Fine-Tuned MobileNetV2 for Multi-Fruit Ripeness Classification Using Deep Transfer Learning

Md. Shazedur Rahman[1], Rofidul Hasan Ovik[1], Sadman Sadik Khan[1*], AKM Bahalul Haque[2], Sudipta Das Gupta[3] and Md Nyem Hasan Bhuiyan[4]

[1] Dept. of CSE, Daffodil International University, Dhaka, Bangladesh
[2] Dept. of Engineering and Information Technology, Abo Akademi University, Turku, Finland
[3] Dept. of Agricultural and Biosystems Engineering, North Dakota State University, North Dakota, USA
[4] Dept. of CSE, Dhaka International University, Dhaka, Bangladesh

{shazedur15-14056, rofidul15-14160, sadman15-13696}@diu.edu.bd*,
akmbahalul.haque@abo.fi, sudipta.gupta@ndsu.edu, nyem.bhuiayan@gmail.com

Abstract. Accurate fruit ripeness detection plays a vital role in ensuring food quality, minimizing waste, and supporting automation in agriculture. This research investigates deep learning techniques for classifying the ripeness of three fruits—apples, bananas, and oranges—into fresh, unripe, and rotten categories. Four pre-trained convolutional neural network models—MobileNetV2, InceptionV3, ResNet50, and DenseNet121—were applied to both individual fruit datasets and a combined dataset. All models performed exceptionally well on single-fruit classification, with MobileNetV2 achieving perfect accuracy. When evaluated on the combined dataset, MobileNetV2 maintained strong performance with 97% accuracy, which further improved to 99% after fine-tuning. The fine-tuning process involved freezing the base model, training 75 layers, and adding a global average pooling layer, dropout, early stopping, and a learning rate scheduler. These results demonstrate the effectiveness of a lightweight, fine-tuned CNN model for robust and high-accuracy fruit ripeness classification, making it suitable for real-world deployment.

Keywords: Fruit Ripeness Classification, Deep learning, Transfer Learning, MobileNetV2, Agricultural automation, Food quality control, Computer vision.

1 Introduction

Ripeness detection of fruits is a crucial process in the agricultural and food supply chain with a direct influence on the timing of harvesting, ripening, storage, dispatch, and consumer satisfaction. Conventional methods of assessing ripeness are more manual, subjective, and tedious, hence unsuitable for mass-scale adoption. Therefore, in recent years, major interest has been put on the automated systems utilizing computer vision and deep learning [1], [2].

M. S. Arefin et al. (eds.), *Proceedings of the International Conference on Intelligent Data Analysis and Applications (IDAA 2025)*, Advances in Intelligent Systems Research 206,
https://doi.org/10.2991/978-94-6239-664-7_54

Deep learning has brought about a revolution in image classification tasks, thanks to the intervention of convolutional neural networks (CNNs), which are capable of extracting hierarchical features from raw pixel data. CNNs have proven their worth as state-of-the-art algorithms in various visual recognition tasks, including MobileNetV2, InceptionV3, ResNet50, and DenseNet121 [3], [4]. Since these models were initially trained on large datasets, such as ImageNet, they can be applied to very specific domain problems. The technique of transfer learning ensures high accuracy for these models with very little data [5].

This study focuses on classifying fruit ripeness in three commonly eaten fruits: apples, bananas, and oranges. The fruit is classified into three stages of ripeness. We first evaluate the performance of four pre-trained CNN models individually on fruit datasets. All models reach 100% classification accuracy, thus demonstrating their efficiency in constrained classification tasks. We carry on our experimentation on a single dataset comprising all fruits to represent a more challenging and realistic scenario. Among the models evaluated, MobileNetV2 achieves the best performance with 97% accuracy. Subsequently, fine-tuning is applied so as to gain better model generalization, and accuracy improves to 99% thereafter.

Our results indicate that while pre-trained models perform extremely well in simpler classification tasks, it is fine-tuning that boosts performance rates in complex multi-class cases. This work advances smart agriculture by providing a practical and scalable solution for the automated detection of fruit ripeness.

2 Literature Review

Fruit automation classification and ripeness detection are pivotal areas of precision agriculture, directed towards improving post-harvest handling, reducing human labor, and enhancing consistency in quality control. The majority of the early research in this area used classical image processing techniques on handcrafted features. For example, Bhosale and Patil [1] focused on fruit recognition based on color, shape, and texture features. However, such approaches provided reasonable performance in controlled environments, but not in real-world scenarios where ambient lighting, occlusions, and natural variations in appearance affect fruit recognition.

To overcome these shortcomings, researchers then shifted to machine learning methods that integrated both pipelines for feature extraction and classification. Support Vector Machine (SVM), k-Nearest Neighbors (k-NN), Decision Tree, and other classifiers were usually coupled with histogram features such as Histogram of Oriented Gradients (HOG), Local Binary Patterns (LBP), and color histograms. For instance, Sahu et al. [6] reported an SVM-based classification of banana ripeness using hue-based color features, yielding satisfactory results. Nevertheless, these traditional approaches also required careful feature selection and preprocessing, which limited their scalability potential.

Furthermore, with the significant advances in deep learning, particularly in convolutional neural networks (CNNs), substantial improvements in accuracy and generalization have been achieved in numerous computer vision applications, including fruit classification. Unlike traditional methods, which require manual engineering of features, CNNs automatically learn hierarchical spatial features from raw images. Rahnemoonfar and Sheppard [2] outlined a deep simulated learning model regarding fruit counting in orchards. This model has been found to generalize well to real-world fruit distributions, despite being trained on a synthetic dataset. Likewise, Fuentes et al. [7] employed deep learning to detect and classify diverse fruits simultaneously under natural atmospheres, indicating the promise of CNNs in contrast-distinguished classification.

Several of these studies are focused on ripeness classification using deep learning. Arakeri and Lakshmi [8] employed a CNN-based architecture for determining mango ripeness, indicating that excellent classification performance could be achieved when RGB images were classified into maturity stages. This study's premise is that deep models bring better performance in identifying minor color and texture alterations associated with fruit ripening. Kamilaris and Prenafeta-Boldú [9] conducted a comprehensive survey of deep learning in agriculture, finding that CNNs were rapidly gaining popularity in tasks such as fruit and plant disease detection, ripeness classification, and yield estimation.

Transfer learning, a promising area for certain applications in agriculture, where models initially trained on large datasets (such as ImageNet) are further fine-tuned on domain-specific data, has been found to be effective. Amara et al. [10] applied a MobileNet-based transfer learning approach for banana disease and ripeness detection, which outperformed existing methods with decreased training time. The lightweight architecture, such as MobileNet or EfficientNet, has become popular for that reason-its adoption is appropriate for characterizing real-time and embedded systems due to the balance between accuracy and low computational demand.

Encouraging results have also been generated by InceptionV3, ResNet-50, and DenseNet-121 when applied to agricultural functions. Mohanty et al. [11] demonstrated that ResNet and Inception networks are effective in classifying plant diseases. Similarly, Sa et al. [12] utilized DenseNet with hyperspectral imaging for detecting tomato ripeness. These models learned to capture deeper, abstract features that would aid in fine-grained separation between stages of ripeness.

Other efforts have been done with ensemble and hybrid approaches. Bhange and Hingoliwala [13] proposed a hybrid model that combines CNNs and SVM classifiers to enhance classification performance. Mureşan and Oltean [14] proposed a large dataset of fruit images specific for fruit classification and used deep CNNs to set the benchmark for multi-class recognition of fruits later on.

Most of the currently available studies now focus on either single fruit types or a few classes. Even fewer are common in their assessments of recognized ripeness classification across different fruits in a unified model, one more in line with real-world cases such as mixed-fruit quality assessments in warehouses or markets. Moreover, the advantages of fine-tuning pre-trained models for multi-fruit, multi-class classification remain largely unexplored.

This research effectively compares pre-trained models in the state-of-the-art, namely MobileNetV2, InceptionV3, ResNet50, and DenseNet121, for each fruit individually and in combination, across datasets of three fruits: apples, bananas, and oranges, each with three ripeness classes. It, therefore, illustrates the contributions of fine-tuning before and after comparison, highlighting the necessity of transfer learning for improving complex classification and generalization in real-world applications. This could potentially provide enriching insights for developing scalable systems applicable in real-world contexts for automating fruit ripeness detection.

3 Methodology

This section outlines the methodological approach employed in our study, including data acquisition, preprocessing steps, model architecture, training procedures, and evaluation metrics.

3.1 Data Collection

The dataset analyzed in this work was sourced from a public library on Kaggle [15]. It contains labeled photographs of three types of fruits—namely, apple, banana, and orange. Each fruit has been classified under three stages of ripeness: unripe, ripe, and overripe, thereby constituting a total of nine classes. The dataset comprises color images with varying resolutions and backgrounds that mirror real-life conditions. Furthermore, no synthetic images or augmentation were created at this stage, allowing the dataset to accurately represent the natural variations in lighting and fruit appearance.

Table 1. Data Distribution

Class		Number of images
Apple	freshapples	1693
	rottenapples	2342
	unripe apple	1934
Orange	freshoranges	1466
	rottenoranges	1595
	unripe oranges	1285
Banana	freshbanana	1581
	rottenbanana	2224
	unripe banana	2097
Total		**16217**

3.2 Data Preprocessing

For the deep learning models' compatibility and consistency, all images were resized to a standard input size of 224 × 224 pixels for CNN models, such as MobileNetV2 and ResNet-50 [16]. The images were converted to their RGB computations, and pixels were normalized by the ratio they scale to in the range [0,1]. The dataset was then randomly split into training, validation, and test sets with an 80:10:10 ratio. This consistent split was used for all experiments to ensure a uniform evaluation framework.

3.3 Model Selection

We examined four state-of-the-art deep learning models: MobileNetV2 [17], InceptionV3 [18], ResNet50 [19], and DenseNet121 [20]. The aforementioned architectures are highly popular and have demonstrated their effectiveness in various image classification-related benchmarks, including the ImageNet dataset. All four models utilized pre-trained weights that were readjusted for the present classification task of nine classes (three degrees of ripeness for each of the three fruit types).

3.4 Baseline Experimentation

In Phase 1, the preset models were used to convert them into a fixed feature extractor. All layers of the base models were frozen (non-trainable), and only the last classification head was replaced with our specific 9-class setup, featuring a custom fully connected output layer. This was intended to test how models convert generic visual features into those relevant to fruit ripeness classification. Each of the four models was independently trained and evaluated on each of the three fruit datasets (apple, banana, and orange).

3.5 Model Fine-Tuning

To make the classification task more realistic and complex, we consolidated all three types of fruit in the dataset and set three different ripeness classes for each fruit (9 labels in total). To improve the performance of the model, fine-tuning on MobileNetV2 was conducted with the following strategy:

Freezing Strategy: For the initial phase of training, the entire MobileNetV2 base was frozen. When initial convergence was reached, the top 75 layers of the network were selectively unfrozen for fine-tuning higher-level features pertinent to ripeness classification [21]. This technique retains lower-level features while adjusting the high-level ones, balancing generalization with specialization.

Custom Layers: On top of the MobileNetV2 base, we appended a GAP layer for feature dimension reduction and spatial invariance. This was followed by a Dropout layer (rate=0.3) to prevent overfitting and a Dense layer with Softmax activation appended to classify the input as one of nine output classes

Table 2. Summary of the Fine-Tuned Model Architecture.

Layer (type)	Output Shape	Param #
input_layer_5	(None, 224,224,3)	0
mobilenetv2_1.00_224	(None, 7,7,1280)	2,257,984
global_average_pooling2d_2	(None, 1280)	0
dense_8	(None, 256)	327,936
dropout_4	(None, 256)	0
dense_9	(None, 9)	2,313

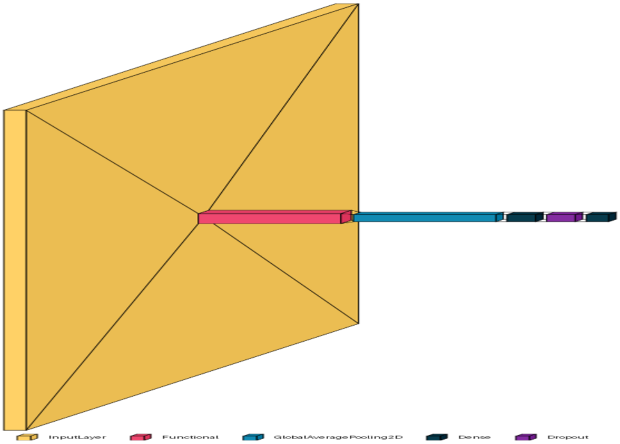


Fig. 1. Model Architecture of the Finetuned MobileNetV2

3.6 Model Fine-Tuning

Optimizer: The Adam optimizer was used, with an initial learning rate of 0.0001.

Learning Rate Scheduler: A ReduceLROnPlateau callback was applied to dynamically reduce the learning rate if the validation loss did not improve over a set number of epochs.

Early Stopping: An early stopping mechanism that monitored validation accuracy was used to prevent overfitting and reduce unnecessary computing, with patience set to 5 epochs.

Loss Function: Since the task is multi-class, the categorical cross-entropy loss function was used.

This training was done for up to 15 epochs with a batch size of 16.

3.7 Model Evaluation

To perform a couple of evaluations on the models, we will use some standard metrics of classification:

Accuracy: The overall proportion of correctly classified images.

Precision: It indicates the proportion of predicted positives that are correct, providing a ratio of true positives to the sum of true positives and false positives.

Recall: This is the ratio between the true positives and the sum of true positives and false negatives, as well. It measures the model's ability to find all relevant instances.

F1 Score: The harmonic mean of precision and recall, which gives an overall measure of performance of the model, especially when one has an imbalanced class condition.

$$Precision = \frac{True\ Positives}{TruePositives+FalsePositives} \tag{1}$$

$$Recall = \frac{True\ Positives}{TruePositives+False\ Negetives} \tag{2}$$

$$F1 = 2 * \frac{Precesion\ and\ Recall}{TruePositives+FalsePositives} \tag{3}$$

4 Result Evaluation

In the fruit ripeness classification experiments, four pre-trained convolutional neural network (CNN) models—MobileNetV2, InceptionV3, DenseNet121, and ResNet50—were evaluated on individual datasets of three fruits: apple, banana, and orange. Each dataset included three ripeness classes, and evaluation metrics included precision, recall, F1 score, and accuracy. The performances of the used algorithms for each fruit are given below in Table 3.

Table 3. Precision, recall, F1 Score and Accuracy of each model for each fruit

		For Apple		
Model	Precision	Recall	F1 Score	Accuracy
InceptionV3	1.00	1.00	1.00	1.00
DenseNet121	1.00	1.00	1.00	1.00
ResNet50	0.81	0.80	0.80	0.81
MobileNetV2	1.00	1.00	1.00	1.00
		For Banana		
Model	Precision	Recall	F1 Score	Accuracy
InceptionV3	1.00	1.00	1.00	1.00
DenseNet121	1.00	1.00	1.00	1.00
ResNet50	0.91	0.91	0.91	0.91

MobileNetV2	1.00	1.00	1.00	1.00
		For Orange		
Model	Precision	Recall	F1 Score	Accuracy
InceptionV3	1.00	1.00	1.00	1.00
DenseNet121	1.00	1.00	1.00	1.00
ResNet50	0.87	0.87	0.87	0.86
MobileNetV2	1.00	1.00	1.00	1.00

For the apple dataset, MobileNetV2, InceptionV3, and DenseNet121 performed perfectly across the board, each getting a score of 1.00 for precision, recall, F1 score, and accuracy. ResNet50, however, fell short significantly, achieving only an accuracy of 0.81 and performing even worse in the other metrics. Perhaps, this suggests that although ResNet50 is indeed a powerful model, its fine-grained classification abilities are challenged on certain datasets, such as apples, without some form of tuning or regularization.

The results for banana classification followed a pattern similar to that for apple classification. InceptionV3, DenseNet121, and MobileNetV2 achieved 100% accuracy again, indicating robust performance in detecting ripening stages for bananas. In comparison to the above study on apples, the accuracy of 0.91 across all metrics here indicated a moderate generalization ability of ResNet50 on this dataset, while still falling behind other models.

The orange dataset followed exactly the same trend, where InceptionV3, DenseNet121, and MobileNetV2 were still classified with perfection, each scoring 1.00 in all the metrics. Once again, ResNet50's performance was comparatively underwhelming, with an accuracy of 0.86, as well as precision and recall scores of 0.87 and an F1 score of 0.87. Since all three fruits receive a similar low score, we can infer that either overfitting or insufficient extraction of the particular features related to ripeness characteristics, which are lacking in the ResNet50 architecture, presents a challenge unless further fine-tuning is employed.

Overall, MobileNetV2 proved to be the best option, as it, along with others, consistently performed well across all three fruit datasets, making it the ideal choice for real-time, lightweight, and accurate fruit ripeness detection systems. InceptionV3 and DenseNet121 also performed exceptionally well; however, they are more computationally complex. On the contrary, ResNet50 has comparatively low classification accuracy and may require fine-tuning or architectural alterations to match the performance of the other models. The performances of the used models on combined dataset are given below in Table 4.

Table 4. Precision, recall, F1 Score and Accuracy of the models on combined dataset

Model	Precision	Recall	F1 Score	Accuracy
InceptionV3	0.95	0.95	0.95	0.95
MobileNetV2	0.97	0.97	0.97	0.97
DenseNet121	0.96	0.96	0.96	0.97
ResNet50	0.63	0.62	0.61	0.63

The combined dataset with all three fruits exhibited MobileNetV2 as the best model out of all the models tested, scoring 0.97 in precision, recall, F1 score, and accuracy. Its consistency in performance at such a high level indicates excellent generalization across different fruits and stages of ripeness.

DenseNet121 also seems strong, punching out an ever-so-slightly lesser 0.96 for precision, recall, and F1 score and 0.97 for the accuracy score. It is likely that its densely connected architecture helped preserve important features that were crucial for classifying the ripeness of different fruits.

InceptionV3 lagged behind with a score of 0.95 across the board; although quite reliable, it may not have captured ripeness-discriminatory features as well as MobileNetV2 and DenseNet121 in the combined setup.

Whereas the accuracy of ResNet50 lagged at 0.63, with low scores across all metrics, indicating a failure to cope with the complexity associated with the mixed dataset, likely due to poor discrimination among features and a lack of fine-tuning. The training and validation accuracy and loss curve of fine-tuned MobileNetV2 on combined dataset is given below in Fig. 2.

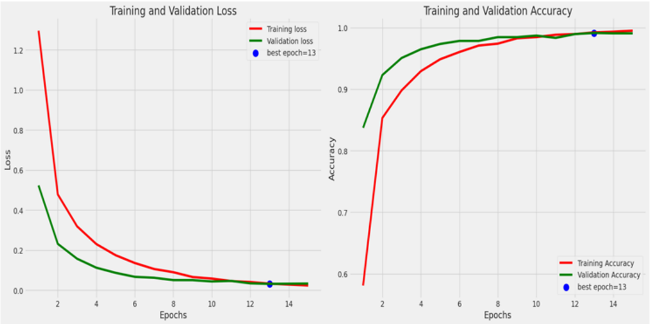


Fig. 2. Training and Validation Accuracy and Loss Curve of Fine Tuned MobileNetV2 on combined Dataset

Indeed, this smooth performance improvement of the fine-tuned MobileNetV2 model continued across 15 epochs. During training, the model achieved an accuracy of 37.13%, and the validation accuracy was 83.72%. This indicates that the top layers began to get meaningful features early in learning, even when the base layers were frozen.

As training progressed, accuracy learning increased quickly. Epoch 5 achieves a training accuracy of 94.73% and a validation accuracy of 97.35%. This is a clear indication of adaptation on the part of the model, resulting from the selective fine-tuning of 75 layers, global average pooling, and a dropout layer for regularization.

In the latter epochs, the model also continued to learn and improve its accuracy, attaining 99.52% training accuracy and 99.08% validation accuracy by the last epoch. This had a very low validation loss of 0.0333. The low constant learning rate and early stopping have resulted in a stable convergence without overfitting.

All these factors together confirm that fine-tuning MobileNetV2 with a well-designed training strategy has resulted in good generalization, yielding excellent performance on the combined fruit ripeness dataset.

Table 5. Classification report of Fine-tuned MobileNetV2 on combined dataset

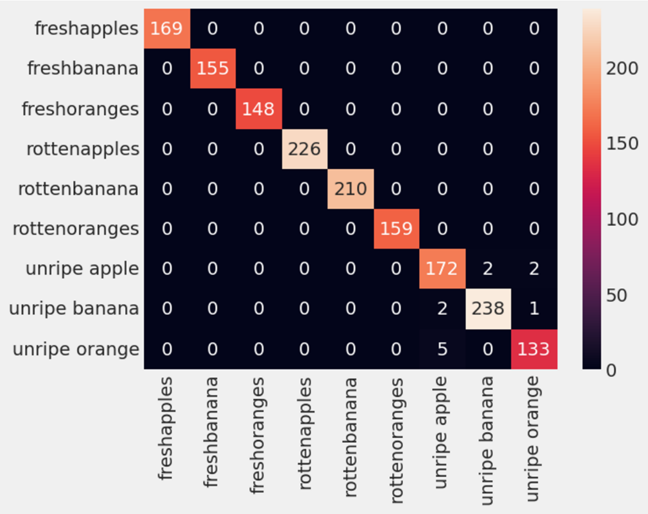
	Precision	Recall	F1 Score
freshapples	1.00	1.00	1.00
freshbanana	1.00	1.00	1.00
freshoranges	1.00	1.00	1.00
rottenapples	1.00	1.00	1.00
rottenbanaba	1.00	1.00	1.00
rotteoranges	1.00	1.00	1.00
unripe apple	0.96	0.98	0.97
unripe banana	0.99	0.99	0.99
unripe orange	0.98	0.96	0.97
accuracy			0.99
macro avg	0.99	0.99	0.99
weighted avg	0.99	0.99	0.99

The fine-tuning of the MobileNetV2 network demonstrated robust performance, achieving an overall accuracy of 0.99 on a combined fruit dataset comprising three ripeness classes for apples, bananas, and oranges. This indicates that the model correctly classified almost all of its test samples.

At the class level, the model attained perfect performance (1.00 precision, recall, and F1) in six categories out of nine: fresh apples, fresh bananas, fresh oranges, rotten apples, rotten bananas, and rotten oranges. This result indicates superior model performance in distinguishing between very well-defined classes, where their visual features are more distinct.

For the immature categories, the model would still have performed quite well. For unripe apples, there was a precision measurement of 0.96, a recall of 0.98, and an F1 score of 0.97, which means somewhat obvious difficulty in determining all real unripe apple samples. Unripe bananas and unripe oranges both achieved almost perfect scores, with 0.99/0.99/0.99 and 0.98/0.96/0.97, respectively. These minor dips may probably result from the closer similarity of colors and textures of unripe and fresh fruits, especially under varying lighting conditions.

Both macro-average and weighted-average metrics are 0.99 for precision, recall, and F1 score. These, again, demonstrate the balanced behavior of the model across all classes, without a preference for the majority classes. These results demonstrate that fine-tuned MobileNetV2, combined with layer freezing, dropout regularization, and learning rate scheduling, is effective in conducting multi-class fruit ripeness classification across various types of fruit. The confusion matrix of the fine-tuned MobileNetV2 is shown in Figure 3.



	freshapples	freshbanana	freshoranges	rottenapples	rottenbanana	rottenoranges	unripe apple	unripe banana	unripe orange
freshapples	169	0	0	0	0	0	0	0	0
freshbanana	0	155	0	0	0	0	0	0	0
freshoranges	0	0	148	0	0	0	0	0	0
rottenapples	0	0	0	226	0	0	0	0	0
rottenbanana	0	0	0	0	210	0	0	0	0
rottenoranges	0	0	0	0	0	159	0	0	0
unripe apple	0	0	0	0	0	0	172	2	2
unripe banana	0	0	0	0	0	0	2	238	1
unripe orange	0	0	0	0	0	0	5	0	133

Fig. 3. Confusion Matrix of the fine-tuned MobileNetV2

5 Conclusion and future Work

This research presents an effective deep learning-based approach for classifying fruit ripeness across three types of fruits—apple, banana, and orange—each categorized into fresh, unripe, and rotten classes. Several pre-trained CNN architectures from the category included MobileNetV2, InceptionV3, ResNet50, and DenseNet121. Every fruit model achieved 100% accuracy for most models, confirming that they had learned specific fruit features. Of these, MobileNetV2 was not only highly accurate in classification at an individual level but also generalized well to the combined dataset after fine-tuning, achieving 99% accuracy. By inserting dropout layers, global average pooling, carefully unfreezing layers, employing early stopping, and adjusting the learning rate, this approach created a solid and generalizable model. The evidence thus shows that lightweight CNNs, such as MobileNetV2, can be efficient and accurate for assessing fruit quality in real-world applications.

The model achieves relatively high-performance accuracy; hence, future improvements can be made in many ways. Increasing the number of fruit types and ripening stages in the database will enhance the model's generalizability. Multi-spectral or hyperspectral imaging would additionally uncover biochemical changes that RGB images miss. One of the very promising possibilities is carrying the trained model directly into real-time systems, such as mobile applications or embedded devices, which would benefit from on-device ripeness detection. It may also be discovered through self-supervised learning or semi-supervised techniques, reducing the reliance on large datasets. This would make the model more scalable and accessible to agricultural use

References

1. M. R. Bhosale and M. S. Patil, "Fruit recognition using color, shape and texture feature," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 4, Art. no. 6, 2015.Author, F., Author, S.: Title of a proceedings paper. In: Editor, F., Editor, S. (eds.) CONFERENCE 2016, LNCS, vol. 9999, pp. 1–13. Springer, Heidelberg (2016).
2. M. Rahnemoonfar and C. Sheppard, "Deep count: Fruit counting based on deep simulated learning," *Sensors*, vol. 17, Art. no. 4, 2017.Author, F.: Contribution title. In: 9th International Proceedings on Proceedings, pp. 1–2. Publisher, Location (2010).
3. A. Howard *et al.*, "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017.
4. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," 2016, pp. 770–778.
5. S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, Art. no. 10, 2010.
6. S. Sahu, A. K. Sahoo, and P. K. Meher, "Banana ripeness detection using machine learning approach," 2020, pp. 1–6.
7. A. Fuentes, S. Yoon, S. C. Kim, and D. S. Park, "A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition," *Sensors*, vol. 17, Art. no. 9, 2017.
8. M. Arakeri and S. Lakshmi, "Computer vision based fruit grading system for quality evaluation of mangoes," 2016, pp. 1104–1109.
9. A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Computers and Electronics in Agriculture*, vol. 147, pp. 70–90, 2018.
10. J. Amara, A. Bouaziz, and A. Algergawy, "A deep learning-based approach for banana leaf diseases classification," 2017, pp. 79–88.
11. S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Frontiers in Plant Science*, vol. 7, p. 1419, 2016.

12. I. Sa, Z. Ge, F. Dayoub, B. Upcroft, T. Perez, and C. McCool, “DeepFruits: A fruit detection system using deep neural networks,” *Sensors*, vol. 16, Art. no. 8, 2016.
13. V. Bhange and D. Hingoliwala, “Fruit quality assessment using image processing,” 2015, pp. 1–4.
14. P. Mureşan and M. Oltean, “Fruit recognition from images using deep learning,” *Acta Universitatis Sapientiae, Informatica*, vol. 10, Art. no. 1, 2018.
15. Kaggle, “Fruit ripeness: Unripe, ripe, and rotten,” 2025.
16. A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” *Communications of the ACM*, vol. 60, Art. no. 6, 2017.
17. A. Howard *et al.*, “MobileNets: Efficient convolutional neural networks for mobile vision applications,” 2017.
18. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the inception architecture for computer vision,” 2016, pp. 2818–2826.
19. K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” 2016, pp. 770–778.
20. G. Huang, Z. Liu, van, and K. Q. Weinberger, “Densely connected convolutional networks,” 2017, pp. 4700–4708.
21. S. J. Pan and Q. Yang, “A survey on transfer learning,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, Art. no. 10, 2010