

## OVERCOMING OCCLUSION AND OVERLAPPING IN WHEAT SPIKE DETECTION: A YOLOV5S-BASED FRAMEWORK FOR ACCURATE PHENOTYPIC EVALUATION

WANG-YAN-JING<sup>1</sup>, YING ZHANG<sup>3</sup>, LINQING WANG<sup>1</sup>, JING LIU<sup>1</sup>, YUNHUI BAI<sup>1</sup>, HECANG ZANG<sup>2\*</sup>

<sup>1</sup>College of Life Sciences, Zhengzhou Normal University, Zhengzhou 450044, China

<sup>2</sup>Institute of Agricultural Information Technology, Henan Academy of Agricultural Sciences, Zhengzhou 450002

<sup>3</sup>Zhengzhou High tech Industrial Development Zone Second Experimental Middle School, Zhengzhou 450003

\*Corresponding author's email: [Zanghechang@163.com](mailto:Zanghechang@163.com)

### Abstract

In different stages of wheat breeding, timely and accurate counting of wheat spikes is essential for phenotypic evaluation and early yield prediction. Manual spike counting, however, is labor-intensive and time-consuming for wheat-producing regions globally. Aiming at the practical challenges including small spike size, dense distribution, and frequent spike occlusion, this study used the wheat cultivar Zhengmai 1860 as experimental material. Wheat spike images were acquired with mobile devices under potted conditions to build a custom image dataset, and a YOLOv5s object-detection model was developed for wheat-spike counting. The trained YOLOv5s model yielded a precision of 86.7%, recall of 94%, and mean average precision of 85.7%, showing favorable detection performance and generalization capacity. Compared with conventional manual counting, the proposed model delivers superior detection performance, enables rapid wheat-spike counting under potted conditions, and offers an effective technical option for automated wheat phenotypic assessment and yield-related estimation.

**Key words:** Wheat phenotype; wheat spike counting; object detection; YOLOv5s

### Introduction

As a major cereal crop, wheat is widely cultivated globally and provides approximately 20% of the world's total dietary caloric intake (Chauhan *et al.*, 2020). In 2021, the global wheat-cultivated area and total wheat output reached approximately 223 million hectares and 776 million tons, respectively (Wang *et al.*, 2025). However, global food security is currently facing severe challenges. It is projected that by 2050, global food demand will need to increase by 69% to meet the unprecedented consumption needs of a global population of 9.7 billion (Jiao, 2024). Factors including the aging population structure and crop cultivated area can affect the formation of wheat yield (Su *et al.*, 2023). In practical agricultural production, wheat spike counting is mainly completed through manual field enumeration. This method suffers from high economic cost, substantial labor input, low counting efficiency, and counting results are easily affected by subjective factors. Timely and accurate estimation of wheat spike count provides critical scientific evidence for wheat yield prediction and variety screening.

In recent years, with the rapid development of artificial intelligence, deep learning-based wheat spike image detection technology has achieved remarkable progress (Madec *et al.*, 2019; He *et al.*, 2020). It has reached the state-of-the-art level in both detection accuracy and inference speed (Bar-Hillel, 2021; Lu *et al.*, 2022). The SSD and YOLO series algorithms belong to the category of single-stage object detection algorithms (Wei *et al.*, 2016) Zhao *et al.*, (2021) proposed an improved method based on YOLOv5, which can detect wheat spikes from UAV images with an average detection accuracy of 94.1%, representing

an increase of 10.8 percentage points compared with the original YOLOv5 model (Zhao *et al.*, 2021). The proposed method addresses the issues of false detection and missed detection of wheat spikes under closed-canopy conditions. Furthermore, Fernandez-Galego *et al.* presented an automatic quantification method for wheat spike count in field environments (Fernandez-Gallego *et al.*, 2018). Lu *et al.* developed a mobile application for the detection and statistics collection of pathogen-affected wheat ears, with an overall detection accuracy reaching 96.6% (Lu *et al.*, 2017). Hasan *et al.* further employed Convolutional Neural Networks (CNN) to implement the detection, statistical analysis and evaluation of wheat spike count, and applied the proposed method to wheat yield prediction (Hasan *et al.*, 2019). Zhang *et al.* proposed a high-fidelity wheat spike detection scheme constructed on the YOLO network framework (Costin-Anton *et al.*, 2009). Sadeghi Tehran *et al.* developed an automated tool named DeepCount, which can realize the detection, recognition and counting of wheat targets in digital images (Sadeghi-Tehran *et al.*, 2019). Fernandez Galego *et al.* constructed a wheat spike image processing platform that can work under field natural conditions through the deployment of thermal imaging equipment (Fernandez-Gallego *et al.*, 2019). Khaki *et al.* constructed the Wheat Net structural framework, which enables accurate and efficient calculation of wheat spike count within images (Khaki *et al.*, 2022).

In recent years, with the continuous iteration and update of the YOLO algorithm series, researchers have further explored the application potential of the latest YOLO models in wheat spike detection tasks. On the basis of improving the YOLOv8 algorithm, Li *et al.* proposed a lightweight wheat

counting model named PSD5-YOLOv8, which effectively solves the detection challenges brought by the dense distribution, small individual size and high degree of occlusion of wheat spikes in UAV-acquired images (Li *et al.*, 2025). This model attained a mean average precision of 96.5%, while simultaneously reducing the overall parameter quantity by 40.6%. On this basis, Lin *et al.* put forward a lightweight detection method termed YOLOv8-FDA, which compresses the model volume to 2.96 MB and enables real-time wheat spike detection and counting for UAV images (Lin *et al.*, 2025). Han *et al.* conducted a systematic comparative analysis on the detection performance of four YOLO variants, namely v5, v6, v8, and v10, in UAV video stream data. The results indicated that YOLOv8 maintains the highest detection stability across conditions of different wheat planting densities and different growth stages, and the coefficient of determination  $R^2$  between the detected wheat spike count and the ground-truth count reached a maximum of 0.90 (Han *et al.*, 2026). To address the detection difficulties caused by small targets and overlapping wheat spikes in dense planting scenes, the DSCA-YOLO model integrates a coordinate attention module and depth wise separable convolution. For wheat spike number prediction at the filling stage and maturity stage, this model obtained  $R^2$  values of 0.9847 and 0.9836 respectively, with a detection speed reaching 119.5 FPS. In current research, wheat spike detection technology has gradually expanded its application scope from basic spike counting to fine-grained phenotypic analysis and interdisciplinary research applications. Sun *et al.* constructed an artificial intelligence-driven wheat phenotyping pipeline named SpikePheno, which combines the ResNet50-UNet network and YOLOv8x-seg model. Through this method, high-throughput extraction of 45 spike and grain traits from 221 wheat varieties was realized (Sun *et al.*, 2025).

Although existing studies have achieved effective feature extraction from wheat spike images, during actual wheat production, factors including high planting density, leaf occlusion, and mutual overlapping between wheat spikes still lead to problems such as false detection and missed detection in existing methods, which further increases the difficulty and complexity of the wheat spike detection task. To effectively address these existing problems, this study optimizes the YOLOv5s model to construct an automatic detection method for wheat spikes, which can accurately identify the position and category of wheat spikes in images.

## Research setting

**Overview of the experimental site:** The wheat under test is "Zhengmai 1860" (bred by the Henan Academy of Agricultural Sciences) and is planted in Henan province. The experimental site is located at N35°0'44" and E113°41'44". It has a temperate continental climate.

**Data collection and dataset construction:** Image Data Collection: The collection period for wheat spike images spanned from May 1, 2024, to May 20, 2025, and image acquisition was conducted every two days. All image collection was carried out during the grain-filling stage of wheat. In this stage, images of potted wheat spikes were captured using smartphones. The smartphone was fixed on

a bracket and positioned 50 cm vertically above the potted wheat for vertical shooting. Meanwhile, the wheat pot was rotated continuously to obtain multi-angle images, so as to ensure the complete and comprehensive collection of wheat spike data. Following the selection criteria of clear image quality, minimal overlapping area between adjacent wheat spikes, and relatively high spike count per image, a total of 300 images were screened as the test set (Fig. 1).

**Data Segregation and Augmentation:** In this study, the whole dataset was partitioned into three subsets, namely the training set, the validation set and the test set, in accordance with the ratio of 6:1:3. To enrich the sample diversity of the wheat spike dataset, this study adopted multiple image augmentation strategies, including vertical flipping, horizontal flipping, noise injection and grayscale conversion. Finally, an augmented dataset containing 3795 images was obtained through the above processing. To meet the input requirements of network training, all images in the standardized wheat spike dataset were uniformly resized to 640×640 pixels. In this study, the Labelling tool was employed to perform detailed manual annotation on the potted wheat spike images. Specifically, each wheat spike was enclosed with a red bounding box and labeled as the category "wheat". After the annotation process was completed, the annotated data was saved in CSV format, which is compatible with the YOLO framework. Subsequently, the augmented dataset was re-partitioned into training, validation and test subsets, with the 6:1:3 partitioning ratio kept consistent with the initial partitioning scheme (Figs. 2 and 3).

**Model construction:** Network Structure of the YOLOv5s Training Model: The official YOLOv5 framework provides four types of object detection networks, namely YOLOv5s, YOLOv5m, YOLOv5l and YOLOv5x. Among the four network variants, YOLOv5s has the smallest network depth and width, and features compact model size and flexible parameter configuration, enabling its deployment on embedded devices. Moreover, the above structural characteristics of YOLOv5s can effectively improve the overall performance of the object detection task. Based on the above analysis, this study constructed a detection model based on the YOLOv5 architecture. The model consists of five components: input module, backbone network, neck module, detection head and output module. The input module undertakes the function of image input for the model, and supports Mosaic data augmentation. This data augmentation approach splices multiple images in a specific way, which can effectively improve the detection performance of the model for small targets. In addition, the model adopts optimized anchor box calculation and adaptive image scaling methods, which helps the model better adapt to the detection of targets of different sizes. Moreover, these optimizations can also reduce the information loss during image preprocessing. The backbone network of the proposed model is composed of Focus, Conv, C3 and SPP modules, which enable the model to extract rich feature information from input images, thereby ensuring higher detection speed and detection accuracy. For the neck component, this study adopts a combined structure of FPN and PAN. This design supports the generation of multi-scale feature maps, and effectively enhances the feature representation capability for targets of different sizes. The detection head component is responsible for predicting the

category probability of each detected object. This module is also capable of generating feature maps of different scales. Backbone Network Modules of YOLOv5s. The backbone network of YOLOv5s contains four core modules, namely Focus, Conv, C3 and SPP. The Conv module consists of convolutional layers, activation functions and BN layers, and its function is to extract local feature information from input images. The Focus module performs slicing processing on input images according to a set stride: it first divides the input image into multiple slices, then aggregates these sliced feature maps, and finally outputs feature maps of different scales. For example, after the slicing, concatenation and convolution operations of the Focus module, an input image with the size of  $3 \times 640 \times 640$  will be converted into a feature map with the size of  $32 \times 320 \times 320$ .



Fig. 1. Photos of potted wheat spikes.



Fig. 2. Original image of potted wheat spikes.



Fig. 3. Annotated image of potted wheat spikes.



Fig. 4. Detection results of wheat spikes images.

The C3 module adopts a dual-branch structure, which splits the input feature map into two separate branches. One branch directly transmits features through the cross-stage connection structure, while the other branch passes through the multiple stacked Conv modules. Furthermore, this module improves the capability of feature propagation and feature fusion via stacked bottleneck structures, and thus optimizes the feature extraction performance of the entire network.

**YOLOv5s Backbone Network Modules:** The backbone network consists of four modules: Focus, Conv, SPP, and C3. The Conv module is composed of batch normalization (BN) layers, convolutional layers and activation functions, and is capable of extracting local features from input images. Subsequently, the Focus module performs slicing processing on the input image with a fixed stride, splits the input into multiple sliced feature subsets, and fuses these subsets via concatenation to generate feature maps at multiple different scales. Taking the input image with a dimension of  $3 \times 640 \times 640$  as an example, after completing the operations of slicing, feature concatenation and convolution, the module can output a feature map with a dimension of  $32 \times 320 \times 320$ . The C3 module of the model is primarily composed of two branches, which split the input feature into two branches. One branch is directly connected across stages via residual shortcut, while the other enhances the propagation and fusion of image features through multiple consecutive Conv modules and stacked bottleneck modules. This structure improves the feature learning capability of the network, enabling the model to extract multi-level semantic features more effectively. The SPP module of the model belongs to the pooling layer, which performs max-pooling on input feature maps with multiple different kernel sizes, including  $5 \times 5$ ,  $9 \times 9$  and  $13 \times 13$ . Subsequently, all pooling results are concatenated along the channel dimension to generate the final output feature map. This design allows the detection model to better process objects with variant shapes and sizes, effectively improving the detection accuracy of the model.

**Model training:** The system uses Intel on the basis of Ubuntu 18.04 LTS ® Train on i7-8700 CPU (3.70 GHz), Nvidia Tesla T4 GPU, and 16 GB RAM. During the training process, the model utilized a batch size of 8000 epochs, a learning rate of 0.001, and an Adam optimizer.

**Model validation:** Ten images were randomly selected from the training set, and the number of wheat ears in each image was manually counted by agronomic experts, with the average of multiple manual counting results taken as the reference standard. The manual counting results were compared with the prediction results output by the model to quantitatively evaluate the detection accuracy of the proposed model.

**Evaluation metrics:** Precision, Recall, and mean Average Precision (mAP) were used to measure the model's performance. The formulas are as follows:  $P = \frac{TP}{TP+FP}$ ;  $R = \frac{TP}{TP+FN}$ ;  $mAP = \int_0^1 P(k) dR(k)$ . Here, TP means the number of correctly detected wheat spikes, FP is false detections number (for example, background areas mistaken for spikes), and FN is the number of missed wheat spikes.  $P(k)$  shows the precision for a given threshold ( $k$ ), and  $dR(k)$  represents change-in-recall as the threshold varies.

## Results and analysis

**Model Detection Performance:** In this study, the YOLOv5s model was employed to conduct detection on wheat spike images, and the detection results are visualized in Fig. 4, where the red bounding boxes mark the wheat spike regions identified by the model. The detection results were compared with the original images to verify the model's recognition performance. The visualization results demonstrate that the model can correctly identify even small-sized wheat spikes, and the positioning of the red bounding boxes for wheat spikes is also highly accurate. This indicates that the model significantly improves the detection performance for small targets, effectively alleviates the detection problems caused by wheat spike occlusion and overlap, and ultimately improves the counting accuracy of wheat spikes.

**Model training results:** To verify the performance of the YOLOv5s model on the self-constructed wheat ear dataset, the trained model was tested, and the results are presented in Fig. 5. The model achieved an average precision (AP) of 86.7%, a recall of 94%, and a mean average precision (mAP) of 85.7%. These results demonstrate that the proposed model maintains high detection accuracy and recall, with a converged loss function that meets the training convergence requirements (Fig. 6).

**Table 1. Comparison between manual counting and YOLOv5s model counting.**

| Image number | Actual value (ears/image) | Predicted value (ears/image) | Relative error (%) |
|--------------|---------------------------|------------------------------|--------------------|
| 1.           | 20                        | 19                           | 5%                 |
| 2.           | 18                        | 18                           | 0                  |
| 3.           | 21                        | 21                           | 0                  |
| 4.           | 22                        | 22                           | 0                  |
| 5.           | 19                        | 19                           | 0                  |
| 6.           | 21                        | 20                           | 4.76%              |
| 7.           | 18                        | 17                           | 5.56%              |
| 8.           | 23                        | 22                           | 4.35%              |
| 9.           | 17                        | 17                           | 0                  |
| 10.          | 19                        | 18                           | 5.26%              |

**Verification of test results:** To verify the performance of the YOLOv5s model for automatic wheat ear counting under real field conditions, this study utilized the trained model to conduct automatic counting on the field-collected wheat ear images. In this study, manual counting and

algorithm-based counting were both performed on the collected wheat ear images to obtain the actual number of wheat ears contained in each test image. Subsequent statistical analysis was performed on the counting results of 10 randomly selected test images. Table 1 presents the partial counting results of the wheat ear test images. The experimental results show that the average relative error of the wheat ear counting results obtained by the YOLOv5s model in this study is approximately 2.49%, and its detection efficiency is significantly higher than that of manual counting.

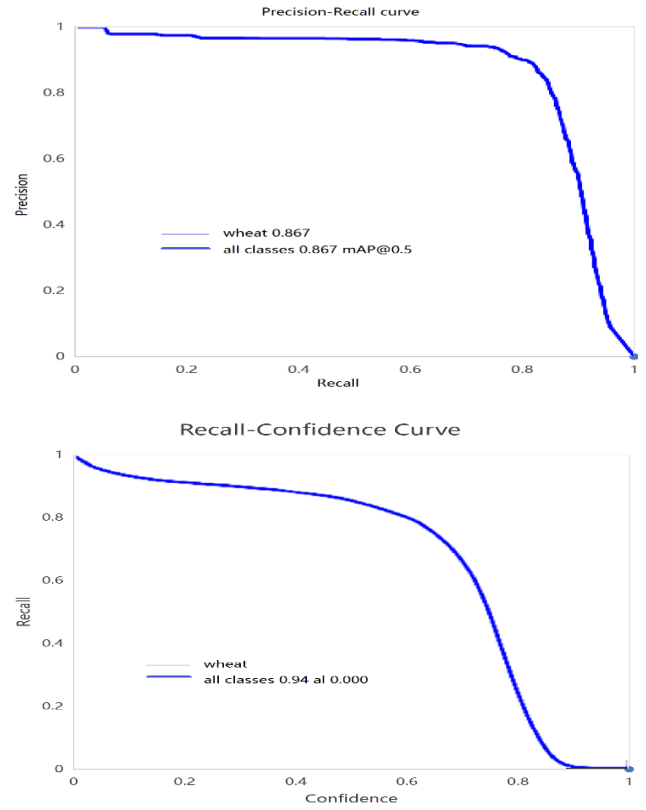


**Fig. 5. Model training results.**

## Discussion and Conclusions

In recent years, wheat ear detection and counting have achieved notable progress and attracted extensive research attention in the academic community. As a typical object detection model, YOLOv5 is a mature integrated iteration of the YOLO series algorithms. It not only delivers outstanding performance in general object detection tasks, but also achieves a fine trade-off between detection speed and inference accuracy. In practical scenarios, the morphological characteristics (including size and shape) of collected wheat ears are variable, due to the influences from growth environment and cultivar difference. Accordingly, when a fixed-scale algorithm is applied to detect wheat ears with varying sizes, shapes and color features, large detection errors frequently occur. The growth process of wheat features extremely high complexity and dynamic changes, which can be effectively characterized by variations in the color of wheat ears and their growth background. In current research, the design and application of wheat spike detection models are mostly oriented to specific growth

stages, with particular focus on early-stage gene fine-mapping-related research; in contrast, there is relatively limited research attention paid to the adaptability of such models to other growth stages or field cultivation conditions. An automatic wheat grain classification system was constructed using the SVM and K-means algorithms, with a classification accuracy of 88.33%. For field wheat images, the SLIC algorithm was adopted to

complete superpixel segmentation, after which morphological contours and color space features of wheat ears were extracted to implement wheat ear counting estimation, and the estimation accuracy reached 94.01%. This paper further proposes a YOLOv5s detection method based on an improved attention mechanism, which is designed for accurate detection of small-target wheat ears in counting tasks.



Fig. 6. Model detection visualization.

In dense wheat population samples, the probability of overlapping between adjacent wheat ears is relatively high. The YOLO-based regression detection framework divides the input image into multiple regular grid units, where each grid is only allowed to predict at most one detection target. When multiple targets coexist in the same grid, it is difficult for the framework to identify all targets, which will lead to degraded detection performance. In wheat ear detection tasks, there are frequently cases where overlapping, adjacent or connected wheat ears in input images are incorrectly identified as a single target, resulting in missed detection. This study proposes an automatic wheat ear counting detection method based on the improved YOLOv5s architecture, which can accurately and effectively capture the dynamic changes in wheat ear number throughout the growth stage from anthesis to maturity, and can replace the labor-intensive traditional manual counting method. Meanwhile, the method proposed in this study can effectively address the counting errors and missed detection caused by mutual occlusion and overlapping of wheat ears in image detection tasks, thereby realizing efficient and accurate counting of wheat ears. This method can provide technical support and methodological reference for field-grown wheat spike population estimation and yield prediction. This study focuses on image analysis of wheat spikes across different growth stages. Although existing object detection models can realize rapid counting of wheat spikes and achieve certain improvements in detection performance, current related research still has certain limitations, and the relevant theoretical system and technical methods require continuous improvement. Future research can be expanded and deepened from the following aspects: first, optimize the wheat spike counting model by integrating single-stage and two-stage object detection algorithms to improve the accuracy of wheat spike population counting; second, construct and expand large-sample wheat spike datasets to improve the balance of sample distribution through model training; third, continuously optimize the wheat spike image processing algorithm and improve the performance of the automatic counting system to enhance the accuracy and stability of counting results.

**Funding:** This research was funded by “Study on the effects of monocotyledonous lactone on physiological metabolism and yield formation of wheat under drought conditions”(21B210012), “Research on wheat spikelet detection method based on YOLO”(Independent innovation project of Henan academy of agricultural sciences 2025ZC77), Innovative Experimental Plan Project for Students of Zhengzhou Normal University (DCZ2024001) and Innovative Experimental Plan Project for Students of Zhengzhou Normal University (DCY2024003).

## Acknowledgments

Thanks to the Wheat Research Institute of Henan Academy of Agricultural Sciences for providing assistance in wheat cultivation and experimentation.

**Conflicts of Interest:** There are no conflicts of interest among all authors.

**Author’s Contribution:** Yanjing Wang, Ying Zhang, Jing Liu and Yunhui Bai performed the majority of the experiments. Linqing Wang reviewed the figures and polished the manuscript language. Hecang Zang provided financial support and supplied all wheat materials used in this study. All authors read and approved the final manuscript.

## References

- Bar-Hillel, A. 2021. Parts-per-object count in agricultural images: Solving phenotyping problems via a single deep neural network. *Remote Sensing*, 13. DOI 10.3390/rs13132496.
- Chauhan, S., R. Darvishzadeh, M. Boschetti and A. Nelson. 2020. Discriminant analysis for lodging severity classification in wheat using radarsat-2 and sentinel-1 data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 164: 138-151. DOI 10.1016/j.isprsjprs.2020.04.012.
- Costin-Anton, B., B. Raducanu and A.C. Spataru. 2009. High-precision orientation and skew detection for texts in scanned documents. In: 2009 IEEE 5th International Conference on Intelligent Computer Communication and Processing. pp. 145-148.
- Fernandez-Gallego, J., M. Buchailot, N. Aparicio Gutiérrez, M. Nieto-Taladriz, J. Araus and S. Kefauver. 2019. Automatic wheat ear counting using thermal imagery. *Remote Sensing*, 11(7): DOI 10.3390/rs11070751.
- Fernandez-Gallego, J.A., S.C. Kefauver, N.A. Gutiérrez, M.T. Nieto-Taladriz and J.L. Araus. 2018. Wheat ear counting in-field conditions: High throughput and low-cost approach using rgb images. *Plant Methods*, 14(1): 22. DOI 10.1186/s13007-018-0289-4.
- Han, H., B. Zhang, S. Fei, L. Li, H. Sun, D. Wang, Y. Meng and Y. Xiao. 2026. Accurate monitoring of wheat spike number based on uav video stream. *Journal of Triticeae Crops (in Chinese)*, 46(2): 264-275. DOI 10.7606/j.issn.1009-1041.2026.02.13.
- Hasan, M.M., J.P. Chopin, H. Laga and S.J. Miklavcic. 2019. Correction to: Detection and analysis of wheat spikes using convolutional neural networks. *Plant Methods*, 15(1): DOI 10.1186/s13007-019-0405-0.
- He, M.X., P. Hao and Y.Z. Xin. 2020. A robust method for wheatear detection using uav in natural scenes. *IEEE Access*, 8: 189043-189053. DOI 10.1109/ACCESS.2020.3031896.
- Jiao, Y. 2024. Winter wheat yield estimation by fusing cnn-malstm deep learning with remote sensing indices. *Agriculture*, 14. DOI 10.3390/agriculture14111961.
- Khaki, S., N. Safaei, H. Pham and L. Wang. 2022. Wheatnet: A lightweight convolutional neural network for high-throughput image-based wheat head detection and counting. *Neurocomputing*. DOI 10.1016/j.neucom.2022.03.017.
- Li, R., X. Sun, K. Yang, Z. He, X. Wang, C. Wang, B. Wang, F. Wang and H. Liu. 2025. A lightweight wheat ear counting model in uav images based on improved yolov8. *Frontiers in Plant Science*, 16: DOI 10.3389/fpls.2025.1536017.
- Lin, Y., X. Xiao and H. Lin. 2025. Yolov8-fda: Lightweight wheat ear detection and counting in drone images based on improved yolov8. *Frontiers in Plant Sci.*, 16: 1682243. DOI 10.3389/fpls.2025.1682243.
- Lu, H., L. Liu, Y.N. Li, X.M. Zhao, X.Q. Wang and Z.G. Cao. 2022. Tasselnetv3: Explainable plant counting with guided upsampling and background suppression. *IEEE Transactions on Geoscience and Remote Sensing* (60-). DOI 10.1109/TGRS.2021.3058962.
- Lu, J., J. Hu, G. Zhao, F. Mei and C. Zhang. 2017. An in-field automatic wheat disease diagnosis system. *Computers and Electronics in Agri.*, 142: 369-379. DOI 10.1016/j.compag.2017.09.012.

- Madec, S., X. Jin, H. Lu, B. De Solan, S. Liu, F. Duyme, E. Heritier and F. Baret. 2019. Ear density estimation from high resolution rgb imagery using deep learning technique. *Agri. and Forest Meteor.*, 264: 225-234. DOI 10.1016/j.agrformet.2018.10.013.
- Sadeghi-Tehran, P., N. Virlet, E.M. Ampe, P. Reyns and M.J. Hawkesford. 2019. Deepcount: In-field automatic quantification of wheat spikes using simple linear iterative clustering and deep convolutional neural networks. *Frontiers in Plant Sci.*, 10: 1176. DOI 10.3389/fpls.2019.01176.
- Su, J., X. Zhu and W.H. Li. 2023. Ai meets uavs: A survey on ai empowered uav perception systems for precision agriculture *Neurocomputing*, 518(Jan.21): 242-270. DOI 10.1016/j.neucom.2022.11.020.
- Sun, F., S. Zheng, Z. Li, Q. Gao and N. Jiang. 2025. Analysis of wheat spike morphological traits by 2d imaging. *Plant Phenomics*, 7(3): 100096. DOI 10.1016/j.plaphe.2025.
- Wang, Y., D. Gu, Y. Zhu, Z. Zhang, J. Cao and H. Zhang. 2025. Reflective clothing detection based on yolov5s fused with se attention mechanism. *International Jol. of Reasoning-based Intel. Sys.*, (1): 17. DOI 10.1504/IJRIS.2025.145054.
- Wei, L., A. Dragomir, E. Dumitru, S. Christian, R. Scott, F. Cheng-Yang and A.C. Berg. 2016. Ssd: Single shot multibox detector. *Springer, Cham*, 21-37. DOI 10.1007/978-3-319-46448-0\_2.
- Zhao, J., X. Zhang, J. Yan, X. Qiu, X. Yao, Y. Tian, Y. Zhu and W. Cao. 2021. A wheat spike detection method in uav images based on improved yolov5. *Remote Sensing*, 13(16): DOI 10.3390/rs13163095.