engineering proceedings

MDPI

Proceeding Paper

Fine-Tuning MobileNet for Durian Variety Classification †

Nyuk Mee Voo [1,*], Tong Ming Lim [2] and Yee Mei Lim [3]

[1] Faculty of Computing and Information Technology, Tunku Abdul Rahman University of Management and Technology, Kuala Lumpur 53300, Malaysia
[2] Centre for Business Incubation and Entrepreneurial Ventures, Tunku Abdul Rahman University of Management and Technology, Kuala Lumpur 53300, Malaysia; limtm@tarc.edu.my
[3] OMIS Consulting Sdn Bhd, Kuala Lumpur 57100, Malaysia; yeemei@omis-consulting.com
* Correspondence: voonm@tarc.edu.my
† Presented at 2025 IEEE International Conference on Computation, Big-Data and Engineering (ICCBE), Penang, Malaysia, 27–29 June 2025.

Abstract

Durian, often referred to as the king of fruits, is widely consumed in Southeast Asia. However, the classification of its varieties is complicated by the lack of distinct visual differences between them. In this study, a fine-tuned MobileNet, a lightweight deep learning model, is applied for the classification of durian varieties. Transfer learning techniques are employed to adapt the MobileNet architecture using a custom dataset of durian images, enabling accurate differentiation between multiple varieties. First, the original MobileNet model is evaluated, which is found to yield low accuracy (8.22%) and a high loss (2.0553). A durian-specific classification layer is then added, and the model is trained for 100 epochs (2 min 21 s), achieving 76.28% training accuracy (0.5985 loss) and 73.20% validation accuracy (0.7606 loss). Further fine-tuning is performed, resulting in 100% training accuracy (4.9623×10^{-4} loss) and 93.69% validation accuracy (0.2281 loss) after 100 epochs (3 min 55 s). The findings demonstrate that the fine-tuned MobileNet model is capable of high classification accuracy while maintaining computational efficiency, making it suitable for real-time durian variety identification in agricultural and commercial settings.

Keywords: fine-tuning; mobilenet; deep learning; computer vision; durian classification

1. Introduction

Durian plays a vital role in Malaysia's agricultural sector and local economy [1–3]. From August to December 2024, the export of 413.61 tons of fresh durian valued at RM24.84 million to China was recorded [4,5]. In Malaysia, as of 2023, 158 durian cultivars have been officially registered under the national crop variety protection system [6]. Among these, premium varieties including Musang King, Black Thorn, and D24 are distinguished by their superior texture and flavour profiles, resulting in significantly higher market prices being commanded. However, the differentiation between durian varieties is complicated by subtle external morphological differences [7,8], making manual classification methods, which are currently employed, both labour-intensive and subjective. Consequently, the development of automated identification systems is necessitated.

Significant advancements in deep learning, particularly in convolutional neural networks (CNNs), have been made in recent years, with transformative applications being demonstrated across various agricultural domains [9–12]. While the potential of CNNs for durian classification has been explored in preliminary studies [13–16], further optimization for practical agricultural implementation remains to be achieved.

Check for updates

Academic Editors: Teen-Hang Meen, Wei Chien and Cheng-Fu Yang

Published: 27 March 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

CNN architectures such as MobileNet [17], ResNet [18], and DenseNet [19] have shown exceptional performance on benchmark datasets [20]. However, their application to specialized agricultural tasks is limited by the requirement for extensive labelled datasets and substantial computational resources [21,22]. To address these limitations, transfer learning [17–19] has been proposed as an effective solution, enabling models pre-trained on large datasets to be adapted for specialized tasks with limited data availability [23–25].

In this study, MobileNet [17,26] is selected due to its lightweight deep neural network architecture employing depthwise separable convolutions [27]. Comparable performance to larger models is achieved while requiring significantly fewer parameters, making MobileNet particularly suitable for deployment in resource-constrained devices, optimized for mobile and embedded vision applications to achieve a good balance between accuracy and computational efficiency. A fine-tuned MobileNet architecture optimized for durian variety classification is proposed to provide an accurate and efficient solution that can be practically implemented in agricultural production and supply chain environments.

2. Literature Review

The development of MobileNet architectures has been driven by the need for efficient deep learning models suitable for mobile vision applications. The original MobileNet architecture [17] was introduced as a lightweight solution employing depthwise separable convolutions, with two global hyperparameters proposed to balance latency and accuracy. This approach was demonstrated to achieve strong performance on ImageNet classification while significantly reducing computational complexity compared to traditional CNNs. Subsequent improvements led to MobileNetV2 [26], which incorporated inverted residual structures with linear bottlenecks, achieving state-of-the-art performance across multiple benchmarks while maintaining efficiency. The architecture was further shown to be effective when adapted for object detection (SSDLite) and semantic segmentation (Mobile DeepLabv3) tasks.

Further optimization was conducted through MobileNetV3 [28], which combined neural architecture search with novel design improvements. The development of both Large and Small variants addressed different resource constraints, with demonstrated improvements in accuracy (3.2–4.6%) and latency reduction (5–15%) compared to MobileNetV2. The introduction of Lite Reduced Atrous Spatial Pyramid Pooling (LR-ASPP) for segmentation tasks was shown to provide 30% faster inference while maintaining accuracy. These successive iterations have established MobileNet as a versatile architecture family capable of delivering efficient performance across various computer vision applications, making it particularly suitable for resource-constrained agricultural implementations.

Recent advances in lightweight deep learning architectures have demonstrated significant potential for agricultural applications, particularly in plant disease detection and classification. MobileNet variants have been extensively modified to enhance performance while maintaining computational efficiency, as evidenced by Real-Time Recognition Lite MobileNetV2 [29], which integrated attention mechanisms to achieve 99.92% accuracy on plant disease datasets while remaining deployable on edge devices. Similar architectural innovations have been explored, such as the fusion of MobileNet with GRU networks for spatiotemporal crop monitoring [30], achieving 93.0% accuracy in remote sensing applications, and the incorporation of MobileNet into ensemble systems like VotTomNet [31], which attained 99.2% diagnostic accuracy through voting-based transfer learning. These studies collectively highlight MobileNet's adaptability to diverse agricultural tasks, though challenges persist in optimizing feature extraction for fine-grained classification, as noted in genetic algorithm-aided feature selection experiments [32].

https://doi.org/10.3390/engproc2026128046

The effectiveness of hybrid and explainable approaches has been further validated in crop-specific applications. Explainable Spectral Extraction-Tomato Disease Classification Neural Network [33] augmented EfficientNetB0 with squeeze-and-excitation blocks and multiscale fusion to achieve 99.11% accuracy in tomato disease classification, outperforming standalone MobileNet (87.44%). Similarly, group learning frameworks [34] demonstrated that MobileNet's 96% baseline accuracy could be elevated to 100% through weighted majority voting, while real-time monitoring systems [35] confirmed its edge deployment capability with 100% accuracy for apple and pepper bell disease detection. Beyond classification, MobileNet's utility has been extended to IoT security in smart agriculture through transfer learning-optimized intrusion detection [36], achieving 99% accuracy in threat identification. These developments underscore MobileNet's dual strengths in task-specific accuracy and operational efficiency, though gaps remain in standardizing performance metrics across heterogeneous agricultural datasets.

The reviewed studies collectively demonstrate that MobileNet-based architectures have been successfully adapted for diverse agricultural applications, achieving high accuracy while maintaining computational efficiency suitable for edge deployment. MobileNet variants have consistently outperformed traditional CNN models in resource-constrained settings. However, the research gaps remain unaddressed, and the MobileNet adaptation to durian variety classification that involves subtler morphological distinctions has not been explored. This study aims to address the gaps by tailoring MobileNet for durian cultivar identification and optimizing the architecture for edge devices in agricultural supply chains.

3. Dataset

Images of different durian varieties were collected from farms, stalls, and distribution centers. A total of 13 varieties and 6487 images were gathered. The collected durian varieties are shown in Table 1. Each image was manually labeled with its corresponding durian variety, with the assistance of durian owners to ensure accurate classification. In this study, only seven durian varieties and 1825 images were selected to maintain an equal number of images per variety, ranging between 100 and 300. The selected durian dataset varieties are presented in Table 2. This helped mitigate class imbalance [37–39], which leads to bias toward the majority classes (D197 and D24) and result in poor performance on underrepresented classes.

Table 1. Collected durian varieties.

D197	**D160**	**D7**	**D144**	**D145**	**D2**	**D13**	**D18**
2895	477	6	7	7	319	7	8
D24	**D101**	**D175**	**D224**	**D200**			
1877	73	141	2113	457			

Table 2. Durian dataset.

D2	**D24**	**D160**	**D175**	**D197**	**D200**	**D224**
300	300	300	141	300	300	214

Data augmentation methods were used to increase the variety of the training dataset and to optimize the ability of the model to generalize. Such techniques include scaling and cropping. Images are resized in a manner that preserves the aspect ratio, thereby introducing variations in scale. Additionally, portions of the image were cropped to focus on the fruit.

Subsequently, the dataset was divided into training and validation datasets. The training dataset constitutes the largest subset and was employed for model training, representing 70% of the total dataset. The validation dataset was employed to calibrate the hyperparameters and assess the performance of the model during the training process, representing 30% of the total dataset. Before feeding the images into the model, data preprocessing was applied. Normalization entails the scaling of pixel values to a range of [0, 1], which converges the model at a more expedient rate. Images were resized to ensure that all images have the same dimensions, which is a prerequisite for processing by the neural networks. The dataset was adjusted using the mean and standard deviation to normalize across the train and validation datasets.

The dataset was organized in a manner that facilitates efficient access and usage. A directory structure was created to separate the training and validation sets, with subdirectories for each durian class. Various tools and platforms were employed to streamline the process of preparing the dataset. Data augmentation libraries, including TensorFlow's image preprocessing utilities and torchvision's image transformations utilities, have been utilized. Cloud storage and a Docker container were employed for the purpose of managing datasets and training the classification models. Sample durians are shown in Figure 1.

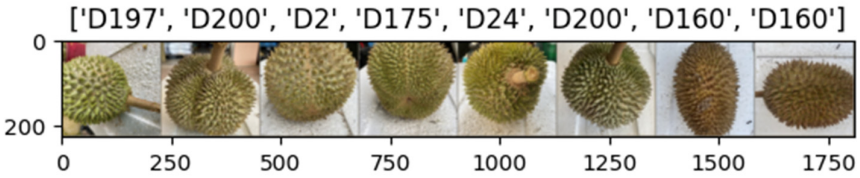


Figure 1. Sample images.

4. Model Architecture

We employed MobileNetv2 as a base model, a lightweight CNN optimized for mobile vision applications. Its architecture includes the following.

1. Depthwise separable convolutions: Standard convolutions are replaced by Depthwise Separable Convolutions which consist of Depthwise Convolution and Pointwise Convolution.
2. Inverted residual blocks: A three-phase dimensional transformation is employed. The input is compressed through a narrow bottleneck layer, subsequently expanded to a higher-dimensional representation, and finally projected back to a compressed output space.
3. Linear activation function: This function is employed at the termination of each residual block instead of the ReLU nonlinearity. This design choice is implemented to prevent information loss in low-dimensional representations, where nonlinear activations have been shown to cause feature collapse.

The architecture processes $224 \times 224 \times 3$ inputs through 19 bottleneck layers (Table 3), containing 3.5 M parameters.

The MobileNetV2 model was pre-trained on the ImageNet dataset, which comprises 1.4 million images spanning 1000 object categories. Although this dataset includes various tropical fruits such as jackfruit (which is correctly classified by the model), durian is notably absent from its categories. When durian images are processed by the base model, they are consistently misclassified as cardoon, echidna, honeycomb, pufferfish, or sea urchin. These incorrect predictions occur because certain textural or morphological similarities are detected, despite fundamental differences between these objects. As illustrated in Figure 2, these misclassifications clearly demonstrate the necessity for domain-specific adaptation of the model.

Table 3. The MobileNetV2 Architecture.

Layer Type	Output Size	Filter Size/Stride
Input Layer	224 × 224 × 3	-
Conv2D + BN + ReLU6	112 × 112 ×32	3 × 3/2
Bottleneck Block (×1)	112 × 112 ×16	3 × 3/1
Bottleneck Block (×2)	56 × 56 × 24	3 × 3/2
Bottleneck Block (×3)	28 × 28 × 32	3 × 3/2
Bottleneck Block (×4)	14 × 14 × 64	3 × 3/2
Bottleneck Block (×3)	14 × 14 × 96	3 × 3/1
Bottleneck Block (×3)	7 × 7 × 160	3 × 3/2
Bottleneck Block (×1)	7 × 7 × 320	3 × 3/1
Conv2D + BN + ReLU6	7 × 7 × 1280	3 × 3/1
Global Average Pooling	1 × 1 × 1280	-
Fully Connected	1 × 1 × num classes	-
Total number of parameters: 3,538,984		

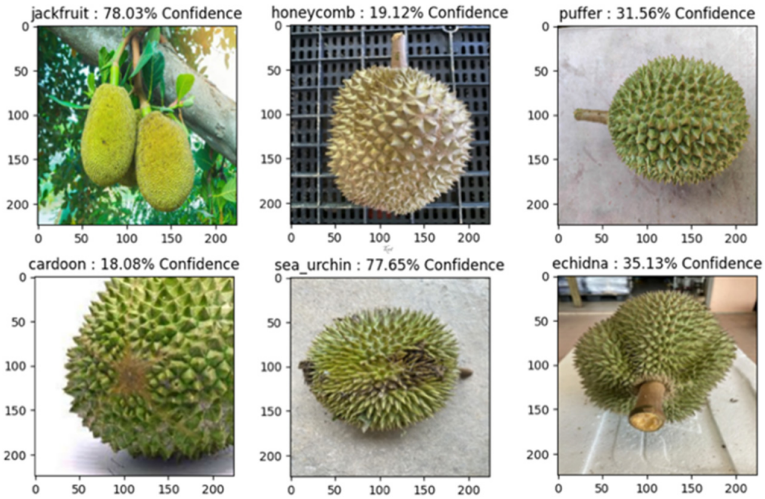


Figure 2. Misclassified samples.

The MobileNetV2 architecture was adopted as the base model in this study. The final classification layer was removed following standard transfer learning practices. Instead, the model was modified to utilize the last layer preceding the flatten operation, referred to as the bottleneck layer as its output. This bottleneck layer was selected because its feature representations maintain greater generalizability compared to the more specialized final classification layer. The modified base model architecture is presented in Table 4, showing the retention of all convolutional and bottleneck layers, the removal of the original classification head, and the preservation of the 7 × 7 × 1280 (dimensional bottleneck features).

Table 4. Base model architecture.

Layer Type	
Input layer	
Conv2D + BN + ReLU6	
Bottleneck block (×17)	
Conv2D + BN + ReLU6	Removed
Global average pooling	Removed
Fully connected	Removed
Total number of parameters: 2,257,984	

5. Feature Extraction

The MobileNetV2 model was pre-trained on the ImageNet dataset, a large-scale image classification task. This pre-training enables the model to serve as a generalized feature extractor, eliminating the need for training from scratch on a new dataset. The learned feature maps from this base convolutional network contain broadly applicable visual features that are commonly utilized for image classification tasks.

For adaptation to durian classification, transfer learning was employed using the pre-trained model. The layer preceding the flatten operation was selected for feature extraction due to its retention of more generalized features compared to the final classification layer. A MobileNetV2 instance was initialized with ImageNet-trained weights, with the top classification layers removed to create an optimal feature extraction architecture. This configuration transforms each input image ($224 \times 224 \times 3$ pixels) into a $7 \times 7 \times 1280$ feature block.

The convolutional base was frozen during feature extraction to preserve the pre-trained weights. A durian-specific classifier layer was appended to the frozen base, enabling training of only the top-level classification components. This freezing mechanism prevents weight updates in the base layers during the training process. Feature processing was implemented as follows.

1. Global average pooling: The $7 \times 7 \times 1280$ feature blocks were reduced to 1280-element vectors by averaging across spatial dimensions.
2. Dense classification layer: The pooled features were processed through a fully connected layer implementing the operation: output = activation(dot(input, kernel) + bias. Rectified linear unit (ReLU) activation was applied, where f(x) = max(0, x).

The complete model architecture was constructed using the Keras Functional Application Programming Interface, integrating data augmentation, input rescaling, the frozen base model, and the new classifier layers. The architectural details of this feature extraction model are presented in Table 5.

Table 5. Feature extraction of durian model architecture.

Layer Type	Output Size	Parameter
Input layer	$224 \times 224 \times 3$	0
mobilenetv2 functional	$7 \times 7 \times 1280$	2,257,984
Global average pooling	$1 \times 1 \times 1280$	-
Fully connected	$1 \times 1 \times$ Num Classes	8967
Total number of parameters: 2,266,951		
Total number of trainable parameters: 8967		
Total number of non-trainable parameters: 2,257,984		

The MobileNet base model's 2,527,984 parameters were frozen during the feature extraction phase, while only 8967 trainable parameters in the dense classification layer were optimized. These parameters were distributed between weight matrices and bias vectors within the layer. Before training, the model was compiled with the following configurations:

1. Optimizer: The Adam optimizer was employed, which implements stochastic gradient descent through adaptive estimation of first- and second-order moments. The default learning rate of 0.001 was maintained.
2. Loss function: Sparse categorical cross-entropy was utilized to compute the loss between predictions and ground truth labels.

3. Metrics: Model performance was evaluated using sparse categorical accuracy, which calculates the frequency of correct predictions matching the true class indices.

The initial evaluation on the validation dataset yielded a loss of 2.04 and an accuracy of 9%. The model was then trained for 100 epochs, during which the following improvements were observed as follows.

- Training accuracy increased from 12.87 to 76.28%
- Training loss decreased from 1.9984 to 0.5985
- Validation accuracy improved from 18.47 to 73.20%
- Validation loss reduced from 1.9409 to 0.7606

The complete training process was completed in 2 min and 21 s. Learning curves depicting the accuracy and loss progression during training are presented in Figure 3.

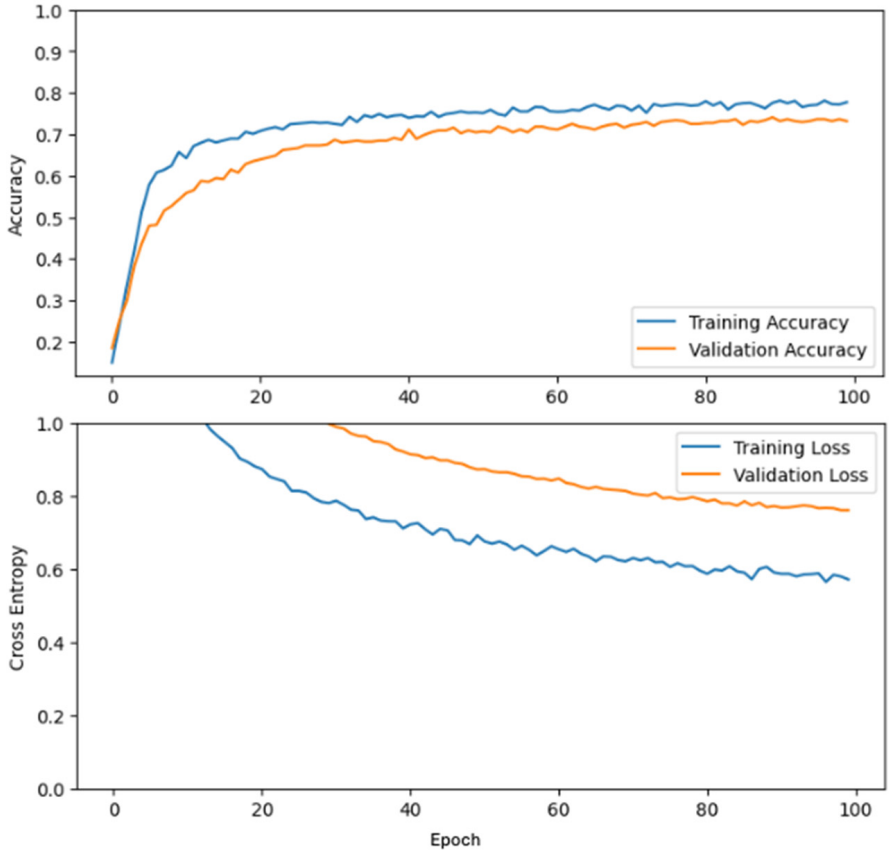


Figure 3. Total accuracy and loss (feature extraction).

6. Fine-Tuning

In the initial feature extraction phase, only the newly added classification layers were trained while the weights of the pre-trained MobileNetV2 network remained frozen. To enhance model performance further, fine-tuning was performed on the upper layers of the base model alongside the durian classifier.

The specialization of features in deep neural networks follows a hierarchical pattern, where lower layers typically learn general visual features (e.g., edges and textures) that are applicable across diverse image types, while higher layers develop more dataset-specific representations. This characteristic was leveraged through selective fine-tuning of only the top 54 layers (layers 100–154) of the 154-layer base model, preserving the generic feature extractors in lower layers while adapting the specialized representations for durian classification. The fine-tuning procedure was implemented as follows.

1. The base model was unfrozen, starting from layer 100.
2. Lower layers (1–99) were kept frozen to maintain their generic feature extraction capabilities.
3. The learning rate was reduced by an order of magnitude (from 0.001 to 0.0001) to prevent rapid overfitting.
4. The model was recompiled with the new trainable layer configuration.
5. Training was resumed using RMSprop optimization.

This approach allowed the model to preserve valuable generic features in early layers, adapt specialized features in higher layers to durian-specific characteristics, and maintain training stability through reduced learning rates. The complete architecture of the fine-tuned model, including the modified layer configurations and parameter counts, is presented in Table 6. The fine-tuning process resulted in significant performance improvements, with validation accuracy increasing from 76.28% to 93.69% while maintaining computational efficiency.

Table 6. Fine-tuned durian model architecture.

Layer Type	Output Size	Parameter
Input layer	$224 \times 224 \times 3$	0
Mobilenetv2 functional	$7 \times 7 \times 1280$	2,257,984
Global average pooling	$1 \times 1 \times 1280$	-
Fully connected	$1 \times 1 \times$ Num Classes	8967
Total number of parameters: 2,266,951 Total number of trainable parameters: 1,870,407 Total number of non-trainable parameters: 396,544		

During the fine-tuning phase, 396,544 parameters in the MobileNet base model were frozen, while 1,870,407 parameters in the dense classification layer were made trainable. The model was compiled before training with the following key components.

1. Optimization: The root mean square propagation (RMSprop) algorithm was employed as the optimizer.
2. Loss function: Sparse categorical cross-entropy was utilized to compute the loss between predictions and ground truth labels.
3. Metrics: Model performance was evaluated using the sparse categorical accuracy metric, which calculated the frequency of correct predictions matching the true labels.

Training progress was monitored at each epoch. The learning curves depicting accuracy and loss metrics are presented in Figure 4. Significant improvements were observed during fine-tuning as follows. The complete fine-tuning process was completed in 3 min and 55 s.

- Training accuracy increased from 41.94 to 100%
- Training loss decreased from 1.8633 to 4.9623×10^{-4}
- Validation accuracy improved from 71.17% to 93.69%
- Validation loss reduced from 0.8296 to 0.2281

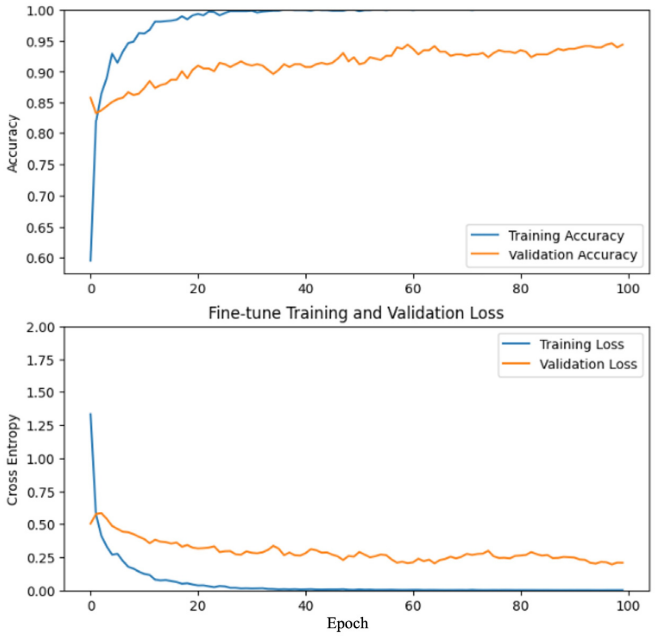


Figure 4. Total accuracy and loss (fine-tuning).

7. Evaluation and Prediction

A dedicated test set was created by allocating 20% of the validation dataset. The model's performance was rigorously evaluated on this unseen test data, achieving an accuracy of 94.64% with a corresponding loss of 0.1478. These metrics demonstrate the model's strong generalization capability to novel input samples. The trained model was subsequently employed to classify seven distinct durian varieties. Representative predictions alongside their ground truth labels are visually presented in Figure 5, illustrating the model's classification performance across different cultivar varieties.

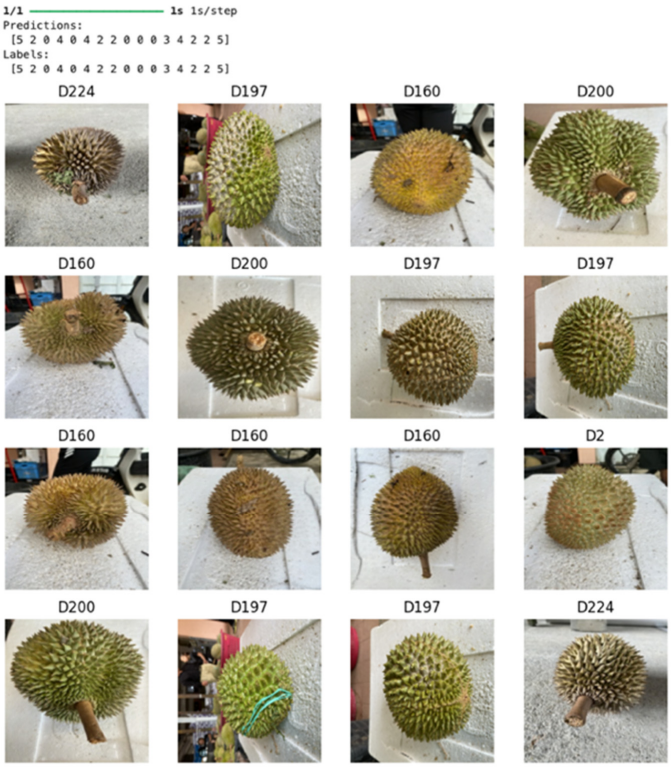


Figure 5. Predictions and labels of durians.

8. Discussion and Conclusions

The developed system presented the effectiveness of MobileNetV2 for durian variety classification, with a validated accuracy of 95.17% being achieved through optimized transfer learning techniques. The proposed framework combines (1) feature extraction from a pre-trained backbone, (2) selective layer freezing to preserve generic visual patterns, and (3) progressive fine-tuning of specialized layers. These approaches maintain computational efficiency while adapting the model to domain-specific requirements, with the final implementation requiring only 14 megabytes of storage space. The following is necessary for further enhancement of the developed system.

1. Architectural suitability: MobileNetV2's depth-wise separable convolutions enabled efficient learning of discriminative features (e.g., spine patterns, fruit shape) with only 1.87 M trainable parameters during fine-tuning.
2. Progressive training strategy: The two-phase transfer learning approach, initial feature extraction followed by selective fine-tuning, prevented catastrophic forgetting while adapting generic ImageNet features to durian-specific characteristics.
3. Dataset expansion: More comprehensive data collection encompassing additional varieties, internal structures (shells, flesh, seeds), and growth stages must be undertaken to enable grade classification.
4. Supply chain integration: The developed model could be incorporated into automated sorting systems to enhance quality control throughout the distribution network.
5. Pathological analysis: The framework may be extended to detect disease conditions by incorporating temporal data on durian development.

The methodology presented in this study provides a replicable blueprint for applying lightweight convolutional neural networks to agricultural classification tasks. Particularly noteworthy is the model's balanced performance in both accuracy and efficiency, making it suitable for deployment in resource-constrained environments.

Author Contributions: Writing, N.M.V.; Supervision, T.M.L. and Y.M.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data from this study cannot be made public at this time due to ongoing research. After the study is completed, the data will be made available to a public repository.

Acknowledgments: The authors thank Go Green Fruits Trading, Sky View Sdn Bhd, Durian Stall 29, Great Fruit Station, and durian farm owners Tan Hua Hong, Liew Soon Cai, Tina Chong, and Chia Teck Chai for their generosity and assistance during data collection.

Conflicts of Interest: Author Yee Mei Lim was employed by the company OMIS Consulting Sdn Bhd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CNN	Convolutional Neural Network
ReLU	Rectified Linear Unit
RMSprop	Root Mean Square Propagation

References

1. *Fruit Crop Statistic*; Department of Agriculture Malaysia: Putrajaya, Malaysia, 2023. Available online: https://www.doa.gov.my/doa/resources/aktiviti_sumber/arkib/statistik_tanaman/2024/statistik_tanaman_buah_2023.pdf (accessed on 28 February 2025).
2. Siddharta, A. Malaysia: Durian Production Volume 2023—Statista. Available online: https://www.statista.com/statistics/1000876/malaysia-durian-production/ (accessed on 28 February 2025).
3. Durian Fruit Market Growth Trends 2025—Industry Analysis Report. Available online: https://www.millioninsights.com/industry-reports/global-durian-fruit-market/ (accessed on 28 February 2025).
4. Bernama. Malaysian Fresh Durians Command Premium Status in China Market. *Business Times*, 25 February 2025. Available online: https://www.nst.com.my/business/economy/2025/02/1180108/malaysian-fresh-durians-command-premium-status-china-market/ (accessed on 25 February 2025).
5. Ho, J. Malaysia's Durian Exports to China Reach rm24.84mil. *The Star*, 25 February 2025. Available online: https://www.thestar.com.my/news/nation/2025/02/25/malaysia039s-durian-exports-to-china-reach-rm2484mil/ (accessed on 25 February 2025).
6. Department of Agriculture Malaysia. Register of Common Crop Varieties–Mypvp. Available online: http://mypvp.doa.gov.my/nvl/plant-database/ (accessed on 28 February 2025).
7. Lim, B.; Valariano, M.; Said, N.H.; Yaacob, Z.; Yusuf, A.; Sodali, S.; Mat Tarmizi, A.; Jamalullail Danial, M.K. *Panduan Pengesahan dan Pencirian Ketulenan Anak Pokok Durian*; Department of Agriculture Malaysia: Putrajaya, Malaysia, 2009. Available online: https://www.doa.gov.my/doa/resources/perkhidmatan/skim_pensijilan/SPBT/panduan_pengesahan_anak_pokok_durian.pdf (accessed on 28 February 2025).
8. National Guidelines for the Conduct of Tests for Distinctness, Uniformity and Stability. Available online: https://plantauthority.gov.in/sites/default/files/teakguideline.pdf (accessed on 28 February 2025).
9. Abubeker, K.M.; Abhijit; Akhil, S.; Kumar, V.K.A.; Jose, B.K. Computer vision-assisted real-time bird eye chili classification using yolo v5 framework. *J. Artif. Intell. Technol.* **2023**, *4*, 265–271. [CrossRef]
10. Yadav, P.K.; Burks, T.; Frederick, Q.; Qin, J.; Kim, M.; Ritenour, M.A. Citrus disease detection using convolution neural network generated features and softmax classifier on hyperspectral image data. *Front. Plant Sci.* **2022**, *13*, 1043712–1043712. [CrossRef] [PubMed]
11. Olugboja, A.; Wang, Z.; Sun, Y. Parallel convolutional neural networks for object detection. *J. Adv. Inf. Technol.* **2021**, *12*, 279–286. [CrossRef]
12. Norval, M.J.; Wang, Z.; Sun, Y. Evaluation of image processing technologies for pulmonary tuberculosis detection based on deep learning convolutional neural networks. *J. Adv. Inf. Technol.* **2021**, *12*, 253–259. [CrossRef]
13. Lim, M.; Chuah, J. Durian types recognition using deep learning techniques. In Proceedings of the 2018 9th IEEE Control and System Graduate Research Colloquium (ICSGRC 2018), Shah Alam, Malaysia, 3–4 August 2018; pp. 183–187.
14. Uy, J.N.; Villaverde, J.F. A durian variety identifier using canny edge and CNN. In Proceedings of the 2021 IEEE 7th International Conference on Control Science and Systems Engineering (ICCSSE 2021), Qingdao, China, 30 July–1 August 2021; pp. 293–297.
15. Zarifie Hashim, N.M.; Azfar Khairil Bahri, M.H.; Abd Ghani, S.M.; Sulistiyo, M.D.; Mohd Kassim, K.A.; Hanin Zahri, N.A. An introduction to a smart durian musang king and durian kampung classification. In Proceedings of the 2022 2nd International Conference on Intelligent Technologies (CONIT 2022), Karnataka, India, 24–26 July 2022; pp. 1–6.
16. Diana, D.; Kurniawan, T.B.; Dewi, D.A.; Alqudah, M.K.; Alqudah, M.K.; Zakari, M.Z.; Fuad, E.F.B.E. Convolutional neural network based deep learning model for accurate classification of durian types. *J. Appl. Data Sci.* **2025**, *6*, 101–114. [CrossRef]
17. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv* **2017**, arXiv:1704.04861. [CrossRef]
18. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2016), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
19. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017), Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.
20. Deng, J.; Dong, W.; Socher, R.; Li, L.; Li, K.; Li, F. Imagenet: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 248–255.
21. Yang, S.; Xiao, W.; Zhang, M.; Guo, S.; Zhao, J.; Shen, F. Image data augmentation for deep learning: A survey. *arXiv* **2022**, arXiv:2204.08610. [CrossRef]
22. Shorten, C.; Khoshgoftaar, T.M. A survey on image data augmentation for deep learning. *J. Big Data* **2019**, *6*, 60. [CrossRef]
23. Yosinski, J.; Clune, J.; Bengio, Y.; Lipson, H. How transferable are features in deep neural networks? In Proceedings of the 28th International Conference on Neural Information Processing Systems, Montreal, QC, Canada, 8–11 December 2014.
24. Pan, S.J.; Yang, Q. A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 1345–1359. [CrossRef]
25. Tan, C.; Sun, F.; Kong, T.; Zhang, W.; Yang, C.; Liu, C. A survey on deep transfer learning. *arXiv* **2018**, arXiv:1808.01974. [CrossRef]

26. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.-C. Mobilenetv2: Inverted residuals and linear bottlenecks. *arXiv* **2019**, arXiv:1801.04381. [CrossRef]
27. Chollet, F. Xception: Deep learning with depthwise separable convolutions. *arXiv* **2019**, arXiv:1610.02357. [CrossRef]
28. Howard, A.; Sandler, M.; Chu, G.; Chen, L.-C.; Chen, B.; Tan, M.; Wang, W.; Zhu, Y.; Pang, R.; Vasudevan, V.; et al. Searching for mobilenetv3. *arXiv* **2019**, arXiv:1905.02244. [CrossRef]
29. Duhan, S.; Gulia, P.; Gill, N.S.; Narwal, E. RTR lite mobilenetv2: A lightweight and efficient model for plant disease detection and classification. *Curr. Plant Biol.* **2025**, *42*, 10045. [CrossRef]
30. Kumar, U.S.; Kapali, B.S.C.; Nageswaran, A.; Umapathy, K.; Jangir, P.; Swetha, K.; Begum, M.A. Fusion of mobilenet and gru: Enhancing remote sensing applications for sustainable agriculture and food security. *Remote Sens. Earth Syst. Sci.* **2025**, *8*, 118–131. [CrossRef]
31. Joshi-Bag, S.; Patil, W.V.; Chavate, S. Vottomnet: Voting-based tomato disease diagnosis with transfer learning. *IAES Int. J. Robot. Autom.* **2025**, *14*, 38–46. [CrossRef]
32. Sharma, R.; Singh, A.; Kumar, P.; Singh, M. Genetic algorithm–aided deep feature selection for improved rice disease classification. *Oper. Res. Forum* **2025**, *6*, 1. [CrossRef]
33. Assaduzzaman, M.; Bishshash, P.; Nirob, M.A.S.; Marouf, A.A.; Rokne, J.G.; Alhajj, R. Xse-tomatonet: An explainable ai based tomato leaf disease classification method using efficientnetb0 with squeeze-and-excitation blocks and multi-scale feature fusion. *MethodsX* **2025**, *14*, 103159. [CrossRef]
34. Javidan, S.M.; Ampatzidis, Y.; Banakar, A.; Asefpour Vakilian, K.; Rahnama, K. An intelligent group learning framework for detecting common tomato diseases using simple and weighted majority voting with deep learning models. *AgriEngineering* **2025**, *7*, 2. [CrossRef]
35. Rahman, K.N.; Banik, S.C.; Islam, R.; Fahim, A.A. A real time monitoring system for accurate plant leaves disease detection using deep learning. *Crop Des.* **2025**, *4*, 1. [CrossRef]
36. Zhou, H.; Zou, H.; Zhou, P.; Shen, Y.; Li, D.; Li, W. CBCTL-IDS: A transfer learning-based intrusion detection system optimized with the black kite algorithm for IOT-enabled smart agriculture. *IEEE Access* **2025**, *13*, 46601–46615. [CrossRef]
37. Thölke, P.; Mantilla-Ramos, Y.-J.; Abdelhedi, H.; Maschke, C.; Dehgan, A.; Harel, Y.; Kemtur, A.; Mekki Berrada, L.; Sahraoui, M.; Young, T.; et al. Class imbalance should not throw you off balance: Choosing the right classifiers and performance metrics for brain decoding with imbalanced data. *NeuroImage* **2023**, *277*, 120253. [CrossRef] [PubMed]
38. Buda, M.; Maki, A.; Mazurowski, M.A. A systematic study of the class imbalance problem in convolutional neural networks. *Neural Netw.* **2018**, *106*, 249–259. [CrossRef] [PubMed]
39. Chen, Z.; Duan, J.; Kang, L.; Qiu, G. Class-imbalanced deep learning via a class-balanced ensemble. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, *33*, 5626–5640. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.