RESEARCH PAPER

A Comprehensive Review of Deep Learning and Computer Vision Techniques for Fruit Quality Assessment and Classification

S. Saleth Shanthi[1*], Dr. M.P. Indra Gandhi[1]

[1]Department of Computer Science, Mother Teresa Women's University, Kodaikanal - 624101, Tamil Nadu, India

[1]Research Scholar (S. Saleth Shanthi) | Email: salethshanthi@gmail.com

[1]Associate Professor (Dr. M.P. Indra Gandhi) | Email: indragandhisridharan@gmail.com

***Corresponding Author:** S. Saleth Shanthi, Research Scholar, Department of Computer Science, Mother Teresa Women's University, Kodaikanal - 624101, Tamil Nadu, India | Email: salethshanthi@gmail.com

Abstract

Fruit image analysis has become an essential research area in precision agriculture, enabling automated fruit classification, quality grading, maturity estimation, disease diagnosis, yield prediction, and intelligent harvesting through advances in computer vision and artificial intelligence (AI). Recent progress in imaging technologies, large-scale benchmark datasets, and deep learning has substantially improved the accuracy, robustness, and scalability of automated fruit analysis systems. This survey presents a comprehensive and systematic review of image processing and AI techniques for fruit image analysis. It covers the complete image analysis pipeline, including image acquisition, preprocessing, enhancement, color space transformation, segmentation, feature extraction, and handcrafted texture descriptors. Furthermore, conventional machine learning algorithms are reviewed alongside state-of-the-art deep learning architectures, including Convolutional Neural Networks (CNNs), Vision Transformers (ViTs), and Vision Mamba models, for applications such as fruit recognition, quality assessment, disease detection, ripeness estimation, and object detection. Publicly available benchmark datasets, evaluation metrics, and performance comparison strategies are also summarized to facilitate objective assessment of existing approaches. In addition, the survey discusses key research challenges, including illumination variations, occlusion, complex orchard environments, limited annotated datasets, computational complexity, and domain adaptation, while highlighting emerging research trends such as self-supervised learning, multimodal learning, explainable AI, lightweight deep networks, foundation models, and edge intelligence for real-time deployment. By providing a unified review of methodologies, datasets, evaluation protocols, challenges, and future research directions, this survey offers a comprehensive reference for researchers and practitioners developing intelligent and scalable fruit image analysis systems for precision agriculture.

Keywords: Fruit Image Analysis, Computer Vision, Deep Learning, Precision Agriculture, Fruit Disease Detection.

How to cite this article: S. Saleth Shanthi, M.P. Indra Gandhi. A Comprehensive Review of Deep Learning and Computer Vision Techniques for Fruit Quality Assessment and Classification. Int J Drug Deliv Technol. 2026;16(79s):994-996. DOI: 10.25258/ijddt.16.79s.171

.Introduction

The rapid advancement of computer vision, image processing, and artificial intelligence (AI) has transformed modern agriculture by enabling automated monitoring, quality assessment, and disease diagnosis of agricultural products. Among agricultural commodities, fruits are economically important due to their nutritional value, market demand, and contribution to global food security. Maintaining fruit quality throughout cultivation, harvesting, storage, and distribution is essential for reducing post-harvest losses and improving productivity. Consequently, automated fruit detection and classification systems have gained significant attention as efficient alternatives to conventional manual inspection [1].

Traditional fruit inspection relies on visual assessment by human experts based on characteristics such as color, shape, size, texture, and surface defects. Although effective to some extent, manual inspection is labor-intensive, time-consuming, subjective, and prone to inconsistencies caused by fatigue and environmental variations. These limitations have motivated the development of machine vision systems capable of providing rapid, objective, and reliable fruit analysis with minimal human intervention [2].

Recent advances in imaging sensors, machine learning (ML), and deep learning (DL) have significantly improved the performance of automated fruit detection systems. Modern AI-based approaches can accurately identify fruit varieties, estimate ripeness, detect diseases, assess quality, and support yield estimation under diverse environmental conditions. Convolutional Neural Networks (CNNs), Vision Transformers (ViTs), and object detection frameworks such as YOLO and Faster R-CNN have demonstrated remarkable accuracy and robustness in fruit recognition tasks, making them valuable tools for precision agriculture, intelligent harvesting, supply chain management, and food quality assurance [3].

This survey presents a comprehensive review of image processing and AI techniques for automated fruit detection and classification. It summarizes conventional image processing methods, ML algorithms, and state-of-the-art DL models, discusses publicly available datasets and evaluation metrics, identifies current research challenges, and highlights future research directions for developing accurate, efficient, and scalable fruit detection systems.

1. Need for Automated Fruit Detection

The increasing demand for high-quality fruit production has highlighted the limitations of manual fruit inspection, which is labor-intensive, time-consuming, and prone to subjective errors. Automated fruit image analysis addresses these challenges by integrating image processing, computer vision, and artificial intelligence to enable rapid and accurate fruit classification, disease detection, ripeness estimation, and quality grading. Recent advances in deep learning have further improved detection accuracy by automatically learning robust visual features from image data. These intelligent systems support precision agriculture by enhancing efficiency, reducing costs, and enabling real-time applications in orchards, food processing, and quality control [4].

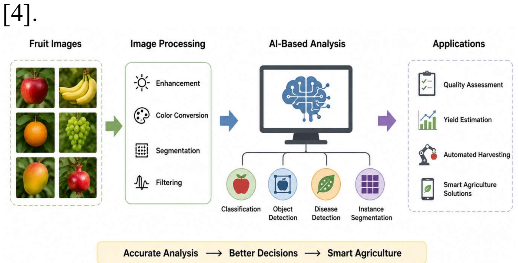

Figure 1. Overview of Fruit Image Analysis and Its Applications

.

1.1. Applications

Automated fruit detection technologies have been successfully adopted across various agricultural and industrial sectors owing to their ability to improve operational efficiency, product quality, and decision-making processes [5]. Major application areas include:

Application Area	Description
Smart Agriculture	Supports fruit detection, yield estimation, ripeness assessment, crop monitoring, and robotic harvesting for precision farming.
Food Industry	Enables automated fruit sorting, grading, quality inspection, and packaging while reducing waste.
Retail Automation	Facilitates image-based fruit recognition for self-checkout systems, improving billing speed and accuracy.
Export Quality Inspection	Ensures compliance with export quality standards through objective grading and defect detection.
Disease Diagnosis	Detects fruit diseases and surface defects at an early stage, enabling timely intervention and minimizing crop losses.

Fruit image analysis integrates computer vision, image processing, and artificial intelligence techniques to automatically identify, classify, detect, and assess fruit characteristics from digital images. It provides the foundation for intelligent agricultural applications such as quality grading, disease diagnosis, yield estimation, and robotic harvesting [6].

2. Computer Vision in Agriculture

Computer vision has become a key technology in modern agriculture by enabling automated analysis of crop and fruit images for monitoring, quality assessment, disease detection, and yield estimation [7]. Combined with artificial intelligence, it supports accurate, efficient, and real-time decision-making, thereby improving agricultural productivity and sustainability.

Table 2. Evolution of Computer Vision Techniques for Fruit Analysis

Generation	Period	Dominant Techniques	Representative Methods	Major Applications	Limitations
Traditional Image Processing	Before 2012	Thresholding, Edge Detection, Morphology	Otsu, Watershed, Canny	Segmentation, Defect Detection	Sensitive to illumination and background
Classical Machine Learning	2012 – 2017	Handcrafted Feature Extraction	GLCM, LBP, HOG + SVM, KNN, RF	Fruit Classification, Disease Detection	Manual feature engineering required
Deep Learning	2017 – 2021	CNN-based Feature Learning	AlexNet, VGG, ResNet, DenseNet, EfficientNet	Classification, Quality Assessment	High computational complexity
Advanced Vision Models	2022 – Present	Transformers & Foundation Models	ViT, Swin Transformer, YOLO v8/11,	Detection, Segmentation, Smart	Large training datasets and compu

		s	SAM, Vision Mamba	Harvesting	tational resources required

Overall, computer vision has transformed fruit analysis from traditional handcrafted feature-based approaches to advanced deep learning and transformer-based models capable of robust and accurate decision-making [8]. These developments have significantly enhanced the automation and reliability of agricultural vision systems, laying the foundation for intelligent fruit monitoring and precision farming applications. Image processing enhances and transforms raw fruit images to improve visual quality and prepare them for accurate analysis. Common operations include resizing, noise removal, contrast enhancement, normalization, and segmentation to facilitate reliable feature extraction and classification.

Figure 1. Workflow of Fruit Image Analysis

A well-designed machine vision system serves as the backbone of automated fruit analysis by integrating image acquisition, preprocessing, feature learning, and intelligent decision-making into a unified framework.

2.1. Fruit Image Acquisition

Fruit image acquisition is the process of capturing high-quality fruit images using digital cameras, smartphones, drones, or imaging sensors for subsequent image processing and analysis [9].

Table 4. Comparison of Imaging Modalities Used in Fruit Analysis

Imaging Modality	Information Captured	Advantages	Limitations	Common Applications
RGB Camera	Color and appearance	Low cost, high availability	Sensitive to illumination	Fruit classification, disease detection
Multispectral Imaging	Selected spectral bands	Improved feature discrimination	Moderate hardware cost	Quality assessment, ripeness estimation
Hyperspectral Imaging	Hundreds of spectral bands	Detects internal defects and early diseases	High computational and acquisition cost	Disease diagnosis, quality inspection
Thermal Imaging	Surface temperature distribution	Identifies physiological stress	Lower spatial resolution	Disease and stress detection
Depth (3D) Camera	Distance and shape information	Robust against background complexity	Limited outdoor performance	Fruit localization, robotic harvesting
UAV/Drone Imaging	Large-area aerial images	Rapid orchard monitoring	Weather-dependent operation	Yield estimation, crop monitoring

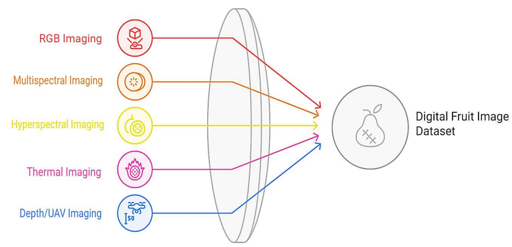

Fig.3 Image Acquisition Techniques

Figure 3 represents common image acquisition techniques employed in automated fruit analysis, illustrating different imaging modalities used to capture fruit characteristics for subsequent preprocessing and intelligent analysis.

2.2. Image Processing Techniques

Image pre-processing enhances the quality of fruit images by reducing noise, correcting illumination variations, and improving contrast to facilitate accurate analysis. It provides reliable input for segmentation, feature extraction, and classification, thereby improving the overall performance of automated fruit image analysis systems [10].

Table 5. Common Image Enhancement Techniques

Technique	Principle	Typical Applications
Median Filter	Replaces pixel with neighborhood median	Fruit surface enhancement
Gaussian Filter	Smooths image using Gaussian kernel	Image denoising
Histogram Equalization (HE)	Redistributes intensity values globally	Fruit recognition
CLAHE	Adaptive local contrast enhancement	Disease detection
Bilateral Filter	Combines spatial and intensity filtering	Defect and lesion enhancement

Adaptive Gamma Correction	Adjusts image brightness adaptively	Outdoor orchard images

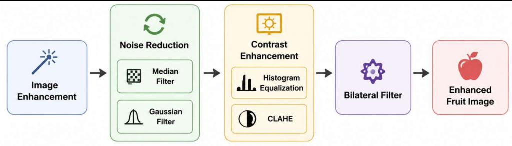

Figure 4. Image Enhancement Techniques Taxonomy

Figure 4 shows the taxonomy of image enhancement techniques commonly employed in fruit analysis for noise removal, contrast improvement, and feature preservation prior to segmentation and classification.

Table 6. Literature Survey on Image Pre-processing Techniques for Fruit Image Analysis

Ref.	Dataset	Preprocessing Techniques
Zhang et al. (2024) [1]	Fruit-360	Image resizing, normalization, data augmentation
Salim et al. (2023) [5]	Fruit-360	Image resizing, normalization, augmentation
Sinha & Dhanalakshmi (2024) [4]	Fruit-360	Contrast enhancement, image normalization, image fusion
Luo et al. (2024) [12]	Peach dataset	Noise removal, ROI extraction, image enhancement
Jrondi et al. (2025) [31]	Citrus fruit dataset	Image resizing (224×224), normalization, augmentation
Khan et al. (2024) [34]	Fruit-360, Fruit Recognition	Image normalization, augmentation, customized pooling
Liu et al. (2025) [35]	Fruit-360, VegFru	Image normalization, resizing

Existing fruit image preprocessing methods are largely static and non-adaptive, limiting robustness and feature preservation under diverse real-world imaging conditions.

2.3. Image Segmentation

Image segmentation partitions a fruit image into meaningful regions to accurately isolate the fruit or diseased areas from the background for further analysis [11].

Table 8. Comparison of Image Segmentation Techniques in Fruit Analysis

Category	Techniques	Common Applications
Threshold-based	Global Threshold, Otsu	Fruit-background separation
Edge-based	Sobel, Canny, Laplacian	Fruit contour extraction
Region-based	Region Growing, Watershed	Fruit and lesion segmentation
Clustering-based	K-means, Fuzzy C-Means	Disease region detection
Graph-based	Graph Cut, GrabCut	Background removal
Deep Learning-based	FCN, U-Net, Mask R-CNN, DeepLabV3+, SAM	Disease segmentation, fruit localization

Traditional segmentation methods remain suitable for controlled laboratory conditions due to their computational efficiency, whereas deep learning-based models have become the preferred choice for real-world fruit analysis because of their superior robustness and segmentation accuracy across diverse environmental conditions [12].

Table 9. Literature Survey on Fruit Image Segmentation Techniques

Reference	Segmentation Category	Dataset	Method Used
Ghazal et al. (2021) [14]	Threshold-based	Citrus fruit images	Otsu Thresholding
Roy et al. (2021) [25]	Edge-based	Apple and Mango images	Canny Edge Detector
Thinh et al. (2021) [15]	Region-based	Fruit-360	Region Growing
Hira & Lande (2022) [11]	Clustering-based	Fruit-360	K-means Clustering
Bhargava & Bansal (2021) [9]	Graph-based	Multiple fruit datasets	Graph Cut
Tunio et al. (2022) [20]	DL-based	Mixed fruit dataset	CNN-based Segmentation
Liang et al. (2024) [24]		Apple dataset	Polar-Net
		Orchard fruit dataset	Occlusion-aware Network

Seo et al. (2024) [22]		Peach dataset	Swin Transformer
Shamrat et al. (2024)		FruitSeg 30	Deep Segmentation Benchmark
Vaghela et al. (2025) [23]		Waxberry dataset	MAWNet
Shang et al. (2026) [18]		Tomato dataset	CNN-based Segmentation
		Orchard RGB-D dataset	DepthCL-Seg
Giménez-Gallego et al. (2024) [19]		On-tree fruit dataset	Instance Segmentation

Existing fruit image segmentation methods have limited robustness and generalization under diverse real-world conditions due to their dependence on specific datasets, extensive annotations, and high computational complexity.

2.4. Feature Extraction

Feature extraction transforms segmented fruit images into representative descriptors that capture discriminative characteristics for classification, quality assessment, and disease diagnosis [13]. Conventional fruit recognition systems rely on handcrafted features derived from color, texture, shape, and statistical properties, with techniques such as Gray-Level Co-occurrence Matrix (GLCM), Local Binary Patterns (LBP), Histogram of Oriented Gradients (HOG), Scale-Invariant Feature Transform (SIFT), and Gabor filters being widely adopted [14].

Table 10. Comparison of Feature Extraction Techniques for Fruit Analysis

Feature Category	Representative Techniques	Features Extracted	Common Applications
Color Features	Color Histogram, Color Moments	Color distribution, intensity	Fruit recognition, ripeness estimation
Texture Features	GLCM, LBP, Gabor Filters	Surface texture, lesion patterns	Disease detection
Shape Features	Hu Moments, Fourier Descriptors	Size, contour, geometric properties	Fruit grading and variety classification
Local Features	SIFT, SURF, ORB	Local keypoints and descriptors	Object matching and detection
Gradient Features	HOG	Edge orientation and gradients	Fruit classification
Deep Features	CNN, ResNet, EfficientNet, Vision Transformer	High-level semantic representations	Classification, detection, segmentation

Handcrafted features remain valuable for lightweight and interpretable machine vision systems, whereas deep feature learning has become the dominant approach due to its superior representation capability and consistent performance across diverse fruit analysis tasks [15].

Table 11. Literature Survey on Feature Extraction Techniques for Fruit Image Analysis

Reference	Dataset	Feature Extraction Method	Major Contribution
Roy et al. (2021) [25]	Fruit-360	Color, Texture and Shape Features	Combined handcrafted descriptors for improved recognition
Salim et al. (2023)		DenseNet-201, Xception	Deep feature extraction using transfer learning
Xue et al. (2023) [26]	Fruit dataset	Hybrid Deep Features	Combined multi-level feature representations
Xiao et al. (2023) [28]	Fruit-360	CNN Feature Learning	Enhanced discriminative feature extraction
Luo et al. (2024) [12]	Peach dataset	Multi-scale Deep Features	Improved fruit quality feature representation
Sinha & Dhanalakshmi (2024) [4]	Fruit-360	Multi-Fused CNN Features	Enhanced feature fusion for classification

Khan et al. (2024) [34]		Enhanced Attention Features (EA-CNN)	Attention-guided feature extraction
Sireesha et al. (2025) [7]	Citrus disease dataset	Vision Transformer Features	Global feature representation for disease recognition
Liu et al. (2025) [35]	Fruit-360, VegFru	Channel Grouping Vision Transformer (CGViT)	Efficient transformer feature extraction
Tang et al. (2025) [25]	Fruit dataset	Multi-scale Attention Features	Improved fine-grained feature learning
Esaki et al. (2025) [27]	Fruit quality dataset	Transformer-based Feature Learning	Robust fruit quality representation
Ergün (2025) [35]	Banana variety dataset	Vision Transformer + SVM Features	High-quality deep feature extraction for variety identification

Existing feature extraction methods have limited generalization and robustness, as they fail to effectively integrate local and global features across diverse real-world fruit datasets [16].

2.5. Classification

2.5.1. Machine Learning-Based Fruit

Machine learning (ML) algorithms have played a significant role in automated fruit classification by utilizing handcrafted features extracted from image processing techniques. Earlier fruit recognition systems commonly combined color, texture, and shape descriptors with supervised classifiers such as Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Decision Trees (DT), Random Forests (RF), Naïve Bayes (NB), and Artificial Neural Networks (ANN) [18]. These algorithms demonstrated satisfactory performance for controlled laboratory datasets due to their low computational requirements and ease of implementation.

Table 14. Comparison of Machine Learning Algorithms for Fruit Classification

Algorithm	Principle	Advantages	Limitations	Common Applications
K-Nearest Neighbors (KNN)	Distance-based classification	Simple implementation, no training phase	High prediction time, sensitive to noise	Fruit variety classification
Support Vector Machine (SVM)	Maximum margin classifier	High accuracy for small datasets	Kernel selection affects performance	Disease detection, fruit recognition
Decision Tree (DT)	Rule-based hierarchical classification	Easy to interpret	Overfitting on complex datasets	Fruit quality grading
Random Forest (RF)	Ensemble of decision trees	High robustness and reduced overfitting	Increased computational cost	Multi-class fruit classification
Naïve Bayes (NB)	Probabilistic classifier	Fast and computationally efficient	Assumes feature independence	Preliminary fruit classification
Artificial Neural Network (ANN)	Multi-layer nonlinear learning	Models complex feature relationships	Requires feature engineering	Fruit recognition and grading

SVM and Random Forest are the most frequently adopted conventional classifiers due to their superior generalization capability. Nevertheless, their dependence on handcrafted features limits scalability, motivating the transition toward end-to-end deep learning models discussed in the following section [19].

Table 15. Literature Survey on Machine Learning Techniques for Fruit Image Analysis

Reference	Dataset	Machine Learning Method	Major Contribution
Ghazal et al. (2021) [14]	Fruit image dataset	Support Vector Machine (SVM)	Improved fruit quality classification
Roy et al. (2021) [25]	Fruit-360	SVM, K-Nearest Neighbor (KNN)	Effective handcrafted feature classification

Bhargava & Bansal (2021) [9]	Multiple fruit datasets	SVM, ANN, Decision Tree	Comprehensive comparison of ML algorithms
Thinh et al. (2021) [15]	Fruit image dataset	Support Vector Machine (SVM)	Accurate fruit recognition with handcrafted features
Hira & Lande (2022) [11]	Fruit-360	Random Forest, KNN	Efficient fruit classification using color and texture features
Mashalkar & Raut (2026)	Fruit image dataset	Support Vector Machine (SVM)	Robust fruit quality grading
Ergün (2025) [35]	Banana variety dataset	Vision Transformer + SVM	Improved banana variety classification using hybrid ML

Existing traditional machine learning methods rely heavily on handcrafted features, limiting their robustness, adaptability, and generalization across diverse fruit datasets and real-world imaging conditions.

2.5.2. Deep Learning for Fruit Analysis

DL has revolutionized fruit analysis by enabling automatic learning of hierarchical visual features directly from raw images, eliminating the dependence on handcrafted descriptors. Convolutional Neural Networks (CNNs) have demonstrated remarkable success in fruit classification, disease detection, quality grading, and ripeness estimation due to their superior feature representation capability [20].

Table 16. DL Techniques Disease Recognition and Classification

CNN Architecture	Key Innovation	Advantages	Limitations	Typical Applications
AlexNet	Deep CNN with ReLU	First successful deep CNN	Large number of parameters	Fruit classification
VGG16/ VGG19	Uniform 3×3 convolutions	High feature extraction capability	Computationally expensive	Disease recognition
GoogLeNet (Inception)	Multi-scale convolution blocks	Efficient feature utilization	Complex architecture	Multi-fruit classification
ResNet	Residual skip connections	Enables very deep networks	Higher inference cost	Disease detection, quality grading
DenseNet	Dense feature connectivity	Improved feature reuse	High memory consumption	Fine-grained classification
EfficientNet	Compound model scaling	High accuracy with fewer parameters	Sensitive to scaling configuration	Fruit classification, mobile deployment
MobileNet	Depthwise separable convolution	Lightweight and fast	Slightly lower accuracy	Edge devices, smartphones
ConvNeXt	Modernized CNN architecture	Strong accuracy with efficient computation	Larger training data preferred	Advanced fruit recognition

Among CNN architectures, EfficientNet, ResNet, and ConvNeXt are the most frequently adopted in recent fruit analysis studies due to their excellent trade-off between classification accuracy, computational efficiency, and scalability across diverse agricultural datasets [21].

Table 17. Literature Survey on CNN-Based Deep Learning Methods for Fruit Image Analysis

Reference	Dataset	CNN Model	Major Contribution
Nithya et al. (2022)	Mango defect dataset	CNN	Automated mango defect detection
Salim et al. (2023)	Fruit-360	DenseNet-201, Xception	Improved fruit classification accuracy
Xue et al. (2023) [26]	Fruit dataset	Hybrid CNN	Enhanced feature representation

Sinha & Dhanalakshmi (2024) [4]	Fruit-360	Multi-Fused CNN	Improved classification through feature fusion
Zhang et al. (2024) [1]	Fruit-360	Lightweight CNN	Efficient multi-label fruit classification
Khan et al. (2024) [34]	Fruit-360	EA-CNN	Attention-guided fruit classification
Shu et al. (2025)	Fruit freshness dataset	CNN	Accurate freshness assessment
Kumar et al. (2026) [8]	Fruit quality dataset	Deep CNN	Automated fruit quality segregation
Rajalingam et al. (2026) [26]	Fruit quality dataset	CNN	Improved post-harvest quality assessment

Existing CNN-based models have limited capability to capture long-range contextual information and often require high computational resources, reducing their efficiency and robustness in complex real-world fruit image analysis.

2.6. Vision Transformers

Transformer-based vision models have emerged as powerful alternatives to CNNs by employing self-attention mechanisms to capture long-range dependencies and global contextual information within images [22]. Unlike convolution-based networks that primarily focus on local receptive fields, Vision Transformers (ViT), Swin Transformer, DeiT, CvT, MaxViT, and Vision Mamba learn comprehensive feature representations that are particularly beneficial for distinguishing visually similar fruit varieties and identifying subtle disease symptoms. Their superior ability to model global relationships has led to increasing adoption in fruit classification, disease diagnosis, and object detection. However, these architectures generally require larger training datasets and greater computational resources than conventional CNNs [23].

Table 18. Comparison of Vision Transformer Architectures

Architecture	Key Innovation	Advantages	Limitations	Common Applications
Vision Transformer (ViT)	Global self-attention	Captures long-range dependencies	Requires large datasets	Fruit classification
DeiT	Data-efficient transformer	Improved performance with limited data	Moderate computational complexity	Disease recognition
Swin Transformer	Shifted window attention	Multi-scale feature learning	More complex architecture	Detection and segmentation
CvT	Convolution + Transformer	Combines local and global features	Higher training complexity	Multi-fruit recognition
MaxViT	Multi-axis attention	Strong multi-scale representation	Computationally intensive	Quality assessment
Vision Mamba	State-space sequence modeling	Linear computational complexity and efficient long-range modeling	Emerging architecture with limited agricultural studies	Large-scale fruit analysis

Transformer-based architectures have become increasingly prominent in fruit analysis because of their superior contextual feature learning. Among them, Swin Transformer and Vision Mamba represent promising directions for future intelligent agricultural vision systems due to their scalability and improved representation capability [24].

Table 19. Literature Survey on Transformer-Based DL Methods for Fruit Image Analysis

Reference	Dataset	Transformer Model	Major Contribution
Sireesha et al. (2025) [7]	Citrus disease dataset	Vision Transformer (ViT)	Improved disease classification
Liu et al. (2025) [35]	Fruit-360, VegFru	Channel Grouping Vision Transformer (CGViT)	Efficient transformer feature learning
Tang et al. (2025)	Fruit dataset	Multi-scale Vision Transformer	Enhanced fine-grained recognition

Esaki et al. (2025) [27]	Fruit quality dataset	Vision Transformer	Robust quality assessment
Ergün (2025) [35]	Banana variety dataset	ViT + SVM	Accurate banana variety identification
Jrondi et al. (2025) [31]	Citrus dataset	Vision Transformer + Phi-3 Mini	Explainable fruit quality assessment

Existing Transformer-based models are computationally intensive and data-hungry, limiting their scalability and generalization across diverse real-world fruit image analysis tasks.

3. Fruit Image Datasets

The availability of large, diverse, and well-annotated datasets is fundamental to developing robust fruit analysis systems. Publicly available datasets facilitate benchmarking, reproducibility, and fair comparison of image processing, machine learning, and deep learning models [25].

- **Fruits-360 Dataset:** Fruits-360 is a widely used benchmark dataset containing 178,812 images across 255 classes of fruits, vegetables, nuts, and seeds. Captured under controlled conditions, it is extensively used for fruit classification, transfer learning, and performance evaluation.
- **PlantVillage Dataset:** PlantVillage contains 54,306 RGB images from 38 disease classes covering 14 crop species. It serves as a standard benchmark for plant disease detection using image processing and deep learning techniques.
- **MinneApple Dataset:** MinneApple comprises 1,000 high-resolution orchard images with over 41,000 annotated apple instances. It is widely used for apple detection, instance segmentation, fruit counting, and yield estimation under real orchard conditions.
- **Banana Classification Dataset:** This dataset includes 13,476 RGB images categorized into four ripeness stages: Unripe, Ripe, Overripe, and Rotten. It is commonly used for banana ripeness classification and fruit quality assessment.
- **FruitNet (Indian Fruits Dataset):** FruitNet contains 19,526 RGB images of six Indian fruit varieties labeled into 18 quality classes. Captured under diverse environmental conditions, it is suitable for fruit classification and quality assessment using deep learning.

Table 24. Publicly Available Benchmark Datasets for Fruit Analysis

Dataset	Fruit/Crop Types	Primary Task	Dataset Size	Annotation Type	Image Environment
Fruits-360	255 fruit, vegetable, nut, and seed classes	Fruit Classification	178,812 images (255 classes)	Class labels	Controlled laboratory
Plant Village (Plant Disease)	14 crop species	Plant Disease Classification	54,306 RGB images (38 classes)	Disease class labels	Controlled laboratory
Minne Apple	Apple	Object Detection, Instance Segmentation, Fruit Counting	41,000+ annotated apple instances	Bounding boxes + Polygon masks	Real orchard
Banana Classification Dataset	Banana (Unripe, Ripe, Overripe, Rotten)	Ripeness Classification	13,476 RGB images (4 classes)	Ripeness labels	Controlled image dataset
FruitNet	Apple, Banana, Guava, Lime, Orange, Pomegranate	Fruit Classification & Quality Assessment	19,526 images (18 classes)	Quality labels	Indoor & outdoor scenes with varying illumination

The selected benchmark datasets collectively encompass a broad spectrum of fruit analysis tasks, including fruit classification (Fruits-360), plant disease recognition (PlantVillage), object detection and instance segmentation (MinneApple), banana ripeness classification (Banana Classification Dataset), and fruit quality assessment (FruitNet: Indian Fruits Dataset with Quality). Their diversity in imaging environments, annotation types, and application domains provides a comprehensive foundation for developing and evaluating robust image processing, machine learning, and deep learning models for intelligent fruit analysis [26].

4. Performance Evaluation Metrics

Performance evaluation metrics provide standardized measures for assessing the effectiveness of fruit analysis models across classification, detection, and segmentation tasks. Classification models are commonly evaluated using Accuracy, Precision, Recall, F1-score, and Receiver Operating Characteristic–Area Under the Curve (ROC-AUC), whereas object detection models additionally employ Intersection over Union (IoU) and mean Average Precision (mAP). For segmentation tasks, Dice Similarity Coefficient (DSC) and IoU are widely used to quantify pixel-level agreement between predicted and ground truth regions. In addition to predictive performance, practical deployment of agricultural vision systems increasingly considers inference time, computational complexity, model size, and memory consumption, particularly for real-time and edge computing applications [28].

Table 25. Performance Evaluation Metrics used for Fruit Segmentation and Classification

Metric	Category	Purpose	Higher/Lower is Better
Accuracy	Classification	Overall prediction correctness	Higher
Precision	Classification	Correctness of positive predictions	Higher
Recall (Sensitivity)	Classification	Ability to detect positive samples	Higher
F1-score	Classification	Balance between precision and recall	Higher
ROC-AUC	Classification	Overall discriminative capability	Higher
IoU (Intersection over Union)	Detection / Segmentation	Overlap between prediction and ground truth	Higher
mAP (mean Average Precision)	Object Detection	Overall detection performance	Higher
Dice Coefficient	Segmentation	Pixel-level similarity	Higher
Inference Time	Efficiency	Time required per prediction	Lower
Model Size	Efficiency	Memory required for deployment	Lower
FLOPs	Computational Complexity	Number of floating-point operations	Lower

Although accuracy remains the most frequently reported metric, recent studies increasingly emphasize mAP, IoU, Dice coefficient, inference time, and model complexity to comprehensively evaluate the effectiveness and deployability of fruit analysis systems [29].

5. Challenges in Fruit Image Analysis

Despite remarkable advances in image processing and artificial intelligence, fruit image analysis continues to face several practical challenges that limit the reliability and large-scale deployment of automated systems.

- **Illumination Variations**: Changes in sunlight, shadows, and weather conditions significantly affect image quality, leading to reduced accuracy in fruit detection, segmentation, and classification [30].
- **Complex Backgrounds and Occlusion**: Leaves, branches, overlapping fruits, and cluttered orchard scenes make it difficult for vision models to accurately distinguish fruits from the background.
- **Intra-class and Inter-class Variability**: Differences in ripeness, disease symptoms, size, and color within the same fruit, along with similarities between different fruit varieties, increase classification complexity.
- **Limited and Imbalanced Datasets**: Many agricultural datasets contain insufficient labeled samples and unequal class distributions, which can result in overfitting and biased model performance.
- **Real-Time Deployment Constraints**: Advanced deep learning models often require high computational resources, limiting their deployment on edge devices, drones, and mobile agricultural platforms [32].
- **Domain Shift Between Laboratory and Field Environments**: Models trained on controlled laboratory images frequently

experience performance degradation when applied to real orchard conditions due to variations in lighting, background, and environmental factors.
- **Model Interpretability and Reliability**: Most deep learning models function as black boxes, making it difficult to explain predictions and gain the trust of farmers and agricultural experts in practical applications [33].

Future fruit image analysis systems should emphasize robustness, computational efficiency, and domain adaptability by integrating advanced vision architectures, lightweight deep learning models, self-supervised learning, and domain adaptation techniques. Addressing these challenges will significantly improve the deployment of intelligent fruit analysis systems in real-world agricultural environments.

6. Conclusion

This survey presented a comprehensive review of image processing and artificial intelligence techniques for automated fruit image analysis, covering the complete workflow from image acquisition and pre-processing to segmentation, feature extraction, classification, and transformer-based vision models. It also summarized benchmark datasets, evaluation metrics, and key challenges affecting real-world deployment. Although deep learning and transformer architectures have significantly improved the accuracy and robustness of fruit analysis systems, challenges such as limited annotated datasets, computational complexity, varying environmental conditions, and model generalization remain unresolved. Future research should focus on lightweight, explainable, and data-efficient AI models capable of robust real-time deployment, thereby supporting the next generation of intelligent and sustainable precision agriculture.

Reference:

1. Zhang, J., Lu, Y., Guo, Y., Wu, C., Liu, H., Yu, Z., & Zhou, J. (2024). Lightweight and Optimized Multi-Label Fruit Image Classification: A Combined Approach of Knowledge Distillation and Image Enhancement. Electronics, 13(16), 3267.
2. Nithya, R., Santhi, B., Manikandan, R., Rahimi, M., & Gandomi, A. H. (2022). Computer Vision System for Mango Fruit Defect Detection Using Deep Convolutional Neural Network. Foods, 11(21), 3483.
3. Y. Shu, J. Zhang, Y. Wang, and Y. Wei, "Fruit Freshness Classification and Detection Based on the ResNet-101 Network and Non-Local Attention Mechanism," Foods, vol. 14, no. 11, Art. no. 1987, 2025, doi: 10.3390/foods14111987.
4. B. B. Sinha and R. Dhanalakshmi, "A Multi-Fused Convolutional Neural Network Model for Fruit Image Classification," International Journal of Cognitive Computing in Engineering, vol. 5, pp. 416–424, 2024, doi: 10.1016/j.ijcce.2024.09.003.
5. F. Salim, F. Saeed, S. Basurra, S. N. Qasem, and T. Al-Hadhrami, "DenseNet-201 and Xception Pre-Trained Deep Learning Models for Fruit Recognition," Electronics, vol. 12, no. 14, Art. no. 3132, 2023, doi: 10.3390/electronics12143132.
6. G. Xue, S. Liu, and Y. Ma, "A Hybrid Deep Learning-Based Fruit Classification Using Attention Model and Convolution Autoencoder," Complex & Intelligent Systems, vol. 9, no. 3, pp. 2209–2219, 2023, doi: 10.1007/s40747-020-00192-x.
7. N. V. Sireesha, G. Rekha, and R. A. Almenweer, "Unassailable Citrus Disease Classification via Multi-Stage Deep Ensemble Learning with Vision Transformers," Scientific Reports, vol. 15, Art. no. 40573, 2025, doi: 10.1038/s41598-025-24416-7.
8. N. Kumar, G. Mishra, S. Kumar, S. Verma, and R. Gavas, "Efficient Real-Time Fruit Quality Segregation Using Deep Image Features," Applied Fruit Science, vol. 68, Art. no. 74, 2026, doi: 10.1007/s10341-026-01782-8.
9. I. Bhargava and A. Bansal, "Fruits and Vegetables Quality Evaluation Using Computer Vision: A Review," Journal of King Saud University – Computer and Information Sciences, vol. 33, no. 3, pp. 243–257, 2021, doi: 10.1016/j.jksuci.2018.06.002.
10. I. R. Hashemi and S. Amiri, "Predicting Fruit Quality with a Hybrid Machine Learning and Image Processing Approach," arXiv preprint arXiv:2606.26165, 2026.
11. S. Hira and S. Lande, "Detection of Fruit Ripeness Using Image Processing," International Journal of Health Sciences, vol. 6, no. S6, pp. 3874–3886, 2022, doi: 10.53730/ijhs. v6nS6.10146.
12. D. Luo, R. Luo, J. Cheng, and X. Liu, "Quality Detection and Grading of Peach Fruit Based on Image Processing Method and Neural Networks in Agricultural Industry," Frontiers in Plant Science, vol. 15, Art. no. 1415095, 2024, doi: 10.3389/fpls.2024.1415095.
13. G. Zhao, L. Wang, J. Zheng, N. Tuerxun, W. Han, and L. Liu, "Optimized Extraction Method of Fruit Planting Distribution Based on Spectral and Radar

Data Fusion of Key Time Phase," Remote Sensing, vol. 15, no. 17, Art. no. 4140, 2023, doi: 10.3390/rs15174140.
14. S. Ghazal, W. S. Qureshi, U. S. Khan, J. Iqbal, N. Rashid, and M. I. Tiwana, "Analysis of Visual Features and Classifiers for Fruit Classification Problem," Computers and Electronics in Agriculture, vol. 187, Art. no. 106267, Aug. 2021, doi: 10.1016/j.compag.2021.106267.
15. N. V. Thinh, A. N. T. Y. Nhi, T. G. Huy, N. H. Khoa, and N. T. Cam, "Fruits Classification by Using Machine Learning—An Experiment Using Popular Approaches on Local Data," in 2021 IEEE International Conference on Machine Learning and Applied Network Technologies (ICMLANT), Soyapango, El Salvador, 2021, pp. 1–6, doi: 10.1109/ICMLANT53170.2021.9690537.
16. P. Zhang and X. Li, "Automatic Fruit Grading System with High Adaptability Using Machine Learning Method," Applied Sciences, vol. 15, no. 22, Art. no. 11866, 2025, doi: 10.3390/app152211866.
17. J. V. Mashalkar and S. D. Raut, "Comparative Study of ML Classifiers for Fruit Detection Using Statistical Features," International Journal of Computer Applications, vol. 187, no. 93, pp. 44–48, Mar. 2026, doi: 10.5120/ijca2026926614.
18. Y. Shang, G. Sun, H. Zhang, X. Li, and Y. Zhang, "DepthCL-Seg: Dual-Stream Feature Fusion for Green Fruit Instance Segmentation Based on Monocular Depth," Agriculture, vol. 16, no. 2, Art. no. 283, 2026, doi: 10.3390/agriculture16020283.
19. J. Giménez-Gallego, J. Martinez-del-Rincon, J. D. González-Teruel, H. Navarro-Hellín, P. J. Navarro, and R. Torres-Sánchez, "On-Tree Fruit Image Segmentation Comparing Mask R-CNN and Vision Transformer Models: Application in a Novel Algorithm for Pixel-Based Fruit Size Estimation," Computers and Electronics in Agriculture, vol. 222, Art. no. 109077, 2024, doi: 10.1016/j.compag.2024.109077.
20. M. H. Tunio, L. Jianping, M. H. F. Butt, I. Memon, and Y. Magsi, "Fruit Detection and Segmentation Using Customized Deep Learning Techniques," in *Proc. 19th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, Chengdu, China, 2022, pp. 151–156, doi: 10.1109/ICCWAMTIP56608.2022.10016600.
21. F. M. J. M. Shamrat, R. Shakil, M. Y. I. Idris, B. Akter, and X. Zhou, "FruitSeg30_Segmentation dataset & mask annotations: A novel dataset for diverse fruit segmentation and classification," Data in Brief, vol. 56, Art. no. 110821, 2024, doi: 10.1016/j.dib.2024.110821.
22. D. Seo, S. K. Lee, J. G. Kim, and I.-S. Oh, "High-Precision Peach Fruit Segmentation under Adverse Conditions Using Swin Transformer," Agriculture, vol. 14, no. 6, Art. no. 903, 2024, doi: 10.3390/agriculture14060903.
23. R. Vaghela, N. Sravya, S. Lal, J. Sarda, A. Thakkar, and S. Patil, "Multi Head Attention Based Deep Learning Framework for Waxberry Fruit Object Segmentation from High Resolution Remote Sensing Images," Scientific Reports, vol. 15, Art. no. 37773, 2025, doi: 10.1038/s41598-025-20073-y.
24. J. Liang, K. Huang, H. Lei, Z. Zhong, Y. Cai, and Z. Jiao, "Occlusion-Aware Fruit Segmentation in Complex Natural Environments Under Shape Prior," Computers and Electronics in Agriculture, vol. 217, Art. no. 108620, 2024, doi: 10.1016/j.compag.2024.108620.
25. Roy, K., Chaudhuri, S.S. & Pramanik, S. Deep learning based real-time Industrial framework for rotten and fresh fruit detection using semantic segmentation. Microsyst Technol 27, 3365–3375 (2021). https://doi.org/10.1007/s00542-020-05123-x
26. K. Rajalingam, S. Srinivasan, M. B. Sudhan, N. Mageshkumar, et al., "Fruit Disease Classification Using Transformer and Convolutional Architectures with Explainable AI," Journal of Experimental Horticulture, vol. 2, Art. no. 100013, 2026, doi: 10.1016/j.jxh.2026.100013.
27. A. Esaki, S. Noma, T. Ban, R. Sultana, and I. Shimizu, "Maturity Classification of Blueberry Fruit Using YOLO and Vision Transformer for Agricultural Assistance," Horticulturae, vol. 11, no. 10, Art. no. 1272, 2025, doi: 10.3390/horticulturae11101272.
28. B. Xiao, M. Nguyen, and W. Q. Yan, "Fruit Ripeness Identification Using Transformers," Applied Intelligence, vol. 53, pp. 22488–22499, 2023, doi: 10.1007/s10489-023-04799-8.
29. J. Tang, Z. Yu, and C. Shao, "Hybrid Attention Transformer Integrated YOLOV8 for Fruit Ripeness Detection,"

Scientific Reports, vol. 15, Art. no. 22652, 2025, doi: 10.1038/s41598-025-04184-0.
30. C. Liu, W. Min, J. Song, Y. Yang, G. Sheng, T. Yao, L. Wang, and S. Jiang, "Channel Grouping Vision Transformer for Lightweight Fruit and Vegetable Recognition," Expert Systems with Applications, 2025, Art. no. 128636, doi: 10.1016/j.eswa.2025.128636.
31. Z. Jrondi, A. Moussaid, and M. Y. Hadi, "Interpretable Citrus Fruit Quality Assessment Using Vision Transformers and Lightweight Large Language Models," AgriEngineering, vol. 7, no. 9, Art. no. 286, 2025, doi: 10.3390/agriengineering7090286.
32. S. Sultana, M. A. M. Tasir, S. M. N. Nobel, M. M. Kabir, and M. F. Mridha, "XAI-FruitNet: An Explainable Deep Model for Accurate Fruit Classification," Journal of Agriculture and Food Research, vol. 18, Art. no. 101474, 2024, doi: 10.1016/j.jafr.2024.101474.
33. S. C. Sugunadevi and B. Uma Maheswari, "Enhanced Fruit Classification Using K-Fold Cross-Validation and Interpretation with Explainable AI (XAI)," in Proc. 2025 6th International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), Tirunelveli, India, 2025, pp. 151–157, doi: 10.1109/ICICV64824.2025.11085564.
34. Z. A. Khan, M. Waqar, K. M. Cheema, A. A. B. Mahmood, Q. Ain, N. I. Chaudhary, A. Alshehri, S. S. Alshamrani, and M. A. Zahoor Raja, "EA-CNN: Enhanced Attention-CNN with Explainable AI for Fruit and Vegetable Classification," Heliyon, vol. 10, no. 23, Art. no. e40820, 2024, doi: 10.1016/j.heliyon.2024.e40820.
35. Ergün, E. High precision banana variety identification using vision transformer based feature extraction and support vector machine. Sci Rep 15, 10366 (2025).