Food and Bioprocess Technology (2025) 19:80
https://doi.org/10.1007/s11947-025-04173-1

RESEARCH

Intelligent Postharvest Sorting of Bananas Using Thermal Imaging and Deep Neural Network Models

Tsega Y. Melesse[1]

Received: 4 July 2025 / Accepted: 8 December 2025 / Published online: 19 December 2025

Abstract

The assessment of postharvest fruit quality under non-invasive and accurate methodologies plays a significant role in enhancing sorting efficiency and loss minimization under precision agriculture. In this study, a deep learning framework is introduced in which a light weight CNN (LW-CNN) and thermal imaging are combined for banana classification under four distinct quality groups: fresh, ripe, overripe, and rotten. A specific set of 4336 thermal images was captured using a FLIR ONE Gen 3 infrared camera operating under thermal-only mode under controlled ambient setups. The set was subjected to balancing and augmentation procedures for better generalization capabilities, and the CNN was trained under such inputs for the identification of thermal signatures corresponding to ripeness. The proposed model demonstrated a high level of accuracy and robust performance in varied measures, marking its ability to distinguish effectively across all categories of quality. Testing under a confusion matrix and a precision-recall curve also supported the effective performance of the classifier under differing confidence thresholds. These results encourage the combination of deep learning and thermal imaging as a feasible, economically viable, and non-invasive real-time postharvest quality evaluation methodology. The proposed methodology forms the basis for smart sorting infrastructure and decision-support systems under data-informed sustainable postharvest management.

Keywords Thermal imaging · Food waste reduction · Convolutional neural networks · Post-harvest technology

Introduction

One of the greatest challenges of the world's agri-food sector is ensuring fruit quality after it is harvested. Post-harvest loss of fruits and vegetables in developing nations exceeds 45%, mainly caused by poor handling practices, delay in quality assessment, and spoilage during transportation and storage (FAO). The losses not only impose considerable economic losses on producers and distributors but also exacerbate issues related to food security and environmental strain due to the water, energy, and agricultural input losses. Hence, the incorporation of smart sensing and artificial intelligence (AI) technologies has gained increasing attention in creating advanced post-harvest quality monitoring systems that facilitate sustainable practice in precision agriculture (Wolfert et al., 2017; Zhang & Kovacs, 2012).

Conventional quality assessment techniques, including visual grading, penetrometric firmness measurement, and chemical analysis, are beset with numerous drawbacks: they are invasive, time-consuming, operator-skill-dependent, and poorly suited for high-throughput processing systems (Nicolaï et al., 2007, 2014). Their scalability is therefore impaired, particularly in automated packaging lines or cold chain distribution systems. Recent efforts have hence been directed at integrating non-invasive imaging techniques with machine learning algorithms to create intelligent and real-time monitoring systems for post-harvest applications (Lorente et al., 2012).

Among the different imaging techniques, thermal imaging, or infrared thermography, has been attracting growing interest since it can detect spatial temperature differences associated with the physiological condition of agricultural produce. Fruit radiates infrared energy proportional to its surface temperature, which mirrors important biological processes like transpiration, respiration, and internal

✉ Tsega Y. Melesse
tsegayenew.melesse@unica.it

[1] Department of Mechanical, Chemical and Material Engineering, University of Cagliari, Via Marengo 2, 09134 Cagliari, Italy

moisture movement (Patras et al., 2010). These temperature gradients are known to correlate with ripeness, spoilage, or structural damage (e.g., bruising) (Melesse et al., 2022; Roy et al., 2016). Compared to red–green–blue (RGB) imaging and hyperspectral imaging, thermal cameras are less sensitive to ambient lighting, require simpler setups, and offer faster image acquisition times, making them more practical for industrial and greenhouse applications (Jones, 2004). Thermal imaging is a valuable non-invasive signal for the characterization of fruit; however, its analysis has depended conventionally on manual features and statistical measures. The advent of deep learning and, in particular, CNNs has transformed image classification endeavors by facilitating the automatic learning of complex features from raw data (Kamilaris & Prenafeta-Boldú, 2018; LeCun et al., 2015).

CNNs have found successful applications in plant disease detection, assessment of fruit maturity, and surface defect classification (Arakeri and Lakshmana 2016; Mohanty et al., 2016). However, most of these models are computationally demanding (e.g., VGGNet, ResNet), which limits their application in real-time constrained agricultural environments such as handheld grading stations or embedded IoT devices. Moreover, banana-specific LW-CNNs for CPU-alone inference and operating threshold adaptation are little studied. When compared with hyperspectral and RGB imaging, the light-insensitive nature and monitoring of the ripening-related physiological gradients make the use of thermal imaging less sensitive. Recent studies confirm the increasing relevance of thermal imaging in the tracking of fruit quality (Chuquimarca et al., 2024, 2025; Zhang et al., 2018).

To fill this gap, in this research, we suggest a tailored LW-CNN architecture designed to be optimized for fruit classification using thermal images. The model architecture was designed to have a trade-off between classification accuracy and low computational expense to be deployed on edge devices. The classification action is focused on the four most important post-harvest fruit quality stages: fresh, ripe, overripe, and rotten. Banana was selected as the trial crop because of its worldwide commercial importance, climacteric nature of ripening, and susceptibility to post-harvest spoilage (Arunima et al., 2024; Thompson et al., 2019).

A novel thermal image dataset was collected under controlled conditions to train and test the CNN. Data were preprocessed and augmented to improve generalization, and class weights were employed to negate imbalance among quality classes. Architectural design prioritizes low parameterization, shallow depth, and fast inference times to retain applicability in low-resource environments prevalent in precision agriculture systems.

This study fills a big gap in the existing literature by integrating the precision of thermal sensing and the effectiveness of LW deep learning models to classify post-harvest fruit quality. The suggested framework aids real-time decision-making in smart post-harvest systems and is aligned with the goals of precision agriculture by facilitating non-contact, automatic, and scalable quality assessment.

Methodology

Thermal Image Acquisition and Dataset Description

For the development of a non-destructive real-time method for postharvest fruit quality evaluation, thermal images of banana samples were obtained using a FLIR ONE Gen 3 thermal camera. The device operates in the long-wave infrared (LWIR) region (8–14 μm) and possesses a native temperature resolution of 160×120 pixels. The emissivity was kept at 0.95, and the reflected temperature was calibrated for ambient environmental conditions. The camera was given a 5-min pre-imaging warming-up routine, along with non-uniformity correction for the sake of ensuring readable consistencies. For the reduction of reflection and the achievement of geometric standardization, a matte black background and constant jig distance were employed. Although it possesses a reduced sensitivity level in comparison with industrial-grade sensors, the consumer-grade FLIR ONE Gen 3 offers a cost-effective and LW alternative, making it a suitable option for wide-scale applications in resource-restricted agricultural settings (Raju & Chitra, 2024; Yang et al., 2024).

A total of 4336 thermal images were taken in a controlled indoor environment where the room temperatures were maintained between 22 and 22–25 °C to reduce external thermal interference. Thermal imaging captures ripening gradients and internal defects independently of orientation, which explains the strong overall performance. However, in practical settings, bananas appear in random orientations, making viewpoint coverage necessary for deployment. In this study, viewpoint labels were not stored, though orientation robustness was partly addressed through augmentation (rotations, flips, and translations). Future datasets will include viewpoint labels to enable detailed analysis of orientation effects.

The fruits were methodically graded into four clearly defined postharvest quality stages: fresh, ripe, overripe, and rotten (Fig. 1). The ripening was done based on visual observation, odor indicators, and expert opinion, according to pre-established maturation standards for climacteric fruits (Melesse et al., 2022; K et al. 2023).

All the thermal images were saved in JPEG format and hierarchically arranged in subfolders according to their respective categories. While trying to maintain the representation across categories evenly balanced during model development, an 80:20 stratified split was employed to separate the dataset into training and validation sets.

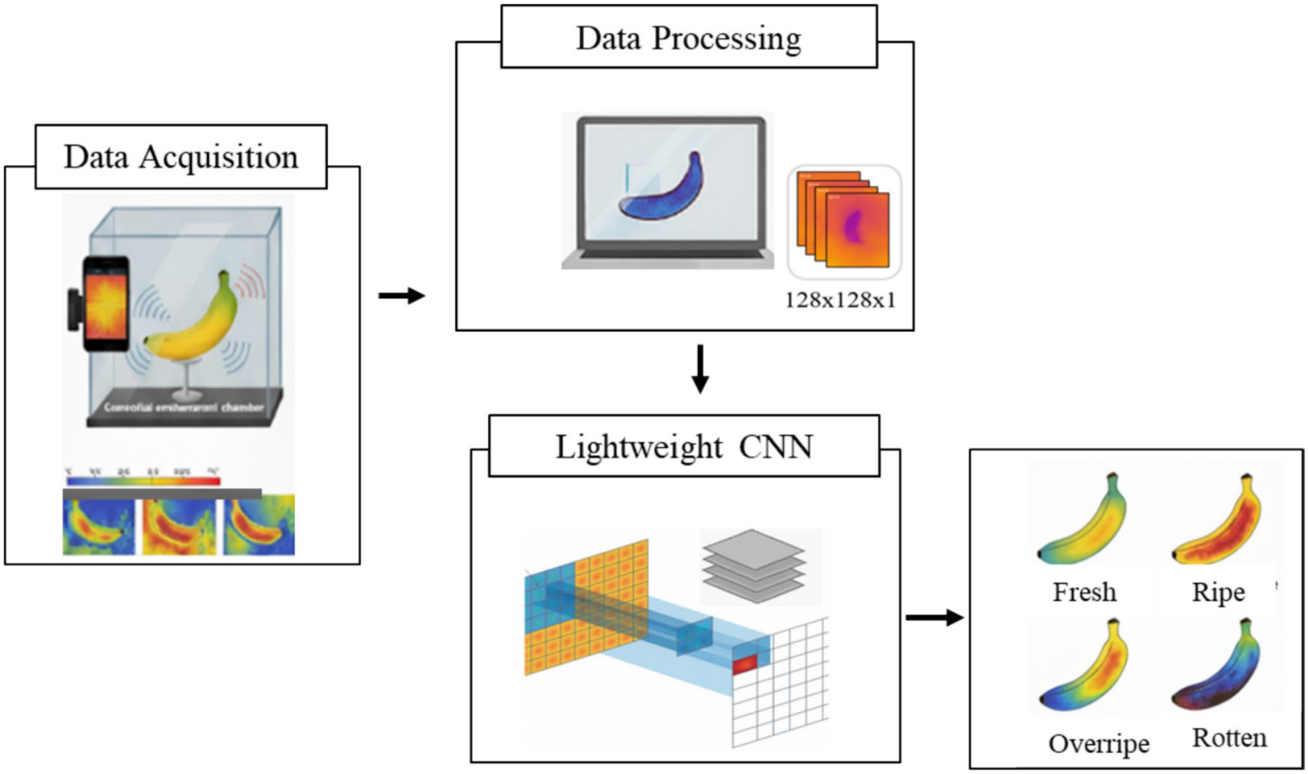

Fig. 1 Experimental setup for banana sorting using thermal imaging

Preprocessing and Data Augmentation

All images were resized to 128 × 128-pixel dimensions to satisfy the input requirements of the convolutional neural network. The pixel values were normalized to the [0,1] range by dividing by 255.0, thereby normalizing thermal intensity over the full dataset. Online data augmentation was applied during training via Keras's ImageDataGenerator to minimize overfitting and encourage generalization. The augmentations consisted of random ± 15° rotations, horizontal flips, zoom changes of ± 10%, and translation of up to 10% in height and width. These transformations simulate variability in position and perspective, thus helping the model preserve robustness against scale and orientation differences during deployment. The training set was further normalized by using computed class weights to counteract differences in the numbers of samples over the quality categories (fresh, ripe, overripe, rotten). Importantly, no methods of normalizing per-image basis were applied, as global scaling was used to preserve relative thermal gradients upon which a proper classification is dependent. The validation set was processed similarly; however, augmentation was not applied to ensure a fair evaluation of how well a deployed model would perform. These transformations were applied to simulate variability in camera angle, positioning of fruits, and lighting, and thereby enhance the generalization capacity of the CNN model. It has already been determined in earlier studies that such augmentation techniques could be efficient in boosting classification accuracy in postharvest image processing activities (Shorten & Khoshgoftaar, 2019).

CNN Architecture Design and Rationale

CNN model design for thermal image classification requires special care, as thermal images depict temperature gradients rather than texture or color, and therefore, pretrained models like MobileNetV2 or ResNet50 might not be as efficient. Recent research demonstrates that small, custom CNNs that learn to adapt to the particular nature of thermal data can outperform traditional deep models, especially if computational expenses are a factor (Li et al., 2024b). The LW-CNN model, which is tailored to thermal images, performed better in classification than VGG16 on various datasets, thereby emphasizing the merit of LW architectures optimized for radiometric features instead of RGB patterns (Fulare et al., 2023; Raju & Chitra, 2024). Additionally, techniques that take advantage of edge information or parallel kernel patterns have also shown improved performance in thermal image feature extraction, i.e., edge-conditioned CNN and parallel multiple kernel size networks (Lee et al., 2021; Li et al., 2021). Such application-tailored or reduced architectures not only reduce the computational demands but also better represent the spatial and temperature-dependent structures in thermal images, for which they are specifically well-suited to applications like quality inspection or medical diagnosis (Civilibal et al., 2023).

The architecture design included three convolutional blocks characterized by progressively deeper filter sizes of 32, 64, and 128. Each block used a 3 × 3 convolutional kernel followed by a rectified linear unit (ReLU) activation function and a subsequent 2 × 2 max pooling operation. The feature maps obtained from the last convolutional layer were then flattened and fed through a dense layer consisting of 128 neurons. A dropout layer of rate 0.5 was employed to prevent the risk of overfitting. The last output layer employed softmax activation to provide probability scores for the four quality classes. This minimal architecture was employed to enhance accuracy along with inference speed so that the model could be easily deployed on CPU-only embedded hardware like Raspberry Pi and Jetson Nano. A schematic diagram of the CNN is provided in Fig. 2.

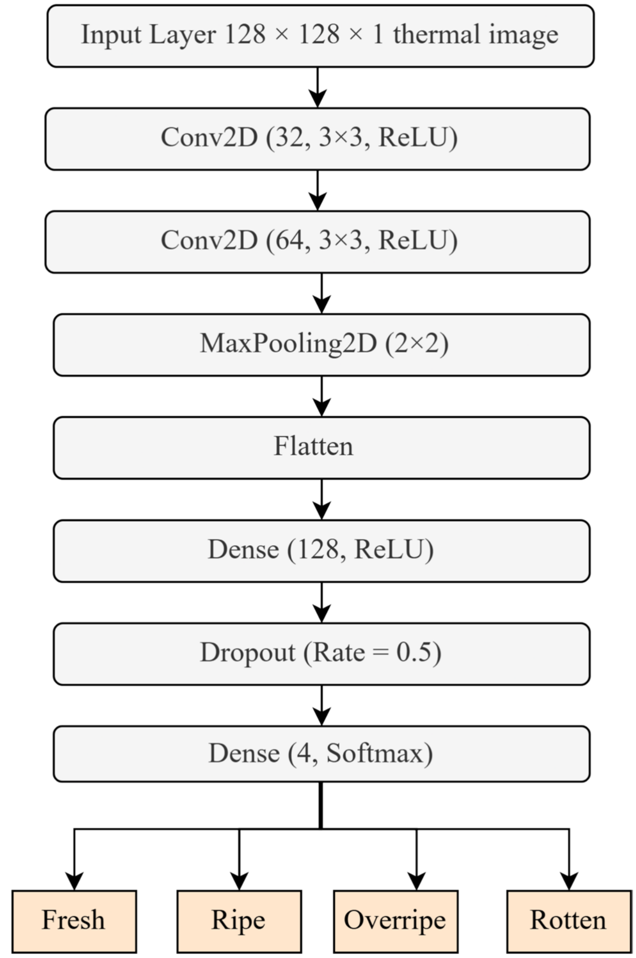

Fig. 2 Architecture of the custom CNN

Training Configuration

Training of the model was done with the Adam optimizer, with an initial learning rate of 0.001. Training was done in batches of size 32 for 10 epochs. To counter the impact of class imbalance, a weighted categorical cross-entropy loss function was employed. Class weights were computed from the inverse frequency of occurrence of samples in each of the classes present in the training dataset.

Two callback functions were utilized to achieve model convergence optimization and overfitting prevention. Early stopping with three-epoch patience was utilized, halting training if accuracy on the validation set did not increase. The ReduceLROnPlateau function was also utilized with two-epoch patience and a factor of 0.5 to adaptively reduce the learning rate when validation performance plateaued.

The training process was conducted on a stock Windows 11 notebook with an Intel Core i5 CPU and 16 GB of RAM, with no utilization of GPU acceleration. Every experimental training session was replicated three times for reliability. The model with the best validation accuracy was used for subsequent testing.

Evaluation Metrics and Visualization

The performance of the models was evaluated on the validation set with various measures of evaluation, such as accuracy, precision, recall, and F1-score. To deal with class imbalance, both macro-averaged and weighted-average values were reported. The misclassification patterns between the four quality classes were also investigated by a confusion matrix.

For a more comprehensive performance evaluation, receiver operating characteristic (ROC) curves and the respective area under the curve (AUC) scores were computed utilizing the One-vs-Rest (OvR) approach. This approach is quite appropriate for multi-class classification issues where one class is considered at a time (Behrisch et al., 2018; Sharma & Jalal, 2021). Moreover, precision-recall (PR) curves were plotted for each class to analyze the sensitivity of the model at different threshold values. Furthermore, a precision-confidence curve was constructed to investigate the relationship between classification coverage and prediction confidence. This investigation was utilized to guide potential tuning of the classification threshold at deployment time, based on the balance between false positives and false negatives desired.

All performance metrics were calculated with Scikit-learn version 1.4, while training history visualization, precision-recall (PR) curves, and receiver operating characteristic (ROC) curves visualization were done with Matplotlib.

Results and Discussion

Model Learning Behavior

The training dynamics throughout the model training process reflect an optimally developed and stable process. As can be seen in Fig. 3, through the accuracy and loss curves, the training and validation accuracy consistently increased across the epochs, ultimately converging at around 97%, with the corresponding loss values displaying an equivalent decrease. The minimal divergence between training and validation curves suggests strong generalization capability and limited overfitting. This performance reflects the combined effectiveness of dropout regularization, dynamic learning rate scheduling, early stopping, and real-time data augmentation applied during the training phase.

Surprisingly, the model recorded high accuracy without GPU acceleration, a testament to its LW architecture. Recent studies indicate that small CNNs optimized for thermal image classification record phenomenal accuracy in fruit quality evaluation, even when applied to CPU-targeted or resource-constrained systems. For example, fine-tuned CNN models have achieved as much as 98% accuracy in bruise detection and classification in strawberries based on thermal images, outperforming more computationally demanding pretrained models and showing their promise for real-time on-site postharvest quality monitoring (Guo et al., 2022). Similarly, transfer learning with compact models like SqueezeNet has enabled quick and accurate classification of a number of fruit varieties from thermal images with over 96% top-1 accuracy and constrained training time, but again supporting the potential for edge deployment (Mishra et al., 2020). Furthermore, multimodal deep learning frameworks that fuse thermal imaging with additional information sources have also demonstrated potential, precisely attaining high accuracy in variety discrimination of pineapple and physicochemical change tracking for quality (Mohd Ali et al., 2023). Field evaluations determine that CNNs, even the resource-efficient ones, are being more utilized for external fruit quality inspection with various imaging modalities and outperform traditional machine learning approaches in general (Chuquimarca et al., 2024; Naranjo-Torres et al., 2020). These findings affirm the viability of implementing thermal CNNs on embedded or edge devices for intelligent, resource-aware fruit quality monitoring in agricultural environments (Chuquimarca et al., 2024; Mishra et al., 2020).

Confusion Matrix Analysis

The confusion matrix in Fig. 4 gives an overall picture of the model's classification performance for the four quality stages considered. Each class, fresh, ripe, overripe, and rotten, records a high true positive detection rate with only small misclassifications between neighboring stages. The predominance of the principal diagonal of the matrix is an assurance that the model can effectively differentiate between various stages of ripeness and spoilage, even in the presence of thermal profile transitional overlaps.

Most misclassifications were observed between the ripe and overripe classes, a finding that mirrors their near physiological similarities and incremental differences in thermal features. This is a trend to be expected in postharvest imaging, where differentiation between ripeness stages tends to be nonlinear and may encompass overlapping features, e.g., slight temperature increments in line with increased respiration. Despite this, overall classification accuracy was

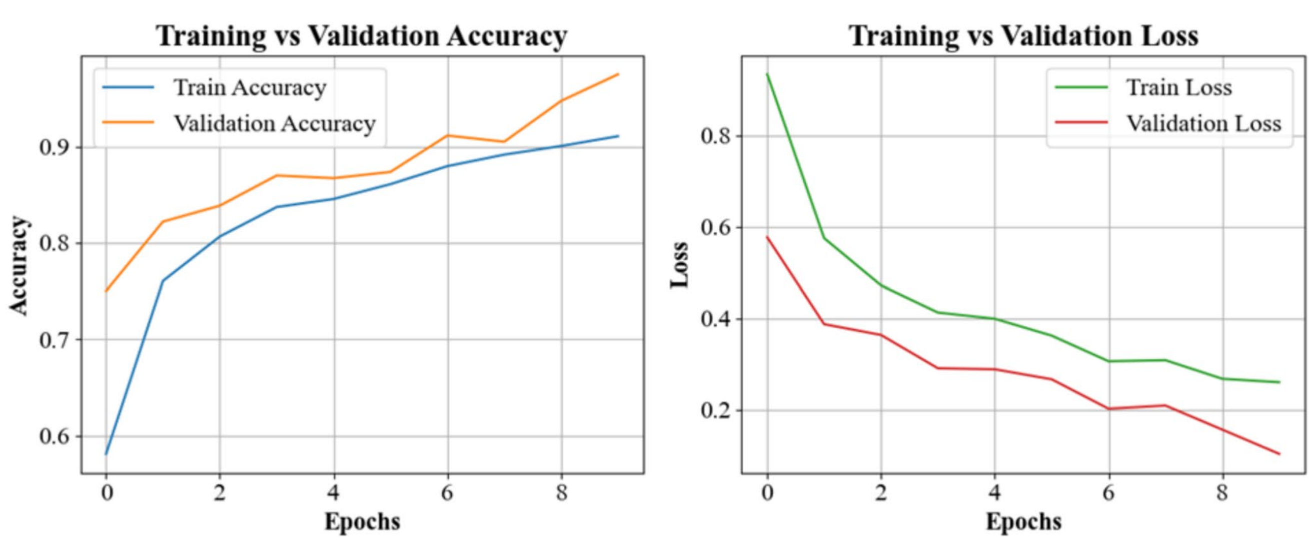

Fig. 3 Comparison of training and validation performance across epochs

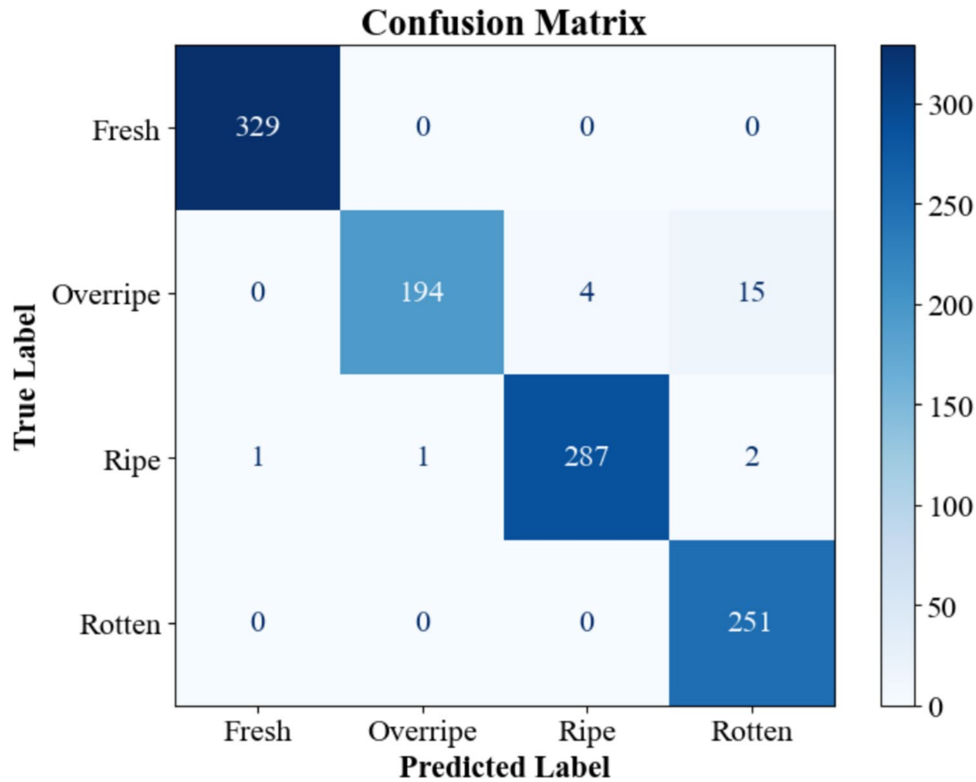

Fig. 4 Confusion matrix for multi-class fruit classification

still very high for all classes, hence further reaffirming the model's feature extraction viability.

The confusion matrix also vindicates the efficacy of the class-balancing approach taken in training. The comparatively balanced spread of accurate predictions, despite the dataset's moderate class imbalance, shows that the model did not overfit to the more prominent categories. This finding validates the use of weighted loss functions and stratified sampling for multi-class classification tasks with fruits, especially in limited data scenarios.

Furthermore, the novelty of class distinction in the confusion matrix illustrates the strength of thermal imaging in picking up underlying physiological variation that would not be apparent from visual features alone. This is a further argument for the applicability of thermal convolutional networks to real-time assessment of postharvest produce in real-world use.

Class-Wise Performance

The per-class precision, recall, and F1-scores in Table 1 indicate the strong classification power of the proposed custom CNN. The fresh class achieved perfect precision, recall, and F1-score, indicating that this class contains a highly distinctive thermal signature that the model was consistently able to recognize. The ripe, overripe, and rotten classes also scored high F1-scores of 0.98, 0.97, and 0.98, respectively. The results show that the model can distinguish between the subtle ripening phases and rot, despite their identical thermal properties. A macro-averaged F1-score of 0.98, along with a weighted average of 0.99, confirms that the model performed optimally overall across all the classes, despite the existence of a slight imbalance in the dataset.

Table 1 Classification performance metrics for banana sorting

Heading level	Precision	Recall	F1-score
Fresh	1.00	1.00	1.00
Ripe	0.99	0.98	0.98
Overripe	0.96	0.98	0.97
Rotten	0.98	0.98	0.98
Accuracy	–	–	0.97
Macro avg	0.98	0.98	0.98
Weighted avg	0.99	0.99	0.99

These findings are comparable to benchmark studies of deep learning-based fruit quality prediction using RGB and hyperspectral imaging modalities (Nicolaï et al., 2007). Unlike RGB imaging, which is inherently sensitive to ambient lighting conditions and visual noise, thermal imaging provides more stable input data across various environmental conditions. The inherent resilience demonstrated within this research most likely contributed importantly to the stable performance observed at all ripeness levels, thus emphasizing the advantages of using infrared thermal data for real-time purposes in postharvest operations.

In addition, thermal-specific convolutional neural networks have performed more effectively compared to general architectures borrowed from pre-trained RGB models like ResNet or MobileNet. This is because they can exploit domain-specific thermal characteristics, i.e., emissivity patterns, localized thermal signatures, and temperature gradients, not addressed by conventional RGB-based models (Lee et al. 2017; Rivadeneira et al., 2020). Current research strengthens the use of domain-specific models that purposefully employ these characteristics, leading to superior detection and classification performance on varied agricultural and industrial thermal imaging tasks.

Apart from the adaptation of architectural components, domain separation methods distinguishing between shared and exclusive characteristics of thermal and RGB modalities have become effective tools for encouraging model generalization, especially under data-sparse conditions (Cho et al., 2025; Shi et al., 2024). These approaches boost the representational capacity of models by retaining modality-specific information while facilitating transfer learning from large RGB datasets. Likewise, unsupervised and semi-supervised approaches such as CycleGAN-based super-resolution, domain adaptation, and style transfer networks have shown encouraging outcomes in bridging the gap between unlabeled thermal datasets and labeled RGB datasets (Li et al., 2024a). The approaches enable knowledge transfer, enhance segmentation and detection performance, and resolve the scarcity of annotated thermal datasets (Rivadeneira et al., 2022). The body of evidence stresses the need for developing architectures with special responses to the characteristics of thermal images, rather than merely adopting RGB-based models, as required for attaining state-of-the-art results in thermal vision-related tasks. In postharvest fruit classification, such targeted development not only increases accuracy but also guarantees robustness and real-world applicability (Kim et al., 2021).

ROC and AUC Analysis

The receiver operating characteristic (ROC) curves in Fig. 5 illustrate excellent class separability, with all four classes of fruit quality having an area under the curve (AUC) measure of 1.00. These near-perfect outcomes attest to the model learning very distinctive thermal features specific to each class, therefore being able to separate well even slight physiological differences among banana samples at different stages of ripeness.

The AUC values also reveal a consistently low false-positive rate across all classes. This is especially important in real-world applications like automated sorting or grading because the errors in classification could result in postharvest loss, mislabeling of products, or reduced consumer confidence in product quality. The high AUC performance across all thresholds indicates that the model is not sensitive to constrained decision boundaries but rather demonstrates robust generalization in the entire decision space.

Similar AUC values have been seen in state-of-the-art fruit quality assessment systems that employed thermal or hyperspectral imaging modalities. For example, earlier research on apples, grapes, and tomatoes has attained AUC values of well over 0.95, which are typically indicative of state-of-the-art performance for agricultural image classification problems (Zhang et al., 2024).

In general, the AUC and ROC values show that the proposed CNN not only achieves high accuracy but also is threshold-independent across a wide range of classification thresholds. This is critical for practical applications where models are expected to operate under varying conditions

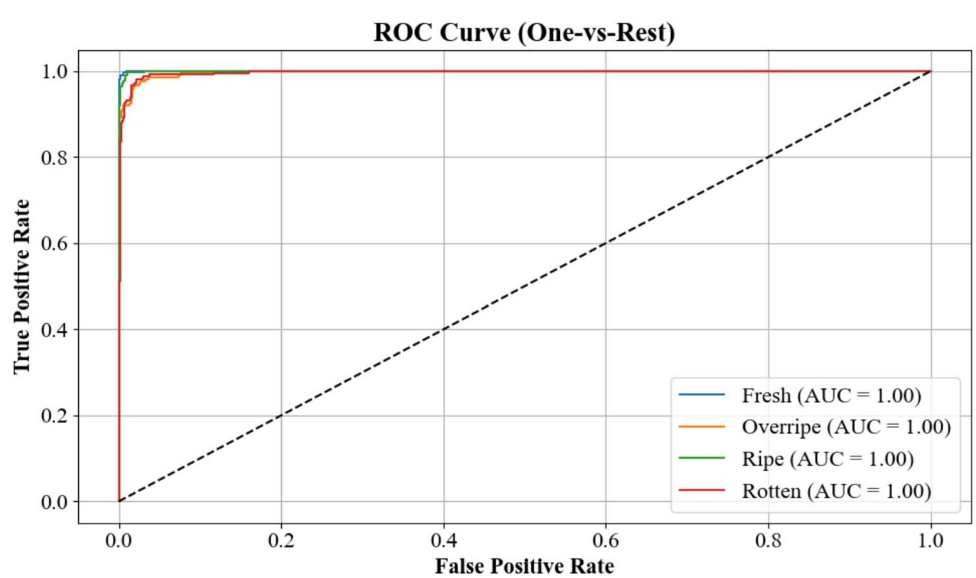

Fig. 5 ROC curves for multi-class fruit classification using One-vs-Rest strategy

and tolerance for errors. The findings show how the model is poised to be incorporated into AI-based quality control in postharvest processing and precision agriculture pipelines.

Precision-Recall and Confidence Threshold Analysis

The precision-recall (PR) curves in Fig. 6 provide a more realistic description of the model's classification performance under class-imbalanced conditions. The Fresh class exhibits high precision across the full range of complete recall, which means that the model accurately detects Fresh samples with low false positives, irrespective of the decision threshold used.

Conversely, the ripe class portrays monotonically decreasing precision at a growing recall. This is likely caused by thermal overlap with the overripe class, with continuous physiological attributes and identical surface temperature profiles. The overripe class possesses the most variability in precision-recall performance due to its intermediate thermal characteristics that share attributes with the ripe and rotten classes, thus causing the difficulty in establishing decision boundaries. Concurrently, the rotten class features a steep and smooth PR curve, which shows the model’s ability to detect advanced decay at high precision and low false-positive rates. Such behavior is particularly of interest in postharvest quality control, where not detecting rotten produce can result in drastic economic losses, batch contamination, or product recalls.

These PR trends are in agreement with the latest advancement in fruit defect detection literature, which shows that thermal imaging combined with deep learning detects late degradation stages more accurately than early ripening since there is higher temperature contrast in rotting tissues (Zhang et al., 2024). The results confirm that the proposed CNN architecture achieves consistent classification across all quality stages, with especially strong performance in the detection of commercially important conditions such as spoilage.

Figure 7 depicts the correlation between model confidence and classification precision. As expected, elevated confidence thresholds are associated with a rise in classification precision, reaching values nearing 1.0 when the model’s predicted confidence surpasses 0.90. This observation suggests that the model’s predictions made with high confidence are nearly always accurate, underscoring its dependability within rigorous decision limits.

But this accuracy gain is achieved at the expense of reduced recall, revealing the inherent trade-off between classification confidence and sample coverage. Lowering the confidence threshold boosts coverage but is accompanied by more false positives, which may be undesirable in certain operational environments. For instance, in automatic reject systems of rotten fruit and vegetables, minimization of false positives is crucial to avoid rejecting good ones. On the other hand, in upstream sorting operations or quality assessment applications, recall maximization guarantees that all defective products are identified.

Tuning the threshold to values ranging from 0.75 to 0.85 appears to find an effective balance between precision and recall, thus an effective compromise for real-time use cases. The observation is aligned with previous studies on fruit defect detection using convolutional neural networks, where the use of adaptive thresholding improved the model’s reliability, especially in edge AI environments and embedded vision systems (Cong et al., 2025). These findings justify the application of confidence-based threshold tuning and probabilistic filtering as core techniques for performance enhancement in operational postharvest classification setups.

Fig. 6 Precision-recall curves

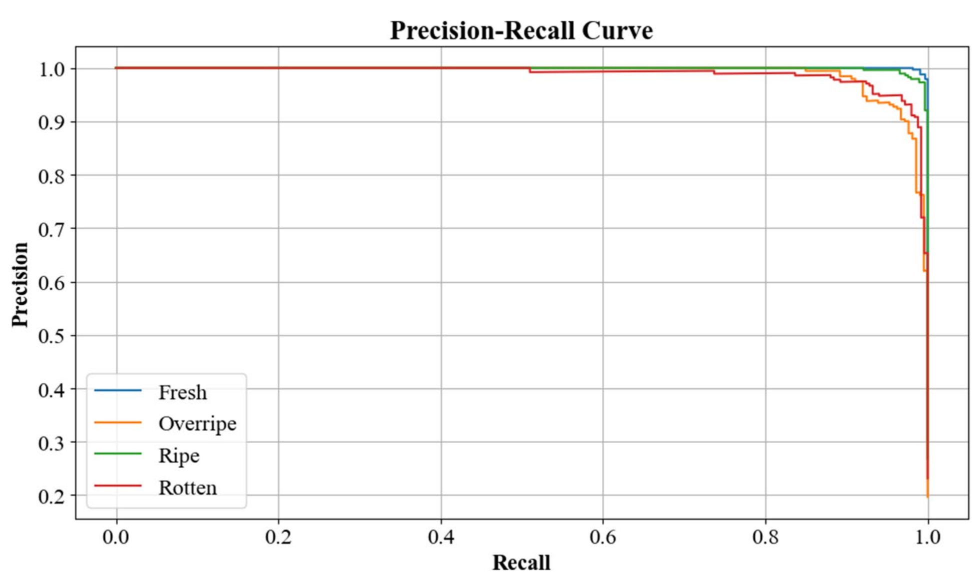

Fig. 7 Relationship between precision and confidence scores

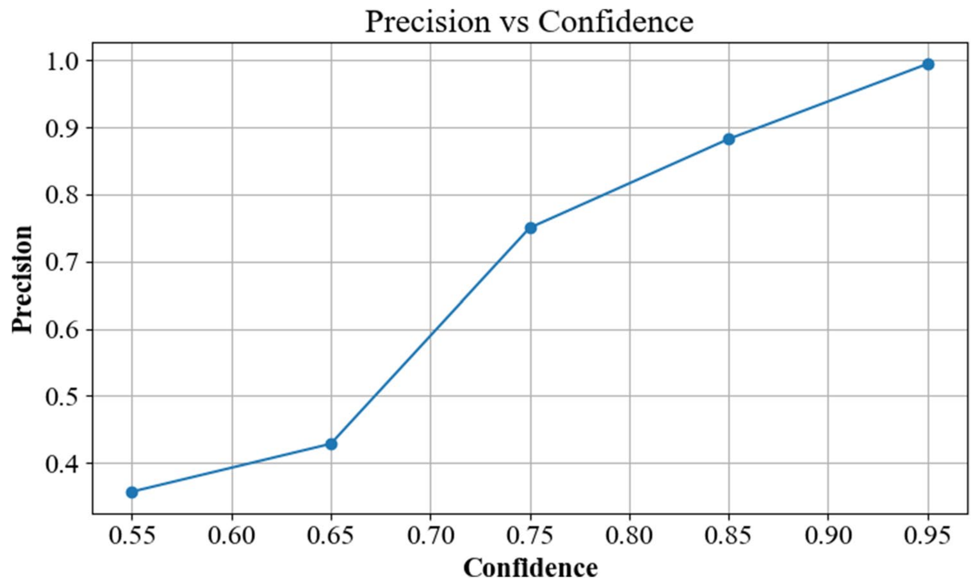

Table 2 Average test results comparing the baseline CNN and the proposed LW-CNN model

Model	Accuracy	Precision	Recall	F1-score
CNN-baseline	0.876	0.972	0.969	0.970
Proposed LW-CNN	~0.97	~0.99	~0.99	~0.99

Comparative Analysis

The proposed thermal classification system based on convolutional neural networks (CNN) offers several advantages over conventional postharvest monitoring practices, including those based on RGB imaging and transfer learning-based models. To numerically evaluate these advantages, a baseline CNN model without thermal-specific architectural designs was also trained on the same data. The baseline model achieved an average accuracy of 87.6%, precision of 97.2%, recall of 96.9%, and F1-score of 97.0% (see Table 2). In comparison, the proposed LW-CNN achieved around 97% accuracy and a weighted F1-score of around 0.99, thus demonstrating the performance gains through task-specific architectural design and thermal-specific optimization.

The proposed CNN-based thermal classification model has many advantages over conventional postharvest monitoring systems, including RGB imaging and transfer learning-based models. While RGB-based systems are widely used in the commercial grading of fruit, they can be affected by variable lighting conditions, surface gloss, and limited ability to detect underlying physiological processes. These issues restrict their use in detecting spoilage at early stages or small ripening variations (Lu et al., 2022; Patel et al., 2012). In contrast, thermal imaging is a non-invasive method that records physiologically meaningful heat signatures related to respiration, evaporation of water, and microbial activity. It enables earlier and more sensitive detection of postharvest changes and hence has an advantage over surface-reliant imaging methods (Roy et al., 2016).

Table 3 compares a lightweight thermal-only CNN architecture against conventional RGB and hyperspectral approaches. For the scope of the four-class validation set, our proposed architecture achieved 97% accuracy, macro-F1 = 0.9962, weighted-F1 = 0.9963, macro-AUC ≈ 1.000, and macro-mAP = 0.999, which indicates the mean average precision calculated from class-wise precision-recall curves. More recent explorations employing RGB technology have shown robust accuracies when challenged by controlled lighting, whereas hyperspectral methods display respectable performance but rely upon higher-order, multi-band sensor technologies. The thermal setting, however, continually offers higher, threshold-conscious performance by a variety of metrics, based solely upon a single inexpensive camera and desktop CPU-only inference, and so implies practical scalability toward applications in postharvest sortation. Because datasets and protocols differ from experiment to experiment, these comparisons are indicative, not direct contests; however, our results underscore the promise of thermal imaging in discerning ripening-related physiological gradients based upon a lightweight, compact CNN.

Compared to state-of-the-art pre-trained models such as MobileNetV2, EfficientNet-Lite, and ResNet50 fine-tuned for natural RGB image classification, the designed LW-CNN used in this study achieved similar accuracy at a significantly lower computational cost. The pre-trained models contain redundant parameters and require extensive fine-tuning to adapt to thermal or agricultural domain-specific data. Recent research has shown that compact, task-oriented CNNs tailored to thermal images can outperform transfer learning

Table 3 Comparison of banana ripeness classification approaches

Study/source	Imaging modality	Task/classes	Metrics
Proposed LW-CNN	**Thermal imaging**	Ripeness (4 classes)	Accuracy; macro-F1; Weighted-F1; macro-AUC; macro-mAP
Arunima et al., 2024	**RGB imaging**	Ripeness (4 classes)	Accuracy
Shuprajhaa et al., 2023	**RGB imaging**	Ripeness (4 stages)	Accuracy; precision; recall; F1
Zhang et al., 2024	**RGB imaging**	Ripeness classification: 7 (coarse)/14 (fine) stages	Accuracy
S et al., 2022	**Combined RGB + hyper-spectral imaging (visible-near infrared)**	Grading/quality evaluation	Accuracy; F1
Wang et al., 2023	**Hyperspectral imaging (visible-near infrared range)**	Quality/composition analysis	Performance varied by model and trait; no single accuracy metric provided
Hespeler et al., 2021	**Thermal imaging**	Classification task	Accuracy

techniques in fruit bruise detection, decay classification, and spoilage prediction (Yadav & Tandan, 2025; Zhang et al., 2018).

Also, the model described in this study was designed to run on CPU platforms alone, making it compatible with low-power devices and ready for real-time application in the field or supply-chain settings, regardless of high-end GPU availability. This kind of hardware optimality is especially useful in smallholder farm setups, mobile postharvest testing units, and off-grid storage warehouses where computing capacity could be constrained. Previous applications of edge AI in agriculture have demonstrated the same benefits, especially in disease detection, yield estimation, and maturity classification using LW neural networks (Ariza-Sentís et al., 2024).

Another key benefit of the proposed approach is its interpretability and integrability. The reduced model architecture enhances the model's usability in more integrated digital twins and IoT ecosystems. This makes it straightforward to integrate real-time classification outcomes in predictive analytics systems, decision-support systems, and intelligent alert systems. The described features are in line with modern trends in postharvest automation, which increasingly seek to incorporate AI-powered sensing technologies directly into infrastructure to enable quality monitoring, traceability, and automatic sorting (Wolfert et al., 2017).

Contribution to Postharvest Digitalization

The suggested thermal classification model based on CNN contributes to the further digitalization of postharvest quality control following the principles of Industry 4.0. It offers a viable and scalable solution to real-time fruit quality analysis for postharvest quality control towards enabling automation, traceability, and data-driven decision-making, smart agriculture, and digital food systems' fundamental pillars. As agri-food supply chains worldwide compete on higher efficiency, transparency, and sustainability, the coupling of AI with sensor-based technologies is growing in importance (Rose & Chilvers, 2018).

The system enhances operational responsiveness in the supply chain through early detection of fruit quality deterioration, particularly during fruit transition to overripe or spoiled stages. This anticipatory insight allows producers and logistics personnel to perform preemptive sorting, dedicated routing, and flexible cold storage practices, thus preventing deterioration from spreading across an entire batch. These characteristics are particularly important in climacteric fruits like bananas, tomatoes, and peaches, which experience pronounced physiological changes following harvest and account for the majority of the world's postharvest losses (Hodges et al., 2011; FAO).

The addition of thermal imaging adds a novel dimension of physiological sensing beyond the traditional RGB-based inspection approaches. Unlike RGB imaging, which is limited to the evaluation of surface color and texture, thermal imaging identifies concealed evidence of inner biological activity, such as elevated respiration rates, moisture loss, and incipient microbial growth. Such markers are prognostic of spoilage and enable preemptive action before defects become visible. Furthermore, thermal sensors are also insensitive to varying light conditions and are well adapted to hygiene-critical applications like cold storage and automated packaging lines (Roy et al., 2016).

When integrated within IoT platforms or digital twin systems, the system can evolve from a classification tool into a predictive engine, whereby future quality states can be predicted from time-series sensor data. Current research has shown that integration of AI-powered sensing with real-time monitoring platforms has a profound impact on enhancing cold chain management, decision-making accuracy, and reducing food rejection at downstream distribution and retail nodes (Verdouw et al., 2016).

From a broader vision of policy and sustainability, this approach is consistent with United Nations Sustainable Development Goal (SDG) 12.3, which aims to reduce food waste worldwide by half by the year 2030. The model's consistency with cost-effective thermal imaging technology, publicly available software libraries, and CPU-only operating environments enhances its usability for deployment in decentralized contexts. This enhances the accessibility of smallholder farmers, cooperatives, and processors in low- and middle-income nations, thus ensuring an equal distribution of digital postharvest technologies and reducing the digital divide. Through the integration of low-cost sensing, advanced analytics, and plug-and-play solutions, this initiative plays a key role in creating smarter and more resilient food systems.

Conclusion

This study presents a novel and efficient convolutional neural network (CNN) approach based on thermal imaging technology for non-invasive and real-time evaluation of postharvest fruit quality. Employing bananas as a typical example of climacteric fruit, the proposed system excellently classified samples into four main quality phases: fresh, ripe, overripe, and rotten, with high accuracy, robustness against class imbalance, and computational efficiency. Using thermal imaging as the main sensing modality, the proposed approach takes advantage of temperature-related physiological markers of ripening and spoilage, which are more resistant to external illumination conditions and surface noise compared to traditional RGB-based approaches.

One of the key contributions of this work is the design of a task-focused, thermally efficient CNN architecture custom-built for deployment on CPU-only hardware, thereby obviating the need for high-end GPUs. This feature ensures applicability under resource-constrained environments, such as small-scale agricultural farms, distributed cold store facilities, and mobile inspection systems for postharvest assessment. This proof-of-concept addressed single fruits; industrial scalability will require segmentation and tracking of bunches on conveyors. The lean design of the model enables scalability, integration within IoT frameworks, and possible extension to prediction digital twin models, thereby providing a pathway to more sophisticated, data-driven postharvest management systems aligning with Industry 4.0 ideas.

Future efforts will scale up the suggested framework to cover more types of fruits and vegetables, evaluate its performance on actual fruit storage and distribution areas, and incorporate temporal streams of data to enable predictive analytics and remaining shelf-life estimation. The consumer-grade FLIR ONE Gen 3 device and controlled indoor setting limit generalization. Deployment claims are restricted to controlled lab scenarios. Future work will test conveyor-based acquisition, multiple cultivars, and variable ambient conditions. Moreover, current research endeavors to fuse thermal data with other non-destructive sensing modalities and to create advanced domain adaptation methods that will further boost accuracy and generalizability. These guidelines will underpin the function of intelligent, AI-driven sensing systems as key enablers for smarter, more durable, and equitable food systems. As a conclusion, this research illustrates an efficient, large-scale, and significant method of modernizing postharvest quality inspection using contemporary technology, thereby ensuring a firm basis for prospective advancement in precision agriculture and sustainable food supply chains.

Author Contribution T.Y.M.: Conceptualization, methodology, investigation, data curation, formal analysis, writing—original draft, writing—review and editing, visualization.

Funding Open access funding provided by Università degli Studi di Cagliari within the CRUI-CARE Agreement.

Data Availability The dataset used in this study is available upon reasonable request.

Declarations

Ethics Approval and Consent to Participate Not applicable.

Consent for Publication Not applicable.

Competing interests The authors declare no competing interests.

References

Arakeri, M. P. (2016). Computer vision based fruit grading system for quality evaluation of tomato in agriculture industry. *Procedia Computer Science*, *79*, 426-433.

Ariza-Sentís, M., Vélez, S., Martínez-Peña, R., et al. (2024). Object detection and tracking in precision farming: A systematic review. *Computers and Electronics in Agriculture, 219*, Article 108757.

Arunima, P. L., Gopinath, P. P., Geetha Lekshmi, P. R., & Esakkimuthu, M. (2024). Digital assessment of post-harvest Nendran banana for faster grading: CNN-based ripeness classification model. *Postharvest Biology and Technology, 214*, Article 112972.

Behrisch, M., Blumenschein, M., Kim, N. W., et al. (2018). Quality metrics for information visualization. *Computer Graphics Forum, 37*, 625–662.

Cho, B., Choi, S., & Lee, S. (2025, January). Domain Separation Aided Data Efficiency Enhancement for Thermal Object Detection. In *2025 International Conference on Electronics, Information, and Communication (ICEIC)* (pp. 1-4). IEEE.

Chuquimarca, L. E., Vintimilla, B. X., & Velastin, S. A. (2024). A review of external quality inspection for fruit grading using CNN models. *Artificial Intelligence in Agriculture, 14*, 1–20.

Chuquimarca, L., Vintimilla, B., & Velastin, S. (2025). Banana ripeness level classification using a simple cnn model trained with real and synthetic datasets. *arXiv preprint arXiv:2504.08568*.

Civilibal, S., Cevik, K. K., & Bozkurt, A. (2023). A deep learning approach for automatic detection, segmentation and classification of breast lesions from thermal images. *Expert Systems with Applications, 212*, Article 118774.

Cong, P., Wang, K., Liang, J., et al. (2025). TQVGModel: Tomato quality visual grading and instance segmentation deep learning model for complex scenarios. *Agronomy, 15*, 1273.

FAO The State of Food and Agriculture 2019: Moving forward on food loss and waste reduction. Rome: FAO. Available at: https://openknowledge.fao.org/server/api/core/bitstreams/11f9288f-dc78-4171-8d02-92235b8d7dc7/content (Accessed 09 June 2025).

Fulare A, Raghavendra Prasad K, Johri P, et al (2023) Convolutional neural network based image classification and it's comparison with Vgg-16 to measure accuracy. In: 2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE). IEEE, Greater Noida, India, pp 367–370

Guo, B., Li, B., Huang, Y., et al. (2022). Bruise detection and classification of strawberries based on thermal images. *Food and Bioprocess Technology, 15*, 1133–1141.

Hespeler, S. C., Nemati, H., & Dehghan-Niri, E. (2021). Non-destructive thermal imaging for object detection via advanced deep learning for robotic inspection and harvesting of chili peppers. *Artificial Intelligence in Agriculture, 5*, 102–117.

Hodges, R. J., Buzby, J. C., & Bennett, B. (2011). Postharvest losses and waste in developed and less developed countries: Opportunities to improve resource use. *Journal of Agricultural Science, 149*, 37–45.

Jones, H. G. (2004). Application of thermal imaging and infrared sensing in plant physiology and ecophysiology. In *Advances in botanical research* (Vol. 41, pp. 107-163). Academic Press.

Kurinji, K., Methira, R., Senthilarasi, M., & Mohamed Mansoor Roomi, S. (2023, December). Banana Ripeness Classification via Thermal Imaging with Transfer Learning. In *2023 International Conference on Energy, Materials and Communication Engineering (ICEMCE)* (pp. 1-5). IEEE.

Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture, 147*, 70–90.

Kim, Y.-H., Shin, U., Park, J., & Kweon, I. S. (2021). MS-UDA: Multi-spectral unsupervised domain adaptation for thermal image semantic segmentation. *IEEE Robotics and Automation Letters*. https://doi.org/10.1109/LRA.2021.3093652

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*, 436–444.

Lee, J., An, Y., Kim, M., & Pan, S. (2021). Comparison of CNN architecture for thermal face emotion classification. *2021 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia)* (pp. 1–4). Gangwon, Korea, Republic of.

Lee, K., Lee, J., Lee, J., et al. (2017). Brightness-based convolutional neural network for thermal image enhancement. *IEEE Access, 5*, 26867–26879.

Li, C., Xia, W., Yan, Y., et al. (2021). Segmenting objects in day and night: Edge-conditioned CNN for thermal image semantic segmentation. *IEEE Transactions on Neural Networks and Learning Systems, 32*, 3069–3082.

Li, G., Yu, Z., Yang, K., et al. (2024a). Exploring feature selection with limited labels: A comprehensive survey of semi-supervised and unsupervised approaches. *IEEE Transactions on Knowledge and Data Engineering, 36*, 6124–6144.

Li, H., Pinto, G., Piscitelli, M. S., et al. (2024b). Building thermal dynamics modeling with deep transfer learning using a large residential smart thermostat dataset. *Engineering Applications of Artificial Intelligence, 130*, Article 107701.

Lorente, D., Aleixos, N., Gómez-Sanchis, J., et al. (2012). Recent advances and applications of hyperspectral imaging for fruit and vegetable quality assessment. *Food and Bioprocess Technology, 5*, 1121–1142.

Lu, S., Chen, W., Zhang, X., & Karkee, M. (2022). Canopy-attention-YOLOv4-based immature/mature apple fruit detection on dense-foliage tree architectures for early crop load estimation. *Computers and Electronics in Agriculture, 193*, Article 106696.

Melesse TY, Bollo M, Pasquale VD, et al (2022) Machine learning-based digital twin for monitoring fruit quality evolution. In: Longo F., Affenzeller M., Padovano A. (eds) Procedia Comput. Sci. Elsevier B.V., pp 13–20

Mishra, D. K., Dey, N., Deora, B. S., & Joshi, A. (2020). *ICT for competitive strategies: Proceedings of 4th International Conference on Information and Communication Technology for Competitive Strategies (ICTCS 2019), December 13th-14th, 2019, Udaipur, India* (First edition). CRC Press, Taylor & Francis Group.

Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for image-based plant disease detection. *Frontiers in Plant Science, 7*, 1419.

Mohd Ali, M., Hashim, N., Abd Aziz, S., & Lasekan, O. (2023). Utilisation of deep learning with multimodal data fusion for determination of pineapple quality using thermal imaging. *Agronomy, 13*, 401.

Naranjo-Torres, J., Mora, M., Hernández-García, R., et al. (2020). A review of convolutional neural network applied to fruit image processing. *Applied Sciences, 10*, Article 3443.

Nicolaï, B. M., Beullens, K., Bobelyn, E., et al. (2007). Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biology and Technology, 46*, 99–118.

Nicolaï, B. M., Defraeye, T., De Ketelaere, B., et al. (2014). Nondestructive measurement of fruit and vegetable quality. *Annual Review of Food Science and Technology, 5*, 285–312.

Patel, K. K., Kar, A., Jha, S. N., & Khan, M. A. (2012). Machine vision system: A tool for quality inspection of food and agricultural products. *Journal of Food Science and Technology, 49*, 123–141.

Patras, A., Brunton, NigelP., O'Donnell, C., & Tiwari, B. K. (2010). Effect of thermal processing on anthocyanin stability in foods; Mechanisms and kinetics of degradation. *Trends in Food Science & Technology, 21*, 3–11.

Raju PS, Chitra P (2024) A weighted ensemble approach for enhanced brain tumor detection and classification using CNN and VGG16*. In: 2024 International Conference on Communication, Control, and Intelligent Systems (CCIS). IEEE, Mathura, India, pp 1–5.

Rivadeneira R, Sappa A, Vintimilla B (2020) Thermal image super-resolution: A novel architecture and dataset: In: Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications. SCITEPRESS - Science and Technology Publications, Valletta, Malta, pp 111–119

Rivadeneira, R. E., Sappa, A. D., Vintimilla, B. X., & Hammoud, R. (2022). A novel domain transfer-based approach for unsupervised thermal image super-resolution. *Sensors, 22*, 2254.

Rose, D. C., & Chilvers, J. (2018). Agriculture 4.0: Broadening responsible innovation in an era of smart farming. *Frontiers in Sustainable Food Systems, 2*, Article 87.

Roy SD, Das DH, Bhowmik MK, Ghosh AK (2016) Bruise detection in apples using infrared imaging. In: 2016 9th International Conference on Electrical and Computer Engineering (ICECE). IEEE, Dhaka, Bangladesh, pp 118–122.

S, R., Ganguli, S., Selvan, P. T., et al. (2022). Deep learning based dual channel banana grading system using convolution neural network. *Journal of Food Quality, 2022*, 1–9.

Sharma, H., & Jalal, A. S. (2021). A survey of methods, datasets and evaluation metrics for visual question answering. *Image and Vision Computing, 116*, Article 104327.

Shi, C., Zheng, Y., & Chen, Z. (2024). Domain adaptive thermal object detection with unbiased granularity alignment. *ACM Transactions on Multimedia Computing, Communications, and Applications, 20*, 1–23.

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data, 6*, 60.

Shuprajhaa, T., Mathav Raj, J., Paramasivam, S. K., et al. (2023). Deep learning based intelligent identification system for ripening stages of banana. *Postharvest Biology and Technology, 203*, Article 112410.

Thompson, A. K., Supapvanich, S., & Sirison, J. (2019). Fruit Ripening. *Banana Ripening* (pp. 25–55). Springer International Publishing.

Verdouw, C. N., Wolfert, J., Beulens, A. J. M., & Rialland, A. (2016). Virtualization of food supply chains with the internet of things. *Journal of Food Engineering, 176*, 128–136.

Wang, H., Lv, L., Li, X., et al. (2023). A safety management approach for Industry 5.0′s human-centered manufacturing based on digital twin. *Journal of Manufacturing Systems, 66*, 1–12.

Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M.-J. (2017). Big data in smart farming – A review. *Agricultural Systems, 153*, 69–80.

Yadav, K. K., & Tandan, G. (2025). FruitNet: Lightweight CNN for high-throughput image-based fruit yield estimation. *SHS Web of Conferences, 216*, 01054.

Yang, B., Hung, Y.-C., Kumar, G. D., et al. (2024). From sunburn detection to optimal cooling: A review of recent applications of thermal imaging to improve preharvest and postharvest handling of fruit and vegetables. *Scientia Horticulturae, 337*, Article 113527.

Zhang, C., & Kovacs, J. M. (2012). The application of small unmanned aerial systems for precision agriculture: A review. *Precision Agriculture, 13*, 693–712.

Zhang, C., Liu, C., Zeng, S., et al. (2024). Hyperspectral imaging coupled with deep learning model for visualization and detection of early bruises on apples. *Journal of Food Composition and Analysis, 134*, Article 106489.

Zhang, Y., Lian, J., Fan, M., & Zheng, Y. (2018). Deep indicator for fine-grained classification of banana's ripening stages. *EURASIP Journal on Image and Video Processing, 2018*, Article 46.