International Journal for Innovative Engineering and Management Research

PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL

www.ijiemr.org

A COMPREHENSIVE REVIEW OF DEEP LEARNING AND COMPUTER VISION TECHNIQUES FOR AUTOMATED FRUIT RIPENESS DETECTION

Chandan Prasad[1)*] , Dr. Shailesh Kumar[2)], Dr. Sunil Damodar Rathod[3)]

[1)]Computer Engineering Department, JJT University, Jhunjhunu, Rajasthan, India, Country

[1)] chandansatyendraprasad@gmail.com

[2)] Computer Science and Engineering, Gopalan College of Engineering and Management,Bengaluru, Karnataka,India

[2)] shaileshkumar1610@gmail.com

[3)] Indira College of Engineering and Management, Savitribai Phule Pune University, Pune,India

[3)] sunil6kr@gmail.com

Abstract

Precise determination of ripeness in fruits is essential for maximizing harvesting choices, optimizing supply chain efficiency, and reducing post-harvest loss. Conventional human-based ripeness evaluation tools are time-consuming, prone to errors, and non-scalable. Advances in deep learning and computer vision have transformed the classification of ripeness via computer-based, non-destructive methods. This review discusses cutting-edge artificial intelligence (AI) models, such as convolutional neural networks (CNNs), vision transformers, and deep hybrid learning architectures, for ripeness detection. The research also explores object detection techniques like Faster R-CNN, YOLOv8, and Mask R-CNN, and sophisticated segmentation models for precision agriculture. Multi-modal AI approaches such as hyperspectral, near-infrared, and thermal imaging have also improved classification accuracy. Self-supervised and few-shot learning methods are also explored as viable solutions for model training with sparse labeled data, enhancing the adaptability of AI-based fruit ripeness detection systems. This extensive review showcases the promise of AI in smart farming applications, with a focus on future research directions such as dataset standardization, real-time inference, and edge AI deployment for large-scale agricultural automation.

Keywords: - Smart Farming, Fruit Quality, Automated Sorting & Grading, Fruit Ripeness, Artificial Intelligence.

1. INTRODUCTION

Fruit ripening is a pivotal phase in the fruit growth cycle when they reach the best texture, flavor, color, aroma, and nutritional quality. Ripening is a process of physiological and biochemical changes that include chlorophyll breakdown, carotenoid and anthocyanin buildup (responsible for color changes), and enzymatic pectin breakdown that softens the fruit. Chemical changes, such as the increase in sugar content, are also part of the process. content (Brix level) and reduction of acidity, which further affect palatability of fruits. Furthermore, climacteric fruits such as bananas, apples, and tomatoes release ethylene, a plant hormone that induces ripening at a faster rate, whereas non-climacteric fruits such as citrus and strawberries follow slow ripening without a drastic ethylene burst. Fruit ripeness needs to be sensed for agriculture, food supply chain, and customer satisfaction. Commercial agriculture requires correct ripeness evaluation optimizes harvest timing, promoting better quality yield and avoiding post-harvest losses. Ripeness determines processing methods, storage requirements, and product quality in the food industry, affecting sustainability and profitability. Ripeness affects freshness, flavor, and nutritional content for consumers, influencing purchase decisions. Inefficient harvesting timing, either early or late, may cause economic losses, food losses, and unfavorable marketability, creating efficient ripeness detection as a requirement in contemporary agriculture.

Fruit ripeness detection has been based in the past on manual inspection, which is human-dependent and thus prone to errors and inconsistencies. Environmental conditions such as sunlight exposure, temperature, humidity, and soil conditions further complicate this process. Fruits such as mangoes, avocados, and melons need invasive techniques for detection of ripeness, which causes damage. Manual inspection is slow, labor-consuming, and expensive, particularly in large-scale farming. With the global shortage of agricultural labor on the rise, human inspection is no longer sustainable. Ripeness detection needs to be automated to enhance accuracy, efficiency, and scalability in agriculture.

Artificial Intelligence (AI) and computer vision are transforming the detection of fruit ripeness beyond the constraints of human inspection. Deep learning, especially Convolutional Neural Networks (CNNs), enables real-time, precise classification of ripeness through color, texture, and shape analysis. AI-based computer vision systems employ image processing algorithms to scan fruit features across various stages of ripening, identifying subtle signs of ripeness. This AI integration within smart farming improves precision farming, streamlines the supply chain, and facilitates robotic fruit-picking technology, decreasing dependency on manpower and enhancing post-harvest processing.

This review explores deep learning and computer vision techniques for automated fruit ripeness detection. It discusses state-of-the-art models like CNNs, Vision Transformers, and hybrid AI architectures, compares different techniques like YOLO, Faster R-CNN, and Mask R-CNN, and explores multi-modal approaches like hyperspectral, near-infrared, and thermal imaging. The

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

study also discusses real-world applications of AI in smart farming, automated sorting, and consumer-facing ripeness detection apps. Key challenges include dataset limitations, real-time processing constraints, and model interpretability. The paper provides valuable insights for researchers and industry professionals.

This article is written in a manner to give a broad overview of AI-based detection of fruit ripeness. Section 2 examines the biological and chemical determinants of ripeness, looking into the changes that occur physiologically while a fruit ripens and their relevance to detection mechanisms. Section 3 goes into the architectures of deep learning and computer vision methodologies, how the AI model categorizes ripeness stages very accurately. Section 4 goes over datasets and benchmarking approaches, emphasizing publicly known datasets and main evaluation metrics adopted in AI-based classification. Section 5 elaborates on real-world uses, such as IoT-based monitoring, AI-enabled sorting systems, and smartphone-based ripeness measurement. Section 6 outlines the main challenges and future research areas, focusing on issues of data scarcity, computational efficiency, and the demand for explainable AI (XAI) in agriculture. Finally, Section 7 provides an overview of conclusions and possible advances, stressing how AI continues to innovate fruit ripeness detection and intelligent agriculture.

2. Biological, Chemical, and Spectral Indicators of Fruit Ripeness

Fruit maturity is based on a series of biological and chemical markers that determine its texture, taste, and general quality as indicated in Figure 1. Biological indicators are color change, softening of texture, aroma development, plant detachment, and flavor enhancement, all of which are phenotypic features that indicate maturation.

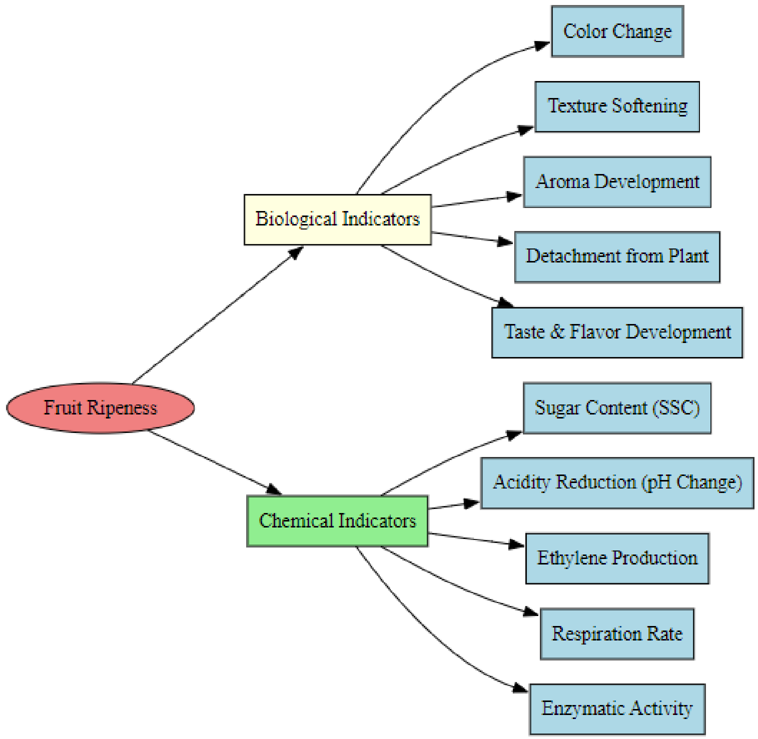

Fig.1: Classification of Fruit Ripeness Indicators: Biological vs. Chemical Factors

Chemical indicators, on the other hand, involve internal compositional changes such as increased sugar content (SSC), acidity reduction (change in pH), ethylene formation, rate of respiration, and enzymatic activity, that control the ripening process at the molecular level. It is important to understand these indicators in agriculture, food processing, and post-harvest management to obtain maximum fruit quality and shelf life.

A. Biological and Visual Ripeness Indicators

Ripeness of the fruit is a vital quality determinant that influences both consumer acceptance and postharvest handling. Multiple biological and visual parameters express the physiological alterations experienced in the course of maturation. Amongst them, color change is among the most noteworthy, wherein the breakdown of chlorophyll causes the formation of carotenoids (yellow-orange pigments) or anthocyanins (red-purple pigments). Texture alteration, spurred by enzymatic degradation, plays a role in the softening and acceptability of the fruit. Size, shape, and odour also develop, with some fruits releasing volatile organic compounds (VOCs) that signal ripeness. While these markers are crucial, the variability of the markers in different species and environmental conditions does not make the detection of ripeness easy through manual measures, warranting the application of AI to create automated tools. Ripening is mainly ethylene-regulated, a plant hormone influencing texture, color, and metabolic process changes in fruit. Research has

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

demonstrated that ETH-treated mangoes have higher malondialdehyde (MDA) concentrations and lower firmness, ETH treatment fostering ripening by activating ACS and ACO enzyme activity [1]. In addition to hormonal regulation, RNA splicing is also important in pigmentation and softening of fruit, illustrating evolutionary flexibility through genetic adjustment [2]. Tomato ripening was investigated in a study, which identified the pectinase enzymes as being responsible for texture alteration, and indicated that total soluble solids (TSS), weight loss, respiration rate, and redness hue all rose with maturity. The results indicate that optimal picking should be two to three weeks after pollination, and keeping it at 16°C will ensure freshness [3]. The same study in melting-flesh peaches singled out polygalacturonase (PG) genes as playing the central regulatory role for pectin solubilization and softening of fruits. Down-regulation of PpPG21 and PpPG22 inhibited ripening, and safeguarding firmness prolonged shelf life [4]. Strategies for managing postharvest ripening have searched for natural treatments such as 1% procyanidin solution (treatment with PA). This procedure efficiently preserves pulp firmness, chlorophyll content, and peel color and minimizes the accumulation of TSS, the production of ethylene, and enzymatic activities [5]. Likewise, tests on sweet cherry cultivars indicate three firmness phases, evidencing a great correlation between skin color, firmness, and mass. They indicate that through the monitoring of a single variable, it becomes possible to provide precise ripeness prediction and enhanced quality control [6]. Through the combination of biochemical markers and AI-based analytical instruments, the automation of ripeness determination is becoming more and more possible, with potential for increased efficiency in harvest management, supply chain logistics, and postharvest preservation.

Volatile organic compounds (VOCs) are important indicators of fruit ripeness and quality. Although artificial scent screening systems replicate the mammalian scent system, with low sensitivity and pattern recognition, their uses are limited for widespread applications. To overcome these limitations, a portable ripeness prediction system integrates colorimetric sensing and deep convolutional neural networks (DCNNs). With the use of gas chromatography-mass spectrometry (GC-MS), it detects individual VOCs from mango, peach, and banana at various stages of ripening. Using 25 gas-sensitive dyes, the system creates distinctive smell fingerprints, processed through DenseNet, with 97.39% validation accuracy and 82.20% test accuracy. Such highly accurate, non-destructive, and affordable system improves real-time ripeness sensing [8]. In addition to fruit ripeness, VOC monitoring is also crucial for agricultural quality control, e.g., mildew detection in stored grains. Though GC-MS is the standard, its prohibitive cost and low speed of detection make it imperative to adopt E-nose and AIdriven sensory analysis, enhancing real-time observation and cost-effectiveness in smart agriculture [9]. AI-driven computer vision methods are transforming fruit maturity testing, substituting manual checks with NMR, NIR, thermal imaging, and hyperspectral scanning as shown in Table 1. The technologies are non-invasive, providing accuracy and scalability in food processing. Combining biosensors with AI makes it even more accurate, allowing automated, real-time maturity detection. Future innovations will need standardization, data privacy practices, and field deployment strategies, making agricultural practices efficient and sustainable [10].

Table 1 Comprehensive Fruit Ripeness Indicators

Indicator	Biochemical Process	Detection Methods	Impact on Ripeness	Influencing Factors	Reference
Color Transformation	Chlorophyll degradation, carotenoid and anthocyanin synthesis	Spectroscopy, image analysis	Indicates transition from immature to ripe state	Light exposure, genetic traits, storage conditions	[1]
Texture Changes	Enzymatic breakdown of cell walls and softening	Texture analyzers, compression tests	Softening signals increased palatability	Enzyme activity, humidity, temperature	[2]
Size & Shape	Cell expansion and morphological evolution	Morphometric analysis, 3D imaging	Growth trends define harvest readiness	Genetic traits, growth environment	[3]
Aroma	Emission of volatile organic compounds (VOCs)	Gas chromatography, e-nose technology	Determines fruit flavor and consumer appeal	Fruit variety, ripening stage, storage duration	[4]
Ethylene Influence	ACS & ACO enzyme activation, ethylene biosynthesis	Ethylene sensors, biochemical assays	Triggers faster ripening and reduced firmness	Ethylene concentration, storage temperature	[5]
RNA Splicing & Gene Control	Alternative splicing influencing	RNA sequencing, gene	Regulates fruit pigmentatio	Genetic mutations, environmental stress	[6]

	gene expression	expression profiling	n and softening		
Pectinase Activity	Pectin degradation affecting fruit firmness	Enzyme activity assays, pectin quantification	Affects shelf life and transportation stability	Pectinase enzyme activity, fruit type	[7]
Polygalacturonase (PG) Expression	Regulation of pectin solubilization and depolymerization	Molecular assays, transcriptomic analysis	Delays or accelerates fruit texture changes	Gene expression regulation, storage conditions	[8]
Procyanidin Treatment Effects	Inhibition of oxidative stress and enzymatic activity	Chemical assays, spectroscopy	Extends storage time while maintaining firmness	Antioxidant content, post-harvest treatments	[9]
Firmness & Sweet Cherry Development	Firmness fluctuation correlated with color and mass	Texture analyzers, colorimeters	Defines maturity phases for better quality control	Genetic variation, environmental adaptation	[10]
Aroma & Gas Emissions	Gas emission variability, sensor accuracy	GC-MS, AI-based scent recognition	Enhances non-invasive ripeness assessment	Gas emission variability, sensor accuracy	[8]
Computer Vision for Maturity Index	AI-driven spectral and thermal analysis	NMR, NIR, thermal imaging, computer vision	Enables automated maturity assessment	Advancements in AI, imaging technology	[10]

B. Chemical and Internal Ripeness Indicators

1) Sugar-to-Acid Ratio (Brix Measurement)

The sugar-to-acid ratio (Brix measurement) is a critical determinant of fruit flavor and quality, directly influencing consumer preference and breeding strategies. In plums (Prunus salicina and Prunus domestica), glucose, fructose, malic acid, and quinic acid were identified as primary contributors to sweetness and acidity. Among these, sucrose plays a dominant role in flesh sweetness, whereas the peel contains 5.5 times more phenolics, contributing to its astringency. Principal Component Analysis (PCA) identified eight key flavor factors, leading to the development of an integrated flavor rating system for plum breeding optimization [11]. Similarly, in 'Xiahui 6' peach trees, reducing the fruit load was found to enhance the sugar-to-acid ratio, resulting in better fruit quality and earlier ripening. A 50% fruit load was determined to be optimal for balancing yield and quality under field conditions [12]. In spine grape (Vitis davidii Foëx), PCA and cluster analysis classified 15 cultivars based on their total sugar, organic acid, and phenolic content, which ranged from 81.80 to 154.89 mg/g, 8.02 to 15.48 mg/g, and 5.58 to 20.12 mg/g, respectively. The 'Red Xiangzhenzhu' cultivar exhibited the highest quality, whereas 'Hongjiangci10' and 'Ziluolan' ranked lowest. Further, cluster analysis grouped cultivars into three quality categories, providing a structured evaluation system for grape breeding and utilization [13].

2) Ethylene Production in Climacteric Fruits

Climacteric fruits undergo ethylene-mediated ripening, which poses challenges in shelf-life extension and quality preservation. Ethylene, a crucial plant hormone, triggers increased respiration rates, accelerating ripening in bananas and tomatoes, while non-climacteric fruits (e.g., oranges) remain unaffected [14]. Various postharvest strategies have been developed to delay ethylene-induced ripening, including ethylene inhibitors, adsorbents, and catalytic oxidation-based scavengers. These strategies influence shelf life, sensory attributes, and volatile compound retention, ensuring improved postharvest quality. A study on open and closed storage environments found that bananas and tomatoes deteriorated within a week, regardless of the storage conditions, confirming their climacteric nature. However, tomatoes stored in a closed environment exhibited greater longevity. Additionally, chemical agents such as carbide were examined for their role in accelerating orange ripening, where higher concentrations (10g) resulted in faster yellowing. Bananas were identified as significant ethylene producers, influencing the ripening of nearby fruits. These findings highlight the crucial role of ethylene in postharvest fruit management and the importance of optimized storage techniques and ethylene control strategies [15].

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

3) pH Variation Across Ripening Stages

Advancements in food science and technology have led to the development of biopolymer-incorporated organic dye indicators for monitoring fruit ripeness. In this study, methylcellulose films containing pH-sensitive dye-based indicators were prepared to detect CO_2 levels inside fruit packaging. As the fruit ripened, the metabolic activity released CO_2, causing pH variations, which were reflected in the color changes of the indicator labels. In 'Nam Dok Mai Si Thong' mangoes, this system successfully correlated CO_2 levels with ripening stages at different storage temperatures. Over time, the indicator labels changed color from blue (unripe), to green (half-ripe), and then to yellow (fully ripe), demonstrating a real-time monitoring system for fruit ripeness. Additionally, firmness decreased from 44.54 to 2.01 N, and titratable acidity (TA) reduced from 2.84% to 0.21%, while soluble solid content (SSC) increased from 10.70% to 18.26%, confirming fruit ripening trends [16]. To better understand the relationship between different ripening indicators, Figure 2 presents an integrated analysis of the sugar-to-acid ratio, ethylene production, pH variation, and firmness retention in various fruits.

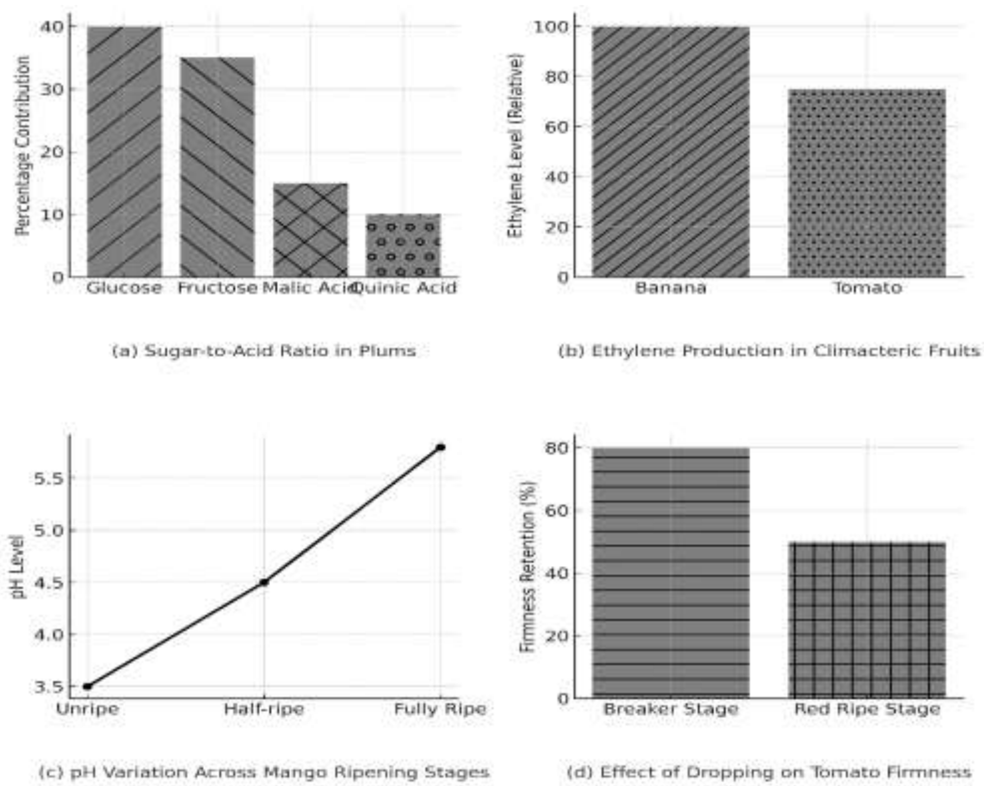

Fig. 2: Multi-Factor Analysis of Fruit Ripening Indicators- (a) Sugar-to-Acid Ratio in Plums, (b) Ethylene Production in Climacteric Fruits, (c) pH Variation Across Mango Ripening Stages, (d) Effect of Dropping on Tomato Firmness

Figure 2 is a detailed breakdown of fruit ripening markers, showing major biochemical and physical alterations during to be present in the cells. Subfigure (a) shows the ratio of sugar to acid in plums, illustrating the percentage to the total sweetness and acidity, which are important for flavor creation and breeding purposes. Subfigure (b) plots the levels of ethylene production in climacteric fruits, such as bananas and tomatoes, highlighting their function in hastening ripening and postharvest deterioration, with bananas particularly affecting the ripening of surrounding fruits, highlighting the requirement for controlled storage conditions. Subfigure (c) depicts pH change throughout mango ripening stages, indicating a progressive rise in pH levels when mangoes go from unripe to completely ripe, paralleling a decrease in acidity and increase in sweetness, which is pivotal for the setting of harvest times and storage conditions. Subfigure (d) looks at the influence of mechanical damage (dropping) on tomato firmness, showing that retention of firmness greatly decreases as tomatoes ripen, highlighting the necessity for gentle handling to reduce damage, retain shelf life, and keep fruit quality. Figure 2 collectively presents an integrated visual overview of biochemical and physical ripening markers, aiding postharvest management, breeding programs, and fruit quality evaluation for commercial production. Novel analytical methods have been investigated in recent studies to precisely identify ripening stages. Electronic Nose (e-nose), ATR-FTIR Spectroscopy, and Image Analysis (IA) were used to differentiate between half-red and fully red strawberries (cv Sabrosa, commercially named Candonga). Principal Component Analysis (PCA) of e-nose, ATR-FTIR, and IA data showed different clustering patterns, proving the effectiveness of these non-destructive analytical tools [17]. Likewise, a novel aggregative index (AQI) for grape quality evaluation was established using visible-near-infrared (Vis-NIR) spectroscopy and chemometric methods. PLSR, SVR, and CNN models were used for predictive analysis, with high accuracy ($Rp^2 = 0.972$ for Cabernet Sauvignon and $Rp^2 = 0.989$ for Muscat Kyoho grapes). These are efficient, non-destructive procedures for checking stages of maturity in fruit and scheduling harvest optimally, most especially in winemaking industry [18]. Tomatoes are perishable crops that suffer mechanical injury during handling and picking, including falling off and bruising. Research measuring Vanessa F1 Hybrid tomatoes of varying stages of ripeness (Breaker (II) and Red Ripe (VI)) indicated that fallen tomatoes had more internal bruising and quicker decay. Firmness, total soluble solids (TSS), titratable acidity, respiration rate, vitamin C, and mineral contents (K and Ca) decreased markedly over 15 days of storage, especially at the Red Ripe stage. Storage at 5°C enhanced fruit retention and alleviated damage severity. The present study points out the necessity of proper postharvest handling methods to prevent mechanical damage and improve tomato shelf-life and marketability [19].

3. Non-Destructive Spectral Techniques

a. Near-Infrared (NIR) and Hyperspectral Imaging.

Near-Infrared (NIR) Hyperspectral Imaging (HSI) is a potentially non-destructive method for estimating the protein content and water content in some agricultural commodities. In soybean seeds, NIR-HSI was used to estimate protein content in 1491 seed samples of three types with low, medium, and high protein concentrations. Partial least square regression (PLSR) was used to construct a calibration model with 70% spectral information, calibrated against the remaining 30%, obtaining an R^2 of 0.92 and an RMSE of 1.08%, validating its ability for fast assessment in processing lines [20]. Likewise, NIR-HSI was utilized to estimate potato flour noodles for protein content. The optimized PLSR model using orthogonal signal correction (OSC) and competitive adaptive reweighted sampling (CARS) showed good predictive accuracy with an R^2 of 0.9606 and 0.8925 in the calibration and prediction sets, respectively. The visualization of protein distribution also added to its power as a non-destructive analytical technique [21]. For the prediction of moisture content (MC) in peanut kernels, hyperspectral data in the 900–1700 nm range were used, along with PLSR modeling. Optimized regression model yielded R^2 values of 0.9357, 0.9133, and 0.9445 for calibration, validation, and prediction, respectively, showing sound performance. A map of moisture content distribution was obtained through pixel-wise hyperspectral analysis, highlighting the capabilities of NIR-HSI in the estimation of moisture in food and agriculture applications [22].

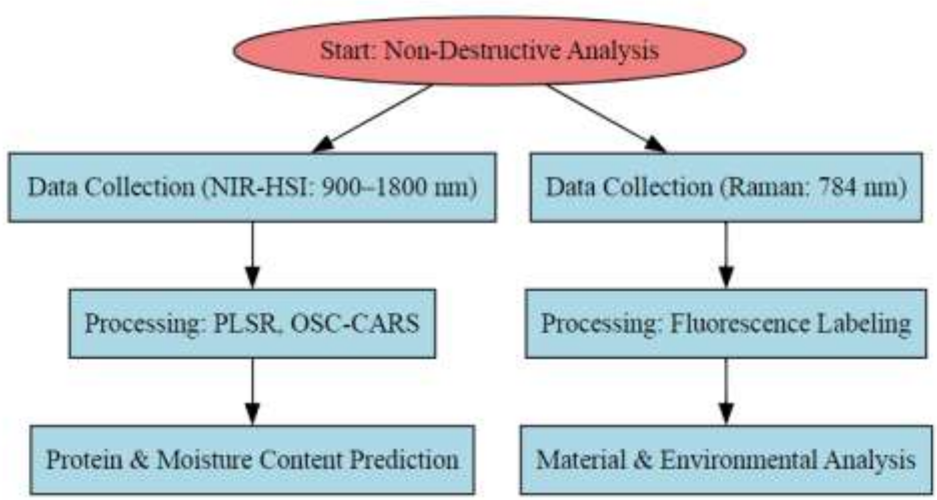

Fig.3: Workflow of Non-Destructive Imaging and Spectroscopy Techniques

Figure 3 shows the workflow of non-destructive imaging and spectroscopy methods, with emphasis on Near-Infrared Hyperspectral Imaging (NIR-HSI) and Raman Spectroscopy. The workflow is separated into two different paths, each designed for particular uses in food quality analysis, material recognition, and environmental monitoring. In the NIR-HIS path, data acquisition is done in the 900–1800 nm range, recording hyperspectral information from agricultural samples. This spectral information is processed using sophisticated processing with Partial Least Squares Regression (PLSR), and optimization methods like Orthogonal Signal Correction (OSC) and Competitive Adaptive Reweighted Sampling (CARS). The data is then refined for protein and moisture content prediction, which is crucial for determining the quality of food items like soybeans, potato flour noodles, and peanut kernels. Alternatively, the Raman Spectroscopy route entails data acquisition at 784 nm, focusing on material composition. The resulting data is additionally optimized by using fluorescence labeling to reduce background noise and increase detection precision. This analyzed data is then utilized in diverse material and environmental analyses such as identifying microplastics in seawater, verifying tea adulteration, and conservation research on temporal paper degradation. Generally, Figure 3 is an organized depiction of how non-destructive imaging and spectroscopy methods are employed for prompt and precise analysis. These methodologies provide very high precision and efficiency and are therefore irreplaceable in scientific research, food safety, environmental monitoring, and material conservation

b. Raman Spectroscopy and Fluorescence Imaging.

Raman spectroscopy and fluorescence imaging have been used extensively in food authenticity and environmental monitoring. For Tieguanyin tea, a fluorescence hyperspectral technique was employed to identify adulteration by Benshan tea at varying concentrations (0%–50%). Through the use of SG-CARS-SVM, the 2-class model reached 100% accuracy, and the 6-class model achieved 94.27% accuracy, validating the method's capacity for detecting and quantifying adulteration effectively [23]. Micro-Raman and luminescence spectroscopy were used integrated with morphological analysis to evaluate degradation patterns in books dating from 1873 to 2021. Highresolution Raman and fluorescence mapping allowed for micron-scale imaging of aging markers, making it particularly useful for analyzing ancient and fragile materials without causing damage [24]. Raman spectroscopy was also used for monitoring marine pollution, particularly detection of microplastics in seawater. Confocal Raman spectroscopy based on fluorescent labeling allowed quick microplastic (PE, PP, PS) identification in the 60–500 μm size range, enhancing detection efficiency and accuracy. Dual-wavelength laser excitation (784/785 nm) and differential Raman spectroscopy effectively eliminated fluorescence interference, making efficient microplastic screening possible in seawater. The technique is beneficial for marine pollution control and ecological risk measurement [25]. Such sophisticated spectroscopic and imaging technologies offer quick, nondestructive, high-accuracy measurements of food quality determination, materials preservation, and monitoring of environmental changes, enabling better efficiency and accuracy in many sectors.

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

4. Challenges in Fruit Ripeness Assessment and Monitoring

a. Variability in Biological Indicators Across Fruit Types

Fruit is an important global agricultural produce that feeds millions. The fruit produce supply chain necessitates high standards of quality tests to ascertain freshness, flavor, and safety. One of the key factors for determining fruit quality is its level of ripening, which was classically sorted manually by subject matter experts in the field. This time-consuming and human error-prone methodology requires automation for improved efficiency and accuracy. There have been new developments in machine learning (ML) and deep learning have greatly enhanced machine-based fruit ripeness classification. In contrast to conventional feature engineering methods, deep learning models are capable of processing raw data without the necessity for intricate, crop-specific engineered features. Computer vision-based classification methods using image descriptors and spectral analysis have been investigated in several studies to improve fruit ripeness assessment [26]. The use of computer vision methods has transformed the food processing sector, substituting manual maturity evaluation with automated systems. Computational breakthroughs like Nuclear Magnetic Resonance (NMR), Near-Infrared Spectroscopy (NIR), thermal imaging, and hyperspectral imaging have been widely used to ascertain fruit and vegetable maturity indices. These non-destruction methods enhance precision and effectiveness, while the incorporation of biosensors and artificial intelligence (AI) further enhances maturity evaluation processes. With advancements in the field, cooperation between specialists in Diverse fields will become critical for standardization, considerations of data privacy, and global adoption of AI driven maturity assessments [27].

b. Challenges in Automated Harvesting

In recent decades, intelligent fruit harvesting robots have been developed to bridge the gap between food demand and labor shortages. However, their commercial adoption remains limited, primarily due to technical challenges in system performance, visual perception, and fruit detachment mechanisms. Studies analyzing existing harvesting robots have highlighted limitations in Table 2, shows adaptability issues to orchard environments and the efficiency of robotic grasping mechanisms. Future research directions emphasize the need for improved robotic vision, dexterous manipulation, and AI-powered decision-making to enhance the effectiveness of automated harvesting systems [28]. Traditional ripeness estimation relies on manual sampling and chemical analyses, which are time-consuming, costly, and destructive. The advent of machine vision techniques has introduced faster, non-invasive, and cost-effective methods for large-scale ripeness assessment. While these methods have been widely applied, particularly in grape ripeness evaluation, further advancements are needed to enhance real-time, in-field maturity monitoring. Recent studies have demonstrated the potential of machine vision-integrated grape harvesting robots, capable of on-the-spot ripeness evaluation to optimize harvesting efficiency [29]. Beyond agriculture, AI and ML have found significant applications in biological treatment technologies such as anaerobic digestion, composting, and insect farming. The complexity of biological treatment processes introduces challenges in process stability and efficiency, which ML models address by enabling real-time monitoring, predictive modeling, and optimization of treatment parameters. Studies indicate that artificial neural networks, tree-based models, support vector machines, and genetic algorithms have shown strong performance in biological treatment prediction and process enhancement [30]. In commercial fruit ripening, AI-driven models have been explored to enhance the uniformity and quality of large-scale ripening processes. A study investigating banana ripening in refrigerated marine containers demonstrated that ML-based predictive models could optimize peel color uniformity and pulp temperature by controlling CO_2 and O_2 gas concentrations. These data-driven approaches significantly improved process monitoring and cost-effectiveness, demonstrating strong correlations between atmospheric gas levels and ripening consistency. For the first time, ML algorithms were used to predict oxygen levels based on atmospheric conditions, providing an alternative to continuous monitoring through direct gas measurements. The application of Long Short-Term Memory (LSTM) regression resulted in low root-mean-square errors (0.033 and 0.202) and high R^2 values (0.999 and 0.959), proving its robustness in banana ripening prediction [31].

table 2 Challenges in Ripeness Indicators

Challenge Category	Specific Challenge	Key Issues	Proposed Solutions	Applications & Impact	References
Variability in Biological Indicators	Traditional ripeness classification is labor-intensive and prone to errors.	Dependence on human expertise, inconsistency, inefficiency in large-scale farms.	AI and deep learning models for automated classification.	Enhances fruit quality control in food supply chains, reduces human error.	[26]
Advancements in Computational Techniques	Automating maturity index assessment using NMR, NIR, and machine vision.	Limited accuracy, high cost, requirement of standardized imaging techniques.	Integrating biosensors and AI to improve real-time analysis.	Faster, cost-effective ripeness detection in food processing and agriculture.	[27]

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

Limitations in Automated Harvesting	Intelligent fruit-picking robots are underdeveloped for commercial use.	Performance limitations, environmental adaptability, fruit detachment complexity.	AI-powered robotic vision, improved grasping mechanisms.	Addresses labor shortages, increases efficiency in fruit harvesting.	[28]
Limitations of Chemical Assessments	Ripeness estimation via chemical analysis is time-consuming and destructive.	Requires manual sampling, expensive testing, inconsistency in large-scale monitoring.	Machine vision techniques for non-destructive analysis.	Accelerates decision-making in vineyards, improves efficiency in large farms.	[29]
Biological Treatment and AI Integration	AI-enhanced biological treatment of organic waste for sustainability.	Complexity of biological processes, inconsistent outputs, environmental impact.	Machine learning models for real-time monitoring and optimization.	Enhances waste management, increases renewable energy production.	[30]
AI in Large-Scale Ripening	Predictive AI models for banana ripening in refrigerated containers.	Lack of automation in ripening monitoring, difficulty in achieving uniformity.	AI-driven atmospheric control, gas concentration monitoring.	Optimizes fruit preservation in logistics, reduces post-harvest losses.	[31]

5. Advances in Deep Learning for Fruit Ripeness Detection

a. Overview of Deep Learning Architectures

Deep learning architectures have significantly transformed image classification, natural language processing, and predictive analytics. Convolutional Neural Networks (CNNs), Vision

Transformers (ViTs), and Hybrid Deep Learning models are among the most prominent advancements. Each of these architectures has specific strengths, making them suitable for different applications such as medical diagnosis, agriculture, remote sensing, sentiment analysis, and disaster prediction.

i. Convolutional Neural Networks (CNNs)

CNN architectures such as ResNet, EfficientNet, and MobileNet have become the foundation of image classification across multiple sectors. These networks extract hierarchical features, enabling robust medical, agricultural, and industrial applications. CNNs have been particularly effective in medical diagnosis, where a hybrid model combining EfficientNet-B0, EfficientNet-B2, and ResNet50 achieved 99.14% accuracy in skin disease classification using the Kaggle Skin Diseases Image Dataset (27,153 images) [33]. This model leveraged CNN-based feature extraction and fusion mechanisms to enhance classification precision, demonstrating strong potential for automated dermatological diagnosis. In agriculture, CNNs play a crucial role in disease detection and classification. A comparative study tested six CNN architectures (DenseNet121, InceptionV3, MobileNetV2, ResNeXt101, ResNet152V, and SE-ResNeXt101) on a dataset of nine rice diseases in Bangladesh. The study also introduced transfer learning and an ensemble model (DEX: DenseNet121, EfficientNetB7, Xception), where the ensemble approach achieved 98% accuracy, significantly outperforming individual models. Additionally, transfer learning increased classification accuracy by 17%, making CNNs highly effective in real-time agricultural disease detection [34]. CNN-based models continue to evolve, improving classification accuracy, feature extraction, and computational efficiency. Their applications in medical imaging, precision farming, and industrial automation highlight their ability to process large datasets efficiently, making them indispensable in various fields [32].

ii. Transformer Models (Vision Transformers, Swin Transformer)

While CNNs have reigned supreme, Vision Transformers (ViTs) are starting to prove themselves as strong contenders for computer vision tasks. While CNNs rely on local spatial hierarchies, ViTs employ self-attention mechanisms to examine global interactions in images. Yet, transformers tend to be hungry for large-scale training sets and are weak at local feature extraction, resulting in the emergence of Hybrid Vision Transformers (CNN-Transformer models). Hybrid architectures merge CNN-based local feature extraction with transformer-based global attention mechanisms greatly enhancing performance on image classification and object detection [35].

The medical community has also started to embrace transformers since they have been shown to perform at or above CNN levels in many medical imaging tasks. Nonetheless, transformers require large labeled datasets, which complicate their use in specialized medical imaging. One of the solutions suggested is self-supervised learning via pretraining on large-scale, unlabeled medical datasets with subsequent fine-tuning for individual imaging tasks. This method has the ability to transcend data sparsity, increase model generalization, and enhance diagnostic accuracy [36]. Transformers have attracted attention in remote sensing, with more than 60 studies implementing ViTs in Very High-Resolution (VHR) imagery, hyperspectral processing, and synthetic aperture radar (SAR) imagery. These models are capable of more advanced long-range spatial feature learning, and as such, are beneficial for earth observation, geospatial intelligence, and machine vision-based automated satellite image interpretation [37]. The continuous innovation of transformers, especially hybridization with CNNs, provides evidence of increasing influence in all medical imaging, remote sensing, and industrial vision tasks. As computational performance continues to improve and pretraining methods improve, transformers are likely to take a dominant position in deep learning-based vision tasks.

iii. Hybrid Deep Learning Models (CNN-RNN, Attention Mechanisms)

Hybrid deep learning models integrating CNNs, Recurrent Neural Networks (RNNs), and attention have been designed to deal with sequential data, deepen predictive analytics, and enhance NLP-based applications. The models are particularly adept at sentiment analysis, text classification, multilingual processing, and disaster prediction. For sentiment analysis, a hybrid RecogNet-LSTM + CNN model incorporating attention mechanisms revealed higher performance in aspect-based sentiment classification. By combining explicit knowledge from external databases, the model greatly enhanced aspect categorization accuracy, providing useful insights for opinion mining and consumer feedback analysis [38]. Hybrid CNN-RNN models also benefit text classification. Two hybrid models, CBAO (Convolutional Bi-LSTM with Attention) and CABO (Convolutional Attention Mechanism with BiLSTM), were evaluated on several datasets. The CBAO model attained 92.72% accuracy on the IMDB dataset, outperforming traditional methods. These models exhibited strong learning ability in natural language processing, document classification, and entity recognition [39]. In multilingual text recognition, a CNN-RNN hybrid model with an attention mechanism was especially proposed for Arabic image text recognition. Due to the complexity of Arabic script, involving variance in font, orientation, and segmentation, this method effectively enhanced text extraction and recognition accuracy. With the integration of sequential feature learning with attention-based processing, the model significantly performed better than conventional CNN-only architectures [40]. Hybrid models have also proved effective in earthquake forecasting, where sequential dependencies and long-term learning of features are paramount. A CNN-BiLSTM model, with attention mechanism and zero-order hold (ZOH) preprocessing, was used on earthquake data from nine Chinese regions. The model was able to accurately predict earthquake magnitude and frequency and show better forecasting ability compared to traditional statistical approaches [41]. These combined architectures increase pattern recognition, predictive analysis, and real-time decision-making in a variety of disciplines. Their

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

combination with attention mechanisms and recurrent processing guarantees enhanced context-awareness and sequential modeling, which makes them crucial in natural language processing, geophysics, and disaster prediction.

b. Object Detection & Image Segmentation for Ripeness

Object detection and image segmentation are central in precision agriculture, especially for computerized fruit harvest, ripeness grade, and quality inspection. Top deep learning networks like YOLOv8, Faster RCNN, and Mask R-CNN have all greatly improved on precision, recall, and the speed of real-time inference and are therefore are crucial in applications of smart farming. Such models enable automated determination of stages of fruit maturity, to streamline harvesting schedules, and enhance farm productivity.

i. Faster R-CNN, YOLOv8, and Mask R-CNN for Instance Segmentation

Instance segmentation is essential to accurately identify individual fruits and tree structures to enable robotic harvesting and precision pruning. YOLOv8 and Mask R-CNN were compared in a study on two orchard datasets. The first dataset, taken during the winter season, consisted of images of apple tree branches and trunks, while the second dataset, taken during the early growing season, included apple tree canopies with immature green fruit (fruitlets). The results indicated that YOLOv8 performed much better than Mask R-CNN. For the first dataset, YOLOv8 had 0.90 precision and 0.95 recall, as opposed to 0.81 for both measures by Mask R-CNN. In the second dataset, YOLOv8 had 0.93 precision and 0.97 recall, better than Mask R-CNN's 0.85 precision and 0.88 recall. Furthermore, YOLOv8 showed quicker inference times—10.9 ms for multi-class segmentation and 7.8 ms for single-class segmentation, whereas 15.6 ms and 12.8 ms are achieved by Mask R-CNN, respectively. The results in Figure 4 indicate that YOLOv8 is more appropriate for real-time orchard operation, especially robotic harvesting and thinning of fruit [42].

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

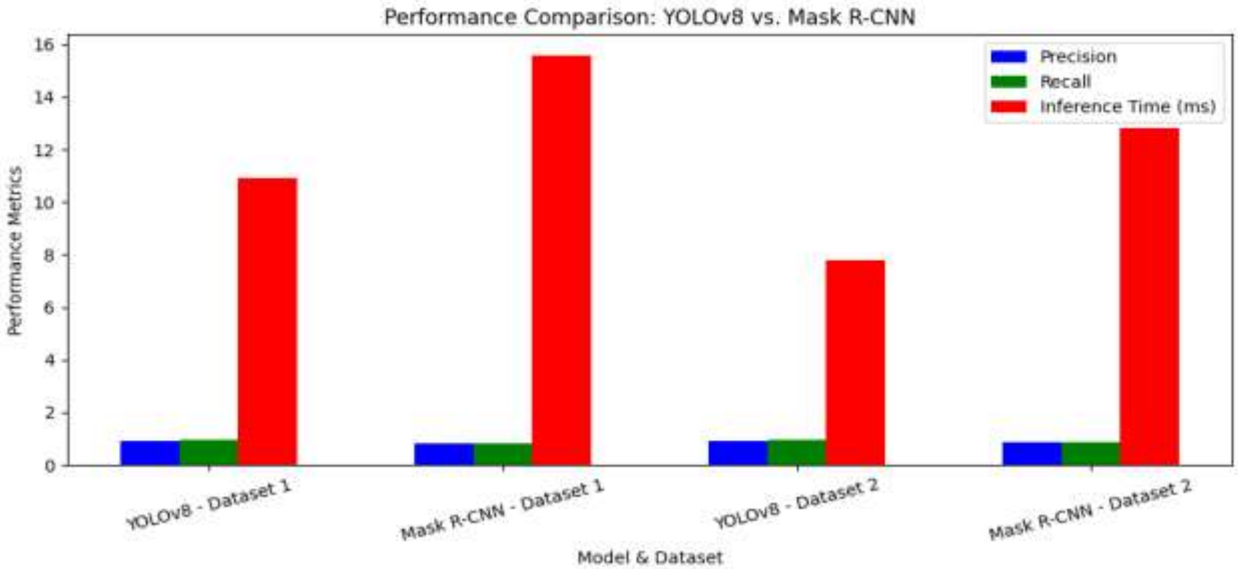

Fig 4: Performance Comparison Between YOLOv8 and Mask R-CNN

ii. Deep Learning for Automated Fruit Identification

Fruit recognition is critical in international agriculture and food processing but is time-consuming and susceptible to human error with traditional manual categorization. Automation with deep learning enhances sorting, grading, and quality inspection. Faster R-CNN and YOLOv8 were tested for recognizing five fruit species: Apple, Cashew Apple, Banana, Mango, and Orange, and results are shown in Figure 5. Although both models had high classification accuracy, Faster R-CNN had greater Mean Average Precision (mAP), while YOLOv8 performed better in real-time processing and recorded higher precision for individual fruits, particularly Cashew Apple and Banana. The results identify that Faster R-CNN is best suited for high-accuracy classification, whereas YOLOv8 is more suitable for real-time sorting purposes in food processing sectors [43].

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

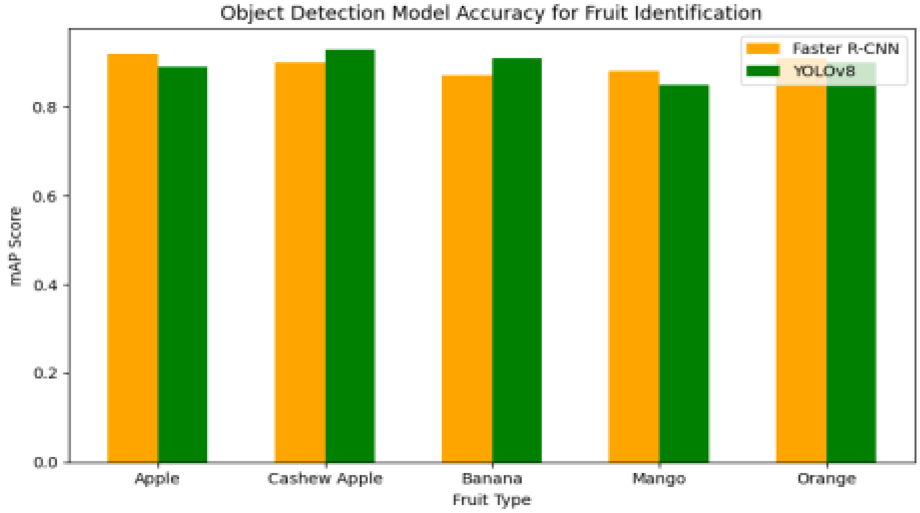

Fig 5: Comparison of Faster R-CNN vs. YOLOv8 Accuracy in Fruit Identification

iii. Ripeness Detection in Tomatoes Using YOLOv8+

Tomato maturity grading is necessary for selective picking, when only perfectly ripe fruits are picked while immature ones are left to continue developing. Yet, variations in lighting and leaf occlusion make precise detection troublesome. To address this, a better YOLOv8+ model was built, with the addition of RCA-CBAM (Region & Color Attention) and BiFPN (Bidirectional Feature Pyramid Network) for improved feature extraction, as depicted in Figure 6. The research indicated in Figure 7 that YOLOv8+ performed better than the baseline YOLOv8, with 95.8% precision and 91.7% accuracy. In addition, incorporating the Inner-FocalerIoU loss function enhanced class balance and decreased incorrect classifications. The results confirm that YOLOv8+ is very effective for real-time tomato ripeness classification and therefore suitable for automatic harvesting purposes [44].

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

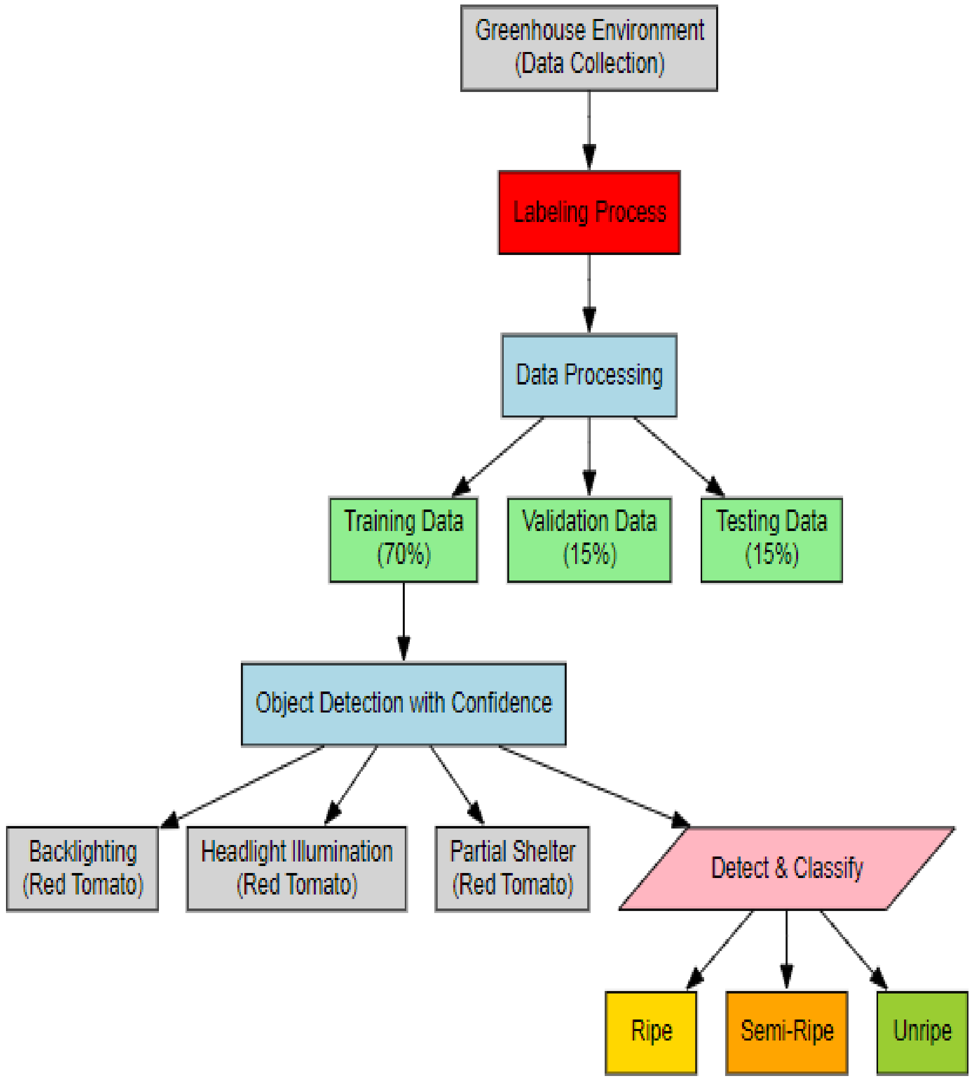

Fig 6. Ripeness Detection in Tomatoes Using YOLOv8

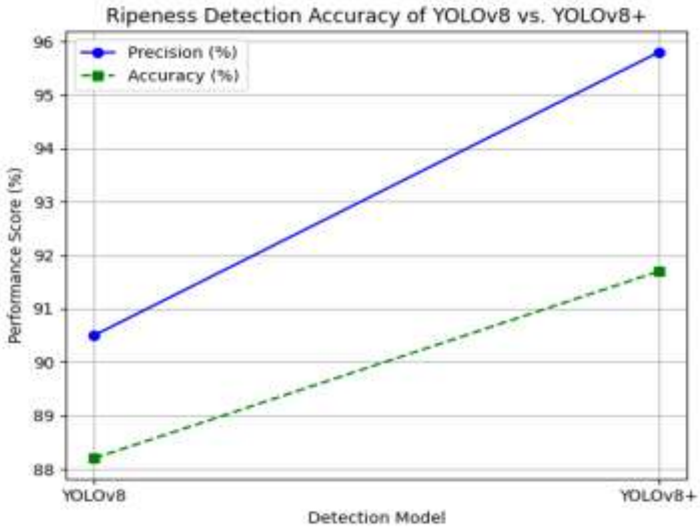

Fig 7: Ripeness Detection Accuracy of YOLOv8 vs. YOLOv8+

iv. Advanced Segmentation Methods for Ripeness Classification and Autonomous Harvesting

In contemporary farming, object segmentation methods are essential in autonomous ripeness classification, fruit counting, defect detection, and selective harvesting. Instance segmentation can be applied to identify single fruit instances, while semantic segmentation classifies image areas in terms of ripeness categories. New developments in deep learning architectures such as Mask R-CNN, UNet, CNN-Transformers, and OccluInst have substantially enhanced accuracy in actual orchard automation and food processing. A comparative study of segmentation models on various agricultural applications indicates the merits and demerits of various approaches. Research has shown that one-stage segmentation models can be superior to Mask R-CNN, which is conventional, with higher accuracy and improved boundary predictions [45]. Likewise, AI-based fruit counting based on highly annotated datasets (e.g., RipSetCocoaCNCH12) has enhanced yield estimation, fertilization planning, and harvesting timetables [46]. Also, hybrid Transformer-CNN models are becoming powerful tools in robotic harvesting under occlusion-dominant situations [48].

Table 3 Comparative Analysis of Advanced Segmentation Models for Agriculture

Application	Dataset	Model Used	Key Features	Performance (Accuracy/AP/ mAP50)	Inference Speed (FPS/ ms)	Best Use Case	Reference

Peach Ripeness Classification	NinePeach Dataset	Mask R-CNN vs. One-Stage Segmentation	Handles lighting variations, occlusions, and multiple fruit adhesions	Mask R-CNN: 69.91% AP One-Stage: 72.12% AP	-	Orchard automation & selective harvesting	[45]
AI-Driven Fruit Counting & Ripeness Detection	RipSetCocoaCNCH12 (Cocoa Pods)	Instance & Semantic Segmentation	4,116 images labeled across four ripeness stages	Improves fruit counting accuracy for yield estimation	-	Harvest scheduling & yield forecasting	[46]
Quality Assessment in the Food Industry	Rotten vs. Fresh Apple Dataset	UNet vs. Enhanced UNet (En-UNet)	Semantic segmentation of peel defects	UNet: 95.36% accuracy En-UNet: 97.54% accuracy, 0.866 IoU	-	Automated fruit sorting in food processing	[47]
Selective Harvesting with Robotic Systems	Broccoli Harvesting Dataset	CNN-Transformer (OccluInst) vs. Mask R-CNN	RGB-depth fusion, occlusion handling	OccluInst: 86.2% mAP50, 83.5% mAR	51.4 FPS	Autonomous harvesting robots under occlusion	[48]

The integration of deep learning-based segmentation models is revolutionizing precision agriculture as shown in Table 3. While Mask R-CNN remains effective, newer one-stage segmentation models, CNN-Transformers, and enhanced UNet architectures provide better

International Journal for Innovative Engineering and Management Research
PEER REVIEWED OPEN ACCESS INTERNATIONAL JOURNAL
www.ijiemr.org

efficiency, speed, and accuracy in fruit detection, defect analysis, and automated harvesting. Future advancements should focus on improving dataset annotations, optimizing real-time inference on edge AI devices, and deploying these models in large-scale farming operations. By leveraging AI-driven segmentation models, the agricultural sector can enhance yield predictions, reduce food waste, and optimize selective harvesting operations, ensuring sustainable and profitable farming practices.

6. MULTI-MODAL AI FOR RIPENESS CLASSIFICATION AND MONITORING

The fusion of several sensing modalities has greatly improved fruit ripeness classification, quality evaluation, and crop monitoring. Conventional RGB-based fruit classification techniques are frequently hampered by lighting condition changes, occlusions, and environmental noise. In order to mitigate these limitations, scientists have established multi-modal AI methods that combine thermal, hyperspectral, depth, and sensor-based information to enhance classification performance and stability. Cross-modal learning methods have also been used for plant health monitoring, preharvest yield prediction, and food safety uses. Figure 8, depicts an AI-based framework for fruit ripeness classification, sorting, quality monitoring, and post-harvest storage optimization. It combines multi-modal data sources, deep learning models, and classification methods to improve agricultural automation. Data collection is the beginning, with RGB, Thermal, Hyperspectral, and UAV imaging collecting visual data, sensor data in temperature, humidity, and gas levels offering real-time observations. Preprocessing methods like noise removal and image restoration smooth out the data prior to analysis. Cross-modal fusion combines several data sources, making it more accurate.

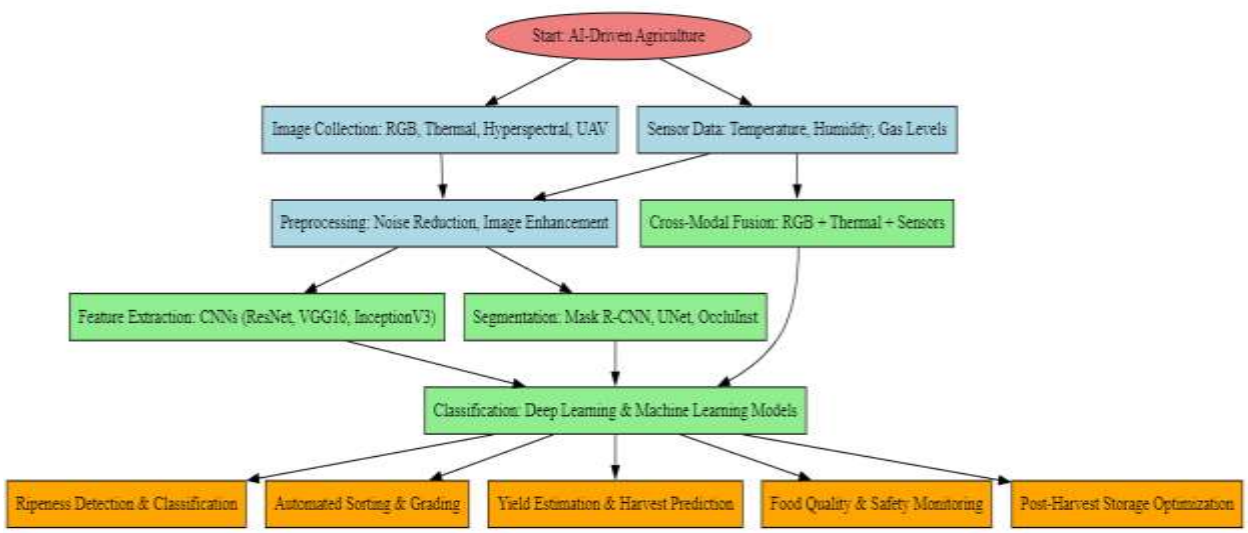

Fig.8. Multi-Modal AI Workflow for Ripeness Classification and Post-Harvest Optimization

During processing, CNN models (ResNet, VGG16, InceptionV3) extract fruit features, while segmentation techniques (Mask R-CNN, UNet, OccluInst) identify boundaries and levels of ripeness. Deep learning classifiers examine the extracted features to classify ripeness, quality, and storage requirements. The last stage comprises real-world applications, such as ripeness detection and classification, automated sorting and grading, yield estimation and harvest forecasting, food quality and safety inspection, and post-harvest storage optimization. With sensor fusion, deep learning, and image segmentation integration, the multi-modal AI system improves efficiency, minimizes postharvest loss, and supports improved decision-making in agriculture.

a. Fusing RGB, Thermal, Hyperspectral, and Sensor Data

The quality of fruit is a significant aspect affecting consumer acceptance and market value. Research examined the application of thermal imaging and deep learning models to distinguish between various pineapple types on the basis of physicochemical changes [49]. The study employed a multi-modal data fusion approach merging thermal imaging with deep learning structures like ResNet, VGG16, and InceptionV3. The outcomes proved that thermal imaging fusion characteristics with CNN-based models substantially enhanced classification accuracy with a high accuracy rate of 0.9687. This highlights the role of thermal imaging in real-time monitoring of fruit quality. A systematic review further compared machine vision systems and AI algorithms for autonomous detection and harvesting of fruits [51]. The review contrasted different 3D imaging modalities, vision sensors, and AI-based fruit detection strategies, emphasizing the need for multi-modal sensor fusion for improved fruit classification and ripeness tracking. The research determined that the use of multiple imaging modalities (e.g., RGB, hyperspectral, and LiDAR) enhances detection accuracy, minimizes false positives, and maximizes automated harvesting efficiency. A bibliometric review of fruit sorting and grading research also validated these results by examining 129 machine vision studies published between 2011 and 2023 [52]. The research found a high growth rate of deep learning models for fruit classification over the last five years, with pre-trained transfer learning models yielding the best accuracy. Research directions in the future focus on the use of multiple sensors, enhanced system robustness, and the creation of standardized evaluation metrics for fruit classification models.

b. Cross-Modal Learning in Fruit and Crop Monitoring

As climate change and resource scarcity pose challenges to agricultural sustainability, integrating artificial intelligence (AI) with multi-modal data sources is becoming crucial for crop health monitoring, plant stress detection, and preharvest yield estimation. A study introduced PA-RDFKNet (Plant Age RGB-Depth Fusion Knowledge Distillation Network) for cross-modal learning in plant growth monitoring [53]. The proposed model fused RGB and depth images during training, but relied solely on RGB images for inference, making it accessible for farmers lacking depth cameras. The approach significantly reduced mean squared error in plant age prediction from 2 weeks to 0.14 weeks, demonstrating high potential for real-time crop growth monitoring. Unmanned Aerial Vehicles (UAVs) have also emerged as key tools for preharvest crop yield estimation. A systematic review of UAV-based yield estimation research analyzed 76 studies on wheat, corn, rice, and soybeans [54]. The findings revealed that machine learning models trained on UAV-collected remote sensing data delivered highly accurate yield predictions. The study emphasized the importance of feature selection, multi-modal fusion (multispectral + thermal imaging), and deep learning models (CNNs and Random Forests) for optimal results. Future research should focus on data augmentation, feature engineering, and real-time UAV-based yield estimation.

c. AI for Food Monitoring and Safety

Food monitoring is an essential practice for nutrient tracking, dietary management, and personal health monitoring. AI-based food recognition systems have been developed to analyze food composition and quality based on smartphone images. However, existing tools suffer from low accuracy and limited contextual understanding. A new study proposed an optimized food recognition model, leveraging machine learning, natural language processing (NLP), and contextual ingredient data from online recipes [55]. The system, named FoodInsight, demonstrated improved ingredient recognition accuracy and was successfully integrated into an Android-based food tracking app. The study suggests that combining food image recognition with AI-driven nutrient databases can significantly improve dietary monitoring applications, opening new opportunities for automated food safety checks and personalized nutrition planning.

7. SELF-SUPERVISED & FEW-SHOT LEARNING IN RIPENESS CLASSIFICATION

a. Training Models with Limited Labeled Data

Few-shot image classification is designed to classify unseen categories using only a limited number of labeled samples. Traditional deep learning models struggle with generalization under such constraints. Recent advancements leverage meta-learning and self-supervised learning (SSL) to improve the adaptability of models. One study proposed an SSL-based embedding network to enhance feature learning, outperforming baseline models in MiniImageNet and CUB datasets. Further evaluation across four cross-domain few-shot learning datasets demonstrated state-of-the-art results [56]. Another approach, SSL-ProtoNet, integrates self-supervised learning, Prototypical Networks, and knowledge distillation to enhance few-shot learning performance. The method consists of three key stages: pre-training, fine-tuning, and self-distillation, significantly reducing overfitting. Experimental results on miniImageNet, tieredImageNet, and CIFAR-FS confirmed the superiority of SSL-ProtoNet over conventional few-shot learning approaches [57]. In remote sensing applications, scene classification remains a challenge due to insufficient labeled data and the complexity of overlapping objects. A self-supervised learning framework (RS-FewShotSSL) was proposed to address these issues using contrastive learning and EfficientNet-B3 for feature extraction. By leveraging large amounts of unlabeled remote sensing images, the model demonstrated superior classification performance on three public remote sensing datasets, outperforming standard approaches such as SimCLR, MoCo, BYOL, and IDSSL [58].

b. Synthetic Data Generation and Augmentation

Self-supervised learning (SSL) has gained significant attention in natural language processing and image classification, particularly for handling small datasets. A novel composite rotation-based auxiliary task was developed to enhance representation learning, enabling deep learning models to extract generalized and discriminative features from few labeled samples. The suggested method exhibited state-of-the-art performance on various benchmarks [59]. In wireless network security, FS-SEI based on Radio Frequency Fingerprinting (RFF) is challenging because labeled samples are limited. A new self-supervised and adversarial augmentation (SA2SEI) method was proposed to enhance feature learning and model robustness. Experiments on ADS-B and Wi-Fi datasets indicated that SA2SEI greatly improves classification accuracy, even with just five samples per device [60]. Lastly, Few-Shot Class Incremental Learning (FSCIL) also poses further challenges because of overfitting and catastrophic forgetting. Scholars proposed a feature fusion-based self-supervised learning method, which fused representations from supervised and self-supervised models. The approach outperformed the current classifiers on CUB200, miniImageNet, and CIFAR100 benchmarks, pushing the state-of-the-art in FSCIL research [60]. The improvements in self-supervised learning and few-shot learning have revolutionized AI-based classification models to make them more efficient, flexible, and generalizable in situations with sparse labeled data. Through the use of meta-learning, contrastive learning, and data augmentation, scientists have improved model robustness across a wide range of domains, such as remote sensing, agriculture, security, and incremental learning. These methods provide promising avenues for real-world AI deployment, especially in automated fruit ripeness classification and quality evaluation.

8. CONCLUSION

This study has undertaken a detailed review of the current advances in AI-based fruit ripeness detection, focusing on deep learning-based classification, object detection, segmentation, and multi-modal fusion methods. The results underscore the revolutionary potential of AI in precision agriculture to accurately and efficiently predict ripeness while minimizing the reliance on human inspection. CNNs, transformers, and hybrid models have shown high classification efficiency with models such as YOLOv8 performing better than conventional segmentation techniques in real-time scenarios. Furthermore, multi-modal techniques combining hyperspectral, thermal, and sensor-based data have also enhanced classification resilience. Nevertheless, some challenges still persist. The access to high-quality, varied datasets for AI training remains a bottleneck, calling for the creation of standardized, annotated databases for various fruit types and ripeness levels. Real-time inference and deployment on resource-limited devices still pose technical hurdles, calling for research into lightweight deep learning algorithms and edge computing. Additionally, explainability and interpretability of AI outcomes in agricultural uses have to be improved to facilitate higher farmer uptake and regulatory acceptance. Future research has to target improved scalability of AI-based ripeness detection by fusing IoT-based smart agriculture solutions, building more resilient models that generalize well under varied environmental settings, and enhancing AI explainability for improved trust and useability. Through tackling these issues, AI-based ripeness detection systems can transform farming practices, enhancing food quality, minimizing waste, and optimizing supply chain Efficiency

9. References

[1] M. Rienth *et al.*, "Grape Berry Secondary Metabolites and their Modulation by Abiotic Factors in a Climate Change Scenario–A Review," *Frontiers in Plant Science*, vol. 12, Mar. 2021, doi: 10.3389/fpls.2021.643258.

[2] M. Chen, H. Gu, L. Wang, Y. Shao, R. Li, and W. Li, "Exogenous ethylene promotes peel color transformation by regulating the degradation of chlorophyll and synthesis of anthocyanin in postharvest mango fruit," *Frontiers in Nutrition*, vol. 9, May 2022, doi: 10.3389/fnut.2022.911542.

[3] S. Gonzali and P. Perata, "Fruit colour and novel mechanisms of genetic regulation of pigment production in tomato fruits," *Horticulturae*, vol. 7, no. 8, p. 259, Aug. 2021, doi: 10.3390/horticulturae7080259.

[4] A. A. Gebregziabher, S. Supriyadi, S. Indarti, and L. Setyowati, "Texture profile and pectinase activity in tomato fruit (Solanum lycopersicum, servo F1) at different maturity stages and storage temperatures," *Planta Tropika*, vol. 9, no. 1, pp. 20–34, Feb. 2021, doi: 10.18196/pt.v9i1.9139.

[5] M. Qian *et al.*, "The downregulation of PpPG21 and PpPG22 influences peach fruit texture and softening," *Planta*, vol. 254, no. 2, Jul. 2021, doi: 10.1007/s00425-021-03673-6.

[6] J. Chen, Y. Li, F. Li, K. Hong, and D. Yuan, "Effects of procyanidin treatment on the ripening and softening of banana fruit during storage," *Scientia Horticulturae*, vol. 292, p. 110644, Oct. 2022, doi: 10.1016/j.scienta.2022.110644.

[7] G. Muskovics, J. Felföldi, E. Kovács, R. Perlaki, and T. Kállay, "Changes in physical properties during fruit ripening of Hungarian sweet cherry (Prunus avium L.) cultivars,"

Postharvest Biology and Technology, vol. 40, no. 1, pp. 56–63, Mar. 2006, doi: 10.1016/j.postharvbio.2005.12.007.

[8] M. Zhao, Z. You, H. Chen, X. Wang, Y. Ying, and Y. Wang, “Integrated fruit Ripeness Assessment System based on an artificial olfactory sensor and deep learning,” *Foods*, vol. 13, no. 5, p. 793, Mar. 2024, doi: 10.3390/foods13050793.

[9] M. Wang, Y. Zhou, and X. Li, “Detection Strategies for Volatile Fragrance Released from Agricultural Products: Progress and Prospects,” *Advanced Sensor Research*, Jun. 2024, doi: 10.1002/adsr.202400044.

[10] N. Anjali *et al.*, “State-of-the-art non-destructive approaches for maturity index determination in fruits and vegetables: principles, applications, and future directions,” *Food Production Processing and Nutrition*, vol. 6, no. 1, Feb. 2024, doi: 10.1186/s43014-023-00205-5.

[11] Q. Xiao *et al.*, “Soluble sugar, organic acid and phenolic composition and flavor evaluation of plum fruits,” *Food Chemistry X*, vol. 24, p. 101790, Aug. 2024, doi: 10.1016/j.fochx.2024.101790.

[12] X. Wang *et al.*, “Effects of fruit load on photosynthetic characteristics of peach leaves and fruit quality,” *Scientia Horticulturae*, vol. 299, p. 110977, Mar. 2022, doi: 10.1016/j.scienta.2022.110977.

[13] X. Zhang *et al.*, “Varietal Differences Among the Fruit Quality Characteristic of 15 Spine Grapes (Vitis davidii Foëx),” *HortScience*, vol. 57, no. 10, pp. 1282–1288, Oct. 2022, doi: 10.21273/hortsci16702-22.

[14] R. Alonso-Salinas, S. López-Miranda, A. J. Pérez-López, and J. R. Acosta-Motos, “Strategies to Delay Ethylene-Mediated Ripening in Climacteric Fruits: Implications for shelf life extension and postharvest quality,” *Horticulturae*, vol. 10, no. 8, p. 840, Aug. 2024, doi: 10.3390/horticulturae10080840.

[15] N. I. D. Arista and S. A. Ardiningtyas, “Comparative analysis of ethylene-induced ripening in climacteric and non-climacteric fruits: implications for post-harvest management,” *Social Agriculture, Food System, and Environmental Sustainability*, vol. 1, no. 2, pp. 90–100, Aug. 2024, doi: 10.61511/safses.v1i2.2024.1202.

[16] D. Noiwan, P. Suppakul, and P. Rachtanapun, “Preparation of Methylcellulose Film-Based CO2 Indicator for Monitoring the Ripeness Quality of Mango Fruit cv. Nam Dok Mai Si Thong,” *Polymers*, vol. 14, no. 17, p. 3616, Sep. 2022, doi: 10.3390/polym14173616.

[17] M. Palumbo *et al.*, "Rapid and Non-Destructive techniques for the discrimination of ripening stages in candonga strawberries," *Foods*, vol. 11, no. 11, p. 1534, May 2022, doi: 10.3390/foods11111534.

[18] X. Zhou *et al.*, "Aggregation and assessment of grape quality parameters with visible-near-infrared spectroscopy: Introducing a novel quantitative index," *Postharvest Biology and Technology*, vol. 218, p. 113131, Aug. 2024, doi: 10.1016/j.postharvbio.2024.113131.

[19] Md. S. Bapary, Md. N. Islam, N. Kumer, M. H. Tahery, M. A. A. Noman, and M. Mohi-Ud-Din, "Postharvest physicochemical and nutritional properties of tomato fruit at different maturity stages affected by physical impact," *Applied Food Research*, p. 100636, Nov. 2024, doi: 10.1016/j.afres.2024.100636.

[20] R. Aulia *et al.*, "Non-destructive prediction of protein contents of soybean seeds using near-infrared hyperspectral imaging," *Infrared Physics & Technology*, vol. 127, p. 104365, Oct. 2022, doi: 10.1016/j.infrared.2022.104365.

[21] J. Zhang *et al.*, "A non-destructive determination of protein content in potato flour noodles using near-infrared hyperspectral imaging technology," *Infrared Physics & Technology*, vol. 130, p. 104595, Feb. 2023, doi: 10.1016/j.infrared.2023.104595.

[22] J. D. Rabanera, J. D. Guzman, and K. F. Yaptenco, "Rapid and Non-destructive measurement of moisture content of peanut (Arachis hypogaea L.) kernel using a near-infrared hyperspectral imaging technique," *Journal of Food Measurement & Characterization*, vol. 15, no. 4, pp. 3069–3078, Mar. 2021, doi: 10.1007/s11694-021-00894-x.

[23] Y. Hu and Z. Kang, "The Rapid Non-Destructive Detection of Adulteration and its degree of tieguanyin by fluorescence hyperspectral technology," *Molecules*, vol. 27, no. 4, p. 1196, Feb. 2022, doi: 10.3390/molecules27041196.

[24] S. Botti, F. Bonfigli, V. Nigro, A. Rufoloni, and A. Vannozzi, "Evaluating the Conservation State of Naturally Aged Paper with Raman and Luminescence Spectral Mapping: Toward a Non-Destructive Diagnostic Protocol," *Molecules*, vol. 27, no. 5, p. 1712, Mar. 2022, doi: 10.3390/molecules27051712.

[25] Q. Xue, G. Yu, F. Lu, and Y. Dong, "Fluorescent labelling combined with confocal differential Raman spectroscopy to detect microplastics in seawater," *Spectrochimica Acta Part a Molecular and Biomolecular Spectroscopy*, vol. 320, p. 124591, Jun. 2024, doi: 10.1016/j.saa.2024.124591.

[26] M. Rizzo, M. Marcuzzo, A. Zangari, A. Gasparetto, and A. Albarelli, "Fruit ripeness classification: A survey," *Artificial Intelligence in Agriculture*, vol. 7, pp. 44–57, Mar. 2023, doi: 10.1016/j.aiia.2023.02.004.

[27] N. Anjali *et al.*, "State-of-the-art non-destructive approaches for maturity index determination in fruits and vegetables: principles, applications, and future directions," *Food Production Processing and Nutrition*, vol. 6, no. 1, Feb. 2024, doi: 10.1186/s43014-023-00205-5.

[28] H. Zhou, X. Wang, W. Au, H. Kang, and C. Chen, "Intelligent robots for fruit harvesting: recent developments and future challenges," *Precision Agriculture*, vol. 23, no. 5, pp. 1856–1907, Jun. 2022, doi: 10.1007/s11119-022-09913-3.

[29] E. Vrochidou, C. Bazinas, M. Manios, G. A. Papakostas, T. P. Pachidis, and V. G. Kaburlasos, "Machine Vision for ripeness estimation in viticulture automation," *Horticulturae*, vol. 7, no. 9, p. 282, Sep. 2021, doi: 10.3390/horticulturae7090282.

[30] L. Chen, P. He, H. Zhang, W. Peng, J. Qiu, and F. Lü, "Applications of machine learning tools for biological treatment of organic wastes: Perspectives and challenges," *Circular Economy*, vol. 3, no. 2, p. 100088, May 2024, doi: 10.1016/j.cec.2024.100088.

[31] R. Kanjilal, J. E. Saenz, and I. Uysal, "Large-scale data-driven uniformity analysis and sensory prediction of commercial banana ripening process," *Postharvest Biology and Technology*, vol. 219, p. 113203, Sep. 2025, doi: 10.1016/j.postharvbio.2025.113203.

[32] S. Deepa, J. L. Zeema, and S. Gokila, "Exploratory Architectures Analysis of various pre-trained image classification models for deep learning," *Journal of Advances in Information Technology*, vol. 15, no. 1, pp. 66–78, Jan. 2024, doi: 10.12720/jait.15.1.66-78.

[33] M. Alruwaili and M. Mohamed, "An Integrated Deep Learning Model with EfficientNet and ResNet for Accurate Multi-Class Skin Disease Classification," *Diagnostics*, vol. 15, no. 5, p. 551, Feb. 2025, doi: 10.3390/diagnostics15050551.

[34] M. T. Ahad, Y. Li, B. Song, and T. Bhuiyan, "Comparison of CNN-based deep learning architectures for rice diseases classification," *Artificial Intelligence in Agriculture*, vol. 9, pp. 22–35, Jul. 2023, doi: 10.1016/j.aiia.2023.07.001.

[35] A. Khan *et al.*, "A survey of the vision transformers and their CNN-transformer based variants," *Artificial Intelligence Review*, vol. 56, no. S3, pp. 2917–2970, Oct. 2023, doi: 10.1007/s10462-023-10595-0.

[36] M. J. Willemink, H. R. Roth, and V. Sandfort, "Toward foundational deep learning models for medical imaging in the new era of transformer networks," *Radiology Artificial Intelligence*, vol. 4, no. 6, Nov. 2022, doi: 10.1148/ryai.210284.

[37] A. A. Aleissaee *et al.*, "Transformers in Remote Sensing: a survey," *Remote Sensing*, vol. 15, no. 7, p. 1860, Mar. 2023, doi: 10.3390/rs15071860.

[38] S. L. Ramaswamy and J. Chinnappan, "RecogNet-LSTM+CNN: a hybrid network with attention mechanism for aspect categorization and sentiment classification," *Journal of Intelligent Information Systems*, vol. 58, no. 2, pp. 379–404, Jan. 2022, doi: 10.1007/s10844-021-00692-3.

[39] S. K. Prabhakar, H. Rajaguru, and D.-O. Won, "Performance Analysis of Hybrid Deep Learning Models with Attention Mechanism Positioning and Focal Loss for Text Classification," *Scientific Programming*, vol. 2021, pp. 1–12, Oct. 2021, doi: 10.1155/2021/2420254.

[40] H. Butt, M. R. Raza, M. J. Ramzan, M. J. Ali, and M. Haris, "Attention-Based CNN-RNN Arabic Text Recognition from Natural Scene Images," *Forecasting*, vol. 3, no. 3, pp. 520–540, Jul. 2021, doi: 10.3390/forecast3030033.

[41] P. Kavianpour, M. Kavianpour, E. Jahani, and A. Ramezani, "A CNN-BiLSTM model with attention mechanism for earthquake prediction," *The Journal of Supercomputing*, vol. 79, no. 17, pp. 19194–19226, May 2023, doi: 10.1007/s11227-023-05369-y.

[42] R. Sapkota, D. Ahmed, and M. Karkee, "Comparing YOLOv8 and Mask R-CNN for instance segmentation in complex orchard environments," *Artificial Intelligence in Agriculture*, vol. 13, pp. 84–99, Jul. 2024, doi: 10.1016/j.aiia.2024.07.001.

[43] K. Sawant *et al.*, "Fruit Identification using YOLO v8 and Faster R-CNN –A Commparative Study," *Sawant*, pp. 1–6, Mar. 2024, doi: 10.1109/icdcot61034.2024.10515677.

[44] Z. Yang, Y. Li, Q. Han, H. Wang, C. Li, and Z. Wu, "A method for tomato ripeness recognition and detection based on an improved YOLOV8 model," *Horticulturae*, vol. 11, no. 1, p. 15, Dec. 2024, doi: 10.3390/horticulturae11010015.

[45] Z. Zhao, Y. Hicks, X. Sun, and C. Luo, "Peach ripeness classification based on a new one-stage instance segmentation model," *Computers and Electronics in Agriculture*, vol. 214, p. 108369, Oct. 2023, doi: 10.1016/j.compag.2023.108369.

[46] J. F. Restrepo-Arias, M. I. Salinas-Agudelo, M. I. Hernandez-Pérez, A. Marulanda-Tobón, and M. C. Giraldo-Carvajal, "RIPSETCOCOACNCH12: labeled dataset for ripeness stage detection, semantic and instance segmentation of cocoa pods," *Data*, vol. 8, no. 6, p. 112, Jun. 2023, doi: 10.3390/data8060112.

[47] K. Roy, S. S. Chaudhuri, and S. Pramanik, "Deep learning based real-time Industrial framework for rotten and fresh fruit detection using semantic segmentation," *Microsystem Technologies*, vol. 27, no. 9, pp. 3365–3375, Nov. 2021, doi: 10.1007/s00542-020-05123-x.

[48] S. Kang, J. Fan, Y. Ye, C. Li, D. Du, and J. Wang, "Maturity recognition and localisation of broccoli under occlusion based on RGB-D instance segmentation network," *Biosystems Engineering*, vol. 250, pp. 270–284, Jan. 2025, doi: 10.1016/j.biosystemseng.2025.01.007.

[49] M. M. Ali, N. Hashim, S. A. Aziz, and O. Lasekan, "Utilisation of Deep Learning with Multimodal Data Fusion for Determination of Pineapple Quality Using Thermal Imaging," *Agronomy*, vol. 13, no. 2, p. 401, Jan. 2023, doi: 10.3390/agronomy13020401.

[50] A. Fuller, Z. Fan, C. Day, and C. Barlow, "Digital Twin: enabling technologies, challenges and open research," *IEEE Access*, vol. 8, pp. 108952–108971, Jan. 2020, doi: 10.1109/access.2020.2998358.

[51] B. O. Olorunfemi, N. I. Nwulu, O. A. Adebo, and K. A. Kavadias, "Advancements in machine visions for fruit sorting and grading: A bibliometric analysis, systematic review, and future research directions," *Journal of Agriculture and Food Research*, vol. 16, p. 101154, Apr. 2024, doi: 10.1016/j.jafr.2024.101154.

[52] S. Bansal, M. Singh, S. Barda, N. Goel, and M. Saini, "PA-RDFKNet: Unifying Plant Age Estimation through RGB-Depth Fusion and Knowledge Distillation," *IEEE Transactions on AgriFood Electronics*, vol. 2, no. 2, pp. 226–235, Jul. 2024, doi: 10.1109/tafe.2024.3418818.

[53] J. Yuan, Y. Zhang, Z. Zheng, W. Yao, W. Wang, and L. Guo, "Grain crop yield prediction using Machine Learning Based on UAV Remote Sensing: A Systematic Literature review," *Drones*, vol. 8, no. 10, p. 559, Oct. 2024, doi: 10.3390/drones8100559.

[54] Z. Hakguder, Y. He, W. Chai, and J. Cui, "Smart Diet Management through Food Image and Cooking Recipe Analysis," *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 2603–2610, Dec. 2022, doi: 10.1109/bibm55620.2022.9994888.

[55] D. Chen, Y. Chen, Y. Li, F. Mao, Y. He, and H. Xue, "Self-Supervised learning for Few-Shot image classification," *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1745–1749, May 2021, doi: 10.1109/icassp39728.2021.9413783.

[56] J. Y. Lim, K. M. Lim, C. P. Lee, and Y. X. Tan, "SSL-ProtoNet: Self-supervised Learning Prototypical Networks for few-shot learning," *Expert Systems With Applications*, vol. 238, p. 122173, Oct. 2024, doi: 10.1016/j.eswa.2024.122173.

[57] N. Alosaimi, H. Alhichri, Y. Bazi, B. B. Youssef, and N. Alajlan, "Self-supervised learning for remote sensing scene classification under the few shot scenario," *Scientific Reports*, vol. 13, no. 1, Jan. 2023, doi: 10.1038/s41598-022-27313-5.

[58] P. Kumar and D. Toshniwal, “Self-Supervison with data-augmentation improves few-shot learning,” *Applied Intelligence*, vol. 54, no. 4, pp. 2976–2997, Feb. 2024, doi: 10.1007/s10489-024-05340-1.

[59] C. Liu *et al.*, “Overcoming data limitations: a Few-Shot specific emitter identification method using Self-Supervised learning and adversarial augmentation,” *IEEE Transactions on Information Forensics and Security*, vol. 19, pp. 500–513, Oct. 2023, doi: 10.1109/tifs.2023.3324394.

[60] T. Ahmad *et al.*, “Few-Shot Class Incremental Learning leveraging Self-Supervised features,” *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Jun. 2022, doi: 10.1109/cvprw56347.2022.00435.