






Cite this: *Sustainable Food Technol.*,
2025, **3**, 2215

Deep learning-based approach for classifying mandarin orange maturity

Raj Singh, ^a C. Nickhil, ^{*a} Konga Upendar,^b Poonam Mishra ^a
and Sankar Chandra Deka ^a

The precise prediction of fruit maturity is essential for determining the optimal harvest time. It helps to reduce postharvest losses and maintain consistent fruit quality for consumers. Traditional methods for assessing maturity depend largely on manual inspection. This process is subjective, time-consuming, and prone to human error. Deep learning approaches, particularly convolutional neural networks (CNNs), offer a promising alternative by automating classification with high precision and consistency. This research seeks to identify the most effective deep-learning algorithms for predicting the maturity of mandarin oranges. In this study, the performance of four convolutional neural network architectures (EfficientNet-B0, ResNet50, VGG16, and a Custom CNN) was investigated for the classification of mandarin oranges based on their maturity levels: unripe, ripe, and overripe. The primary dataset comprised 1095 images, with each category containing 365 images. The deep learning models achieved the best accuracy rates of 98% for both EfficientNet-B0 and ResNet50, 83% for VGG16, and an impressive 99% for the Custom CNN, considering primary images. By comparing these models on a balanced dataset, this work offers a practical guide for researchers and practitioners on selecting models for assessing fruit maturity. Notably, EfficientNet-B0, ResNet50, and the Custom CNN exhibited significantly higher success rates compared to VGG16 and existing models, making them particularly recommendable for the development of an efficient automated system for harvesting and sorting mandarin oranges in the near future. The results aim to identify potential applications for improving agricultural practices, quality assessment, and overall efficiency in the food industry.

Received 21st July 2025
Accepted 18th September 2025

DOI: 10.1039/d5fb00408j

rsc.li/susfoodtech

Sustainability spotlight

This research introduces a deep learning-based classification model for determining the maturity stages of mandarin oranges, enabling accurate, real-time, and non-destructive assessment. By enhancing harvest timing and reducing post-harvest losses, the approach supports resource-efficient practices, minimizes food waste, and contributes to a more sustainable and data-driven citrus supply chain.

1. Introduction

Khasi mandarin (*Citrus reticulata* Blanco), a premium seeded mandarin orange cultivar originating from India, stands out as a widely acclaimed and economically significant fruit globally.¹ In the Indian context, Khasi mandarin constitutes a substantial 43.6% of the total citrus fruit production, encompassing nearly 38.2% of the overall citrus cultivation area.² Its primary cultivation area is concentrated in the north-eastern part of India, particularly in Assam and the surrounding foothills. The high commercial importance of Khasi mandarin is underscored by its significant contributions to both production and export potential.^{3–5} Khasi mandarin, recognized for its easy peeling

and vibrant deep orange color when ripe, is globally appreciated for its dietary and nutritional attributes.⁶ Whether consumed fresh or as juice, the fruit's peel is commonly discarded as waste.

Nevertheless, the peel constitutes a substantial portion, approximately 40% to 50% of the wet fruit mass, representing a promising reservoir of bioactive elements, such as ascorbic acid, carotenoids, phenolic compounds, and pectins, with flavonoids being particularly concentrated. Surprisingly, the peel surpasses the juice in vitamin C content, as indicated by the USDA National Nutrient Database, and displays remarkable antioxidant properties compared to other fruit parts.⁷ Despite citrus peels often being treated as agricultural waste, they prove to be a potential source of valuable secondary plant metabolites and essential oils.⁸ Beyond its economic value, the Khasi mandarin holds significance due to its nutritional and medicinal contributions to both human and domestic animal health.

^aDepartment of Food Engineering and Technology, School of Engineering, Tezpur University, Napaam, Assam, India. E-mail: nickhil@tezu.ernet.in

^bSmart Farm Machinery, Centurion University, Visakhapatnam, Andhra Pradesh, India



Given the nutritional importance of citrus fruits in the human diet, efforts have been made to determine the optimal fruit maturity for enhanced nutritional value, considering factors such as internal changes in fruit flesh and external coloration of the peel during development, growth, and maturity.^{9,10}

The mandarin orange is characterized by its high fragility and a brief shelf life of 1–2 weeks at room temperature, making postharvest management challenging and resulting in substantial losses for farmers and the overall economy.¹¹ Effective temperature control is crucial during postharvest processes to preserve the overall quality of fresh citrus fruits, playing a pivotal role in their postharvest performance.¹² Inappropriate harvesting maturity leads to physiological disorders during storage, negatively impacting their shelf life and quality. Accurate determination of the fruit's maturity stage at harvest is essential for minimizing postharvest losses.^{13,14} Achieving the proper maturity stage is critical for optimal harvesting and shelf-life preservation. The international market underscores the significance of both external and internal quality standards for Khasi mandarin fruits.¹¹ Fruits are typically harvested when they reach harvestable maturity, a period determined through various computational and physical methods. This maturity stage is specific to each variety, with a fixed duration from full bloom to harvesting, allowing for the establishment of calendar dates for fruit plucking in orchards. While color change is a visible indicator of maturity, environmental factors such as temperature, relative humidity, and sunlight can influence the ripening timeline. Meeting market demands often necessitates delayed maturity. The final and crucial phase of fruit development is its maturity stage, characterized by heightened biological activity, intense metabolic processes, and cellular changes that manifest in alterations to texture, color, aroma, and flavor. Notably, the determination of commercial maturity indices in citrus fruits proves highly variable, contingent on factors such as the cultivation region, market demand, and specific varieties.¹⁵ Ensuring the fruit reaches the appropriate harvestable maturity extends its storage life and enhances postharvest storage quality.¹⁶

Deep learning models have emerged as a transformative tool in the field of image recognition. Their ability to autonomously learn hierarchical features from extensive datasets positions

them as formidable candidates for complex visual classification tasks, particularly in precisely distinguishing different ripeness stages of fruits.^{17–19,44} By training the model on a diverse set of mandarin images representing various maturity levels, the deep learning system can generalize patterns and make accurate predictions, overcoming the limitations inherent in human-dependent sorting methods. The integration of deep learning into agricultural practices not only streamlines the sorting process but also addresses scalability challenges. As the demand for fruits continues to increase, automated systems based on deep learning can handle large volumes, ensuring a consistent and efficient classification process. This not only benefits farmers and producers but also contributes to minimizing post-harvest losses and enhancing overall supply chain management.^{20,21} Fig. 1, shows the foundational framework of the CNN model designed for fruit maturation classification. In this model, every input image undergoes processing through the first, second, and third fully connected layers before ultimately reaching the output layer.

Several studies have examined machine learning and deep learning techniques for fruit classification and maturity detection (Table 1). Early methods relied on traditional image processing and feature extraction techniques, such as color, texture, and shape analysis. However, these methods often struggled with varying lighting conditions and background differences. More recent research has introduced deep learning methods, especially CNNs, to improve accuracy and reliability. Nonetheless, many existing studies either concentrate on single model architectures, use smaller or less varied datasets, or fail to assess the real-world effects of classification errors in agricultural settings.

Many studies have explored image processing and deep learning for orange fruit maturity stage detection. The successful implementation of a deep learning-based system for orange maturity classification holds profound significance. It promises to streamline the supply chain, reduce labor costs and human errors, and contribute to sustainable agriculture practices.^{27,28} Enhanced accuracy in maturity assessment enables precise harvesting, minimizing waste, optimizing resource allocation, and ultimately improving the quality of oranges

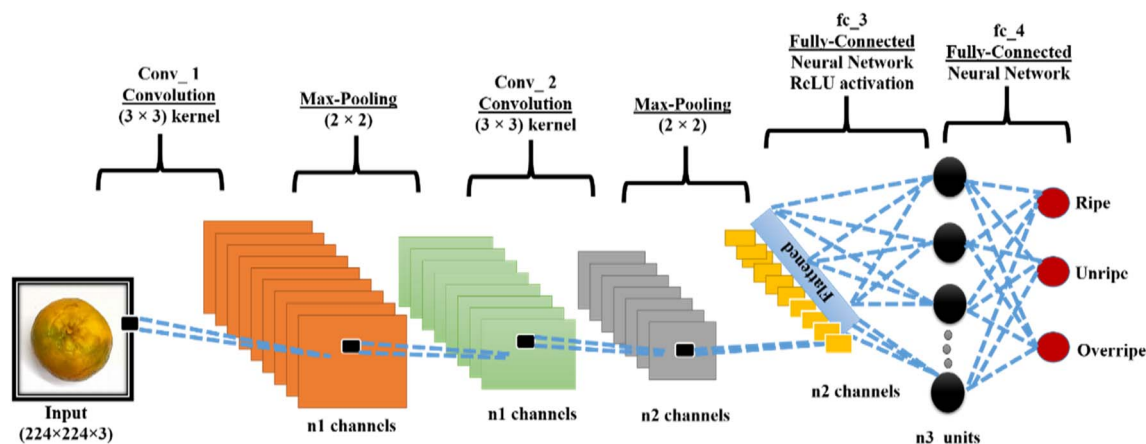


Fig. 1 The fundamental structure of the CNN model for fruit maturity classification.



Table 1 Comparative summary of existing work on various deep learning models in classifying various fruits

Crop/Fruit	Method/Model	Accuracy	Limitations	References
Cherry	CNN model	99.40%	Limited generalization	22
Multi-class fruit detection	R-CNN model	95.00%	Sensitive to background	23
Citrus fruit	CNN model	94.55%	Lacked augmentation	24
Papaya	KNN with HOG feature	99.50%	Small dataset	25
Banana	EfficientNet model	98.60%	Limited generalization	26

reaching consumers. Carolina *et al.* used image processing techniques to identify the degree of maturity in oranges,²⁹ while Saha *et al.*³⁰ applied a deep learning approach to classify and identify diseases in orange fruit. Using multi-modal input data, a deep learning approach, “Deep Orange”, was developed to detect and segment individual fruits, including oranges.⁴⁰ Asriny *et al.*³¹ proposed a classification model using Convolutional Neural Networks (CNNs) to classify orange images, achieving an accuracy of 96% with the ReLU activation function. These studies collectively demonstrate the potential of image processing and deep learning for accurately and efficiently detecting the maturity stage of the orange fruit. Through this research, we aim to contribute valuable insights into developing and deploying deep learning models for the classification of oranges by maturity, paving the way for more accurate, scalable, and cost-effective solutions in the agricultural landscape.

Unlike previous studies in unrelated fields, our research addresses a practical agricultural problem: the accurate classification of fruit maturity. This work aims to reduce postharvest losses and improve sorting efficiency. The objectives of the present research are to explore the feasibility of leveraging deep learning algorithms such as EfficientNet-B0, ResNet50, VGG16, and a Custom CNN to classify Khasi mandarin by maturity and then develop a robust and accurate deep learning model capable of discerning the maturity stage of Khasi mandarin. In this study, we examined four CNN architectures—EfficientNet-B0, ResNet50, VGG16, and a Custom CNN—to determine the best model for classifying mandarin orange maturity. These networks were chosen because they follow different design philosophies. EfficientNet-B0 focuses on parameter efficiency through compound scaling, ResNet50 incorporates residual learning for deeper architectures, VGG16 serves as a popular

benchmark for image classification tasks, and the Custom CNN was created to fit the specific needs of the dataset. By comparing these models using a balanced dataset, this work offers both a benchmark and practical guidance for researchers and practitioners in choosing models for assessing fruit maturity.

The uniqueness of this study is in its direct comparison of various deep learning architectures for classifying mandarin orange maturity under uniform experimental conditions. Unlike previous studies that often examine just one architecture, our approach shows the trade-offs in accuracy, efficiency, and reliability among different CNN designs. The results show that EfficientNet-B0, ResNet50, and the Custom CNN perform better than traditional models, such as VGG16. They also indicate their potential for future use in automated harvesting and sorting systems. This work assists readers by offering clear recommendations for selecting algorithms and highlighting the importance of AI in promoting sustainable farming practices.

2. Materials and methods

The system leverages advanced algorithms to analyze images of mandarin oranges, determining their maturity level based on visual cues. Utilizing machine learning models, the system continuously refines its ability to assess ripeness accurately over time. Furthermore, real-time communication technologies enable the prompt dissemination of guidance to farmers, allowing them to make informed decisions about harvesting based on the analyzed data.

2.1. Image acquisition

A machine vision system was created to capture images of mandarin oranges, as depicted in Fig. 2. The experimental

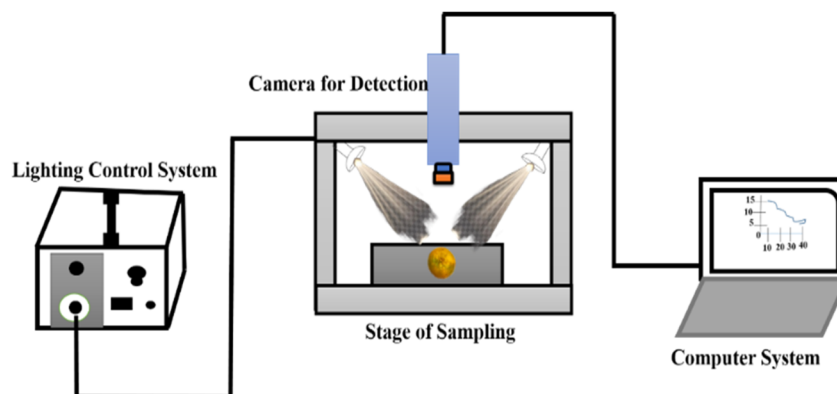
**Fig. 2** Image acquisition setup for determining the ripening stage of mandarin oranges.

Table 2 Logitech C525 webcam specification

Image sensor	C525 webcam
Frame rate	30fps @ 640 × 480 pixels
Connector	1x USB 2.0
System requirements	
Computer	512 MB RAM or more 200 MB hard drive space USB 1.1 port (2.0 recommended)
Operating system	Windows XP (SP2 or higher), windows vista or windows 7 (32 bit or 64 bit)

dataset for this study, consisting of images of mandarin oranges, was acquired using a digital color camera. A Logitech (C525) webcam was used as the weed detection sensor (Table 2). The mandarin orange images were captured off-field, with the fruits positioned on a plain white background during the photo-capturing process.

2.2. Image pre-processing

The collected dataset encompasses images of mandarin oranges. During the image acquisition phase, images with a resolution of 640 × 480 pixels were captured. Additionally, to expedite the training time of convolutional neural network (CNN) models, all images were rescaled to 224 × 224 × 3 using Python code implemented with the OpenCV framework (Mahmood *et al.*, 2022).

2.3. Dataset

For the training and testing of our pre-trained models, a concise dataset comprising 1095 images was created by capturing actual mandarin oranges representing three maturity classes. Each of the collected images was systematically labeled based on its maturity level, distinguishing between unripe, ripe, and over-ripe categories, with 365 images in each category. Fig. 3 visually represents the distinct categories of mandarin. These approaches were used to develop a robust and tailored model suitable for the specific characteristics of the custom dataset. The details of the primary dataset are presented in Table 3. This table provides comprehensive information about the primary dataset, capturing key characteristics and attributes essential for the research.

Table 3 Primary dataset information during image processing

Category	Primary datasets	
	Training	Testing
Unripe (0)	292	73
Ripe (1)	292	73
Overripe (2)	292	73
Total images	876	219

2.4. Model architecture

Deep Learning (DL) incorporating Convolutional Neural Networks (CNNs) is a prevalent choice for image processing tasks, given the ability of convolutions or filters to grasp intricate symmetrical structures, locate objects anywhere within an image, and extract abstract visual concepts through the capture of progressively intricate hierarchies.³⁹ This study employs four DL architectures: EfficientNet-B0, ResNet50, VGG16, and a Custom CNN. The proposed methodology is implemented using Keras and TensorFlow Python libraries. The custom dataset was used for the initial stages of training and validation.

2.4.1. EfficientNet-B0. As illustrated in Fig. 4, the EfficientNet-B0 architecture comprises mobile inverted bottle-neck convolution (MBConv) blocks akin to those found in MobileNetV2.⁴³ These blocks exhibit shortcut connections linking the initial and final segments, and a 1 × 1 Conv. layer expands the input block to augment the channel count or feature map depth. In contrast, Depthwise Conv. 3 × 3 and Pointwise Conv. 1 × 1 layers decrease the channels in the output block. Narrow layers featuring fewer channels are connected by shortcut connections, while wider layers are positioned between these shortcuts. This design significantly reduces both the parameter count and operational load. The EfficientNet-B0 model employs AdaptiveAvgPool2d layers to identify crucial data features and minimize training parameters. A dropout layer diminishes interdependent neuron learning. The data regression layer adopts linear regression. Despite its extensive layer count (337 layers), EfficientNet-B0 boasts merely 4 million parameters and 7.3 GFLOPs.⁴¹ This study utilizes the EfficientNet-B0 model to classify the maturity stages of mandarin oranges.

2.4.2. ResNet50. The CNN architecture introduced by He *et al.*⁴² known as ResNet, was designed to push the limits of

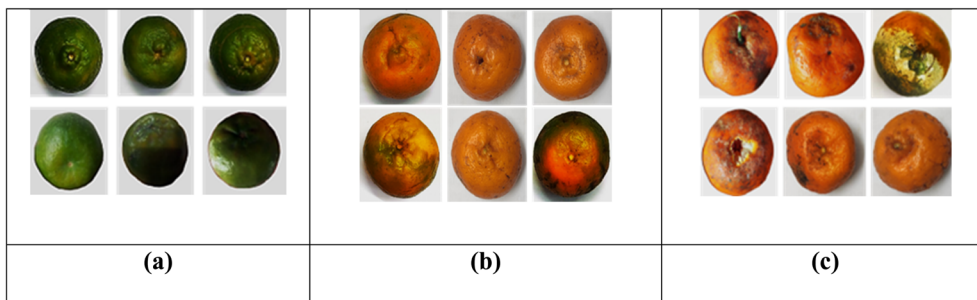


Fig. 3 The ripening stages of mandarin oranges are classified into three categories: (a) unripe, (b) ripe, and (c) over-ripe.



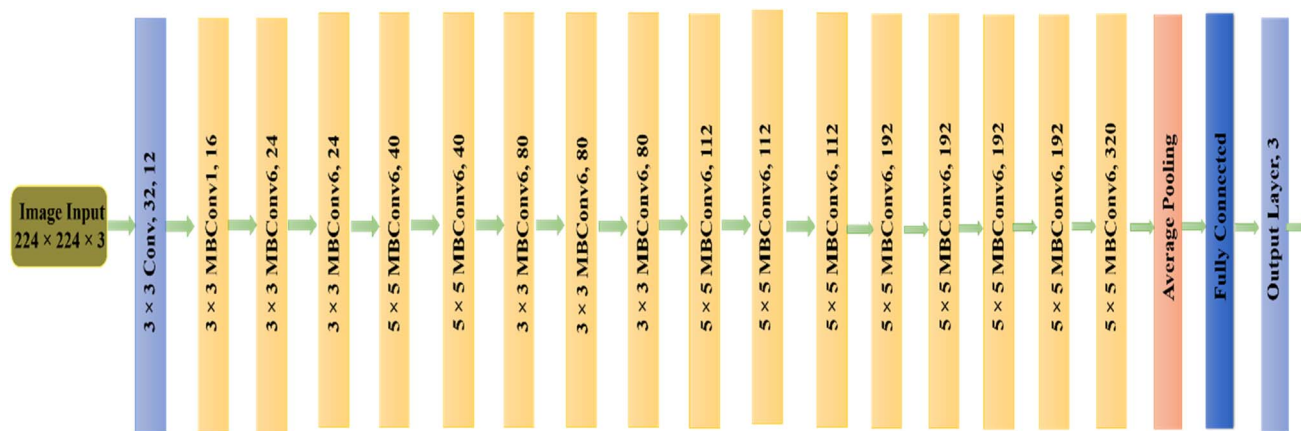


Fig. 4 A schematic representation of the EfficientNet-B0 architecture.

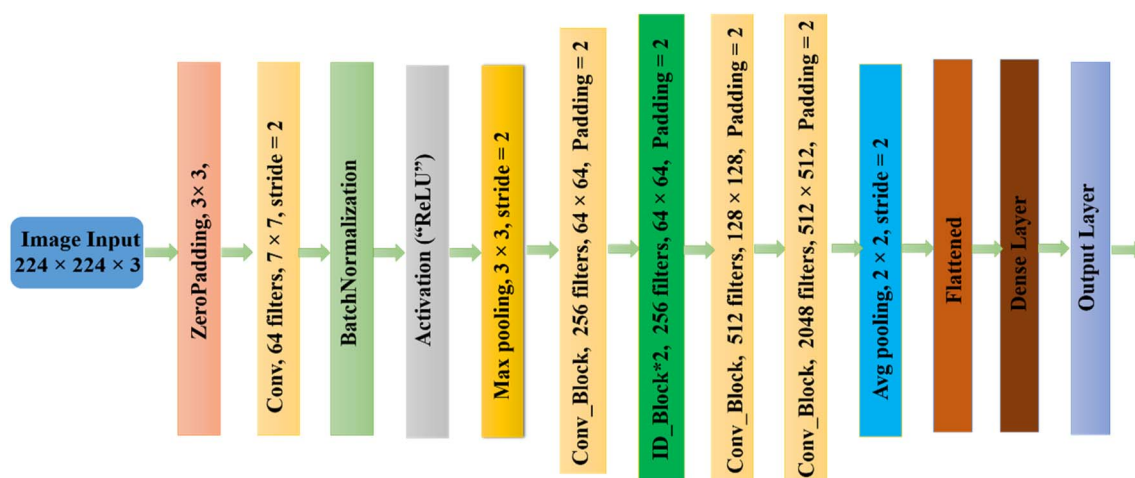


Fig. 5 Illustration of the structure of the ResNet50 model.

convolutional network depth. ResNet incorporates a network-in-network (NIN) architecture, theoretically enabling it to achieve infinite depth without compromising accuracy. In practical terms, ResNet can comprise up to 152 layers by employing 'residual blocks' throughout the network. These residual blocks, inspired by network-in-network (NIN) structures, utilize fewer convolutional layers but feature more intricate micro-neural networks. This design enables the network to focus on smaller receptive fields, thereby enhancing feature extraction compared to traditional convolutional networks that use linear filters to scan input images. ResNet's architecture comprises stacked residual blocks, combining convolution and pooling layers. Despite its similarity to AlexNet, ResNet is approximately 20 times deeper, successfully addressing the degradation problem (vanishing gradient). This study implements ResNet-50 explicitly.³² The structure of ResNet-50 is shown in Fig. 5.

2.4.3. VGG16. The configuration of the CNN-VGG16 model used in this research is presented in Fig. 6, encompassing essential components such as convolutional layers, pooling layers, dropout layers, flattening layers, and dense layers. The CNN-VGG16 model incorporates a distinctive feature by

introducing a dropout layer following the pooling layer for each convolution process. The chosen activation function is ReLU (rectified linear unit). The kernels for each convolution are sized at 3×3 , while the subsequent pooling layer is 2×2 . Specifically, 64 filters are employed for both the first and second convolutions. In the third and fourth convolutions, 128 filters are employed, while the fifth, sixth, and seventh convolutions utilize 256 filters. Subsequently, 512 filters are employed with a two-step process from the eighth to the thirteenth convolutions. The first part involves a 3×3 kernel, a 2×2 max pooling operation, and dropout regularization for the eighth, ninth, and tenth convolution processes. Following this convolution process, the second part reuses 512 filters for the eleventh, twelfth, and thirteenth layer convolutions under the same conditions. The hidden layers consist of 25 088, 4096, and 4096 nodes of neurons, implemented in sequential steps for the fully connected layers. The classification task involves three classes, utilizing a softmax classifier to simplify the process by computing the probability of all labels and delivering intuitive results, as depicted in Fig. 6. Finally, the research model is implemented using a Python script with the Keras and TensorFlow libraries.³³



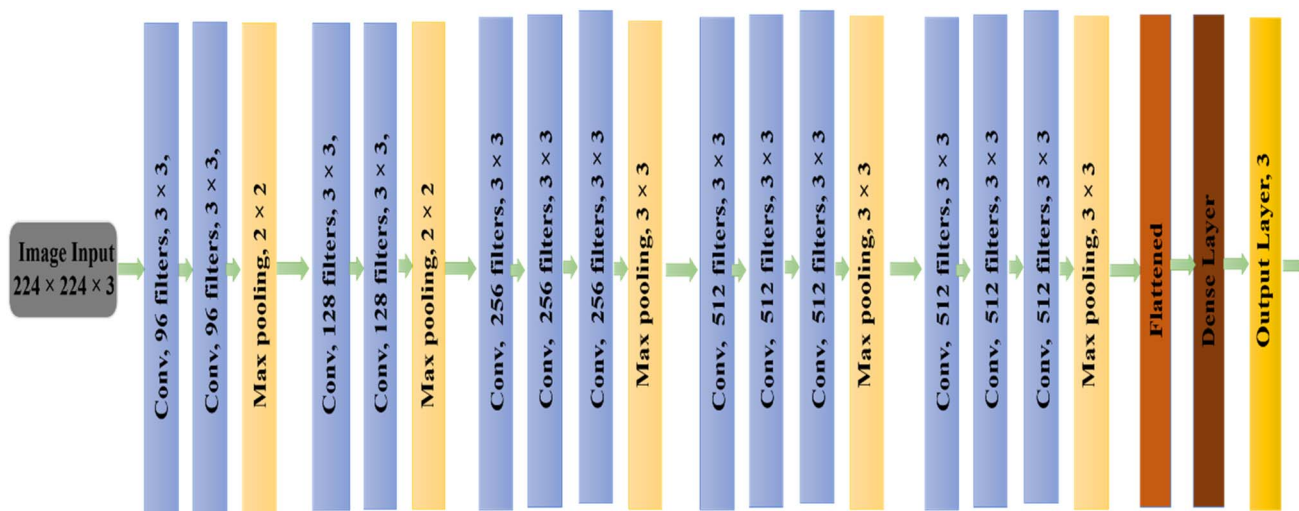


Fig. 6 Illustration of the structure of the VGG16 model.

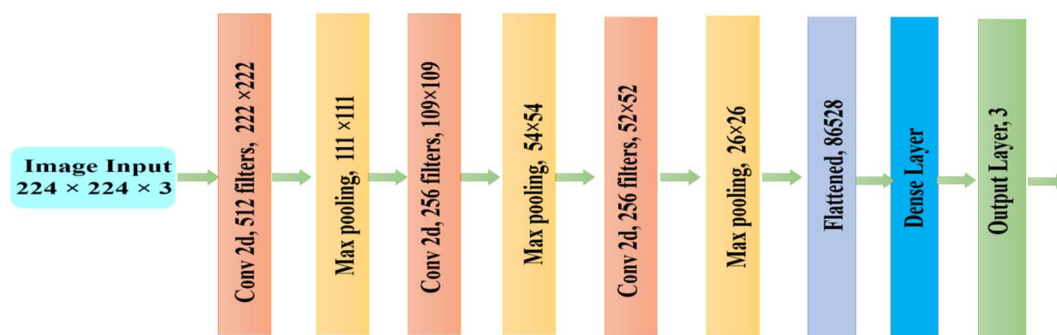


Fig. 7 The architecture of the proposed CNN model.

2.4.4. Proposed CNN model. A Convolutional Neural Network (CNN) is a type of deep learning method comprising neurons designed to process input images by assigning trainable weights and biases. CNNs can consist of tens or even hundreds of layers, and each layer is trained to recognize different aspects of an image. The architecture of a CNN enables it to learn spatial hierarchies autonomously and adaptively within data through backpropagation. This learning process involves vital building blocks such as convolution layers, pooling layers, and fully connected layers. During training, each image undergoes filtering at various resolutions, with the result of each convolved image serving as input to the subsequent layer. Starting with basic properties like brightness and borders, the filters progressively become more complex until they identify specific object characteristics. Max pooling layers extract the maximum value from the image area covered by the kernel.

This study devised a Custom CNN model with three convolutional layers and three max-pooling layers. The input image size is specified as $224 \times 224 \times 3$ for the primary dataset. The chosen loss function is the cross-entropy function, and the Adam optimizer is selected for its stability in weight and offset updates. Accuracy is incorporated to strike a balance between

training and validation. The final output layer employs the softmax activation function. The schematic representation of the proposed CNN model is depicted in Fig. 7.

2.5 Performance metrics of models

When assessing the performance of object detection models, various metrics are commonly used to gauge their accuracy, precision, recall, and overall effectiveness. These metrics provide insights into how well the model identifies and localizes objects within an image. One fundamental metric is accuracy, representing the overall correctness of the model's predictions. It is calculated as the number of correct predictions divided by the total number of predictions made, as shown in eqn (1). Precision measures the proportion of true positive predictions among all positive predictions made by the model, helping to evaluate its ability to avoid false positives, as shown in eqn (2). Recall, also known as sensitivity or true positive rate, assesses the proportion of true positive predictions among all actual positive instances in the dataset, as shown in eqn (3). The F_1 score combines precision and recall into a single metric, balancing the two, respectively, as shown in eqn (4).



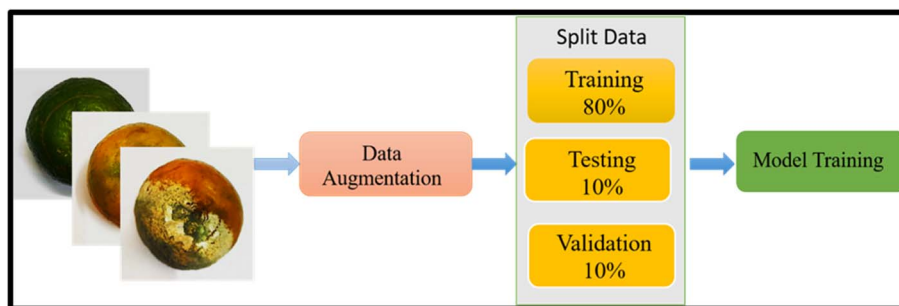


Fig. 8 A flowchart is depicted illustrating the process of dividing the data augmented dataset for model training.

Accuracy =

$$\frac{\text{true positive} + \text{true negative}}{\text{true positive} + \text{true negative} + \text{false positive} + \text{false negative}} \quad (1)$$

$$\text{Precision} = \frac{\text{true positive}}{\text{true positive} + \text{false positive}} \quad (2)$$

$$\text{Recall} = \frac{\text{true positive}}{\text{true positive} + \text{false negative}} \quad (3)$$

$$\text{F1score} = 2 \times \frac{(\text{precision} \times \text{recall})}{(\text{precision} + \text{recall})} \quad (4)$$

2.6. Data training

The primary dataset was partitioned into training, testing, and validation sets in an 80 : 10 : 10 ratio, as shown in Fig. 8. Google Colab was used as the training system. Table 4 presents the details of the corresponding training parameters, respectively.

To address the relatively small dataset size, data augmentation was applied during training, including random rotations ($\pm 20^\circ$), horizontal and vertical flips, zooming (up to 20%), brightness adjustments ($\pm 15\%$), and random shifts (up to 10%). Hyperparameters were optimized through trial and error using a validation split of the training data. The final configuration utilized the Adam optimizer with a learning rate of 0.0001, a batch size of 8, and categorical cross-entropy as the loss function. Training was performed for 200 epochs, with early stopping implemented to prevent overfitting.

3. Results and discussion

In the present study, convolutional neural network (CNN) based pre-trained models like EfficientNet-B0, ResNet50, and VGG16

were trained using the SGD optimizer with a modest batch size of 8 over 200 epochs. Cross-entropy was chosen as the loss function, promoting effective learning, and a conservative learning rate of 0.0001 was set to ensure stable convergence. Additionally, the activation function ReLU was consistently applied across all models in this investigation. These models were implemented using the TensorFlow and Keras deep learning frameworks, with pre-trained weights obtained from the primary dataset.

In evaluating the performance of the employed architecture, this research relies on precision, recall, F1-score, and accuracy metrics derived from experiments conducted within the study. Table 5 provides a comprehensive summary of the performance, which extends to four distinguished architectures: EfficientNet-B0, ResNet50, and VGG16, alongside a customized CNN which is tailored explicitly for the task. The reported accuracy rates provide a nuanced understanding of each architecture's efficacy. The evaluation of different CNN architectures reveals intriguing performance variations: EfficientNet-B0 and ResNet-50 both demonstrate high accuracy rates of 98%, highlighting their effectiveness in capturing nuanced features related to the maturity of mandarin oranges. ResNet50 also performed well, particularly in identifying ripe and overripe fruits, with near-perfect accuracy and F1 scores. The residual connections in ResNet50 likely helped it learn deep hierarchical features effectively, reducing vanishing gradient problems during training. This suggests that ResNet50 is ideal when accuracy is more important than computational cost.

VGG16, although exhibiting a slightly lower accuracy of 83%, still demonstrates commendable performance. In contrast, VGG16 showed significant variability in its performance across different classes. While it did reasonably well with unripe and overripe fruits, its performance for ripe fruits dropped sharply, with recall as low as 46%. This indicates that it often misclassifies ripe fruits, possibly due to its deeper yet less efficient architecture, which lacks residual or scaled connections. These results highlight the limitations of VGG16 in managing fine intra-class differences compared to those of more modern architectures.

Interestingly, the Custom CNN model outshines the others with an impressive accuracy rate of 99%, emphasizing the importance of tailored architectures for specific classification tasks. The proposed Custom CNN model, comprising three

Table 4 Configuration of training parameters

Hyperparameter	Value
Epoch	200
Batch size	8
Optimizer	SGD
Learning rate	0.0001
Image size	224 × 224



Table 5 Performance comparison of orange maturity stages

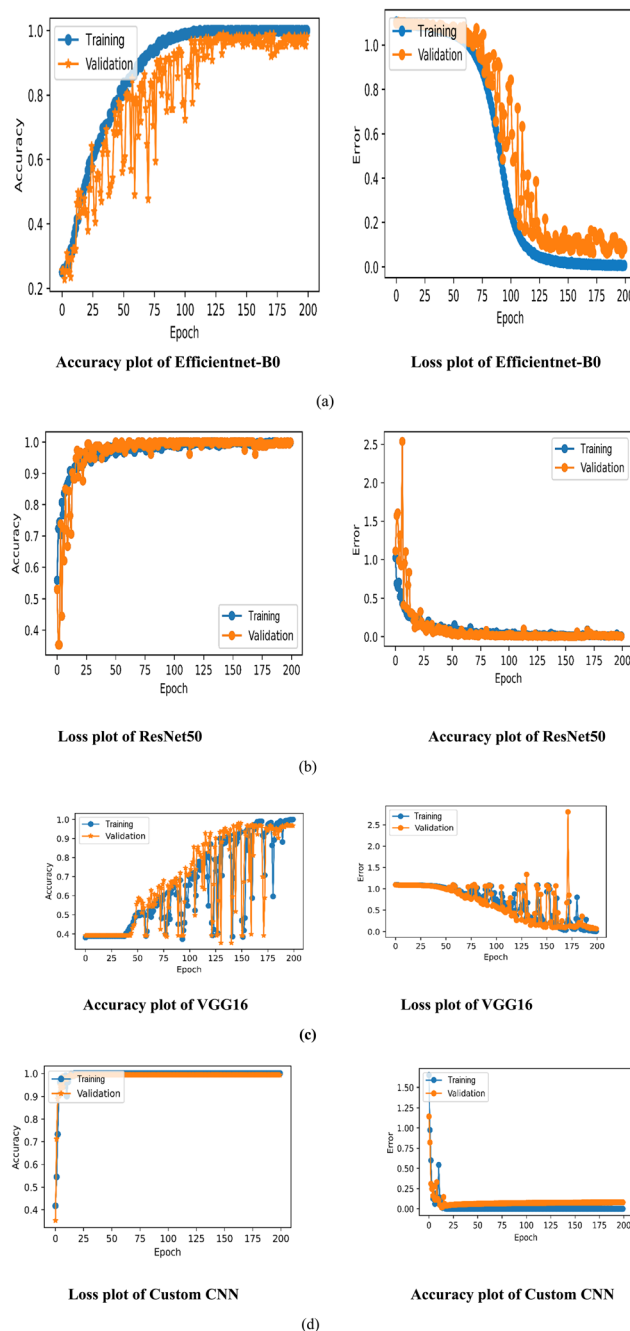
CNN architecture	Class	Precision	Recall	F1-score	Accuracy
EfficientNetB0	Unripe	98%	98%	97%	98%
	Ripe	100%	97%	99%	
	Overripe	97%	100%	98%	
ResNet50	Unripe	97%	100%	99%	98%
	Ripe	100%	98%	98%	
	Overripe	99%	100%	100%	
VGG16	Unripe	84%	91%	88%	83%
	Ripe	95%	46%	62%	
	Overripe	79%	100%	88%	
Custom CNN	Unripe	100%	100%	100%	99%
	Ripe	100%	97%	99%	
	Overripe	98%	100%	99%	

convolutional layers with 3×3 filters in each layer, incorporated 2×2 max pooling layers, a softmax classifier for image data classification, and five hidden layers. Its performance is similar to that of EfficientNetB0. However, the Custom CNN is a very small network with few convolutional layers and few dense layers; it showed better results compared to EfficientNetB0, ResNet50, and VGG16. The Custom CNN achieved perfect classification in the unripe category and high precision and recall across all classes, similar to EfficientNetB0 and ResNet50. This shows the promise of tailored lightweight architectures for specific applications. The overall accuracy of 99% suggests that with careful design, even simpler CNNs can perform as well as advanced pre-trained networks, which can be beneficial for use in devices with limited resources.

As it has fewer training parameters than EfficientNetB0, ResNet50, and VGG16, the model takes less time to train and exhibits lower latency during testing. This customized approach highlights the importance of domain-specific design choices in achieving optimal results and underscores the nuanced intricacies of agricultural image classification.

The loss curves and accuracy curves, meticulously plotted with the x-axis representing the number of epochs and the y-axis portraying corresponding values, provide a visual narrative of the training process and model evolution shown in Fig. 9. These curves offer insights into convergence patterns, optimal training epochs, and the impact of hyperparameters on the learning dynamics of the models. The loss curves illustrate the model's ability to minimize error during training, showcasing trends and potential challenges in the optimization process. Simultaneously, the accuracy curves reveal the model's proficiency in correctly classifying mandarin oranges at various maturity stages over epochs.

In addition to accuracy, precision, recall, and F1-scores, we calculated the area under the receiver operating characteristic curve (AUC) to provide a better measure of classification performance. The AUC values for EfficientNet-B0, ResNet50, and the Custom CNN were all above 0.97. VGG16 achieved an AUC of 0.89. These results further confirm the strong performance and reliability of EfficientNet-B0, ResNet50, and the Custom CNN in distinguishing maturity stages of mandarin oranges. The confusion matrices for each model were analyzed

**Fig. 9** Training and validation performance graphs for: (a) EfficientNet-B0, (b) ResNet50, (c) VGG16, and (d) Custom CNN.

to examine misclassification patterns. The results showed that most errors occurred between the ripe and overripe categories. Misclassifications between unripe and other classes were relatively rare. This suggests that visual differences between ripe and overripe fruits are more subtle, leading to occasional overlap in classification. From an agricultural viewpoint, misclassifying fruit ripeness can lead to serious problems. When ripe fruits are wrongly labeled as overripe, they may be rejected from the market too soon, causing waste after harvest. On the other hand, if overripe fruits are seen as ripe, they enter the



supply chain with a shorter shelf life. This can lead to quality issues and complaints from consumers. Additionally, if unripe fruits are mistakenly identified as ripe, they might be picked or sold too early. This results in poor taste, less satisfaction for consumers, and ultimately lower profits for producers. In comparison to previously discussed studies, our proposed system employing EfficientNet-B0, ResNet50, VGG16, and a Custom CNN has demonstrated superior accuracy levels. The comparison of the four architectures shows clear differences in how well they classify fruit ripening stages. EfficientNetB0 consistently achieved high precision, recall, and F1-scores for unripe, ripe, and overripe stages. This demonstrates its ability to balance feature extraction and computational efficiency. Its lightweight design, with compound scaling, helps it capture detailed visual features without overfitting, which explains its stable performance. For instance, Arampongsanuwat *et al.*³⁴ reported accuracy rates of 77.50% for VGG16, 75% for AlexNet, 79% for ResNet50, and 78% for InceptionV3 in mangosteen ripeness classification. However, our proposed system achieved notably higher accuracy rates of 83% for VGG16 and an impressive rate of 98% for ResNet50, showcasing the superior performance of our approach in this context. Similarly, Kusakunniran *et al.*³⁵ attained a remarkable accuracy rate of 100% for DenseNet121, EfficientNetB0, ResNet50, and VGG16. Nonetheless, our proposed system yielded slightly lower accuracy rates of 98% for EfficientNetB0 and ResNet50, and 83% for VGG16. While these figures indicate slightly lower performance compared to the referenced study, they still demonstrate competitive results for our proposed system. Furthermore, Al-Masawabe *et al.*³⁶ achieved a perfect accuracy of 100% in classifying bananas based on ripeness categories. In contrast, our proposed model achieved an accuracy of 83% for VGG16, indicating comparable performance in this specific classification task. Additionally, Mahmood *et al.*³⁷ reported high accuracy rates for AlexNet and VGG16 in classifying jujube fruits into maturity categories. Despite their reported accuracies of 94.17% and 97.65% for actual and augmented images, respectively, our proposed system achieved a lower accuracy of 83% for VGG16 in a similar classification task. Moreover, Nasiri *et al.*³⁸ achieved an impressive accuracy rate of 98.4% using VGG16 for classifying mandarin oranges based on maturity levels. In contrast, our proposed system achieved an accuracy rate of 83% for the same task, indicating comparatively lower performance. While our proposed system may exhibit slightly lower accuracy rates compared to some referenced studies, it still demonstrates competitive performance across various classification tasks. It's worth noting that the standard accuracy of our models, computed by averaging accuracy over actual and augmented datasets, provides a comprehensive measure of their overall performance. Comparatively, the methods reported in other studies also underwent similar standard accuracy computations, facilitating a meaningful comparison across different approaches. Overall, the results indicate that modern, scalable architectures like EfficientNet and ResNet, along with Custom CNNs designed for specific domains, are more effective for classifying fruit ripening stages than older architectures, such as VGG16. In addition, the near-perfect accuracies imply that

these proposed models are ready for real-world applications, such as automated grading systems in post-harvest management and supply chain monitoring. Future research could focus on testing how well these models generalize across different lighting conditions, backgrounds, and types of fruit to further establish their practical use.

4. Conclusion

In this comprehensive study, the efficacy of three renowned pre-trained deep learning models, EfficientNet-B0, ResNet50, and VGG16, as well as a Custom CNN, was explored for the crucial task of classifying mandarin oranges based on their maturity levels. Leveraging the concept of transfer learning, these models were trained on a dataset comprising 876 actual images of mandarin oranges, categorized into three distinct maturity classes. The EfficientNet-B0 model achieved an impressive accuracy of 98%, showcasing its efficiency in discerning different maturity stages of mandarin oranges. Similarly, the ResNet50 model also exhibited a commendable accuracy of 98%, emphasizing its robust performance in image classification tasks. The VGG16 model, while slightly trailing behind with an 83% accuracy, still demonstrated a significant capability in distinguishing maturity levels. Furthermore, the Custom CNN outperformed all other models with an exceptional accuracy of 99%, highlighting the effectiveness of a tailored convolutional neural network in this specific application. The findings not only highlight the prowess of these pre-trained models in fruit maturity classification but also emphasize the practical implications of such advancements in the context of automated agricultural processes, showcasing their potential for real-world applications in the fruit industry. The best-performing models, ResNet50, EfficientNet-B0, and the Custom CNN, have implications for enhancing agricultural practices, quality assessment, and broader applications in the food industry. The method has practical uses in automated sorting and grading systems for fruit processing industries. Rapid and accurate ripeness detection can decrease the need for labor and reduce human error. It can also be used in supply chain and storage management to minimize post-harvest losses by ensuring fruits are transported and stored at the right ripening stages. Additionally, this method has potential for smart retail systems, allowing consumers to assess ripeness objectively. It can also benefit precision agriculture, where IoT-based decision support systems can help guide optimal harvesting. The findings reinforce that this method is not only technically sound but also has broad industrial and commercial importance.

Ethical approval

This study does not present any ethical concerns.

Author contributions

Raj Singh: conceptualization, data curation, and writing – review and editing; C. Nickhil: investigation, writing and editing; Konga Upendar: investigation; Poonam Mishra: review and



editing; Sankar Chandra Deka: investigation and final checking of the manuscript.

Conflicts of interest

The authors declare that there is no conflict of interest.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

- 1 A. K. Singh, N. T. Meetei, B. K. Singh and N. Mandal, Khasi mandarin: Its importance, problems and prospects of cultivation in North-eastern Himalayan region, *Int. J. Agric. Environ. Biotechnol.*, 2016, **9**(4), 573–592.
- 2 J. Tariang, D. Majumder and H. Papang, Efficacy of native *Bacillus subtilis* against postharvest *Penicillium* rot pathogen *Penicillium* sp. of Khasi mandarin oranges in Meghalaya, India, *Int. J. Curr. Microbiol. Appl. Sci.*, 2018, **7**(12), 447–460.
- 3 N. A. Deshmukh, R. K. Patel, H. Rymbai, A. K. Jha and B. C. Deka, Fruit maturity and associated changes in Khasi mandarin (*Citrus reticulata*) at different altitudes in humid tropical climate, *Indian J. Agric. Sci.*, 2016, **86**(7), 854–859.
- 4 C. Sangma, T. Anbazhagan and D. J. Rajkhowa, Influence of soil health and age of orchard on fruit quality of Khasi mandarin in Nagaland, in *27th Natl Conf Soil Conserv Soc India on Sustainable Management of Soil and Water Resources for Doubling Farmers' Income*, AAU, Jorhat, 2018.
- 5 R. Sanabam, N. S. Singh, P. J. Handique and H. S. Devi, Disease-free Khasi mandarin (*Citrus reticulata* Blanco) production using in vitro microshoot tip grafting and its assessment using DAS-ELISA and RT-PCR, *Sci. Hortic.*, 2015, **189**, 208–213.
- 6 T. K. Hazarika, Citrus genetic diversity of north-east India, their distribution, ecogeography and ecobiology, *Genet. Resour. Crop Evol.*, 2012, **59**, 1267–1280.
- 7 K. Sharma, N. Mahato and Y. R. Lee, Extraction, characterization and biological activity of citrus flavonoids, *Rev. Chem. Eng.*, 2019, **35**(2), 265–284.
- 8 D. A. Zema, P. S. Calabrò, A. Folino, V. Tamburino, G. Zappia and S. M. Zimbone, Valorisation of citrus processing waste: A review, *Waste Manag.*, 2018, **80**, 252–273.
- 9 K. Kashyap, D. Kashyap, M. Nitin, N. Ramchiary and S. Banu, Characterizing the nutrient composition, physiological maturity, and effect of cold storage in Khasi mandarin (*Citrus reticulata* Blanco), *Int. J. Fruit Sci.*, 2020, **20**(3), 521–540.
- 10 R. Singh, C. Nickhil, R. Nisha, K. Upendar, B. Jithender and S. C. Deka, A comprehensive review of advanced deep learning approaches for food freshness detection, *Food Eng. Rev.*, 2025, **17**(1), 127–160.
- 11 P. R. Rokaya, D. R. Baral, D. M. Gautam, A. K. Shrestha and K. P. Paudyal, Effect of altitude and maturity stages on quality attributes of mandarin (*Citrus reticulata* Blanco), *Am. J. Plant Sci.*, 2016, **7**(6), 958.
- 12 Z. Tietel, E. Lewinsohn, E. Fallik and R. Porat, Importance of storage temperatures in maintaining flavor and quality of mandarins, *Postharvest Biol. Technol.*, 2012, **64**(1), 175–182.
- 13 A. Saikumar, C. Nickhil and L. S. Badwaik, Physicochemical characterization of elephant apple (*Dillenia indica* L.) fruit and its mass and volume modeling using computer vision, *Sci. Hortic.*, 2023, **314**, 111947.
- 14 A. Saikumar, A. Sahal, S. M. Mansuri, A. Hussain, P. M. Junaid, C. Nickhil, *et al.*, Assessment of physicochemical attributes and variation in mass-volume of Himalayan pears: Computer vision-based modeling, *J. Food Compos. Anal.*, 2025, **137**, 106955.
- 15 J. Y. Zhang, Q. Zhang, H. X. Zhang, Q. Q. Ma, J. Q. Lu and Y. J. Qiao, Characterization of polymethoxylated flavonoids in the peels of Shatangju mandarin (*Citrus reticulata* Blanco) by online high-performance liquid chromatography coupled to photodiode photodiode array detection and electrospray tandem mass spectrometry, *J. Agric. Food Chem.*, 2012, **60**(36), 9023–9034.
- 16 M. A. H. Talukder, M. M. Rahman, M. Hossain, M. A. K. Mian and Q. A. Khaliq, Determination of maturity indices in mandarin orange, *Ann. Bangladesh Agric.*, 2015, **19**, 33–42.
- 17 R. Singh, C. Nickhil, R. Nisha, K. Upendar and S. C. Deka, Investigating the effect of oxygen, carbon dioxide, and ethylene gases on Khasi mandarin orange fruit during storage, *ACS Agric. Sci. Technol.*, 2024, **4**(11), 1206–1215.
- 18 R. Singh, C. Nickhil, K. Upendar, S. C. Deka and R. Nisha, Non-destructive estimation of mandarin orange fruit quality during the ripening stage using machine-learning-based spectroscopic techniques, *J. Food Meas. Char.*, 2025, **19**(2), 862–875.
- 19 P. Chu, *Robust Fruit Detection and Localization for Robotic Harvesting Dissertation*, Michigan State University, 2023.
- 20 S. K. Chakraborty, A. Subeesh, K. Dubey, D. Jat, N. S. Chandel, R. Potdar and D. Kumar, Development of an optimally designed real-time automatic citrus fruit grading-sorting machine leveraging computer vision-based adaptive deep learning model, *Eng. Appl. Artif. Intell.*, 2023, **120**, 105826.
- 21 S. W. Chen, S. S. Shivakumar, S. Dcunha, J. Das, E. Okon, C. Qu and V. Kumar, Counting apples and oranges with deep learning: A data-driven approach, *IEEE Rob. Autom. Lett.*, 2017, **2**(2), 781–788.
- 22 M. Momeny, A. Jahanbakhshi, K. Jafarnejhad and Y. D. Zhang, Accurate classification of cherry fruit using deep CNN based on hybrid pooling approach, *Postharvest Biol. Technol.*, 2020, **166**, 111204.
- 23 S. Wan and S. Goudos, Faster R-CNN for multi-class fruit detection using a robotic vision system, *Comput. Netw.*, 2020, **168**, 107036.
- 24 A. Khatkhat, M. U. Asghar, U. Batool, M. Z. Asghar, H. Ullah, M. Al-Rakhami, *et al.*, Automatic detection of citrus fruit and leaves diseases using deep neural network model, *IEEE Access*, 2021, **9**, 112942–112954.



- 25 S. K. Behera, A. K. Rath and P. K. Sathy, Maturity status classification of papaya fruits based on machine learning and transfer learning approach, *Inf. Process. Agric.*, 2021, **8**(2), 244–250.
- 26 N. Ismail and O. A. Malik, Real-time visual inspection system for grading fruits using computer vision and deep learning techniques, *Inf. Process. Agric.*, 2022, **9**(1), 24–37.
- 27 C. Nickhil, S. M. Mansuri, A. Saikumar, P. M. Junaid, R. Nisha, L. S. Badwaik, *et al.*, Estimation of mass and volume of freshly harvested Assam lemon (*Citrus limon* Burm L.) using computer vision: Exploring changes on different storage days, *Appl. Fruit Sci.*, 2025, **67**(3), 141.
- 28 S. M. Mansuri, P. V. Gautam, D. Jain and C. Nickhil, Computer vision model for estimating the mass and volume of freshly harvested Thai apple ber (*Ziziphus mauritiana* L.) and its variation with storage days, *Sci. Hortic.*, 2022, **305**, 111436.
- 29 C. P. D. Carolina and N. T. D. David, Classification of oranges by maturity, using image processing techniques, in *2014 III International Congress of Engineering Mechatronics and Automation (CIIMA)*, IEEE, 2014, pp. 1–5.
- 30 R. Saha, Orange fruit disease classification using deep learning approach, *Int. J. Adv. Trends Comput. Sci. Eng.*, 2020, 2297–2301.
- 31 D. M. Asriny, S. Rani and A. F. Hidayatullah, Orange fruit images classification using convolutional neural networks, *IOP Conf. Ser. Mater. Sci. Eng.*, 2020, **803**(1), 012020.
- 32 S. H. M. Ashtiani, S. Javanmardi, M. Jahanbanifard, A. Martynenko and F. J. Verbeek, Detection of mulberry ripeness stages using deep learning models, *IEEE Access*, 2021, **9**, 100380–100394.
- 33 J. Pardede, B. Sitohang, S. Akbar and M. L. Khodra, Implementation of transfer learning using VGG16 on fruit ripeness detection, *Int. J. Intell. Syst. Appl.*, 2021, **13**(2), 52–61.
- 34 S. Arampongsanuwat and O. Chaowalit, Application of deep convolutional neural networks for mangosteen ripeness classification, *ICIC Express Lett.*, 2021, **15**, 649–657.
- 35 W. Kusakunniran, T. Imaromkul, K. Aukkapinyo, K. Thongkanchorn and P. Somsong, Automatic classification of mangosteens and ripe status in images using deep learning-based approaches, *Multimed. Tool. Appl.*, 2023, 1–16.
- 36 M. M. Al-Masawabe, L. F. Samhan, A. H. AlFarra, Y. E. Aslem and S. S. Abu-Naser, Papaya maturity classifications using deep convolutional neural networks, *Int. J. Eng. Inf. Syst.*, 2021, **5**(12), 60–67.
- 37 A. Mahmood, S. K. Singh and A. K. Tiwari, Pre-trained deep learning-based classification of jujube fruits according to their maturity level, *Neural Comput. Appl.*, 2022, **34**(16), 13925–13935.
- 38 A. Nasiri, A. Taheri-Garavand and Y. D. Zhang, Image-based deep learning automated sorting of date fruit, *Postharvest Biol. Technol.*, 2019, **153**, 133–141.
- 39 N. Aherwadi, U. Mittal, J. Singla, N. Z. Jhanjhi, A. Yassine and M. S. Hossain, Prediction of fruit maturity, quality, and its life using deep learning algorithms, *Electronics*, 2022, **11**(24), 4100.
- 40 P. Ganesh, K. Volle, T. F. Burks and S. S. Mehta, Deep orange: Mask R-CNN-based orange detection and segmentation, *IFAC-Papersonline*, 2019, **52**(30), 70–75.
- 41 Q. H. Phan, V. T. Nguyen, C. H. Lien, T. P. Duong, M. T. K. Hou and N. B. Le, Classification of tomato fruit using Yolov5 and convolutional neural network models, *Plants*, 2023, **12**(4), 790.
- 42 K. He, X. Zhang, S. Ren and J. Sun, Deep residual learning for image recognition, in *Proc IEEE Conf Comput Vis Pattern Recognit*, 2016, pp. 770–778.
- 43 M. Sandler, A. Howard, M. Zhu, A. Zhmoginov and L. C. Chen, Mobilenetv2: Inverted residuals and linear bottlenecks, in *Proc IEEE Conf Comput Vis Pattern Recognit*, 2018, pp. 4510–4520.
- 44 R. Singh, R. Nisha, R. Naik, K. Upendar, C. Nickhil and S. C. Deka, Sensor fusion techniques in deep learning for multimodal fruit and vegetable quality assessment: A comprehensive review, *J. Food Meas. Char.*, 2024, **18**(9), 8088–8109.

