



Mango Ripeness Classification Using Deep Learning-Based Convolutional Neural Network Models

Intan Adha Maharani¹, Rahma Aliya^{2*},
Nidithia Putri Rahrahima³, Elfani Mardhatillah⁴

^{1,2}Department of Information System, Faculty of Science and Technology,
Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia

³Department of Syari'ah Islamiyah, Faculty of Dirost Islamiyah, Al-Azhar University, Egypt

⁴Department of Ushulluddin, Faculty of Dirost Islamiyah, Al-Azhar University, Egypt

E-Mail: ¹rahmaalya@gmail.com, ²intanadhamaharani@gmail.com,
³nindithiaputri@gmail.com, ⁴elfanimardhatillah@gmail.com

Received Jun 12th 2025; Revised Feb 01st 2026; Accepted Jul 04th 2026; Available Online Aug 01st 2026

Corresponding Author: Rahma Aliya

Copyright © 2026 by Authors, Published by Institute of Research and Publication Indonesia (IRPI)

Abstract

Manual assessment of mango ripeness is often subjective, time-consuming, and labor-intensive, making it inefficient for large-scale applications. This study investigates a deep learning-based approach for mango ripeness classification using Convolutional Neural Networks (CNN) and a publicly available image dataset obtained from Kaggle. Three CNN architectures, ResNet50V2, InceptionV3, and MobileNetV2, were evaluated using transfer learning and optimized with Adam and RMSprop to classify mangoes as unripe, ripe, or rotten. Data augmentation techniques were applied to improve model generalization and reduce overfitting. Model performance was evaluated using accuracy, precision, recall, and F1-score metrics. Experimental results show that all models achieved competitive performance, with ResNet50V2 obtaining the highest test accuracy of 81.8%, while InceptionV3 demonstrated more stable validation results. These findings indicate that CNN-based transfer learning is effective for mango ripeness classification and can serve as a baseline reference for further development of automated fruit classification systems.

Keywords: Convolutional Neural Network, Data Augmentation, Deep Learning, Mango Ripeness, ResNet50V2

1. INTRODUCTION

Mango (*Mangifera indica*) is one of the key horticultural commodities in tropical countries such as Indonesia, possessing high economic value in both local and international markets [1]. The quality and ripeness of mangoes are key determinants of consumer satisfaction and distribution chain efficiency [2]. However, manual processes for classifying ripeness and fruit condition face significant challenges, such as inconsistency between assessors, time constraints, and high labor costs [3].

Conventional methods for classifying ripeness levels or detecting rot in mangoes, such as visual inspection or manual pressure testing, are often inaccurate and highly dependent on human subjective perception [4]. Differences in lighting, background, and mango varieties also pose challenges in the visual classification process [5]. Furthermore, in industrial settings where thousands of fruits must be processed daily, these manual approaches become inefficient and prone to errors [6]. This problem highlights the need for automated classification systems based on image processing and artificial intelligence.

Advances in artificial intelligence, particularly deep learning with Convolutional Neural Network (CNN) architecture, have opened up significant opportunities for automatic fruit image classification [7]. CNNs have been proven capable of extracting complex visual features from images, such as color, texture, and shape, without the need for manual feature extraction [5]. Several previous studies have shown that CNNs can be used effectively for classifying mango varieties and ripeness levels with a fairly high degree of accuracy. In addition, transfer learning with pre-trained models such as VGG16 has been shown to improve classification performance compared to simple CNN models. A study by Abu Mettleq also shows that CNNs can be used to classify mango varieties with high accuracy [1]. Furthermore, the transfer learning approach using pre-trained architectures such as VGG16 has also been proven to improve fruit classification performance [8]. By combining data augmentation, dropout, and model refinement techniques, the classification performance of mango conditions can be significantly improved.



Research by Varga uses hyperspectral imaging and deep learning to measure fruit ripeness, achieving high accuracy in detecting chemical changes during ripening. They developed a mango ripeness classification system using machine learning techniques with a visual approach, obtaining satisfactory accuracy [9]. Additionally, the use of simple image processing methods to distinguish mango ripeness stages shows potential for practical applications. Based on this review, it can be concluded that there are still several limitations in previous studies, namely the use of a limited local dataset, the classification of mango conditions into three categories (unripe, ripe, and rotten), which has not been widely studied, and the lack of comprehensive model performance evaluation using various evaluation metrics. In addition, few studies have tested the performance of CNN models on images with uncontrolled background and lighting variations, which better reflect real conditions in the fruit processing industry. However, studies that specifically classify mango conditions into three categories (unripe, ripe, and rotten) using a CNN approach with comprehensive evaluation (accuracy, precision, recall, F1-score) are still limited, especially in a local context with uncontrolled background variations and lighting conditions [10].

This study aims to develop and evaluate a CNN-based classification model for mango conditions (unripe, ripe, rotten) using a local dataset. The main innovation of this study lies in the use of a combination of data augmentation techniques, dropout regularization, and a transfer learning approach to improve model generalization. Model performance will be evaluated using accuracy, precision, recall, and F1-score metrics and compared with the baseline approach. It is expected that the resulting model can be applied in automatic sorting systems in the mango processing industry and serve as a reference for the development of similar systems for other horticultural commodities [6].

2. MATERIAL AND METHOD

This study aims to develop an automatic mango ripeness classification system (categories: unripe, ripe, and rotten) using a deep learning approach, specifically a Convolutional Neural Network (CNN) architecture. Using a local image dataset, the model was built through comprehensive stages, including data acquisition, preprocessing, augmentation, architecture design, optimizer selection, and performance evaluation with confusion-matrix-based metrics. The flowchart of the research methodology is shown in Figure 1.

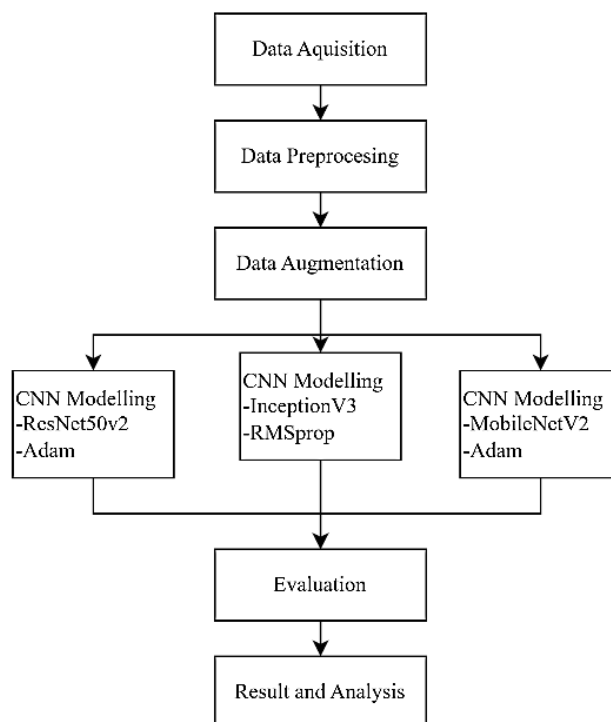


Figure 1. Flowchart of Research Methodology

2.1. Data Acquisition

Data acquisition is the process of gathering information that will later be processed and analyzed. Data collection can be conducted through various methods, such as direct interviews, questionnaire distribution, field observations, or by utilizing secondary data from existing sources. The choice of collection method should be aligned with the research objectives and hypotheses to be tested. Additionally, data can also be obtained from official sources, such as websites, government agencies, or health institutions like hospitals [11].

2.2. Data Preprocessing

Data preprocessing is a crucial step in deep learning modeling to improve data quality, model performance, and ease of implementation. This step involves increasing the number of images through augmentation techniques to address overfitting issues caused by a small dataset [12], [13].

2.3. Data Augmentation

Data augmentation is a method that enables practitioners to significantly increase the variation of training data without the need to collect new data. In the context of machine learning, this technique is crucial for improving model performance and generalization ability. Generally, data augmentation is used in deep learning to increase the volume of data, allowing the model to be trained more optimally and produce more accurate predictions [14].

2.4. Convolutional Neural Networks (CNN)

Convolutional Neural Networks (CNNs) are a type of artificial neural network that mimic how the human brain extracts local features from input data and combines them into more complex features. CNNs are designed to process two-dimensional data and are commonly used in image processing [15]. CNN consists of several layers, including convolutional, pooling, and fully connected layers. The convolutional layers produce feature maps that represent the unique characteristics of an image, while the pooling layers reduce the model's complexity by combining neighboring pixels. Common types of pooling include max pooling and average pooling. CNNs have been used in various applications such as image classification, detection, and segmentation [16][17].

CNNs analyze input data, particularly images, using various convolutional layers. This allows CNN to process spatiotemporal features at high resolution and then transform them into more complex features at a lower resolution. The basic formula used in CNN for the convolution process is shown in Equation 1 [13].

$$\alpha_{i,j} = \sum_{m=0}^s \sum_{n=0}^s W_{m,n} x_{i+m,j+n} \quad (1)$$

The selection of the CNN architectures in this study is based on their demonstrated effectiveness in image classification tasks, particularly those involving fruit and other natural objects. ResNet50V2 is chosen for its ability to learn deep hierarchical features and mitigate vanishing gradient issues through residual learning, which has proven effective in various classification domains, including agricultural image tasks such as tomato maturity classification. Comparing ResNet50V2 and MobileNetV2 performance [18]. InceptionV3 is included for its capacity to capture multiscale visual patterns through mixed convolutional filters, a characteristic beneficial for differentiating subtle texture and shape variations in fruit images, as observed in comparative studies of fruit image classification models, where Inception-based architectures performed [19]. MobileNetV2 is selected due to its lightweight and computationally efficient design, making it suitable for practical deployment scenarios with limited resources, and has proven effective in fruit image classification tasks with transfer learning [19], [20]. Thus, combining these architectures allows a comprehensive evaluation of both high-capacity and resource-efficient CNN models for mango ripeness classification.

2.5. ResNet-50v2

ResNet-50v2 is the second version of the Residual Network architecture with 50 layers, designed to improve the training of deep networks without experiencing performance degradation. By utilizing residual learning, this architecture addresses the training challenges in very deep networks [21]. ResNet-50v2 improves the ordering of convolution and batch normalization operations to enhance efficiency and stability. It is widely used in image classification applications, such as plant disease classification and pneumonia detection from X-ray images, achieving excellent accuracy [22], [23].

2.6. Inception V3

Inception V3 is used as a deep learning model capable of recognizing patterns in digital images. The Inception V3 architecture is chosen for its ability to process complex visual features by combining different filter sizes within a single layer. With this structure, the model can effectively distinguish between various types of images based on the pattern differences formed from binary malware data [14].

2.7. Adam

Adam (Adaptive Moment Estimation) is an optimization algorithm widely used in deep learning model training due to its ability to combine the advantages of Momentum and RMSProp methods. Adam adaptively adjusts the learning rate for each parameter by utilizing estimates of the first moment (mean) and the second moment (variance) of the gradients. This allows Adam to handle sparse gradients and noisy data more efficiently, as well as accelerate model convergence [24].

In the study by Janarthanan and Ganesh (2023), the use of Adam in image classification demonstrated improved accuracy following proper hyperparameter tuning [25]. Furthermore, a study by Muflikhah (2024) employed Adam in an LSTM network for hypertension risk prediction based on SNP data, achieving an accuracy of 89% and an F1-score of 92%, demonstrating Adam's stability and reliability in handling complex biological data [26].

2.8. RMSprop

RMSprop is a gradient-based optimization algorithm developed to accelerate model training by adaptively adjusting the learning rate through a moving average of squared gradients, resulting in more stable convergence, especially in non-convex optimization problems [27][28][29]. RMSprop works by dividing the learning rate for each parameter by the square root of the exponential moving average of the squared gradients. This makes RMSprop highly effective at handling non-stationary objectives and suitable for use in large and deep neural networks, such as Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs).

2.9. MobileNetV2

MobileNetV2 is a convolutional neural network (CNN) architecture designed for high efficiency on resource-constrained devices, such as smartphones and IoT devices. Introduced by Sandler et al. in 2018, MobileNetV2 features two main innovations: inverted residual structures and linear bottlenecks [30]. MobileNetV2 utilizes inverted residual structures and linear bottlenecks to maintain rich feature representations with low complexity. The linear bottleneck, without nonlinear activation, prevents the loss of important information in narrow dimensions. Additionally, this architecture employs depthwise separable convolutions to reduce parameters and computational load, making it highly efficient for real-time applications on power-constrained devices [31].

2.10. Evaluation

Model evaluation is a crucial stage in the development process for machine learning or artificial neural network models, aiming to measure the model's accuracy and effectiveness in performing the assigned task. This process plays an important role in model validation and calibration, making the selection of appropriate evaluation metrics essential [13]. One of the commonly used metrics is accuracy [14], which calculates the proportion of correct predictions to the total number of predictions, as formulated in Equation 2.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

In addition to accuracy, several other important metrics are used to provide a more comprehensive assessment of model performance, especially in multi-class classification cases. Precision measures the proportion of correctly predicted positive instances and is defined as in Equation 3.

$$\text{Precision} = \frac{TP}{TP+FN} \quad (3)$$

Meanwhile, recall (sensitivity) measures the extent to which the model detects all actual positive instances, as formulated in Equation 4.

$$\text{Recall} = \frac{TP}{TP+FN} \quad (4)$$

To balance precision and recall, the F1-score is used, the harmonic mean of the two, as shown in Equation 5.

$$\text{F1 - Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

The evaluation presented in this study includes all four metrics to provide a comprehensive overview of the model's performance in classifying mango ripeness levels. Using multiple evaluation metrics enables more accurate identification of the model's strengths and weaknesses and supports informed decision-making for future performance improvements.

3. RESULTS AND DISCUSSION

3.1. Data Acquisition

The image dataset used in this study comprises mango fruit images collected from various sources and is divided into three main categories: rotten, unripe, and ripe mangoes. Each image has been annotated based

on the mango's distinct visual condition to support the automatic classification process. Data collection was conducted systematically to ensure representative variation in condition and lighting across each class. After collection, each image underwent manual inspection and annotation to ensure that the class labels matched the corresponding visual characteristics. Visual representations of the mango image dataset for each category are shown in Figure 2.



Figure 2. Mango Dataset

3.2. Data Preprocessing

After data collection, image preprocessing was performed by resizing all images to a uniform size and normalizing pixel values to the range 0–1 to ensure consistency and processing efficiency. Subsequently, data augmentation was applied using Keras's ImageDataGenerator to increase the diversity of the training data and reduce the risk of overfitting. Augmentation techniques included random rotations up to 30°, 15% zoom, horizontal and vertical shifts up to 20%, 15% shear, and random horizontal and vertical flips. Empty pixels resulting from transformations were filled with the nearest pixel values to preserve the image structure. The results of the augmented images are shown in Figure 3.

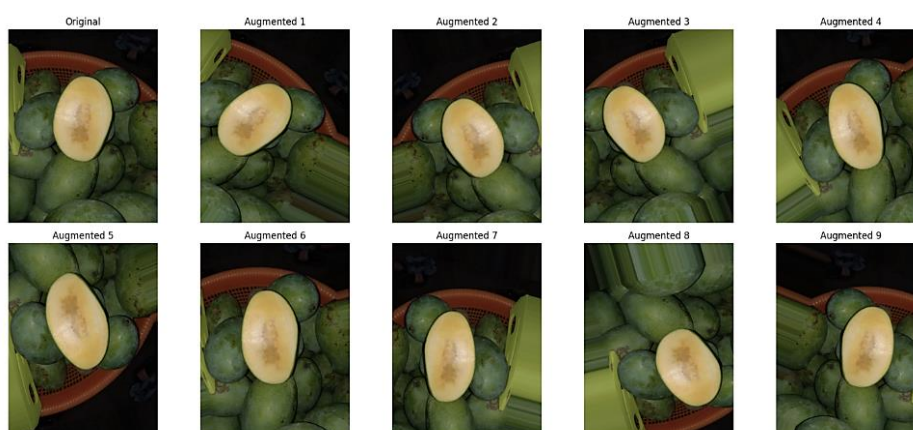


Figure 3. Augmentation Results

3.3. CNN Model

This study employs a transfer learning approach using three CNN architectures: ResNet50V2, InceptionV3, and MobileNetV2, each combined with a different optimization algorithm to classify mango images into three categories: rotten, unripe, and ripe. ResNet50V2 is used with the Adam optimizer due to the advantages of its residual structure and adaptive learning rate adjustment. InceptionV3 is paired with RMSprop to effectively handle multiscale features and visual complexity. Meanwhile, the lightweight and efficient

MobileNetV2 also uses Adam for fast, stable training. These combinations are expected to produce optimal classification results for mango ripeness levels.

3.4. CNN Modeling with ResNet50V2 Architecture and Adam Optimizer

Training the CNN model with the ResNet50V2 architecture and the Adam optimizer for 15 epochs showed a consistent increase in training accuracy, reaching up to 88%. However, validation accuracy only reached approximately 64% and remained stagnant after the 6th epoch. This discrepancy indicates the presence of overfitting. The training loss decreased significantly from around 1.4 to below 0.4, while the validation loss decreased more slowly and slightly increased toward the end of training. These trends suggest that although the model learns effectively from the training data, its performance on unseen data deteriorates. This behavior is illustrated in Figure 4.

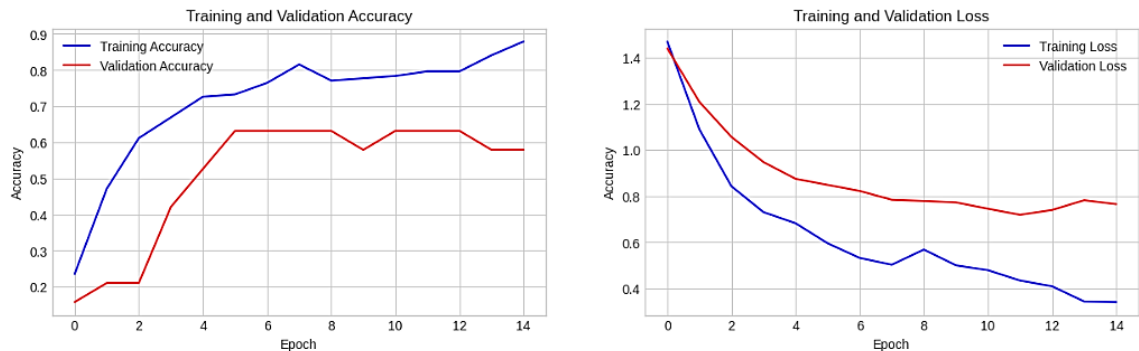


Figure 4. Training and Loss Graph of ResNet50V2 and Adam

As shown in Table 1, the image classification model for mango varieties demonstrated strong performance on both training and testing data, achieving accuracies of 85.35% and 81.82%, respectively. However, validation accuracy was only 63.16%, accompanied by a relatively high validation loss, suggesting potential overfitting. Overall, the model is capable of generalizing to new data reasonably well and performs stable classification of mango images, as reflected in Table 1.

Table 1. Accuracy Results of ResNet50V2 and Adam

	Train	Validation	Testing
Loss	0.358	0.718	0.429
Accuracy	0.853	0.631	0.818

To improve validation performance and further reduce overfitting, it is recommended to incorporate regularization techniques such as dropout, data augmentation, or L2 regularization, and to explore early stopping or fine-tuning deeper layers of the model.

3.5. CNN Modeling with InceptionV3 Architecture and RMSprop Optimizer

The CNN model, using the InceptionV3 architecture and the RMSprop optimizer, demonstrated a consistent increase in training accuracy, exceeding 80%. However, validation accuracy showed fluctuations and began to decline after the 5th epoch, indicating signs of overfitting. This is further supported by the steady decrease in training loss, while validation loss increased during the second half of training. These trends suggest that although the model learns effectively on the training data, its performance on unseen data degrades, indicating limited generalization. The training and validation patterns are illustrated in Figure 5.

The evaluation results summarized in Table 2 show that the model achieved a training accuracy of 87.26% with a training loss of 0.296. On the validation set, performance dropped to 78.95% accuracy with a loss of 0.590. Nevertheless, the model maintained relatively stable generalization when tested on unseen data, with a test accuracy of 77.27% and a loss of 0.558. Overall, despite the indications of overfitting, the model demonstrates adequate classification performance on the test data.

Table 2. Accuracy Results of InceptionV3 and RMSprop

	Train	Validation	Testing
Loss	0.296	0.590	0.558
Accuracy	0.872	0.789	0.772

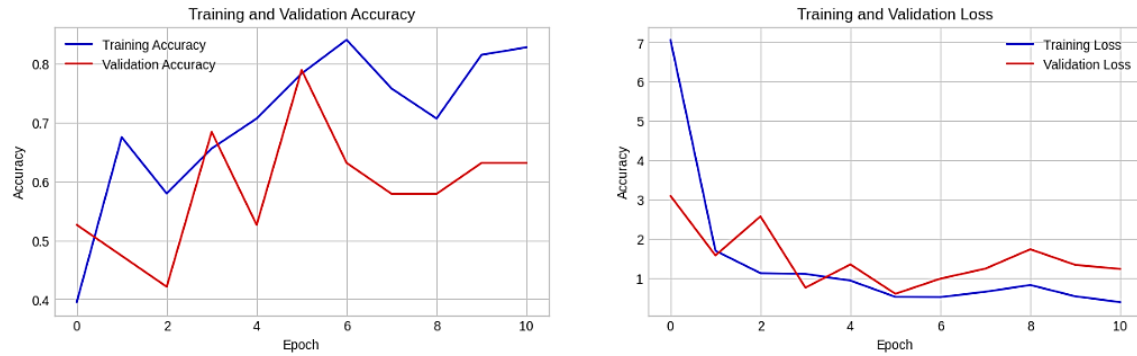


Figure 5. Training and Loss Graph of InceptionV3 and RMSprop

To address overfitting and enhance the model's generalization ability, additional strategies should be considered, such as applying regularization techniques (e.g., dropout or data augmentation), implementing early stopping, or fine-tuning the final layers of the InceptionV3 model.

3.6. CNN Modeling with MobileNetV2 Architecture and Adam Optimizer

The CNN model, employing the MobileNetV2 architecture and optimized with the Adam optimizer, demonstrated a substantial improvement in training accuracy, increasing from approximately 40% in the initial epoch to over 90% by the 14th epoch. In contrast, validation accuracy improved more slowly and exhibited noticeable fluctuations throughout training. This behavior was mirrored in the loss trends, where the training loss declined significantly to around 0.339, while the validation loss decreased more gradually and displayed variability across epochs.

The disparity between training and validation performance suggests a mild degree of overfitting. However, the model maintained relatively stable performance overall. The training and validation accuracy and loss trends are visualized in Figure 6.

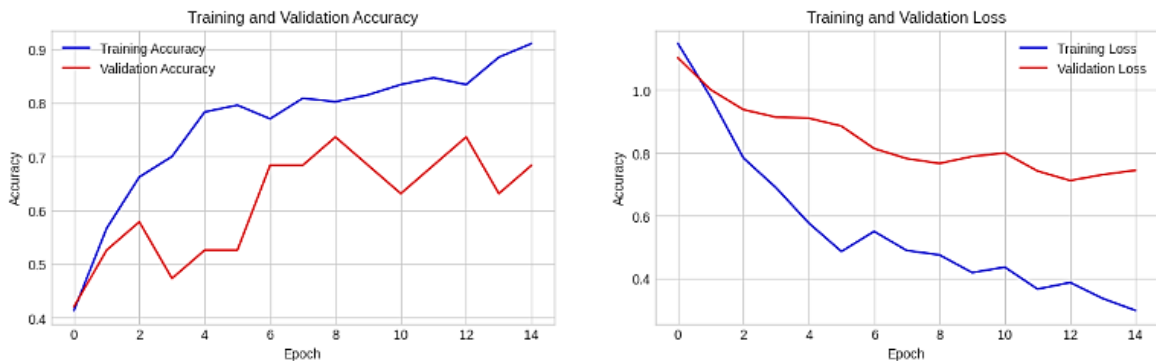


Figure 6. Training and Loss Graph of MobileNetV2 and Adam Optimizer

The evaluation metrics presented in Table 3 indicate that the model performed well on the training data, achieving an accuracy of 86.62% and a training loss of 0.339. On the validation dataset, the model's performance declined, with an accuracy of 73.68% and a loss of 0.712, suggesting possible overfitting. Despite this, the model generalized reasonably well to unseen data, achieving a testing accuracy of 77.27% and a loss of 0.540.

Table 3. Accuracy Results of MobileNetV2 and Adam Optimizer

	Train	Validation	Testing
Loss	0.339	0.712	0.540
Accuracy	0.866	0.736	0.772

In conclusion, MobileNetV2 trained with the Adam optimizer effectively learned the classification task and maintained stable performance on the test data, despite exhibiting some overfitting during validation. To enhance generalization and reduce overfitting, future work may incorporate regularization strategies such as dropout, L2 regularization, or advanced data augmentation techniques.

3.7. Confusion Matrix

Next, the model was evaluated based on its performance in classifying data using a Confusion Matrix. The confusion matrix results for the ResNet50V2 architecture with the Adam optimizer are presented in Figure 7. The confusion matrix for the InceptionV3 architecture with the RMSprop optimizer is shown in Figure 8. Meanwhile, the confusion matrix for the MobileNetV2 architecture using the Adam optimizer is shown in Figure 9.

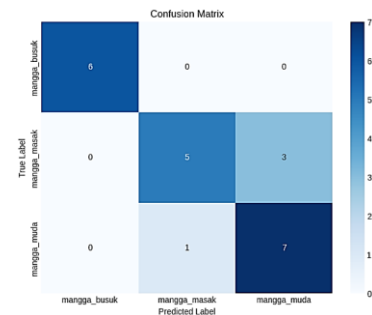


Figure 7. Confusion Matrix of ResNet50V2 Architecture and Adam Optimizer

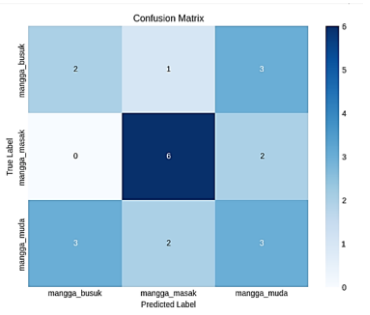


Figure 8. Confusion Matrix of InceptionV3 Architecture and RMSprop Optimizer

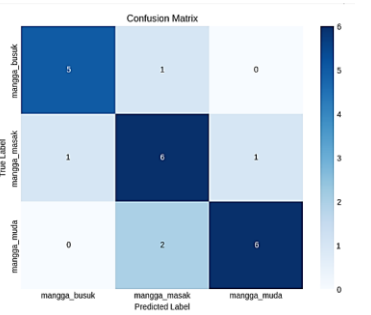


Figure 9. Confusion Matrix of MobileNetV2 Architecture and Adam Optimizer

Figures 7, 8, and 9 present the confusion matrices obtained from three different classification models for identifying rotten (mangga_busuk), ripe (mangga_masak), and unripe (mangga_muda) mangoes. A comparison of these matrices reveals substantial differences in classification capability among the evaluated models. Figure 6 demonstrates the strongest performance, where all rotten mangoes were correctly classified (6/6), while ripe and unripe mangoes achieved correct classifications of 5/8 and 7/8, respectively. The only noticeable errors occurred between ripe and unripe mangoes, indicating that these two maturity stages share similar visual characteristics, such as overlapping color distributions and surface textures. In contrast, Figure 7 exhibits the weakest classification performance, particularly for rotten and unripe mangoes. Only two rotten samples and three unripe samples were correctly recognized, while the remaining samples were frequently confused with other classes. This pattern suggests that the extracted image features were insufficient to clearly separate the three categories, leading to a high degree of overlap in the feature space. Figure 8, however, shows a considerable improvement over Figure 7, with correct classifications of 5 rotten, 6 ripe, and 6 unripe mangoes. The reduction in classification errors indicates that the corresponding model learned more representative discriminative features, resulting in better class separation and improved overall predictive performance.

From a machine learning perspective, the confusion matrices provide valuable insights into the strengths and weaknesses of the classification models beyond overall accuracy. The consistent misclassification observed between ripe and unripe mangoes in Figures 6 and 8 suggests that the transition from unripe to ripe is characterized by gradual changes in color and texture, making these classes inherently difficult to distinguish. Conversely, rotten mangoes were generally identified with higher accuracy because decay introduces more distinctive visual features, including dark discoloration, irregular texture, and structural deterioration, which are easier for the model to recognize. The poor performance shown in Figure 7 suggests either insufficient feature extraction or limited model generalization, leading to substantial confusion across all classes. In comparison, Figures 6 and 8 demonstrate that improved feature representation significantly reduces false classifications and produces a more balanced distribution of true positives across categories. These findings indicate that enhancing image preprocessing, increasing dataset diversity, applying data augmentation, or employing more advanced deep learning architectures could further improve classification performance. Overall, the confusion matrix analysis confirms that the model represented by Figure 6 provides the most reliable classification performance, while Figure 8 offers competitive results with slightly lower accuracy, and Figure 7 requires further optimization before practical deployment.

3.8. Evaluation

Figure 10 presents a comparison of the accuracy of the three CNN models used in this study, ResNet50V2, InceptionV3, and MobileNetV2, across the training, validation, and testing stages.

During the training phase, all three models demonstrated good performance, with accuracies of 85.3% (ResNet50V2), 87.2% (InceptionV3), and 86.6% (MobileNetV2), reflecting their ability to learn patterns from the training data. In the validation phase, InceptionV3 achieved the highest accuracy at 78.9%, followed by MobileNetV2 at 73.6%, while ResNet50V2 showed a significant drop to 63.1%, indicating signs of overfitting. However, during testing, ResNet50V2 achieved the highest accuracy of 81.8%, followed by InceptionV3 and

MobileNetV2 at 77.2% each. This indicates that although ResNet50V2 had lower validation performance, it generalized well to unseen data.

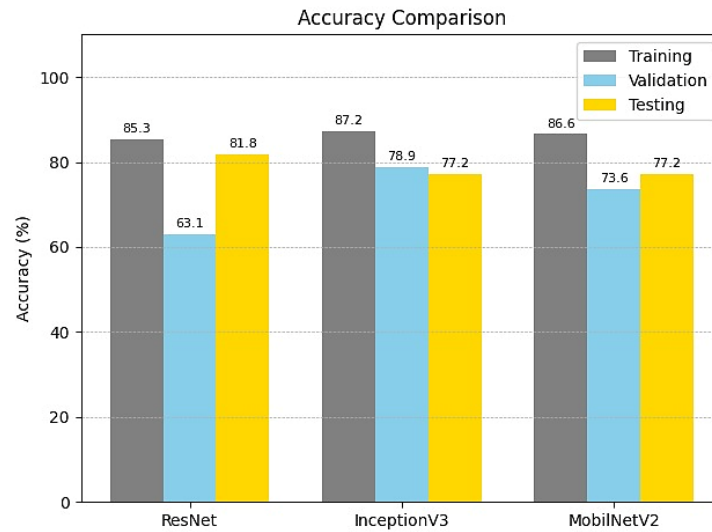


Figure 10. Accuracy Comparison

In addition to overall accuracy, confusion matrix analysis provides deeper insight into model performance across individual mango classes. The results show that rotten mangoes are generally easier to identify due to distinct visual features such as dark spots and surface damage. In contrast, ripe and unripe mangoes exhibit similar color and texture characteristics, which leads to a higher rate of misclassification between these two classes. Based on the confusion matrices, most misclassification errors occurred between ripe and unripe mango categories, indicating that distinguishing gradual transitions in ripeness remains challenging, particularly under varying lighting conditions.

Overall, each CNN model demonstrates distinct strengths. InceptionV3 shows the most consistent performance during training and validation, MobileNetV2 provides stable and efficient classification with lower computational complexity, while ResNet50V2 achieves the highest testing accuracy, indicating superior generalization capability on unseen data. These results suggest that ResNet50V2 is the most suitable model for real-world mango ripeness classification, while further improvements can be achieved through enhanced feature extraction and regularization techniques.

3.9. Discussion

The results indicate that all three CNN models ResNet50V2, InceptionV3, and MobileNetV2 successfully classified mango ripeness levels, particularly on the training dataset. These findings are consistent with previous studies reporting the effectiveness of transfer learning-based CNN architectures for fruit classification and ripeness detection. The use of pre-trained models allows the extraction of discriminative visual features such as color, texture, and surface patterns, which are essential for mango ripeness classification.

Among the evaluated models, InceptionV3 demonstrated the most consistent performance, achieving the highest validation accuracy (78.9%). This stability can be attributed to its multi-scale feature extraction mechanism, which enables the model to capture both local and global visual patterns effectively. MobileNetV2 also showed stable and efficient performance, making it suitable for applications with limited computational resources. Although ResNet50V2 exhibited signs of overfitting during validation, it achieved the highest test accuracy (81.8%), indicating strong generalization to unseen data. This behavior may be explained by the residual learning structure of ResNet50V2, which facilitates deeper feature learning and improves robustness.

From a practical perspective, the results demonstrate the potential of CNN-based models for implementation in automated mango sorting systems. Accurate classification of mango ripeness can support post-harvest quality control, reduce reliance on manual inspection, and improve sorting consistency. In particular, the high testing performance of ResNet50V2 suggests that this model is well-suited for real-world deployment where robustness to new data is required.

Despite the promising results, several limitations should be acknowledged. The dataset size used in this study is relatively limited, which may restrict the model's ability to generalize to broader variations in mango appearance. Additionally, similarities in visual characteristics between unripe and ripe mangoes, as well as variations in lighting conditions, may contribute to misclassification. Future research should consider expanding the dataset, increasing image diversity, and applying additional regularization techniques to reduce overfitting and improve validation stability.

4. CONCLUSION

This study successfully developed a mango ripeness classification model using a deep learning approach based on Convolutional Neural Networks (CNN). Three CNN architectures, ResNet50V2, InceptionV3, and MobileNetV2, were evaluated and demonstrated the ability to classify mangoes into three main categories: rotten, ripe, and unripe, with relatively high accuracy. ResNet50V2, combined with the Adam optimizer, achieved the best performance on the test data, while InceptionV3 showed the most consistent results during training and validation. These findings confirm that applying transfer learning and data augmentation techniques effectively improves image classification performance.

Despite these promising results, several limitations should be noted. The dataset used in this study was relatively limited in size and diversity, which may restrict the model's ability to generalize to broader variations in mango appearance. In addition, the experiments were conducted under relatively controlled conditions, and the model's performance in real-world environments has not yet been fully evaluated.

Future research should focus on expanding the dataset to include greater diversity in lighting conditions and visual characteristics, as well as on implementing the model in real-world automated sorting systems. Moreover, exploring additional regularization techniques and optimization strategies may further improve model robustness and stability. With further development, the proposed system has strong potential for practical implementation in agricultural industries, particularly for automated mango sorting, and could be extended to other horticultural commodities.

REFERENCES

- [1] A. A. Elsharif, I. M. Dheir, and S. S. Mettleq, Alaa Soliman Abu Abu-Naser, "Potato Classification Using Deep Learning Classification using Deep Learning," *Int. J. Acad. Pedagog. Res. IJAPR*, vol. 3, no. 12, pp. 1–8, 2019.
- [2] M. Han and C. Yi, "Deep Convolutional Neural Networks for Palm Fruit Maturity Classification," pp. 1–9, 2025.
- [3] L. A. Varga, J. Makowski, and A. Zell, "Measuring the Ripeness of Fruit with Hyperspectral Imaging and Deep Learning," *Proc. Int. Jt. Conf. Neural Netw.*, vol. 2021-July, 2021, doi: 10.1109/IJCNN52387.2021.9533728.
- [4] D. Worasawate, P. Sakunasinha, and S. Chiangga, "Automatic Classification of the Ripeness Stage of Mango Fruit Using a Machine Learning Approach," *AgriEngineering*, vol. 4, no. 1, pp. 32–47, 2022, doi: 10.3390/agriengineering4010003.
- [5] M. S. Morshed, S. Ahmed, T. Ahmed, M. U. Islam, and A. B. M. Ashikur Rahman, "Fruit Quality Assessment with Densely Connected Convolutional Neural Network," *12th Int. Conf. Electr. Comput. Eng. ICECE 2022*, pp. 461–464, 2022, doi: 10.1109/ICECE57408.2022.10088873.
- [6] P. K. Verma, P. Pathak, B. Kumar, Himani, and Preety, "Automatic Optical Imaging System for Mango Fruit using Hyperspectral Camera and Deep Learning Algorithm," *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 11, no. February, pp. 112–117, 2023, doi: 10.17762/ijritcc.v11i5s.6635.
- [7] A. Zakaria et al., "Improved maturity and ripeness classifications of Magnifera Indica cv. harumanis mangoes through sensor fusion of an electronic nose and acoustic sensor," *Sens. Switz.*, vol. 12, no. 5, pp. 6023–6048, 2012, doi: 10.3390/s120506023.
- [8] C. Velásquez, F. Prieto, L. Palou, S. Cubero, J. Blasco, and N. Aleixos, "New model for the automatic detection of anthracnose in mango fruits based on Vis/NIR hyperspectral imaging and discriminant analysis," *J. Food Meas. Charact.*, vol. 18, no. 1, pp. 560–570, 2024, doi: 10.1007/s11694-023-02173-3.
- [9] S. Kienzle, P. Sruamsiri, R. Carle, S. Sirisakulwat, W. Spreer, and S. Neidhart, "Harvest maturity specification for mango fruit (*Mangifera indica* L. 'Chok Anan') in regard to long supply chains," *Postharvest Biol. Technol.*, vol. 61, no. 1, pp. 41–55, 2011, doi: 10.1016/j.postharvbio.2011.01.015.
- [10] F. S. Mim, S. M. Galib, M. F. Hasan, and S. A. Jerin, "Automatic detection of mango ripening stages – An application of information technology to botany," *Sci. Hortic.*, vol. 237, no. March, pp. 156–163, 2018, doi: 10.1016/j.scienta.2018.03.057.
- [11] W. J. Sari et al., "Performance Comparison of Random Forest, Support Vector Machine and Neural Network in Health Classification of Stroke Patients," *Public Res. J. Eng. Data Technol. Comput. Sci.*, vol. 2, no. 1, pp. 34–43, 2024, doi: 10.57152/predatecs.v2i1.1119.
- [12] A. R. Rizalde and H. A. Mubarak, "A Deep Learning Approach Based on Classification to Detect Facial Skin Defect," vol. 2, no. January, pp. 93–106, 2025.
- [13] V. Wulandari, W. J. Sari, Z. H. Al-sawaff, and S. Manickam, "Comparative Analysis of Weather Image Classification Using CNN Algorithm with InceptionV3 , DenseNet169 and NASNetMobile Architecture Models," vol. 2, no. January, pp. 81–92, 2025.
- [14] M. Ahmed, N. Afreen, M. Ahmed, M. Sameer, and J. Ahamed, "An inception V3 approach for malware classification using machine learning and transfer learning," *Int. J. Intell. Netw.*, vol. 4, no. September 2022, pp. 11–18, 2023, doi: 10.1016/j.ijin.2022.11.005.

- [15] S. R. Suartika E. P, I Wayan, Wijaya Arya Yudhi, "Klasifikasi Citra Menggunakan Convolutional Neural Network (CNN) Pada Caltech 101," J. Tek. ITS, vol. 5, no. 1, p. 76, 2016.
- [16] A. Nawrocka, M. Nawrocki, and A. Kot, "Research study of image classification algorithms based on Convolutional Neural Networks," Proc. 2023 24th Int. Carpathian Control Conf. ICC3 2023, pp. 299–302, 2023, doi: 10.1109/ICC357093.2023.10178933.
- [17] N. Sharma, V. Jain, and A. Mishra, "An Analysis of Convolutional Neural Networks for Image Classification," Procedia Comput. Sci., vol. 132, no. Iccids, pp. 377–384, 2018, doi: 10.1016/j.procs.2018.05.198.
- [18] R. R. Santosa, A. Sasikirana, Z. L. Ramadhan, Y. P. B. Satria, and A. Nugroho, "Perbandingan Kinerja MobileNetV2 dan ResNet50V2 Dalam Klasifikasi Tingkat Kematangan Tomat," vol. 3, no. 1, 2025.
- [19] "Analisis Perbandingan Kinerja Model EfficientNetV2-S, InceptionV3, dan ResNet50 dalam Klasifikasi Gambar Buah-Buahan," J. Ilm. Komputasi, vol. 24, no. 3, Sep. 2025, doi: 10.32409/jikstik.24.3.3842.
- [20] A. C. Hartono and A. R. Muslikh, "Penerapan Transfer Learning MobileNetV2 Pada Klasifikasi Citra Jenis Buah-Buahan," vol. 3, no. 2, 2025, doi: 10.26905/jisad.v3i2.16187.
- [21] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 2016-December, pp. 770–778, 2016, doi: 10.1109/CVPR.2016.90.
- [22] D. Hastari, S. Winanda, A. R. Pratama, N. Nurhaliza, and E. S. Ginting, "Application of Convolutional Neural Network ResNet-50 V2 on Image Classification of Rice Plant Disease," Public Res. J. Eng. Data Technol. Comput. Sci., vol. 1, no. 2, pp. 71–77, 2024, doi: 10.57152/predatecs.v1i2.865.
- [23] M. Agil Izzulhaq and Alamsyah, "Penerapan Algoritma Convolutional Neural Network Arsitektur ResNet50V2 Untuk Mengidentifikasi Penyakit Pneumonia," Indones. J. Math. Nat. Sci., vol. 47, no. 1, pp. 12–22, 2024.
- [24] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," 3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc., pp. 1–15, 2015.
- [25] J. Sekar and G. T. Kumar, "Hyperparameter Tuning in Deep Learning-Based Image Classification to Improve Accuracy using Adam Optimization," Int. J. Perform. Eng., vol. 19, no. 9, pp. 579–586, 2023, doi: 10.23940/ijpe.23.09.p3.579586.
- [26] L. Muflikhah et al., "Single nucleotide polymorphism based on hypertension potential risk prediction using LSTM with Adam optimizer," Indones. J. Electr. Eng. Comput. Sci., vol. 33, no. 2, pp. 1126–1139, 2024, doi: 10.11591/ijeecs.v33.i2.pp1126-1139.
- [27] F. Zou, L. Shen, Z. Jie, W. Zhang, and W. Liu, "A Sufficient Condition for Convergences of Adam and RMSProp," no. 1.
- [28] S. De, A. Mukherjee, and E. Ullah, "Convergence guarantees for RMSProp and ADAM in non-convex optimization and an empirical comparison to Nesterov acceleration," pp. 1–33.
- [29] D. Zhou, "On the Convergence of Adaptive Gradient Methods for Nonconvex Optimization," no. 2014, pp. 1–25, 2024.
- [30] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., pp. 4510–4520, 2018, doi: 10.1109/CVPR.2018.00474.
- [31] L. Bai, Y. Zhao, and X. Huang, "A CNN Accelerator on FPGA Using Depthwise Separable Convolution," IEEE Trans. Circuits Syst. II Express Briefs, vol. 65, no. 10, pp. 1415–1419, 2018, doi: 10.1109/TCSII.2018.2865896.