agronomy

Review

A Review of Fruit Tree Canopy Branch Feature Extraction and 3D Reconstruction Algorithms

Yong Jiang [1,2,3], Jing Chen [1,2,3] and Shengyi Zhao [1,2,3,*]

1 School of Agricultural Engineering, Jiangsu University, Zhenjiang 212013, China; 1000006668@ujs.edu.cn (Y.J.); 2222416005@stmail.ujs.edu.cn (J.C.)
2 National Digital Agricultural Equipment (Artificial Intelligence and Agricultural Robotics) Innovation Sub-Centre, Jiangsu University, Zhenjiang 212013, China
3 Key Laboratory of Modern Agricultural Equipment and Technology, Ministry of Education, Jiangsu University, Zhenjiang 212013, China
* Correspondence: 1000006619@ujs.edu.cn

Abstract

Accurate perception and 3D reconstruction of fruit tree branch structures are fundamental to smart orchard development, with broad applications in intelligent harvesting, crop phenotyping, and precision management. However, the slender and highly branched morphology, multi-scale distribution, weak surface texture, and severe occlusion inherent to fruit tree branches pose substantial challenges to high-fidelity modeling. This paper systematically reviews advances in branch feature extraction and 3D reconstruction for fruit tree canopies. A structured literature search was conducted using the Web of Science, Scopus, and Google Scholar databases, with search terms including "fruit tree branch", "point cloud reconstruction", "3D canopy modeling", "branch feature extraction", and "agricultural robotics". Studies published between 2000 and 2025 were considered, with inclusion criteria requiring relevance to branch structure perception, reconstruction accuracy, or orchard application; non-peer-reviewed sources and studies lacking quantitative evaluation were excluded. We trace the evolution of feature extraction from classical 2D image processing and geometric fitting, through point cloud segmentation and skeleton extraction, to modern deep learning approaches and multimodal perception techniques. For 3D reconstruction, we compare active and passive sensing strategies alongside both explicit and implicit scene representation methods, discussing their respective strengths and applicable scenarios. A five-dimensional evaluation framework is also proposed, encompassing geometric accuracy, structural consistency, feature stability, computational efficiency, and generalization capability. Finally, we identify key bottlenecks in fine-grained structure recovery, occlusion handling, and cross-scene generalization, and highlight future directions in structural prior integration, multimodal collaborative modeling, and lightweight neural representations—offering a structured reference for advancing 3D perception research in smart orchards.

Keywords: fruit tree canopy; branch feature extraction; three-dimensional reconstruction; point cloud processing; deep learning

Check for updates

Academic Editor: Valerio Cristofori

Received: 29 May 2026
Revised: 26 June 2026
Accepted: 29 June 2026
Published: 1 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

With the rapid advancement of smart agriculture, agricultural production is undergoing a broad transformation toward digitalization, intelligence, and automation. By integrating sensing technologies, computer vision, and artificial intelligence, this transformation centers on information perception, intelligent decision-making, and automated execution to enable precision management throughout the production process. Against this backdrop, the growing deployment of agricultural robots, UAV-based inspection systems, and automated harvesting equipment has placed increasingly demanding requirements on the ability to acquire structural information from crops [1]. In orchard environments particularly, the complexity of tree architecture and multi-scale spatial distribution make it difficult for conventional 2D vision to fully capture plant morphology. Three-dimensional structural data, by contrast, offers a more accurate description of spatial geometry and topological relationships, providing essential support for intelligent equipment deployment, growth monitoring, and orchard management [2,3].

For a long time, fruit tree structural information was typically obtained through manual measurement or empirical judgment, relying on tools such as measuring tapes, clinometers, and total stations. These approaches contributed to the quantitative study of tree morphological parameters and laid an early foundation for structural analysis. However, as smart agriculture has advanced, their limitations have become increasingly apparent [4]. Manual measurement requires branch-by-branch recording, making it labor-intensive, time-consuming, and poorly suited to the rapid data collection demands of large-scale orchards. The process is also susceptible to subjective variability, compromising data consistency and reliability [5]. Furthermore, the intricate architecture of fruit tree canopies—where overlapping branches and foliage frequently occlude interior structures—makes it difficult to capture complete structural information using traditional methods, particularly under high-density planting conditions [6]. As orchards grow in scale and the push toward intelligent management intensifies, the demand for rapid, high-precision acquisition of 3D fruit tree structural data has become increasingly urgent. Automated extraction and 3D reconstruction of canopy branch structures through advanced sensing technologies and computer vision methods has thus emerged as a prominent research direction [7,8].

The fruit tree canopy comprises a branch system and a foliar system, with the branch structure serving as the skeletal framework that determines the spatial form of the canopy and the distribution of fruit. It also exerts significant influence on light interception, ventilation, and nutrient transport. Accurate description and modeling of branch structures therefore constitutes a foundational requirement for fruit tree structural analysis and precision management. In the context of smart agriculture, 3D structural information plays a critical supporting role across multiple application scenarios, most notably in intelligent equipment operation, crop growth monitoring, and automated orchard management [9–11]. In equipment applications, without reliable perception of the spatial topology of branches, robotic arms are prone to collisions during harvesting, reducing operational efficiency and potentially damaging fruit [12]. High-fidelity 3D branch models provide robots with dependable spatial environment information to support path planning, obstacle avoidance, and grasp pose adjustment—forming an essential basis for intelligent harvesting and autonomous robotic operation [13]. In crop growth monitoring, phenotypic parameters such as branch length, branching angle, branch density, and spatial distribution are widely regarded as key indicators of tree vigor, carrying significant implications for variety breeding, cultivation optimization, and growth assessment [14]. By constructing 3D models of canopy branch structures, these parameters can be automatically extracted and quantitatively analyzed, substantially improving the efficiency of phenotyping research. In automated orchard management, 3D structural data provides a foundation for precision production decision-making. Canopy branch parameters can be used to build yield prediction models and to inform targeted fertilization and irrigation strategies. Structural anomalies can also serve as early indicators for pest and disease monitoring and tree health assessment, elevating the intelligence and precision of orchard management.

Driven by these application demands, a complete technical pipeline for 3D perception of fruit tree canopy branches has progressively taken shape—spanning data acquisition, feature extraction, and 3D modeling through to application-level decision-making. To systematically organize the key processes and hierarchical relationships within this field, this paper presents a structured framework for fruit tree canopy branch feature extraction, 3D reconstruction, and representative applications, as illustrated in Figure 1. The framework unfolds across a data layer, a feature extraction layer, a 3D reconstruction layer, and an application layer, clearly capturing the interconnections among multi-source data acquisition, 3D structural modeling, and typical application scenarios—providing a structured basis for the review and analysis of related methods in the sections that follow.

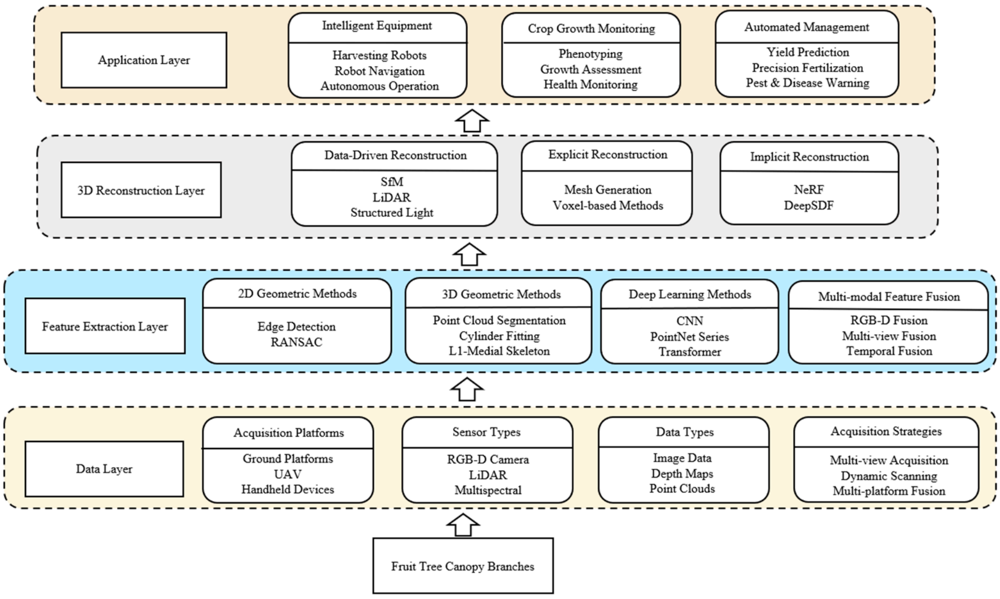


Figure 1. Framework for Fruit Tree Canopy Branch Extraction and 3D Reconstruction.

Despite growing interest in fruit tree branch perception, several critical gaps persist in the existing literature. First, most prior reviews focus narrowly on either feature extraction or 3D reconstruction, without systematically addressing the complete pipeline from data acquisition through structural modeling to application-level evaluation. Second, existing works seldom compare methods under standardized orchard conditions across multiple fruit species, limiting the practical transferability of reported findings. Third, the lack of a unified multi-dimensional evaluation framework makes cross-method performance comparison inconsistent and incomplete. These gaps motivate the present review, which aims to provide a comprehensive, critically comparative, and practically oriented synthesis of the field. The literature search was conducted using Web of Science, Scopus, and Google Scholar, with keywords including "fruit tree branch reconstruction", "canopy point cloud", "branch feature extraction", "3D orchard modeling", and "agricultural robot perception". Studies published between 2000 and 2025 were considered; inclusion required relevance to branch structural analysis or orchard-context 3D reconstruction, while conference abstracts lacking full methodology and non-peer-reviewed sources were excluded.

Based on the foregoing background, this paper systematically reviews research on fruit tree branch feature extraction and 3D reconstruction, tracing methodological developments, analyzing mainstream reconstruction algorithms and their optimization strategies, and consolidating commonly used evaluation metrics. We further examine key technical bottlenecks in smart orchard applications and outline future research directions. Compared with existing reviews, this work places greater emphasis on branch structure reconstruction under realistic orchard conditions, highlighting the interplay among complex occlusion, multi-scale structural representation, and agricultural robot applications—offering a systematic technical reference for 3D perception research in smart orchards.

2. Fruit Tree Branch Feature Extraction Techniques

In the 3D perception of fruit trees, branch feature extraction serves as the critical link between data acquisition and structural modeling, with its outcomes directly shaping the accuracy and stability of subsequent 3D reconstruction. However, fruit tree branches are characterized by slender, highly branched morphology, wide scale variation, and severe occlusion, and often exhibit weak texture, uneven coloration, and local data absence—making them prone to confusion with foliage and background clutter in complex scenes, which significantly increases the difficulty of feature extraction.

From a technical development perspective, fruit tree branch feature extraction has evolved through a progression of 2D visual methods, 3D geometric approaches, deep learning methods, and multimodal collaborative perception techniques. The taxonomic framework of fruit tree branch feature extraction technologies is illustrated in Figure 2. Organized along the axis of technical evolution, the framework categorizes existing methods into four broad classes, 2D data-based extraction methods, 3D geometry-driven methods, data-driven deep learning methods, and multimodal collaborative perception techniques. The following sections discuss the theoretical foundations, representative techniques, and applicability of each category under complex environmental conditions, and further explore strategies for robustness optimization and computational efficiency improvement in feature extraction.

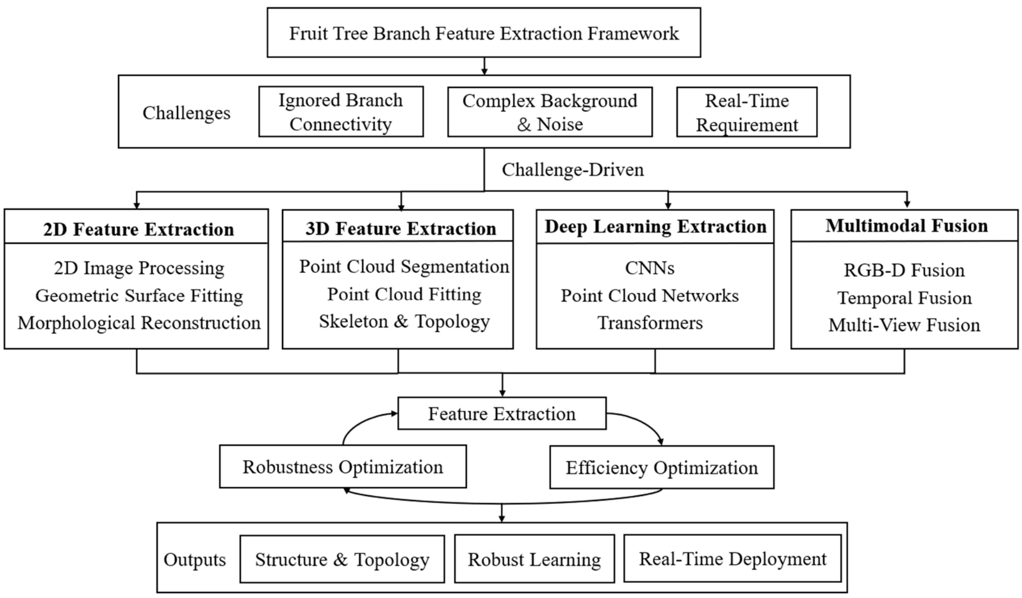


Figure 2. Taxonomic framework of fruit tree branch feature extraction techniques.

2.1. Branch Feature Extraction Methods Based on 2D Data

Branch feature extraction methods based on 2D data primarily rely on low-level visual features and geometric models, achieving branch detection and structural modeling through explicit rules [15]. These methods offer relatively high computational efficiency and strong interpretability, and have served as an important foundation for early fruit tree modeling.

2.1.1. 2D Image Processing-Based Methods

Branch detection from 2D images represents one of the earliest research directions in this field. Early studies drew heavily from general computer vision, particularly edge detection and image segmentation. The Canny edge detection algorithm, as a representative example, provided a foundational tool for extracting elongated structures and was subsequently widely applied to trunk and branch contour detection [16]. Its core principle relies on gradient information to delineate the boundaries of slender structures, enabling effective detection of trunks and coarse branches [17]. However, methods at this stage were highly sensitive to illumination changes and complex backgrounds. In real orchard environments, leaf texture, shadows, and background clutter can generate strong gradient responses, leading to false detections.

Researchers subsequently introduced grayscale thresholding and region growing approaches, with Qin Zhang et al. among the first to demonstrate the feasibility of vision-based methods for fruit tree structural analysis [18]. LBP texture descriptors were further incorporated to enhance the discrimination between branches and foliage, marking a shift from single edge features toward multi-feature fusion [19]. Nevertheless, the limited adaptability of 2D image methods to occlusion, illumination variation, and fine branch structures gradually drove research toward 3D approaches that incorporate depth information. Nevertheless, the limited adaptability of 2D image methods to occlusion, illumination variation, and fine branch structures has gradually driven research toward 3D approaches that incorporate depth information. From a critical perspective, these 2D image processing methods were evaluated predominantly on dormant-season apple and pear datasets under controlled outdoor illumination conditions, where branch–sky contrast is relatively high. Under such conditions, reported branch detection performance is generally acceptable, with F1-scores typically ranging from 0.70 to 0.85 for primary branches, as reported in the literature under specific experimental settings. However, performance degrades substantially under more production-realistic conditions: changes to overcast illumination can reduce consistency by 15–25%, while partial or full leaf occlusion may lead to F1-score drops exceeding 20%, primarily due to the loss of clear branch–background contrast. In addition, 2D-based approaches inherently lack depth information and are unable to recover key structural attributes such as branch diameter, spatial orientation, and 3D bifurcation angles, which are essential for phenotypic analysis and robotic arm trajectory planning.

2.1.2. Geometry-Based Surface Fitting Methods

As the demand for parametric description of fruit tree structures grew, researchers began modeling branches as parameterizable geometric primitives. Early work was influenced by geometric modeling concepts from computer vision, approximating trunks as cylindrical structures. Building on this, the RANSAC algorithm provided a robust solution for model estimation in noisy environments, achieving reliable structural parameter recovery through random sampling and consistency constraints, and demonstrating strong noise resistance in branch axis extraction and point cloud fitting [20]. In orchard scenes, branch structures can generally be approximated as locally linear or low-order curves, enabling RANSAC to effectively eliminate outliers introduced by leaves, occlusions, and background mismatches. It can also extract boundary points for canopy contour delineation [21], thereby improving the robustness of structural estimation.

As research shifted from global model estimation toward finer structural representation, local geometric fitting methods were introduced to model curved branches and complex bifurcation structures through neighborhood point set analysis. By combining statistical optimization with local geometric constraints, these methods enable stable representation of complex branch structures through continuous local fitting, making them particularly well-suited to curved branches and irregular growth patterns [22]. RANSAC provides robust initialization of global structural models, while local surface estimation achieves fine-grained morphological recovery through neighborhood geo-metric analysis. Together, they form a modeling paradigm that unites global robust estimation with local geometric refinement. This combination not only improves the stability of structural extraction in noisy environments, but also lays an important theoretical foundation for subsequent 3D branch point cloud reconstruction, skeleton ex-traction, and topological analysis—gradually establishing itself as a key technical bridge in the transition from 2D visual methods to 3D geometric perception. In terms of applied performance, geometry-based fitting methods have been primarily evaluated on peach and walnut orchards for trunk and primary scaffold branch modeling. When applied to dormant-season single-view or stereo images under sparse canopy conditions, RANSAC-based cylinder fitting achieves errors generally below 5 mm for branch diameters above 20 mm. However, performance degrades markedly for secondary branches thinner than 10 mm, where the assumption of ideal cylindrical primitives breaks down due to irregular surface texture and taper. Curved branches and complex Y-shaped bifurcation regions require multi-segment fitting, increasing computational overhead and parameter sensitivity. Moreover, without depth sensors, 2D-based geo-metric fitting cannot resolve branch overlap or spatial adjacency, making it unsuitable for high-density orchard systems such as trellis-trained pear or apple varieties.

2.1.3. Mathematical Morphology-Based Reconstruction Methods

The introduction of skeletonization methods shifted the research focus from geometric shape description toward topological structure analysis. Methods grounded in mathematical morphology use operations such as erosion and dilation for noise suppression and structural enhancement, enabling the extraction of elongated structures while preserving connectivity, thereby providing a stable basis for branch skeleton extraction. Building on this, skeletonization and thinning methods were further developed to represent branch structures as centerlines, capturing their topological connectivity [23].

To improve the stability of skeleton extraction, researchers introduced distance transforms and graph-based structural modeling, refining skeleton positioning through medial axis constraints and abstracting branch structures into graph models to represent branching relationships and hierarchical organization. More recent work has further enhanced skeleton stability through optimization strategies and structural constraints, though these methods remain sensitive to noise and data incompleteness, and their robustness in natural orchard environments still requires improvement.

Mathematical morphology and skeletonization methods achieve topological correctness above 85–90% for primary branches in open-canopy apple and pear orchards under dormant conditions, under specific conditions reported in the literature. However, in high-density canopies where branch density exceeds 150 segments per square meter, false connection rates can exceed 20%. Wind-induced point cloud noise further causes skeleton fractures in fine branches thinner than 8 mm. These results indicate that skeleton-based methods are more suitable for coarse structural analysis in dormant-season pruning tasks, and should be complemented by learning-based approaches for fine branch recovery in leafy or high-density conditions. In leafy-season scenarios, connectivity errors may still exceed 15% even under controlled settings. Moreover, these methods inherently cannot recover branch diameter or 3D orientation, limiting their applicability in robotic trajectory planning and phenotypic analysis.

Table 1 summarizes representative 2D data-based branch feature extraction methods for fruit tree modeling. These methods mainly rely on visual features, geometric models, and topological morphology to achieve branch detection and structural reconstruction. In terms of practical applicability, 2D image processing methods have been applied primarily to dormant apple and pear trees under controlled or low-wind outdoor conditions, where branch-background contrast is relatively high; reported branch detection F1-scores in such settings typically range from 0.70 to 0.85, but degrade substantially under dense canopy occlusion or variable illumination. Geometry-based fitting methods using RANSAC have shown strong robustness for trunk and primary branch reconstruction in peach and walnut trees, with cylinder fitting errors generally below 5 mm under sparse canopy conditions, though performance degrades for secondary branches thinner than 10 mm. Morphology-based skeletonization methods provide clear topological representations for dormant-season apple datasets, but exhibit connectivity errors exceeding 15% in leafy-season scenarios due to sensitivity to noise and data incompleteness. Although these approaches established an important foundation for early fruit tree branch modeling, their limited robustness in complex orchard environments—particularly for fine branches, leafy canopies, and high-density plantings—has gradually driven the development of 3D perception and reconstruction methods.

Table 1. Comparison of 2D data-based branch feature extraction methods for fruit tree branches.

Method Category	Techniques	Target Species	Application Condition	Reported Performance	Limitations	References
2D Image Processing	Edge detection, thresholding, region growing	Apple, pear (dormant season)	Controlled illumination, low wind, open canopy	high F1 in dormant season	Illumination-sensitive; fails under occlusion; no depth information	[17]
Geometry-Based Fitting	RANSAC, cylinder/cone fitting, B-spline curves	Peach, walnut (trunk and scaffold branches)	Sparse canopy, dormant season, single-view 2D images	low robustness under occlusion	Assumes idealized primitives; poor for curved or bifurcated secondary branches; noise-sensitive	[20]
Mathematical Morphology	Erosion/dilation, thinning, distance transform, graph skeletonization	Apple (dormant season datasets)	Controlled background, leafless canopy, uniform lighting	moderate geometric accuracy for primary branches	Fragile topology under noise; no branch diameter info; breaks under occlusion or leaf coverage	[23]

2.2. Feature Extraction Methods Based on 3D Data

With the advancement of 3D perception technologies, point cloud-based feature extraction of fruit tree branches has gradually become a research hotspot. Compared with traditional 2D approaches, 3D data can directly capture the spatial structural information of branches, offering greater stability and accuracy under complex occlusion conditions.

2.2.1. Point Cloud Segmentation Methods

Point cloud segmentation forms the foundational step in 3D branch extraction. By leveraging spatial adjacency relationships and geometric features, 3D data enables structural parsing to separate branches from background or other structural components [24]. Early studies employed region growing and Euclidean clustering for structural partitioning. Region growing operates under the assumption of local geometric consistency, progressively expanding regions based on the similarity of normal vectors and curvature within neighborhoods, which helps preserve the continuity of branch structures. Euclidean clustering, by contrast, achieves point set grouping through spatial distance constraints, making it suitable for separating spatially distinct branch structures with relatively high computational efficiency.

Building on this, density-based segmentation methods—most notably DBSCAN—were further developed to identify structural regions through local density distributions, offering adaptability to irregular shapes and a degree of noise suppression [25]. However, these methods are sensitive to parameter selection and tend to produce over- or under-segmentation when point cloud density is unevenly distributed, with computational costs rising significantly as point cloud scale increases. Graph-theoretic methods have since been introduced to address these limitations, representing point clouds as weighted graphs and applying global optimization strategies to achieve more consistent and stable segmentation results, though their relatively high computational complexity constrains their applicability in large-scale point cloud scenarios [26]. This progression marks a transition from 2D morphological operations to 3D geometric and graph-based modeling, and has become an essential foundation for modern structural analysis and skeleton reconstruction of fruit tree point clouds. In terms of practical performance under orchard conditions, region growing and DBSCAN-based segmentation methods have been mainly evaluated on TLS and MLS point clouds of apple and peach orchards reported in the literature. Under moderate canopy density and dormant season conditions, primary branch separation performance is generally satisfactory, with recall typically above 80 percent under controlled settings. However, performance degrades significantly for fine branches thinner than approximately 15 mm due to insufficient point density, and in leafy season scans where branch and leaf points interleave, leading to over segmentation and loss of structural continuity. Graph-based segmentation methods show improved consistency in complex scenes, but their computational cost increases rapidly with point cloud size and becomes inefficient for large-scale TLS datasets exceeding one million points. A key limitation is the lack of cross species validation, limiting generalization to structurally different canopies such as cherry, citrus, and high-density trellis trained apple orchards.

2.2.2. Geometry-Based Structure Fitting from Point Clouds

Point cloud-based geometric fitting methods inherit and extend the geometric modeling ideas developed for 2D data. Early research applied RANSAC cylinder fitting to trunk identification, achieving robust estimation of model parameters through random sampling and consistency constraints, and enabling stable recovery of branch axis positions and scale parameters under noisy and occluded conditions—providing an important basis for early tree structure reconstruction [27].

As research progressed, single-scale models proved insufficient for accurately describing complex branch structures, leading to the introduction of multi-scale geometric fitting approaches. By constructing local point sets at different spatial scales, these methods separately estimate the trunk and fine branches, balancing overall morphological representation with local detail. To address structural discontinuities arising from local fitting, researchers further incorporated skeleton constraints and structural priors. Skeletons, serving as centerline representations of branches, provide global topological information to constrain model orientation and connectivity. Structural priors integrate biological growth patterns—such as branch continuity and radius tapering—into the model optimization process, improving the plausibility and stability of reconstruction results. This trajectory has driven tree modeling from local geometric approximation toward a structured reconstruction framework that accounts for both topological consistency and biological regularity, laying important theoretical groundwork for subsequent learning-based methods and data-driven models. From a critical standpoint, 3D point cloud-based geometric fitting methods show clear advantages over 2D image-based approaches in both accuracy and applicability. Cylinder fitting RMSE is generally below 3 mm for trunk and primary scaffold branches in dormant season TLS scans of walnut and cherry orchards, while 2D image-based methods typically exhibit errors in the range of 5 to 10 mm under comparable conditions reported in the literature. However, several limitations remain unresolved. Branch coverage decreases significantly in leafy season scans due to occlusion, reducing effective point density and leading to fitting instability or diameter overestimation. In addition, models calibrated on specific species with thick primary branches do not generalize well to species with fine branch architectures such as peach and cherry, where branch diameters are often below 5 mm. Finally, multi-scale fitting methods require higher computational cost, typically taking several minutes per full canopy TLS scan, which limits their applicability in near real time robotic systems.

2.2.3. Skeleton Extraction and Topological Reconstruction

Skeleton extraction aims to recover the central structure of a tree from point clouds, and is a key step in branch topological analysis. Representative methods such as L1-Medial Skeleton progressively approximate branch centers through optimization-based contraction, enhancing robustness to noise and uneven sampling while preserving local geometric features, thereby enabling stable skeleton extraction [28]. Theoretically, skeleton extraction is grounded in medial axis theory, recovering central structures through distance field analysis to effectively characterize branch connectivity and branching morphology, abstracting complex tree bodies into centerline representations.

Building on this foundation, subsequent research has combined geodesic distances and adjacency graphs to achieve branch-level stratification and structural separation. Local cross-sectional fitting is further applied to obtain skeleton positions and scale information, with topological connections established to enable the conversion from raw point clouds to structured branch models. In recent years, data-driven approaches have also been introduced into skeleton extraction to improve the recovery of branch topology under conditions of occlusion, noise, and data incompleteness.

Overall, skeleton extraction strengthens branch structural representation and complements point cloud segmentation and geometric fitting methods, forming a cornerstone of 3D branch structural analysis and an indispensable technical component in the complete pipeline from point cloud acquisition to structured tree modeling. However, a critical examination reveals important performance limitations. Skeleton extraction methods such as L1 Medial Skeleton perform reliably for main branch orders in open-canopy apple and pear orchards under dormant conditions, where point clouds are relatively complete and noise levels are low. In these settings, topological correctness for primary branches typically exceeds 85 to 90 percent. However, performance degrades significantly in high-density canopies, where spatial proximity leads to false connections above 20 percent due to merged skeleton paths during medial axis estimation. Wind-induced noise further reduces stability, causing fractures in fine branches thinner than 8 mm. Distance-based stratification methods can partially reduce false connections, but depend on accurate branch ordering, which is often unreliable under incomplete point coverage. Overall, skeleton-based methods are suitable for coarse structural analysis in dormant season pruning tasks, while fine branch recovery in leafy or high-density conditions requires complementary learning-based approaches.

Compared with the 2D data-based methods summarized in Table 1, the approaches listed in Table 2 further incorporate depth and spatial geometric information, enabling more accurate representation of branch morphology and topology in complex orchard environments. These methods gradually shift branch feature extraction from low-level visual analysis toward 3D structural perception. In terms of practical performance under orchard conditions, point cloud segmentation methods such as DBSCAN and region growing have demonstrated branch separation recall rates above 80% for primary branches in apple and peach orchards when canopy density is moderate, but accuracy drops notably for branches thinner than 15 mm or under heavy leaf occlusion. Geometry-based fitting methods achieve mean cylinder fitting RMSE below 3 mm for trunk and scaffold branches in dormant-season LiDAR scans of walnut and cherry trees, while their performance is constrained in leafy-season scans by incomplete point coverage. Skeleton extraction methods enable reliable topological reconstruction for main branch orders in open-canopy systems, but are prone to false connections or skeleton fractures in high-density apple and pear canopies, where branch density exceeds 150 segments per square meter. Together, these approaches establish the fundamental technical framework for 3D fruit tree structural reconstruction and provide important support for subsequent learning-based modeling methods.

Table 2. Comparison of 3D data-based branch feature extraction methods for fruit tree branches.

Method Category	Techniques	Target Species	Application Condition	Reported Performance	Limitations	References
Point Cloud Segmentation	Region growing, Euclidean clustering, DBSCAN	Apple, peach (LiDAR point clouds)	Moderate canopy density; dormant or early-leaf stage	high recall for primary branches in dormant season	High computational cost; parameter-sensitive; fails for fine or overlapping branches	[26]
Geometric Structure Fitting (3D)	RANSAC, multi-scale cylinder fitting, least-squares fitting	Walnut, cherry (trunk and scaffold branches)	Dormant-season TLS scans; sparse to moderate canopy	high geometric accuracy for trunk and scaffold branches under sparse canopy	Limited to primary/secondary branches; fails for irregular shapes; no semantic output	[27]
Skeleton Extraction and Topological Reconstruction	L1-Medial Skeleton, geodesic distances, adjacency graphs	Apple, pear (LiDAR, open-canopy orchards)	Open-canopy systems; dormant season; low-wind conditions	Reliable topology for main branches; false connections/fractures in high-density canopies	Sensitive to noise and missing data; lacks branch attributes (diameter); topology errors under occlusion	[28]

2.3. Deep Learning-Based Feature Extraction Methods

With the advancement of data-driven approaches, deep learning has gradually become the dominant technical paradigm for fruit tree branch feature extraction [29–32]. Unlike traditional methods that rely on hand-crafted features, deep models can automatically learn multi-level representations through end-to-end training, demonstrating greater robustness under complex backgrounds and occlusion conditions [33,34].

2.3.1. Convolutional Neural Networks

As one of the most representative deep learning models in computer vision, convolutional neural networks (CNNs) achieve hierarchical feature modeling through local receptive fields and weight sharing, progressively learning from low-level edges and textures to high-level semantic information. CNNs thus exhibit notable advantages in image analysis tasks—effectively capturing local spatial patterns, reducing training complexity through parameter sharing, and maintaining strong robustness against complex backgrounds [35,36]. In orchard scenes, CNNs are commonly applied to branch detection and segmentation, extracting discriminative features in cluttered environments and improving the stability of target recognition [37].

As network architectures evolved, models such as AlexNet, VGGNet, and ResNet progressively enhanced feature representation capacity, enabling broad application of CNNs in semantic segmentation and object detection [38,39]. However, due to their reliance on local receptive fields, CNNs have inherent limitations in modeling long-range dependencies and complex topological structures, which constrains their performance in fine branch structure recovery and spatial relationship modeling. These limitations have driven research further toward 3D networks and global modeling approaches to improve structural representation and reconstruction in complex natural scenes [40]. CNN-based methods including U-Net and Mask R-CNN achieve mean IoU of 0.72–0.85 on dormant-season apple and pear under controlled illumination, but cross-species F1 drops 15–25% due to bark texture and canopy differences. They also lack long-range spatial modeling, limiting fine branch recovery under severe occlusion.

2.3.2. Point Cloud Deep Learning Networks

With the rapid proliferation of RGB-D cameras and LiDAR sensing technologies, direct learning on 3D point clouds has emerged as an important research direction for fruit tree structural perception. Unlike 2D image-based methods, point cloud data directly encodes the true spatial geometry of objects, making it better suited for modeling slender 3D structures such as branches. To address the unordered and irregular nature of point cloud data, researchers proposed PointNet as a representative approach for direct modeling of 3D point sets [41]. Through point-level feature learning and global feature aggregation, such methods can effectively extract spatial structural information from branches.

However, PointNet has limitations in local structural representation. PointNet++ addressed this by introducing hierarchical structures and local neighborhood modeling, significantly enhancing the ability to represent complex geometric forms and advancing point cloud learning from global feature modeling toward joint local-global modeling [42]. Further developments include dynamic graph convolutional networks (DGCNN), which improve local topological modeling through dynamic construction of adjacency relationships, and sparse convolution methods that improve computational efficiency for large-scale point cloud processing [43]. In orchard scenes, these methods can directly leverage 3D geometric information, demonstrating high stability in branch segmentation and structural extraction tasks [44]. Nevertheless, the recognition of fine branches remains challenging due to point cloud sparsity and occlusion. Point cloud networks such as PointNet and PointNet++ achieve over 85% segmentation accuracy for primary branches but recall below 55% for branches thinner than 10 mm under leaf occlusion. They require dense point clouds and degrade in leafy scans; cross-species generalization remains unvalidated [42].

2.3.3. Transformer Architectures

The introduction of Transformer architectures has provided a new technical paradigm for fruit tree structural analysis. Originally developed to achieve global feature modeling through self-attention mechanisms, Transformers offer a compelling framework for complex structural analysis. Compared with CNNs that rely on local receptive fields, Transformers can establish feature associations across the full spatial extent, making them better suited to describing branch bifurcations and topological relationships [45].

The Vision Transformer subsequently extended this architecture to visual tasks for the first time [46,47]. While early vision Transformers were primarily applied to 2D image classification and segmentation, researchers quickly recognized that their global modeling capability was well-suited to handling plant canopy structures with complex topological relationships—enabling the modeling of inter-branch associations across the entire scene and more accurate characterization of bifurcation structures. In the 3D vision domain, Transformers have progressively expanded to point cloud tasks. Models such as Point Transformer learn spatial neighborhood weights through self-attention mechanisms, achieving joint modeling of global semantic information and local geometric features—marking a transition in 3D learning from convolution-dominated to attention-driven approaches. However, these methods typically involve high computational complexity and strong data scale dependency, requiring a practical trade-off between accuracy and efficiency in real-world applications. Transformer methods like Point Transformer offer superior global topology recovery but require 5 to 15 s per scene on high-end GPUs, precluding real-time use. They depend on large annotated datasets, which are scarce, and their performance in dense canopies is not systematically evaluated [48,49].

With the continued development of deep learning, data-driven feature extraction methods have been widely adopted for fruit tree branch structural analysis. Representative approaches including U-Net, PointNet, and Vision Transformer advance feature representation from different perspectives—2D semantic modeling, 3D structural learning, and global relational modeling respectively—yet each involves certain trade-offs in geometric information utilization, data dependency, and computational complexity, as summarized in Table 3. Different deep learning methods reflect distinct priorities between representational capacity and computational efficiency, and how to effectively integrate multi-source information while improving model robustness for complex branch structures remains an important direction for future research.

Table 3. Comparison of Deep Learning Methods for Fruit Tree Branch Feature Extraction.

Method Category	Techniques	Target Species	Application Condition	Reported Performance	Limitations	References
CNN Methods	U-Net, Mask R-CNN	Apple, pear	Dormant season, controlled outdoor illumination	high IoU for trunk and scaffold, reduced under cross-species transfer	Lacks 3D geometric information	[35]
Point Cloud Learning	PointNet, PointNet++	Apple, peach	Open-canopy LiDAR, dormant or early-leaf	high accuracy for primary branches, low recall under occlusion for fine branches	High data requirements	[42]
Transformer Methods	Point Transformer	Apple	Dormant canopy, multi-view datasets	good topology recovery but data-dependent	High computational overhead	[46]

2.4. Multimodal Feature Fusion Techniques

As research has advanced, single-sensor approaches have proven insufficient for reliably characterizing complex orchard environments, and multimodal fusion has gradually emerged as an important direction in fruit tree structural perception. RGB images provide texture and color information, while depth data captures spatial geometric structure—the two modalities exhibit strong complementarity. Early methods largely relied on data-level

https://doi.org/10.3390/agronomy16131274

concatenation for fusion, but with limited robustness. This gave way to feature-level fusion frameworks that jointly model visual and depth information, significantly improving branch detection stability [50–52]. Representative work by Gupta et al. on RGB-D representation learning demonstrated the effectiveness of depth information in compensating for gaps in visual semantics [50].

In terms of fusion strategies, attention mechanisms have been introduced to enable adaptive weighting of multimodal information, improving perceptual performance under complex illumination and occlusion conditions. Multi-view fusion has also emerged as an important technical approach for UAV-based orchard modeling, effectively mitigating branch–foliage occlusion through geometric consistency constraints [53,54]. Temporal fusion, meanwhile, exploits dynamic information from continuous observation sequences, combining temporal models with SLAM techniques to enable ongoing structural updates and robust reconstruction, further enhancing modeling stability in complex environments. In practice, each fusion strategy presents distinct trade-offs. RGB-D fusion methods based on feature-level integration have been mainly validated on greenhouse or controlled orchard datasets, achieving relatively high detection performance under moderate illumination, with F1-scores typically reported around 0.80–0.88. However, degrade in field conditions due to depth noise and texture instability. Multi-view fusion can reduce occlusion effects in sparse canopies, but its benefit is limited in high-density trellis systems and is constrained by high acquisition overhead, making real-time deployment difficult. Temporal fusion improves registration stability under low-wind conditions, but is sensitive to wind-induced motion, which reduces accuracy in dynamic environments. A key limitation is the lack of cross-species and cross-season validation, leaving generalization ability insufficiently demonstrated.

Building upon the methods summarized in Tables 1–3, the approaches in Table 4 further emphasize multimodal collaborative perception for fruit tree structural analysis. By integrating RGB images, depth data, multi-view observations, and temporal information, these methods improve structural representation and reconstruction robustness in complex orchard environments. Among them, RGB-D fusion enhances semantic and geometric complementarity, multi-view fusion improves structural completeness, and temporal fusion strengthens reconstruction continuity and stability.

Table 4. Comparison of multimodal feature fusion methods.

Method Category	Techniques	Target Species	Application Condition	Reported Performance	Limitations	References
RGB-D Fusion	Semantic and geometric complementarity	Apple, peach	Greenhouse or controlled outdoor, moderate illumination	degraded performance under strong sunlight and wet bark	Dependent on sensor synchronization	[50]
Multi-view Fusion	Geometry-consistent structural recovery	Apple, pear	Sparse canopy, dormant or early-leaf	partial occlusion reduction, limited in high-density canopies	Complex data acquisition	[51]
Temporal Fusion	Dynamic structural discrimination	Apple, peach	Low wind, stable illumination	improved registration accuracy, unstable under wind conditions	High model complexity	[54]

Systematically compare the reconstruction performance across different technical approaches, Figure 3 presents a unified visual reference illustrating the three major technical pathways for fruit tree branch feature extraction. The 2D feature extraction pathway employs traditional computer vision methods, capturing planar contours through image preprocessing and edge detection. The 3D feature extraction pathway leverages point cloud or mesh data, combining point cloud segmentation, geometric structure fitting, and skeleton extraction with topological reconstruction through geometric techniques such as least squares fitting, RANSAC, and region growing, enabling precise acquisition of diameter, length, and branching angle. The deep learning extraction pathway, meanwhile, draws on advanced architectures including CNNs, Transformers, and PointNet to achieve automated feature learning spanning from semantic segmentation to keypoint estimation. Taken together, the figure clearly illustrates the technical evolution from conventional pixel-level analysis, through high-dimensional geometric reconstruction, to intelligent semantic understanding.

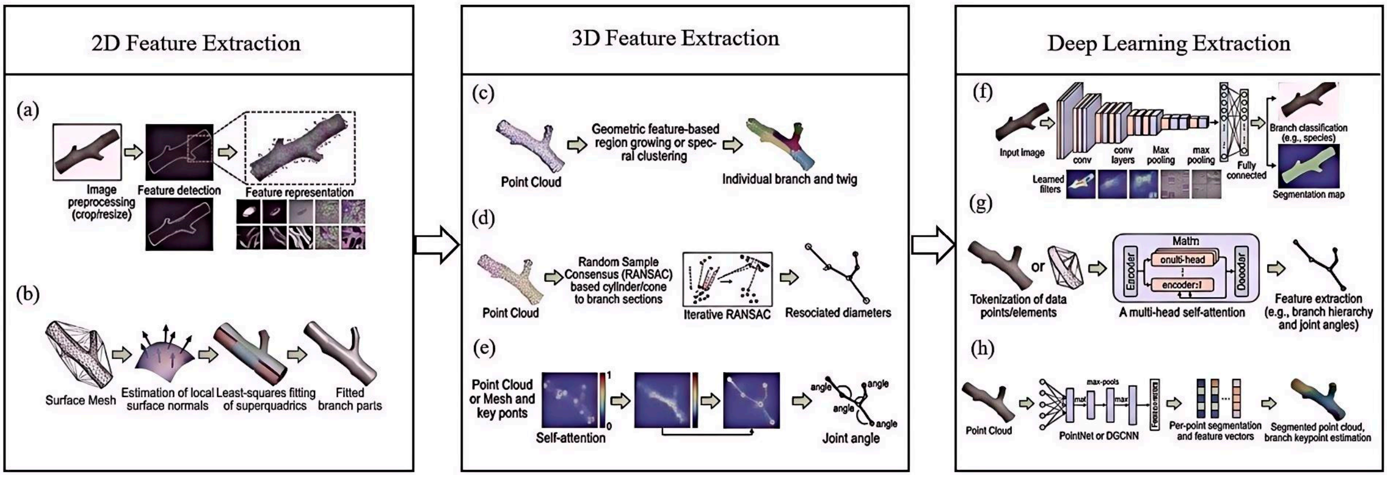


Figure 3. Schematic illustration of branch feature extraction algorithm principles. (**a**) 2D Image Processing; (**b**) Geometry-Based Surface Fitting; (**c**) Point Cloud Segmentation Methods; (**d**) Point Cloud-Based Geometric Structure Fitting; (**e**) Skeleton Extraction and Topological Reconstruction; (**f**) Neural Radiance Fields (NeRF); (**g**) Transformer Architectures; (**h**) Point Cloud Deep Learning Networks.

3. 3D Reconstruction Algorithms for Fruit Tree Branches

3D reconstruction technology employs computer graphics and image processing techniques to extract the geometric and topological information of target objects from 2D image data, constructing three-dimensional mathematical models for virtual reconstruction. In the context of fruit trees, 3D branch reconstruction aims to recover the true spatial geometry and topological relationships of tree structures in complex natural environments, providing a foundational basis for intelligent equipment applications, crop growth monitoring, and automated orchard management [55]. Compared with architectural or industrial scenes, fruit tree structures exhibit pronounced irregularity, multi-scale branching, and severe occlusion—posing significant challenges for conventional 3D reconstruction methods. Fine branch regions in particular are characterized by weak texture, small scale, and frequent leaf occlusion, often resulting in structural fractures and topological gaps in reconstruction outcomes.

From a technical development perspective, 3D reconstruction of fruit tree branches has progressively followed a trajectory from 3D data acquisition and explicit geometric modeling through to implicit representation-based reconstruction. The technical framework for fruit tree branch 3D reconstruction algorithms is illustrated in Figure 4, reflecting the evolution from explicit structural representations such as point clouds and meshes, toward implicit representations integrated with deep learning and multimodal reconstruction.

This chapter examines the performance of various algorithms in handling irregular branch structures, addressing occlusion-induced fractures, and achieving real-time dynamic reconstruction, with focused discussion on key challenges including multi-source data fusion and reconstruction accuracy optimization.

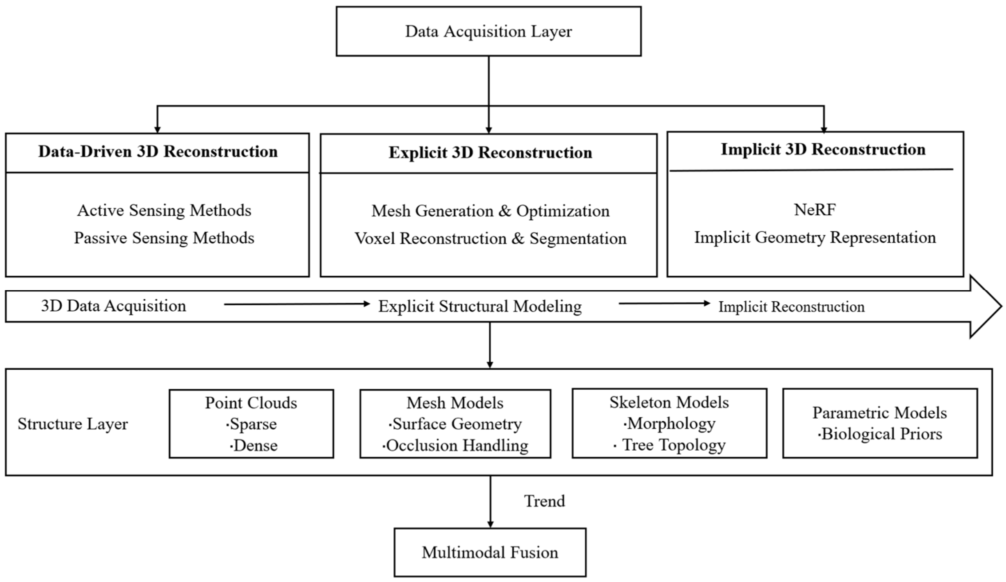


Figure 4. Technical framework of 3D reconstruction algorithms for fruit tree branches.

3.1. 3D Reconstruction Methods Based on Data Acquisition Approaches

According to the manner in which 3D data is acquired, reconstruction methods can be broadly divided into passive vision-based approaches and active sensing-based approaches. The former relies primarily on the geometric relationships between multi-view images to recover 3D structure, with stereo vision and multi-view geometry as representative techniques. The latter directly acquires scene depth information through active sensing devices such as laser scanners or structured light systems.

3.1.1. Active Sensing-Based Reconstruction Methods

As an active ranging technology, LiDAR can directly acquire 3D point cloud data with precise spatial coordinates. In early research, Lefsky et al. applied LiDAR to forest ecological measurement, systematically validating its effectiveness in vegetation vertical structure extraction and canopy parameter estimation, laying the groundwork for subsequent plant 3D modeling [56]. Compared with passive imaging methods, LiDAR does not rely on texture information and offers greater stability in complex vegetation environments, giving it a clear advantage in branch structure extraction [57].

With the advancement of sensing equipment, laser scanning has gradually diversified into multiple modalities. Terrestrial laser scanning (TLS) enables high-density point cloud acquisition suited to fine characterization of trunks and primary branches. Mobile laser scanning (MLS), combined with positioning systems, enables continuous data collection and improves orchard-scale data acquisition efficiency. UAV-based LiDAR scanning, by leveraging aerial perspectives through canopy gaps, can effectively capture structural information in the upper canopy regions.

Structured light and time-of-flight (ToF) depth cameras acquire depth information through active light pattern projection, and are commonly used for local branch modeling and grasp region reconstruction during fruit harvesting operations, thereby improving operational precision and environmental perception [58]. However, LiDAR remains susceptible to sampling deficiencies at branch terminations due to scan angle limitations, and data gaps can arise from occlusion within dense canopies [59]. While structured light and ToF cameras can acquire depth information at high frame rates, they are prone to interference from natural light and have limited range, making them unsuitable for large-scale orchard modeling. They are therefore predominantly used as auxiliary sensors in combination with other modalities, commonly deployed in agricultural robots for local branch modeling and close-range perception tasks.

3.1.2. Passive Vision-Based Reconstruction Methods

Beyond active 3D reconstruction, passive approaches offer an alternative path to 3D reconstruction by exploiting naturally available visual information—primarily images—to recover scene structure without active sensing devices, offering the advantages of low cost and high flexibility. Early methods relied mainly on feature matching and geometric constraints to recover sparse 3D structures.

A representative technique is Structure from Motion (SfM), which reconstructs sparse point clouds from multiple images of the same scene captured from different viewpoints [60]. The Photo Tourism system proposed by Snavely et al. was among the first to extend this idea to large-scale image-based reconstruction [61]. Building on this, the Multi-View Stereo (MVS) algorithm proposed by Furukawa and Ponce achieves dense point cloud generation by searching for pixel correspondences across multiple views [62], jointly considering geometric and photometric consistency to expand the 3D model from sparse structure to dense geometric surfaces—forming the now-classic "SfM + MVS" reconstruction pipeline.

In agricultural applications, SfM-based techniques have been evaluated for orchard canopy modeling, enabling 3D reconstruction of tree structures. However, the weak and repetitive texture of branches leads to poor feature matching stability and limited fine branch recovery, while natural lighting variation and occlusion significantly affect reconstruction accuracy. Recent efforts have sought to improve canopy data acquisition through multi-view path planning and viewpoint optimization, yet the inherent limitations of 2D image information continue to constrain complete structural reconstruction.

3.2. 3D Reconstruction Methods Based on Explicit Representations

Explicit representation methods directly describe 3D geometry through discrete structures such as point clouds, meshes, or voxels, and represent the dominant expression forms in traditional 3D reconstruction.

3.2.1. Point Cloud Processing and Surface Mesh Generation

Point clouds constitute the most fundamental form of explicit 3D representation, typically obtained through LiDAR or image-based reconstruction [63–65]. Due to the presence of raw noise, uneven density, and local data absence, point clouds are not directly suited for structural analysis and typically require further conversion into continuous surface mesh models.

Triangular meshes are the most widely used explicit 3D representation. The Ball-Pivoting Algorithm (BPA) proposed by Bernardini et al. enables stable mesh generation from point clouds, simulating the rolling of a sphere across the point cloud surface to generate triangular facets, and performing reliably for regular structure reconstruction [66]. Subsequent work incorporated normal consistency optimization and hole-filling algorithms to improve model completeness [66]. Least-squares surface fitting and radial basis function (RBF) methods have also been applied to recover locally continuous surfaces from sparse point clouds, offering practical value in branch topological modeling and local morphological representation [67].

3.2.2. Voxel-Based Reconstruction and Segmentation

Voxel-based modeling discretizes space into a 3D grid, representing object structure through occupancy probabilities. In fruit tree branch scenes, voxel representations can transform discrete point clouds into continuous spatial occupancy structures, facilitating holistic expression of complex branching morphology [68–71]. Through voxelization, issues of uneven point cloud density and local data absence can be effectively mitigated, enabling slender branches to form spatially connected regions and improving structural completeness [72–78]. Building on this, voxel models lend themselves to 3D connectivity analysis and morphological operations, providing support for branch skeleton extraction and topological structure recovery [68,69].

Multi-view or multi-temporal voxel fusion methods can progressively reconstruct occluded structural regions, partially alleviating data gaps caused by branch–foliage occlusion in orchard environments [72,73]. However, due to the small scale of fine branches, insufficient voxel resolution tends to produce structural blurring or fractures, while high-resolution modeling significantly increases storage and computational overhead, limiting real-time applicability in large-scale orchard scenes [74,75].

3.3. 3D Reconstruction Methods Based on Implicit Representations

With the development of neural representations, 3D reconstruction has progressively shifted from explicit geometric modeling toward learning-driven approaches. Methods based on neural representations can be broadly categorized according to scene representation type into implicit geometric representations, radiance field methods, and explicit representation methods.

3.3.1. Neural Radiance Fields (NeRF)

Mildenhall et al. introduced Neural Radiance Fields (NeRF) in 2020 [79], using neural networks to learn color and density at arbitrary spatial positions for continuous 3D representation. NeRF models the 3D scene as a continuous implicit function, directly predicting color and volumetric density from spatial position and viewing direction inputs. A volumetric rendering mechanism then integrates the continuous field along each viewing ray to generate pixel values. This approach enables the network to implicitly recover scene structure without explicit geometric representation, achieving high-fidelity expression of complex 3D morphology. NeRF marked the entry of 3D reconstruction into the neural representation era, and compared with traditional methods, its ability to generate plausible completions in view-missing regions offers potential advantages for occluded fruit tree canopy scenes.

Instant-NGP, subsequently proposed by Thomas Müller et al., significantly improved training efficiency [80], bringing neural reconstruction closer to practical applicability. Its core contribution lies in addressing the long training times and high computational costs of early NeRF models. By introducing multi-resolution hash encoding, spatial coordinates are efficiently mapped to compact feature representations, substantially reducing network parameter count and improving both training and inference efficiency. Overall, Instant-NGP marks an important transition for neural radiance fields from high-quality but computationally prohibitive to real-time capable, laying a critical foundation for the practical deployment of learning-driven 3D reconstruction methods.

3.3.2. Implicit Geometric Representations

Implicit representation methods describe 3D shapes by learning a Signed Distance Function (SDF) [81–84]. The DeepSDF framework proposed by Park et al. achieves high-quality geometric representation by reformulating geometric reconstruction as the learning of a continuous function that predicts the signed distance from any point in space to the object surface [84]. In this framework, the sign of the function value distinguishes whether a point lies inside or outside the object, while zero-crossings correspond to the object surface. This implicit representation enables the model to characterize complex geometric structures in continuous space, achieving fine-grained and compact 3D reconstruction. Compared with explicit mesh methods, implicit representations naturally accommodate complex topological structures, offering advantages for continuous branch modeling, and have been increasingly applied in plant structure reconstruction research—though they remain strongly dependent on training data and optimization procedures.

3.4. Multi-View Data Fusion and Reconstruction Accuracy Optimization

With the advancement of multi-source sensing technologies, single-modality data has proven insufficient to meet the demands of high-precision fruit tree modeling, making the fusion of multi-view and multi-modal data an increasingly important trend in 3D reconstruction [81–83]. RGB images provide rich texture information, LiDAR data offers high-precision spatial geometry, and temporal data captures the continuous evolution of structure over time. Through unified calibration and registration frameworks, data from different modalities can be jointly optimized across spatial and temporal dimensions, compensating for the inherent limitations of individual sensors. Cross-platform acquisition strategies—such as the coordinated use of UAVs and ground-based laser scanners—can simultaneously cover both the upper canopy and interior regions, significantly improving structural visibility under occlusion [85]. Further research has introduced temporal consistency constraints, stabilizing reconstruction results through continuous observation and reducing structural fluctuations caused by random noise. While multi-source fusion substantially improves reconstruction completeness and accuracy, it also introduces new challenges including complex data alignment, increased computational overhead, and system calibration difficulties. Achieving an effective balance between accuracy and efficiency therefore remains a critical issue in the development of smart orchard 3D perception systems. In terms of suitability for specific orchard conditions, TLS-based LiDAR reconstruction delivers the highest geometric accuracy, with Chamfer Distance typically below 5 mm for dormant season apple and walnut trees in open row systems, but suffers from significant occlusion gaps in dense canopies and is unsuitable for real time robotic applications due to data acquisition constraints. SfM plus MVS passive reconstruction is cost effective and widely applicable to UAV-based orchard surveys, but weak branch texture leads to matching failures for branches thinner than 8 mm, and reconstruction completeness in heavily occluded inner canopy regions typically falls below 60 percent. NeRF-based implicit methods show promise for handling occluded and view incomplete regions and have demonstrated improved fine branch recovery in controlled apple tree experiments, but training times of several hours per scene and the need for dense multi-view input currently limit field deployment. Voxel-based methods offer computational predictability and support multi-view fusion, but resolutions above 5 mm result in structural blurring of fine branches, while higher resolutions become computationally infeasible for full canopy scenes. These trade offs highlight the need for hybrid reconstruction pipelines tailored to specific application scenarios, such as TLS for phenotyping research and lightweight NeRF variants for robotic guidance. Nevertheless, fine branch recovery and complex topology preservation remain key research bottlenecks, motivating further work on algorithm optimization and evaluation strategies.

Figure 5 presents a systematic comparative illustration of 3D reconstruction techniques. Subplots (a–g) all use the same fruit tree branch as the reconstruction subject, ensuring visual consistency in morphological characteristics across methods. The figure traces the full pipeline from front-end data-driven acquisition—including LiDAR scanning, structured light projection, and SfM-based initial point cloud recovery—through mid-level explicit geometric reconstruction using Delaunay triangulation, mesh optimization, and voxel fusion for topological representation, to cutting-edge implicit 3D reconstruction via NeRF-based neural radiance field rendering and implicit geometric representation through SDF learning. It should be noted that the point cloud visualizations shown serve as conceptual illustrations of reconstruction principles rather than actual measurement outputs generated by LiDAR, structured light, or SfM algorithms.

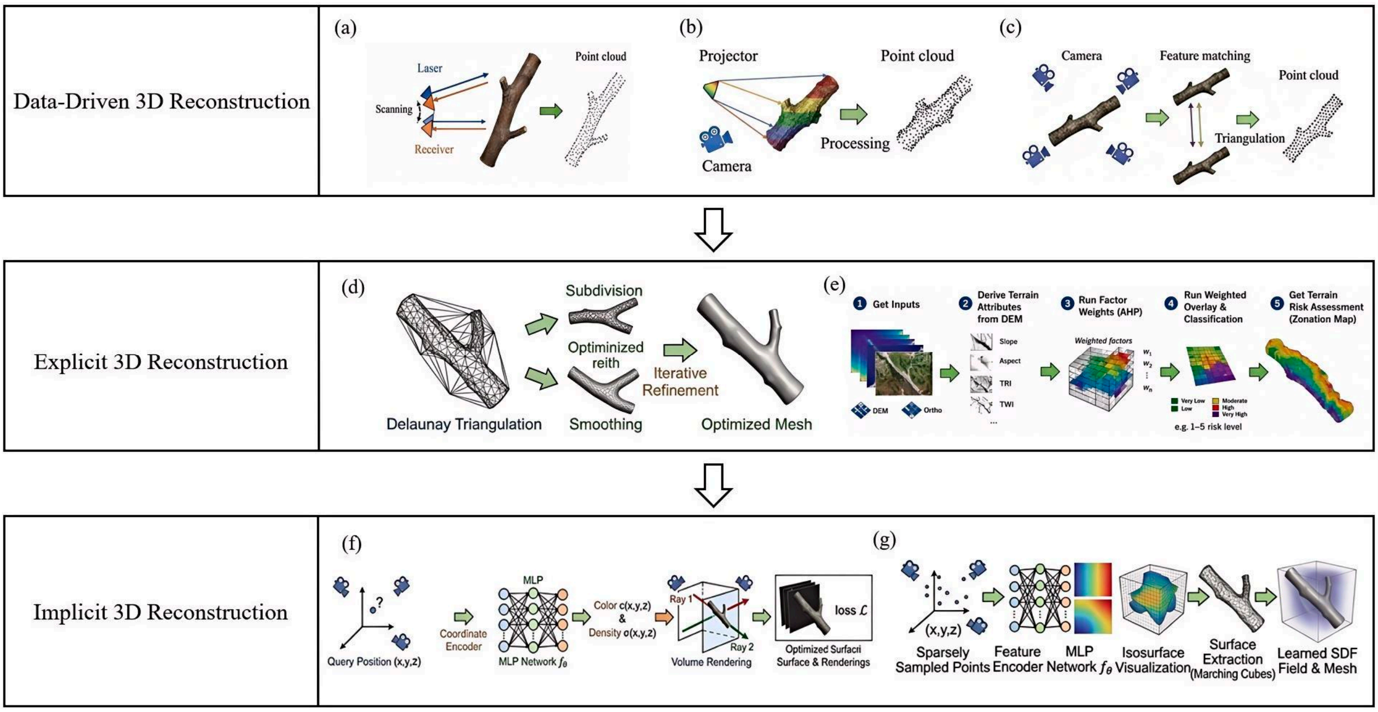
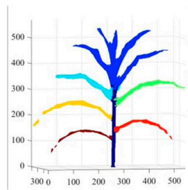


Figure 5. Comparison of different 3D reconstruction techniques and their corresponding reconstruction processes. (**a**) LiDAR; (**b**) Structured Light; (**c**) Structure from Motion; (**d**) Triangle mesh generation and optimization; (**e**) Voxel-based reconstruction and segmentation; (**f**) Neural Radiance Fields (NeRF); (**g**) Implicit geometric representation.

Building on this, to further extend the discussion from an intuitive comparison at the "methodological pipeline level" to a systematic summary at the "method attribute level," it is necessary to comprehensively analyze different 3D reconstruction techniques in terms of key dimensions such as data source, representation form, reconstruction accuracy, and application scenarios. Accordingly, focusing on the task of fruit tree branch reconstruction, this study categorizes representative methods and provides a comparative summary of their characteristics and applicability under a unified evaluation framework. Table 5 presents a systematic comparison of different 3D reconstruction methods in terms of data type, representation, accuracy, and methodological characteristics, thereby more clearly revealing their intrinsic differences and developmental progression.

Table 5. Comparison of 3D reconstruction methods for fruit tree branches.

Method Category	Techniques	Representation	Suitable Scenarios	Characteristics	Exhibition	References
Passive Sensing	SfM/MVS	Point cloud/Mesh	Open canopies; UAV surveys; texture-rich dormant trees	Mature and stable, but texture-dependent		[72]
LiDAR Scanning	LiDAR	Point cloud	Open-row orchards; high-precision phenotyping	High precision for complex structures		[59]
Depth Camera	RGB-D/ToF	Point cloud	Near-range robotic perception; local reconstruction	Strong real-time capability for local perception		[73]
Triangular Mesh Generation and Optimization	BPA	Mesh	Dense, high-quality point cloud surfaces	Dense intersections, prone to errors		[74]
Voxel Methods	OctoMap	Voxel	Large-scale reconstruction; multi-view fusion	Facilitates multi-view data fusion		[75]
Neural Radiance Fields	NeRF	Implicit field	Occluded canopies; fine branch reconstruction (offline)	Strong structural representation & completion		[76]
Passive Sensing	DeepSDF	SDF	Shape completion; offline branch reconstruction	Well-suited for complex topological structures		[77]

4. Metric Evaluation and Performance Analysis

In research on fruit tree canopy branch feature extraction and 3D reconstruction, algorithm performance evaluation is a critical component for validating model effectiveness and application potential. Early evaluation frameworks in the 3D vision field centered primarily on geometric error, measuring algorithm performance by comparing spatial differences between reconstruction outputs and ground truth models [86]. However, unlike regular industrial objects, fruit trees exhibit highly non-rigid morphology, multi-scale branching, and complex occlusion structures, making single geometric metrics insufficient to comprehensively reflect model quality. As plant phenotyping research has progressed, there has been growing recognition of the need to establish a comprehensive evaluation framework spanning multiple dimensions—including geometric accuracy, structural consistency, computational efficiency, and cross-scene generalization [87,88]. This chapter therefore presents a systematic review of evaluation methodologies for fruit tree branch 3D modeling, organized around evaluation metric frameworks, experimental design approaches, and performance analysis strategies. To systematically illustrate the evaluation process and the relationships among different assessment dimensions, Figure 6 presents the overall flowchart of the 3D reconstruction evaluation framework. The framework integrates geometric accuracy, structural consistency, computational efficiency, and generalization capability into a unified evaluation pipeline, providing a comprehensive basis for performance analysis of fruit tree branch reconstruction methods under complex orchard environments.

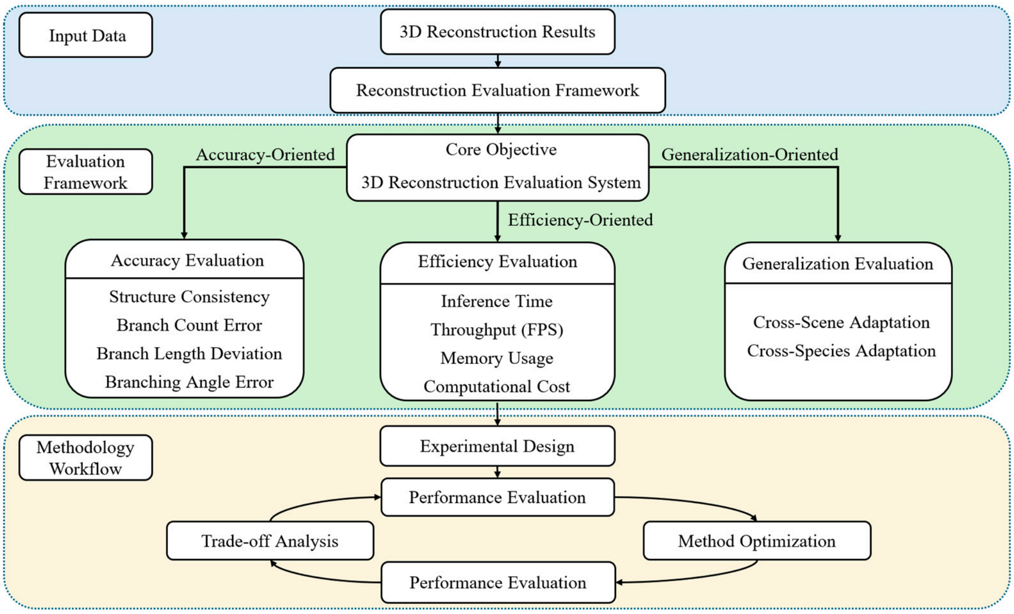


Figure 6. Flowchart of the 3D reconstruction evaluation framework.

4.1. Development and Selection of Evaluation Metrics

Evaluation metrics for 3D reconstruction originally emerged from geometric reconstruction research in computer vision. Seitz et al. emphasized geometric error as the core standard for algorithm comparison, a principle that became an important foundation for subsequent research [89]. However, as plant modeling research advanced, scholars increasingly recognized that relying solely on geometric distance metrics was insufficient to accurately assess plant structural recovery, prompting the evaluation framework to expand toward structure-aware assessment. Well-chosen evaluation metrics form the basis of meaningful algorithm comparison, and the prevailing view holds that a sound 3D reconstruction evaluation should simultaneously consider spatial geometric consistency and structural semantic recovery, enabling comprehensive analysis of algorithm performance.

In this study, all evaluation metrics are normalized to a unified scale of [0, 1] using min–max normalization to enable fair comparison across heterogeneous indicators. For metrics where lower values indicate better performance (e.g., Chamfer Distance, RMSE, inference time), inverse normalization is applied to ensure consistency in scoring direction.

Beyond defining evaluation metrics, recent studies have increasingly emphasized selecting evaluation metrics according to the characteristics of different reconstruction algorithms. Geometry-based reconstruction methods are primarily evaluated using geometric accuracy metrics, whereas learning-based segmentation and reconstruction methods additionally require semantic prediction and robustness evaluation. Consequently, comprehensive multi-dimensional evaluation has gradually become the standard practice for comparing fruit tree branch reconstruction algorithms.

(1) Geometric Reconstruction Accuracy

Commonly used metrics include Chamfer Distance and Hausdorff Distance, which respectively capture overall error and maximum deviation, with the latter being more sensitive to fine branch absence [90,91]. RMSE evaluates reconstruction stability from the perspective of overall coordinate error, but is less sensitive to localized structural inaccuracies. These metrics are individually normalized and averaged to obtain a unified geometric accuracy score.

Different categories of reconstruction methods exhibit distinct characteristics under these metrics. LiDAR-based explicit reconstruction methods generally achieve stable RMSE owing to accurate spatial measurements, whereas image-based SfM/MVS approaches are more sensitive to illumination variation and weak branch texture, often resulting in larger geometric deviations in fine branches. Recently developed implicit neural representation methods, such as NeRF, frequently demonstrate improved Chamfer Distance through continuous surface modeling in partially occluded regions, although local structural errors may still occur under severe occlusion.

(2) Structural Consistency Evaluation

In response to the specific demands of plant modeling, evaluation frameworks have progressively incorporated topological and morphological metrics—such as branch count error, branch length deviation, and bifurcation angle error—to characterize structural recovery capability [92]. Compared with traditional geometric metrics, these measures place greater emphasis on biological morphological semantics, but depend on high-quality annotations and remain limited in robustness under occlusion and scale variation. Each metric is normalized separately and aggregated to form the structural consistency score.

Compared with purely geometry-based evaluation, these topology-aware metrics better reflect the biological characteristics of fruit trees. Skeleton-based reconstruction algorithms generally preserve branch connectivity more effectively than surface-based methods, while graph-constrained reconstruction methods have recently shown improved bifurcation recovery by incorporating structural priors.

(3) Feature Extraction Robustness and Stability Evaluation

In complex orchard environments, illumination changes, occlusion, and background interference significantly affect feature extraction performance. To address this, metrics such as Precision, Recall, and F1-score have been introduced to evaluate branch detection and segmentation [93]. High recall facilitates fine branch detection, while high precision improves noise resistance. Cross-scene experiments are additionally employed to assess model stability across varying environmental conditions. The final robustness score is obtained by averaging the normalized values of these metrics.

These metrics are widely adopted in deep learning-based branch segmentation algorithms, including CNN-, PointNet-, and Transformer-based models. CNN-based approaches generally achieve high Precision under relatively simple backgrounds, whereas Transformer architectures improve Recall under dense canopy occlusion by capturing long-range contextual information. Point cloud learning networks such as PointNet++ further improve three-dimensional feature representation.

(4) Computational Efficiency and Real-Time Performance Evaluation

https://doi.org/10.3390/agronomy16131274

As applications increasingly demand real-time operation, evaluation frameworks have expanded to encompass computational efficiency, primarily including inference time, frame rate (FPS), and resource consumption. These three aspects jointly reflect system responsiveness and engineering deployability, and exist in trade-off with reconstruction accuracy—high precision typically comes with greater computational overhead, while lightweight methods may sacrifice fine detail representation [94]. Multi-metric co-optimization has therefore become a key focus of current research. Computational efficiency is evaluated using inference time, FPS, and memory consumption.

Traditional geometry-based reconstruction algorithms generally require relatively low computational cost but involve considerable manual preprocessing. Deep learning-based reconstruction methods usually achieve higher reconstruction accuracy at the expense of longer inference time. Recent lightweight neural rendering methods, such as Instant-NGP, have substantially improved computational efficiency, making learning-based reconstruction increasingly suitable for real-time orchard applications.

(5) Generalization and Adaptability Evaluation

Given the significant variability across tree species and growth stages, generalization capability has emerged as an important evaluation dimension [95]. Relevant studies typically validate models through cross-dataset, cross-season, and multi-scene experiments, and employ domain adaptation and multi-species modeling strategies to improve transferability. Models with strong generalization can reduce data dependency and enhance practical application value. The generalization score is obtained by normalizing and inverting the degradation metric across cross-domain experiments.

Recent studies increasingly validate reconstruction models across different orchards, tree species, seasons, and acquisition platforms. Compared with traditional geometry-based methods, deep learning models generally achieve higher accuracy on training datasets but often experience performance degradation under unseen environments. Consequently, domain adaptation, self-supervised learning, and multimodal fusion have become important strategies for improving model generalization.

Overall, recent studies increasingly recommend reporting multiple complementary evaluation metrics rather than relying on a single criterion. Geometric accuracy, structural consistency, feature robustness, computational efficiency, and generalization capability collectively provide a more comprehensive understanding of algorithm performance and facilitate fair comparison among different categories of fruit tree branch reconstruction methods. A comparison of the various evaluation metrics is presented in Table 6.

Table 6. Performance evaluation metric framework for 3D reconstruction.

Metric Category	**Metric**	**Evaluation Target**	**Advantage**	**Disadvantage**	**References**
Geometric Accuracy	► Chamfer Distance ► FHausdorff Distance ► RMSE	★ Average point cloud deviation ★ Maximum structural error ★ Global geometric consistency	• Stable for deep models • Fracture-sensitive • Physically interpretable	• Outlier-insensitive • Noise-sensitive • No topology	[90]
Structural Consistency	► Branch Count Error ► Branch Length Deviation ► Bifurcation Angle Error	★ Topology ★ Scale accuracy ★ Structure	• Biologically relevant • Morphology-aware • Phenotypic analysis	• Needs annotation • Skeleton-sensitive • Expensive	[91]

Table 6. *Cont.*

Metric Category	Metric	Evaluation Target	Advantage	Disadvantage	References
Feature Stability	► Precision ► Recall ► F1-score	★ False positives ★ Completeness ★ Overall	• Fewer FP • Sensitive to weak parts • Balanced	• Misses details • More FP • No location info	[92]
Computational Efficiency	► Inference Time ► FPS ► Resource	★ Speed ★ Real-time ★ Cost	• Comparable • Intuitive • Practical	• Implementation-dependent • Deployment-ready • No standard	[93]
Generalization	► Cross-scene ► Cross-species	★ Robustness ★ Transferability	• Real-world • Deployment-ready	• Costly • Data-limited	[94]

To facilitate comparison across methods, the proposed multi-dimensional evaluation results are visualized using a radar chart. Each axis represents a normalized evaluation metric, enabling a compact representation of overall performance across geometric accuracy, structural consistency, robustness, efficiency, and generalization, as shown in Figure 7.

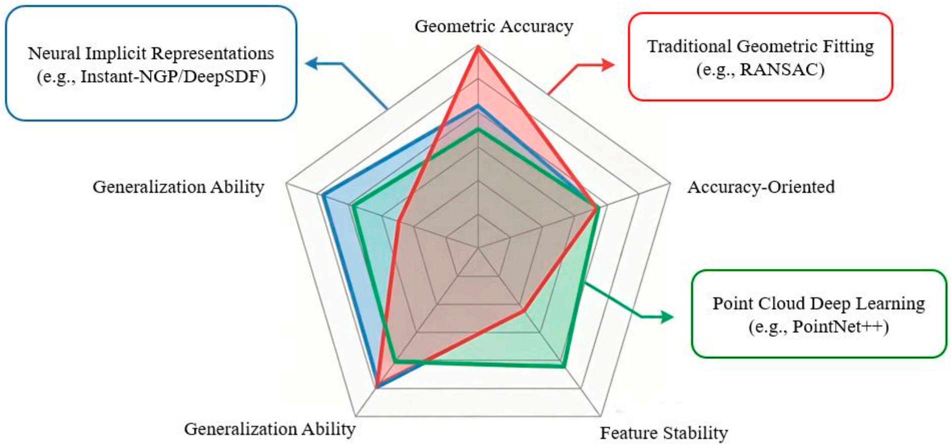


Figure 7. Five-Dimension Evaluation Metric Radar Chart of 3D Reconstruction Algorithms for Fruit Tree Branches.

4.2. Evaluation Methods and Experimental Design Paradigms

Rigorous and well-designed experimentation is essential for enabling fair algorithm comparison. Early multi-view reconstruction studies lacked unified testing standards, making results difficult to reproduce. Knapitsch et al. established the first high-quality public benchmark for 3D reconstruction evaluation through the Tanks and Temples benchmark, providing a standardized reference for subsequent research [96].

In the agricultural domain, due to the absence of unified datasets, researchers have largely relied on UAV multi-view imagery, ground-based LiDAR scans, and RGB-D data for experimentation. In recent years, the gradual emergence of public plant point cloud datasets has made algorithm comparison more systematic. Typical experimental workflows include unified input data, multi-algorithm comparison, multi-metric comprehensive evaluation, and statistical significance analysis. For fair comparison, algorithms should be evaluated on common datasets with identical preprocessing pipelines, consistent parameter initialization, and repeated experiments across at least three independent orchard scenes or growth-stage conditions to ensure cross-scene validity. Regarding statistical significance analysis, the most commonly recommended approaches in the plant modeling literature include paired Wilcoxon signed-rank tests for non-parametric comparison of algorithm outputs across multiple scenes, and analysis of variance (ANOVA) followed by Tukey's post hoc test for multi-algorithm comparisons. Effect sizes (Cohen's d) should also be reported alongside *p*-values to distinguish statistically significant from practically meaningful performance differences. Ablation studies can effectively analyze the contribution of individual network components to overall performance, thereby revealing the rationale behind model design decisions [97].

At the tooling level, CloudCompare is widely used for point cloud error analysis, MeshLab for mesh quality assessment, and Open3D provides an extensible 3D data processing framework—collectively improving experimental reproducibility and evaluation efficiency [98]. These tools enable visualization of error distributions, helping researchers identify regions where reconstruction fails.

4.3. Performance Analysis and Optimization Directions Based on Evaluation Metrics

A review of experimental results across existing studies reveals significant performance differences among 3D reconstruction methods on various evaluation metrics, reflecting the key bottlenecks of current algorithms in complex orchard scenes.

From the perspective of geometric accuracy metrics, small-scale branch regions consistently exhibit larger reconstruction errors. As their dimensions approach the resolution limits of sensors, insufficient point cloud sampling density or noise interference leads to notably increased Chamfer Distance and point-to-surface error, indicating that existing methods remain limited in fine-grained structural representation. Regarding structural consistency metrics, occluded regions frequently cause branch count errors and topological connection failures. While multi-view fusion can partially reduce structural absence, its effectiveness depends on acquisition density and viewpoint distribution—when observations are insufficient, models still struggle to accurately recover internal structures. From the standpoint of computational efficiency, deep learning methods, despite their accuracy advantages, typically incur high inference times and computational resource demands, limiting their applicability in real-time agricultural scenarios. Traditional geometric methods, by contrast, offer greater computational efficiency but fall short in complex structural recovery. Furthermore, in generalization evaluations, notable performance fluctuations across different tree species and acquisition conditions indicate that current models remain sensitive to distributional shifts and lack stable cross-scene adaptability.

Based on the above metric analysis, current research has converged on several directions: introducing structural priors and topological constraints to improve structural consistency; combining multimodal fusion and self-supervised learning to enhance model robustness; and applying lightweight networks and temporal modeling to optimize computational efficiency and dynamic adaptability. These directions, each grounded in different evaluation dimensions, offer important pathways toward the coordinated improvement of accuracy, efficiency, and generalization [98,99].

5. Technical Challenges and Development Trends in Fruit Tree Canopy Branch Feature Extraction and 3D Modeling

Despite the significant progress made in fruit tree feature extraction and 3D modeling, substantial challenges remain on the path toward practical deployment, including branch–foliage occlusion, fine structure recovery, and dynamic environment interference [100–102]. At the same time, emerging paradigms such as multimodal fusion, end-to-end modeling, and self-supervised learning are reshaping the technical landscape of this field, driving fruit tree perception systems toward higher levels of automation and intelligence. This chapter presents a systematic analysis of these core challenges and development trends.

https://doi.org/10.3390/agronomy16131274

5.1. Core Technical Challenges

To systematically summarize the relationship between key technical challenges and corresponding methodological paradigms in fruit tree canopy feature extraction and 3D reconstruction, the overall framework is illustrated in Figure 8.

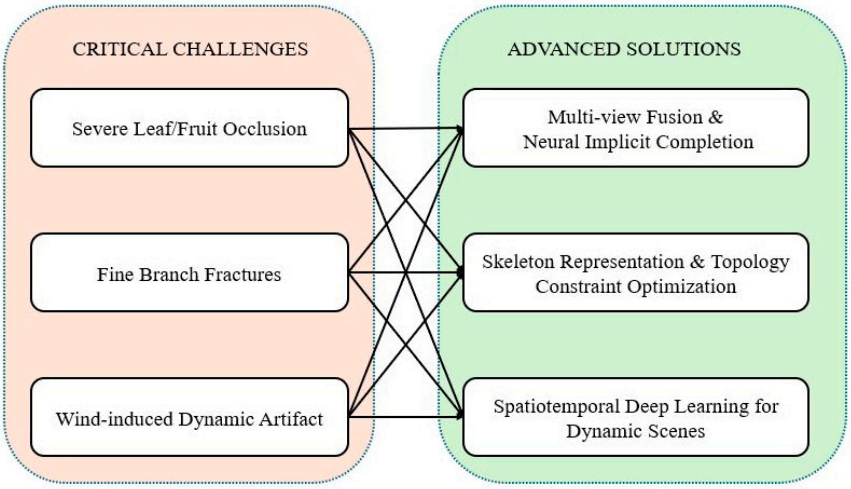


Figure 8. Framework of Technical Challenges and Solution Strategies for Fruit Tree Canopy Reconstruction.

5.1.1. Branch–Foliage Occlusion and Structural Absence

Branch–foliage occlusion is the primary factor affecting the quality of fruit tree feature extraction and 3D reconstruction [100,102]. Since leaves vastly outnumber branches, visual and LiDAR sensor data predominantly captures leaf surface information, leaving branch structures severely underrepresented in point clouds. Single-viewpoint observation is insufficient for capturing the internal structure of the canopy, making structural recognition and information recovery in occluded regions a critical challenge [103].

Multi-view data fusion has been established as an effective strategy for mitigating occlusion [104], progressively reconstructing invisible regions by acquiring data from multiple observation directions, thereby improving branch feature continuity and model completeness. In practice, circumferential ground-based acquisition, coordinated UAV and ground sensor observation, and multi-height scanning strategies have been widely adopted to enhance coverage of canopy interior structures. Multi-view geometric optimization methods have further improved reconstruction stability under complex conditions through joint estimation of camera poses and spatial structure. Building on this, structural priors and probabilistic modeling have been introduced to infer missing structures from plant growth patterns or characterize uncertain regions through spatial probability distributions. More recently, neural implicit representation methods have enabled prediction of occluded regions by learning continuous spatial density distributions, assisting in branch region segmentation and structural completion—marking an important transition in occlusion handling from deterministic geometric inference toward learning-driven modeling. However, multi-view fusion, while improving structural completeness, also substantially increases data redundancy and registration complexity [105,106]. How to reduce computational cost while maintaining reconstruction quality remains an important open problem.

5.1.2. High-Precision Recovery of Fine Branches

Fine branches, with scales approaching sensor resolution limits, are prone to loss during sampling, segmentation, and reconstruction, representing a key bottleneck in improving structural detail [107]. The weak texture and sparse point cloud density in fine branch regions further lead to missed detections, fractures, and topological errors during feature extraction. While increasing data acquisition resolution can partially improve detection capability, it also significantly raises computational burden, making it difficult to scale to large-scale applications [108].

Researchers have therefore increasingly turned to optimizing structural representations. Skeleton-based methods, by preserving topological structure rather than surface detail, enable effective representation of fine branches under sparse data conditions and have emerged as an important technical pathway [109,110]. Building on this, deep models combining point cloud learning and multi-scale feature fusion can predict missing structures from data, achieving fine-grained structural completion and improving branch segmentation and node recognition [111]. In recent years, topological consistency constraints and graph-based structural optimization have been incorporated into model training, effectively reducing branch fractures and erroneous connections by constraining node connectivity. Despite these advances in detail recovery, achieving stable and reliable fine branch reconstruction under complex occlusion conditions remains an open challenge.

5.1.3. Modeling Instability Under Wind-Induced Motion

Wind-induced branch and foliage motion in natural orchard environments causes geometric inconsistencies across multi-frame data, adversely affecting point cloud registration and model fusion quality. Traditional 3D reconstruction methods typically assume static scenes and therefore perform poorly in dynamic environments. Branch and leaf movement also causes variations in edge and texture information across temporal image sequences, reducing the consistency of branch detection and segmentation results.

Early approaches to this problem primarily employed motion detection and weight adjustment to reduce the influence of dynamic regions on reconstruction, combined with temporal consistency constraints to improve model stability [112]. As temporal modeling methods have matured, spatiotemporal feature modeling based on deep learning has become a growing research focus [113–115], improving structural recovery and feature extraction robustness under dynamic conditions through the integration of multi-frame temporal information [116–118].

5.2. Major Development Trends

As orchard environments become increasingly complex, single-sensor approaches are no longer sufficient for comprehensive acquisition of structural and semantic information, making multimodal fusion an important development direction [119,120]. By integrating visual imagery, LiDAR, UAV data, and depth information, complementary representation of geometric and semantic features can be achieved [121–123]. Research has gradually shifted from early data-level fusion toward feature-level and representation-level fusion, leveraging attention mechanisms for dynamic cross-modal selection and collaborative modeling.

At the same time, traditional multi-stage 3D reconstruction pipelines often suffer from cumulative error propagation, leading to the rapid emergence of end-to-end learning frameworks that directly predict 3D structural representations from raw sensory data. End-to-end models incorporating geometric constraints and structural priors have demonstrated stronger robustness in complex natural environments [124–126], enabling more efficient and stable mapping from sensory input to structural models.

In addition, limited annotated datasets in agricultural scenarios have further promoted the development of self-supervised learning. By constructing pretraining tasks based on multi-view consistency and temporal prediction, models can learn effective spatial structural representations from unlabeled data and achieve improved generalization across different orchard environments and tree species [127].

https://doi.org/10.3390/agronomy16131274

With the rapid advancement of agricultural robots and unmanned systems, real-time performance and edge deployment capability have also become critical requirements for 3D modeling algorithms. Progress in lightweight network design, model compression, and edge computing has enabled 3D perception systems to operate efficiently on mobile devices, accelerating the transition of applications such as automated harvesting, path planning, and intelligent inspection from experimental research to practical production environments [128].

The emergence of digital twin technology has introduced a new paradigm for fruit tree modeling by enabling continuous crop monitoring and intelligent decision support through virtual models synchronized with real orchards [129,130]. Within this framework, 3D reconstruction technology is shifting from one-time structural recovery toward long-term spatiotemporal dynamic modeling, providing continuous data updates and decision support for smart orchard management.

Overall, future research in fruit tree 3D reconstruction is expected to move toward more intelligent, robust, and integrated systems, enabling accurate perception, efficient modeling, and long-term dynamic management in complex orchard environments. To achieve these goals, several concrete recommendations are warranted. First, the field urgently needs publicly available benchmark datasets covering multiple fruit species (apple, pear, peach, cherry, walnut), multiple growth stages (dormant, flowering, full-leaf, harvest), and representative orchard management systems (trellis, open-center, high-density), with standardized annotation protocols for branch topology, branch length, diameter, and bifurcation angle. Second, standardized evaluation protocols—specifying preprocessing pipelines, parameter initialization ranges, test-train splits, and minimum scene diversity requirements—should be established to ensure reproducibility and fair cross-algorithm comparison. Third, cross-species and cross-orchard validation should be adopted as a required evaluation component in future studies, rather than treated as an optional supplement. Fourth, open-source algorithm implementations and calibration tool standards for multi-sensor systems (LiDAR + RGB-D + IMU) would substantially lower barriers to reproducibility. Finally, field trials conducted across different growth stages, canopy architectures, and geographic climatic conditions are essential to validate the practical deployment readiness of proposed systems beyond controlled laboratory or single-orchard settings.

6. Conclusions

This paper systematically reviews fruit tree canopy branch feature extraction and 3D reconstruction, tracing the methodological evolution from classical 2D image processing and geometry-based fitting, through point cloud segmentation and skeleton extraction, to deep learning and multimodal fusion. Active sensing- and passive vision-based reconstruction strategies are compared alongside explicit and implicit scene representation methods, with implicit neural approaches such as NeRF showing particular promise for handling structural incompleteness caused by branch–foliage occlusion. A five-dimensional evaluation framework—encompassing geometric accuracy, structural consistency, feature stability, computational efficiency, and generalization capability—is further proposed to support more comprehensive algorithm assessment beyond conventional geometric error metrics. Several key findings emerge from this review. First, no single reconstruction method dominates across all orchard conditions: TLS-based LiDAR excels in geometric precision for open-canopy systems, while learning-based implicit methods are better suited for occluded inner-canopy recovery. Second, the transition from dormant-season to leafy-season conditions consistently represents the most critical performance boundary for all method categories, highlighting the need for more robust feature extraction and

3D reconstruction algorithms under dense canopy conditions. Third, the lack of standardized benchmark datasets and evaluation protocols remains the primary bottleneck for rigorous cross-algorithm comparison, more so than algorithmic limitations themselves. The specific contribution of this review lies in providing the first comprehensive side-by-side critical comparison of 2D, 3D geometric, deep learning, and multimodal approaches under a unified fruit-tree-specific evaluation lens, and in proposing concrete directions—including cross-species benchmarking, open-source calibration tools, and standardized evaluation protocols—to accelerate the transition from laboratory methods to deployable smart orchard systems.

Despite these advances, three core challenges remain: branch–foliage occlusion causing persistent structural absence, the difficulty of reliably recovering fine branches near sensor resolution limits, and modeling instability induced by wind-driven dynamic deformation. Future progress will require the integration of structural priors and topological constraints, end-to-end learning frameworks that reduce cumulative error propagation, self-supervised strategies to overcome annotated data scarcity, and lightweight architectures suitable for real-time edge deployment. The convergence of these directions with digital twin technology is expected to drive fruit tree 3D perception systems toward continuous spatiotemporal orchard monitoring, providing a robust foundation for the next generation of intelligent orchard management.

Author Contributions: Conceptualization, J.C. and S.Z.; methodology, J.C. and S.Z.; investigation, J.C.; writing—original draft preparation, J.C.; visualization, J.C.; writing—review and editing, Y.J. and S.Z.; supervision, Y.J. and S.Z.; project administration, Y.J.; funding acquisition, Y.J. All authors have read and agreed to the published version of the manuscript.

Funding: The work was supported by Young Scientists Fund of the National Natural Science Foundation of China (No. 32401693), National Key Research and Development Program of China (No. 2025YFE0211700), Jiangsu Province Modern Agricultural Machinery Equipment and Technology Demonstration and Promotion Project (No. NJ2024-26), Priority Academic Program Development of Jiangsu Higher Education Institutions (No. PAPD2023-87) and Jiangsu University Talent Introduction and Research Launch Project (No.5501200009).

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: During the preparation of this manuscript, the authors used ChatGPT (version 5.5) for language editing, academic writing assistance, and text refinement. The authors reviewed and edited the output and take full responsibility for the content of this publication.

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

Abbreviations

Abbreviation	Definition
3D	Three-Dimensional
UAV	Unmanned Aerial Vehicle
2D	Two-Dimensional
AI	Artificial Intelligence
CNN	Convolutional Neural Network
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
DGCNN	Dynamic Graph Convolutional Neural Network
L1	Least Absolute Deviations
LBP	Local Binary Pattern
LiDAR	Light Detection and Ranging
PointNet	Point Network

PointNet++	Hierarchical Point Network
RANSAC	Random Sample Consensus
ResNet	Residual Network
RGB	Red Green Blue
SLAM	Simultaneous Localization and Mapping
U-Net	U-Shaped Convolutional Network
VGGNet	Visual Geometry Group Network
ViT	Vision Transformer
BPA	Ball-Pivoting Algorithm
MLS	Mobile Laser Scanning
MVS	Multi-View Stereo
NeRF	Neural Radiance Field
RBF	Radial Basis Function
SDF	Signed Distance Function
SfM	Structure from Motion
TLS	Terrestrial Laser Scanning
ToF	Time-of-Flight
FPS	Frames Per Second
RMSE	Root Mean Square Error

References

1. Urgo, M.; Terkaj, W. Integrating digital factory twin and AI for monitoring manufacturing systems through synthetic data generation and vision transformers. *CIRP Ann.* **2025**, *74*, 639–643. [CrossRef]
2. Valada, A.; Mohan, R.; Burgard, W. Self-supervised model adaptation for multimodal semantic segmentation. *Int. J. Comput. Vis.* **2020**, *128*, 1239–1285. [CrossRef]
3. Hess, C.; Bienert, A.; Härdtle, W.; Von Oheimb, G. Does tree architectural complexity influence the accuracy of wood volume estimates of single young trees by terrestrial laser scanning? *Forests* **2015**, *6*, 3847–3867. [CrossRef]
4. Doersch, C.; Gupta, A.; Efros, A.A. Unsupervised visual representation learning by context prediction. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 11–18 December 2015; pp. 1422–1430. [CrossRef]
5. Tu, H.Y.; Huang, D.; Huang, X.; Aheto, J.H.; Yi, R.; Yu, W.; Liu, J.; Liu, S.; Niu, M.; Xu, M. Detection of browning of fresh-cut potato chips based on machine vision and electronic nose. *J. Food Process Eng.* **2021**, *44*, e13627. [CrossRef]
6. Xu, Q.; Cai, J.; Zhang, W.; Bai, J.; Li, Z.; Tan, B.; Sun, L. Detection of citrus Huanglongbing (HLB) based on the HLB-induced leaf starch accumulation using a home-made computer vision system. *Biosyst. Eng.* **2022**, *218*, 163–174. [CrossRef]
7. Hu, Y.; Asante, E.A.; Lu, Y.; Mahmood, A.; Buttar, N.A.; Yuan, S. Review of air disturbance technology for plant frost protection. *Int. J. Agric. Biol. Eng.* **2018**, *11*, 21–28. [CrossRef]
8. Memon, M.S.; Shen, B.; Liang, R.; Tang, Z.; Wang, S.; Zhou, W.; Memon, N. Automatic visual recognition, detection and classification of weeds in cotton fields based on machine vision. *Crop Prot.* **2024**, *187*, 106966. [CrossRef]
9. Huang, X.; Wang, W.; Li, Z.; Wang, Q.; Zhu, C.; Chen, L. Design method and experiment of machinery for combination of seed fertilizer and herbicide. *Int. J. Agric. Biol. Eng.* **2019**, *12*, 63–71. [CrossRef]
10. Huang, X.; Yu, S.; Xu, H.; Aheto, J.H.; Bonah, E.; Ma, M.; Wu, M.; Zhang, X. Rapid and nondestructive detection of freshness quality of postharvest spinaches based on machine vision and electronic nose. *J. Food Saf.* **2019**, *39*, e12708. [CrossRef]
11. Bing, L.; Sun, J.; Yang, N.; Wu, X.; Zhou, X. Identification of tea white star disease and anthrax based on hyperspectral image information. *J. Food Process Eng.* **2020**, *44*, e13584. [CrossRef]
12. Zhang, Z.; Yang, M.; Pan, Q.; Jin, X.; Wang, G.; Zhao, Y.; Hu, Y. Identification of tea plant cultivars based on canopy images using deep learning methods. *Sci. Hortic.* **2024**, *339*, 113908. [CrossRef]
13. Zhao, J.; Fan, S.; Zhang, B.; Wang, A.; Zhang, L.; Zhu, Q. Research status and development trends of deep reinforcement learning in the intelligent transformation of agricultural machinery. *Agriculture* **2025**, *15*, 1223. [CrossRef]
14. Tai, S.; Tang, Z.; Li, B.; Wang, S.; Guo, X. Cumin-harvesting mechanization of the Xinjiang cotton-cumin intercropping system: Review of the problem status and solutions. *Agriculture* **2025**, *15*, 809. [CrossRef]
15. Zhu, Y.; Zhang, S.; Tang, S.; Gao, Q. Research progress and applications of artificial intelligence in agricultural equipment. *Agriculture* **2025**, *15*, 1703. [CrossRef]
16. Canny, J. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **1986**, *8*, 679–698. [CrossRef]
17. Tai, S.; Tang, Z.; Li, B.; Wang, S.; Guo, X. Intelligent recognition and automated production of chili peppers: A review addressing varietal diversity and technological requirements. *Agriculture* **2025**, *15*, 1200. [CrossRef]

18. Pajares, G.; García-Santillán, I.; Campos, Y.; Montalvo, M.; Guerrero, J.M.; Emmi, L.; Romeo, J.; Guijarro, M.; Gonzalez-de-Santos, P. Machine-vision systems selection for agricultural vehicles: A guide. *J. Imaging* **2016**, *2*, 34. [CrossRef]
19. Ojala, T.; Pietikainen, M.; Maenpaa, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 971–987. [CrossRef]
20. Fischler, M.A.; Bolles, R.C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM* **1981**, *24*, 381–395. [CrossRef]
21. Chen, Z.; Yin, J.; Farhan, S.M.; Liu, L.; Zhang, D.; Zhou, M.; Cheng, J. A comprehensive review of obstacle avoidance for autonomous agricultural machinery in multi-operational environment. *Artif. Intell. Agric.* **2025**, *16*, 139–163. [CrossRef]
22. Edelsbrunner, H. *Geometry and Topology for Mesh Generation*; Cambridge University Press: Cambridge, UK, 2001. [CrossRef]
23. Haralick, R.M.; Sternberg, S.R.; Zhuang, X. Image analysis using mathematical morphology. *IEEE Trans. Pattern Anal. Mach. Intell.* **1987**, *9*, 532–550. [CrossRef] [PubMed]
24. Rusu, R.B.; Cousins, S. 3D is here: Point Cloud Library (PCL). In Proceedings of the IEEE International Conference on Robotics and Automation, Shanghai, China, 9–13 May 2011; pp. 1–4. [CrossRef]
25. Wang, Y.; Zhang, Z.; Jia, W.; Ou, M.; Dong, X.; Dai, S. A review of environmental sensing technologies for targeted spraying in orchards. *Horticulturae* **2025**, *11*, 551. [CrossRef]
26. Felzenszwalb, P.F.; Huttenlocher, D.P. Efficient graph-based image segmentation. *Int. J. Comput. Vis.* **2004**, *59*, 167–181. [CrossRef]
27. Jiang, L.; Xu, B.; Husnain, N.; Wang, Q. Overview of agricultural machinery automation technology for sustainable agriculture. *Agronomy* **2025**, *15*, 1471. [CrossRef]
28. Su, Z.; Li, S.; Liu, H.; He, Z. Tree skeleton extraction from laser scanned points. In Proceedings of the IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019; pp. 6091–6094. [CrossRef]
29. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]
30. Bai, J.; Xiao, H.; Ma, H.; Zhou, C. Artificial neural network modeling of drying kinetics and color changes of Ginkgo biloba seeds during microwave drying process. *J. Food Qual.* **2018**. [CrossRef]
31. Huan, J.; Cao, W.; Liu, X. A dissolved oxygen prediction method based on K-means clustering and the ELM neural network: A case study of the Changdang Lake, China. *Appl. Eng. Agric.* **2017**, *33*, 461–469. [CrossRef]
32. Zhao, Z.; Jin, M.; Tian, C.; Yang, S. Prediction of seed distribution in rectangular vibrating tray using grey model and artificial neural network. *Biosyst. Eng.* **2018**, *175*, 194–205. [CrossRef]
33. Wang, Y.; Zhang, X.; Ma, G.; Du, X.; Shaheen, N.; Mao, H. Recognition of weeds at asparagus fields using multi-feature fusion and backpropagation neural network. *Int. J. Agric. Biol. Eng.* **2021**, *14*, 190–198. [CrossRef]
34. Guo, Z.; Zou, Y.; Sun, C.; Jayan, H.; Jiang, S.; El-Seedi, H.R.; Zou, X. Nondestructive Determination of Edible Quality and Watercore Degree of Apples by Portable Vis/NIR Transmittance System Combined with CARS-CNN. *J. Food Meas. Charact.* **2024**, *18*, 4058–4073. [CrossRef]
35. Zhang, T.; Zhou, J.; Liu, W.; Yue, R.; Shi, J.; Zhou, C. SN-CNN: A lightweight and accurate line extraction algorithm for seedling navigation in ridge-planted vegetables. *Agriculture* **2024**, *14*, 1446. [CrossRef]
36. Qiu, D.; Guo, T.; Yu, S.; Liu, W.; Li, L.; Sun, Z. Classification of apple color and deformity using machine vision combined with CNN. *Agriculture* **2024**, *14*, 978. [CrossRef]
37. Wang, B.; Deng, J.; Jiang, H. Markov transition field combined with convolutional neural network improved the predictive performance of near-infrared spectroscopy models for determination of aflatoxin B1 in maize. *Foods* **2022**, *11*, 2210. [CrossRef] [PubMed]
38. Pan, Y.; Jin, H.; Gao, J.; Rauf, H.T. Identification of buffalo breeds using self-activated-based improved convolutional neural networks. *Agriculture* **2022**, *12*, 1386. [CrossRef]
39. Zhu, Y.; Ma, Z.; Han, M.; Li, Y.; Xing, L.; Lu, E.; Gao, H. Quantitative damage detection of direct maize kernel harvest based on image processing and BP neural network. *Food Sci. Technol.* **2022**, *42*, e54322. [CrossRef]
40. Xu, J.; Liu, H.; Shen, Y.; Zeng, X.; Zheng, X. Individual nursery trees classification and segmentation using a point cloud-based neural network with dense connection pattern. *Sci. Hortic.* **2024**, *328*, 112902. [CrossRef]
41. Ge, L.; Ren, Z.; Yuan, J. Point-to-point regression PointNet for 3D hand pose estimation. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 475–491. [CrossRef]
42. Chen, X.; Jiang, K.; Zhu, Y.; Wang, X.; Yun, T. Individual tree crown segmentation directly from UAV-borne LiDAR data using the PointNet of deep learning. *Forests* **2021**, *12*, 131. [CrossRef]
43. Wang, Y.; Sun, Y.; Liu, Z.; Sarma, S.E.; Bronstein, M.M.; Solomon, J.M. Dynamic graph CNN for learning on point clouds. *ACM Trans. Graph.* **2019**, *38*, 146. [CrossRef]
44. Guo, Y.; Wang, H.; Hu, Q.; Liu, H.; Liu, L.; Bennamoun, M. Deep learning for 3D point clouds: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *43*, 4338–4364. [CrossRef] [PubMed]
45. Tao, T.; Wei, X. STBNA-YOLOv5: An improved YOLOv5 network for weed detection in rapeseed field. *Agriculture* **2025**, *15*, 22. [CrossRef]

46. Zhu, W.; Sun, J.; Wang, S.; Shen, J.; Yang, K.; Zhou, X. Identifying field crop diseases using transformer-embedded convolutional neural network. *Agriculture* **2022**, *12*, 1083. [CrossRef]
47. Ji, W.; Zhai, K.; Xu, B.; Wu, J. Green apple detection method based on multidimensional feature extraction network model and transformer module. *J. Food Prot.* **2024**, *88*, 100200. [CrossRef] [PubMed]
48. Han, J.; Li, Y.; Liang, J.; Chen, L.; Yu, X.; Chen, L.; Otu, W.; Hafida, W.; Yang, H.; Zhou, C. Based on the umami-transformer machine learning model: Screening of umami peptides from ultrasound-assisted enzymatic hydrolysis of shiitake mushrooms and study on their flavor mechanism. *Food Res. Int.* **2026**, *228*, 118379. [CrossRef] [PubMed]
49. Gao, X.; Gao, J.; Qureshi, W.A. Applications, trends, and challenges of precision weed control technologies based on deep learning and machine vision. *Agronomy* **2025**, *15*, 1954. [CrossRef]
50. Gupta, S.; Girshick, R.; Arbeláez, P.; Malik, J. Learning rich features from RGB-D images for object detection and segmentation. In Proceedings of the European Conference on Computer Vision, Zurich, Switzerland, 6–12 September 2014; pp. 345–360. [CrossRef]
51. Chen, X.; Ma, H.; Wan, J.; Bo, L.; Tian, X. Multi-view 3D object detection network for autonomous driving. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1907–1915. [CrossRef]
52. Feng, D.; Haase-Schütz, C.; Rosenbaum, L.; Hertlein, H.; Gläser, C.; Timm, F. Deep multi-modal object detection and semantic segmentation for autonomous driving: Datasets, methods, and challenges. *IEEE Trans. Intell. Transp. Syst.* **2020**, *22*, 1341–1360. [CrossRef]
53. Guo, Y.; Gao, J.; Mazhar Hussain, T.; Wang, L. Study on the identification of mildew disease of cuttings at the base of mulberry cuttings by aeroponics rapid propagation based on a BP neural network. *Agronomy* **2023**, *13*, 106. [CrossRef]
54. Wang, W.; Li, C.; Xi, Y.; Gu, J.; Zhang, X.; Zhou, M.; Peng, Y. Research progress and development trend of visual detection methods for selective fruit harvesting robots. *Agronomy* **2025**, *15*, 1926. [CrossRef]
55. Elsherbiny, O.; Gao, J.; Guo, Y.; Tunio, M.H.; Mosha, A.H. Fusion of the deep networks for rapid detection of branch-infected aeroponically cultivated mulberries using multimodal traits. *Int. J. Agric. Biol. Eng.* **2025**, *18*, 75–88. [CrossRef]
56. Lefsky, M.A.; Cohen, W.B.; Parker, G.G.; Harding, D.J. Lidar remote sensing for ecosystem studies. *BioScience* **2002**, *52*, 19–30. [CrossRef]
57. Gu, W.; Wang, W.; Wu, W.; Zheng, C.; Lu, X.; Chang, W.; Xiao, P.; Guo, X. 3D reconstruction of wheat plants by integrating point cloud data and virtual design optimization. *Agriculture* **2024**, *14*, 391. [CrossRef]
58. Salvi, J.; Fernandez, S.; Pribanic, T.; Llado, X. A state of the art in structured light patterns for surface profilometry. *Pattern Recognit.* **2010**, *43*, 2666–2680. [CrossRef]
59. Tang, Z.; Jiang, T.; Wang, Y.; Sun, X. LiDAR: A new player in analyzing plant phenotypes. *Trends Plant Sci.* **2024**, *29*, 1383–1384. [CrossRef] [PubMed]
60. Furukawa, Y.; Ponce, J. Accurate, dense, and robust multiview stereopsis. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *32*, 1362–1376. [CrossRef] [PubMed]
61. Snavely, N.; Seitz, S.M.; Szeliski, R. Photo tourism: Exploring photo collections in 3D. *ACM Trans. Graph.* **2006**, *25*, 835–846. [CrossRef]
62. Li, T.; Zhang, Y.; Hu, L.; Zhao, Y.; Cai, Z.; Yu, T.; Zhang, X. Multi-trait phenotypic analysis and biomass estimation of lettuce cultivars based on SFM-MVS. *Agriculture* **2025**, *15*, 1662. [CrossRef]
63. Lin, Y. LiDAR: An important tool for next-generation phenotyping technology of high potential for plant phenomics? Comput. *Electron. Agric.* **2015**, *119*, 61–73. [CrossRef]
64. Ge, X.; Wu, S.; Wen, W.; Shen, F.; Xiao, P.; Lu, X.; Liu, H.; Zhang, M.; Guo, X. LettuceP3D: A tool for analysing 3D phenotypes of individual lettuce plants. *Biosyst. Eng.* **2025**, *251*, 73–88. [CrossRef]
65. Schönberger, J.L.; Frahm, J.M. Structure-from-motion revisited. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 4104–4113. [CrossRef]
66. Bernardini, F.; Mittleman, J.; Rushmeier, H.; Sliva, C.; Taubin, G. The ball-pivoting algorithm for surface reconstruction. *IEEE Trans. Vis. Comput. Graph.* **1999**, *5*, 349–359. [CrossRef]
67. Zhang, J.; Zhang, C.; Xiao, R.; Chi, B. A nonconforming surface mesh generation method by binary tree. *Eng. Anal. Bound. Elem.* **2024**, *166*, 105846. [CrossRef]
68. Song, C.; Pang, Z.; Jing, X.; Xiao, C. Distance field guided L1-median skeleton extraction. *Vis. Comput.* **2018**, *34*, 243–255. [CrossRef]
69. Paturkar, A.; Sen Gupta, G.; Bailey, D. Making use of 3D models for plant physiognomic analysis: A review. *Remote Sens.* **2021**, *13*, 2232. [CrossRef]
70. Paulus, S. Measuring crops in 3D: Using geometry for plant phenotyping. *Plant Methods* **2019**, *15*, 103. [CrossRef] [PubMed]
71. Wang, J.; Zhang, Y.; Gu, R. Research status and prospects on plant canopy structure measurement using visual sensors based on three-dimensional reconstruction. *Agriculture* **2020**, *10*, 462. [CrossRef]
72. Peng, S.; Niemeyer, M.; Mescheder, L.; Pollefeys, M.; Geiger, A. Convolutional occupancy networks. In Proceedings of the European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020; pp. 523–540. [CrossRef]

73. Yang, T.; Zhou, S.; Ye, J.; Xu, A.; Wu, X. A method for 3D reconstruction of trees via SfM guidance and depth estimation. *Comput. Electron. Agric.* **2026**, *247*, 111674. [CrossRef]
74. Qian, M.; Ma, B.; Li, X.; Jiang, H. Three-dimensional reconstruction and experiment of dormant walnut trees based on Kinect V3. *J. Chin. Agric. Mech.* **2023**, *44*, 168. [CrossRef]
75. Ge, J.; Li, H.; Zhang, X.; Li, H. An adaptive plant branch reconstruction method. *J. Syst. Simul.* **2013**, *25*, 855–860.
76. Das Choudhury, S.; Maturu, S.; Samal, A.; Stoerger, V.; Awada, T. Leveraging image analysis to compute 3D plant phenotypes based on voxel-grid plant reconstruction. *Front. Plant Sci.* **2020**, *11*, 521431. [CrossRef] [PubMed]
77. Li, J.; Qi, X.; Nabaei, S.H.; Liu, M.; Chen, D.; Sun, Q.; Zhang, X.; Yin, X.; Li, Z. A survey on 3D reconstruction techniques in plant phenotyping: From classical methods to neural radiance fields (NeRF), 3D Gaussian Splatting (3DGS), and beyond. *Plant Phenomics* **2025**, *2025*, 100137. [CrossRef] [PubMed]
78. Qiu, T.; Moon, J.; Cheng, L.; Kuehn, L.; Jiang, Y. Neural implicit representation-based tree crop plant 3D reconstruction. In Proceedings of the 2023 ASABE Annual International Meeting, Omaha, NE, USA, 9–12 July 2023; p. 1. [CrossRef]
79. Mildenhall, B.; Srinivasan, P.P.; Tancik, M.; Barron, J.; Ramamoorthi, Y.; Ng, R. NeRF: Representing scenes as neural radiance fields for view synthesis. In Proceedings of the European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020; pp. 405–421. [CrossRef]
80. Müller, T.; Evans, A.; Schied, C.; Keller, A. Instant neural graphics primitives with a multiresolution hash encoding. *ACM Trans. Graph.* **2022**, *41*, 102. [CrossRef]
81. Li, W.; Wang, J.; Xu, B.; Zhang, T. Apple grading based on multi-dimensional view processing and deep learning. *Foods* **2023**, *12*, 2117. [CrossRef] [PubMed]
82. Peng, Z.; Zhu, L.; Zhou, R.; Ntakirutimana, T.; Wei, B.; Wang, D. Transforming apple grading: Standards survey and deep learning insights. *Food Bioprocess Technol.* **2025**, *18*, 5954–5969. [CrossRef]
83. Xu, J.; Liu, H.; Shen, Y. Image and point cloud-based neural network models and applications in agricultural nursery plant protection tasks. *Agronomy* **2025**, *15*, 2147. [CrossRef]
84. Park, J.J.; Florence, P.; Straub, J.; Newcombe, R.; Lovegrove, S. DeepSDF: Learning continuous signed distance functions for shape representation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 165–174. [CrossRef]
85. Zhang, J.; Wang, X.; Meng, H.; Huang, J.; Zhang, X.; Zhou, K.; Li, Y.; Peng, H. Semantically guided 3D reconstruction and body weight estimation method for dairy cows. *Agriculture* **2026**, *16*, 182. [CrossRef]
86. Seitz, S.M.; Curless, B.; Diebel, J.; Scharstein, D.; Szeliski, R. A comparison and evaluation of multi-view stereo reconstruction algorithms. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, New York, NY, USA, 17–22 June 2006; Volume 1, pp. 519–528. [CrossRef]
87. Förstner, W.; Wrobel, B. *Photogrammetric Computer Vision*; Springer: Cham, Switzerland, 2016. [CrossRef]
88. Zhang, H.; Wang, L.; Jin, X.; Bian, L.; Ge, Y. High-throughput phenotyping of plant leaf morphological, physiological, and biochemical traits on multiple scales using optical sensing. *Crop J.* **2023**, *11*, 1303–1318. [CrossRef]
89. Fan, H.; Su, H.; Guibas, L.J. A point set generation network for 3D object reconstruction from a single image. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 605–613. [CrossRef]
90. Roshan, T.R.; Jafari, M.; Golami, M.; Kazemi, M. Evaluating geometric measurement accuracy based on 3D model reconstruction of nursery tomato plants by Agisoft photoscan software. *Comput. Electron. Agric.* **2024**, *221*, 109000. [CrossRef]
91. Girshick, R. Fast R-CNN. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 11–18 December 2015; pp. 1440–1448. [CrossRef]
92. Kok, E.; Wang, X.; Chen, C. Obscured tree branches segmentation and 3D reconstruction using deep learning and geometrical constraints. *Comput. Electron. Agric.* **2023**, *210*, 107884. [CrossRef]
93. Newcombe, R.A.; Izadi, S.; Hilliges, O.; Molyneaux, D.; Kim, D.; Davison, A.J. KinectFusion: Real-time dense surface mapping and tracking. In Proceedings of the 2011 IEEE International Symposium on Mixed and Augmented Reality, Basel, Switzerland, 26–29 October 2011; pp. 127–136. [CrossRef]
94. Han, M.; Wang, L.; Xiao, L.; Zhang, H.; Zhang, C.; Xie, X.; Zheng, S.; Dong, J. FuseFPS: Accelerating farthest point sampling with fusing KD-tree construction for point clouds. In *Proceedings of the 2024 29th Asia and South Pacific Design Automation Conference (ASP-DAC)*; IEEE: New York, NY, USA, 2024; pp. 238–243. [CrossRef]
95. Tsaftaris, S.A.; Minervini, M.; Scharr, H. Machine learning for plant phenotyping needs image processing. *Trends Plant Sci.* **2016**, *21*, 989–991. [CrossRef] [PubMed]
96. Knapitsch, A.; Park, J.; Zhou, Q.Y.; Koltun, V. Tanks and Temples: Benchmarking large-scale scene reconstruction. *ACM Trans. Graph.* **2017**, *36*, 78. [CrossRef]
97. Yang, R.; Kalnis, P.; Tung, A.K.H. Similarity evaluation on tree-structured data. In Proceedings of the ACM SIGMOD International Conference on Management of Data, Baltimore, MD, USA, 14–16 June 2005; pp. 754–765. [CrossRef]

98. Saurer, M.; Sahlstedt, E.; Rinne-Garmston, K.T.; Lehmann, M.M.; Oettli, M.; Gessler, A.; Treydte, K. Progress in high-resolution isotope-ratio analysis of tree rings using laser ablation. *Tree Physiol.* **2023**, *43*, 694–705. [CrossRef] [PubMed]
99. Zhang, D.; Huang, S.; Sun, X.; Zou, X.; Battino, M.; Katona, J.; Shen, L. Advances in 2D and 3D machine vision technologies for morphological characterization of granular food products: From laboratory to application. *J. Food Meas. Charact.* **2025**, *19*, 9292–9318. [CrossRef]
100. Lohrey, M.; Maneth, S.; Mennicke, R. Tree structure compression with repair. In *Proceedings of the 2011 Data Compression Conference*; IEEE: New York, NY, USA, 2011; pp. 353–362. [CrossRef]
101. Liu, D.; Wang, P.; Zhang, Z.; Lu, Y.; Wang, B.; Hu, Y. Robust detection of dense small tea shoots across cultivars under occlusion and bud-leaf similarity for intelligent selective harvesting. *Sci. Hortic.* **2025**, *353*, 114499. [CrossRef]
102. Sun, J.; He, X.; Ge, X.; Wu, X.; Shen, J.; Song, Y. Detection of key organs in tomato based on deep migration learning in a complex background. *Agriculture* **2018**, *8*, 196. [CrossRef]
103. Shi, Q.; Pan, Y.; He, B.; Zhu, H.; Liu, D.; Shen, B.; Mao, H. The airflow field characteristics of UAV flight in a greenhouse. *Agriculture* **2021**, *11*, 634. [CrossRef]
104. Sun, Y.; Luo, Y.; Zhang, Q.; Xu, L.; Wang, L.; Zhang, P. Estimation of crop height distribution for mature rice based on a moving surface and 3D point cloud elevation. *Agronomy* **2022**, *12*, 836. [CrossRef]
105. Liu, X.; Wu, C.; Zhang, C.; Wang, L. A memory-tree driven network for multi-view fusion anomaly detection. *Pattern Recognit.* **2026**, *170*, 112106. [CrossRef]
106. Maver, J.; Bajcsy, R. Occlusions as a guide for planning the next view. *IEEE Trans. Pattern Anal. Mach. Intell.* **1993**, *15*, 417–433. [CrossRef]
107. Rosell, J.R.; Sanz, R. A review of methods and applications of the geometric characterization of tree crops using laser scanning. *Sensors* **2012**, *81*, 124–141. [CrossRef]
108. Akpo, H.A.; Atindogbé, G.; Obiakara, M.C.; Adjinanoukon, A.B.; Gbedolo, M.; Lejeune, P.; Fonton, N.H. Image data acquisition for estimating individual trees metrics: Closer is better. *Forests* **2020**, *11*, 121. [CrossRef]
109. Au, O.K.C.; Tai, C.L.; Chu, H.K.; Cohen-Or, D.; Lee, T.Y. Skeleton extraction by mesh contraction. *ACM Trans. Graph.* **2008**, *27*, 44. [CrossRef]
110. Bucksch, A. A practical introduction to skeletons for plant analysis. *Plant Physiol.* **2014**, *2*, 1400005. [CrossRef] [PubMed]
111. Qi, C.R.; Su, H.; Kaichun, M.; Guibas, L.J. PointNet: Deep learning on point sets. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017. [CrossRef]
112. Horn, B.K.P.; Schunck, B.G. Determining optical flow. *Artif. Intell.* **1981**, *17*, 185–203. [CrossRef]
113. Zhou, H.; Jia, W.; Li, Y.; Ou, M. Method for estimateing canopy thickness using ultrasonic sensor technology. *Agriculture* **2021**, *11*, 1011. [CrossRef]
114. Zhou, Q.; Xue, X.; Chen, C.; Cai, C.; Jiao, Y. Canopy deposition characteristics of different orchard pesticide dose models. *Int. J. Agric. Biol. Eng.* **2023**, *16*, 1–6. [CrossRef]
115. Liu, X.; Jia, W.; Ruan, C.; Zhao, D.; Gu, Y.; Chen, W. The recognition of apple fruits in plastic bags based on block classification. *Precis. Agric.* **2018**, *19*, 735–749. [CrossRef]
116. Gao, Z.; Ma, J.; Hu, W.; Wang, K.; Liu, K.; Chen, J.; Wang, T.; Dong, X.; Qiu, B. Wind-induced bending characteristics of crop leaves and their potential applications in air-assisted spray optimization. *Horticulturae* **2025**, *11*, 1002. [CrossRef]
117. Xi, T.; Li, C.; Qiu, W.; Wang, H.; Lv, X.; Han, C.; Ahmad, F. Droplet deposition behavior on a pear leaf surface under wind-induced vibration. *Appl. Eng. Agric.* **2020**, *36*, 913–926. [CrossRef]
118. Shi, Q.; Liu, D.; Mao, H.; Shen, B.; Li, M. Wind-induced response of rice under the action of the downwash flow field of a multi-rotor UAV. *Biosyst. Eng.* **2021**, *203*, 60–69. [CrossRef]
119. Li, Z.; Niklaus, S.; Snavely, N.; Wang, O. Neural scene flow fields for space-time view synthesis. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 6498–6508. [CrossRef]
120. Liu, H.; Zhu, H. Evaluation of a laser scanning sensor in detection of complex-shaped targets for variable-rate sprayer development. *Trans. ASABE* **2016**, *59*, 1181–1192. [CrossRef]
121. Underwood, J.P.; Huang, C.; Whelan, B.; Sukkarieh, S. Mapping almond orchard canopy volume, flowers, fruit and yield using LiDAR and vision sensors. *Comput. Electron. Agric.* **2016**, *130*, 83–96. [CrossRef]
122. Chehreh, B.; Moutinho, A.; Viegas, C. Latest trends on tree classification and segmentation using UAV data—A review of agroforestry applications. *Remote Sens.* **2023**, *15*, 2263. [CrossRef]
123. Tsouros, D.C.; Bibi, S.; Sarigiannidis, P.G. A review on UAV-based applications for precision agriculture. *Information* **2019**, *10*, 349. [CrossRef]
124. Xie, G.; Qi, Y.; Zhang, P.; Yang, C.; Li, Y.; Wang, L. Deep learning-driven 3D reconstruction of internal wood defects from CT images. *J. For. Res.* **2026**, *37*, 80. [CrossRef]

125. Wang, Y.; Zhao, Y.; Li, F.; Wang, J.; Wang, M.; Li, L.; Huang, M.; Wang, Z.; Pan, S. Graph-to-Tree: Topological Decomposition for Self-Supervised Learning. In Proceedings of the ACM Web Conference 2026, Dubai, United Arab Emirates, 29 June–3 July 2026; pp. 891–900. [CrossRef]
126. Luo, Y.; Wei, L.; Xu, L.; Zhang, Q.; Liu, J.; Cai, Q.; Zhang, W. Stereo-vision-based multi-crop harvesting edge detection for precise automatic steering of combine harvester. *Biosyst. Eng.* **2022**, *215*, 115–128. [CrossRef]
127. Zhang, R.; Zhu, H.; Chang, Q.; Mao, Q. A comprehensive review of digital twins technology in agriculture. *Agriculture* **2025**, *15*, 903. [CrossRef]
128. Shao, L.; Gong, J.; Fan, W.; Zhang, Z.; Zhang, M. Cost comparison between digital management and traditional management of cotton fields—Evidence from cotton fields in Xinjiang, China. *Agriculture* **2022**, *12*, 1105. [CrossRef]
129. Verdouw, C.; Tekinerdogan, B.; Beulens, A.; Woldert, S. Digital twins in smart farming. *Agric. Syst.* **2021**, *189*, 103046. [CrossRef]
130. Pylianidis, C.; Osinga, S.; Athanasiadis, I.N. Introducing digital twins to agriculture. *Comput. Electron. Agric.* **2021**, *184*, 105942. [CrossRef]

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