

Computer vision and image segmentation algorithms in agricultural robotics applications

Tingrui Zhan

College of Mechanical Engineering, Taiyuan University of Technology, Taiyuan, China

rock_0119@qq.com

Abstract. Agriculture is a cornerstone of global food security, yet it faces significant challenges in the modern era, including labor shortages, environmental pressures, and the need for sustainable practices. Traditional farming methods are often labor-intensive and inefficient, leading to increased costs and reduced productivity. To address these challenges, the integration of advanced technologies into agricultural practices has become essential. Among these technologies, computer vision has emerged as a powerful tool, offering precise and automated solutions that can significantly enhance farming efficiency and sustainability. This paper presents a comprehensive review of computer vision technologies and their algorithmic applications in agricultural robotics, highlighting their transformative role in modernizing traditional farming practices. The study underscores computer vision as a pivotal driver for scalable, efficient, and sustainable agriculture, with future potential in resource optimization and food security. Proposals for future computer vision-enabled agricultural robots in the fields of picking and weeding, crop identification, path planning, pest detection, seed screening, etc.

Keywords: agriculture, computer vision, image segmentation, robotics

1. Introduction

Traditional agriculture predominantly depends on human and animal labor, rendering it highly susceptible to natural factors. The sector is typically characterized by a singular structure, limited production scale, and inadequate resilience to disasters. With the development of industrialization, agricultural machinery such as tractors and harvesters have greatly improved efficiency, making large-scale production possible. However, due to complex terrain and the lack of intelligence in mechanical equipment, agricultural outcomes remain unsatisfactory. In recent years, Convolutional Neural Networks (CNNs) have been widely applied in fields like image segmentation, model classification, and crop detection. Meanwhile, YOLO and its derivative models play an indispensable role in areas such as weed and crop identification, pest and disease detection, and resource detection. It is evident that computer-empowered agricultural machinery is an inevitable path for agricultural development.

This paper reviews the application scope and development prospects of computer vision-related technologies in the field of agricultural robots. By examining these technologies, the study aims to highlight their potential to revolutionize agricultural practices, enhance productivity, and address the pressing challenges of labor shortages and environmental sustainability in modern farming.

2. An overview of image segregation

Image segmentation technology is a computer vision technique whose core is to divide a digital image into several independent regions with similar attributes according to certain rules, such that pixels within each region are consistent, while there are obvious differences between different regions. This technology provides a foundation for machines to understand image content and is widely used in medical imaging, autonomous driving and agriculture. The agricultural application scenarios of robots can be roughly divided into three categories: planting, breeding, and special scenario applications [1]. It can accurately capture the growth dynamics of fruits and provide clear operational targets for robots. Zhang et al. used traditional image segmentation algorithms such as Otsu, Simple Linear Iterative Clustering (SLIC), watershed algorithm, concave point analysis algorithm, and edge detection algorithm to analyze the development of crop seeds [2]. Wang et al. believe that image segmentation is a process of dividing an image into multiple unique sub-regions based on features such as gray scale, color, shape, and texture. They used U-Net and DeepLabv3+ models for segmentation to determine the leaf disease status of wheat and corn [3]. In the picking of

clustered crops such as grapes and tomatoes, segmentation technology can further distinguish "mature fruits" from "immature fruits". Robots segment fruit regions of different colors (e.g., the red region of mature tomatoes and the green region of immature ones) and only plan picking paths for mature regions to avoid mistakenly picking immature fruits, which increases the picking qualification rate by 20%-30%. Therefore, image segmentation enables machines to better understand structures and contents, providing a basis for crop condition analysis.

3. The application of computer visualization in agriculture machine

3.1. Crop recognition and picking

Computer vision algorithms can be widely applied to crop identification and picking. Image recognition usually involves the following steps: collecting corresponding data, performing segmentation processing on crop images, extracting crop features, and detecting the required targets [4]. Taking grape picking as an example, segmentation technology effectively isolates complete grape clusters from the background, which is composed of vines and leaves. This is achieved by leveraging differences in color and morphology. The technology then precisely delineates the spatial position and contour of each individual bunch of grapes. Subsequently, it assesses maturity through color thresholds to meet the needs of graded picking. In an experiment conducted in Flevoland, the Netherlands, Momchil Yordanov et al. extracted phenological information and crop types from street-level images using TensorFlow and convolutional neural networks for classification. Meanwhile, in research on real-time recognition of apple picking methods by robots, Yan Bin et al. enhanced the recognition of apple targets in the current planting row by improving the feature map cross-connection fusion method of the input medium-sized target detection layer and the initial anchor box size in the original YOLOv5m architecture. This provides a good solution to the problem of avoiding damage when picking apples blocked by branches [5]. The application of this technology has broken the efficiency bottleneck of traditional manual picking. Through the closed loop of "precision recognition - intelligent decision-making - automated operation", large-scale and precise crop picking has been successfully realized. In the future, combining multi-modal perception (such as fusing visual and tactile data) with cross-scenario transfer learning can further improve the model's adaptability to complex environments and promote agricultural picking robots to develop towards higher efficiency and lower cost.

3.2. Pest and disease detection

In addition, computer vision and its algorithms can also be applied in fields such as pest and disease detection. It can provide fast and accurate detection results, helping to take timely prevention and control measures and reduce losses. Through deep learning models such as Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN), machines can perform advanced analysis on images and sensor data to recognize crop growth status, diseases, and pests in real time [6]. Different pests have significantly different damage characteristics. Image segmentation technology can help distinguish pest types by extracting morphological features of damaged areas. For example, in tomato field pests, cotton bollworms cause irregular holes in leaves, while aphids cause leaf curling. Instance segmentation models such as Mask R-CNN can achieve a recognition accuracy of over 90%. In the experiment of pest and disease detection on tomato leaves, Hu et al. used computer vision-related technologies and proposed an improved model Real-Time Transfer Learning for Target Recognition (RT-TLTR) based on the Real-Time Detection Transformer (RT-DETR) framework. This improved model showed good spatial positioning and category judgment capabilities in actual detection, especially for small target features, but there were still cases of misjudgment of some healthy leaf areas when facing overlapping leaves [7]. Zhu et al. applied the improved Contextual Receptive Field Dilated Convolution-YOLOv8 (CRDM-YOLOv8) to greatly improve the detection accuracy of litchi pests and diseases in natural orchard environments [8]. Deng optimized the YOLOv5 model and created an exclusive dataset covering 9 types of pests and diseases and normal fruits, providing strong technical support for jujube pest and disease prevention [9]. This technology, combined with multi-spectral image segmentation (such as utilizing the difference in reflectance between pests and crops in the near-infrared band) and lightweight model deployment (such as integrating algorithms into low-cost sensors), can further enhance the adaptability of the technology in complex field environments and promote the upgrading of pest control towards "intelligence and greenization".

3.3. Fruit and vegetable picking

The growth of weeds is widely recognized as a significant factor that can greatly affect crop yields. The empowerment and application of computer vision technology have also made weeding easier. Some crops and weeds differ in color, spectrum, and resolution. The chlorophyll content of healthy crop leaves is generally high, and their reflectance in the near-infrared band is significantly higher than that of weeds. There are also differences in color depth and vein distribution between crops and weeds at the seedling stage. Image recognition combined with RGB + near-infrared images collected by multi-spectral cameras enhances recognition differences through color space conversion, improving accuracy. Commonly used laser wavelengths for

weeding robots are 1064nm near-infrared laser and 532nm green laser. The 1064nm laser has strong penetration ability, can penetrate into weeds and destroy their cell structure, and is suitable for removing weeds with deep roots [10]. With the core technology of computer vision, weeds and crops can be accurately identified. Crops and weeds can be distinguished through features such as leaf shape, color, texture, and growth morphology, and even weed species can be identified. Wang et al. used the improved MSBlock-ELA-SIoU-YOLO (MES-YOLO), which is quite effective in dealing with lamb's-quarters [11]. Reenul Reedha et al. applied Vision Transformer technology to plant classification in UAV images, solving problems such as high computing cost and the need for large-scale labeling in weed and crop classification [12].

3.4. Road planning

In terms of path planning for agricultural robots, image segmentation technology also shows its excellence. Traditional field navigation relies on GPS or set paths, which are easily affected by crop occlusion and difficult to adapt to programmed paths. Computer vision technology can avoid these problems to a certain extent. Using edge detection technology, crop areas and weed areas can be separated from field images. During driving, the crop offset is monitored in real time and fed back to the path planning module. Path planning algorithms are the core technology in agricultural robot navigation systems. The ant colony algorithm is an optimization method based on swarm intelligence, which finds the optimal solution through a pheromone positive feedback mechanism. Li Juan et al. improved the algorithm using the self-attention mechanism of the differential Transformer. By using the distributed parallel search ability of the ant colony method to generate the initial global path, and integrating the differential Transformer model to replace the traditional pheromone update method, the approach achieves accurate capture of long-distance dependencies and nonlinear features between path nodes. This also enables efficient update of pheromones, thus better overcoming obstacles in complex environments [13].

3.5. The selection of seed

Image segmentation technology plays a crucial role in seed screening. By accurately extracting seed regions and analyzing their features, it enables automation, high efficiency, and precision in seed screening. This effectively replaces the traditional screening method that relies on manual visual inspection, significantly improving screening efficiency and quality.

In specific applications, industrial cameras first collect images of bulk seeds to obtain original images containing seeds. At this point, the images may be mixed with impurities, debris, or background areas, and image segmentation technology comes into play here: using threshold segmentation, edge detection, or deep learning segmentation algorithms (such as U-Net), seeds are separated from complex backgrounds, and the contours of each seed are accurately outlined to obtain target regions containing only seeds.

After segmentation, key features of the seeds can be extracted based on the segmentation results, such as size (calculated by contour area), shape (e.g., circularity, aspect ratio), and color (to determine if the color is uniform and if there are mold spots). By comparing these features with the preset standards for high-quality seeds, the system can automatically screen out healthy seeds that meet the requirements and eliminate deformed, damaged, pest-infested, or underdeveloped seeds.

This application is widely applicable to the screening of various crop seeds, whether it is grain seeds (such as wheat and rice), oil seeds (such as rapeseed and peanut), or vegetable seeds. Image segmentation technology can achieve efficient screening, provide high-quality seeds for agricultural production, ensure the germination rate and growth quality of crops from the source, reduce labor costs, and promote the standardization and modernization of agricultural planting.

4. Challenges and countermeasures of computer visualization in agricultural machinery

Computer empowerment has injected unprecedented vitality into traditional agriculture. Empowering digital decision-making, computer technology has brought unprecedented intelligent changes to the agricultural field. Empowering automated operations has taken smart agriculture to a new level. Empowering information management helps the high-quality development of the agricultural industry [14]. It accelerates the computer's ability to process information and operate continuously, breaks through the physiological limitations of human labor, significantly improves production efficiency, and enables large-scale operations. This technology also synergistically enhances multi-modal perception. For example, the SwishFormer model combines RGB images with DIGIT tactile sensor data, achieving an accuracy rate of over 96% in avocado maturity detection. It has wide applications in the refined management of the planting industry, the expansion of breeding and special scenarios, and cross-industry technology migration and innovation.

However, the current application of computer vision technology in agriculture still has certain potential risks. Firstly, at the technical level, because computer vision relies on algorithms to identify crops, pests, etc., if there are errors in the algorithms, it may lead to recognition mistakes, which in turn affect the accuracy of fertilization, irrigation, etc. Secondly, the data collected by agricultural robots may be maliciously tampered with or illegally collected, thus causing security risks. In many cases, the

existing legal and regulatory system has not fully kept up with the pace of technological development, especially in terms of data security, privacy protection, and the use of intelligent equipment, where there are no clear and specific regulations for standardized management [15]. In addition, the high cost of this technology may lead to problems such as low cost performance.

For computer-empowered agricultural robots, improvements can be made in three key aspects: technology research and development, cost control, and talent training. In terms of technology research and development, continuous efforts can be devoted to improving and updating existing technologies. Through cross-scenario transfer learning, the feature extraction of crop models can be gradually enhanced. Technical updates can be achieved through interdisciplinary joint research. In terms of cost control, lessons can be drawn from DJI Agriculture's "three-electrical system integrated design" to reduce the cost of core components through mass production. Alternatively, unnecessarily expensive parts can be replaced with more cost-effective alternatives. Regarding talent training, models employed by enterprises such as Huawei can be emulated by establishing corresponding research institutes and laboratories to attract top-tier talents.

5. Conclusion

This paper discusses the application status, technical paths, and development prospects of computer image segmentation technology and its algorithms in the field of agricultural robots. Computer vision technology, leveraging core methods such as image segmentation, target detection, and deep learning models, has demonstrated significant value in key agricultural applications, including crop identification and picking, pest and disease detection, weed control, and path planning. In crop picking, segmentation and recognition technologies based on models like U-Net and the YOLO series can achieve precise positioning of mature fruits, thereby increasing the picking qualification rate by 20% to 30%. In the field of pest and disease detection, improved models such as RT-TLTR and CRDM-YOLOv8 can achieve a recognition accuracy of over 90%, providing robust support for rapid prevention and control efforts. In weed recognition, the combination of multi-spectral images with advanced technologies like MES-YOLO and Vision Transformer has effectively addressed the challenge of accurately distinguishing between crops and weeds. In path planning, the integration of differential Transformer models with ant colony algorithms has overcome traditional limitations, such as occlusion and poor path adaptability, thereby enhancing navigation capabilities in complex environments.

The current application of the technology still faces three challenges: algorithm errors may lead to deviations in agricultural operations, lack of regulations for data security and privacy protection, and high cost of core technologies restricting large-scale promotion. In this regard, optimization should be promoted in three aspects: improving model performance through cross-scenario transfer learning and interdisciplinary research at the technical level; reducing thresholds through mass production integrated design and cheap component replacement strategies at the cost level; and cultivating interdisciplinary teams relying on professional laboratories and research institutions at the talent level. Computer vision technology provides a core driving force for the intelligent upgrading of agricultural robots, and its in-depth integration with agricultural scenarios is an inevitable path to realize agricultural scale, precision, and efficiency. In the future, through technological iteration and industrial collaboration, this field will play a more critical role in improving agricultural production efficiency, reducing resource consumption, and ensuring food security.

References

- [1] Cao, J. (2025). Application and impact of robots in agriculture. *Hebei Agricultural Machinery*, (10), 12–14. <https://doi.org/10.15989/j.cnki.hbnjzss.2025.10.038>
- [2] Zhang, W., Wang, F., Sun, X., Wang, J., Liu, H., & Wang, X. (2024). Research progress of traditional image segmentation algorithms in crop seed testing. *Journal of Chinese Agricultural Mechanization*, 45(2), 280–287. <https://doi.org/10.13733/j.jcam.issn.2095-5553.2024.02.040>
- [3] Wang, C., & Qiu, X. (2025). Application of artificial intelligence in image segmentation of crop leaf diseases. *Seed Industry Guide*, (2), 76–80.
- [4] Shi, X., Wang, S., Liu, H., Wang, S., Qu, H., Li, N., & Yang, H. (2024). Research status and prospects of fruit-picking robots. *Agricultural Equipment & Vehicle Engineering*, 62(6), 1–7.
- [5] Yan, B., Fan, P., Wang, M., Shi, S., Lei, X., & Yang, F. (2022). Real-time recognition of apple-picking methods for harvesting robots based on improved YOLOv5m. *Transactions of the Chinese Society for Agricultural Machinery*, 53(9), 28–38+59.
- [6] Guo, S. (2024). Application and development of intelligent robots in agriculture. *Hebei Agricultural Machinery*, (10), 42–44. <https://doi.org/10.15989/j.cnki.hbnjzss.2024.10.041>
- [7] Hu, C., Huang, L., Zheng, Q., Li, J., Wang, B., & Huang, Z. (2024). RT-TLTR: A method for detecting pests and diseases in tomato leaves. *Computer Engineering and Applications*, 1–20.
- [8] Zhu, J., Chen, G., Jiang, L., Yin, J., Zhan, H., Lan, Y., ... Xu, H. (2025). Lightweight improved YOLOv8n for litchi pest and disease detection. *Transactions of the Chinese Society of Agricultural Engineering*, 1–9.
- [9] Deng, B. (2025). Jujube pest and disease detection system: Engineering application and deployment practice of YOLOv5. *Journal of Xiangnan University*, 46(2), 45–51. <https://doi.org/10.20277/j.cnki.xnxyxb.2025.02.007>

- [10] Li, L., Zhang, W., & Xu, L. (2025). Design and optimization of the mechanical structure of intelligent laser weeding robots. *Hebei Agricultural Machinery*, (6), 12–14. <https://doi.org/10.15989/j.cnki.hbnjzss.2025.06.041>
- [11] Wen, T., Wang, T., Huang, S., & Zhou, J. (2024). MES-YOLO: An improved YOLOv8-based algorithm for crop and lamb's-quarters detection. *Computer Engineering & Science*, 1–13.
- [12] Reedha, R., Dericquebourg, E., Canals, R., & Hafiane, A. (2022). Transformer neural network for weed and crop classification of high-resolution UAV images. *Remote Sensing*, 14(3), 592.
- [13] Li, J., & Zhang, Z. (2025). Path planning for agricultural robots using ant colony algorithm and differential Transformer. *Journal of Chinese Agricultural Mechanization*, 46(7), 164–172.
- [14] Li, F. (2023). Opportunities and challenges of computer technology empowering smart agriculture. *China Fruits*, (12), 155.
- [15] Li, L., Hong, S., & Huang, X. (2024). Challenges and breakthroughs in the application of computer networks in intelligent rice agriculture. *Northern Rice*, 54(5), 76–78. <https://doi.org/10.16170/j.cnki.1673-6737.2024.05.006>