the feature fusion method, the ability of the network to extract and utilize apple features is enhanced.

1. Compared with YOLOv5s, the C3-Light model has a slight improvement in mAP-0.5, and the improvement effect is better after adding CBAM attention mechanism, which can also maintain the effect in more complex network structures.

2. The C3-Light model can effectively reduce the use of hardware resources, which can reduce about 17% in GFLOPs and about 11% in GPU memory calls.

This research successfully optimizes the C3light lightweight algorithm based on YOLOv5 for apple-picking recognition. Through improvements in network structure and training strategy, the performance of the algorithm is significantly improved. The experimental results show that the optimized algorithm has better hardware resource scheduling ability than the original algorithm and other mainstream methods, and can maintain approximate accuracy, and has a certain robustness in different scenarios.

## 6. Future work

This research only indirectly analyzes the smoothness of the application on the single chip microcomputer through parameters, and then migrates the algorithm to the single chip microcomputer for experiments in the later stage, and initially chooses Raspberry PI, STM32 and Arduino for environment deployment and construction, which may face problems such as data migration between different systems and camera selection. After successful deployment, the algorithm will be improved for a variety of fruit and picking scenarios.

**Table 2**. Comparison of experimental results of different network models.

| Models | GFLOPs | Parameters | G-memory | Layers |
|---|---|---|---|---|
| 5s | 16.0 | 7027720 | 2.08 | 270 |
| 5s-C3L | 13.2 | 5970376 | 1.86 | 215 |
| 5s-CBAM | 16.0 | 7060586 | 2.09 | 281 |
| 5s-CBAM-C3L | 13.3 | 6003242 | 1.84 | 226 |
| 5s-LSE | 17.3 | 7093320 | 2.09 | 277 |
| 5s-LSE-C3L | 14.5 | 6035976 | 1.86 | 222 |

## Author Contributions

Kecheng SHAN and Quanhong FENG improved the algorithm and analysed the data, Xiaowei LI, Xianglong MENG, Hongkuan LYU (LV) and Chenfeng WANG collected the data of apple, Liyang MU drew the flow chart, Xin LIU did the typesetting and suggested changes.

## Acknowledgements

This work was supported by the Funding for Visiting Scholar project of ordinary undergraduate universities in Shandong Province in 2024, Youth Fund of Shandong Agriculture and Engineering University (QNKJZ202301), and Shandong Agriculture and Engineering University Start-Up Fund for Talented Scholars (BSQJ-202301). Kecheng SHAN and Quanhong FENG contributed equally to this work.

## References

[1] M. Hussain. "YOLOv1 to v8: Unveiling Each Variant-A Comprehensive Review of YOLO." *IEEE Access*. **2024**, *12*, 42816-42833.

[2] J. Tian, N. Ma, Y. Zhou, Y. Lv, W. Chi, L. Sun. "An Efficient End-to-End Lightweight Object Detection Method Based on YOLOv5 for Intelligent Sweeping Robots." *In IECON 2023-49th Annual Conference of the IEEE Industrial Electronics Society, IEEE*. **2023**, 10311823.

[3] J. An, Y. Lee, J. Kim, A. Jo, P. K. "Efficient Multi-Receptive Pooling YOLOv5 with Coordinate Attention Module for Object Detection on Drone."

*In 2023 IEEE 32nd International Symposium on Industrial Electronics*, IEEE, **2023**, 10227913.

[4] J. Shi, J. Yang, Y. Zhang, "Research on Steel Surface Defect Detection Based on YOLOv5 with Attention Mechanism." *Electronics,* **2022**, *11*, 3735.

[5] J. Cao, F. Han, M. Wang, X. Zheng, H. Gao. "A novel YOLOv5-based Hybrid Underwater Target Detection Algorithm Combining with CBAM and CloU." *Journal of Physics: Conference Series,* **2023**, 2560, 012001.

[6] L. Sun, G. Hu, C. Chen, H. Cai, C. Li, S. Zhang, J. Chen. "Lightweight Apple Detection in Complex Orchards Using YOLOV5-PRE." *Horticulturae,* **2022**, *8*, 1169.

[7] M. Li, J. S. Aviles, "Improvement of Remote sensing image target detection algorithm based on YOLO V5." *Journal of Physics: Conference Series,* **2023**, 2560, 012001.

[8] Z. Guo, C. Wang, G. Yang, Z. Huang, G. Li. "MSFT-YOLO: Improved YOLOv5 Based on Transformer for Detecting Defects of Steel Surface." *Sensors,* **2022**, *22*, 3467.

[9] Y. Zhang, Z. Guo, J. Wu, Y. Tian, H. Tang, X. Guo. "Real-Time Vehicle Detection Based on Improved YOLO v5." *Sustainability,* **2022**, *14*, 12274.

[10] Xu, N. Fang, N. Liu, F. Lin, S. Yang, J. Ning, "Visual recognition of cherry tomatoes in plant factory based on improved deep instance segmentation." *Computers and Electronics in Agriculture,* **2022**, *197*, 106991.

[11] J. Qi, X. Liu, K. Liu, F. Xu, H. Guo, X. Tian, M. Li, Z. Bao, Y. Li, "An improved YOLOv5 model based on visual attention mechanism: Application to recognition of tomato virus disease." *Computers and Electronics in Agriculture,* **2022***, 194, 106780.*

[12] Z. Guan, H. Li, Z. Zuo, L. Pan, "Design a robot system for tomato picking based on YOLOv5." *IFAC-PapersOnLine,* **2022**, *55*, 166-171.

[13] Y. Jian, Q. Zhen, Z. Yanjun, Y. Qin, H. Liao, "Real-time recognition of tomatoes in complex environments based on improved YOLOv4-tiny." *Transactions of the Chinese Society of Agricultural Engineering,* **2022**, *38*, 215-221.

[14] Z. Ning, L. Luo, J. Liao, H. Wen, H. Wei, Q. Lu, "Recognition and the optimal picking point location of grape stems based on deep learning." *Transactions of the Chinese Society of Agricultural Engineering,* **2021**, *37*, 222-229.

[15] R. Pérez-Zavala, M. Torres-Torriti, F. A. Cheein, G. Troni, "A pattern recognition strategy for visual grape bunch detection in vineyards." *Computers and Electronics in Agriculture,* **2018**, *151*, 136-149.

[16] H. Li, C. Li, G. Li, L. Chen, "A real-time table grape detection method based on improved YOLOv4-tiny network in complex background." *Biosystems Engineering,* **2021,** *212*, 347-359.

[17] Y. Jin, J. Liu, J. Wang, Z. Xu, Y. Yuan, "Far-near combined positioning of picking-point based on depth data features for horizontal-trellis cultivated grape." *Computers and Electronics in Agriculture,* **2022,** *194*, 106791.

[18] A. Olenskyj, B. Sams, Z. Fei, "End-to-end deep learning for directly estimating grape yield from ground-based imagery." *Computers and Electronics in Agriculture,* **2022**, *198*, 107081.

[19] L. Yu, J. Xiong, X. Fang, Z. Yang, Y. Chen, X. Lin, S. Chen, "A litchi fruit recognition method in a natural environment using RGB-D images." *Biosystems Engineering,* **2021**, *204*, 50-63.

[20] C. Liang, J. Xiong, Z. Zheng, Z. Zhong, Z. Li, S. Chen, Z. Yang, "A visual detection method for nighttime litchi fruits and fruiting stems." *Computers and Electronics in Agriculture,* **2020**, *169*, 105192.

[21] Y. Osako, H. Yamane, S.-Y. Lin, -A. Chen, R. Tao, "Cultivar discrimination of litchi fruit images using deep learning." *Scientia Horticulturae,* **2020**, *269*, 109360.

[22] S. Sun, M. Jiang, D. He, Y. Long, H. Song, "Recognition of green apples in an orchard environment by combining the Grab Cut model and Ncut algorithm." *Biosystems Engineering,* **2019**, *187*, 201-213.

[23] D. Wang, D. He, "Fusion of Mask RCNN and attention mechanism for instance segmentation of apples under complex background." *Computers and Electronics in Agriculture,* **2022**, *196*, 106864.

[24] H. Zhao, Y. Qiao, H. Wang, Y. Yue, "Apple fruit recognition in complex orchard environment based on improved YOLOv3." *Transactions of the Chinese Society of Agricultural Engineering.* **2021**, *37*, 127-135.

[25] L. He, W. Fang, G. Zhao, Z. Wu, L. Fu, R. Li, Y. Majeed, J. Dhupia, "Fruit yield prediction and estimation in orchards: A state-of-the-art comprehensive review for both direct and indirect methods." *Computers and Electronics in Agriculture,* **2022**, *195*, 106812.

[26] R. Syazwani, H. Asraf, M. Amin, N. Dalila. "Automated image identification, detection and fruit counting of top-view pineapple crown using machine learning." *Alexandria Engineering Journal,* **2022**, *61*, 1265-1276.

[27] L. Fu, J. Duan, X. Zou, G. Lin, S. Song, B. Ji, Z. Yang, "Banana detection based on color and texture features in the natural environment." *Computers and Electronics in Agriculture,* **2019**, *167*, 105057.