Proceedings of the 4th International Conference on Mechatronics and Smart Systems
DOI: 10.54254/2755-2721/2026.33737

Imaging Modalities and Deep Learning Approaches for Citrus Huanglongbing Detection

Boyang Hu

School of Software, East China University of Technology, Nanchang, China
2024211458@ecut.edu.cn

Abstract. Citrus Huanglongbing remains one of the most difficult citrus diseases to diagnose reliably in commercial orchards. Early symptoms are often faint, uneven within the canopy, and similar to nutrient deficiency or other field stresses, which makes both manual inspection and image-based diagnosis uncertain. This review examines how deep learning has been used to improve Huanglongbing diagnosis and where its practical value is still limited under real orchard conditions. Recent studies were compared from three aspects: imaging modality and sensing scale, learning task, and model architecture. The reviewed work includes visible-light imaging, multispectral and hyperspectral sensing, low-altitude aerial remote sensing, classification, detection, segmentation, convolutional networks, multimodal fusion, and attention-based models. The comparison shows that visible-light images are useful for low-cost and rapid screening, but their performance depends strongly on lighting, background complexity, and symptom clarity. Spectral sensing provides richer disease-related information and is more helpful for early recognition, although equipment cost, calibration, and data handling remain practical barriers. Aerial sensing extends diagnosis from leaves or single trees to orchard-level mapping, while detection and segmentation can provide more usable spatial information than simple image classification. Overall, the reviewed evidence suggests that reliable Huanglongbing diagnosis cannot depend on one sensor or one model alone. More practical systems should combine affordable sensing, complementary image information, robust model design, and field-based evaluation to support earlier and more stable disease management in citrus orchards.

Keywords: Citrus Huanglongbing, deep learning, image-based diagnosis, multispectral and hyperspectral imaging, orchard remote sensing

1. Introduction

Citrus Huanglongbing (HLB) is one of the most destructive diseases affecting citrus production worldwide [1]. Its consequences are severe: yield reduction, fruit quality deterioration, canopy decline, and even tree death in many cases. For key citrus cultivation areas, the disease is as much a threat to production stability as to tree health to long-term stability of production. The broader agricultural context is equally serious. According to the Food and Agriculture Organization (FAO), plant pests and diseases account for annual crop losses of up to 40% worldwide, with associated

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.33737

economic damage exceeding USD 220 billion each year [2]. Within this context, HLB stands out as a sustained burden on surveillance and disease control [1].

Current HLB diagnosis still depends largely on field scouting, visual inspection, and laboratory assays. Useful, yes - but also labor-intensive, subjective, and difficult to scale for continuous orchard monitoring. A further complication lies in symptom expression: HLB can resemble nutrient deficiency and other physiological stresses, especially in early stages [1]. Reliable diagnosis therefore becomes difficult precisely when early intervention would matter most. Under orchard conditions, the problem becomes sharper still; illumination changes, canopy occlusion, and background stress signals may all obscure subtle symptoms and weaken the consistency of manual judgment.

These limitations have motivated image-based deep learning approaches that can learn HLB-related visual and spectral features from RGB, hyperspectral, multispectral, and UAV imagery across leaf-to-orchard scales. Yang et al. recently demonstrated that combining RGB with hyperspectral data can push leaf-level HLB classification accuracy to 97.89%, outperforming either modality alone [3]. Yet a leaf-level assay, however precise, still leaves open the question of how such a model would perform under the variable lighting and canopy occlusion of a commercial orchard. That question is taken up by Ye et al. [4], who moved the sensing platform to UAVs and showed that multispectral imagery, rather than high-resolution RGB, is more effective for tree scale disease recognition and achieves an mIoU of 0.89, an mPA of 0.94, a precision of 0.95, and a recall of 0.94. Together, these studies illustrate a shift from classification at leaf level to sensing at orchard scale: not just a progression in accuracy, but a conceptual move toward operational, cross-scale diagnosis that merges spectral richness with field practicality.

Many existing models are developed on controlled datasets and therefore do not fully reflect the heterogeneity of real orchards, where cultivars, canopy structures, illumination conditions, and acquisition protocols may vary substantially [1-4]. Cross-study comparison is also complicated by inconsistencies in imaging modalities, label definitions, learning tasks, and evaluation metrics. Therefore, this review focuses specifically on deep learning studies for HLB diagnosis rather than plant disease detection in general. It examines how imaging modalities, sensing scales, model architectures, data conditions, and deployment settings influence the reliability and practical usability of HLB diagnosis in orchard environments. The remainder of this paper is organized as follows. Section 2 reviews deep learning applications in HLB diagnosis, including imaging modalities, learning tasks, and representative model architectures. Section 3 discusses current challenges and future research directions. Section 4 concludes the paper.

2. Deep learning for HLB diagnosis

2.1. Imaging modalities and sensing scales

2.1.1. RGB imaging under controlled and natural conditions

RGB imaging is the earliest visual tool for intelligent HLB diagnosis. Its low cost and field usability made leaf- and canopy-scale visible-light analysis central to early research [5-7]. It differentiates healthy, HLB-infected and non-HLB samples, laying a baseline for automated disease identification [5, 6]. Pourreza et al. hit 92.7% accuracy for real-time orchard detection [5]; Deng et al. boosted illumination robustness via a color-texture framework [6]; Gómez-Flores et al. cut lighting-related error with intensity-invariant texture analysis [7]. All RGB methods rely entirely on visible symptoms, vulnerable to background clutter and symptom confusion from nutrient deficiency or

DOI: 10.54254/2755-2721/2026.33737

abiotic stress. Suited only for rapid preliminary screening, not confirmatory diagnosis, high-specificity HLB detection will require multi-source data fusion [5-8].

2.1.2. Multispectral and hyperspectral imaging

Unlike RGB imaging, multispectral and hyperspectral imaging capture invisible HLB-induced changes in leaf pigments, water status and internal tissue structure, making them ideal for early HLB detection [9]. Deng et al. built a non-destructive hyperspectral reflectance detection method, innovating band selection via entropy distance and sequential backward selection; their SVM model hit 90.8% three-group classification accuracy and 96.0% healthy-symptomatic differentiation, with 13 optimal characteristic bands identified to balance diagnostic value and efficiency. Large-scale orchard use remains limited by high equipment costs, data redundancy, and calibration/real-time processing challenges under fluctuating sunlight and complex field backgrounds [9].

2.2. Learning tasks in HLB diagnosis

2.2.1. Detection and localization

Object detection and localization pinpoint the exact spatial location of HLB lesions, a key advantage over image classification that only returns whole-image class probabilities. This fits core orchard needs: field scouting and variable-rate operations. Liu et al. built MSHLB-DETR from an optimized RT-DETR architecture, targeting multi-scale HLB symptoms and foliage occlusion, bringing transformer-based detectors into practical HLB field deployment for the first time [10]. These models output precise diseased region bounding boxes, delivering direct guidance for infected tree annotation, automated scouting path planning, and prescription map generation [10].

2.2.2. Segmentation and canopy-scale mapping

Segmentation pushes HLB diagnosis to the pixel level and aims to delineate diseased regions and their spatial distribution within the canopy, making it an important link between single-tree diagnosis and orchard-scale mapping [4]. A representative study is the MGA-UNet model proposed by Ye et al. [4] for early HLB recognition. By introducing channel and spatial attention into the skip connections, the model improved its ability to detect and segment subtle diseased areas under complex lighting and textured backgrounds. Compared with classification and detection, pixel-level segmentation provides finer information on disease coverage and spatial distribution, which is more useful for estimating infection extent, assessing orchard-level damage, and supporting variable-rate spraying. Its limitations are also clear: pixel-level annotation is expensive, and model generalization across orchards and imaging conditions is still limited [4].

2.3. Model architectures

2.3.1. Cnn-based models

CNNs help because they offer a much stronger visual feature backbone for HLB diagnosis. Their value lies in the ability to learn local texture, edges, and spatial structure directly from image data, which greatly reduces the dependence on hand-crafted features. Gómez-Flores et al. introduced deep neural networks and transfer learning into orange-tree HLB recognition, showing clearly how this line of work was shifting from shallow feature engineering to end-to-end deep representation

DOI: 10.54254/2755-2721/2026.33737

learning [8]. The same pattern appears in most RGB-based studies. Even when the task is framed as classification or multimodal diagnosis, convolutional feature extraction still sits at the core; in the four-class setting reported by Yang et al., discrimination among HLB, suspected HLB, zinc deficiency, and healthy samples still depended on effective visual feature learning [3]. CNNs remain the basic framework supporting HLB visual diagnosis [3, 8].

2.3.2. Multimodal fusion and transformer-based architectures

A single image stream is often not enough for HLB diagnosis. Leaf yellowing and mottling caused by HLB can look very close to zinc or manganese deficiency, not only in RGB images but also in spectral responses, so models built on one modality alone are easily confused. This is where fusion models become useful. Yang et al. designed a four-class framework that separated HLB, suspected HLB, zinc deficiency, and healthy samples by coupling RGB and hyperspectral feature extraction with bilinear fusion, showing that complementary spatial and spectral cues can improve discrimination when symptoms overlap [3]. A similar idea appears at the orchard scale: Deng et al. combined UAV hyperspectral data with a multi-input neural network for stress detection across large planting areas, indicating that multi-source remote sensing can support HLB identification under more complex field conditions [11]. Transformer-based models push this direction further. In Liu et al.'s MSHLB-DETR, self-attention was used to handle multi-scale targets and occlusion from dense branches and leaves, which are common in commercial citrus orchards [10]. CNNs still dominate the field, but fusion and transformer-based hybrids may improve robustness when symptom similarity, canopy complexity, and background interference occur together [3, 10, 11].

3. Discussions

3.1. Challenges

Low-cost RGB imaging, as discussed in Section 2.1.1, remains attractive for orchard deployment, yet its dependence on visible phenotypes leaves early HLB screening vulnerable to symptom confusion with zinc or manganese deficiency and to instability under variable sunlight, shadow, and background clutter [5-8]. Spectral sensing partly eases this problem by capturing physiological responses beyond visible color, but the gain in sensitivity comes with a different burden: expensive instruments, redundant bands, and demanding calibration under field conditions [9]. The same tension appears at orchard scale. UAV platforms extend observation from individual leaves to entire orchard blocks, but changes in flight height, canopy density, viewing angle, and illumination can shift image statistics enough to weaken consistency across orchards and seasons [4, 11, 12].

This gap between promising experiments and routine use also runs through task design and model architecture. Classification is efficient, but it stops at image-level labels; detection adds location, whereas segmentation offers finer canopy mapping at the cost of labor-intensive annotation [4, 10]. CNN-based models still provide the main feature backbone, multimodal fusion improves separation of HLB from look-alike stresses, and Transformer-based detectors handle occlusion and scale variation more effectively [3, 8, 10, 11]. Yet each improvement introduces a new constraint, whether data demand, model complexity, or reduced deployability. What still matters most is not accuracy on a single dataset, but whether these systems can remain stable, interpretable, and operational in commercial orchards.

DOI: 10.54254/2755-2721/2026.33737

3.2. Future prospects

Future work will likely move toward systems that are easier to use in orchards rather than models that only perform well on isolated datasets. RGB imaging should remain the first choice for routine screening because it is cheap and easy to deploy, but its dependence on visible symptoms means that higher-specificity diagnosis will need support from spectral sensing or UAV-based observation, especially when early HLB, nutrient stress, and uneven field illumination appear together [3, 4, 9, 11, 12]. What seems more promising is not a single "best" sensor, but a practical combination of sensing scale and modality.

The same trend applies to model design. CNNs will probably continue to serve as the basic feature extractor, while multimodal fusion will be useful where symptom overlap is difficult to resolve from one image stream alone [3, 8, 11]. Transformer-based and hybrid models may become more important if they can keep their advantage under canopy occlusion and scale variation without sacrificing efficiency [10]. In the end, the value of these methods will depend on whether they can support real decisions in orchards, including scouting, targeted spraying, and infected-tree removal, rather than only reporting strong accuracy under controlled conditions.

4. Conclusion

Deep learning is most useful for Huanglongbing diagnosis when it is treated as a field tool rather than only an accuracy problem. RGB images give a simple starting point, but the disease is too easy to mix up with nutrient stress when only visible symptoms are used. Spectral images and UAV data make the diagnosis richer, especially at canopy and orchard scales, and models such as fusion networks or detection models can turn image data into more practical information. Still, this paper does not think one method can solve the problem alone. A useful system should be affordable, stable under real orchard conditions, and able to give growers information they can act on. For this reason, the real value of deep learning is not just higher numbers in experiments, but earlier and more reliable support for HLB management.

References

[1] Thakuria, D., Chaliha, C., Dutta, P., Sinha, S., Uzir, P., Singh, S.B., Hazarika, S., Sahoo, L., Kharbikar, L.L. and Singh, D. (2023) Citrus Huanglongbing (HLB): Diagnostic and management options. Physiological and Molecular Plant Pathology, 125, 102016.

[2] Food and Agriculture Organization of the United Nations. (2026) About FAO's work on plant production and protection. FAO, Rome, Italy. Retrieved Mar. 27, 2026, from https: //www.fao.org/plant-production-protection/about/en

[3] Yang, D., Wang, F., Hu, Y., Lan, Y. and Deng, X. (2021) Citrus Huanglongbing detection based on multi-modal feature fusion learning. Frontiers in Plant Science, 12, 809506.

[4] Ye, N., Mai, W., Qin, F., Yuan, S., Liu, B., Li, Z., Liu, C., Wan, F., Qian, W., Wu, Z. and Qiao, X. (2025) Early detection of Citrus Huanglongbing by UAV remote sensing based on MGA-UNet. Frontiers in Plant Science, 16, 1503645.

[5] Pourreza, A., Lee, W.S., Ehsani, R., Schueller, J.K. and Raveh, E. (2015) An optimum method for real-time in-field detection of Huanglongbing disease using a vision sensor. Computers and Electronics in Agriculture, 110, 221-232.

[6] Deng, X., Lan, Y., Hong, T. and Chen, J. (2016) Citrus greening detection using visible spectrum imaging and C-SVC. Computers and Electronics in Agriculture, 130, 177-183.

[7] Gómez-Flores, W., Garza-Saldaña, J.J. and Varela-Fuentes, S.E. (2019) Detection of Huanglongbing disease based on intensity-invariant texture analysis of images in the visible spectrum. Computers and Electronics in Agriculture, 162, 825-835.

[8] Gómez-Flores, W., Garza-Saldaña, J.J. and Varela-Fuentes, S.E. (2022) A Huanglongbing detection method for orange trees based on deep neural networks and transfer learning. IEEE Access, 10, 116686-116696.

DOI: 10.54254/2755-2721/2026.33737

[9] Deng, X., Huang, Z., Zheng, Z., Lan, Y. and Dai, F. (2019) Field detection and classification of citrus Huanglongbing based on hyperspectral reflectance. Computers and Electronics in Agriculture, 167, 105006.
[10] Liu, Z., Wu, D., Xu, F., Du, Z., Luo, R. and Li, C. (2025) MSHLB-DETR: Transformer-based multi-scale Citrus Huanglongbing detection in orchards with aggregation enhancement. Horticulturae, 11, 1225.
[11] Deng, X., Zhu, Z., Yang, J., Zheng, Z., Huang, Z., Yin, X., Wei, S. and Lan, Y. (2020) Detection of Citrus Huanglongbing based on multi-input neural network model of UAV hyperspectral remote sensing. Remote Sensing, 12, 2678.
[12] Lan, Y., Huang, Z., Deng, X., Zhu, Z., Huang, H., Zheng, Z., Lian, B., Zeng, G. and Tong, Z. (2020) Comparison of machine learning methods for citrus greening detection on UAV multispectral images. Computers and Electronics in Agriculture, 171, 105234.