Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Deep learning for non-destructive quality assessment and intelligent grading of postharvest fruits: A comprehensive review

Yikun Huang
Fujian Agriculture and Forestry University
Yongchun Zhan
Fujian Agriculture and Forestry University
Yunzhong Zhou
Fujian Agriculture and Forestry University
Riqing Chen
Fujian Agriculture and Forestry University
Pang Jie
pang37219410163.com
Fujian Agriculture and Forestry University

Research Article

Keywords: Deep learning, Postharvest fruits, Non-destructive quality assessment, Intelligent grading, Industrial deployment

Posted Date: November 4th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11350704/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Deep learning for non-destructive quality assessment and intelligent grading of postharvest fruits: A comprehensive review

Yikun Huang[1,2], Yongchun Zhan[1], Yunzhong Zhou[1], Riqing Chen[1,3*], Jie Pang[1*]

[1]Fujian Agriculture and Forestry University, Fuzhou, Fujian, China.
[2]Concord University College, Fujian Normal University, Fuzhou, Fujian, China.
[3]Fujian Key Lab of Agricultural IOT Applications, Sanming University, Sanming, Fujian, China.

*Corresponding author(s). E-mail(s): riqing.chen@fafu.edu.cn; pang3721941@163.com;

Abstract

Postharvest fruit loss presents a critical global challenge. Although deep learning (DL) has revolutionized quality assessment, achieving high accuracy in classification and defect detection, existing reviews predominantly focus on isolated technical aspects, lacking a systematic whole-chain perspective. To bridge this gap, this paper establishes a comprehensive review framework for postharvest fruit quality assessment and intelligent grading, encompassing the entire process from non-destructive data acquisition to industrial deployment. Following a "task-architecture-data" thread, we systematically compare the information characteristics and application scenarios of eight non-destructive sensing modalities (e.g., red-green-blue (RGB), hyperspectral, near-infrared, and X-ray imaging) based on DL data requirements. Furthermore, this review evaluates the applicability of fundamental DL architectures and synthesizes their recent advancements across three core domains: integrated appearance and internal quality assessment, pest/disease early warning, and intelligent grading. Crucially, we delve into the key engineering challenges of transitioning from laboratory settings to industrial sorting lines, addressing domain shift adaptation, model

lightweighting, multi-sensor integration, and flexible grasping. Finally, we identify four current bottlenecks—data scarcity, insufficient generalization, the engineering deployment gap, and model interpretability—and propose corresponding future developmental directions. This review aims to foster the deep integration of advanced sensing technologies with neural networks, ultimately enhancing the efficiency of the global fruit and vegetable supply chain and reducing food loss.

Keywords: Deep learning, Postharvest fruits, Non-destructive quality assessment, Intelligent grading, Industrial deployment

1 Introduction

The global food system is confronted with severe challenges regarding postharvest losses. Accurate quality assessment and intelligent grading represent critical technical interventions for mitigating postharvest losses. By enabling the timely identification of defective fruits and the precise classification of quality grades, these technologies facilitate optimized sorting decisions, directing fruits of varying quality to the most appropriate market channels or processing routes, thereby maximizing their commercial value and minimizing unnecessary waste.

For decades, postharvest fruit quality assessment has relied on a combination of destructive physicochemical analysis, manual sensory evaluation, and conventional machine vision. Although earlier fruit quality assessment methods improved non-destructiveness and inspection efficiency, recent reviews further confirm that destructive sampling, manual inspection, and conventional machine-vision pipelines remain limited by subjectivity, low scalability, and weak robustness under complex imaging and postharvest conditions (Fu and Sun 2026; Liu et al. 2026). Destructive physicochemical methods can quantify internal quality indicators such as soluble solids content (SSC) and firmness, but damage or consume samples, limiting their suitability for non-destructive inspection of every fruit. Conventional RGB-based machine vision primarily captures external appearance and requires complementary sensing modalities to assess internal quality.

The advent of deep learning (DL) has fundamentally altered this landscape, marking a paradigm shift from "feature engineering" to "feature learning" (Alzubaidi et al. 2021). By circumventing the limitations of handcrafted features and unlocking the capability to extract internal quality information from non-visible data modalities, DL has rapidly become the mainstream technical route for postharvest quality assessment. The detailed evolution of this paradigm shift, along with representative milestone studies, is systematically reviewed in Section 2.

In recent years, several reviews have explored technological advancements in this field from various perspectives. Although recent reviews have summarized fruit quality evaluation and deep learning-based image analysis from perspectives such as nondestructive devices, sensing modalities, classification, detection, and segmentation (Wang et al. 2026; Ma et al. 2026), they still provide limited organization of the full technical chain from data acquisition and architecture selection to grading decisions and industrial deployment. While providing valuable insights, these reviews exhibit common limitations: they predominantly focus on isolated technical dimensions, such as specific DL architectures or external quality assessment. Consequently, there is a lack of systematic discussion regarding critical aspects such as non-destructive internal quality evaluation (e.g., SSC, firmness, and internal disorders), early identification and risk warning for pests and diseases, the strategic evolution from single-index to multi-index combined grading, and industrial deployment challenges involving model lightweighting, multi-sensor integration, and flexible grasping. In short, existing literature often examines isolated links in the technical chain rather than providing a holistic analysis from data acquisition to industrial deployment.

To bridge the aforementioned research gaps, this review provides a comprehensive, whole-chain framework encompassing the entire process from non-destructive data acquisition to industrial deployment, as conceptually illustrated in Fig. 1. The remainder of this paper is structured as follows: Section 2 systematically reviews the

evolutionary trajectory of postharvest quality assessment from destructive testing to deep learning, clarifying the technical drivers behind this paradigm shift. Based on the data requirements of DL modeling, Section 3 structurally compares the characteristics and limitations of eight non-destructive sensing modalities. Following a "task–architecture–data" thread, Section 4 systematically outlines fundamental DL architectures and cross-task enhancement strategies. Section 5 synthesizes the technical routes of representative studies across three integrated core applications: combined external and internal quality evaluation, pest and disease early warning, and intelligent grading decisions. Section 6 delves into the critical engineering challenges of transitioning from the laboratory to the production line, including domain shift adaptation, model lightweighting, and flexible grasping. Finally, Sections 7 and 8 summarize current technical bottlenecks and highlight future research directions.

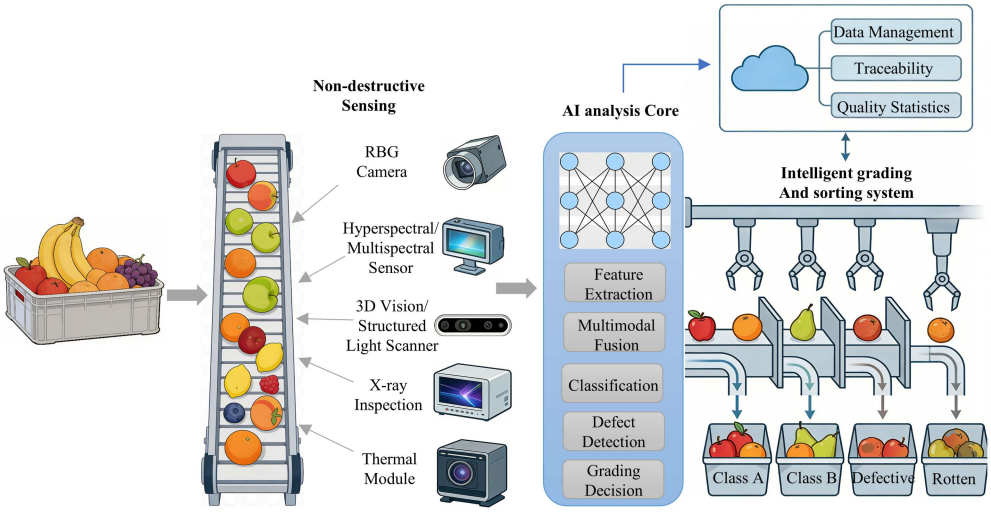


Fig. 1 Conceptual framework of deep learning-driven postharvest fruit quality assessment and intelligent grading. The framework integrates four sequential layers: (i) non-destructive multimodal sensing (RGB, hyperspectral/multispectral, 3D vision, X-ray, and thermal imaging), (ii) an artificial intelligence (AI) analysis core performing feature extraction, multimodal fusion, classification, defect detection, and grading decision, (iii) an intelligent grading and sorting system that physically routes fruits into quality classes, and (iv) a cloud-based platform for data management, traceability, and quality statistics. This whole-chain perspective constitutes the structural backbone of this review

2 Evolution of Postharvest Quality Assessment: From Traditional Methods to AI-Driven Paradigms

Over the past decades, this field has progressed through four distinct stages: destructive physicochemical analysis, manual sensory evaluation, automated machine vision, and deep learning-driven intelligent assessment. Each technological leap has been driven by the need to systematically overcome the limitations of the preceding stage regarding non-destructiveness, consistency, inspection efficiency, and generalization capability. This section follows this evolutionary timeline to review the core methods, application scenarios, and inherent limitations of each stage, elucidating the technical motivations for the integration of DL and laying the methodological groundwork for the remainder of this review.

2.1 Destructive Testing and Sensory Evaluation

Destructive physicochemical analysis represents the earliest and most direct approach to postharvest quality assessment. These methods typically involve extracting juice samples to measure SSC via refractometry, titratable acidity (TA) through titration, and flesh firmness (Wang et al. 2025a). However, their inherent sample consumption and tedious pre-processing limit their application to representative sampling rather than online total inspection. Similarly, sensory evaluation relies on trained panels to evaluate multidimensional attributes, but remains heavily affected by evaluator fatigue, physiological conditions, and long training cycles (Khan et al. 2025). Collectively, the deficiencies of these two approaches in terms of non-destructiveness, efficiency, and scalability served as the primary drivers for the development of automated non-destructive testing technologies.

2.2 Traditional Machine Vision and Shallow Machine Learning

Machine vision is characterized by the "manual feature extraction + shallow classifier" paradigm. In this approach, engineers design feature extraction algorithms based on prior knowledge—such as Otsu thresholding for segmentation, Canny edge detection, and texture descriptors like Histogram of Oriented Gradients (HOG) and Local Binary Patterns (LBP) followed by classification or regression using models like Support Vector Machines (SVM) or Random Forests. This stage achieved non-destructive, non-contact automated appearance assessment and demonstrated engineering feasibility under structured conditions (e.g., fixed lighting and single cultivars) (Xu et al. 2025).

However, the performance of traditional machine vision depends heavily on manually designed features and fixed decision thresholds, resulting in an inherent lack of generalization capability. Low-level features extracted by traditional algorithms exhibit weak robustness when facing common interferences such as lighting fluctuations, fruit occlusion, and varietal variations in actual industrial environments (Hou et al. 2023; Lv et al. 2025). Comparative studies have shown that traditional methods yield significantly lower accuracy than DL in fruit defect detection tasks (Fan et al. 2020). Conventional RGB-based machine vision primarily captures surface characteristics, whereas internal quality assessment requires sensing modalities sensitive to tissue composition or structure.

2.3 Deep Learning: Paradigm Shift from Feature Engineering to Feature Learning

The introduction of DL has shifted quality assessment from "feature engineering" to a new "feature learning" paradigm, offering a systematic response to the aforementioned technological bottlenecks. This paradigm shift is reflected in three dimensions. First, regarding the workflow, DL employs an end-to-end mode to map image inputs

directly to quality outputs, replacing multi-stage pipelines and eliminating cumulative errors while reducing reliance on domain expertise (Guo et al. 2025). Second, in terms of representation capability, Convolutional Neural Networks (CNNs) automatically learn hierarchical features—ranging from low-level textures to high-level semantics—via multiple layers of non-linear transformations, enabling the detection of subtle defects and nuanced maturity differences. Third, concerning data adaptability, pre-training and transfer learning strategies allow visual representations learned from large-scale general datasets (e.g., ImageNet) to be transferred to small-sample agricultural tasks, mitigating the scarcity of labeled data in postharvest scenarios.

These breakthroughs have established DL as the mainstream technical route for postharvest quality assessment. To date, DL-based systems have achieved exceptional accuracy (typically exceeding 90%) in fruit cultivar classification and surface defect detection (Jahanbakhshi et al. 2020). Cutting-edge research has further expanded application depth: Yang et al. (2023) integrated CNNs with the LW-Swin Transformer architecture to achieve high-precision image recognition of strawberry ripeness. Beyond classification accuracy, DL provides flexible approaches for learning complex relationships between spectral, X-ray, or acoustic measurements and internal fruit quality attributes. Concurrently, advancements in lightweight network design (e.g., MobileNet, EfficientNet) and model compression techniques (e.g., knowledge distillation, quantization) are facilitating the transition of DL from laboratory settings to real-time deployment on edge devices, narrowing the gap between laboratory accuracy and industrial throughput. Detailed discussions regarding specific DL architectures, task adaptation, and performance in the postharvest sector will follow in Sections 4 and 5.

3 Non-Destructive Data Acquisition Modalities

The performance ceiling of DL models is largely dictated by the richness and quality of the input data. In postharvest quality assessment, diverse non-destructive sensing modalities capture fruit appearance features, chemical compositions, and internal structures across multiple physical dimensions, providing heterogeneous data inputs for DL algorithms. This chapter outlines the informational characteristics, typical applications, and respective limitations of primary non-destructive acquisition modalities. A comparative summary of their core parameters is provided at the end of the chapter to establish the data-level context for subsequent architecture selection (Section 4) and application analysis (Section 5).

3.1 RGB Imaging and 3D Vision

Red-green-blue (RGB) cameras capture reflectance signals in the visible spectrum to directly characterize the color, surface texture, and macroscopic morphological features of fruits, representing the most cost-effective data acquisition method in postharvest assessment (Santelices et al. 2025; Abebe et al. 2023). When coupled with deep learning, hierarchical visual features can be automatically extracted from RGB images, demonstrating superior performance in tasks such as maturity detection, surface defect identification, and cultivar classification (Ma et al. 2024b; Goyal and Lakhwani 2025). Nevertheless, RGB imaging faces two fundamental constraints: high sensitivity to ambient lighting, where unstable illumination significantly compromises image quality and grading precision (Khan et al. 2025); and the inability to penetrate the peel for internal quality inspection due to its surface-level sensing nature.

3D vision technologies—utilizing structured light, binocular stereo vision, or Time-of-Flight principles—acquire point cloud data of fruit surfaces, offering a geometric informational dimension distinct from RGB textures. Recent reviews on three-dimensional analysis of fruit and vegetables indicate that structured light, stereo

vision, time-of-flight, and related 3D sensors can provide geometric traits such as volume, size, shape, and surface morphology for grading and categorization, but their practical use must still consider sensor cost, reconstruction accuracy, and processing efficiency (Li et al. 2026). However, the computational latency and cost associated with 3D reconstruction hinder its industrial deployment; for instance, Timprae et al. (2025) reported that generating a single 3D tomato model requires approximately 1.5 hours, which fails to meet real-time sorting requirements. In summary, while RGB imaging and 3D vision form the data foundation for external quality assessment, their lack of chemical sensitivity necessitates complementarity with spectroscopic techniques.

3.2 Hyperspectral and Multispectral Imaging

Hyperspectral imaging (HSI) detects selective absorption and scattering caused by chemical constituents within fruit tissues, thereby enabling the non-destructive extraction of internal physicochemical and safety-related information; recent Spectral-AI reviews further show that hyperspectral imaging combined with ML/DL can capture spatial biochemical signatures and support high-throughput screening for quality and spoilage-related hazards (Mohammadi et al. 2026). For instance, Chen et al. (2022) utilized HSI in conjunction with neural networks to predict the moisture content of persimmons and generate spatial distribution maps, comprehensively highlighting the unique advantages of spatial-spectral joint modeling.

Regarding defect detection, HSI is capable of identifying early-stage latent damages that remain invisible to the naked eye. The underlying principle is that many postharvest defects initially trigger alterations in the tissue's chemical composition, and these physiological changes manifest in specific spectral bands prior to the appearance of visible macroscopic symptoms. For example, Feng et al. (2025) combined visible-near-infrared hyperspectral imaging with a CNN-BiLSTM model to detect early subtle bruising in strawberries.

However, HSI is constrained by massive data volumes and exorbitant equipment costs, making its direct application on high-speed sorting lines highly challenging. Multispectral imaging (MSI) addresses this by employing effective wavelength selection algorithms to isolate a limited number of representative characteristic bands, thereby reducing data dimensionality while preserving critical predictive information. Therefore, MSI serves as an important pathway for transitioning hyperspectral technology into industrial scenarios. Furthermore, snapshot hyperspectral systems can now capture the complete data cube in a single exposure (Kothawade et al. 2026). These technological advancements indicate that spectral imaging is steadily progressing toward miniaturization, cost-effectiveness, and real-time operational capability.

3.3 Near-Infrared Spectroscopy

Near-infrared spectroscopy (NIRS) primarily exploits the overtone and combination absorption properties of hydrogen-containing groups (e.g., O-H, C-H, N-H) within organic molecules of fruit tissues in the 780–2500 nm spectral range, enabling the non-destructive quantitative evaluation of internal physicochemical indicators such as SSC, TA, and firmness. Unlike the three-dimensional data structure of HSI, NIRS typically acquires one-dimensional spectral signals via point measurements (lacking spatial resolution). However, owing to its exceptional hardware maturity and rapid acquisition speed, NIRS remains the only spectroscopic sensing technology deployed on a large scale within industrial sorting lines.

The detection efficacy of NIRS is heavily constrained by optical penetration depth and the applied measurement mode. Research indicates that within the 720–880 nm wavelength range, the penetration depth of near-infrared light in fruit tissues generally varies between 11 and 21 mm, depending on the excitation wavelength and the fruit's optical properties (Liu et al. 2025b). Consequently, NIRS has evolved into three distinct detection modes—reflectance, transmittance, and interactance—to flexibly

accommodate diverse assessment requirements (Xiao et al. 2025). Regarding internal quality evaluation, SSC prediction constitutes the most mature application: Zeng et al. (2024) employed CNNs coupled with NIR spectra to predict apple SSC, achieving R^2 of 0.92, significantly outperforming traditional partial least squares regression (PLSR) models. In contrast, the prediction accuracy for acidity and firmness is often limited; for instance, O'Brien et al. (2024) reported that the R^2 for mango firmness prediction using NIRS was below 0.5.

The integration of deep learning is effectively overcoming the performance bottlenecks of traditional PLSR models in deciphering high-dimensional, non-linear spectral data. Recent reviews on multidimensional CNN architectures in food quality assessment indicate that 1D-CNNs, 2D-CNNs, and 3D-CNNs can process spectral, image, and spectral-spatial fused data, respectively, enabling deep feature extraction and multimodal representation learning while still facing challenges related to data quality, model complexity, interpretability, computational cost, and generalization (Deng et al. 2026). Zhang et al. (2025b) developed a dual-attention-enhanced 1D-CNN model, successfully realizing the simultaneous and precise prediction of SSC and TA in melons. For peaches, Bu et al. (2025) combined visible/near-infrared (Vis/NIR) spectroscopy with a multi-task CNN to predict multiple internal quality attributes across varieties. Similarly, the 1D-Inception-ResNet architecture constructed by Yu et al. (2025) demonstrated superior performance over conventional single-branch architectures in spectral quantitative analysis tasks involving multiple fruit cultivars. These cutting-edge studies signify that deep learning is driving a fundamental transition in NIRS technology—from conventional single-parameter linear regression toward multi-parameter, non-linear joint evaluation.

3.4 Other Sensing Modalities

In addition to mainstream optical imaging and spectroscopic technologies, sensing modalities such as X-ray/computed tomography (CT), acoustic response, and thermal imaging exhibit unique advantages in specific application scenarios, providing highly differentiated feature inputs for deep learning models.

X-ray and CT imaging can penetrate deep fruit tissues, enabling the visualization of severe internal physiological disorders such as bitter pit, chilling injury, and internal browning (Wood et al. 2025). Van De Looverbosch et al. (2021) utilized machine learning algorithms to analyze CT images of pears, successfully distinguishing healthy from defective fruits with an accuracy exceeding 90%. However, prohibitive equipment costs, slow scanning velocities, and underlying radiation safety concerns severely restrict their application in industrial sorting lines. Currently, these technologies are predominantly confined to the sampling inspection of premium fruits and fundamental scientific research.

Acoustic response technology evaluates texture and firmness by analyzing the vibration signals generated after a mechanical impact on the fruit. Tian et al. (2022) reported a firmness prediction model for kiwifruit based on this technology, achieving R^2 of 0.96. Although this modality provides unique mechanical feature inputs for deep learning, it requires high morphological regularity in fruits and is highly susceptible to environmental temperature and humidity fluctuations (Ding et al. 2021).

Thermal imaging constructs surface temperature distribution maps based on the thermophysical differences between damaged and healthy tissues, demonstrating robust immunity to external illumination conditions. Melesse (2026) developed an intelligent banana sorting system combining a lightweight CNN with thermal imaging, which not only achieved a high classification accuracy of 97% but also successfully executed real-time inference on general-purpose processors.

In summary, the depth and hierarchy of internal information probed by different sensing technologies vary significantly. RGB and 3D vision technologies are restricted to characterizing surface texture and macroscopic morphology. HSI and MSI can effectively capture early subsurface latent damage. NIRS is primarily utilized for the quantitative evaluation of bulk physicochemical indicators such as SSC and firmness. Conversely, X-ray and CT imaging can deeply penetrate the fruit to precisely detect profound physiological disorders, including internal browning and watercore disease. These varying penetration capabilities naturally necessitate the multimodal data fusion strategies discussed next.

3.5 Multimodal Data Fusion

Because a single sensing modality can only capture a unidimensional aspect of fruit quality, multimodal data fusion integrates sensory information derived from diverse physical principles to achieve a more comprehensive and robust quality assessment. Based on the hierarchical level of information integration, fusion strategies are primarily categorized into three types. Early fusion concatenates multimodal raw data at the data level following spatial registration before feeding it into the network. This approach maximally preserves complementary information but imposes stringent requirements on spatiotemporal registration precision (Zhao et al. 2024; Liang et al. 2025). Zhao et al. (2025) employed a dual-branch model to fuse near-infrared spectroscopy, HSI, and RGB images, achieving precise discrimination of growth stages across multiple navel orange cultivars. Late fusion involves independent modeling for each modality, followed by weighted integration at the decision level. This strategy is characterized by a high degree of modularity and strong fault tolerance. He et al. (2026b) adopted this strategy for predicting the freshness of fragrant pears, resulting in a 47.44% increase in accuracy compared to single-modality models. Hybrid fusion utilizes cross-modal attention mechanisms to facilitate deep information interaction

at both the feature and decision levels. The ITF-WPI model developed by Dai et al. (2023) exemplifies this strategy, achieving a high accuracy of 97.98% in the recognition of wolfberry pests.

The structural comparisons of the three fusion strategies are illustrated in Fig. 2. Taking the combination of RGB images and spectral data as an example: in early fusion, RGB and spectral data are aligned and concatenated at the data level before being fed into a unified neural network for end-to-end learning. In late fusion, each data modality is processed in parallel by independent neural networks to output individual predictive confidence scores (e.g., 90% and 85%); these branch outputs are subsequently integrated via a weighted fusion mechanism to yield a comprehensive decision. In hybrid fusion, preliminary features are first extracted independently from the RGB and spectral images. Bidirectional information interaction is then established at the feature level using cross-modal attention mechanisms, leading to a final determination output after deep feature fusion.

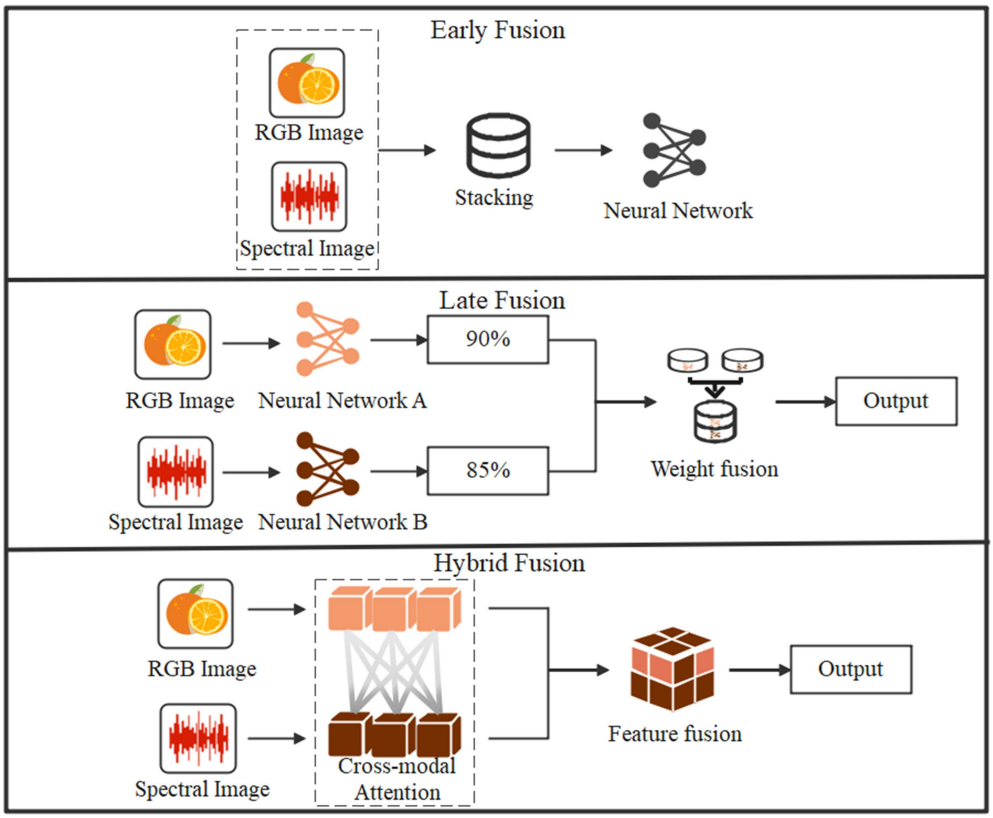


Fig. 2 Topological architectures of early, late, and hybrid multimodal fusion strategies

In a related traceability application, Gao et al. (2025) combined NIR spectra and RGB images to identify the geographical origin of apples, illustrating the application of multimodal feature fusion beyond quality prediction.

To overcome the insensitivity of single modalities (e.g., RGB images) to latent damage, the scope of multimodal fusion has expanded significantly beyond traditional optical-spectral combinations. In the spatial dimension, fusing 3D geometry with hyperspectral features enables comprehensive grading of pears (Song et al. 2024). Across volatile and penetrative dimensions, the integration of metal-oxide-semiconductor electronic noses, volatile organic compound (VOC) sensing, and X-ray

imaging has been successfully applied to kiwifruit maturity prediction, strawberry shelf-life evaluation, and the localization of internal disorders in pears (Du et al. 2019; Do et al. 2024; Tempelaere et al. 2023). Furthermore, hybrid networks incorporating visual-tactile scanning (e.g., SwishFormer) have broadened the technical pathways for fruit firmness recognition (Mohsan et al. 2025).

Currently, multimodal data fusion is rapidly evolving toward lightweight design and real-time processing. Fan et al. (2025) demonstrated the feasibility of deploying lightweight fusion strategies on resource-constrained edge devices. Furthermore, Zou et al. (2025) integrated multimodal perception with digital twin technology and long short-term memory (LSTM) networks, successfully constraining the shelf-life estimation error of highly perishable fruits to within ± 1.2 days.

3.6 Comparison and Summary of Sensing Modalities

Table 1 systematically compares the aforementioned non-destructive sensing modalities across key dimensions, including data type, detectable quality indicators, inspection speed, equipment cost, and industrial deployment maturity. As evidenced by Table 1, no single sensing modality can independently satisfy all the requirements for a comprehensive postharvest quality assessment. Leveraging their high technological maturity and cost-effectiveness, RGB imaging and NIRS have pioneered large-scale industrial deployment. In contrast, although HSI and X-ray/CT technologies offer unique advantages in terms of informational probing depth, their current applications remain predominantly confined to laboratory research. Consequently, multimodal data fusion has emerged as an inevitable trend to enhance the comprehensiveness and robustness of quality evaluations. However, in practical applications, the selection of a fusion strategy must strike a rigorous balance among informational integrity, computational complexity, and industrial feasibility. The selection of deep learning architectures and their specific task adaptations corresponding to these data modalities will be discussed in detail in Section 4.

Table 1 Comparison of sensing modalities for postharvest fruit NDT

Modality	Data Type	Indicators	Speed	Cost	Deployment	DL Pros	Limitations
RGB Imaging	2D color image	Color, defects, shape, maturity	Fast	Low	High	Pre-trained models; easy transfer learning	Only surface info; light-sensitive; no internal quality
3D Vision	Point cloud / depth map	Volume, area, shape, deformity	Slow–Med	Med–High	Low	Geometric features; volume regression	Slow reconstruction; high cost; poor real-time
HSI	3D spectral cube	SSC, moisture, acidity, latent damage	Slow	High	Low–Med	Spatial-spectral modeling; multi-task	Big data; complex preprocessing; expensive
MSI	Multi-channel image	SSC, maturity, defects	Med–Fast	Medium	Medium	Key spectral features; fast deployment	Band selection relies on prior knowledge; less info than HSI
NIRS	1D spectrum	SSC, acidity, firmness	Fast	Medium	High	1D-CNN/LSTM compatible	No spatial resolution; limited accuracy
X-ray / CT	Grayscale / 3D voxel	Internal browning, cavities, pest damage	Slow	Very High	Low	Deep internal defect detection; 3D segmentation	High cost; slow; radiation risk
Acoustic	1D vibration signal	Firmness, texture, moisture	Fast	Low	Low–Med	Mechanical feature input	Sensitive to shape and noise
Thermal Imaging	Temperature map	Subsurface damage, early lesions	Med–Fast	Medium	Low–Med	Illumination-robust; multimodal fusion ready	Contrast depends on environment; limited alone

Note: NDT = Non-Destructive Testing.

4 Deep Learning Frameworks and Datasets for Postharvest Contexts

Postharvest fruit quality assessment presents specific industrial vision challenges, including lighting fluctuations, strong specular reflections, fruit occlusion, and blurred boundaries of weak defects. Traditional shallow models often exhibit insufficient generalization across cultivars and imaging conditions, whereas DL and transfer-learning strategies can improve representation robustness when source-target similarity, domain mismatch, and negative transfer are carefully considered, especially in small-sample image classification tasks (LeCun et al. 2015; Plested et al. 2026). Tailored to specific assessment requirements, fundamental visual tasks can be categorized into object detection for defect localization, image segmentation for area quantification, and classification/regression for grading or physicochemical parameter prediction. Furthermore, cross-task enhancement strategies, such as attention mechanisms, elevate representational capacity in complex scenarios (Dosovitskiy et al. 2021). This chapter structurally reviews mainstream DL frameworks and open-source data resources in the postharvest domain following a "task-architecture-data" trajectory.

4.1 Object Detection Architectures

Object detection is primarily deployed to extract candidate regions for surface defects and lesions. In postharvest applications, one-stage detectors (typified by the YOLO series) serve as the preferred choice for real-time screening on high-speed sorting lines due to their exceptional inference speed. Recent iterations have further fortified their performance in complex environments: for instance, DBH-YOLO achieved rapid localization of surface defects in blue honeysuckle (Xue et al. 2026). In contrast, two-stage detectors (e.g., Faster R-CNN) provide a higher accuracy baseline in complex scenes with co-occurring multiple defect classes (Wan and Goudos 2020) but are better suited for offline quality inspection due to substantial inference overhead.

Additionally, Transformer-based detectors (e.g., SCORE-DETR) bypass cumbersome post-processing steps via global attention mechanisms, demonstrating robust efficacy against dense citrus occlusion and complex backgrounds (Zhang et al. 2025a), although their high training costs and sensitivity to data scale restrict broader deployment.

4.2 Image Segmentation Architectures

When grading standards rely on "defect area proportion" rather than binary defect presence, pixel-level image segmentation becomes indispensable. Semantic segmentation (e.g., U-Net-based SPMUNet) directly outputs defect masks on citrus surfaces, providing precise area quantification metrics for grading (Xu et al. 2024). For irregular boundaries or co-existing multiple lesions, instance segmentation further isolates individual instances of the same defect class. For example, Mask R-CNN has been successfully applied to quantify strawberry bruise severity under varying illumination (Zhou et al. 2022), while YOLO-ALDS combined with active learning achieved tomato defect segmentation and grading, representing a shift toward low-annotation-cost engineering (Xue et al. 2025). Recent studies indicate that although Transformer-based foundation models show zero-shot potential in pomegranate segmentation, conventional Mask R-CNN currently maintains an advantage in inference time and overall precision (Giménez-Gallego et al. 2024).

4.3 Classification and Regression Architectures

Whole-fruit grading and physicochemical parameter prediction predominantly rely on classification and regression architectures. Regarding image classification backbones, researchers must trade off between representation capability and deployment efficiency: heavyweight networks (e.g., ResNet50) exhibit higher stability against lighting or varietal shifts, whereas lightweight networks (e.g., MobileNet, EfficientNet) are better tailored for edge deployment (Lee et al. 2025; Arunima et al. 2024). The two-stage

classification system constructed by Tapia-Mendez et al. (2023), based on tandem MobileNetV2 models, further validates the engineering scalability of lightweight backbones in multi-task scenarios.

For sequential regression architectures handling 1D spectral signals (e.g., 1D-CNNs and LSTMs), please refer to the detailed discussion in Section 3.3. Furthermore, to counteract textural interference and strong reflections in close-range imaging, attention mechanisms effectively reinforce the representation of critical regions by dynamically adjusting feature weights. Embedding Transformer modules into CNNs (e.g., YOLOv8 fused with LW-Swin) has emerged as an effective strategy to balance local feature extraction efficiency with global dependency modeling.

4.4 Public Datasets

High-quality open-source data is foundational for driving methodological comparability and reproducibility in postharvest quality assessment. Table 2 summarizes representative datasets oriented toward defect recognition and freshness discrimination. However, current data resources exhibit significant limitations: the majority are restricted to RGB classification labels, lacking non-visible modalities (e.g., hyperspectral and X-ray imaging) and pixel-level segmentation annotations; furthermore, a distinct gap exists between controlled laboratory acquisition and actual industrial conditions. Consequently, constructing multimodal, finely annotated benchmark datasets that mirror real-world industrial scenarios represents a critical bottleneck for future research. The specific performance and evolutionary trends of these DL architectures across core postharvest applications will be detailed in Section 5.

Table 2 Representative public datasets in the field of postharvest fruit quality assessment

Dataset	Sample Size	Image Size	No. of Classes	Fruit Types	Fresh Label	Rotten Label	Supported Tasks
Fruit 360 (Mureşan and Oltean 2018)	90,483	100 × 100	131	83	✓	–	Classification
Fruit Freshness (Kang and Gwak 2022)	6,300	224 × 224	14	7	✓	✓	Classification
(Hussain et al. 2018)	44,406	320 × 258	15	10	✓	–	Classification
Papaya Fruit (Behera et al. 2021)	300	227 × 227	3	1	✓	–	Classification
(Kalluri 2018)	10,901	416 × 416	6	3	✓	✓	Classification
(Seth 2020)	4,320	618 × 618	36	10	✓	✓	Recognition
Tomato (da Costa et al. 2020)	43,843	100 × 100	2	1	✓	–	Classification
(Mukhiddinov et al. 2022)	12,000	Various	20	5	✓	✓	Detection
Chili (Junyang et al. 2026)	1,080	2976 × 3968	2	1	✓	✓	Classification / Detection / Segmentation
Fruit Classification (Kaizen 2020)	22,495	100 × 100	33	33	–	–	Classification
Healthy–Defective–Fruits (Pacheco et al. 2023)	20,000	Various	4	2	✓	✓	Classification
(Cofone 2025)	29,091	Various	2	14	✓	✓	Classification

Note: For the Papaya dataset, the sample size refers to the 300 original images reported by Behera et al. (2021).

5 Core Applications in Quality Assessment and Intelligent Grading

This chapter shifts the focus from "technological tools" to "application logic," systematically analyzing the technical evolutionary patterns and common bottlenecks of DL across four scenarios: external assessment, internal inspection, pest/disease warning, and intelligent grading. The discussion is supported by two complementary tables: Table 3 provides a subtask-level breakdown of representative studies on appearance quality, while Table 4 offers an application-domain overview of milestone breakthroughs across all four scenarios.

5.1 Integrated Assessment of External and Internal Quality

External quality is the primary criterion for market access, while internal physicochemical indicators (e.g., SSC, firmness) dictate the actual eating quality of the fruit. Synthesized analysis of recent progress highlights three predominant evolutionary trends:

First, the paradigm shift from classification to pixel-level quantification. To align with industrial grading standards centered on the "percentage of damaged area," research focus has transitioned from binary "defect presence" classification to pixel-level segmentation. Architectures such as Mask R-CNN (Zhou et al. 2022) and YOLO-ALDS (Xue et al. 2025) enable precise quantification of bruise and lesion areas, directly feeding into area-ratio-based grading rules.

Second, the leap from RGB single-modality to multi-source information fusion. To address latent damages or internal constituents poorly characterized by RGB images, multimodal architectures (as detailed in Section 3.5) have become the mainstream solution. On an application level, sensing dimensions have expanded from optical-spectral enhancements (e.g., HSI, NIRS) into non-optical domains, seamlessly

integrating visuotactile sensors for firmness prediction (Ma et al. 2024a) or VOC sensing for metabolic monitoring into the overall quality evaluation framework.

Third, DL-driven replacement of traditional chemometrics. In internal quality prediction, end-to-end 1D-CNNs and transfer learning architectures are systematically overcoming the accuracy ceilings and generalization bottlenecks of traditional PLSR models by capturing complex non-linear mapping relationships. Wang et al. (2025b) demonstrated cross-cultivar generalization for the simultaneous prediction of SSC, firmness, and moisture content using a transfer learning network, while CT-based feature extraction has been validated for internal browning detection using only a single 2D slice—drastically reducing scan time relative to full 3D acquisition (Wood et al. 2024).

Table 3 Performance comparison of representative studies on appearance quality assessment

Subtask	Study and method	Data and target	Key performance
Surface defect detection	YOLO-ALDS with active learning and instance segmentation (Xue et al. 2025)	RGB; tomato	Integrated defect segmentation and automatic grading with reduced annotation cost.
	VIS–NIR polarized HSI with chemometric feature analysis (Faqeerzada et al. 2024)	VIS–NIR HSI; sweet pepper	Enhanced surface defect characterization under complex surface conditions.
Mechanical damage detection	Mask R-CNN instance segmentation (Zhou et al. 2022)	RGB, multi-illumination; strawberry	Bruise region segmentation and area-ratio-based grading.
	Spectral–spatial enhanced 3D CNN with region-based hyperspectral analysis (Liu et al. 2025a)	VIS–NIR HSI; apple	Detection of early mild bruises at the RGB-invisible stage.
Shape and size measurement	Reconstructed 3D hyperspectral imaging (Song et al. 2024)	HSI + 3D; pear	Comprehensive grading based on volume and texture traits.
	SfM-based 3D reconstruction with kaolin particle film (Masoudi et al. 2024)	3D imaging; tomato	Improved geometric reconstruction accuracy under highly reflective conditions.
Color and maturity assessment	LS-YOLOv8s with LW-Swin Transformer (Yang et al. 2023)	RGB; strawberry	Maturity detection under occlusion and clustered-fruit scenarios.
	CNN-based multi-trait prediction model (Masuda et al. 2023)	RGB; peach	Prediction of skin color, firmness, and sugar content from RGB images.
Multi-fruit appearance grading	FruitVision based on MobileNetV3 (Hayat et al. 2024)	RGB; multiple fruits	Reported 97.96%–99.50% grading accuracy across eight fruit categories.

5.2 Pest and Disease Detection: From Classification to Early Warning

Pest and disease detection focuses on quality deterioration caused by pathogen infection or pest activity. Unlike simple external surface defects (such as mechanical bruises), pests and diseases often exhibit strong latency (indistinct early symptoms) and transmissibility (capable of spreading rapidly to adjacent fruits during storage and transportation). Therefore, the commercial and academic value of early detection and risk warning far exceeds that of post-symptomatic severe deterioration classification. To capture these subtle early signals, relying solely on RGB imaging is highly prone to missed detections; consequently, the research focus in this field has comprehensively shifted toward a synergistic strategy of "physical imaging enhancement coupled with deep models."

The evolution of postharvest pests and diseases presents typical temporal characteristics. In the initial stage of infection (the latent phase), since the surface tissue remains intact, it is imperative to leverage hyperspectral or multimodal imaging technologies to extract invisible biochemical variance signals. As the disease progresses to the mild symptom phase, the research focus shifts toward precisely localizing early lesions through lightweight CNNs and image segmentation techniques.

Guided by this evolutionary logic, the integration of non-destructive sensing and deep learning in postharvest early decay detection is driving the technology from feasibility verification to stable application. For instance, Li et al. (2023) pioneered the use of Structured-Illumination Reflectance Imaging (SIRI) to detect early decay in navel oranges. Addressing the practical production challenges of cross-cultivar generalization and sample scarcity, Cai et al. (2024) proposed a universal classification model for various citrus types to enhance cross-cultivar transferability.

Spectral-spatial modeling provides an approach to identifying postharvest fruit diseases when visible symptoms are limited. Guo et al. (2024) developed a dual-branch

selective attention capsule network to classify kiwifruit soft rot using hyperspectral images. The network extracted spectral and spatial features separately and fused them through a selective attention mechanism, illustrating the potential of spectral-spatial feature integration for postharvest disease identification.

Internal pest damage can be difficult to identify from surface appearance alone. He et al. (2026a) combined X-ray radiography with deep transfer learning to detect codling moth larval damage in 'Conference' pears. The study used samples collected across multiple growing seasons and simulation-based data augmentation, providing evidence for non-destructive internal pest inspection under the evaluated conditions.

5.3 Intelligent Grading Strategies: Closing the Loop between Perception and Execution

Intelligent grading is the ultimate engineering objective, transforming multi-source perception results into executable grading decisions. Technical focus has shifted from single-index sorting to multi-index joint grading. Realistic commodity grading requires a unified framework to process both discrete attributes (classification labels) and continuous attributes (SSC values), while flexibly configuring weights according to market demands. Furthermore, online real-time grading serves as the final test, requiring not only model lightweighting but also high-speed coordination between sensor arrays and actuators.

Table 4 Representative studies on non-destructive fruit quality assessment and intelligent grading (2024–2026)

Application Domain	Representative Study	Sensing Modality	Algorithm & Architecture	Key Breakthrough
Appearance Quality	Xue et al. (2025)	RGB	YOLO-ALDS (Active Learning + segmentation)	Integrated pixel-level defect segmentation and grading with significantly reduced annotation costs.
	Liu et al. (2025a)	VIS-NIR HSI	Spectral-spatial 3D CNN	Successfully detected early mild bruises at the RGB-invisible stage.
	Song et al. (2024)	HSI + 3D Vision	Reconstructed 3D HSI	Achieved comprehensive grading by fusing volume geometry and spectral texture traits.
Internal Quality	Wang et al. (2025b)	HSI	Transfer Learning Network	Realized cross-variety generalization for multiple indices (SSC, firmness, moisture).
	Ma et al. (2024a)	Vision + Tactile	Visuotactile Neural Network	Developed a non-destructive firmness prediction method adaptable for robotic flexible sorting.
	Wood et al. (2024)	2D CT Slice	Feature Extraction + Logistic regression	Proved effective internal browning detection using only a single CT slice, drastically reducing scan time.
Pest & Disease	Cai et al. (2024)	SIRI	General Classification Model	Achieved stable cross-variety transfer capability for early decay detection across multiple citrus cultivars.
	Yan et al. (2025)	Fluorescence	Multi-light-source Analysis	Amplified biochemical differences for rapid and non-destructive HLB diagnosis.
	He et al. (2026a)	X-ray Transmission	Transfer Learning	Non-destructive detection of codling moth damage in pears using multi-season samples and simulation-based data augmentation.
Intelligent Grading	Chen et al. (2026); Xue et al. (2025)	Camera images	Lightweight and active-learning YOLO models	Real-time banana freshness grading and tomato defect-based automatic grading

In propelling these application scenarios toward industrialization, three cross-directional challenges remain prominent: (1) the "high-cost annotation" dilemma due to the destructive nature of internal quality labeling; (2) "model performance collapse" triggered by cross-cultivar and cross-year variability; and (3) the "engineering gap" between high-precision algorithms and millisecond-level inference on sorting lines. These technical pain points will be thoroughly explored as core challenges of industrial deployment in Section 6.

6 From Laboratory to Industrial Sorting Lines

High precision in laboratory settings does not directly translate to industrial viability. Transitioning from offline evaluation under controlled conditions to stable operation on high-speed sorting lines requires overcoming a substantial engineering gap. This chapter explores the "last-mile" challenges, systematically discussing the differences between laboratory and industrial scenarios, model lightweighting, system integration, flexible grasping, and cost-benefit analysis.

6.1 Differences between Laboratory and Industrial Scenarios

As illustrated in Fig. 3, the gap between laboratory and industrial environments can be categorized into four dimensions. Regarding imaging conditions, industrial lines must handle illumination fluctuations and motion blur during high-speed movement, which are absent in static laboratory setups (van Dael et al. 2019). Regarding throughput, commercial sorting lines demand processing 10–12 fruits per second, requiring the entire chain—from imaging to execution—to be completed within 80–100 ms (Kumar Pothula et al. 2023). Regarding data diversity, models must adapt to systematic domain shifts across cultivars, batches, and production regions (Gidado et al. 2026). Lastly, regarding system robustness, long-term performance degradation due to light source aging and lens contamination must be addressed.

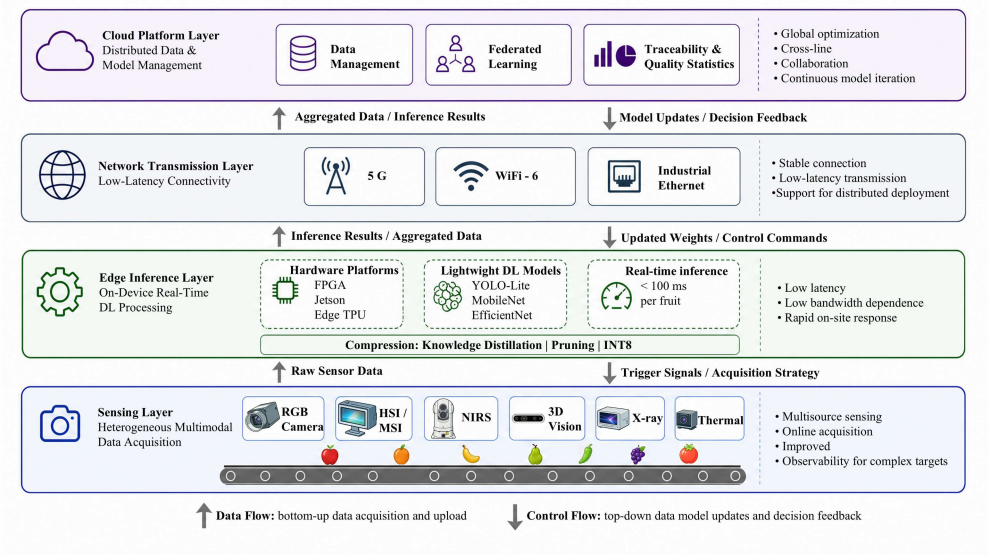


Fig. 3 Hierarchical Internet of Things (IoT) architecture for a flexible postharvest sorting system, illustrating the data and control flow across the sensing layer (RGB/HSI/NIR sensors), the edge inference layer (real-time DL models on field-programmable gate array (FPGA)/Jetson devices), the network transmission layer, and the cloud platform layer (data management, federated learning, traceability). This architecture forms the technical foundation for bridging laboratory algorithms and industrial throughput requirements

6.2 Model Lightweighting and Edge Deployment

Under stringent throughput constraints, deep learning models must achieve real-time inference on resource-limited edge devices. Current technical pathways to optimize this "speed-accuracy trade-off" operate at three interconnected levels. First, efficient architecture design utilizes lightweight backbones (e.g., MobileNetV3, EfficientNetV2) to inherently reduce computational overhead (Howard et al. 2019; Tan and Le 2021). Melesse (2026) demonstrated this by deploying a lightweight CNN for real-time banana sorting on general-purpose processors. Second, model compression techniques—such as knowledge distillation, structural pruning, and INT8 quantization (achieving 2–4× acceleration)—further shrink models without architectural redesign, drastically reducing deployment costs.

Finally, hardware platform adaptation is essential for edge deployment. Chen et al. (2026) implemented a lightweight YOLO-based banana freshness grading system on NVIDIA Jetson Orin NX and Jetson TX2 devices, demonstrating real-time inference for portable postharvest assessment. Moving forward, integrating "hardware-aware neural architecture search"—where device latency and power consumption act as direct optimization constraints during the search phase—represents a critical direction for achieving true algorithm-hardware collaborative design.

6.3 Multi-sensor System Integration

Industrial intelligent sorting lines can be regarded as Industry 4.0-oriented smart food systems that integrate sensor arrays, edge computing nodes, data-driven monitoring, and adaptive control modules, where sensor evaluation under real and risky environments is essential for quality, safety, sustainability, and traceability (Braggio et al. 2026). A primary challenge is precise spatiotemporal alignment: temporal synchronization relies on photoelectric triggers, while spatial registration requires pixel-level alignment across modalities to avoid compromising multimodal fusion (He et al. 2025). Furthermore, modular and reconfigurable designs are essential for lowering investment thresholds for small and medium enterprises (SMEs) by enabling "plug-and-play" sensor expansion.

6.4 Vision-guided and Flexible Grasping

To prevent secondary mechanical damage caused by traditional pneumatic pushers, flexible grasping has become a focal point for premium fruit sorting. Deep learning facilitates this by predicting grasp points via instance segmentation and dynamically adjusting gripping force based on predicted firmness. Liu et al. (2024) developed a soft bionic gripper that ensures damage-free handling through visuo-tactile closed-loop control. However, a significant throughput gap remains between flexible grasping

(1–3 s per fruit) and high-speed sorting requirements (0.1 s per fruit), confining its use primarily to high-value-added scenarios (Kumar Pothula et al. 2023).

6.5 Cost-Benefit Analysis and Technological Maturity Summary

Economic feasibility dictates the widespread adoption of intelligent grading systems. Costs encompass hardware investment, expensive annotation, and maintenance, while benefits include labor savings, improved grading consistency, and reduced postharvest losses (Zou et al. 2025). Currently, basic sorting systems based on RGB/NIR have reached high industrial maturity. Multi-sensor synchronization and edge deployment are technologically feasible but require scenario-specific optimization. Conversely, vision-guided flexible grasping and cross-cultivar adaptation are still in early transition stages. From an industrial perspective, future deployment should move beyond isolated grading equipment toward supply-chain-level quality monitoring systems that integrate non-invasive sensing, lightweight edge AI, multimodal data fusion, digital twins, and IoT infrastructure to improve traceability, sustainability, and decision-making (Mao et al. 2026). The technological gaps and maturity variations across these engineering domains provide the foundation for discussing future research directions in the subsequent sections.

7 Current Challenges and Future Directions

Despite the breakthrough progress of DL in postharvest fruit quality assessment, the field still faces three systemic barriers. First, data scarcity: the reliance on destructive analysis for internal quality labels results in a severe lack of ImageNet-scale benchmark datasets in the postharvest domain (Lu and Young 2020). Second, insufficient cross-domain generalization: models are highly susceptible to performance collapse when encountering new cultivars or production regions (Gidado et al. 2026). Third, the

deployment and trust gap: challenges ranging from throughput bottlenecks in edge computing to trade trust crises triggered by the "black-box" nature of models severely restrict industrialization. To overcome these barriers, future research should focus on the following three frontier directions:

7.1 Overcoming Data Barriers and Latent Internal Disease Detection

An effective approach to combat data scarcity involves reducing reliance on large-scale annotation. Leveraging the self-supervised pre-trained features or interactive segmentation capabilities of large-scale vision foundation models (such as the Segment Anything Model and DINOv2) can rapidly establish detection models for new cultivars with minimal annotations. Furthermore, a physical challenge far more daunting than data scarcity is the early detection of latent internal decay (e.g., moldy core, internal browning). Unlike surface defects, internal decay lacks visible external symptoms in its early stages. Existing deep-probing technologies (e.g., X-ray, CT) face industrial bottlenecks due to prohibitive equipment costs and exceedingly slow scanning speeds, making them incompatible with high-speed sorting lines; meanwhile, conventional NIRS is constrained by insufficient spatial resolution, often failing to detect localized lesions. Future research urgently needs to explore novel deep-penetrating sensing modalities (such as spatially offset Raman spectroscopy or acousto-optic multimodal fusion) coupled with 3D deep learning architectures to achieve low-cost, real-time, and non-destructive detection of internal decay.

7.2 Rebuilding Industrial Trust: Explainable AI and Standardization

Grading scenarios that directly influence commodity pricing have a rigid demand for model interpretability. Future AI frameworks must incorporate uncertainty awareness

(triggering human intervention for low-confidence samples) and embed physiological knowledge corresponding to specific spectral bands to generate auditable decision logs. Furthermore, multidisciplinary standardization is urgently needed. This includes constructing large-scale, open-source multimodal benchmark datasets covering diverse cultivars and regions; establishing unified model evaluation protocols and cross-domain testing procedures; and formulating industrial acceptance standards for intelligent sorting systems.

7.3 Reconstructing Business Logic: Edge-Cloud Collaboration and Full-Chain Integration

Next-generation sorting systems will evolve from isolated quality-control nodes into distributed data engines. Architecturally, they will adopt a collaborative model of "real-time edge inference + cloud-based federated learning/global optimization." Commercially, grading data should extend beyond immediate sorting: tracking upstream orchard harvest quality to guide agricultural management, and coupling with downstream cold-chain monitoring to establish quality degradation kinetic models for dynamic shelf-life prediction and logistics scheduling (Zou et al. 2025).

8 Conclusions

The introduction of deep learning has brought a profound methodological paradigm shift to postharvest fruit quality assessment. Through end-to-end feature learning, detection systems have completely eliminated reliance on manually engineered features. This has not only achieved exceptional precision in external defect segmentation but also broken the performance ceilings of traditional non-destructive testing in internal multi-parameter collaborative prediction and early warning of latent diseases.

Consequently, intelligent grading is steadily progressing from coarse single-index classification to multimodal, multi-attribute joint decision-making aligned with realistic market standards.

Looking ahead, translating laboratory "theoretical high precision" into "industrial efficacy" across global fruit and vegetable supply chains necessitates overcoming the systemic engineering gaps of data scarcity, fragile generalization, and edge deployment. Postharvest intelligent assessment systems are rapidly evolving into IoT-driven core data engines. By leveraging multimodal foundation models to bypass sample limitations, breaking through the dual hardware and algorithmic bottlenecks of real-time latent internal decay detection, rebuilding trade trust through explainable AI, and deeply integrating front-end harvesting, mid-end sorting, and back-end cold-chain logistics, a full-lifecycle quality data closed loop will be established. Realizing this ambitious vision urgently requires profound interdisciplinary integration among agronomy, computer science, and optical engineering, ultimately providing the definitive solution to reducing global food loss and ensuring food security.

Statements and Declarations

Author contributions. Yikun Huang conceived the review, defined its scope and structure, analyzed and synthesized the literature, and critically revised the manuscript for important intellectual content. Yongchun Zhan, Yunzhong Zhou, and Wenkai Wang conducted the literature search and screening, prepared the figures and tables, and wrote the initial draft. Jie Pang and Riqing Chen supervised the work. All authors read and approved the final manuscript.

Funding. This research was funded by the National Natural Science Foundation of China (No. 32660418), the Natural Science Foundation of Fujian Province (No. 2026J001384), the Fujian Agriculture and Forestry University Science and Technology Innovation Fund (No. KFB24043A).

Competing Interests. The authors declare no conflicts of interest.

References

Abebe AM, Kim Y, Kim J, et al (2023) Image-Based High-Throughput Phenotyping in Horticultural Crops. Plants 12:2061. https://doi.org/10.3390/plants12102061

Alzubaidi L, Zhang J, Humaidi AJ, et al (2021) Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. Journal of Big Data 8:53. https://doi.org/10.1186/s40537-021-00444-8

Arunima P, Gopinath PP, Geetha Lekshmi P, et al (2024) Digital assessment of post-harvest Nendran banana for faster grading: CNN-based ripeness classification model. Postharvest Biology and Technology 214:112972. https://doi.org/10.1016/j.postharvbio.2024.112972

Behera SK, Rath AK, Sethy PK (2021) Maturity status classification of papaya fruits based on machine learning and transfer learning approach. Information Processing in Agriculture 8:244–250. https://doi.org/10.1016/j.inpa.2020.05.003, URL https://data.mendeley.com/datasets/rcy6y8fhcn/1, dataset repository; accessed 11 October 2026

Braggio M, Melesse TY, Arena S, et al (2026) Sensor technologies in the food industry under industry 4.0: A systematic review. Food Control 187:112183. https://doi.org/10.1016/j.foodcont.2026.112183

Bu Y, Luo J, Tian Q, et al (2025) Nondestructive detection of internal quality in multiple peach varieties by Vis/NIR spectroscopy with multi-task CNN method. Postharvest Biology and Technology 227:113579. https://doi.org/10.1016/j.postharvbio.2025.113579

Cai Z, Sun C, Zhang H, et al (2024) Developing universal classification models for the detection of early decayed citrus by structured-illumination reflectance imaging coupling with deep learning methods. Postharvest Biology and Technology 210:112788. https://doi.org/10.1016/j.postharvbio.2024.112788

Chen L, Zhu J, Gui Y, et al (2026) Real-time banana freshness grading: A portable end-to-end detection system with high precision. Postharvest Biology and Technology 231:113904. https://doi.org/10.1016/j.postharvbio.2025.113904

Chen X, Jiao Y, Liu B, et al (2022) Using hyperspectral imaging technology for assessing internal quality parameters of persimmon fruits during the drying process. Food Chemistry 386:132774. https://doi.org/10.1016/j.foodchem.2022.132774

Cofone L (2025) Fruit and Vegetable (Freash vs Rotten). Kaggle, https://doi.org/10.34740/KAGGLE/DSV/12010423, URL https://www.kaggle.com/datasets/zlatan599/fruitquality1/versions/1, accessed 11 October 2026

da Costa AZ, Figueroa HEH, Fracarolli JA (2020) Computer vision based detection of external defects on tomatoes using deep learning. Biosystems Engineering 190:131–144. https://doi.org/10.1016/j.biosystemseng.2019.12.003

van Dael M, Verboven P, Zanella A, et al (2019) Combination of shape and X-ray inspection for apple internal quality control: *in silico* analysis of the methodology based on X-ray computed tomography. Postharvest Biology and Technology 148:218–227. https://doi.org/10.1016/j.postharvbio.2018.05.020

Dai G, Fan J, Dewi C (2023) ITF-WPI: Image and text based cross-modal feature fusion model for wolfberry pest recognition. Computers and Electronics in Agriculture 212:108129. https://doi.org/10.1016/j.compag.2023.108129

Deng W, Qin Q, Zhao J, et al (2026) Advances and challenges in multidimensional architectural applications of 1D/2D/3D convolutional neural networks in food quality assessment. Food Chemistry 505:147961. https://doi.org/10.1016/j.foodchem.2026.147961

Ding C, Feng Z, Wang D, et al (2021) Acoustic vibration technology: Toward a promising fruit quality detection method. Comprehensive Reviews in Food Science and Food Safety 20:1655–1680. https://doi.org/10.1111/1541-4337.12722

Do E, Kim M, Ko DY, et al (2024) Machine learning for storage duration based on volatile organic compounds emitted from 'Jukhyang' and 'Merry Queen' strawberries during post-harvest storage. Postharvest Biology and Technology 211:112808. https://doi.org/10.1016/j.postharvbio.2024.112808

Dosovitskiy A, Beyer L, Kolesnikov A, et al (2021) An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. https://doi.org/10.48550/arXiv.2010.11929

Du D, Wang J, Wang B, et al (2019) Ripeness Prediction of Postharvest Kiwifruit Using a MOS E-Nose Combined with Chemometrics. Sensors 19:419. https://doi.org/10.3390/s19020419

Fan C, Song K, Zou B, et al (2025) DA-Distill: Dual-Alignment Distillation for Multimodal Knowledge Transfer on Edge Devices. IEEE Transactions on Consumer Electronics 71:8055–8066. https://doi.org/10.1109/TCE.2025.3598319

Fan S, Li J, Zhang Y, et al (2020) On line detection of defective apples using computer vision system combined with deep learning methods. Journal of Food Engineering 286:110102. https://doi.org/10.1016/j.jfoodeng.2020.110102

Faqeerzada MA, Kim YN, Kim H, et al (2024) Hyperspectral imaging system for pre- and post-harvest defect detection in paprika fruit. Postharvest Biology and Technology 218:113151. https://doi.org/10.1016/j.postharvbio.2024.113151

Feng R, Han X, Lan Y, et al (2025) Detection of Early Subtle Bruising in Strawberries Using VNIR Hyperspectral Imaging and Deep Learning. Vibrational Spectroscopy 138:103786. https://doi.org/10.1016/j.vibspec.2025.103786

Fu Y, Sun DW (2026) Advancing Fruit Quality Monitoring with Artificial Intelligence-Augmented Non-Destructive Technologies. Food Engineering Reviews 18:9. https://doi.org/10.1007/s12393-026-09437-w

Gao A, Zeng Y, Gong Y (2025) Multimodal feature fusion for geographical origin identification of apples. Journal of Food Composition and Analysis 146:107884. https://doi.org/10.1016/j.jfca.2025.107884

Gidado MJ, Nagoor Gunny AA, Gopinath SCB, et al (2026) Artificial intelligence and machine learning in postharvest fruit quality assessment: Current challenges, recent advances, and future prospects. Journal of Stored Products Research 116:102959. https://doi.org/10.1016/j.jspr.2026.102959

Giménez-Gallego J, Martinez-del Rincon J, González-Teruel JD, et al (2024) On-tree fruit image segmentation comparing Mask R-CNN and Vision Transformer models. Application in a novel algorithm for pixel-based fruit size estimation. Computers and Electronics in Agriculture 222:109077. https://doi.org/10.1016/j.compag.2024.109077

Goyal A, Lakhwani K (2025) Integrating advanced deep learning techniques for enhanced detection and classification of citrus leaf and fruit diseases. Scientific Reports 15:12659. https://doi.org/10.1038/s41598-025-97159-0

Guo Q, Zhao M, Liu Y, et al (2025) Developing an attention-enhanced deep learning approach for impurity detection in *Camellia oleifera* seeds. Journal of Food Composition and Analysis 148:108148. https://doi.org/10.1016/j.jfca.2025.108148

Guo Z, Ni Y, Gao H, et al (2024) A dual-branch selective attention capsule network for classifying kiwifruit soft rot with hyperspectral images. Scientific Reports 14:10664. https://doi.org/10.1038/s41598-024-61425-4

Hayat A, Morgado-Dias F, Choudhury T, et al (2024) FruitVision: A deep learning based automatic fruit grading system. Open Agriculture 9:20220276. https://doi.org/10.1515/opag-2022-0276

He J, Belien T, Alhmedi A, et al (2026a) Deep transfer learning for codling moth damage detection in 'Conference' pear using X-ray radiography. Food Control 183:111923. https://doi.org/10.1016/j.foodcont.2025.111923

He W, Huang W, Li Y, et al (2025) Multimodal information fusion and precision harvesting system for fruit growth driven by flexible optoelectronic sensing and hierarchical attention networks. Computers and Electronics in Agriculture 238:110784. https://doi.org/10.1016/j.compag.2025.110784

He Z, Yu J, Zhou X, et al (2026b) Multi-sensor fusion with optimized machine learning for non-destructive freshness monitoring of stored Korla fragrant pears. Food Control 181:111692. https://doi.org/10.1016/j.foodcont.2025.111692

Hou G, Chen H, Jiang M, et al (2023) An Overview of the Application of Machine Vision in Recognition and Localization of Fruit and Vegetable Harvesting Robots. Agriculture 13. https://doi.org/10.3390/agriculture13091814

Howard A, Sandler M, Chen B, et al (2019) Searching for MobileNetV3. In: 2019 IEEE/CVF International Conference on Computer Vision (ICCV). IEEE, Seoul,

Korea (South), pp 1314–1324, https://doi.org/10.1109/ICCV.2019.00140

Hussain I, He Q, Chen Z (2018) Automatic Fruit Recognition Based on DCNN for Commercial Source Trace System. International Journal on Computational Science & Applications 8:01–14. https://doi.org/10.5121/ijcsa.2018.8301, URL https://zenodo.org/records/1310165, dataset repository; accessed 11 October 2026

Jahanbakhshi A, Momeny M, Mahmoudi M, et al (2020) Classification of sour lemons based on apparent defects using stochastic pooling mechanism in deep convolutional neural networks. Scientia Horticulturae 263:109133. https://doi.org/10.1016/j.scienta.2019.109133

Junyang W, Maskat MY, Mohd Ali M (2026) Non-destructive monitoring of postharvest quality changes in chili (Capsicum frutescens) during storage using image processing coupled with machine learning. Postharvest Biology and Technology 234:114140. https://doi.org/10.1016/j.postharvbio.2025.114140

Kaizen (2020) Fruit Classification. Kaggle, URL https://www.kaggle.com/datasets/sshikamaru/fruit-recognition, accessed 11 October 2026

Kalluri SR (2018) Fruits fresh and rotten for classification. Kaggle, URL https://www.kaggle.com/datasets/sriramr/fruits-fresh-and-rotten-for-classification, accessed 11 October 2026

Kang J, Gwak J (2022) Ensemble of multi-task deep convolutional neural networks using transfer learning for fruit freshness classification. Multimedia Tools and Applications 81:22355–22377. https://doi.org/10.1007/s11042-021-11282-4

Khan Z, Rawat S, Kumar S, et al (2025) Using computer vision for automated grading and sorting of fruits and vegetables. International Journal of Agriculture and Food Science 7:1056–1065. https://doi.org/10.33545/2664844X.2025.v7.i8j.674

Kothawade GS, Ranjan R, Sharrer K, et al (2026) Snapshot hyperspectral imaging for production improvements in Atlantic salmon farming: A proof-of-concept study. Smart Agricultural Technology 13:101724. https://doi.org/10.1016/j.atech.2025.101724

Kumar Pothula A, Zhang Z, Lu R (2023) Evaluation of a new apple in-field sorting system for fruit singulation, rotation and imaging. Computers and Electronics in Agriculture 208:107789. https://doi.org/10.1016/j.compag.2023.107789

LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521:436–444. https://doi.org/10.1038/nature14539

Lee A, Baek I, Kim J, et al (2025) Deep learning approaches for bruised mandarin orange classification by fluorescence hyperspectral imaging. Postharvest Biology and Technology 230:113724. https://doi.org/10.1016/j.postharvbio.2025.113724

Li J, Lu Y, Lu R (2023) Detection of early decay in navel oranges by structured-illumination reflectance imaging combined with image enhancement and segmentation. Postharvest Biology and Technology 196:112162. https://doi.org/10.1016/j.postharvbio.2022.112162

Li S, Zanchin A, Marinello F, et al (2026) Three-dimensional analysis of fruit and vegetables: Sensors and methods. Computers and Electronics in Agriculture 245:111497. https://doi.org/10.1016/j.compag.2026.111497

Liang C, Cappelletto J, Massot-Campos M, et al (2025) Self-supervised learning with multimodal remote sensed maps for seafloor visual class inference. The International Journal of Robotics Research 44:2340–2364. https://doi.org/10.1177/02783649251343640

Liu C, Wu X, Yang W, et al (2025a) Early detection of apple bruises using spectral-spatial enhanced 3D CNN and region-based hyperspectral analysis. Food Research International 222:117630. https://doi.org/10.1016/j.foodres.2025.117630

Liu Q, Bai L, Suhaiza Binti Sulaiman P, et al (2026) A Systematic Review On Deep Learning Techniques For Maturity Classification of Tomato Fruit On Plants. Smart Agricultural Technology p 102155. https://doi.org/10.1016/j.atech.2026.102155

Liu Y, Zhang J, Lou Y, et al (2024) Soft bionic gripper with tactile sensing and slip detection for damage-free grasping of fragile fruits and vegetables. Computers and Electronics in Agriculture 220:108904. https://doi.org/10.1016/j.compag.2024.108904

Liu Y, Chen J, Wang G, et al (2025b) Simulation and sensitivity analysis of light propagation in watermelon tissues for enhanced Vis/NIR spectroscopic quality assessment. Measurement 253:117342. https://doi.org/10.1016/j.measurement.2025.117342

Lu Y, Young S (2020) A survey of public datasets for computer vision tasks in precision agriculture. Computers and Electronics in Agriculture 178:105760. https://doi.org/10.1016/j.compag.2020.105760

Lv M, Xu YX, Miao YH, et al (2025) A Comprehensive Review of Deep Learning in Computer Vision for Monitoring Apple Tree Growth and Fruit Production. Sensors 25. https://doi.org/10.3390/s25082433

Ma C, Ying Y, Xie L (2024a) Visuo-tactile sensor development and its application for non-destructive measurement of peach firmness. Computers and Electronics in Agriculture 218:108709. https://doi.org/10.1016/j.compag.2024.108709

Ma J, Li M, Fan W, et al (2024b) State-of-the-Art Techniques for Fruit Maturity Detection. Agronomy 14. https://doi.org/10.3390/agronomy14122783

Ma J, Wan Y, Min W, et al (2026) Frontiers and advances of deep learning-based fruit and vegetable image analysis. Computers and Electronics in Agriculture 241:111256. https://doi.org/10.1016/j.compag.2025.111256

Mao Y, Zhang M, Law CL, et al (2026) The role of artificial intelligence in enhancing non-invasive quality monitoring in fresh food products supply chains: A comprehensive review. Food Research International 232:118927. https://doi.org/10.1016/j.foodres.2026.118927

Masoudi M, Golzarian MR, Lawson SS, et al (2024) Improving 3D reconstruction for accurate measurement of appearance characteristics in shiny fruits using post-harvest particle film: A case study on tomatoes. Computers and Electronics in Agriculture 224:109141. https://doi.org/10.1016/j.compag.2024.109141

Masuda K, Uchida R, Fujita N, et al (2023) Application of deep learning diagnosis for multiple traits sorting in peach fruit. Postharvest Biology and Technology 201:112348. https://doi.org/10.1016/j.postharvbio.2023.112348

Melesse TY (2026) Intelligent Postharvest Sorting of Bananas Using Thermal Imaging and Deep Neural Network Models. Food and Bioprocess Technology 19:80. https://doi.org/10.1007/s11947-025-04173-1

Mohammadi MA, Tajdar-oranj B, Esmaeili E (2026) Spectral-AI technologies for food safety: Advances, challenges, and future directions. Food Control 187:112144. https://doi.org/10.1016/j.foodcont.2026.112144

Mohsan MM, Hasanen BB, Hassan T, et al (2025) SwishFormer for robust firmness and ripeness recognition of fruits using visual tactile imagery. Postharvest Biology

and Technology 225:113487. https://doi.org/10.1016/j.postharvbio.2025.113487

Mukhiddinov M, Muminov A, Cho J (2022) Improved Classification Approach for Fruits and Vegetables Freshness Based on Deep Learning. Sensors 22:8192. https://doi.org/10.3390/s22218192, URL https://www.kaggle.com/datasets/muhriddinmuxiddinov/fruits-and-vegetables-dataset, dataset repository; accessed 11 October 2026

Mureşan H, Oltean M (2018) Fruit recognition from images using deep learning. Acta Universitatis Sapientiae, Informatica 10:26–42. https://doi.org/10.2478/ausi-2018-0002, URL https://github.com/Horea94/Fruit-Images-Dataset, dataset repository, version 2020.05.18.0; accessed 11 October 2026

O'Brien C, Falagán N, Kourmpetli S, et al (2024) Non-destructive methods for mango ripening prediction: Visible and near-infrared spectroscopy (visNIRS) and laser Doppler vibrometry (LDV). Postharvest Biology and Technology 212:112878. https://doi.org/10.1016/j.postharvbio.2024.112878

Pacheco R, González P, Chuquimarca L, et al (2023) Fruit defect detection using CNN models with real and virtual data. In: Proceedings of the 18th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, vol 4. SciTePress, pp 272–279, https://doi.org/10.5220/0011679800003417, URL https://github.com/luischuquim/Healthy-Defective-Fruits, dataset repository; accessed 11 October 2026

Plested J, Phiri M, Gedeon T (2026) Deep transfer learning for image classification: a survey. Artificial Intelligence Review 59:100. https://doi.org/10.1007/s10462-026-11491-z

Santelices IR, Cano S, Moreira F, et al (2025) Artificial Vision Systems for Fruit Inspection and Classification: Systematic Literature Review. Sensors 25. https://doi.org/10.3390/s25051524

Seth K (2020) Fruits and Vegetables Image Recognition Dataset. Kaggle, URL https://www.kaggle.com/datasets/kritikseth/fruit-and-vegetable-image-recognition, accessed 11 October 2026

Song JY, Qin ZS, Xue CW, et al (2024) Fruit grading system by reconstructed 3D hyperspectral full-surface images. Postharvest Biology and Technology 212:112898. https://doi.org/10.1016/j.postharvbio.2024.112898

Tan M, Le Q (2021) EfficientNetV2: Smaller Models and Faster Training. In: Proceedings of the 38th International Conference on Machine Learning. PMLR, pp 10096–10106

Tapia-Mendez E, Cruz-Albarran IA, Tovar-Arriaga S, et al (2023) Deep Learning-Based Method for Classification and Ripeness Assessment of Fruits and Vegetables. Applied Sciences 13:12504. https://doi.org/10.3390/app132212504

Tempelaere A, Van De Looverbosch T, Kelchtermans K, et al (2023) Synthetic data for X-ray CT of healthy and disordered pear fruit using deep learning. Postharvest Biology and Technology 200:112342. https://doi.org/10.1016/j.postharvbio.2023.112342

Tian S, Wang J, Xu H (2022) Firmness measurement of kiwifruit using a self-designed device based on acoustic vibration technology. Postharvest Biology and Technology 187:111851. https://doi.org/10.1016/j.postharvbio.2022.111851

Timprae W, Sagawa T, Baar S, et al (2025) Tomato Growth Monitoring and Phenological Analysis Using Deep Learning-Based Instance Segmentation and 3D Point

Cloud Reconstruction. Sustainability 17. https://doi.org/10.3390/su172210120

Van De Looverbosch T, Raeymaekers E, Verboven P, et al (2021) Non-destructive internal disorder detection of Conference pears by semantic segmentation of X-ray CT scans using deep learning. Expert Systems with Applications 176:114925. https://doi.org/10.1016/j.eswa.2021.114925

Wan S, Goudos S (2020) Faster R-CNN for multi-class fruit detection using a robotic vision system. Computer Networks 168:107036. https://doi.org/10.1016/j.comnet.2019.107036

Wang C, Li X, Zhang Z, et al (2025a) Nondestructive Quality Detection of Characteristic Fruits Based on Vis/NIR Spectroscopy: Principles, Systems, and Applications. Agriculture 15. https://doi.org/10.3390/agriculture15202167

Wang D, Li L, Jiang X, et al (2026) Toward robust in-field fruit quality evaluation: A critical review of emerging nondestructive technologies and devices. Food Research International 225:118058. https://doi.org/10.1016/j.foodres.2025.118058

Wang G, Wang Y, Liu D, et al (2025b) Hyperspectral-based assessment of internal quality indicators of peach and quantitative transfer learning model analysis. Postharvest Biology and Technology 230:113824. https://doi.org/10.1016/j.postharvbio.2025.113824

Wood RM, Schut DE, Trull AK, et al (2024) Detecting internal browning in apple tissue as determined by a single CT slice in intact fruit. Postharvest Biology and Technology 211:112802. https://doi.org/10.1016/j.postharvbio.2024.112802

Wood RM, Schut DE, Balk PA, et al (2025) Assessing the development of internal disorders in pome fruit with X-ray CT before, during and after controlled atmosphere storage and shelf life. Food Control 168:110970. https://doi.org/10.1016/j.

foodcont.2024.110970

Xiao Y, Zhou L, Zhao Y, et al (2025) Deep learning-based regression of food quality attributes using near-infrared spectroscopy and hyperspectral imaging: A review. Food Chemistry 493:145932. https://doi.org/10.1016/j.foodchem.2025.145932

Xu R, Li Y, Zhang Z, et al (2025) Automatic Fruit Size Evaluation System Based on LabVIEW. Horticulturae 11. https://doi.org/10.3390/horticulturae11111341

Xu X, Xu T, Li Z, et al (2024) SPMUNet: Semantic segmentation of citrus surface defects driven by superpixel feature. Computers and Electronics in Agriculture 224:109182. https://doi.org/10.1016/j.compag.2024.109182

Xue L, Cong H, Xu D, et al (2026) DBH-YOLO: A novel algorithm for surface defects detection of postharvest blue honeysuckle berry fruit. Postharvest Biology and Technology 233:114045. https://doi.org/10.1016/j.postharvbio.2025.114045

Xue S, Li Z, Wang D, et al (2025) YOLO-ALDS: an instance segmentation framework for tomato defect segmentation and grading based on active learning and improved YOLO11. Computers and Electronics in Agriculture 238:110820. https://doi.org/10.1016/j.compag.2025.110820

Yan K, Fang X, Yang W, et al (2025) Multiple light sources excited fluorescence image-based non-destructive method for citrus Huanglongbing disease detection. Computers and Electronics in Agriculture 237:110549. https://doi.org/10.1016/j.compag.2025.110549

Yang S, Wang W, Gao S, et al (2023) Strawberry ripeness detection based on YOLOv8 algorithm fused with LW-Swin Transformer. Computers and Electronics in Agriculture 215:108360. https://doi.org/10.1016/j.compag.2023.108360

Yu Y, Huang J, Wang L, et al (2025) A 1D-inception-ResNet based global detection model for thin-skinned multifruit spectral quantitative analysis. Food Control 167:110823. https://doi.org/10.1016/j.foodcont.2024.110823

Zeng S, Zhang Z, Cheng X, et al (2024) Prediction of soluble solids content using near-infrared spectra and optical properties of intact apple and pulp applying PLSR and CNN. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 304:123402. https://doi.org/10.1016/j.saa.2023.123402

Zhang B, Zhang F, An S, et al (2025a) SCORE-DETR: An efficient Transformer-based network for small and occluded citrus detection. Computers and Electronics in Agriculture 238:110843. https://doi.org/10.1016/j.compag.2025.110843

Zhang J, Bai X, Wu J, et al (2025b) Nondestructive detection method for soluble solids content and titratable acidity content in pepino melons based on Vis/NIR spectroscopy and dual-attention enhanced 1D-CNN. Journal of Food Composition and Analysis 148:108232. https://doi.org/10.1016/j.jfca.2025.108232

Zhao C, Ren Z, Li Y, et al (2025) Growth Stages Discrimination of Multi-Cultivar Navel Oranges Using the Fusion of Near-Infrared Hyperspectral Imaging and Machine Vision with Deep Learning. Agriculture 15. https://doi.org/10.3390/agriculture15141530

Zhao F, Zhang C, Geng B (2024) Deep Multimodal Data Fusion. ACM Comput Surv 56:216:1–216:36. https://doi.org/10.1145/3649447

Zhou X, Ampatzidis Y, Lee WS, et al (2022) Deep learning-based postharvest strawberry bruise detection under UV and incandescent light. Computers and Electronics in Agriculture 202:107389. https://doi.org/10.1016/j.compag.2022.107389

Zou Y, Wu J, Meng X, et al (2025) Digital twin integration for dynamic quality loss control in fruit supply chains. Journal of Food Engineering 397:112577. https://doi.org/10.1016/j.jfoodeng.2025.112577