

Editorial

Computer Vision and Artificial Intelligence Driving the Advancement of Agricultural Intelligence in Dynamic Environments

Xiuguo Zou ^{1,*} , Xiaochen Zhu ², Wentian Zhang ³, Yan Qian ¹ and Yuhua Li ¹

¹ College of Artificial Intelligence, Nanjing Agricultural University, Nanjing 210031, China; qianyan@njau.edu.cn (Y.Q.); lyhresearch@njau.edu.cn (Y.L.)

² School of Applied Meteorology, Nanjing University of Information Science and Technology, Nanjing 210044, China; xiaochen.zhu@nuist.edu.cn

³ Faculty of Engineering and Information Technology, University of Technology Sydney, Sydney, NSW 2007, Australia; wentian.zhang@alumni.uts.edu.au

* Correspondence: zouxiuguo@njau.edu.cn; Tel.: +86-25-58606585

The rise of agricultural digitalization is progressively reshaping the conventional extended management model through the profound integration of intelligent sensing technology and artificial intelligence. The efficient transfer and application of these technologies therefore remains a major challenge, since modern agricultural settings are dynamically complicated and mostly unstructured. In response, researchers have been refining these technologies with a focus on lightweight design and adaptability, aiming to facilitate the large-scale application of smart sensing in agriculture.

This Special Issue examines the application of and breakthroughs in computer vision and artificial intelligence in agriculture. It features fifteen articles reporting advances in areas such as crop growth monitoring, fruit grading detection, animal health and behavior recognition, and remote sensing image analysis. Table 1 illustrates the essential information contributed by these fifteen papers. These studies employ a range of advanced technologies, including multi-scenario real-time detection via YOLO-based frameworks, Transformer architecture-based approaches, methods employing generative adversarial networks (GANs), and other algorithmic strategies. Researchers have made significant strides in a variety of agricultural tasks and contexts. Field tests have successfully shown that these technologies are feasible in complex and dynamic agricultural settings, offering technical support for the expanded use of computer vision and artificial intelligence in agriculture.

The YOLO series of object detection algorithms is a typical framework in computer vision. Because of its high accuracy and real-time performance, it has quickly developed across a variety of fields and is now a robust baseline for many downstream tasks. In the agricultural domain, researchers have introduced several targeted improvements to the YOLO architecture in response to challenges such as complex background interference, small object detection, and the need for deployment on edge devices, focusing particularly on enhancing real-time detection capabilities and optimizing model efficiency. Xin et al. [1] developed a mobile system for the detection and retrieval of dead chickens utilizing an enhanced YOLOv6 model, with a success rate of 81.3% indicated in field trials. Sun et al. [2] created a lightweight YOLOv8-Pearpollen model for evaluating the germination vigor of pear tree pollen; through the application of knowledge distillation and model pruning, they achieved a detection speed of 147.1 FPS. Chen et al. [3] addressed the challenges posed by



Received: 21 September 2025

Accepted: 23 September 2025

Published: 11 October 2025

Citation: Zou, X.; Zhu, X.; Zhang, W.; Qian, Y.; Li, Y. Computer Vision and Artificial Intelligence Driving the Advancement of Agricultural Intelligence in Dynamic Environments. *Agriculture* **2025**, *15*, 2112. <https://doi.org/10.3390/agriculture15202112>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

the small size and frequent occlusion of Chinese bayberry fruits with an improved YOLOv7-Tiny algorithm, which integrates partial convolutions, the SimAM attention mechanism, and a SIOU loss function. This optimized model retains a compact design and high efficiency, achieving a mere 4% miss rate and a 3% false detection rate in field tests. To tackle the issue of efficient citrus fruit detection in intricate orchard environments, Jing et al. [4] developed the YOLOv7-Tiny-BVP network by incorporating the BiFormer attention mechanism, the VoVGSCSP module, and partial convolutions into the YOLOv7-Tiny framework. This methodology produced a lightweight and highly precise citrus detection model, achieving enhancements of 0.9% in recognition accuracy, 2.02 FPS, and 1% in F1 score. Lin et al. [5] presented AG-YOLO, an enhanced algorithm based on the NextViT backbone and a Global Context Fusion Module. Specifically designed for occlusion-aware citrus detection, the research also involved the establishment of a diverse and comprehensive dataset, resulting in a testing performance of 83.2% mAP@0.5 and a detection speed of 34.22 FPS.

Table 1. Key information contributed by each paper.

Authors	Objects	Models	Contributions
Xin et al. [1]	Dead yellow-feather broilers	YOLOv6	Real-time detection and robotic grasping of dead yellow-feather broilers
Sun et al. [2]	Pollen germination vigor in pear trees	YOLOv8	Real-time detection of pollen germination vigor in pear trees
Chen et al. [3]	Bayberry	YOLOv7-Tiny	Real-time detection and picking of Chinese bayberries
Jing et al. [4]	Citrus	YOLOv7-Tiny	Citrus fruit recognition under varying occlusion scenarios and lighting conditions
Lin et al. [5]	Citrus	YOLO, NextViT	Citrus fruit detection in complex environments
Jin et al. [6]	Wheat	Transformer, Symmetric Diffusion	Precise monitoring of wheat growth status and yield prediction in high-density agricultural environments
Huang et al. [7]	Pear	RT-DETR	Rapid detection of Xinli No. 7 fruit in natural environments
Huo et al. [8]	Tomato	StyleGAN3, Transformer	Recognition of growth stages in greenhouse tomato cultivation
Li et al. [9]	Weed	Semi-Supervised Diffusion Model	Weed detection in agricultural scenarios
Wang et al. [10]	Apple	Edge Attention Mechanism	Extraction of apple phenotypic features and recognition of growth abnormalities
Chen et al. [11]	Oudemansiella raphanipes	CNN, Three-Teacher Knowledge Distillation	Quality grading of Oudemansiella raphanipes
Cho et al. [12]	Dairy cows	Weakly Supervised Representation Learning	Classification and detection of mastitis in dairy cows
Yang et al. [13]	Yak	YOLOv7-pose	Detection and classification of behavior patterns in yaks
Wu et al. [14]	Zhuji Torreya	Machine Learning	Suitable habitats for Torreya in Zhuji City
Wu et al. [15]	Walnut	w-YOLO	Identification and counting of small walnut fruits in UAV remote sensing images
Pacioni et al. [16]	Vineyard shoots	Mask R-CNN (ResNet50 backbone), YOLOv8	Direct detection of vine shoot cutting areas for intelligent pruning

By leveraging self-attention mechanisms to model long-range dependencies in both sequential and spatial data, Transformer architectures enable high-precision crop phenotyping, yield forecasting, and pest-and-disease detection in agricultural settings, thereby markedly enhancing intelligent monitoring and decision-making. Jin et al. [6] presented SDS-Net, a Symmetric Diffusion Segmentation Network utilizing a Transformer backbone. By effectively combining Transformer blocks with symmetric diffusion modules and a symmetric attention mechanism, SDS-Net significantly enhances wheat instance segmentation accuracy in high-density planting environments, achieving an F1 score of 0.89 for segmentation and 0.92 for growth measurement. Huang et al. [7] addressed the challenge of identifying Xinli No. 7 pears in unstructured outdoor environments by proposing a lightweight Transformer-based detection framework. Their optimized model achieved a 48.47% reduction in total parameter count, a 56.2% decrease in computational overhead (FLOPs), and a 48.31% reduction in memory usage, all while maintaining robust detection accuracy.

However, the efficacy of deep learning models is fundamentally dependent on the availability of high-quality and diverse training data. In agricultural contexts, data collection often lacks standardized protocols, resulting in imbalanced datasets and challenges in achieving both efficiency and accuracy in recognition. Furthermore, traditional augmentation methods generate images of limited fidelity and diversity, rendering them insufficient for the rigorous requirements of effective model training. To address these challenges, Huo et al. [8] introduced an innovative framework that merges the capabilities of the StyleGAN3 generative adversarial network with Vision Transformer techniques, effectively addressing the shortcomings of traditional data collection and recognition approaches. Through the innovative use of StyleGAN3 to create high-fidelity images of tomato growth stages, they effectively tackled challenges related to data scarcity and class imbalance. This approach achieved a remarkable accuracy of 98.39% on the test set, with an average inference time of just 9.5 ms. Li et al. [9] utilized diffusion models for weed detection by creating a semi-supervised, diffusion-generated recurrent network that excels at generating high-quality synthetic data. By incorporating generation-aware attention modules and a hybrid diffusion loss function, alongside real-world images, this approach significantly lessens the dependence on extensive annotated datasets. It also boosts model adaptability and detection performance in challenging, data-scarce agricultural settings.

The studies discussed have mainly concentrated on object detection within conventional agricultural environments, investigating the use of YOLO-based models, Transformer architectures, and GANs in these contexts. With an increasing need for advanced agricultural practices, there is a notable trend towards collaborative efforts that encompass multiple tasks. In response, extensive work has been conducted to enhance algorithmic frameworks and introduce a range of innovations suited to diverse application scenarios.

In plant phenotyping and quality grading, Wang et al. [10] considered environmental variables like climate and soil by merging instance segmentation with natural language processing. This cross-modal fusion facilitated the integration of multidimensional data for the extraction of phenotypic traits in apples and the detection of growth anomalies, resulting in a more thorough framework for identifying anomalies and assessing quality. Similarly, Chen et al. [11] enhanced the automated grading of *Oudemansiella raphanipes* in resource-limited settings using a three-teacher knowledge distillation approach, showing that parallel, standard cascade, and residual-connected architectures each markedly improve the accuracy of compact models. Turning to livestock health monitoring, Cho et al. [12] introduced a weakly supervised classification approach for detecting mastitis utilizing representation learning techniques. This method utilizes a one-dimensional convolutional neural network autoencoder, incorporating a classifier branch within the latent

space of the autoencoder. This facilitates efficient representation learning of milking data and the weakly supervised identification of mastitis symptoms. Current image-based techniques for detecting animal behavior face challenges in accurately differentiating between closely related behavioral patterns in yaks, including standing, walking, and excreting, as these actions exhibit similar leg postures. In response to this issue, Yang et al. [13] introduced a method for detecting yak behavior utilizing an enhanced YOLOv7-pose model. The integration of the Mish activation function, the SPPFCSPC feature extraction module, and a dynamic head module notably improved the model's capacity to identify yak behavioral patterns. The method demonstrated a precision of 89.9% and an mAP@0.5 of 90.4%. In the domain of remote sensing and habitat assessment, Wu et al. [14] employed various machine learning algorithms, such as Random Forest (RF), Support Vector Machine (SVM), and Naive Bayes (NB), to conduct suitability zoning for *Torreya grandis* cultivation in Zhuji City. The researchers integrated ecological factor raster data with the geographic coordinates of current *Torreya grandis* distribution points to pinpoint areas conducive to its growth. The results indicate that suitable locations for *Torreya* cultivation are primarily confined to mountainous and hilly terrains. The central basin and northern river plain zones are considered inappropriate. Among the variables examined, edaphic and topographic factors impose more significant constraints on *Torreya* distribution than climatic factors do. To address the challenge of detecting small walnut fruits in imagery obtained from UAVs, Wu et al. [15] developed an efficient detection network called w-YOLO. In direct comparisons with leading object detection frameworks, w-YOLO achieved a mAP@0.5 of 97% and an F1 score of 92%. The model demonstrated a 52.3% decrease in parameter count when compared to YOLOv8s. This demonstrates remarkable precision and effectiveness, even under the difficult lighting conditions commonly found in Yunnan Province. Finally, to resolve the issue of intelligent vineyard pruning via cutting area detection, Pacioni et al. [16] developed a labeled VidPrune Dataset and tested various segmentation models. The optimal model achieved mAP50 = 0.883 (shoot class) and a 55 ms inference time using Jetson AGX Orin, enabling efficient direct cutting area detection.

Intelligent detection in dynamic agricultural environments faces challenges such as varying lighting conditions, occlusion, limited computational resources, and the need for real-time performance, as well as the integration of multimodal data. The studies published in this Special Issue address these challenges from multiple technical perspectives, demonstrating breakthroughs in adapting to dynamic agricultural scenarios. These works validate the feasibility of applying next-generation artificial intelligence and computer vision techniques to smart agriculture. Researchers have utilized techniques such as pruning, quantization, knowledge distillation with cascade design, and Efficient Transformer architectures to reduce computational complexity in model lightweighting. The studies have integrated diffusion models, generative adversarial networks, and semi-supervised learning to address the challenges of limited data diversity and class imbalance in agricultural datasets. The integration of visual and historical information in application scenarios facilitates multimodal and comprehensive perception of environmental conditions. This Special Issue compiles various studies that align with the advanced techniques mentioned, aiming to inspire new directions in agricultural intelligence and promote ongoing research and advancements in related areas. Ongoing technological refinement and interdisciplinary collaboration are expected to drive the advancement of smart agriculture, resulting in concurrent improvements in productivity and sustainability.

Acknowledgments: We are thankful to QiuHong Zhang, Hongfei Chen, and Jinlin Liu, who have contributed to helping with paper information sorting.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Xin, C.; Li, H.; Li, Y.; Wang, M.; Lin, W.; Wang, S.; Zhang, W.; Xiao, M.; Zou, X. Research on an Identification and Grasping Device for Dead Yellow-Feather Broilers in Flat Houses Based on Deep Learning. *Agriculture* **2024**, *14*, 1614. [\[CrossRef\]](#)
2. Sun, W.; Chen, C.; Liu, T.; Jiang, H.; Tian, L.; Fu, X.; Niu, M.; Huang, S.; Hu, F. YOLOv8-Pearpollen: Method for the Lightweight Identification of Pollen Germination Vigor in Pear Trees. *Agriculture* **2024**, *14*, 1348. [\[CrossRef\]](#)
3. Chen, Z.; Qian, M.; Zhang, X.; Zhu, J. Chinese Bayberry Detection in an Orchard Environment Based on an Improved YOLOv7-Tiny Model. *Agriculture* **2024**, *14*, 1725. [\[CrossRef\]](#)
4. Jing, J.; Zhai, M.; Dou, S.; Wang, L.; Lou, B.; Yan, J.; Yuan, S. Optimizing the YOLOv7-Tiny model with multiple strategies for citrus fruit yield estimation in complex scenarios. *Agriculture* **2024**, *14*, 303. [\[CrossRef\]](#)
5. Lin, Y.; Huang, Z.; Liang, Y.; Liu, Y.; Jiang, W. Ag-yolo: A rapid citrus fruit detection algorithm with global context fusion. *Agriculture* **2024**, *14*, 114. [\[CrossRef\]](#)
6. Jin, Z.; Hong, W.; Wang, Y.; Jiang, C.; Zhang, B.; Sun, Z.; Liu, S. A Transformer-Based Symmetric Diffusion Segmentation Network for Wheat Growth Monitoring and Yield Counting. *Agriculture* **2025**, *15*, 670. [\[CrossRef\]](#)
7. Huang, Z.; Zhang, X.; Wang, H.; Wei, H.; Zhang, Y.; Zhou, G. Pear Fruit Detection Model in Natural Environment Based on Lightweight Transformer Architecture. *Agriculture* **2024**, *15*, 24. [\[CrossRef\]](#)
8. Huo, Y.; Liu, Y.; He, P.; Hu, L.; Gao, W.; Gu, L. Identifying Tomato Growth Stages in Protected Agriculture with StyleGAN3-Synthetic Images and Vision Transformer. *Agriculture* **2025**, *15*, 120. [\[CrossRef\]](#)
9. Li, R.; Wang, X.; Cui, Y.; Xu, Y.; Zhou, Y.; Tang, X.; Jiang, C.; Song, Y.; Dong, H.; Yan, S. A Semi-Supervised Diffusion-Based Framework for Weed Detection in Precision Agricultural Scenarios Using a Generative Attention Mechanism. *Agriculture* **2025**, *15*, 434. [\[CrossRef\]](#)
10. Wang, Z.; Cui, W.; Huang, C.; Zhou, Y.; Zhao, Z.; Yue, Y.; Dong, X.; Lv, C. Framework for Apple Phenotype Feature Extraction Using Instance Segmentation and Edge Attention Mechanism. *Agriculture* **2025**, *15*, 305. [\[CrossRef\]](#)
11. Chen, H.; Huang, H.; Peng, Y.; Zhou, H.; Hu, H.; Liu, M. Quality Grading of *Oudemansiella raphanipes* Using Three-Teacher Knowledge Distillation with Cascaded Structure for LightWeight Neural Networks. *Agriculture* **2025**, *15*, 301. [\[CrossRef\]](#)
12. Cho, S.H.; Lee, M.; Lee, W.H.; Seo, S.; Lee, D.H. Mastitis Classification in Dairy Cows Using Weakly Supervised Representation Learning. *Agriculture* **2024**, *14*, 2084. [\[CrossRef\]](#)
13. Yang, Y.; Deng, Y.; Li, J.; Liu, M.; Yao, Y.; Peng, Z.; Gu, L.; Peng, Y. An Effective Yak Behavior Classification Model with Improved YOLO-Pose Network Using Yak Skeleton Key Points Images. *Agriculture* **2024**, *14*, 1796. [\[CrossRef\]](#)
14. Wu, L.; Yang, L.; Li, Y.; Shi, J.; Zhu, X.; Zeng, Y. Evaluation of the Habitat Suitability for Zhuji Torreya Based on Machine Learning Algorithms. *Agriculture* **2024**, *14*, 1077. [\[CrossRef\]](#)
15. Wu, M.; Yun, L.; Xue, C.; Chen, Z.; Xia, Y. Walnut Recognition Method for UAV Remote Sensing Images. *Agriculture* **2024**, *14*, 646. [\[CrossRef\]](#)
16. Pacioni, E.; Abengózar, E.; Macías Macías, M.; García-Orellana, C.J.; Gallardo, R.; González Velasco, H.M. Towards Intelligent Pruning of Vineyards by Direct Detection of Cutting Areas. *Agriculture* **2025**, *15*, 1154. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.