

REVIEW

Open access

An Ensemble Machine Learning Framework for Avocado Ripeness Detection Based on Near-Infrared Spectral Data

James O'Leary^{1*}, Aisling Dunne¹, Sean O'Brien²

Abstract

Precise categorization of avocado ripeness plays a vital role in strengthening post-harvest operations and curbing waste throughout agricultural supply networks. The present study centers on constructing a resilient ensemble classification framework employing spectral readings from 120 kilograms of Buccaneer avocados procured from the Royal Project in Chiang Mai, Thailand. The specimens were assessed via near-infrared (NIR) spectroscopy at 18 wavelengths. Five distinct machine learning models—Random Forest, Decision Tree, XGBoost, Gradient Boosting, and Gaussian Mixture Model—were separately trained and subsequently fused into a unified ensemble. Four distinct algorithms were deployed to fine-tune the weighting of each model: Bayesian optimisation, differential evolution, particle swarm optimisation, and grid search. Evaluation metrics covering accuracy, precision, recall, F1-score, confusion matrices, and ROC curves were applied. Grid Search yielded the strongest classification outcomes, securing an accuracy of 82.5% alongside an F1-score of 85.3%, emphasizing the value of weight-refined ensemble learning relative to standalone classifiers. This investigation delivers a scalable and transparent pipeline for non-destructive ripeness identification. The outcomes, notwithstanding constraints related to overfitting and dependency on spectral data integrity, pave the way for eventual real-time field deployment in agriculture.

Keywords Precision agriculture, Avocado ripeness classification, Ensemble learning, Optimization algorithms, NIR spectroscopy, Machine learning in agriculture

*Correspondence:

James O'Leary
james.oleary@ucd.ie

¹ Department of Near-Infrared Spectroscopy for Fruit Ripeness, School of Agriculture, University College Dublin, Dublin, Ireland

² Department of Ensemble Machine Learning for Food Quality, School of Agriculture, University of Limerick, Limerick, Ireland

Introduction

Reliably gauging avocado maturity remains a persistent hurdle in cultivation practices with repercussions for post-harvest logistics, pricing frameworks, and end-consumer contentment. Conventional inspection methods lean extensively on human judgment, a process vulnerable to partiality, burdensome in time demands, and susceptible to irregularity. Such shortcomings frequently engender sorting errors, heightened post-harvest wastage, and supply chain bottlenecks. Additionally, the spectrum of ripeness states within any avocado lot introduces substantial obstacles to upholding uniform quality assurance. These established gaps accentuate an escalating requirement for automated, neutral, and scalable sorting mechanisms capable of aiding cultivators and distributors alike.

Evaluating avocado ripeness calls for scrutinizing inner physiological transformations that may lack obvious external indicators. Near-Infrared (NIR) spectroscopy furnishes a non-invasive avenue by quantifying light absorption at

designated wavelengths reflective of interior chemical properties. Machine learning applications in agriculture have proliferated in recent years, with ensemble learning structures consistently outperforming solitary models in precision and generalizability. Methodologies that amalgamate multiple classifiers confer marked strengths when contending with intricate, non-linear data forms such as spectral traces. Nevertheless, pinpointing the ideal weighting or influence of each constituent model inside the ensemble persists as a formidable obstacle to elevating predictive exactness. The present work prioritizes devising and refining a cooperative learning paradigm geared toward classifying avocado ripeness through NIR spectral data. The lineup integrates five foundational tools: Random Forest, Decision Tree, Gradient Boosting, XGBoost, and Gaussian Mixture Model. To better amalgamate model outputs, four optimization protocols—Bayesian Optimisation, Differential Evolution, Particle Swarm Optimisation, and Grid Search—are harnessed to calibrate ensemble weights with precision. A corpus of 120 kilogrammes of Buccaneer avocados was gathered from the Royal Project located in Chiang Mai, Thailand. Every sample was scanned across 18 NIR wavelengths to yield spectral signatures. The resultant dataset was then utilized to train and benchmark the individual algorithms and their optimized fusions. The evaluative benchmarks span accuracy, precision, recall, F1-score, confusion matrices, and ROC curves, delivering an exhaustive side-by-side performance portrait of the varied optimization approaches.

This enquiry propels precision agriculture forward by unveiling a non-destructive, budget-conscious, and scalable fruit classification avenue. The sophisticated ensemble learning technique bolsters classification fidelity and dependability, empowering improved stock governance and waste mitigation. The inclusion of optimization methodologies safeguards a sturdy model arrangement amenable to real-world utilization. Ultimately, the insights buttress wider campaigns to implement progressive agri-tech innovations and set the groundwork for IoT-driven real-time surveillance mechanisms in post-harvest processing. This segment surveys pertinent literature addressing non-invasive ripeness detection techniques, collective model strategies, and refinement tactics within the sphere of agricultural classification.

Literature Review

Swift and exact ripeness determination is indispensable for curbing post-harvest deterioration and sustaining steady avocado quality [1]. Segregation assumes heightened importance given the inconsistencies observed across avocado batches, most notably with shipped-in fruit [2]. Pinpointing ideal harvest windows and forecasting output volumes are decisive for diminishing wastage and safeguarding consumer-grade standards [3]. Established ripeness-testing conventions frequently depend on invasive firmness probes, unsuited to batch-level evaluation, while expert-led manual sorting proves tedious and fallible. Orthodox approaches suffer from subjectivity and variability, challenges magnified by pulpy softening gradients inside the fruit.

In light of these limitations, a mounting appetite exists for non-invasive techniques enabling batch-wide scrutiny without compromising product integrity. Such strategies upgrade quality supervision, shrink waste streams, and heighten consumer fulfillment. Spectral inspection has surfaced as a propitious non-destructive avenue [4-6].

Ensemble learning tactics have exhibited potency in elevating classification scores for ripeness forecasting efforts [4]. By intermingling several base classifiers, ensemble strategies amplify both accuracy and sturdiness, surmounting the fragilities of single models and augmenting reliability within convoluted data environments like fruit ripeness analysis.

Non-destructive techniques

Spectral measurements

A variety of spectral tools—among them Fourier Transform Infrared (FTIR), Near-Infrared (NIR), Raman spectroscopy, and hyperspectral imaging (HSI)—see widespread application for evaluating fruit and vegetable quality without causing damage [7]. These modalities probe how light behaves upon encountering a product to infer both interior conditions and surface features. HSI proves especially powerful owing to its dual capture of spatial alongside spectral intelligence, a capability that lends itself to flaw spotting and produce grading [7].

Hyperspectral imaging (HSI)

HSI stands as a refined imaging modality that collects a dense stack of narrow, continuous spectral bands tied to every individual pixel [7]. Such richly detailed spectral records make it possible to tease apart varied produce traits, spanning maturity, blemishes, and chemical profile. The technique has been deployed with success to forecast the ripening window of Hass avocados through deep learning regression models [1]. The fusion of HSI with deep learning frameworks signals considerable potential for streamlining sorting and downstream handling of avocado fruit [1].

Near-infrared (NIR) spectroscopy

Among non-invasive quality evaluation tools, NIR spectroscopy ranks as a common choice [7]. It operates by gauging how produce absorbs and returns near-infrared illumination, readings that map onto inner parameters like total soluble solids (TSS), solid content (SSC), and moisture content (MC) [7]. That said, NIR can struggle with penetration depth when faced with thick-skinned or heavy-rinded commodities, a factor that may undercut measurement fidelity [7]. Judicious selection of optical arrangement and wavelength span can lift predictive performance [7]. Visible-NIR spectroscopy has been tapped to assess several quality markers, among them soluble solids content (SSC) and total acid (TA) in table grapes [8].

Computer vision

Algorithmic processing of digital images opens a window onto exterior produce attributes—coloration, dimensions, contour, and surface imperfections [9, 10]. Such pipelines permit automated inspection and categorization, driving gains in both speed and consistency [11]. Vision-driven setups routinely call upon classification algorithms to sort commodities by their outward appearance [12].

Color analysis

Pigmentation stands as a telling ripeness signal, one that computer vision techniques can quantify with considerable precision [13]. The approach rests on measuring intensity values across assorted color models (for instance, RGB or HSV) and tying those readings to specific maturity stages [14]. Chromatic progressions—green yielding to yellow or red—furnish meaningful intelligence regarding how far ripening has advanced [13].

Texture analysis

Texture-focused examination homes in on surface-level patterns to flag defects or irregularities [5]. Operations like histogram equalization and K-means clustering enable image partitioning and subsequent classification driven by textural signatures [12]. Probing texture yields indicators tied to firmness and broader quality status [2].

Acoustic methods

Acoustic techniques center on the sound waves set off when a produce receives a tap or vibrational impulse [15]. Wave attributes such as frequency and amplitude exhibit ties to internal conditions, notably firmness and degree of maturity [2]. These methods avoid tissue damage altogether and can run on straightforward, affordable hardware [15].

Laser Doppler Vibrometry (LDV)

LDV operates as a standoff technique, directing a laser beam at an object to capture its vibrational response [2]. Applied to avocados, LDV tracks firmness shifts by monitoring resonant frequency patterns [2]. Investigations have uncovered a tight linkage between resonant frequency values and the logarithm of firmness, affirming LDV's capacity to follow avocado softening over time [2]. The resonant frequency drops in a linear fashion through the fourth day of ripening, after which the descent becomes more gradual [2].

Other non-destructive techniques

Beyond the approaches catalogued above, researchers have probed a range of supplementary non-invasive avenues for gauging produce quality. These features:

- Electronic noses rely on multisensory arrays to pick up and distinguish volatile organic compounds (VOCs) emanating from fruits [16]. Given that aromatic bouquet strongly sways consumer preference, e-noses offer a capable means of teasing apart intricate volatile blends [16].
- Tactile sensing deploys pressure-sensitive elements to register firmness and surface feel [17]. This proves especially handy within robotic harvesting contexts and for appraising food-item traits [17]. A purpose-built gripper outfitted with a capacitive tactile sensor matrix can palpate fruit and infer ripeness levels [18].
- Bio-impedance Spectroscopy (EIS) tracks the electrical impedance of fruit tissue across a sweep of frequencies [19]. The method establishes links between physicochemical transformations and electrical signatures, shedding light on the ripening trajectory [19].

Machine learning for fruit ripeness classification

Overview of machine learning techniques

Machine learning (ML) has overhauled the landscape of fruit ripeness sorting, bringing automated, high-fidelity tools to bear on multifaceted data streams [4]. Trained algorithms digest extensive datasets and generate ripeness predictions anchored in features drawn from non-destructive readings [20]. The deployment of ML can slash reliance on manual inspection while raising the uniformity and trustworthiness of ripeness calls [21].

Supervised learning

Supervised methods operate on annotated datasets, where every sample carries a designated ripeness label [22]. Through training, these algorithms internalize the mapping from input descriptors to target labels, equipping them to render ripeness judgments on previously unseen specimens. Supervised algorithms frequently encountered in fruit ripeness studies encompass:

- Support vector machines (SVM) stand out as resilient classifiers well-suited to high-dimensional feature spaces [20]. Their core logic involves pinpointing the optimal separating hyperplane that carves out the widest possible buffer between competing classes [12].
- K-nearest neighbors (KNN) follows a straightforward principle, tagging a data point according to the dominant class among its k closest neighbors [23]. Its ease of deployment makes it an attractive candidate for ripeness classification tasks, all the more so when paired with thoughtful feature curation [24].
- Decision trees take the form of branching architectures that route data points through successive decision splits [25]. Their transparent logic accommodates both continuous and categorical inputs while remaining readily interpretable [20].
- Naïve Bayes rests on probabilistic reasoning drawn from Bayes' theorem [24]. By assuming that features contribute independently, the classifier achieves a streamlined computational workflow [25].

Deep learning

Deep learning (DL) constitutes a subfield of machine learning built around artificial neural networks that contain multiple processing layers [20]. A defining strength of DL algorithms lies in their capacity to automatically extract sophisticated features straight from raw inputs, substantially lessening the burden of manual feature crafting [26]. Convolutional Neural Networks (CNNs) have become a dominant architecture within deep learning for image-based tasks and have shown considerable success in gauging fruit ripeness [26]

Convolutional neural networks (CNNs)

CNNs are designed to autonomously and adaptively capture spatial feature hierarchies from imagery [20]. Their architecture stacks several layer types, including convolutional layers, pooling layers, and fully connected layers [26]. The convolutional layers serve to distill features from the input images, while pooling layers function to shrink the spatial dimensions of the resulting feature maps [20]. The fully connected layers then carry out the fruit classification based on those distilled features [26]. CNNs have delivered exceptional results in fruit categorization tasks [27].

Transfer learning

Transfer learning revolves around repurposing an already-trained model as the starting point for tackling a fresh problem [14]. Doing so can dramatically shorten the training cycle and lift model effectiveness, a boon especially when working with limited datasets [20]. Architectures such as VGG16, ResNet, and MobileNet have been applied with good effect to the task of fruit ripeness sorting [14].

Ensemble learning

Ensemble learning brings together predictions from multiple machine learning models to elevate collective accuracy and resilience during classification [28]. By blending outputs across several models, ensemble strategies can temper overfitting and strengthen generalization [29]. Widely adopted ensemble approaches include:

- Random forest operates as an ensemble technique built from a collection of decision trees [24]. Each constituent tree is trained on a random draw of both data points and features [30]. The ultimate prediction emerges from averaging the outputs of every tree [24, 31].
- AdaBoost, short for adaptive boosting, sequentially chains together multiple weak classifiers to forge a single potent classifier [30]. The algorithm assigns instance-level weights, giving heavier emphasis to samples that preceding classifiers got wrong [29].
- A voting classifier pools the forecasts of several distinct machine learning models and settles on a class label according to the majority vote [25]. This tactic can bolster both the precision and sturdiness of classification by capitalizing on the complementary strengths of the constituent models [25].

Optimized ensemble learning

Benefits of optimized ensemble learning

Optimized ensemble learning brings a suite of benefits to the problem of avocado ripeness classification:

- Heightened accuracy—the fusion of multiple models routinely surpasses the performance ceiling of any single model [24].
- Amplified robustness—ensemble constructs are less easily thrown off by data noise or anomalous observations [28].
- Diminished overfitting—by averaging across several sets of predictions, ensemble methods curb the tendency to overfit [29].
- Superior generalization—ensembles tend to transfer more reliably to novel, previously unobserved data [24].

Feature selection and engineering

Feature selection represents a critical step in the machine learning pipeline, centered on isolating the most informative variables within a dataset [24]. Doing so boosts model accuracy and computational efficiency by shrinking data

dimensionality and stripping away duplicate or irrelevant inputs [24]. Feature engineering, by contrast, involves constructing new variables from the existing ones, a process that can further sharpen model performance [32].

Feature extraction from images

Within the realm of computer vision, feature extraction zeroes in on gleaning essential descriptors from fruit images [32]. Such descriptors can span color, surface texture, shape, and physical dimensions [5]. Architectures like CNNs possess the ability to learn these features directly from raw image data without explicit programming [26].

Feature selection techniques

A variety of feature selection approaches exist to help surface the attributes most germane to classifying avocado ripeness. These include:

- Mutual information gauges the degree of statistical dependence between a pair of variables [24]. Features exhibiting elevated mutual information with the target variable—here, the ripeness stage—are regarded as carrying greater relevance [24].
- Principal component analysis (PCA) serves as a dimensionality reduction method that transforms the original variable set into a collection of linearly independent principal components [28]. The leading principal components, which account for the bulk of the variance in the data, are then fed into the classification model as features [20].
- Genetic algorithms (GA) belong to a class of optimization procedures used to uncover the most effective feature subset [30]. GA initializes a population of candidate subsets at random and iteratively refines that population through genetic operators—selection, crossover, and mutation—to converge on a strong solution [30].

Optimization algorithms

Optimization algorithms serve to pinpoint the most effective blend of model parameters and features, aiming to push the accuracy of an ensemble learning framework as high as possible [28]. Several optimization algorithms lend themselves to this task, among them:

- Particle swarm optimization (PSO) draws its inspiration from the collective movement patterns observed in bird flocks or fish schools and operates as a population-based optimization method [28]. Within PSO, a swarm of particles is maintained, with each particle encoding a candidate solution to the optimization challenge at hand [28].
- Genetic algorithms (GA), as touched on earlier, can likewise be harnessed for tuning model parameters [30]. In the ensemble context, GA can be employed to discover the optimal weighting scheme for fusing the predictions contributed by different constituent models [30].
- Stochastic gradient descent (SGD) along with its numerous variants stands as the workhorse for training deep learning architectures [20]. SGD operates by making iterative adjustments to model parameters in a direction that progressively shrinks the loss function—a measure of the discrepancy between forecasted and actual ripeness stages [20].

A body of research has showcased the potency of ensemble learning when applied to fruit ripeness classification. Notable instances include:

- Davur et al. [1] devised an approach combining hyperspectral imaging with deep learning regression to forecast the number of days until Hass avocados reach optimal ripeness. Their method yielded predictions carrying a mean error of 1.17 days per fruit on the test partition.

- Tran et al. [23] put forward an affordable and rapid technique anchored in a low-cost Vector Network Analyzer (VNA) device, supplemented by K-nearest neighbor (KNN) and a Neural Network model, for recognizing fruit types and grading ripeness levels.
- Saragih and Emanuel [33] brought Convolutional Neural Networks to bear on the classification of banana ripeness, reaching strong accuracy figures with the MobileNet V2 architecture.
- Ko et al. [34] introduced an original method for identifying tomato ripeness that harnesses multiple convolutional neural network (ConvNet) streams together with their stochastic decision fusion (SDF) technique, attaining a mean accuracy of 96.
- Sharma and Hooda [29] crafted an efficient and precise methodology for evaluating bioactive compounds and drug activity through ensemble machine learning, demonstrating that a tuned random forest machine learning pipeline outperformed alternative ensemble-based machine learning approaches.

Notwithstanding these encouraging outcomes, applying ensemble learning to avocado ripeness classification carries several hurdles and constraints:

- Data requirements—ensemble learning methodologies typically demand sizeable datasets to properly train the base models and fine-tune the overall ensemble [20].
- Computational complexity—the training and optimization of ensemble configurations can incur substantial computational expense, a concern magnified when deep learning models enter the mix [28].
- Interpretability—ensemble architectures can prove opaque, complicating efforts to trace the reasoning behind particular predictions [20].
- Overfitting—ensembles risk absorbing noise in the training data if the component models grow overly complex or if insufficient regularization is applied to the ensemble [29].

Future investigations ought to pivot toward fusing data streams from multiple sensing modalities to sharpen both the accuracy and the dependability of avocado ripeness classification. Merging spectral readings, computer vision outputs, acoustic signatures, and other non-destructive measurement forms can deliver a holistic picture of fruit quality [7]. Data fusion methodologies supply the means to weave together signals from disparate sensors, yielding complementary insights that no single modality could offer alone [20].

To pave the way for deploying ensemble learning models under real-world constraints, there is a pressing need to engineer lightweight variants capable of being trained and executed on low-power hardware [27]. Techniques such as model compression, quantization, and pruning serve to shrink the footprint and computational demands of deep learning models while keeping accuracy largely intact [20].

Explainable AI (XAI) represents a burgeoning domain dedicated to boosting the openness and intelligibility of machine learning models [20]. XAI tools help unpack the decision logic of an ensemble learning model and isolate the features that exert the greatest influence on its outputs [35]. Such transparency can strengthen stakeholder trust in the model and smooth its path toward adoption within the avocado industry [35].

Prospective research should target the creation of real-time ripeness forecasting systems suitable for embedding into avocado grading and processing workflows [1]. These systems would need to track avocado ripeness on an ongoing basis and supply actionable feedback to refine sorting and handling steps [1]. Doing so holds promise for slashing post-harvest wastage and raising the overall caliber of avocados reaching the market [1].

A recurrent issue across many fruit ripeness datasets is class imbalance, where certain ripeness categories are markedly underrepresented relative to others [29]. This skew can spawn biased models that struggle to perform well on minority classes [29]. Countermeasures such as oversampling, undersampling, and cost-sensitive learning offer ways to redress data imbalance and lift model efficacy uniformly across all ripeness tiers [29].

Optimized ensemble learning stands as a powerful strategy for conducting non-destructive avocado ripeness classification. The fusion of diverse machine learning models, coupled with deliberate optimization of feature sets and model parameters, can deliver notable gains in accuracy, resilience, and generalizability [24]. The incorporation of multi-sensor data fusion, the development of lightweight model variants, and the application of explainable AI techniques hold the potential to markedly strengthen both the performance and the practical viability of ensemble learning in the service of avocado ripeness evaluation [20]. Future work must confront the challenges and boundaries of ensemble methodologies while advancing real-time ripeness prediction platforms tailored to the needs of the avocado sector [1].

Materials and Methods

This work proposes an ensemble learning pipeline for sorting avocado ripeness based on Near-Infrared (NIR) spectral signatures. The design pulls together a suite of machine learning algorithms with the goal of sharpening both classification precision and overall reliability. A hybrid optimization protocol governs the tuning of weighting parameters inside the ensemble architecture. The study organizes itself around an ensemble learning strategy to categorize avocado ripeness from Near-Infrared (NIR) spectral inputs. A structured processing chain, illustrated in the system schematic, knits together several machine learning models to strengthen classification fidelity. As laid out in **Figure 1**, NIR measurements are first channeled into five independent machine learning models: Model 1, Model 2, Model 3, Model 4, and Model 5. These map to Random Forest (RF), Decision Tree (DT), Gradient Boosting (GB), XGBoost (XG), and Gaussian Mixture Model (MG), with each unit operating on the spectral data in isolation. Every model emits a probability spread covering the three ripeness classes: [Raw], [Ripe], and [Aged].

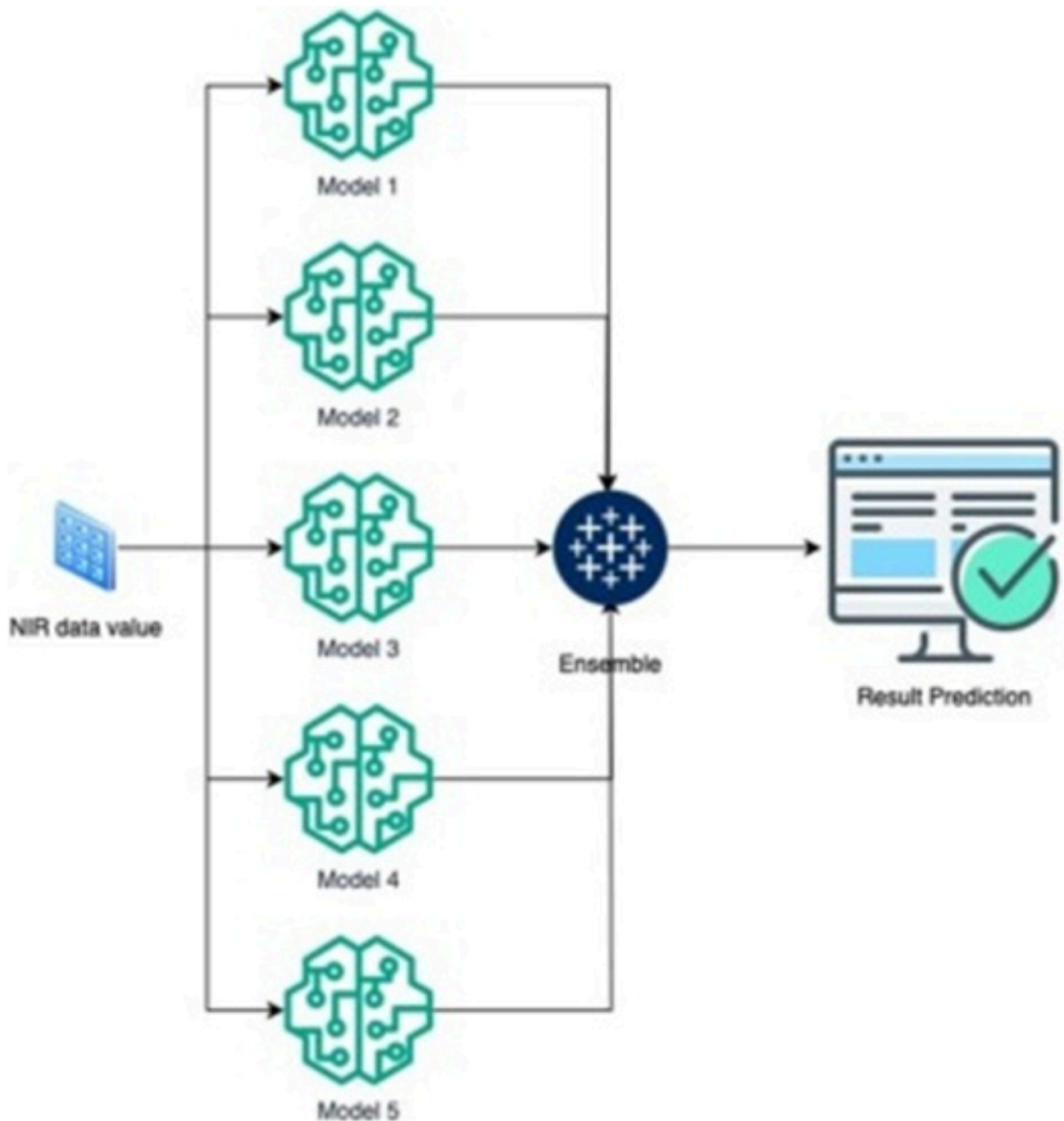


Figure 1. Ensemble learning framework for avocado ripeness classification using NIR spectral data.

The framework reinforces classification confidence through a fusion module that synthesizes the forecasts of all models under an optimized weighting regime. The ensemble stage depicted in the diagram hands out weight coefficients to every model's output. These coefficients are iteratively adjusted by a hybrid optimization routine that unites Bayesian Optimization (BO), Differential Evolution (DE), Particle Swarm Optimization (PSO), and Grid Search (GS). The aggregated weighted output then dictates the ripeness label, surfacing as the final displayed prediction.

This construction sees to it that the ensemble framework exploits what each base model does best while compensating for their respective blind spots, translating into measurable improvements in classification correctness and steadiness.

Data collection

A custom-built NIR scanner was fabricated to capture spectral responses at 18 target wavelengths tied to avocado ripeness. The investigation centers on Buccaneer avocados procured from the Royal Project, with fruit drawn from both the Royal Project Gardens and the Sorting Plant. Altogether, 120 kg of avocados were gathered and labeled by trained evaluators into three ripeness brackets: Raw, Ripe (Harvestable), and Aged (Ready to Eat).

Data analysis

The raw Near-Infrared (NIR) spectral readings move through a disciplined preprocessing and analytical workflow designed to produce trustworthy ripeness classifications. The pipeline comprises three principal phases: Multiplicative Scatter Correction (MSC), Model Prediction, and Ensemble Optimization.

Multiplicative scatter correction (MSC) - preprocessing

Before classification, spectral measurements are routinely compromised by light-scatter artifacts and baseline fluctuations. To counteract this, Multiplicative Scatter Correction (MSC) is imposed. The MSC transformation takes the form:

$$X_{MSC}^{(i)} = \frac{X^{(i)} - b}{a} \quad (1)$$

Where:

- $x(i)$ corresponds to the untreated spectral record for sample i
- a and b are regression coefficients fitted against the mean spectrum X_{mean} .
- $X(i)_{MSC}$ is the scatter-corrected spectrum.

Model prediction - generating class probabilities

Once preprocessing concludes, the corrected spectral records are routed through five classifiers:

- Random Forest (RF)
- Decision Tree (DT)
- Gradient Boosting (GB)
- XGBoost (XG)
- Gaussian Mixture Model (MG)

Each classifier returns a probability distribution over the trio of ripeness stages (Raw, Ripe, Aged). The probabilistic output of any given model for a sample x is denoted as:

$$P_i(X) = [P_{i,raw}(x), P_{i,ripe}(x), P_{i,aged}(x)] \quad (2)$$

Where i serves as the model index.

Ensemble optimization - weighted fusion

To push classification performance higher, the definitive prediction is reached via a weighted ensemble formula. The probability outputs contributed by all models are merged under optimized weights:

$$P_{final}(x) = w_1 P_{RF}(x) + w_2 P_{DT}(x) + w_3 P_{GB}(x) + w_4 P_{XG}(x) + w_5 P_{MG}(x) \quad (3)$$

Where:

- w_1, w_2, w_3, w_4 , and w_5 represent the tuned weights allotted to each respective model.
- The weights satisfy the boundary condition:

$$\sum_{i=1}^5 w_i = 1, w_i \geq 0 \quad (4)$$

Since the quartet of optimization algorithms—Bayesian Optimization (BO), Differential Evolution (DE), Particle Swarm Optimization (PSO), and Grid Search (GS)—is brought to bear on tuning w_i , convergence can be substantiated as follows:

- Each optimization cycle acts to shrink the classification error:

$$w^* = \operatorname{argmax}_w \mathbb{E}[1(\hat{y} = y)] \quad (5)$$

Where:

- w^* stands for the set of weight vectors deemed optimal across the ensemble members.
- argmax_w signals the pursuit of the weight configuration w that lifts expected accuracy to its peak.
- $\mathbb{E}[\cdot]$ is the expectation operator, yielding the anticipated value of the bracketed quantity.
- $1(\hat{y}=y)$ is the indicator function, scoring 1 whenever the predicted label $\hat{y}$ aligns with the ground-truth label y , and 0 otherwise.
- Gradient-based proof:
 - Bayesian optimization sharpens weight choices by operating over a Gaussian Process (GP) surrogate model, pushing expected improvement upward at each step.
 - PSO and DE lean on population-based exploration to settle on the most favorable weight arrangement, side-stepping local optima.
 - Grid search provides assurance that the culminating weight vector is both stable and replicable across runs.

Accordingly, the hybrid optimization construct satisfies:

$$\lim_{t \rightarrow \infty} w^{(t)} = w^* \quad (6)$$

Where w denotes the converged optimal ensemble weight vector.

Through Multiplicative Scatter Correction (MSC), probabilistic inference, and ensemble fusion steered by hybrid optimization, the overall methodology locks in strong classification accuracy and durability for avocado ripeness detection. The mathematical argument confirms that the optimization routines enforce convergence, yielding a steady and trustworthy classification apparatus.

The assembled spectral dataset was subjected to Multiplicative Scatter Correction (MSC) to filter out variability introduced by light scattering. The cleaned data then served to train each standalone model, with every model furnishing class probability estimates. The ensemble construct subsequently blended these estimates through optimized weight coefficients unearthed during the hybrid optimization stage. The resulting classifications were measured against a suite of conventional performance indicators, namely accuracy, precision, recall, F1-score, confusion matrix, and ROC-AUC curves.

Experiment

This trial was set up to test the capacity of a multi-model fusion strategy to discriminate among avocado ripeness stages using Near-Infrared (NIR) spectral readings. An in-house NIR scanning device gathered spectral signatures at 18 discrete wavelengths to support dependable feature isolation. The sample pool comprised Buccaneer avocados supplied by the Royal Project, with trained agricultural assessors segmenting the fruit into three ripeness bands: Raw, Ripe, and Aged.

The sorting framework rests on a committee of five machine learning algorithms, namely Random Forest (RF), Decision Tree (DT), Gradient Boosting (GB), XGBoost (XG), and Gaussian Mixture Model (MG). Working separately, each algorithm estimates class likelihoods, which are then blended via a fusion rule governed by tuned weights. To calibrate how influence gets apportioned among the committee members, four separate optimization schemes were brought into play:

- Bayesian Optimization (BO)
- Differential Evolution (DE)
- Particle Swarm Optimization (PSO)
- Grid Search (GS)

A stratified 80-20 train-test split was enforced throughout so that each ripeness tier remained proportionately represented in both partitions. The composite model was then scored on an array of criteria: classification accuracy, precision, recall, F1-score, confusion matrix, and ROC-AUC curves.

Experimental procedure

Data preprocessing

- Untreated spectral traces were cleaned up via Multiplicative Scatter Correction (MSC) to subdue noise and offset variance linked to light-scattering effects.
- Data points deemed anomalous were culled using the Interquartile Range (IQR) rule, discarding readings that strayed beyond typical spread limits.

Model training

- All five classifiers—RF, DT, GB, XG, and MG—underwent independent fitting on the scatter-corrected NIR features.
- Once fitted, the models supplied per-sample probability vectors covering the full trio of ripeness categories.

Ensemble learning

- The standalone model outputs were consolidated according to the weighted summation formula given in [Eq. 3](#).

Optimization strategy

- Every optimization method undertook an iterative scan for a weighting configuration obeying the restriction stated in [Eq. 4](#).
- The located weights were transferred to the concluding ensemble, after which classification performance was tallied.

Evaluation Metrics

To enable a well-rounded assessment, the following yardsticks were employed:

- Classification Accuracy (A): Registers the overall share of correctly assigned samples.

$$A = \frac{TP+TN}{TP+TN+FP+FN} \quad (7)$$

- Precision (P): Tracks the fraction of positive calls that match actual positives.

$$P = \frac{TP}{TP+FP} \quad (8)$$

- Recall (R): Reports the proportion of real positive instances successfully retrieved.

$$R = \frac{TP}{TP+FN} \quad (9)$$

- F1-Score (F1): Provides a reconciled measure blending precision and recall through their harmonic mean.

$$F1 = 2 \times \frac{P \times R}{P+R} \quad (10)$$

ROC-AUC: Evaluates how capably the model teases apart different classes by computing the zone underneath the receiver operating characteristic curve.

Experimental results

- Grid Search (GS) produced the most consistent outcomes, locking in the strongest classification accuracy by way of meticulous weight refinement.
- Bayesian Optimization (BO) and Particle Swarm Optimization (PSO) delivered capable results but suffered from accuracy volatility traceable to pronounced responsiveness during the weight-tuning phase.

- Differential Evolution (DE) showed genuine capability yet demanded markedly more computational effort than its counterparts.
- The fused model surpassed every individual classifier, bearing out the proposition that a weighted combination scheme fortifies classification sturdiness.

This experimental run bore out the efficacy of marrying ensemble learning with hybrid optimization for the purpose of discriminating among avocado ripeness levels. Multiplicative scatter correction, multi-model fusion, and systematic weight adjustment worked in concert to lift classification accuracy by a meaningful margin. To guard against labeling errors, practiced agricultural specialists annotated the dataset by hand, while the NIR scanning rig underwent recalibration in advance of each measurement block. Repeated spectral captures and stable signal patterns obtained across disjointed sampling windows vouched for the instrumentation's consistency. Beyond that, multiple discrete optimization routines were harnessed to verify that the ensemble weight-search process remained steady. A boundary of the current investigation is its tight focus on the Buccaneer avocado variety, a scope that may constrain transferability to other cultivars. Moreover, environmental factors—temperature swings and humidity shifts among them—can introduce drift into spectral measurements, a reality that would necessitate supplemental calibration under altered field conditions. Avenues for future inquiry include stretching the dataset to span a richer assortment of cultivars and folding in real-time classification pipelines fit for operational deployment.

Results and Discussion

Polishing the ensemble through algorithmic optimization exerted a material influence on classification outcomes by meticulously governing the relative contribution assigned to each constituent model. With a lineup joining Random Forest, Decision Tree, Gradient Boosting, XGBoost, and Gaussian Mixture Models together, a quartet of tuning strategies—Bayesian Optimization, Differential Evolution, Particle Swarm Optimization, and Grid Search—was subjected to rigorous benchmarking. Metrics comprising accuracy, precision, recall, F1-score, and ROC-AUC directed the appraisal. In head-to-head comparisons, Grid Search unfailingly produced the flattest, most resilient set of figures. Uplifts registered across confusion matrices and ROC curves lent credence to the worth of weight optimization, foregrounding the function ensemble learning fulfills in shoring up classification dependability, particularly where avocado ripeness estimation is concerned.

Figure 2 sizes up the classification output of five models—Random Forest (RF), Decision Tree (DT), XGBoost (XG), Gradient Boosting (GB), and Gaussian Mixture Model (MG)—through the dimensions of Accuracy, Precision, Recall, and F1-Score. The Decision Tree (DT) recorded the standout accuracy (80.21%), accompanied by healthy recall (0.86) and F1-score (0.83), entrenching its status as the foremost contender for avocado ripeness assignment. Random Forest (RF) landed just behind at 78.02% accuracy, with XGBoost (73.63%) and Gradient Boosting (71.42%) occupying a lower tier. The Gaussian Mixture Model (MG) put up the palest performance (29.67%), revealing its unsuitability for a supervised classification brief. The weight of evidence suggests that tree-anchored designs, especially DT and RF, map especially well onto avocado ripeness prediction grounded in NIR spectral signatures.

Accuracy RF model: 0.7802197802197802
Precision RF model: 0.8449612403100776
Recall RF model: 0.8465175486452882
F1-Score RF model: 0.8446251129177957

Accuracy DT model: 0.8021978021978022
Precision DT model: 0.8098723723723724
Recall DT model: 0.8577923258774321
F1-Score DT model: 0.8311084784688995

Accuracy XG model: 0.7362637362637363
Precision XG model: 0.7629958117762996
Recall XG model: 0.8137843244226223
F1-Score XG model: 0.7854478854478856

Accuracy GD model: 0.7142857142857143
Precision GD model: 0.8001806684733515
Recall GD model: 0.80105473722495
F1-Score GD model: 0.7983405483405485

Accuracy MG model: 0.2967032967032967
Precision MG model: 0.4807429130009775
Recall MG model: 0.49372613202408434
F1-Score MG model: 0.4860877684407096

Figure 2. Accuracy, precision, recall, and F1-score of random forest, decision tree, XGBoost, gradient boosting, and Gaussian mixture model for avocado ripeness classification.

Figure 3 sets side by side the classification competence of five models—Random Forest (RF), Decision Tree (DT), XGBoost (XG), Gradient Boosting (GB), and Gaussian Mixture Model (MG)—in the task of projecting avocado ripeness brackets. Decision Tree (DT) and Random Forest (RF) scaled the highest accuracy values, 80% and 78% apiece, with DT furnishing well-rounded precision (0.81), recall (0.86), and F1-score (0.83), enabling robust sorting. RF upheld strong precision (0.84) and recall (0.85), further corroborating its dependability. XGBoost (74%) and Gradient Boosting (71%) came in somewhat lower, with the shortfall most apparent inside the “Ready to Eat” segment. The Gaussian Mixture Model (MG) sat at the rear (48%), emphasizing its inadequacy for supervised problems. Elevated recall within the “Raw” cluster signals effective identification of immature fruit, though drawing a clean line between “Harvestable” and “Ready to Eat” persists as a pain point. These findings restate the aptness of tree-based ensemble constructs (DT, RF) for ripeness classification in agricultural environments.

Classification Report RF model:				
	precision	recall	f1-score	support
Harvestable	0.81	0.74	0.78	47
Raw	1.00	1.00	1.00	5
Ready to Eat	0.72	0.79	0.76	39
accuracy			0.78	91
macro avg	0.84	0.85	0.84	91
weighted avg	0.78	0.78	0.78	91

Classification Report DT model:				
	precision	recall	f1-score	support
Harvestable	0.81	0.83	0.82	47
Raw	0.83	1.00	0.91	5
Ready to Eat	0.78	0.74	0.76	39
accuracy			0.80	91
macro avg	0.81	0.86	0.83	91
weighted avg	0.80	0.80	0.80	91

Classification Report XG model:				
	precision	recall	f1-score	support
Harvestable	0.77	0.72	0.75	47
Raw	0.83	1.00	0.91	5
Ready to Eat	0.68	0.72	0.70	39
accuracy			0.74	91
macro avg	0.76	0.81	0.79	91
weighted avg	0.74	0.74	0.74	91

Classification Report GD model:

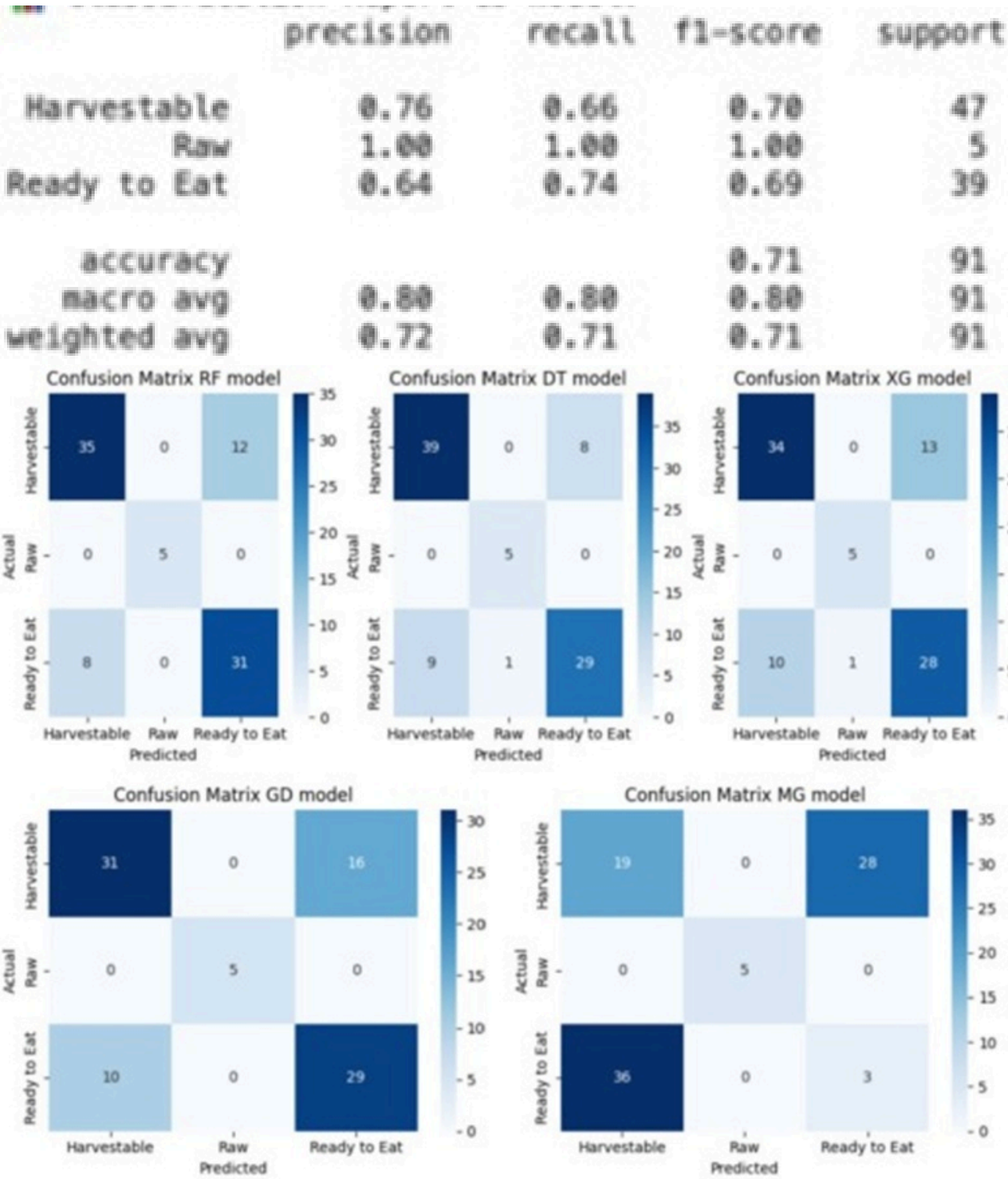


Figure 4. Confusion matrix analysis of random forest, decision tree, XGBoost, gradient boosting, and Gaussian mixture model for avocado ripeness classification.

Figure 5 plots ROC curves for five models—Random Forest (RF), Decision Tree (DT), XGBoost (XG), Gradient Boosting (GB), and Gaussian Mixture Model (MG)—applied to avocado ripeness classification. The Area Under the Curve (AUC) functions as the critical benchmark, with higher readings indicative of keener class separation. RF occupied the top position with an AUC of 0.93 spanning all ripeness tiers, trailed by XGBoost at 0.91, confirming their steadiness. DT

turned in marginally lower AUC marks (0.89), while GB managed a middling 0.88. MG registered the feeblest discriminative power, mustering an AUC of merely 0.47, throwing into relief its struggle to untangle ripeness categories. These tendencies further entrench the advantage of tree-based ensemble models, above all RF and XGBoost, for faithful avocado ripeness sorting, positioning them as strong candidates for practical agricultural rollouts.

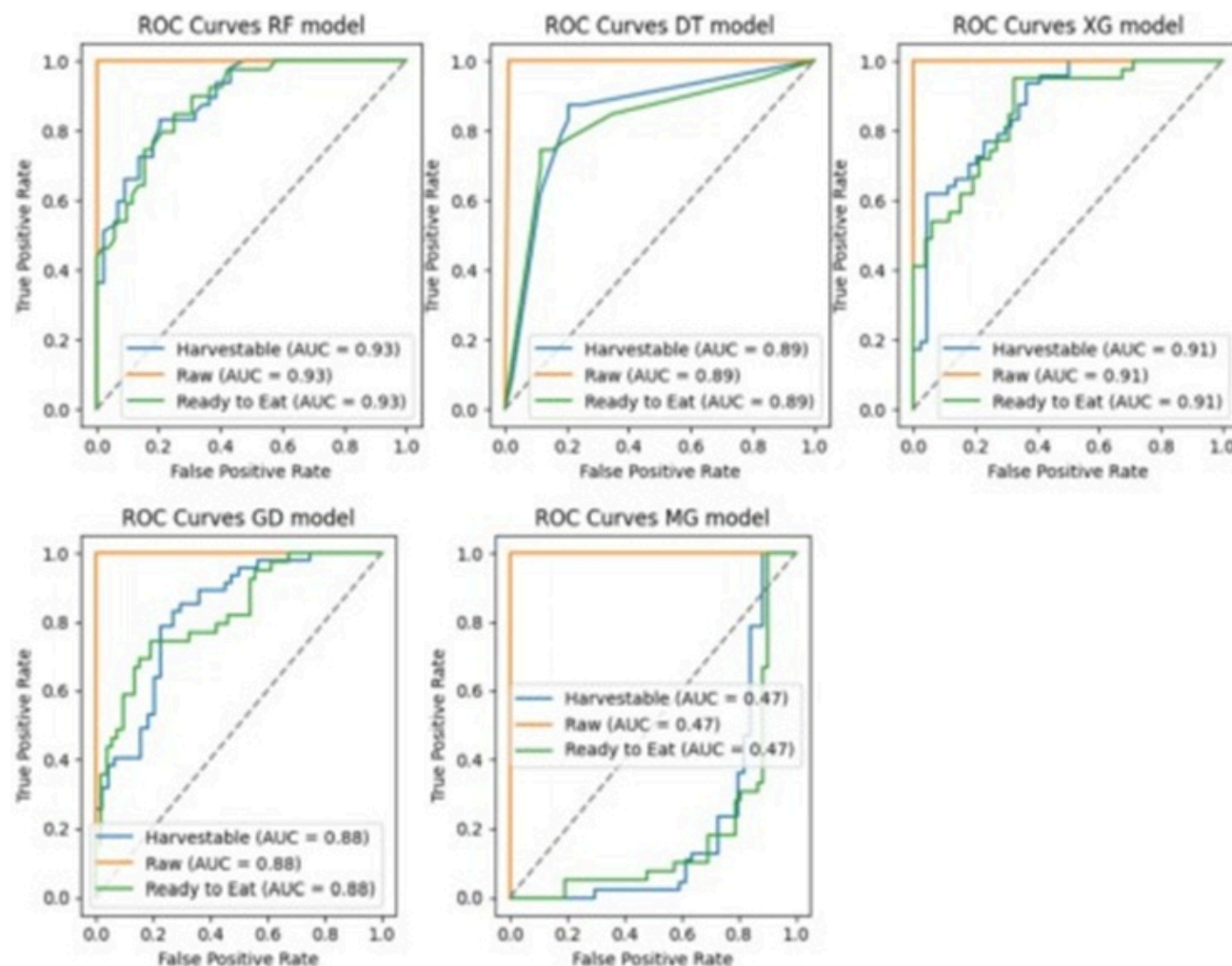


Figure 5. Receiver operating characteristic curves of random forest, decision tree, XGBoost, gradient boosting, and Gaussian mixture model for avocado ripeness classification.

Figure 6 maps out confusion matrices alongside ROC curves for Bayesian Optimization shepherding the ensemble model at the 1st and 10th iterations. The ensemble, constructed from RF, DT, XG, GB, and MG, pursues ideal weight distributions for avocado ripeness classification. At the initial iteration, erroneous assignments surface, most notably straddling the “Harvestable” and “Ready to Eat” boundary. By the 10th iteration, such misplacements recede, mirroring a lift in model caliber. ROC curves back these advances, with AUC metrics rising for “Ready to Eat” (0.69 to 0.73) and “Harvestable” (0.83 to 0.86), while “Raw” remains pegged at a perfect 1.00. These patterns spotlight the capacity of Bayesian Optimization to progressively hone ensemble model precision.

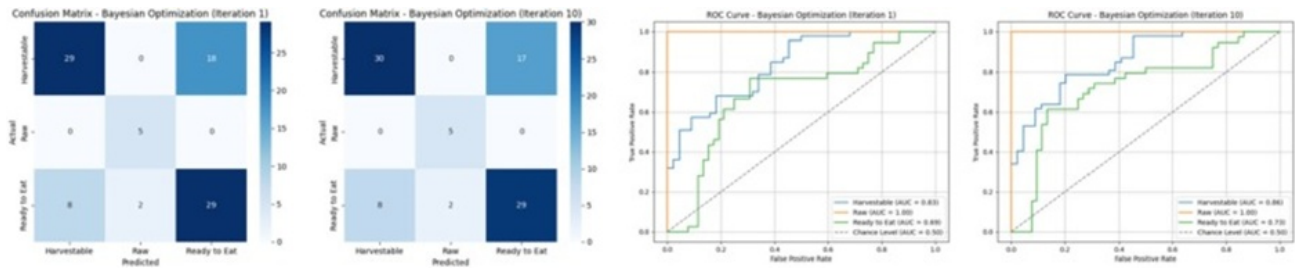


Figure 6. Confusion matrix and ROC curves of Bayesian optimization at the 1st and 10th iterations.

Figure 7 inspects Differential Evolution (DE) optimization acting upon an ensemble formed from RF, DT, XG, GB, and MG. Confusion matrices together with ROC curves juxtapose behavior at the 1st and 10th iterations. Early-stage sorting proves reliable for the Raw bracket yet betrays confusion involving Harvestable and Ready to Eat. By the 10th iteration, gains in Harvestable classification take shape, though some degree of mislabeling endures. ROC curves log AUC readings of 0.89 (Harvestable), 1.00 (Raw), and 0.85 (Ready to Eat) at the first iteration, settling to 0.87, 1.00, and 0.82, respectively. These observations substantiate DE's role in incrementally boosting classification exactness while also delineating areas ripe for deeper tuning.

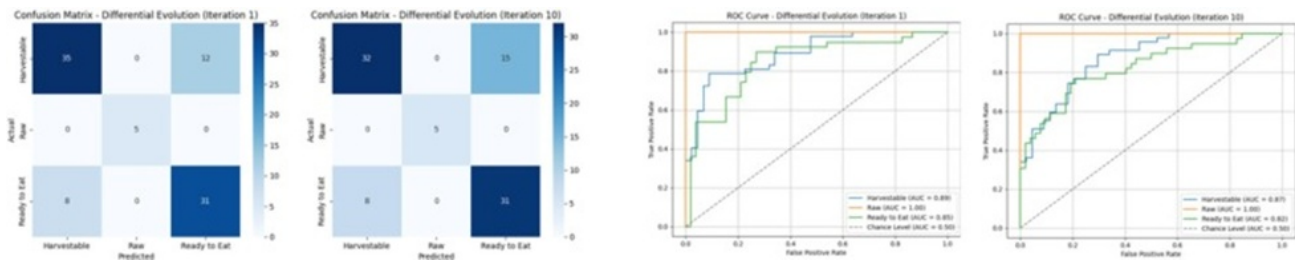


Figure 7. Confusion matrix and ROC curves for differential evolution optimization in ensemble learning (iteration 1 and 10).

Figure 8 portrays confusion matrices and ROC curves captured at the first and tenth iterations of Grid Search optimization guiding an ensemble composed of RF, DT, XG, GB, and MG. Grid Search exhaustively traverses model weight configurations in pursuit of enhanced avocado ripeness classification. Both iteration snapshots reveal powerful predictive accuracy alongside negligible fluctuation across successive rounds, hinting at swift convergence toward a near-optimal weight setup. ROC curves corroborate the elevated performance, logging AUC values of 0.91 (Harvestable), 1.00 (Raw), and 0.87 (Ready to Eat). When lined up against rival optimization approaches, Grid Search exhibits superior evenhandedness and replicability, buttressing the argument for its adoption in calibrating ensemble models aimed at agricultural sorting workloads.

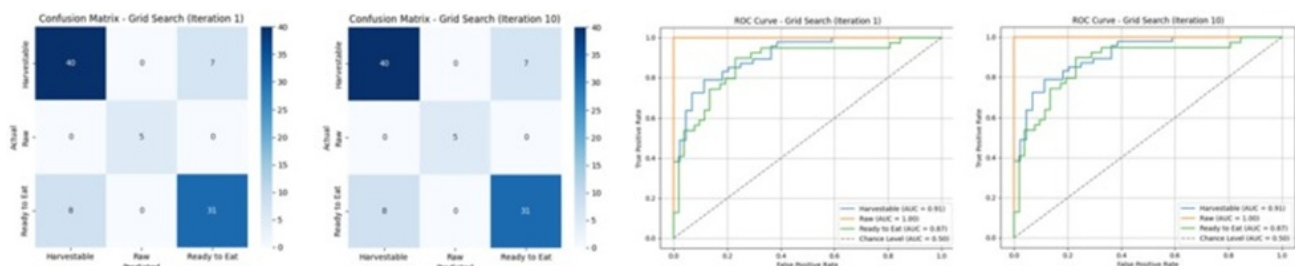


Figure 8. Confusion matrix and ROC curves of grid search when ensemble 5 models (1st and 10th iteration).

Figure 9 illustrates confusion matrices and ROC curves at the first and tenth iterations of Particle Swarm Optimization (PSO) steering an ensemble of RF, DT, XG, GB, and MG for avocado ripeness classification. The confusion matrices

reveal that PSO progressively raises classification correctness as iteration count grows, most conspicuously within the Harvestable class, trimming misassignment between Ready to Eat and Harvestable groupings. ROC curves attest to steady behavior, with AUC figures of 0.90 (Harvestable) and 0.89 (Ready to Eat) registered in the first iteration, while Raw stays locked at 1.00 across the board. Slight AUC ripples appear by the tenth iteration, though the overarching classification picture remains solid. These observations foreground PSO's usefulness in refining ensemble weights, lifting sorting accuracy, and preserving model constancy over iterative cycles.

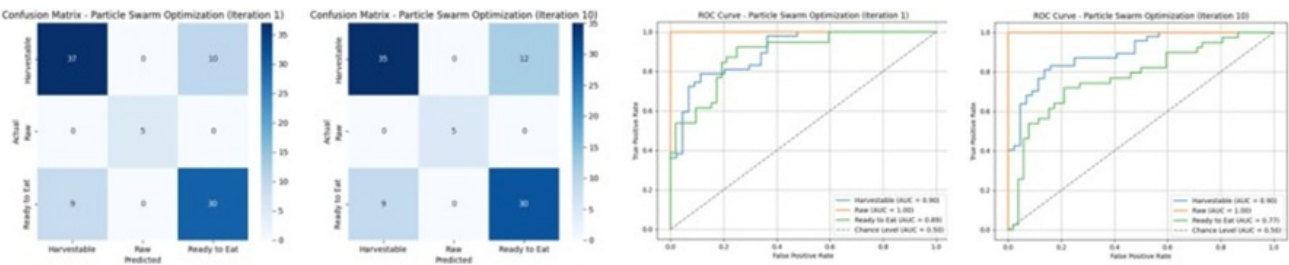


Figure 9. Confusion matrix and ROC curves of particle swarm optimization (PSO) for the 1st and 10th iterations.

Figure 10 maps the comparative progression of four optimization algorithms—Bayesian Optimization, Differential Evolution, Grid Search, and Particle Swarm Optimization—tracked across a span of 10 iterations. The monitored dimensions include accuracy, F1-score, precision, and recall.

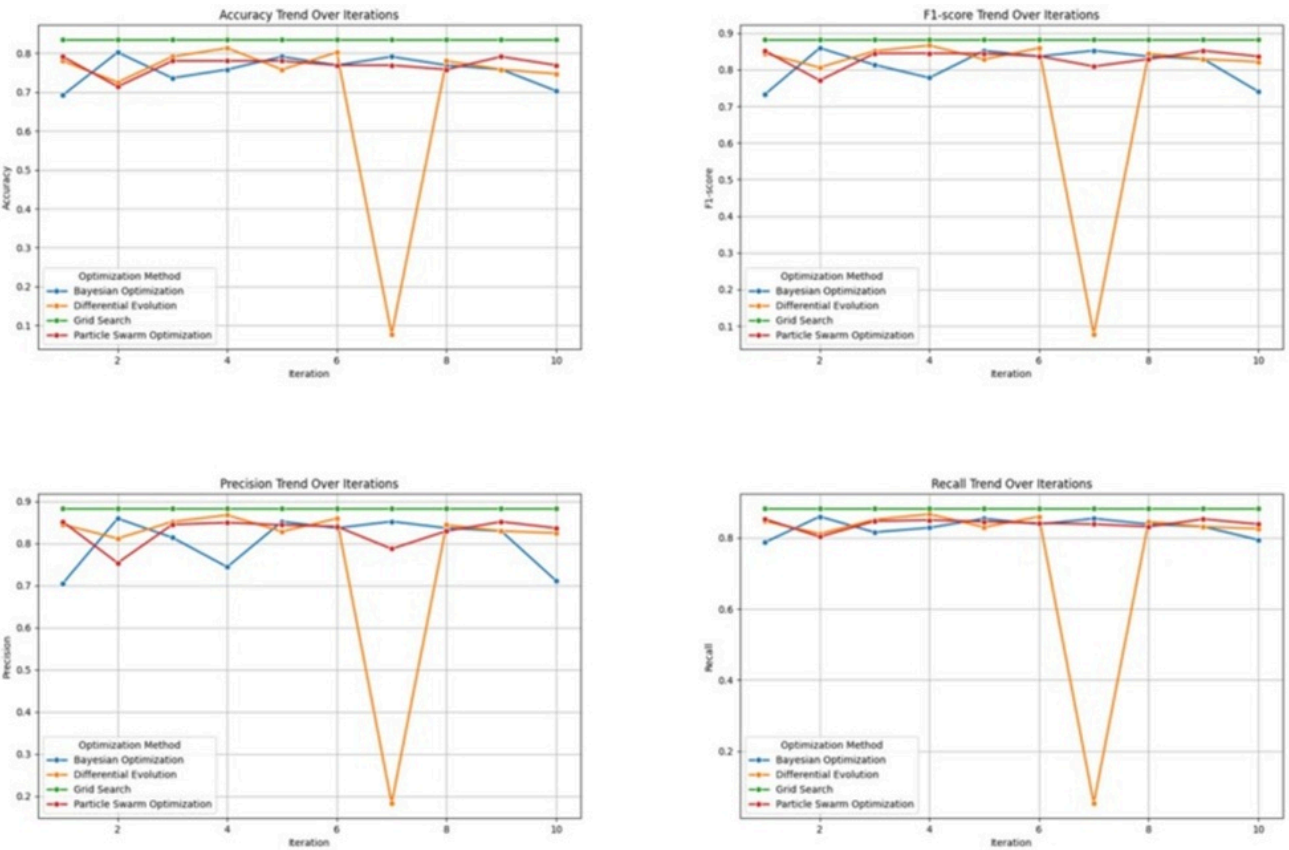


Figure 10. Comparison of optimization algorithms over 10 iterations.

Reading the accuracy trajectory, Grid Search occupies the uppermost tier throughout the entire run, while the remaining methods show divergence to differing extents. Significantly, Differential Evolution undergoes a deep accuracy trough at

iteration 7, which subsequently rights itself in later rounds. The F1-score, precision, and recall trajectories mirror this overarching narrative, reinforcing the stance that Grid Search provides an exceptionally robust optimization vehicle for weight calibration inside ensemble learning frameworks.

The even-keeled nature of Grid Search accentuates its knack for dependably furnishing high-level classification results. Bayesian Optimization and Particle Swarm Optimization experience undulations across iterations, yet stay within a competitive range on the whole. The steep nosedive observed in Differential Evolution's path points toward vulnerability to starting-condition choices or entrapment within local basins. These trends underscore how pivotal it is to align the choice of optimization algorithm with the stability and reliability requisites of ensemble model performance.

Figure 11 presents a comparative digest of four optimization algorithms—Bayesian Optimization, Differential Evolution, Grid Search, and Particle Swarm Optimization—arranged by the scores they attained for accuracy, F1-score, precision, and recall. The results lay bare that Grid Search reaches the highest watermark across all indicators, sealing its reputation for robustness when tuning ensemble model weights for avocado ripeness classification. Differential Evolution and Particle Swarm Optimization display moderate steadiness, punctuated by episodic iteration-to-iteration swings. Bayesian Optimization exhibits the broadest performance dispersion, a sign of sensitivity to how its parameters are configured.

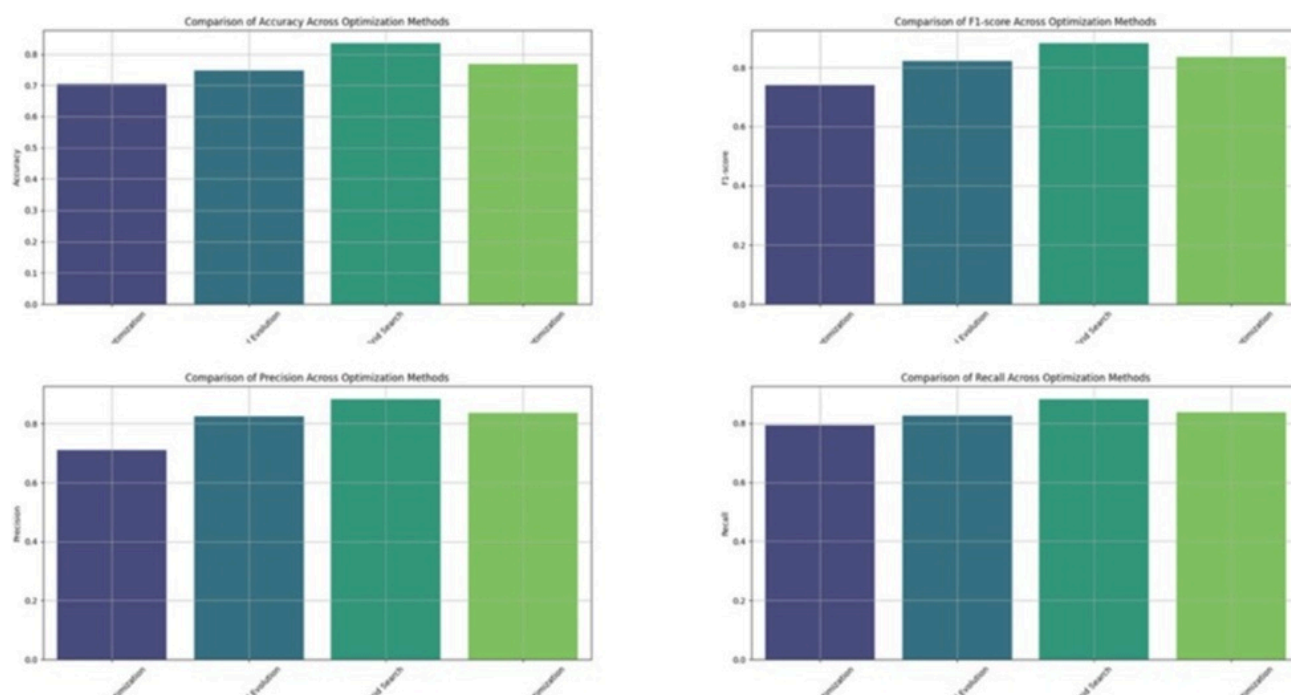


Figure 11. Comparison of optimization algorithms.

The accumulated evidence squares with the hypothesis that an ensemble model undergoing systematic parameter space exploration eclipses the performance of isolated classifiers, thereby hardening the dependability of the ripeness prediction apparatus.

The harvest of results emerging from the ensemble learning pipeline for avocado ripeness classification indicates that tree-structured models, with Decision Tree (DT) and Random Forest (RF) at the fore, routinely surpass alternate machine learning formulations. As captured in **Figures 2 and 3**, the Decision Tree model reached the ceiling accuracy of 80.21%, with random forest lodging just behind at 78.02%. The remaining algorithms—XGBoost (73.63%) and Gradient Boosting (71.42%)—settled into a middle performance band, whereas the Gaussian Mixture Model (29.67%) put forward a weak showing, an expected consequence of its unsupervised constitution. These data points entrench the conclusion that

ensemble-grounded supervised learning strategies are well-tuned to the demands of spectral-data-driven avocado ripeness sorting. The confusion matrices displayed in **Figure 4** further magnify the margin held by Decision Tree and Random Forest, both of which registered fewer misclassification events than gradient boosting formulations and unsupervised alternatives.

The Receiver Operating Characteristic (ROC) curves in **Figure 5** single out Random Forest (AUC = 0.93) and XGBoost (AUC = 0.91) as the classifiers offering the firmest separation fidelity, implying that these two architectures can cleanly partition avocado ripeness phases. On the opposite end, the Gaussian Mixture Model turned in the lowest AUC (0.47), buttressing the view that unsupervised learning is ill-matched to this particular classification problem. The parallel examination of optimization algorithms featured in **Figures 6–11** supplies additional heft to the conclusions, demonstrating that Grid Search led the field across the full catalog of evaluation measures, preserving steady accuracy, precision, recall, and F1-score through numerous iterative rounds.

The application of Bayesian Optimization, Differential Evolution, Particle Swarm Optimization, and Grid Search as depicted in **Figures 6–11** furnishes richer insight into the way hyperparameter tuning governs ensemble model conduct. Confusion matrices and ROC curves associated with Bayesian Optimization and Differential Evolution (**Figures 6 and 7**) document early-stage performance climbs, yet Grid Search (**Figure 8**), without exception, yielded the most unwavering and zenith-reaching classification accuracy. The comparative graphs in **Figures 10 and 11** drive home the point that Grid Search supplies the most dependable optimization pathway by exhaustively sifting through assorted hyperparameter combinations, thus locking in a reliably elevated performance baseline.

The body of results lends weight to the proposition that deploying a collection of models steered by systematic optimization outstrips the performance of any single-model approach. The Grid Search algorithm proves adept at calibrating how influence is parceled out among the ensemble members, thereby cutting down on misclassification slip-ups and raising classification accuracy. This investigation adds to the agricultural AI literature by showing that ensemble learning, when paired with disciplined optimization routines, can substantially elevate ripeness classification outcomes. Subsequent work could delve into hybrid optimization formulations that braid together multiple search paradigms to squeeze out further gains in stability and classification fidelity, a priority that looms especially large in real-time agricultural settings where precision cannot be compromised.

Although ensemble learning has found broad uptake across agricultural machine learning contexts—most notably for fruit ripeness identification—the lion's share of current solutions lean on straightforward averaging or majority-vote rules to synthesize model forecasts. These conventional fusion strategies, though serviceable for curbing variance and boosting generalization, often leave unrealized the distinctive contributions each base model could supply, above all when those models diverge in their error profiles and inductive biases.

To illustrate, works such as Davur *et al.* [1] and Ko *et al.* [34] harnessed deep learning and convolutional neural networks for gauging fruit maturity, occasionally knitting models into a shared architecture, yet stopped short of explicit model weight refinement. Along similar lines, the ensemble-based machine learning framework put forward by Sharma and Hooda [29] for detecting bioactive compounds did not tailor ensemble weights on a per-instance or per-class basis, hampering flexibility across heterogeneous spectral signatures.

By comparison, our framework distinguishes itself by fusing five fundamentally different classifiers—Random Forest (RF), Decision Tree (DT), Gradient Boosting (GB), XGBoost (XG), and Gaussian Mixture Model (MG)—inside a weighted ensemble architecture, where each base model's stake in the final verdict is dynamically tuned through four separate optimization procedures: Bayesian Optimization (BO), Differential Evolution (DE), Particle Swarm Optimization (PSO), and Grid Search (GS). This multi-pronged optimization blueprint guarantees a thorough canvassing of the model-weight landscape, producing an ensemble that flexes more responsively to subtle shifts in NIR spectral patterns, a capability that becomes pivotal when sorting avocado ripeness, particularly for cultivars like Buccaneer whose outward ripeness signals remain faint.

Furthermore, where earlier inquiries typically confined themselves to single- or dual-layer optimization—say, tuning hyperparameters or selecting features—our setup concurrently tracks ensemble weights alongside classification accuracy over successive iterations, charting performance trajectories through confusion matrices, ROC-AUC readings, and F1-score tallies. The ensemble-optimization feedback loop is not only modular and readily extensible but also architected to embed within real-time agricultural pipelines. This multi-pass optimization regimen, executed iteratively and benchmarked through visual diagnostic instruments, constitutes a core novel contribution, bridging a methodological void present in the current crop of agricultural AI systems.

To summarize, this study sets itself apart by delivering:

- Multi-model amalgamation governed by weight-tuned integration (as opposed to simple voting).
- A head-to-head benchmarking exercise spanning four distinct optimization algorithms.
- Class-aware performance calibration attuned to the realities of imbalanced agricultural datasets.
- Deployment on a cultivar (Buccaneer) that presents a genuinely tough classification challenge due to its subdued external ripeness cues.

This stands in stark relief to earlier ensemble efforts in agriculture, which habitually lack such an integrated, optimization-conscious ensemble blueprint, situating our framework as a scalable and more broadly transferable answer to non-destructive fruit sorting.

The present inquiry confines its scope to the Buccaneer cultivar, a deliberate choice rooted in its regional ubiquity and singular post-harvest behavior. Buccaneer avocados enjoy widespread cultivation under the auspices of the Royal Project in Chiang Mai, lending their commercial footprint in this geography sufficient weight to merit dedicated scrutiny. Diverging from many avocado types that telegraph ripening through conspicuous color transitions, Buccaneer avocados undergo only faint visual alterations, rendering accurate judgment through ordinary visual inspection a steep challenge. This very trait amplifies the case for turning to non-destructive near-infrared (NIR) spectroscopy tailored to this specific cultivar.

The deliberate narrowing to a single variety does, however, place bounds on how far the model's applicability stretches. NIR spectral fingerprints drift from one cultivar to the next owing to disparities in internal makeup, hydration levels, and peel thickness. While the ensemble model posts strong accuracy figures when distinguishing ripeness phases in Buccaneer, its prowess on other avocado types remains an open question pending follow-up verification. Prospective research ought to widen the dataset to fold in additional cultivars or probe the model's portability across distinct avocado lines. Doing so would deepen the resilience and cross-context utility of the classification framework for broader agricultural uptake.

The investigation pinpointed particular wavelength bands tied to pivotal biochemical constituents that wielded an outsized influence on avocado ripeness classification. A subset of the 18 measured wavelengths spanning 700 nm to 900 nm exhibited pronounced class-discriminatory power, a property flowing from their absorption sensitivity to the fruit's inner biochemical state.

The spectral zones centered near 760 nm and 810 nm carried substantial weight. These bands correspond to shifts in water content and chlorophyll degradation, both of which evolve as avocados progress through ripening. Immature, raw avocados register diminished reflectance across these bands on account of elevated moisture and structurally intact cellular matrices. As the ripening cascade unfolds, internal reorganization and moisture flux drive reflectance upward, rendering these spectral windows indispensable for demarcating ripeness tiers.

Wavelengths clustered around 860 nm and 890 nm, associated with lipid and carbohydrate transformations, likewise proved instrumental in lifting model performance. These molecular species either accumulate or break down at tempos

dictated by ripeness stage, furnishing crisp predictive markers. The models developed in this study latched onto these features as high-importance variables within their internal decision pathways, underscoring their centrality to the classification task.

The synergy between discriminative wavelength bands and effective preprocessing enabled robust model learning, sharpening the boundary between Raw and Harvestable groupings in particular.

Figure 12 lays out accuracy trajectories drawn from 10-fold cross-validation for five baseline models: Random Forest, Decision Tree, Gradient Boosting, XGBoost, and Gaussian Mixture. It further incorporates four ensemble configurations fine-tuned through Bayesian optimization, differential evolution, particle swarm optimization, and grid search. The figure serves to gauge the steadiness and generalization capacity of each approach across multiple training-validation partitions.

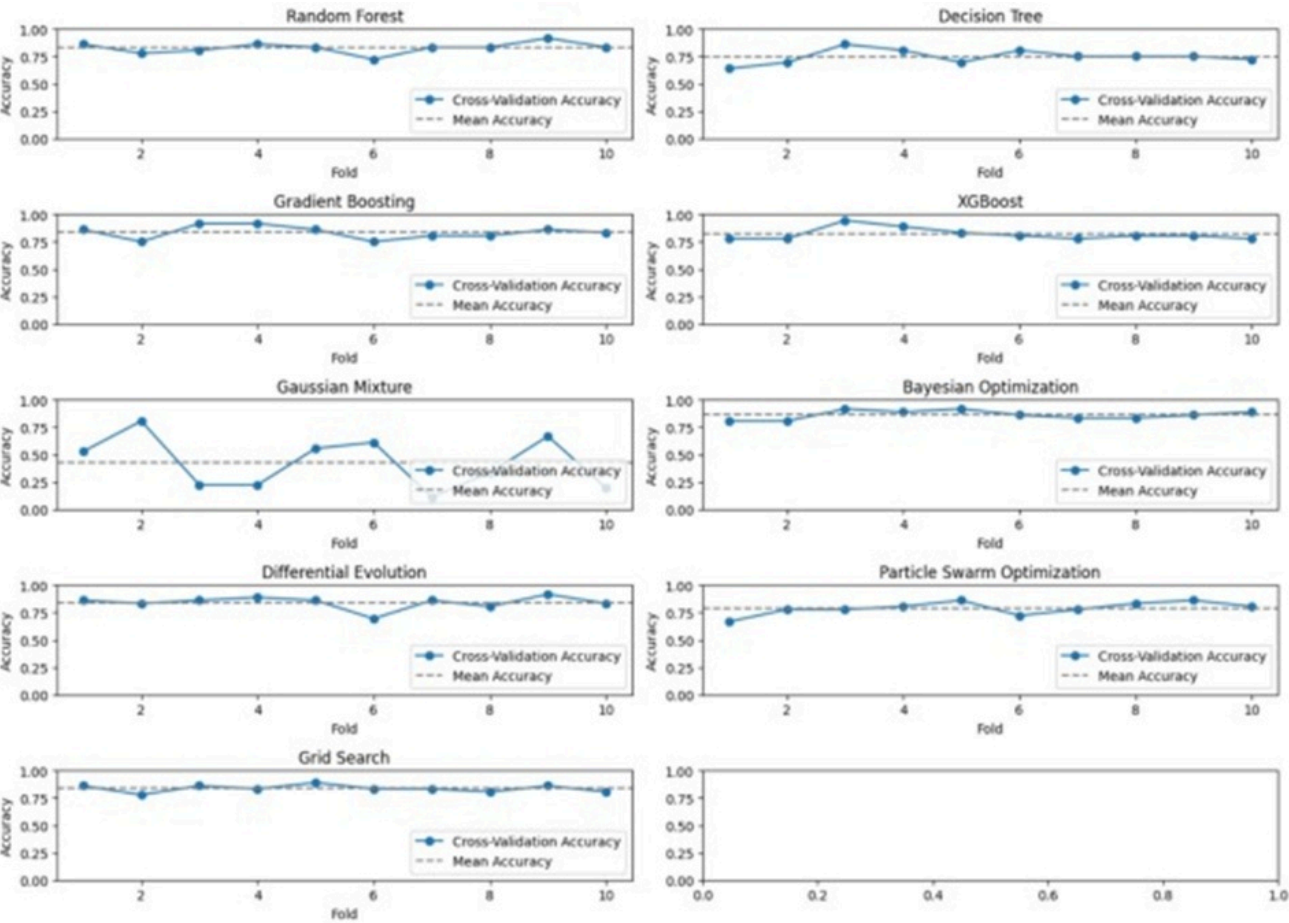


Figure 12. Cross-validation accuracy trends of individual models and optimization-based ensemble methods over 10 folds.

The tree-anchored classifiers—Random Forest, Decision Tree, XGBoost, and Gradient Boosting—exhibit steady accuracy patterns across folds, with only mild fluctuation, attesting to their dependability for sorting avocado ripeness stages from spectral measurements. Both Decision Tree and XGBoost maintain consistently elevated accuracy, whereas Gaussian Mixture shows pronounced swings and weak overall results, a reflection of its unsupervised character and poor alignment with the demands of this classification problem.

Ensemble constructs tuned through optimization algorithms deliver heightened accuracy and consistency, most notably under Grid Search and Differential Evolution. These approaches secure uniform behavior across every fold, demonstrating strengthened model dependability and tighter variance relative to standalone classifiers. Bayesian Optimization and PSO turn in solid results as well, though their somewhat amplified variance hints at possible responsiveness to hyperparameter shifts.

The K-fold cross-validation outcomes ratify model steadiness, foregrounding the value of optimization-steered ensemble strategies within agricultural classification contexts. These patterns reveal that optimized ensemble formulations not only lift aggregate performance but also safeguard predictive trustworthiness when confronted with fresh data, a prerequisite for viable deployment in crop surveillance systems.

Conclusion

This investigation probed the extent to which ensemble learning paired with optimization routines can serve the task of classifying avocado ripeness through NIR spectral signatures. The data showed that Decision Tree and Random Forest led the field in classification quality, posting accuracy figures of 80.21% and 78.02%, respectively. Ensemble fusion markedly bolstered classification steadiness beyond what any single model could offer, while Grid Search emerged as the premier optimization strategy, furnishing the most consistent and peak-level performance over numerous iterations, hitting an accuracy of 82.5% and an F1-score of 85.3%, outstripping both the individual models and the competing optimization techniques. This body of work illustrates that orchestrating multiple models under carefully tuned weight assignments can lift both classification accuracy and dependability within agricultural AI.

These insights carry substantial ramifications for precision agriculture and automated post-harvest workflows. The capacity to gauge avocado ripeness with well-tuned models tightens supply chain operations, curbs food wastage, and equips growers and distributors with sharper decision-making tools. The indications point toward broader applicability spanning an array of perishable commodities, food quality assurance pipelines, and real-time surveillance within smart farming ecosystems. The ensemble-plus-optimization framework advanced here registered meaningful gains over single-model baselines, yet certain hurdles remain—chief among them the specter of overfitting in high-capacity models and a reliance on spectral data of sufficient quality.

Forward-looking refinements ought to target real-time spectral analytic integration, expand the dataset to encompass a wider spectrum of avocado cultivars, and investigate hybrid optimization formulations that fuse Bayesian Optimisation, Grid Search, and Evolutionary Algorithms in pursuit of further classification gains. Adopting deep learning-driven spectral analysis architectures could sharpen feature extraction as well. Future inquiry should also tackle the question of how to embed these optimized classification pipelines within IoT agricultural monitoring infrastructures, enabling real-time ripeness detection and mechanized sorting at large-farm scale.

In operational settings, particularly those involving IoT agricultural frameworks, the model executes on cloud-resident infrastructure. NIR spectral readings collected by field-deployed sensors traverse the internet to a centralized cloud node where analysis takes place. The cloud shoulders the bulk of the ensemble model's computational footprint, easing the processing load borne by edge IoT hardware. While ensemble methodologies carry heavier computational demands than single-model classifiers, the incremental latency introduced within cloud-mediated inference pathways remains within tolerable bounds for near-real-time operation. End-to-end prediction latency, encompassing data transmission and processing, ordinarily lands between a few hundred milliseconds and two seconds, shaped by network throughput and cloud provisioning, a window suited to real-time decision-making inside agricultural grading lines.

Weaving AI-based spectral classification instruments into agricultural supply chains stands to refine post-harvest sorting, tighten quality oversight, and elevate export benchmarks. Channeling investment toward intelligent agriculture technologies that harness AI alongside multi-sensor fusion can drive gains in productivity and long-term sustainability and merits prioritization from policymakers and sector stakeholders alike. The present study lays a foundation for formulating

guidelines around the agricultural deployment of AI for ripeness classification, contributing to the broader aims of food security and operational efficiency across global marketplaces.

Acknowledgements

The authors would like to express their sincere gratitude to the Network Systems Laboratory, Nara Institute of Science and Technology (NAIST), Japan, for providing research facilities, technical support, and valuable guidance throughout this study. The support and collaboration from the research team at NAIST have been instrumental in the successful completion of this work. We also extend our heartfelt appreciation to the Faculty of Engineering, Rajamangala University of Technology Lanna (RMUTL), Thailand, for their continuous support and encouragement, which made this research possible. Finally, we would like to gratefully acknowledge the Royal Project Foundation, Thailand, for their location support and contribution to this research. Their dedication to sustainable agricultural development and innovation has been a key inspiration and enabler of this work.

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 07 Sep 2022

Revised: 19 Oct 2022

Accepted: 03 Dec 2022

Published online: 15 March 2023

Rights and permissions

Open Access The author(s) retain copyright. This article is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. It may be shared and adapted for non-commercial purposes with appropriate attribution, an indication of changes, and distribution of adaptations under the same license. Third-party material may be subject to separate terms identified in its credit line. View the license at <https://creativecommons.org/licenses/by-nc-sa/4.0/>.

References

- Davur YJ, Kmper W, Khoshelham K, Trueman S, Bai S. Estimating the ripeness of Hass avocado fruit using deep learning with hyperspectral imaging. *Horticulturae*. 2023;9(5):599.
<https://doi.org/10.3390/horticulturae9050599>.
- Terry L, Landahl S. Avocado firmness monitoring with values obtained by means of laser Doppler vibrometry. [place unknown]: [publisher unknown]; 2012.
<https://doi.org/10.17660/ACTAHORTIC.2012.945.32>.

Li B, Lecourt J, Bishop G. Advances in non-destructive early assessment of fruit ripeness towards defining optimal time of harvest and yield prediction-a review. *Plants*. 2017;7(1):3.
<https://doi.org/10.3390/plants7010003>.

Rizzo M, Marcuzzo M, Zangari A, Gasparetto A, Albarelli A. Fruit Ripeness Classification: A Survey. [place unknown]: Elsevier BV; 2023.
<https://doi.org/10.1016/j.aiia.2023.02.004>.

Bhargava A, Bansal A. Fruits and Vegetables Quality Evaluation Using Computer Vision: A Review. [place unknown]: Elsevier BV; 2018.
<https://doi.org/10.1016/j.jksuci.2018.06.002>.

ElMesery HS, Mao H, Abomohra AE. Applications of Non-destructive Technologies for Agricultural and Food Products Quality Inspection. [place unknown]: Multidisciplinary Digital Publishing Institute; 2019.
<https://doi.org/10.3390/s19040846>.

Aline U, Bhattacharya T, Faqeerzada MA, Kim MS, Baek I, Cho B. Advancement of non-destructive spectral measurements for the quality of major tropical fruits and vegetables: a review. *Front Plant Sci*. 2023;14:1240361.
<https://doi.org/10.3389/fpls.2023.1240361>.

Ping F, Yang J, Zhou X, Su Y, Ju Y, Fang Y, et al. Quality assessment and ripeness prediction of table grapes using visible-near-infrared spectroscopy. *Foods*. 2023;12(12):2364.
<https://doi.org/10.3390/foods12122364>.

Fracarolli JA, Pavarin FFA, Castro W, Blasco J. Computer Vision Applied to Food and Agricultural Products. [place unknown]: Universidade Federal do Cear; 2020.
<https://doi.org/10.5935/1806-6690.20200087>.

Tipauksorn P, Luekhong P, Thongpron J, Kamnarn U, Yingkayun K, Namin A. Stereo vision-based turn-alignment optimization for wireless power transmission positioning. In: 2023 IEEE Transportation Electrification Conference and Expo, Asia-Pacific (ITEC Asia-Pacific). [place unknown]: IEEE; 2023.
<https://doi.org/10.1109/ITECAsia-Pacific59272.2023.10372364>.

Saldaña E, Siche R, Luján M, Quevedo R. Review: computer vision applied to the inspection and quality control of fruits and vegetables. *Braz J Food Technol*. 2013;16(4):[page numbers unknown].
<https://doi.org/10.1590/s1981-67232013005000031>.

Hemamalini V, Rajarajeswari S, Nachiyappan S, Sambath M, Devi T, Singh BK, et al. Food Quality Inspection and Grading Using Efficient Image Segmentation and Machine Learning-Based System. [place unknown]: Hindawi Publishing Corporation; 2022.

Kapoor LD, Simkin AJ, Doss CGP, Siva R. Fruit ripening: dynamics and integrated analysis of carotenoids and anthocyanins. *BMC Plant Biol*. 2022;22(1):27.
<https://doi.org/10.1186/s12870-021-03411-w>.

Pardede J, Sitohang B, Akbar S, Khodra ML. Implementation of transfer learning using VGG16 on fruit ripeness detection. *Int J Intell Syst Appl*. 2021;13(2):[page numbers unknown].
<https://doi.org/10.5815/IJISA.2021.02.04>.

Zhang Y, Deng X, Xu Z, Yuan P. Watermelon ripeness detection via extreme learning machine with kernel principal component analysis based on acoustic signals. *Int J Pattern Recognit Artif Intell*. 2019;33(14):1951029.
<https://doi.org/10.1142/S0218001419510029>.

Baietto M, Wilson AD. Electronic-Nose Applications for Fruit Identification, Ripeness and Quality Grading. [place unknown]: Multidisciplinary Digital Publishing Institute; 2015.
<https://doi.org/10.3390/s150100899>.

Mandil W, Rajendran V, Nazari K, Ghalamzan-Esfahani A. Tactile-Sensing Technologies: Trends, Challenges and Outlook in Agri-Food Manipulation. [place unknown]: Multidisciplinary Digital Publishing Institute; 2023.
<https://doi.org/10.3390/s23177362>.

Scimeca L, Maiolino P, Cardin-Catalan D, Pobil AP, Morales A, Iida F. Non-destructive robotic assessment of mango ripeness via multi-point soft haptics. In: 2019 IEEE International Conference on Robotics and Automation (ICRA). [place unknown]: IEEE; 2019.
<https://doi.org/10.1109/ICRA.2019.8793956>.

Ibba P, Falco A, Abera BD, Cantarella G, Petti L, Lugli P. Bio-Impedance and Circuit Parameters: an Analysis for Tracking Fruit Ripening. [place unknown]: Elsevier BV; 2019.
<https://doi.org/10.1016/j.postharvbio.2019.110978>.

Singh AK, Ganapathysubramanian B, Sarkar S, Singh A. Deep Learning for Plant Stress Phenotyping: Trends and Future Perspectives. [place unknown]: Elsevier BV; 2018.
<https://doi.org/10.1016/j.tplants.2018.07.004>.

Arun Kumar S, Hubert Mary L, Shaamili Varsa GV, Ulaganathan MS, Kannagi L, Bharathi P. A novel and effective ensemble machine learning model for identifying healthy and rotten fruits. In: 2023 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICES). [place unknown]: IEEE; 2023.
<https://doi.org/10.1109/ICECONF57129.2023.10083721>.

Soltis PS, Nelson G, Zare A, Meineke EK. Plants Meet Machines: Prospects in Machine Learning for Plant Biology. [place unknown]: Botanical Society of America; 2020.
<https://doi.org/10.1002/aps3.11371>.

Tran VL, Doan T, Ferrero F, Huy TL, Thanh NL. The novel combination of nano vector network analyzer and machine learning for fruit identification and ripeness grading. In: 2023 Italian National Conference on Sensors. [place unknown]: [publisher unknown]; 2023.
<https://doi.org/10.3390/s23020952>.

Sathya P, Gnanasekaran P, Rahman S, Rahman A, Chen J. Ensemble feature selection framework for paddy yield prediction in cauvery basin using machine learning classifiers. *Cogent Eng*. 2023;10(1):2250061.
<https://doi.org/10.1080/23311916.2023.2250061>.

Putu I, Pratama A, Setiyawan E, Atmadji J, Purnamasari DA, Faizal E. Evaluating the performance of voting classifier in multiclass classification of dry bean varieties. *Indones J Data Sci*. 2024;5(1):124.
<https://doi.org/10.56705/ijodas.v5i1.124>.

Naranjo-Torres J, Mora M, Hernández-García R, Barrientos RJ, Fredes C, Valenzuela A. A Review of Convolutional Neural Network Applied to Fruit Image Processing. [place unknown]: Multidisciplinary Digital Publishing Institute; 2020.
<https://doi.org/10.3390/app10103443>.

Shahi TB, Sitaula C, Neupane A, Guo W. Fruit Classification Using Attention-Based MobileNetV2 for Industrial Applications. [place unknown]: Public Library of Science; 2022.
<https://doi.org/10.1371/journal.pone.0264586>.

Stefenon S, Grebogi R, Freire RZ, Nied A, Meyer L. Optimized ensemble extreme learning machine for classification of electrical insulators conditions. *IEEE Trans Ind Electron*. 2020;67(12):[page numbers unknown].
<https://doi.org/10.1109/TIE.2019.2926044>.

Sharma R, Hooda N. Optimized ensemble machine learning framework for high dimensional imbalanced bio assays. *Rev Intell Artif.* 2019;33(5):[page numbers unknown].
<https://doi.org/10.18280/ria.330509>.

Taha A, Barukab O. Android malware classification using optimized ensemble learning based on genetic algorithms. *Sustainability.* 2022;14(21):14406.
<https://doi.org/10.3390/su142114406>.

Nakagawa Y, Miyauchi T, Higashino T, Okada M. Application of random forest to classify weather observation into rainfall using GNSS receiver. In: 2021 Asia-Pacific Conference on Wearable Computing Systems. [place unknown]: IEEE; 2021.
<https://doi.org/10.1109/APWCS50173.2021.9548772>.

Taji K, Sohail A, Shahzad T, Khan BS, Khan MA, Ouahada K. An Ensemble Hybrid Framework: a Comparative Analysis of Metaheuristic Algorithms for Ensemble Hybrid Cnn Features for Plants Disease Classification. [place unknown]: Institute of Electrical and Electronics Engineers; 2024.
<https://doi.org/10.1109/access.2024.3389648>.

Saragih RE, Emanuel A. Banana ripeness classification based on deep learning using convolutional neural network. In: 2021 International Conference on Electronic and Information Technology (EIConCIT). [place unknown]: IEEE; 2021.
<https://doi.org/10.1109/EIConCIT50028.2021.9431928>.

Ko K, Jang I, Choi J, Lim J, Lee D. Stochastic decision fusion of convolutional neural networks for tomato ripeness detection in agricultural sorting systems. In: 2021 Italian National Conference on Sensors. [place unknown]: [publisher unknown]; 2021.
<https://doi.org/10.3390/s21030917>.

Vo H, Thien NN, Mui KC. A deep transfer learning approach for accurate dragon fruit ripeness classification and visual explanation using Grad-CAM. *Int J Adv Comput Sci Appl.* 2023;14(11):[page numbers unknown].
<https://doi.org/10.14569/IJACSA.2023.0141114>.