Vol. 14 (2026): Proceedings of the 2nd International Conference on Artificial Intelligence, Modern Engineering and Environmental Sustainability (IC-AIMEES 2026)
E-ISSN: 3105-0522 P-ISSN：3105-0514
DOI: https://doi.org/10.70267/ m1wnbq95

Design of Lightweight Vision System for Agricultural Drones: Real-time Identification of Crop Diseases under Complex Lighting Conditions

Yinlai Zhang*

China Agricultural University, Beijing 100000, China

**Corresponding author：Yinlai Zhang*

Abstract

Global food security is facing multiple pressures such as shrinking arable land and extreme weather, driving agriculture towards high precision and intelligence. Unmanned aerial vehicles (UAVs), as an aerial monitoring platform, provide important technical support for crop disease monitoring with their mobility and imaging capabilities. However, the complex lighting conditions in the field have led to a decline in image quality, and the limited computing power of the embedded platform has severely restricted the accuracy and real-time performance of disease identification. For this reason, this paper systematically reviews the latest progress in lightweight visual model design and light enhancement methods. In terms of model lightweighting, channel pruning, quantization techniques and knowledge distillation strike an effective balance between accuracy and efficiency; In terms of illumination enhancement, from traditional Retinex algorithms to physically guided generative adversarial networks, there is a significant improvement in image quality and model robustness. Future research could focus on areas such as polarization imaging and federated learning to further enhance the generalization ability and reliability of the system in complex environments.

Keywords

agricultural drones, lightweight vision systems, crop disease identification, complex lighting robustness, edge computing deployment

1. Introduction

Global food security is facing multiple pressures such as shrinking arable land and extreme weather, which forces agriculture to move towards high precision and high intelligence. As an aerial detection platform in the field, drones have shown unique value in crop growth detection - their flexible mobility can quickly cover large areas of farmland, and their high-definition imaging capabilities help with precise diagnosis of agricultural conditions. However, complex and variable lighting conditions can sometimes affect the performance of these detection devices: yellow spots of wheat stripe rust are confused with reflections under strong light, and the loss of grain patterns of grape downy mildew is caused by shade, which is particularly prominent in the rice-wheat rotation area measured by the Jiangsu Academy of

Agricultural Sciences. According to the actual test data of Wang Jianjun's team, the rate of missed detection of typical diseases by drones in cloudy weather was as high as 41.2%, and the direct loss caused by incorrect pesticide application exceeded 220 yuan per mu.

An in-depth analysis of field failure cases shows that the core contradictions are concentrated in three aspects:

First, light interference erodes recognition accuracy. The specular reflection of citrus leaves under oblique light blurs the boundary between canker spots and healthy tissue. Second, the computing power shackles restrict the response speed. Zhang's tests showed that when using high-precision models like ResNet50 [1], even on the Jetson Nano platform, the inference latency exceeded 500ms, which simply couldn't match the 10m/s operation speed of the agricultural drone. Third, the crop diversity challenge model is broad. Chen et al. 's comparative experiments revealed that the same algorithm could identify rice sheath blight with an accuracy of 85.7%, but for cotton yellow wilt, it dropped sharply to 63.4% [2].

To systematically promote the resolution of the aforementioned technical dilemma, the research framework of this paper aims to review the key technological breakthroughs in the past five years. First, in the field of model lightweighting, this paper will compare the evolution paths from manual rule-based pruning schemes to neural architecture search-based automatic model compression techniques. Secondly, in the direction of illumination enhancement, this paper will analyze the technological iterations and advantages from the traditional multi-scale Retinex algorithm to the emerging physical-guided generative adversarial network. In addition, this paper will explore optimization schemes in edge deployment practices, such as engine optimization techniques that can significantly compress inference latency to the millisecond level. By comparing the performance boundaries and inherent limitations of these schemes, this paper will ultimately point to frontier directions such as polarized light sensing and federated learning to provide insights for building the next generation of robust, inclusive agricultural drone vision systems.

2. Technical Routes for Lightweight Vision Models

2.1 Model Compression Methods

To accommodate the limited computing resources of field edge devices, model compression techniques significantly reduce the complexity and energy consumption of models through pruning and quantization. In terms of channel pruning, Zhao et al. proposed a method that assesses channel importance by calculating the L1 norm of the convolutional layer weights, then ranks and removes the bottom 50% of channels that contribute less to the lesion features, and finally uses a small number of typical lesion samples for fine-tuning to restore accuracy. In the actual test in Shaanxi Apple Orchard, the model size was compressed from 89MB to 41MB, and the inference speed was increased from 22FPS to 38FPS on the Jetson Nano platform, while the average accuracy of black star disease detection decreased by only 1.8%. On the other hand, INT8 quantization technology performed well in Jiangsu cotton field pest monitoring by retaining small target features through a calibration mechanism, converting the 32-bit floating-point weights of ResNet18 to 8-bit integers, reducing the model size to 4.3MB, enabling farmers to update the model in seconds via mobile terminals, and has been widely applied in production.

The practical significance of the model compression technique lies in its effective resolution of the real-world bottlenecks faced by agricultural drones. By reducing the model size and computational requirements, these methods ensure that the recognition system can keep up with the high-speed operation rhythm of the drones and meet the urgent need for real-time processing in the field. Further, the highly compressed models reduce the reliance on hardware performance, making large-scale, low-cost deployments possible and significantly enhancing the practical value and promotion prospects of agricultural drone vision systems.

2.2 Knowledge Distillation Architecture

Knowledge distillation architectures transfer knowledge from large teacher models to lightweight student models through a teacher-student learning mechanism, enabling efficient edge deployment while ensuring high precision. The cross-modal teacher-student architecture developed by Liu Y et al. is a representative work, which uses high-precision multispectral ResNet50 as the teacher model and lightweight RGB-

MobileNetv3 as the student model at a cost of only one-fifth of the teacher model. The core of the distillation mechanism of the method lies in guiding the student model to imitate the feature response heatmap generated by the teacher model to the edge features of the lesion, such as the radial texture presented by the citrus ulcer lesion.

In agricultural disease identification tasks, the value of knowledge distillation is mainly reflected in the construction of a technical path from a high-cost, high-precision model to a low-cost, high-efficiency model. Rather than simply replicating the final classification results of the teacher model, the student model is required to learn the feature representation of the middle layer of the teacher model, thereby mastering how to effectively capture the subtle differences and structural features between the lesion and the healthy tissue. This mechanism offers two significant advantages. First, the student model can break through the capacity limitations of its own network structure and achieve a significant leap in recognition performance, reaching the design goal of being small but precise. Secondly, the method enables the transfer of knowledge from multispectral modes to common RGB modes, allowing lightweight systems equipped with only cheap RGB cameras to achieve recognition capabilities close to those of high-precision systems that rely on expensive multispectral imaging equipment, significantly lowering the threshold and application cost of advanced agricultural detection technologies.

In the field test in a citrus orchard in Guangdong, the recognition accuracy of the student model adopting this scheme reached 96% of that of the teacher model, and the reasoning speed was as high as 56FPS, fully meeting the requirements of modern agricultural drones for ground imitation flight operations. This practice fully demonstrates that knowledge distillation, through algorithmic knowledge compression, effectively addresses the core contradiction of insufficient recognition accuracy of lightweight models and the difficulty of deploying high-performance models on resource-constrained devices, providing key technical support for achieving high-reliability real-time disease identification on edge computing platforms.

3. Complex Light Enhancement Mechanism

3.1 Traditional Image Processing Techniques

Traditional methods estimate and correct illumination components mainly through physical models or image transformations. The multi-scale Retinex algorithm proposed by the Poznan Polytechnic University team is a representative work. The principle is to break down the image into the illumination component and the reflection component, estimate and separate the illumination through the Gaussian function, and ultimately achieve image enhancement. This scheme effectively suppresses shadows and enhances spot contrast, but it has poor manual parameter generalization ability and is prone to over-enhancement in strong light, resulting in detail loss and halo artifacts.

Homomorphic filtering is another technical approach that addresses uneven illumination by suppressing low-frequency light components and enhancing high-frequency reflection components in the frequency domain. However, these methods have inherent limitations: first, they cannot fix the information loss caused by overexposure or underexposure; Second, they rely heavily on expert experience for parameter tuning, making it difficult to adapt to the ever-changing lighting conditions during drone patrols.

3.2 Deep Learning Enhancement Methods

Deep learning enables end-to-end image enhancement in a data-driven way, with the core being learning the mapping relationship from degraded images to high-quality images, providing key technical support for the reliable operation of agricultural drones in complex lighting environments.

Generative adversarial network frameworks have shown significant advantages in agricultural image enhancement. PhysGAN, developed by Wang et al., introduces a physically-guided mechanism, where the generator follows the pre-embedded physical model and the discriminator simultaneously determines the authenticity and physical rationality of the image [3]. This physical constraint enables the network to generate visually realistic and optically compliant results, improving image availability from 0.71 to over 0.89 on standard agricultural datasets. The core value of the technology on agricultural drones lies in its ability to effectively suppress light interference caused by strong light reflection or shadow coverage, providing highly robust input for subsequent lightweight recognition models and significantly reducing performance fluctuations of the models under different lighting conditions. This feature enables drones to work stably in complex lighting conditions such as dawn, dusk, and overcast, while significantly reducing the cost and difficulty of collecting perfect training samples in the field by generating high-quality, information-complete images.

Regression networks based on encoder-decoder structures have also been widely explored and have shown unique engineering application value in agricultural drone systems. These networks encode low-quality image features and then decode and reconstruct them into high-quality images. While avoiding the training instability of Gans, their performance is highly dependent on a large amount of paired data. In specific scenarios where data is abundant, such as long-term monitoring of a single high-value crop, these networks offer more stable and easily deployable enhancements. Its advantages in agricultural drone applications are mainly reflected in two aspects: on the one hand, it avoids the training instability of GAN models, shortens the research and development and deployment cycle, and enables faster integration into embedded platforms; On the other hand, once trained, its forward inference process is typically less computationally complex, which is crucial for computationally tight drone edge devices, helping to ensure high throughput for real-time processing and meet real-time analysis requirements during high-speed drone flight.

The two deep learning methods address the challenges faced by agricultural drones in complex lighting from different perspectives, providing complementary technical paths for building adaptable and reliable field vision systems.

4. Comparative Analysis of Technical Performance

4.1 Performance Boundaries of Lightweight Models

The core objective of lightweight vision models is to address the core contradiction between the limited computing power of edge devices and the high computational complexity of models. The technical paths vary, with both advantages and limitations coexisting.

The channel pruning technique mainly addresses the problem of model parameter redundancy and excessive size. In the case of black star disease detection in Shaanxi Apple Orchard, the scheme reduced the model size from 89MB to 41MB by pruning redundant channels and increased the inference speed from 22FPS to 38FPS on the Jetson Nano platform. The core advantage lies in directional optimization, which retains the spot-sensitive feature layers through importance assessment, thereby achieving significant acceleration and slimming while controlling the average precision (mAP) loss to just 1.8%, achieving a good balance between efficiency and precision. However, its limitations lie in its weak ability to retain features for small or sparse targets such as bollworm holes, which are prone to missed detections due to cropping.

INT8 quantization techniques are more aggressive in breaking through the bottleneck of model storage and transmission bandwidth. In the application of pest identification in cotton fields in Jiangsu, the technique compressed the model size to 4.3MB by converting 32-bit floating-point weights to 8-bit integers. This fundamentally solved the update problem of the model in low-bandwidth networks, enabling farmers to complete second-level model updates via mobile phone apps, greatly enhancing the practicality and scalability of the system. The advantage lies in the extreme compression ratio and ease of deployment, but performance is highly dependent on a fine calibration process to avoid information loss of small target features during quantification.

The unique value of the knowledge distillation architecture lies in its breaking of the traditional dilemma that "lightweighting is bound to be accompanied by loss of precision." The practice of South China Agricultural University in the Guangdong Citrus orchard shows that by having the lightweight student model imitate the feature response of the high-performance teacher model, the student model achieves 96% recognition accuracy at only 1/5 of the cost of the teacher model and an inference speed of 56FPS. The greatest advantage is that it achieves performance that approximates complex models within a lightweight framework, making it particularly suitable for real-time operation scenarios where both accuracy and speed are demanding. However, the upper limit of the method's performance is constrained by the quality of the teacher model and the effectiveness of knowledge transfer, and its generalization ability across different crops needs further validation.

4.2 Verification of Lighting Enhancement Effect

Complex illumination enhancement techniques aim to address the problem of reduced model recognition due to image quality degradation. Their methods range from traditional image processing to deep learning, with different focuses on adapting to scenarios and enhancing effects.

Traditional methods, such as the multi-scale Retinex algorithm, contribute by estimating and separating the illumination components through physical models, effectively addressing the problem of insufficient image contrast in low light and shadow conditions, and improving image usability by approximately 15%. The advantage of this method lies in its intuitive principle, relatively low computational complexity, and stable performance in scenes with gentle illumination changes. But the disadvantages are equally obvious: halo artifacts are prone to occur at strong light boundaries, and parameter adjustments rely on expert experience, making it difficult to adapt to the drastic dynamic changes in lighting conditions during drone patrols.

Deep learning-based methods, especially physically-guided generative adversarial networks such as PhysGAN, essentially break through the limitations of traditional methods. It learns the end-to-end mapping from degraded images to high-quality images in a data-driven way, and it introduces a physically guided mechanism that forces the generated results to conform to optical laws. This approach significantly enhances the ability to recover from extreme lighting, such as overexposure and underexposure, and boosts image availability to over 0.89 on standard datasets, thereby providing highly robust input for subsequent recognition models. The core advantage is its strong adaptability and excellent enhancement effect. However, its high computational complexity and requirements for the amount of training data pose new challenges for real-time deployment on edge devices and need to be applied in combination with model lightweight technology.

In conclusion, lightweight technology and light enhancement methods are the two key factors for improving the performance of the vision system of agricultural drones. The selection and optimization direction should be closely integrated with specific application scenarios: INT8 quantification has a significant advantage for large-scale monitoring with extremely limited computing power and frequent model updates; For precision plant protection scenarios with extremely high precision requirements and high crop value, a combination of knowledge distillation or pruning and light enhancement may be a better solution. Future research should focus on organically integrating these techniques to build end-to-end optimization solutions.

5. Challenges and Developments

5.1 Existing Technology Challenges

Despite significant progress in lightweight models and light enhancement techniques, the current system still faces several clear challenges in actual deployment. In terms of model lightweighting, pruning and quantization techniques have significantly improved efficiency, but their inherent information loss leads to a decrease in the model's sensitivity to identifying specific types of spots, such as small or sparse spots like cotton bollworm holes, significantly increasing the risk of missed detection. The performance of the knowledge distillation method is highly dependent on the quality and generalization ability of the teacher model, resulting in insufficient adaptability in cross-crop and cross-regional scenarios. In terms of light enhancement, traditional methods represented by multi-scale Retinex have problems such as poor generalization ability and reliance on manual parameter tuning, making it difficult to cope with the intense dynamic changes in field light conditions during drone patrols; Deep learning methods represented by PhysGAN show significant enhancement effects, but their high computational complexity poses new challenges for real-time processing on edge devices with limited computing power, restricting their wide application on agricultural drone platforms.

5.2 Future Technological Advancements

To address these challenges, future research could focus on two frontier directions. One of them is to explore the fusion of new sensing technologies and algorithms, such as polarization imaging, which can directly suppress specular reflection at the physical level and provide models with higher-quality raw data, fundamentally alleviating the interference caused by complex lighting. The second is to draw on distributed learning paradigms such as federated learning to achieve collaborative training of scattered data while strictly guaranteeing the privacy of individual farmers' data, thereby continuously optimizing model parameters and effectively enhancing the generalization ability of the model across different crops, regions and growth stages, ultimately promoting the construction of more inclusive and robust field disease intelligent identification systems.

6. Summary

This paper systematically reviews the research progress of lightweight visual models and light enhancement methods in response to the technical requirements for real-time identification of crop diseases by agricultural drones in complex lighting environments. Through the analysis of model compression methods such as channel pruning, quantization techniques and knowledge distillation, the technical paths to improve model efficiency while maintaining recognition accuracy are identified; At the same time, the performance of traditional Retinex algorithms and deep learning methods in terms of illumination enhancement was compared to verify the effectiveness of techniques such as physically guided generative adversarial networks in improving image quality. The study shows that the combined application of these techniques can significantly improve the performance of unmanned aerial vehicle (UAV) vision systems in complex environments.

The value of this study lies in providing a systematic technical reference for the actual deployment of agricultural drone vision systems. By optimizing the model structure and enhancing the image quality, the accuracy and real-time performance of disease identification can be effectively improved, providing reliable technical support for precise pesticide application, which is of positive significance for reducing agricultural production costs and pesticide use.

With the continuous development of new sensing technologies and distributed learning methods, agricultural drone vision systems will make further breakthroughs in model generalization ability and environmental adaptability. Future research directions include multimodal data fusion, cross-platform collaborative learning, etc. These advancements will drive the continuous improvement of the perception ability of agricultural drones in complex environments and provide more complete technical solutions for the development of modern agriculture.

References

[1] Kong, Xiangxin. (2025). Research and Development of Embodied Intelligent Hybrid Robot Control and AI Vision System Based on Lightweight Network Design. Science and Technology Innovation, (18), 203-206.

[2] Tian Chen. (2024). Research on Disease Identification and Optimization Decision-Making for Drone Plant Protection Operations under Agricultural Outsourcing Models (Master's Thesis, Northwest A&F University). Master'shttps://doi.org/10.27409/d.cnki.gxbnu.2024.001396.

[3] Wang, F. K. (2025). Traffic Object Detection Algorithm Based on YOLOv10n Improvement (Master's Thesis, Nanchang University). Master's https://doi.org/10.27232/d.cnki.gnchu.2025.003135.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.