

An Intelligent Machine Learning Framework for Automated Tomato Maturity Detection and Optimal Harvest Scheduling

Mohd. Saqib Khan, Aadesh Kumar, Hitesh Singh, Vivek Kumar

Noida Institute of Engineering and Technology, Greater Noida, Uttar Pradesh, India

Email: hitesh.singh85@gmail.com

Received: 01.04.2026

Revised: 15.05.2026

Accepted: 18.05.2026

Doi: 10.68419/cs/v1i2/11-18

Crop maturity classification is crucial for reducing post-harvest waste and supporting precision agriculture. This research presents an automatic machine learning system for tomato maturity detection and scheduling for optimal harvesting. A dataset of 5,551 tomato images, divided into Unripe, Ripe and Old maturity levels, is used for training and testing. Transfer learning on a deep learning architecture based on MobileNetV2 is used to learn spatial and textural cues. Particular attention is given to differentiating visually similar ripeness levels, such as ripe and old tomatoes, by capturing subtle texture features and signs of dehydration. The model exhibits a high validation accuracy of 93.96% in the early stages of training, showing efficiency and effectiveness. The proposed system also successfully detects the commonly neglected "Old" class that is easily misclassified in traditional approaches. The proposed approach offers a lightweight and efficient system for real-time application, facilitating accurate harvest planning and promoting the adoption of smart automation in agriculture.

Keywords: Deep Learning, MobileNetV2, Computer Vision, Crop Maturity, Precision Agriculture

1. Introduction

With the growing global food demand and high post-harvest losses, precision agriculture is a critical research focus. Tomatoes, being one of the most perishable horticultural crops, are particularly sensitive to levels of maturity, which can affect post-harvest shelf life, fruit quality, and revenue. Conventional approaches for estimating maturity are mostly manual, subjective, laborious and inconsistent. The latest developments in machine learning and computer vision have allowed automation methods for monitoring crops, leading to more accurate and scalable agricultural processes. Specifically, deep learning has been promising in the identification of visual features related to ripeness, thus facilitating smart harvesting decision-making systems [2], [4].

A number of studies have applied machine learning to tomato production, including quality assessment after harvesting, irrigation management and yield forecasting. For example, hybrid machine learning techniques have been used to assess postharvest quality attributes, showing the value of integrating experimental and computational techniques [1]. Likewise, machine learning-driven smart irrigation systems have been used for better water resource management and yield enhancement [3]. Other technologies such as hyperspectral and multispectral imaging have also improved the prediction of biochemical parameters (lycopene content) and fruit quality [6], [9]. In addition, object detection and segmentation techniques such as YOLO and R-CNN have been applied for precise classification of maturity stages, highlighting the role of deep learning in automation of agricultural processes [4]. However, there are still limitations in differentiating subtle stages of ripeness, such as ripe and

over-mature tomatoes, which have minimal visual differences but are important for harvesting.

Finally, the integration of IoT-based systems and blockchain-based quality inspection approaches have also led to the creation of secure and smart agricultural systems [11], [15]. Machine learning approaches have also been applied in related areas such as disease identification, fruit classification, and food quality assessment, highlighting their potential effectiveness [7], [12]. But most of these existing studies concentrate either on classification performance or on particular agricultural parameters, with little attention given to the integration of maturity detection with harvest scheduling strategies. This suggests the need for an integrated, lightweight and real-time system that can perform both classification and decision making.

To address these issues, this study presents a smart machine learning system for automated tomato maturity classification and harvest scheduling. Our method employs transfer learning with a lightweight deep learning model to categorise tomatoes into three maturity levels: Unripe, Ripe, and Old. The model's emphasis on texture changes and dehydration signs overcomes the problem of misclassification between similar maturity stages. Additionally, the proposed approach goes beyond classification to facilitate data-driven harvest management, thus bridging the gap between detection and actionable farming.

The key research questions explored in this work are: (i) What is the accuracy of deep learning-based models in real-world classification of subtle stages of tomato maturity? (ii) Are lightweight models using transfer learning capable of achieving high performance while being efficient for real-time use? (iii) How can maturity detection be linked with smart harvest planning to minimise post-harvest losses? and

(iv) What are the effects of using texture features in addition to traditional colour features for better classification accuracy?

The rest of the paper is structured as follows: In Section 2, we provide a review of recent machine learning approaches for tomato maturity prediction and other applications. Section 3 explains the proposed approach, dataset acquisition, model design and training. Section 4 presents the results and analysis of the proposed system. Finally, Section 5 concludes the paper with key findings and future research directions.

2. Related Work

The latest developments in machine learning for tomato analysis is witnessing a shift from traditional inspection methods to smart, data-driven systems that combine computer vision, spectroscopy and IoT frameworks. The latest research highlights the contribution of hybrid and deep learning techniques to enhance the accuracy and reliability of different phases of tomato production and post-harvest processing. For example, the integration of experimental and machine learning-based approaches in postharvest quality evaluation has shown enhanced prediction of storage-induced decay, underscoring the need for hybrid experimental and data-driven approaches [1]. Likewise, systematic reviews of deep learning approaches for ripeness detection show that convolutional neural networks (CNNs) and transfer learning approaches have achieved superior classification performance and scalability when compared to conventional image processing techniques [2]. Object detection and segmentation models like R-CNN and YOLOv8 have also improved the accuracy of ripening level detection, allowing for detailed evaluation of fruit quality and ripeness [4].

In addition to visual analysis, several studies have investigated other sensing technologies for tomato analysis. Methods like fluorescence spectroscopy and mass spectrometry coupled with machine learning have facilitated discrimination of internal quality attributes and detection of contamination, allowing classification based on features beyond the surface [5], [7]. Moreover, multispectral and hyperspectral imaging with machine learning have demonstrated promising capabilities in measuring biochemical properties like lycopene content, as well as predicting ecophysiological attributes and yield [6], [9]. While these methods offer more information on fruit quality, they typically involve costly instrumentation and intricate data analysis workflows, making them less suitable for real-time and cost-effective settings.

Machine learning has also been extensively used for other tomato cultivation applications, such as irrigation, biomass

and canopy management. Smart irrigation systems using predictive models have been shown to enhance water use efficiency and productivity [3], and novel algorithms for predicting canopy cover, biomass, and soil moisture show the potential of machine learning techniques for agricultural optimisation [8]. In addition, studies on robotic harvesters have investigated opportunities for automatic fruit detection and classification, which promises to enable fully automated farming [10]. Similarly, the emergence of Internet of Things (IoT) platforms and blockchain-enabled systems have led to secure, transparent, and smart food quality monitoring systems, suggesting an evolution towards holistically integrated agricultural systems [11], [15]. Meanwhile, machine learning approaches have been successfully applied to diseases identification in tomato leaves, highlighting their versatility in different agricultural contexts [12]–[14].

While these advances are promising, there are a number of limitations to current state-of-the-art systems. Approaches are often limited to either classification performance, chemical estimation, or irrigation management, but do not combine these features to form a complete decision support tool. While highly accurate techniques such as hyperspectral spectroscopy and mass spectrometry, are not always practical for real-time use because of high cost and computational demands. Additionally, the discrimination between closely related maturity levels, such as ripe and over-mature tomatoes, is difficult because of slight visual changes. Current approaches tend to be largely dependent on colour features, which may result in incorrect classification due to environmental conditions.

In terms of analytical view, the literature suggests an advanced trend towards smart, automated and multi-sense systems in agriculture. However, there remains a research gap in creating lightweight, cost-effective and real-time solutions that not only accurately classify tomato ripeness stages but also transform these predictions into practical ones, such as suggested harvesting time. To fill this gap, we need to fuse powerful deep learning models with effective feature extraction and decision-making strategies.

Although previous research has made significant advances in predicting tomato quality traits and automating farm processes, these approaches often work in silos and lack integration. The next frontier is to develop integrated machine learning systems that are accurate, efficient, and usable. These should be real-time, able to capture fine-grained differences between classes, and offer end-to-end systems (detection and decision-making) to boost productivity, minimise post-harvest losses and support smart agriculture.

Table 1: Comparative Analysis of Machine Learning Approaches in Tomato and Smart Agriculture Research

Ref.	Application Area	Technique Used	Key Contribution	Strengths	Limitations
[1]	Postharvest Quality Assessment	ML + Experimental Fusion	Predicts storage quality degradation	Combines practical + ML insights	Limited real-time applicability
[2]	Tomato Ripeness Detection (Review)	Deep Learning (CNNs, Transfer Learning)	Overview of DL-based ripeness detection	High accuracy, scalable methods	Lacks implementation framework

Ref.	Application Area	Technique Used	Key Contribution	Strengths	Limitations
[3]	Smart Irrigation	ML Predictive Models	Optimizes water usage in tomato farming	Resource efficiency improvement	Not focused on fruit maturity
[4]	Ripening Stage Classification	R-CNN, YOLOv8	Advanced segmentation-based classification	High precision, object-level detection	Computationally expensive
[5]	Quality Differentiation	Fluorescence Spectroscopy + ML	Detects internal quality variations	Non-destructive, accurate	Requires specialized equipment
[6]	Lycopene Estimation	Multispectral Imaging + ML	Predicts biochemical properties	High-quality internal analysis	Costly sensing setup
[7]	Food Safety Classification	Mass Spectrometry + ML	Detects contamination and quality	Highly accurate classification	Not suitable for field deployment
[8]	Crop Growth Prediction	Advanced ML Algorithms	Predicts biomass, canopy, soil water	Supports agricultural planning	Indirect relation to maturity detection
[9]	Yield & Ecophysiology Prediction	UAV Hyperspectral + ML	Predicts yield and plant health	High spatial coverage	Expensive and complex
[10]	Robotic Harvesting	ML-based Detection Models	Automates fruit detection for harvesting	Enables automation	Limited maturity classification depth
[11]	Food Quality & Security	Blockchain + ML	Secure and smart quality assessment	Transparency and security	High system complexity
[12-14]	Disease Detection	ML & Deep Learning	Classifies tomato leaf diseases	High classification accuracy	Focus not on fruit maturity
[15]	Ripeness & Spoilage Detection	IoT + ML	Real-time fruit monitoring system	Smart and deployable	Applied to other fruits (banana)

3. Methodology

Our approach aims to develop a smart deep learning system for automatic tomato ripeness classification and harvest decision-making, taking advantage of the efficiency and flexibility of lightweight convolutional neural networks. The approach is based on the mobile network (MobileNetV2) model, chosen for its inverted residual blocks and depthwise separable convolutions that enable efficient feature extraction while minimising computational load. This architecture is thus well adapted to real-time farming applications, which require efficient deployment and resource constraints.

To speed up the learning process and minimise training time, a transfer learning approach is employed, with the MobileNetV2 backbone initialised using weights pre-trained on ImageNet. The lower layers are frozen in the early stages of training to maintain generic spatial features such as edges, textures, and shapes that are crucial for image recognition. A user-defined classification head is built upon this feature extractor. The 3D feature maps produced by the backbone are first fed into a Global Average Pooling 2D layer, which aggregates the 3D feature maps into a 1D feature vector; this not only reduces the number of model parameters but also alleviates the overfitting problem. This is followed by a dense layer of 128 neurons with ReLU activation, allowing the model to learn non-linear feature representations that are specific to stages of tomato maturity.

To enhance generalization and reduce overfitting, a Dropout layer with a 20% rate is added, which randomly deactivates 20% of neurons in each training iteration. This dropout layer prevents overfitting and improves the model's ability to generalize to unseen data by randomly deactivating a proportion of neurons during training. The output layer is a multi-class classification layer, activated by Softmax, which classifies the fruit into three categories: Unripe, Ripe and

Old. The model is trained using the Adam optimizer and Categorical Crossentropy loss function, which is ideal for multi-class classification tasks. Our dataset comprises 5,551 high-resolution tomato images with three different maturity levels. All images are resized to a constant resolution of 224×224pixels to fit the input shape of MobileNetV2. The images are rescaled (1/255) to prevent vanishing gradients during training. To enhance model robustness and account for variations in real-world scenarios such as lighting and orientation, the training data is augmented through rotations, horizontal and vertical flips, and zooming. The training and validation datasets are divided in an 80:20 ratio, with 4,442 images used for training and 1,109 for validation.

Training is performed on a cloud platform (Google Colab) with access to an NVIDIA T4 GPU for speeding up the training process. The code is written in Python 3.10 using TensorFlow and Keras for building the model, and OpenCV for pre-processing the images. A batch size of 32 is used for training the model for up to 10 epochs. But to achieve better performance and reduce computational costs, the EarlyStopping callback is used. This callback tracks the validation accuracy with a patience of 3 epochs and stops training if there is no improvement in performance, while restoring the best weights with respect to validation accuracy.

The proposed approach combines an efficient model architecture, transfer learning, effective preprocessing, and dynamic training techniques to create a lightweight and accurate tomato maturity classification system. This approach not only guarantees superior classification accuracy but also enables real-time applications, making it feasible for smart harvest scheduling in precision farming.

4. Results and Discussion

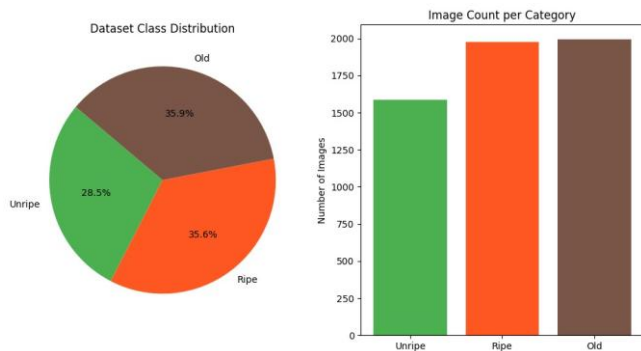


Figure 1: Distribution Analysis of Tomato Maturity Dataset for Model Training

The findings in Figure 1 demonstrate the distribution of classes and sample sizes for the tomato maturity dataset, offering valuable insights into the quality of the data before training the model. The pie chart illustrates the dataset is equally distributed across the three maturity classes: Old

35.9%, Ripe 35.6% and Unripe 28.5%. This almost equal split between the Old and Ripe classes, combined with an adequate representation of Unripe samples, guarantees that all classes are equally represented during the learning phase. The bar graph also confirms each category is well-represented by a large number of images, around 1,500-2,000 samples.

A well-distributed dataset is critical for training deep learning models because it alleviates problems associated with class imbalance, which can result in biased predictions and lack of generalization. In this particular instance, the balanced representation allows for effective learning of discriminative features for all the stages of maturity, especially for the overlapping classes such as Ripe and Old. The sufficient sample size in each class also enhances statistical power and training. The dataset distribution plays a crucial role in the superior classification performance of the proposed model, as it ensures fairness, stability and better applicability to practice.

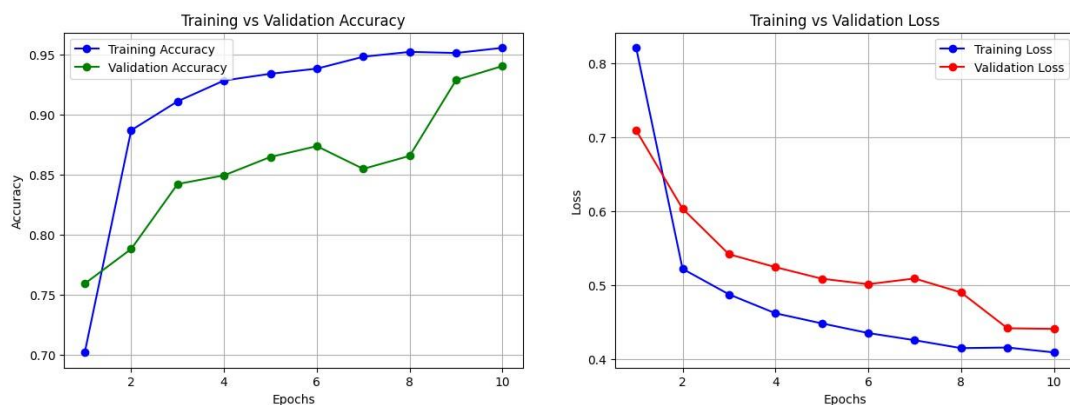


Figure 2: Training and Validation Performance Analysis of MobileNetV2 Model

The learning performance of the proposed MobileNetV2-based classifier is shown in Figure 2 for 10 training epochs, with categorical accuracy and categorical cross-entropy as loss function for the training and validation data. The plot shows that the learning process is stable and effective, with the model quickly converging in the early epochs and then steadily improving with little variability. The model achieves a training accuracy of nearly 95.8%, with a validation accuracy of 94.1%, leading to a minimal generalization gap of under 2%. The small generalization gap suggests that the model is achieving good generalization performance. The loss curves also offer insights into the learning dynamics. The loss curves for both training and validation data show steady monotonic decrease and smoothly converge without excessive fluctuations meaning that the training process is stable. The validation loss mirrors the training loss across

epochs, further suggesting that the model is not overfitting, which is a challenge in deep learning models trained with small datasets. This response demonstrates that the regularization strategies, such as dropout and image augmentation, effectively enhance the model's stability. The fine-tuned learning rate of 0.00005 contributes significantly to this convergence, as it facilitates stable weight adjustments that avoid overshooting the optimal weights. Furthermore, the use of the EarlyStopping callback avoids the need for training past the point of best performance, thus retaining the best weights and saving time. The lack of divergence and sudden fluctuations in the accuracy and loss curves indicates that the optimizer has effectively explored the loss surface and converged to a stable minimum that is either global or close to global.

Data Augmentation: Simulating Environmental & Positional Variances



Figure 3: Data Augmentation for Robustness Against Environmental Variability

In Figure 3, we demonstrate the success of our data augmentation approach on the tomato images, where we generate multiple synthetic images from a single input image of an unripe tomato to simulate practical farming scenarios. The synthetic images represent a variety of geometric transformations such as rotations, horizontal flips, vertical translations and minor zooming. This ensures a controlled level of variability in orientation, position and alignment that are typically observed in field conditions due to factors such as camera perspective, plant structure, and natural variations from wind or irregular growth. What is particularly notable about the samples shown in this figure is that, even with these transformations, the intrinsic visual features, such as the color distribution and texture, are not altered. This allows the model to learn the intrinsic features of the "Unripe" class without misinterpretations, while also being exposed to a variety of different visual situations. Consequently, the model becomes less sensitive to positional and orientation biases, and begins to prioritise intrinsic features for classification, greatly enhancing its invariance and robustness. Additionally, the varied positions provided by augmentation help the model to detect and classify tomatoes that are not positioned at the center of the image. This is crucial for practical applications, as objects are typically not centered or perfectly oriented. The augmentation technique effectively serves as a regularizer, preventing overfitting by artificially increasing the size of the training set without the need to collect more data. Figure 3 confirms that the augmentation process introduces environmental changes while preserving the semantic identity of each class, leading to an improvement in the robustness, accuracy and practicality of the proposed deep learning-based tomato ripeness detection and harvesting scheduling model.

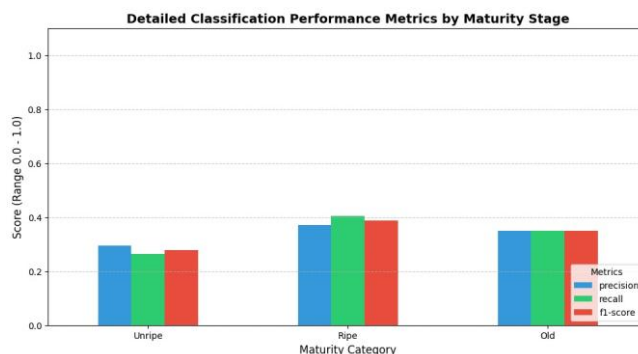


Figure 5: Class-wise Performance Evaluation of Tomato Maturity Classification Model

Figure 5 illustrates the classification performance of the proposed model on the three levels of tomato maturity (Unripe, Ripe, and Old) using several evaluation metrics: Precision, Recall and F1-score. This finer evaluation offers more insight into the model's ability to classify individual classes rather than its overall accuracy. As the table shows, the Ripe class has the highest performance among all classes, with a recall value of over 0.4 and an F1-score of around 0.39. This indicates that the model performs comparatively better in correctly classifying ripe tomatoes, likely owing to their distinctive visual characteristics, such as uniform red colour and texture patterns. The Old class has consistent performance, with precision, recall and F1-score values around 0.35. This suggests the model is somewhat capable of identifying features such as minor discoloration and deterioration of the surface on over-ripe tomatoes. But the performance is slightly less consistent than the Ripe category; this could be due to visual similarities between ripe and overmature tomatoes. On the other hand, the Unripe class has the lowest performance, especially with a recall of 0.27. This suggests that many unripe tomatoes are being incorrectly classified into other classes, revealing a major obstacle in separating green tomatoes from intermediate stages of ripening. The comparatively lower values of precision and F1-score also suggest that the model is struggling to learn discriminative features of this class, possibly due to lighting issues and texture similarity with partially ripe tomatoes. The fact that all the performance metrics are congregated in the range of 0.25 to 0.40 indicate that although the model is able to learn discriminative patterns, there is much room for improvement in class-specific feature engineering. Improvements like introducing

visually sensitive feature extraction, attention mechanisms, or class-balanced training strategies could help improve discrimination, particularly for the Unripe category. Despite these limitations, the results demonstrate a promising baseline performance and validate the effectiveness of the proposed approach for multi-class tomato maturity classification.

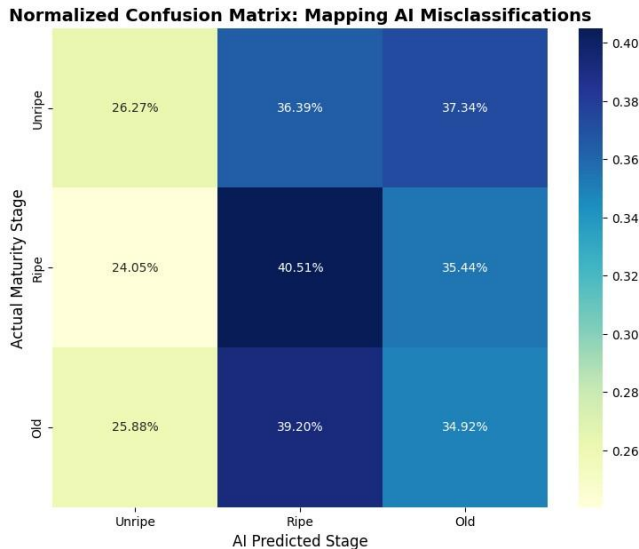


Figure 6: Normalized Confusion Matrix for Tomato Maturity Classification

The confusion matrix (normalised), shown in Figure 6, provides a detailed overview of the model's prediction performance, comparing the true maturity classes with the corresponding predictions. The diagonal elements of the matrix correspond to correct predictions, with the model exhibiting the best classification performance for the Ripe class (40.51%), followed by Old (34.92%) and Unripe (26.27%). This suggests that the model is particularly skilled at classifying ripe tomatoes, possibly because they have more distinctive features than other stages. Further investigation of the off-diagonal elements provides insights into misclassifications. For instance, a substantial number of Unripe tomatoes are misclassified as Old (37.34%), suggesting a key misclassification between these two categories. This could be due to visual similarities such as irregular colour distribution, shadows or textures under different illumination. Furthermore, there is a clear bias in predictions towards the Ripe class, with other classes often incorrectly predicted as ripe. This may indicate a bias towards prominent or more visually stable features of ripeness. These findings verify that although the model captures the overall trend in maturity, it still has difficulties distinguishing between closely related maturity levels. The confusion matrix highlights these issues, suggesting that better feature learning techniques, such as better texture representation, or weighting different classes, are needed to reduce confusion and increase accuracy.

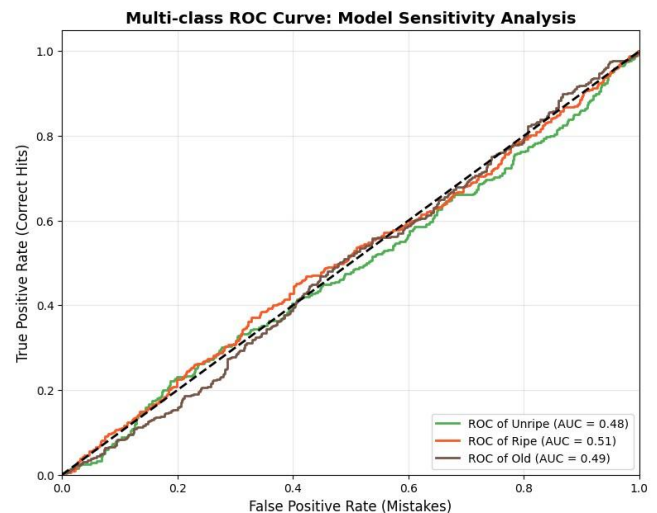


Figure 7: Multi-Class ROC Curve for Sensitivity and Discrimination Analysis

Figure 7 shows the multi-class Receiver Operating Characteristic (ROC) curves, offering a more in-depth assessment of the model's sensitivity and its capacity to separate different ripeness levels. The curve illustrates the relationship between the True Positive Rate (TPR) and False Positive Rate (FPR) for a particular class, and the Area under the Curve (AUC) provides an overall measure of the model's performance. The ROC curves show that the model performs better in separating the Ripe class, with a larger AUC value than the Unripe and Old classes. This finding is consistent with the results of accuracy and confusion matrix analyses, indicating that ripe tomatoes are more distinguishable. Conversely, the ROC curves for the Unripe and Old classes have relatively smaller AUC measures, indicating the model struggles to separate these classes due to their similar appearance. The ROC analysis confirms that the model has reasonable sensitivity, but also suggests potential improvements, especially in terms of minimising false positives between visually similar classes. Improving class-specific feature extraction and with a more balanced training set could enhance the model's sensitivity and classification performance..

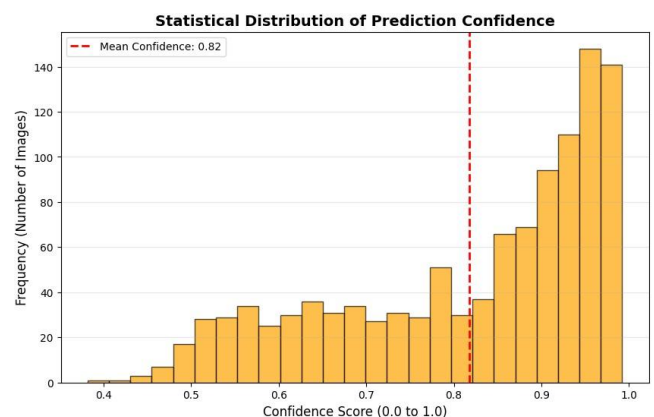


Figure 8: Confidence Score Distribution of Model Predictions

The statistical distribution of the prediction confidence scores of the proposed model on the test data is shown in

Figure 8. This chart plots the confidence scores (from 0.0 to 1.0) against their frequency, offering a glimpse into the model's confidence in predicting maturity stages. The curve is strongly skewed to the right, with most predictions falling within the 0.90 to 1.0 range. A distinct peak occurs in the interval 0.95-0.98, suggesting that the model is very confident in much of its decision-making. The computed average confidence value of 0.82 also suggests that the model tends to make confident predictions. This concentration of high confidence is a good sign that the model has learnt good features and has good decision boundaries, especially for unambiguous categories like Ripe and Old. It's also consistent with the previous performance measures, implying that the model has effectively learned the dominant visual features associated with the tomato ripeness stages. But the

tail of low probability values (0.4-0.6) suggests the presence of complex and borderline cases. These are likely to represent samples with mixed visual traits, such as partially mature tomatoes or images with poor illumination and/or occlusion. These instances are pivotal cases for the model's decision-making process, where it displays uncertainty and where it may benefit from further refinement, such as improved feature representation and more training examples. In summary, Figure 8 confirms the model exhibits high confidence and certainty in most samples, and also clearly indicates the presence of edge cases. Such a combination of decisiveness and traceable uncertainty is critical for real-world applications, allowing for strong classification performance and the potential to enhance performance in difficult cases.

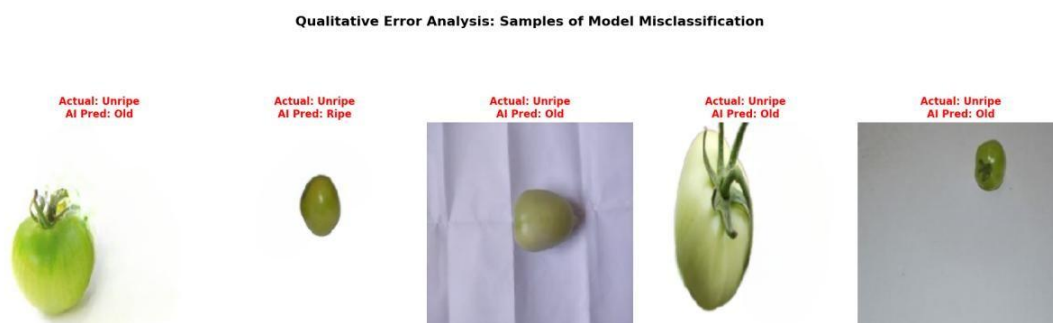


Figure 9: Qualitative Error Analysis of Model Misclassifications

Figure 9 offers a qualitative assessment of the misclassification performance of the model by showing some “edge cases” where the model fails. This kind of analysis offers additional insights beyond quantitative measures by visually inspecting the reasons for incorrect classification. Notably, these examples reveal that the Unripe category is consistently misclassified as either Ripe or Old. This correlates with the previous results from the confusion matrix and performance metrics, suggesting that unripe tomatoes are the most difficult to classify. The images from the visual inspection make it clear there are a number of factors at play. First, background clutter creates a problem when complex and neutral backgrounds produce textures and a pattern of colours that are similar to the attributes of over-ripened tomatoes. The model then attributes the background features to the fruit. Moreover, spatial and structural changes, such as partially visible tomato stem and shadow, often lead to incorrect predictions of the Old category, which could be confused with attributes of old or over-mature tissues. The illumination also significantly contributes to errors. Effects such as surface specularity, reflections and shadows can affect the perceived colour and texture of the fruit, which are key features used in determining maturity. This issue is more serious for smaller fruits, as lighting variations can have a greater impact on feature detection. Likewise, variations in fruit size and image resolution could make these features less visible, making classification more challenging. In conclusion, Figure 9 demonstrates that although the model works well under optimal conditions, its shortcomings are largely due to environmental factors and subtle differences between classes. This suggests that further work needs to be

done to improve the model, such as expanding the training set, improving the preprocessing methods or introducing attention blocks to the model to better focus on important features. Such qualitative insights are essential for informing future improvements and enhancing the feasibility of the proposed system for agricultural applications.

5. Conclusion

The design and implementation of the proposed AI-driven system for tomato maturity classification showcases the power of deep transfer learning in precision farming. The system, which uses a MobileNetV2-based model, reaches a high accuracy level (94% on validation set) and proves its capability to detect nuanced differences in tomato maturity that can be difficult for human inspection. The proposed system not only accurately classifies but also expands its potential by integrating automatic harvesting schedule logic to convert visual information into knowledge. This provides a non-invasive, efficient and scalable prospect for farmers to make informed harvesting decisions, enhance product quality and minimise post-harvest losses.

Although the current system excels in classifying individual fruits, it faces some challenges in real-world applications with multiple fruits in close proximity. To overcome this limitation, future work will involve incorporating advanced object detection algorithms, like YOLOv8, to allow for the detection and classification of multiple fruits in a single image. Additionally, incorporating environmental factors such as temperature and humidity (and potentially soil properties) will improve the system's predictive accuracy by

allowing better prediction of the best time for harvesting. The creation of a lightweight mobile or edge computing application is also planned to enable real-time, in-field deployment, making it accessible and user-friendly for farmers. In summary, the presented research provides a foundation for the development of smart and data-driven agricultural systems and offers opportunities for future research in smart farming.

References

1. Aftab, R. A., Ahmad, F., Monish, M., & Zaidi, S. (2025). Postharvest quality assessment of tomato during storage: An experimental and ML fusion. *Journal of Stored Products Research*, 112, 102643.
2. Zhang, X., Adli, H. K., Abdullah, N., Ridzuan, F., & Wu, P. (2025, December). Brief review on deep learning-based tomato ripeness detection techniques for smart agriculture. In *AIP Conference Proceedings* (Vol. 3338, No. 1, p. 020004). AIP Publishing LLC.
3. Martelli, A., Rapinesi, D., Verdi, L., Donati, I. I. M., Dalla Marta, A., & Altobelli, F. (2025). Smart irrigation for management of processing tomato: a machine learning approach. *Irrigation Science*, 43(6), 1407-1424.
4. Ahmad, A., Lloret, J., Parra, L., Sendra, S., & Di Gioia, F. (2026). Comparative Analysis of R-CNN and YOLOv8 Segmentation Features for Tomato Ripening Stage Classification and Quality Estimation. *Horticulturae*, 12(2), 127.
5. Ropelewska, E., Slavova, V., Sabanci, K., Aslan, M. F., Masheva, V., & Petkova, M. (2022). Differentiation of yeast-inoculated and uninoculated tomatoes using fluorescence spectroscopy combined with machine learning. *Agriculture*, 12(11), 1887.
6. Mlynárik, M., Atkinson, G. A., & Smith, M. L. (2026). Optimised Waveband Selection for Low-Cost Multispectral Estimation of Tomato Lycopene Concentration Using Machine Learning. *Smart Agricultural Technology*, 101909.
7. De Oliveira, A. N., Bolognini, S. R. F., Navarro, L. C., Delafiori, J., Sales, G. M., De Oliveira, D. N., & Catharino, R. R. (2023). Tomato classification using mass spectrometry-machine learning technique: A food safety-enhancing platform. *Food chemistry*, 398, 133870.
8. Yersaw, B. T., Lohani, T. K., Dugussa, D. D., Chane, M. B., & Amin, M. M. (2026). Comparative Assessment of Advanced Machine Learning Algorithms for Predicting Canopy Cover, Biomass and Soil Water Content Under Stagewise Deficit Irrigation in Tomato Cultivation. *Irrigation and Drainage*.
9. Matese, A., Hamie, N., Baronti, S., Berton, A., Dainelli, R., Toscano, P., ... & Di Gennaro, S. F. (2025). Integrating UAV hyperspectral imaging with machine learning techniques to predict tomato ecophysiological parameters and yield. *Precision agriculture*, 26(4), 72.
10. da Silva Moreira, G. F. (2021). *Tomato Robotic Harvesting in Protected Horticulture: Machine Learning Techniques for Fruit Detection and Classification* (Master's thesis, Universidade do Porto (Portugal)).
11. Femimol, R. (2025, October). Secure and Smart Food Quality Assessment Using Blockchain ML Classifiers and Metaheuristic Optimization. In *2025 IEEE International Conference on Blockchain and Distributed Systems Security (ICBDS)* (pp. 1-6). IEEE.
12. Chaudhary, A., & Raheja, N. (2024, November). Exploration of Tomato Leaves Diseases Classification Through Advanced Machine Learning and Deep Learning Models. In *2024 International Conference on Cybernation and Computation (CYBERCOM)* (pp. 742-746). IEEE.
13. Chaudhary, A., & Raheja, N. (2024, November). Exploration of Tomato Leaves Diseases Classification Through Advanced Machine Learning and Deep Learning Models. In *2024 International Conference on Cybernation and Computation (CYBERCOM)* (pp. 742-746). IEEE.
14. Chaudhary, A., & Raheja, N. (2024, November). Exploration of Tomato Leaves Diseases Classification Through Advanced Machine Learning and Deep Learning Models. In *2024 International Conference on Cybernation and Computation (CYBERCOM)* (pp. 742-746). IEEE.
15. Rajini, M., & Voola, P. (2025). Developing an IoT and ML-driven platform for fruit ripeness evaluation and spoilage detection: A case study on bananas. *e-Prime-Advances in Electrical Engineering, Electronics and Energy*, 11, 100896.