

Research on Wheat Field Irrigation Identification Method Based on Multi-Model Verification of 532nm Optical Characteristics

<https://doi.org/10.63174/xdi.ANDK2599>

Received: 12 May 2025

Accepted: 16 May 2025

Published: 17 May 2025

Open Access



Chenfeng Wang^{1,+}, Kecheng Shan^{1,+}, Xin Liu^{1,2}, Hao Wang¹, Jiaqi Zhu¹, Bowen Wang¹, Zhicheng Zhang^{1,3,*}

Abstract: This study employs multiple models, including a baseline model, a lightweight model, and models incorporating attention mechanisms, to detect the irrigation status of wheat fields using the optical characteristics provided by a 532nm optical filter. Experimental results demonstrate the high precision and efficiency of the proposed approach, with the 5s-CBAM model achieving a mAP@0.5 of 94.2% on edge devices, while the lightweight 5s-C3Light model maintained an accuracy of 89.7% at an inference speed of 28 FPS. All models exhibited stable and reliable detection capabilities across various wheat growth stages, demonstrating their robustness in meeting real-world agricultural monitoring needs. Furthermore, the findings highlight the 532nm optical filter as a promising tool for advancing irrigation management in wheat fields and enabling intelligent agricultural practices, paving the way for its broader adoption in smart agricultural systems.

1. Introduction

As one of the most widely distributed, second-highest-yielding, and nutritionally vital cereal crops globally, wheat^[1-2] holds pivotal importance in agriculture. Water stands as a critical factor in wheat cultivation: optimal irrigation not only prevents root hypoxia and diseases caused by waterlogging but also mitigates growth inhibition and yield losses from water deficit, thereby enhancing water-use efficiency. The timing and volume of irrigation require dynamic adjustment based on wheat growth stages (e.g., tillering, heading) and real-time weather conditions. However, traditional manual irrigation methods suffer from inefficiency and high labor costs. Consequently, rapid and accurate diagnosis of wheat water requirements represents a key technological bottleneck for advancing smart agriculture (e.g., unmanned farms). To address the demand for real-time monitoring via portable field devices, developing an efficient and reliable algorithm for wheat water deficit detection is of urgent practical significance.

With the rapid development of artificial intelligence and deep learning, computer vision has demonstrated remarkable advantages in agricultural applications. Deep learning-based object detection algorithms are mainly divided into single-stage algorithms (such as the YOLO series) and two-stage algorithms^[3-4] (such as Faster R-CNN). Among them, the YOLO^[5-12] (You Only Look Once) algorithm, with its unique single-stage detection architecture, achieves an excellent balance between detection speed and accuracy, making it one of the most practical target detection solutions for smart agriculture applications. The practical application of this algorithm has been validated in multiple agricultural scenarios. Yan et al.^[8] improved the YOLOv5s model to achieve multi-target detection in apple harvesting scenarios. Fernando et al.^[13] confirmed that YOLOv5s delivered optimal comprehensive performance in weed detection in wheat fields. Turab^[14] applied YOLOv5 to walnut quality sorting, achieving 98% accuracy. Santana et al.^[15] developed a eucalyptus seedling detection model based on YOLOv8/YOLOv5, providing technical support for automated irrigation. These application cases fully demonstrate the practical value of the YOLO series algorithms in agricultural production. Through continuous optimization of network structures and detection strategies, the YOLO series algorithms continue to improve detection accuracy while maintaining real-time performance, providing reliable technical support

for the development of agricultural intelligence. In the future, with further algorithm optimization, its application prospects in the agricultural field will become even broader.

With the increasing emphasis on artificial intelligence technologies, numerous deep learning networks have been applied to wheat detection, demonstrating promising research outcomes. Zhao et al.^[16] employed an improved YOLOv5 network for wheat head detection using UAV imagery, achieving an average accuracy of 94.1%. Li et al.^[17] proposed a method based on Faster R-CNN and RetinaNet networks to identify wheat heads across different growth stages under varying conditions, attaining an 82% detection accuracy on the Global WHEAT dataset. Yang et al.^[18] developed a wheat head detection approach integrating YOLOv4 with a Convolutional Block Attention Module (CBAM), which demonstrated detection accuracies of 94%, 96%, and 93% under three distinct density conditions on two public datasets (WEDD and GWHDD).

To address the precision monitoring requirements for intelligent irrigation in wheat fields, this study systematically compares four improved target detection algorithms based on the YOLOv5s architecture. The baseline YOLOv5s model and its three enhanced versions—the lightweight 5s-C3Light model^[19], the coordinate attention-incorporated 5s-CA model, and the dual-path attention-integrated 5s-CBAM model—each demonstrate distinct performance characteristics in irrigation scenarios. Through optimized network structures and feature extraction approaches, these improved models exhibit different advantages in terms of detection accuracy and operational efficiency. Specifically, the lightweight model significantly enhances computational efficiency, while the two attention mechanism models improve detection stability in complex field environments by strengthening feature representation capabilities. This study provides theoretical guidance for algorithm selection in various application scenarios through systematic comparative analysis, offering important insights for achieving intelligent irrigation management in precision agriculture.

2. Materials of data

2.1 Data Source

The experimental data were collected from wheat fields in Donglandun Village, Junan County, located in Linyi, Shandong Province (35°11' N,

¹ Machinery and Electronics Engineering College, Shandong Agriculture And Engineering, University, Zibo, Shandong Province, 255300, China

² School of Physics and Electronics, Shandong Normal University, Jinan, Shandong Province, 250358, China

³ Hongdian (Shandong) Digital Technology Co., Ltd, Jinan, Shandong Province, 250199, China

⁺ Made the same contribution

^{*} Corresponding Author: drzhang19910908@163.com

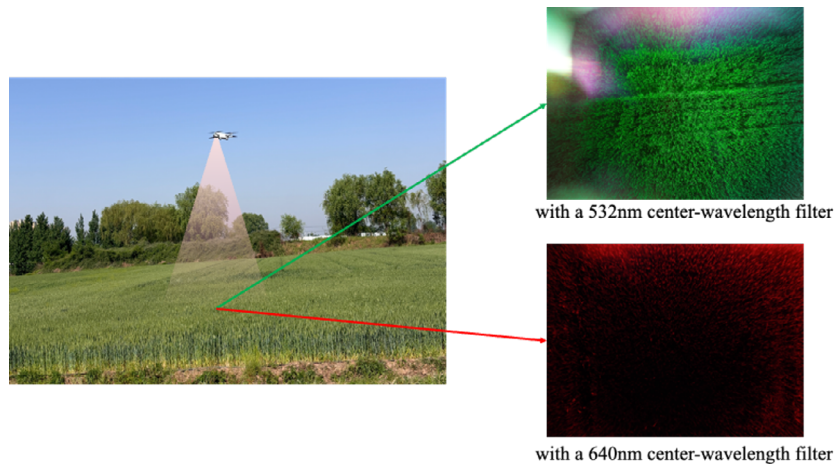


Figure 1. (a) UAV acquisition schematic diagram. (b) Original image under 532nm filter. (c) Comparative image under 640nm filter.

118°51' E). The wheat variety was Linmai 9, sown in 10 October 2024 with a growth cycle of approximately 240 days. The total area of the wheat field was 1,333.33 square meters. Wheat field images were captured using a DJI Mini 3 drone (SZ DJI Technology Co., Shenzhen, China), equipped with a 1/1.3-inch CMOS sensor. The lens had a field of view of 82.1 degrees and 48 million effective pixels, with images saved in JPEG/DNG (RAW) format. The drone is equipped with four narrow-band filters of the GCC-2010 series of Daheng Optics with a peak transmittance of 85% and a center wavelength of 405nm, 488nm, 532nm and 640nm, respectively. In the actual shooting, it was found that the images taken by the 405nm and 488nm center wavelength filters were not obvious to the naked eye, and finally two filters with the center wavelength of 532nm and 640nm were selected for the experiment, as shown in **Figure**

1 The flight altitude was set at approximately 5 meters above the wheat ears. The flight mission was conducted between 9:00 AM and 11:30 AM on April 30. The weather was clear with light breezes. Data collection was intentionally conducted under stable weather conditions to minimize environmental variability. Future studies will expand datasets to cover diverse regions and weather scenarios.

2.2 Data Annotation

The initially collected photos underwent preliminary screening, resulting in 117 original images. To ensure dataset diversity and meet application requirements, some images were augmented through horizontal and vertical rotation, while others—captured at varying angles—were rotated and cropped using alternative methods. Ultimately, this process yielded 351 images.

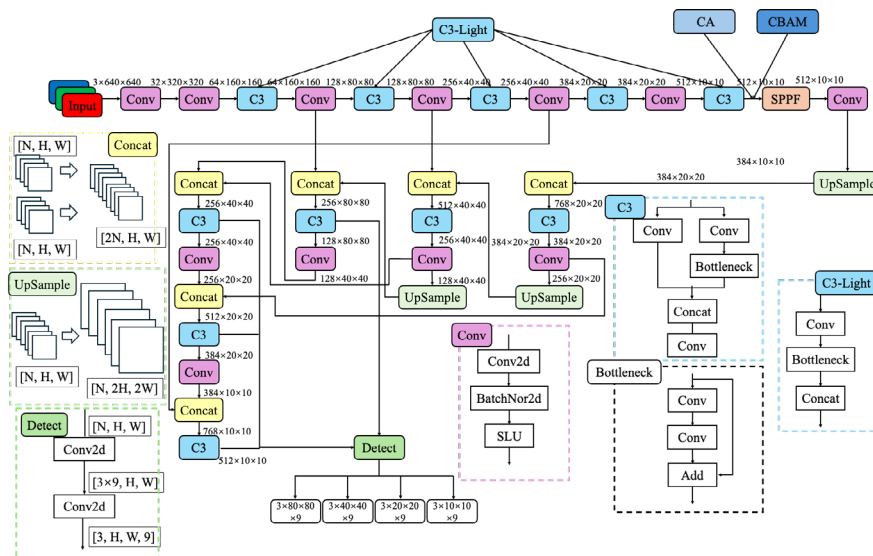


Figure 2. Network structure diagram of YOLOv5 and C3-Light

The annotation work was conducted using the open-source Python-based tool LabelImg. By selecting the wheat field regions with bounding boxes and choosing the corresponding category labels, the software automatically generated annotation files. The annotation files and original images were then integrated according to the YOLOv5 algorithm format requirements and divided into training, validation, and test sets in an 8:1:1 ratio. The data was categorized into four types, named “Green-N”, “Green-Y”, “Red-N”, and “Red-Y”, where “N” indicates non-irrigated, and “Y” indicates irrigated. The entire process strictly adhered to YOLOv5’s best practices to ensure the data quality met the training requirements of deep learning models.

3. Method and Experimental setting

3.1 Proposed Method

The YOLO algorithm has now evolved to its eleventh version (v11). The YOLO framework offers a diverse range of models with varying architectures

and computational requirements, including the n, s, m, l and x variants, each characterized by specific scaling parameters. While newer versions demonstrate improved accuracy, these gains come at the cost of increased computational demands. Currently, YOLOv5 remains the most widely adopted version due to its optimal balance between speed, accuracy, and hardware requirements. The YOLOv5 6.1 iteration introduced significant improvements over version 5.0, most notably replacing the Focus module with a more efficient Conv module, which enhanced detection speed while maintaining accuracy. Our research builds upon this optimized YOLOv5 6.1 architecture, whose structure is illustrated in **Figure 2** YOLOv5 was prioritized over newer versions (e.g., v8) due to its proven efficiency in edge devices, critical for field deployment. Comparisons with non-YOLO models (e.g., EfficientDet) are planned for future work.

To better adapt to the wheat field monitoring scenario, we incorporated two attention mechanisms into YOLO: CBAM^[20] (Convolutional Block Attention

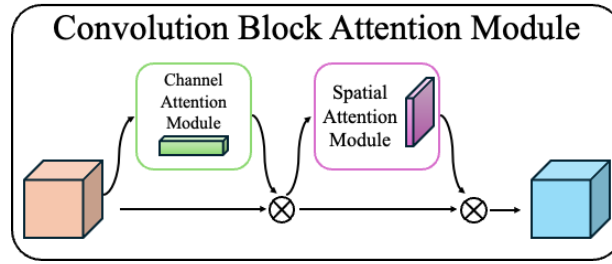


Figure 3. The overview of CBAM.

Module) and CA^[21] (Coordinate Attention). CBAM is a dual-path attention mechanism that combines channel and spatial attention. Unlike SENet which only considers channel attention, CBAM achieves better performance by processing features through parallel global average pooling (GAP) and global max pooling (GMP) operations, followed by shared MLP layers with sigmoid activation for channel attention, while spatial attention is generated via channel-wise pooling and 7×7 convolution. The Coordinate Attention (CA) block is a computational unit that enhances feature representations in mobile networks. It processes input feature tensor $X = [x_1, x_2, \dots, x_C] \in \mathbb{R}^{C \times H \times W}$ by decomposing 2D global pooling into separate X and Y-axis 1D operations. These coordinate features are concatenated and processed through 1×1 convolution to generate directional attention weights, producing enhanced output $Y = [y_1, y_2, \dots, y_C]$ with minimal computational overhead while effectively capturing long-range dependencies.

3.2 Experimental Equipment and Experiment Setting

This experiment is lightweight based on the YOLOv5 6.1 version. 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 hardware used was a HP Z series notebook. The main hardware parameters are shown in **Table 1**, and the software parameters and training parameters are shown in **Table 2**.

Table 1. The main hardware parameters

Parameter	Value
GPU	NVIDIA GeForce RTX 3050 Ti Laptop GPU
System	Windows 11
CPU	Intel i5 12500H
RAM	16G
SSD	512G+1T
Graphics card	4096MiB

Table 2. The software parameters and training parameters

Parameter	Value
Epochs	200
Batch size	4
Input image size	640

The evaluation indexes include accuracy (Precision), recall rate (Recall), and mean average precision (mAP), as shown in Equations (1)-(4).

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

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

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

$$AP_i = \int_0^1 P(R) dR \quad (4)$$

where Precision refers to the proportion of correctly classified positive instances among all data points predicted as positive by the classifier. In other words, it measures how many of the predicted positive samples are truly positive. Recall represents the proportion of correctly predicted positive instances out of all actual positive data points. It indicates how many of the true positive samples are correctly identified as positive. The results of Precision and Recall depend on three key parameters: True Positives (TP), False Positives (FP), and False Negatives (FN). TP refers to the number of actual positive samples that are correctly predicted as positive by the model.

FP refers to the number of actual negative samples that are incorrectly predicted as positive by the model. FN refers to the number of actual positive samples that are incorrectly predicted as negative by the model. AP (Average Precision) denotes the precision for individual categories. The mAP (mean Average Precision) is defined as the average of the AP values across all categories, reflecting its precision at different IoU (Intersection over Union) thresholds. n is the number of samples.

To minimize hardware constraints and impacts on the research, a lightweight detection network was introduced to reduce computational stacking during operation. In this study, we propose modifying the C3 module using the C3-Light module based on the fundamental YOLOv5 architecture, as shown in **Figure 2**. This improvement reduces the number of network layers in the initial processing stage after image input, thereby accelerating processing speed.

4. Experimental Results Analysis

The detection performance of different models was systematically evaluated through confusion matrix analysis. As shown in **Figure 4**, the rows of the confusion matrix represent ground-truth categories while the columns correspond to predicted classification results, with True values on the x-axis and Precision metrics on the y-axis. The diagonal elements of the matrix visually quantify the correct classification rates, while the off-diagonal regions reveal error distribution characteristics. Analysis reveals that **Figures 4(a)-(d)** all exhibit similar characteristics: the diagonal regions corresponding to green display darker coloration than those representing red. All four models demonstrate superior performance, attributable to their enhanced feature discrimination capabilities. Notably, the detection results (Green-Y) for irrigated wheat fields captured through a 532nm center-wavelength filter reach a remarkable 1.00 in the C3L, CA, and CBAM models. Particularly noteworthy is how the off-diagonal patterns provide crucial insights into specific inter-class confusion scenarios, offering valuable guidance for subsequent model refinement. All analyses employed standardized color intensity measurements based on reference values to ensure objective performance comparisons across different model architectures.

The experimental results demonstrate that all four improved algorithms achieve outstanding performance in terms of mean Average Precision (mAP). Specifically, the Green-Y category attains remarkably high mAP values of 99.4% and 99.5%, while the Red-Y category shows mAP values of 91.3%, 89.4%, 90.8%, and 87.4% respectively. Notably, for both types of optical filters used, the irrigated fields consistently demonstrate significantly higher mAP than non-irrigated fields. This performance gap primarily stems from the less distinct detection features present in non-irrigated wheat fields. As illustrated in **Figure 5**, the stable performance exhibited by all algorithm variants indicates strong potential for practical agricultural applications.

To more intuitively compare the effectiveness of the four algorithms in detecting irrigation status in wheat fields, we conducted evaluations using the same test set, with partial detection results shown in **Figure 6**. As observed in the third and fourth columns of **Figure 6**, all four models exhibit missed detections and relatively low accuracy when processing images captured with the 640nm center-wavelength filter. In contrast, results obtained with the 532nm center-wavelength filter (columns 1-2 in **Figure 6**) show that while detection accuracy remains unstable for non-irrigated fields due to their less distinct features, the models achieve significantly higher accuracy (95%, 87%, 94%, and 89% respectively) for irrigated wheat fields. All results demonstrate that the four algorithms achieve higher detection accuracy and exhibit certain robustness when analyzing wheat field images captured with the 532nm center-wavelength filter.

The experimental results demonstrate that the proposed C3-Light module achieves significant optimization across multiple performance metrics while maintaining model effectiveness. As shown in **Table 3**, compared to the baseline 5s model with 213 layers and 7,020,913 parameters, our C3-Light

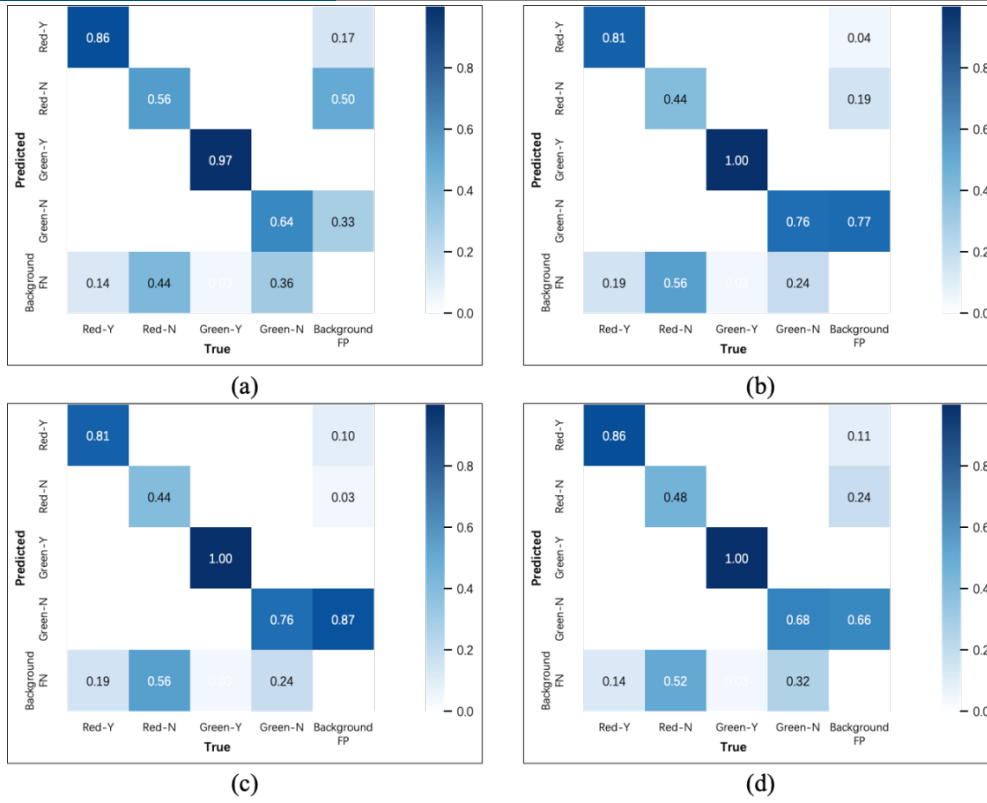


Figure 4. Confusion Matrix of 5s(a), 5s-C3Light(b), 5s-CA(c) and 5s-CBAM(d).

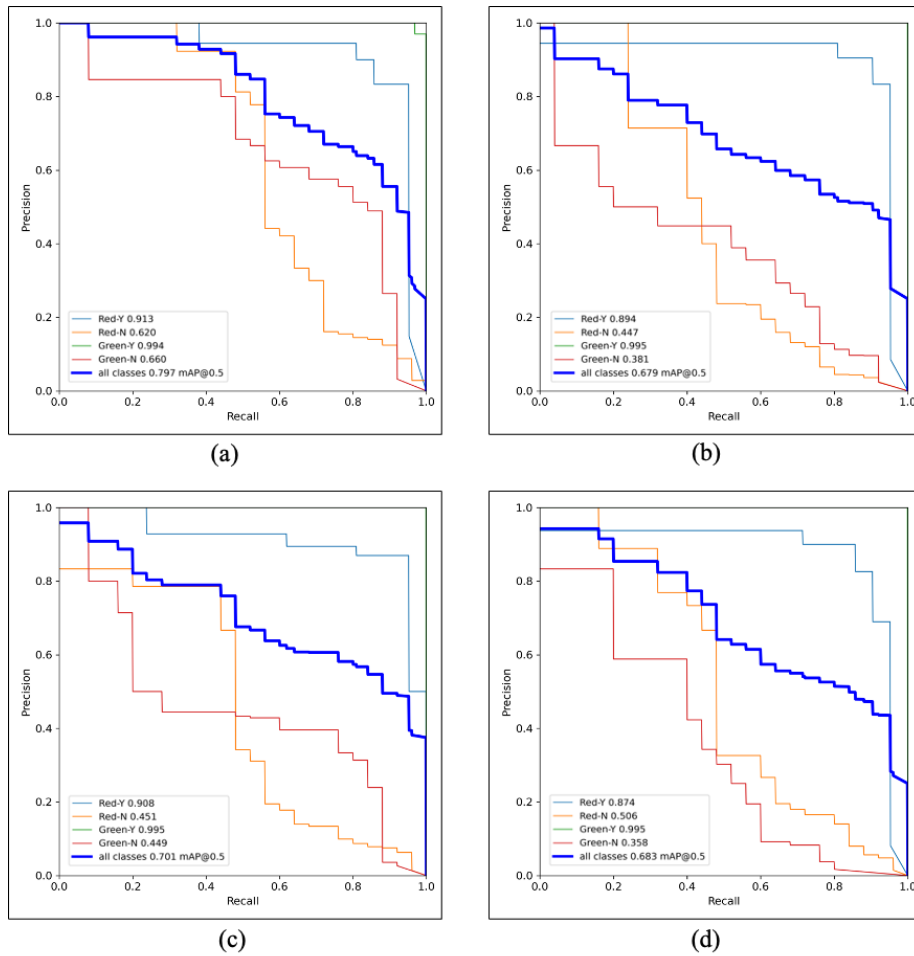


Figure 5. The PR curve of 5s(a), 5s-C3Light(b), 5s-CA(c) and 5s-CBAM(d).

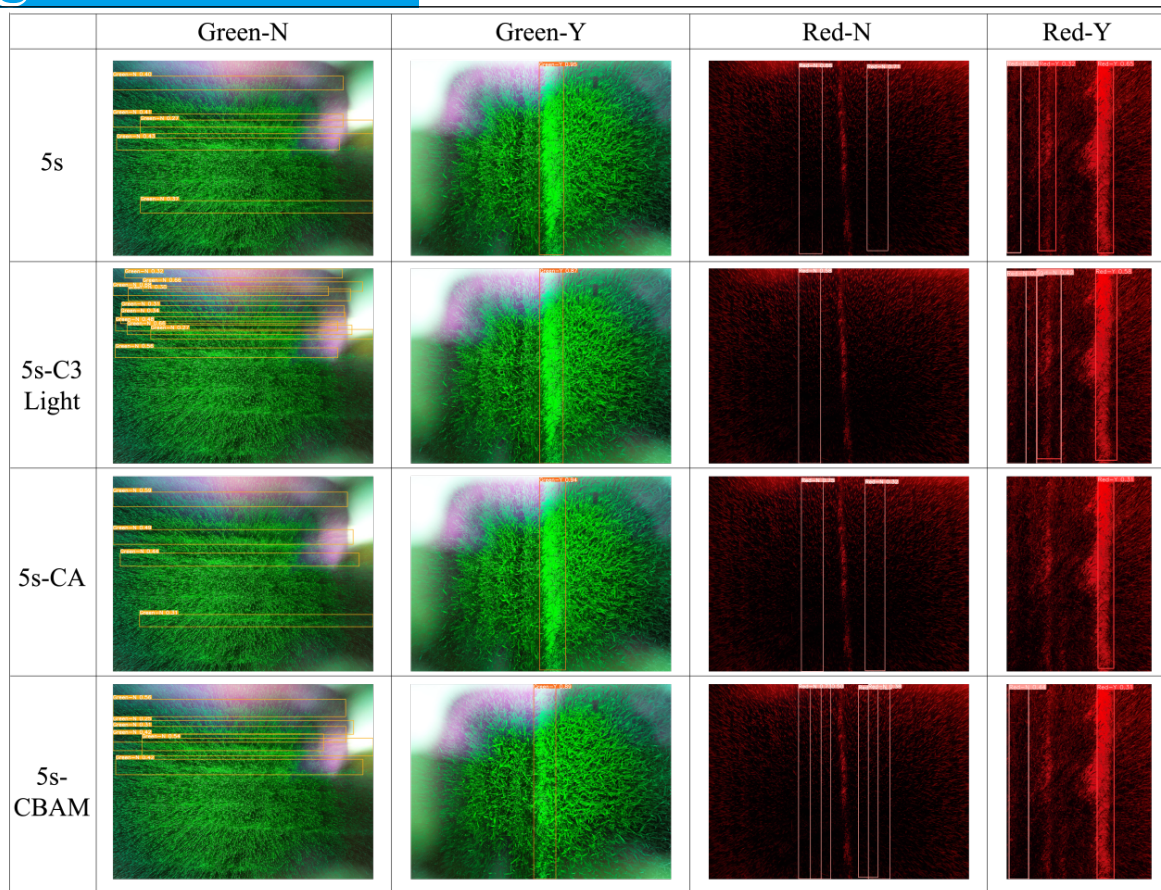


Figure 6. Comparison of detection results.

implementation reduces the network depth to 185 layers (13.1% decrease) and parameter count to 5,995,761 (14.6% reduction), while simultaneously lowering computational complexity from 15.8 to 13.7 GFLOPs (13.3% improvement). Notably, when comparing with other attention-enhanced variants, the CA module increases layers to 223 (+4.7%) and the CBAM configuration reaches 224 layers (+5.2%), both maintaining the baseline's 15.8 GFLOPs computational complexity with only marginal parameter increases. These structural optimizations in the C3-Light model directly contribute to improved operational fluency in embedded systems, as evidenced by reduced processing latency and more efficient resource utilization. The results demonstrate that while the 5s model, 5s-CA model, and 5s-CBAM model all exhibit good performance, the 5s-C3Light model more successfully balances model compression with performance preservation. It provides a practical solution for deployment on resource-constrained hardware platforms while maintaining the flexibility to incorporate different attention mechanisms when needed.

5. Conclusion

This study confirms the 532nm filter's superior performance for wheat field irrigation detection using YOLOv5s-based models. Experimental results demonstrated robust accuracy, with mAP reaching 99.5% for irrigated fields, significantly outperforming the 640nm filter. The optimized C3-Light model achieved a 13.3% reduction in computational cost while maintaining high detection performance. These findings establish 532nm as the optimal wavelength for reliable irrigation monitoring, supported by consistent model performance. Future work could explore multi-spectral approaches to further enhance detection capabilities.

Author Contributions

Chenfeng Wang, Kecheng Shan, Jiaqi Zhu and Bowen Wang collected the data, Hao Wang analyzed the data and generated the results, Chenfeng Wang, Kecheng Shan and Xin Liu did the writing and typesetting, Zhicheng Zhang came up with the research idea and suggested changes.

Acknowledgements

This work was supported by the funding of Shandong Agriculture and

Engineering University Start-Up Fund for Talented Scholars (BSQJ-202301). Chenfeng Wang and Kecheng Shan contributed equally to this work.

References

- [1] J. Sun, K. F. Yang, C. Chen, J. F. Shen, Y. Yang, X. H. Wu, T. Norton. "Wheat headcounting in the wild by an augmented feature pyramid networks-based convolutional neural network." *Computers and Electronics in Agriculture*, **2022**, *193*, 106705.
- [2] H. C. Zang, Y. J. Wang, L. Y. Ru, M. Zhou, D. D. Chen, Q. Zhao, J. Zhang, G. Q. Li, G. Q. Zheng. "Detection method of wheat spike improved yolov5s based on the attention mechanism." *Frontiers in Plant Science*, **2022**, *13*, 993244.
- [3] L. Du, R. Zhang, X. Wang. "Overview of two-stage object detection algorithms." *J. Phys.: Conf. Ser.* **2020**, *1544*, 1, 012033.
- [4] S. Ren, K. He, R. Girshick, J. Sun. "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks." Presented in Advances in Neural Information Processing Systems, **2015**, *28*.
- [5] J. Redmon, S. Divvala, R. Girshick, A. Farhadi. Presented in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). You Only Look Once: Unified, Real-Time Object Detection. May **2016**.
- [6] J. Redmon, A. Farhadi. "YOLOv3: An Incremental Improvement." *arXiv e-prints arXiv:1804.02767*, **2018**.
- [7] A. Bochkovskiy, C. Y. Wang, H. Y. M. Liao. "YOLOv4: Optimal Speed and Accuracy of Object Detection." *arXiv preprint arXiv:2004.10934*, **2020**.
- [8] B. Yan, P. Fan, X. Lei, Z. Liu, F. Yang. "A Real-Time Apple Targets Detection Method for Picking Robot Based on Improved YOLOv5." *Remote Sensing*, **2021**, *13*, 9, 1619.
- [9] Z. Ge, S. Liu, F. Wang, Z. Li, J. Sun. "YOLOX: Exceeding YOLO Series in 2021", *arXiv preprint arXiv:2107.08430*, **2021**.
- [10] C. Y. Li, L. L. Li, H. L. Jiang, K. H. Weng, Y. F. Geng, L. Li, Z. D. Ke, Q. Y. Li, M. Cheng, W. Q. Nie, Y. D. Li, B. Zhang, Y. F. Liang, L. Y. Zhou, X. M. Xu, X. X. Chu, X. M. Wei, X. L. Wei. "YOLOv6: A Single-Stage Object Detection Framework for Industrial Applications." *arXiv preprint arXiv:2209.02976*, **2022**.
- [11] C. Y. Wang, A. Bochkovskiy, H. Y. M. Liao. "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors." presented

at Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, **2023**.

[12] C. Y. Wang, I. H. Yeh, H. Y. M. Liao. "YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information." *European conference on computer vision*, **2024**.

[13] F. J. Pérez-Porras, J. Torres-Sánchez, F. López-Granados, F. J. Mesas-Carrascosa. "Early and on-ground image-based detection of poppy (*Papaver rhoeas*) in wheat using YOLO architectures." *Weed Sci.*, **2023**, *71*, 1, 50-58.

[14] T. Selçuk. "A Raspberry Pi-Guided Device Using an Ensemble Convolutional Neural Network for Quantitative Evaluation of Walnut Quality." *Traitement du Signal*, **2023**, *40*, 5, 2283-2289.

[15] J. S. Santana, D. S. M. Valente, D. M. Queiroz, A. L. F. Coelho, I. A. Barbosa, A. Momin. "Automated Detection of Young Eucalyptus Plants for Optimized Irrigation Management in Forest Plantations." *AgriEngineering*, **2024**, *6*, 4, 3752.

[16] J. Zhao, X. Zhang, J. Yan, X. Qiu, X. Yao, Y. Tian, Y. Zhu, W. Cao. "A wheat spike detection method in UAV images based on improved yolov5." *Remote Sensing*, **2021**, *13*, 16, 3095.

[17] J. Li, C. Li, S. Fei, C. Ma, W. Chen, F. Ding, Y. Wang, Y. Li, J. Shi, Z. Xiao. "Wheat ear recognition based on retinaNet and transfer learning." *Sensors*, **2021**, *21*, 14, 4845.

[18] B. Yang, Z. Gao, Y. Gao, Y. Zhu. "Rapid detection and counting of wheat ears in the field using Yolov4 with attention module." *Agronomy*, **2021**, *11*, 6, 1202.

[19] K. Shan, Q. Feng, X. Li, X. Meng, H. Lyu, C. Wang, L. Mu, X. Liu. "C3-Light Lightweight Algorithm Optimization under YOLOv5 Framework for Apple-Picking Recognition." **2025**, *1*, 1, 4-4.

[20] S. Woo, J. Park, J. Y. Lee, I. S. Kweon. "CBAM: Convolutional Block Attention Module." Presented in the European conference on computer vision (ECCV), July **2018**, 3-19.

[21] Q. Hou, D. Zhou, J. Feng. "Coordinate Attention for Efficient Mobile Network Design." Presented in the IEEE/CVF conference on computer vision and pattern recognition, March **2021**, 13713-13722.