Journal of Applied Informatics and Computing (JAIC)
Vol.10, No.4, August 2026, pp. 3402~3410
e-ISSN: 2548-6861

Banana Ripeness Classification Based on Color and Texture Using a CNN with ResNet50 Architecture

Raafiandy Wirawan Avicenna [1], Christy Atika Sari [2]*, Eko Hari Rachmawanto [3]
Informatics Engineering, Universitas Dian Nuswantoro, Semarang, Indonesia
111202315113@mhs.dinus.ac.id [1], christy.atika.sari@dsn.dinus.ac.id [2], eko.hari@dsn.dinus.ac.id [3]

Article Info

Article history:

Received 2026-07-01
Revised 2026-07-24
Accepted 2026-08-08

Keyword:

Banana Ripeness,
Image Classification,
ResNet50,
Color and Texture Features,
Confusion Matrix.

ABSTRACT

Banana ripeness is one of the main factors affecting fruit quality before distribution and consumption. Conventional ripeness assessment is generally performed through visual observation, which may lead to subjective and inconsistent results. Previous studies have reported promising results in banana ripeness classification. However, distinguishing adjacent ripeness stages remains challenging because of their similar visual characteristics. This study proposes a banana ripeness classification model using a Convolutional Neural Network (CNN) with the ResNet50 architecture. The dataset consisted of four ripeness categories, namely unripe, ripe, overripe, and rotten. Data balancing, preprocessing, and augmentation were applied before model training. A total of 1,120 images were used to train and evaluate the model. Transfer learning with full fine-tuning was employed to adapt pretrained visual features to different banana ripeness levels. The experimental results showed that the proposed model achieved an accuracy of 92.86%, while precision, recall, and F1-score reached 93%. Several misclassifications were observed between adjacent ripeness categories due to similarities in visual characteristics. These results indicate that the proposed ResNet50 model can effectively classify banana ripeness levels on the testing dataset based on color and texture information learned automatically from digital images.

This is an open access article under the CC–BY-SA license.

I. INTRODUCTION

Bananas are among the most widely cultivated horticultural commodities in Indonesia. People commonly consume bananas either as fresh fruit or as raw materials for various processed products. As a result, bananas have considerable economic value. In postharvest activities, fruit quality plays an important role in determining market value and consumer acceptance. Ripeness level is one of the indicators used to evaluate fruit quality. Ripeness is associated with changes in color, texture, taste, aroma, and shelf life. Therefore, ripeness must be considered during storage and distribution processes [1].

Most farmers and traders still determine banana ripeness through visual observation of peel color. This method is practical because it does not require special equipment. However, subjective factors often affect the assessment results. Differences in experience, lighting conditions, and fatigue levels may lead to different evaluations of the same fruit. These limitations become more apparent when the sorting process involves a large number of fruits. Under such conditions, the sorting process requires more time and increases the risk of classification errors [2].

Digital image processing has become an effective approach for automatic object identification. This technology enables computers to recognize visual characteristics such as color, shape, and texture. In agriculture, researchers have applied image processing to plant disease identification, crop quality classification, and fruit ripeness assessment. Compared with manual observation, image processing can provide faster and more consistent classification results [3]. Recent studies have further demonstrated that deep learning-based computer vision can be effectively applied to agricultural image analysis, including automated banana disease detection, thereby improving the accuracy and efficiency of image-based classification systems [4].

Convolutional Neural Network (CNN) is one of the most widely used approaches in image processing. CNN belongs to the deep learning family and can learn visual patterns directly from training data. Unlike conventional methods, CNN does not require manual feature extraction because the learning process occurs automatically during training. This advantage has made CNN a popular method for image classification tasks, including fruit ripeness identification [5]. Previous studies used CNN to identify banana ripeness levels based on peel color changes during the ripening process [6]. Other studies reported that CNN could classify bananas into several ripeness categories using visual information extracted from digital images [7]. Despite these promising results, the classification process remains challenging because the differences between ripeness levels are often subtle. Similar visual characteristics among classes may reduce the model's ability to distinguish objects accurately [8].

Researchers have also applied CNN to agricultural commodities other than bananas. For example, ResNet50 has also been applied to classify spice images, indicating its applicability to various agricultural image classification tasks [9]. Several studies reported that CNN can identify ripeness levels and product quality in different types of crops. Color and texture information extracted from images can help distinguish object conditions [10]. Similar approaches have also produced satisfactory results in studies involving other tropical fruits [11]. These findings indicate that deep learning has strong potential to support quality control in the agricultural sector.

Although CNN has achieved good performance in many image classification studies, deeper network architectures may introduce training difficulties. One of the most common problems is the vanishing gradient issue. This problem occurs when gradient information gradually decreases as it passes through multiple network layers, reducing the effectiveness of the learning process [12]. To address this issue, researchers developed the Residual Network (ResNet) architecture. ResNet applies residual learning through a skip connection mechanism that allows information to pass directly between layers without losing important features. ResNet50 is one of the most widely used variants. This architecture can extract more complex visual features and identify relatively small differences among object classes [10].

ResNet50 has shown good performance in banana ripeness classification [13]. It has also been applied to food quality classification using digital images [14]. The architecture can learn visual characteristics more effectively because changes in banana color and texture occur gradually during the ripening process. Although previous studies reported promising results, some of them used imbalanced datasets, resulting in different numbers of images in each class. This condition may cause the model to recognize certain classes more frequently than others. Moreover, previous studies primarily focused on improving overall classification accuracy while providing limited analysis of classification errors between adjacent ripeness stages, where color and texture characteristics often overlap. Consequently, distinguishing visually similar ripeness categories remains a challenge in banana ripeness classification. Based on these issues, this study applies a Convolutional Neural Network with the ResNet50 architecture to identify banana ripeness levels using color and texture information extracted from digital images. Although color and texture are the primary visual indicators of banana ripeness, these characteristics are learned automatically through the convolutional layers of the ResNet50 model rather than being extracted manually. The dataset consists of four ripeness categories: unripe, ripe, overripe, and rotten. Before training, the number of images in each category was balanced to provide a more representative distribution for every class. The ResNet50 model was initialized using pretrained ImageNet weights and subsequently fine-tuned on the balanced dataset to improve feature learning from a relatively limited number of training images. Model performance was evaluated using accuracy, precision, recall, F1-score, and confusion matrix analysis to provide a more comprehensive assessment of the proposed approach. Unlike previous studies that mainly focused on improving overall classification accuracy, this study emphasizes a more comprehensive evaluation of banana ripeness classification by combining a balanced dataset, full fine-tuning of ResNet50, and detailed confusion matrix analysis to investigate classification errors between adjacent ripeness stages. These contributions provide a more comprehensive understanding of model performance by complementing conventional evaluation metrics with an analysis of prediction errors caused by visual similarities in banana color and texture.

II. METHODS

A. Literature Review

Convolutional Neural Networks (CNNs) have been widely used for banana ripeness classification. Research in [8] applied a CNN-based approach to identify the ripeness of Indonesian banana cultivars and reported an accuracy of 88.67%. The results demonstrated that CNN models were able to learn visual patterns related to ripeness from digital images. Another study [3] utilized CNN architectures to classify banana ripeness and quality attributes using RGB images. The proposed models achieved accuracies of 89.22% and 90.42%, indicating that CNN-based methods can effectively extract color and texture features from banana images.

Despite these results, several challenges remain in banana ripeness classification. Differences in image acquisition conditions, unequal class distribution, and similarities between adjacent ripeness levels may affect model performance. An imbalanced dataset can lead the model to favor certain categories, reducing its ability to classify underrepresented classes accurately. Furthermore, similar color and texture characteristics between ripeness levels may increase the possibility of misclassification.

This study applies the ResNet50 architecture to classify banana ripeness into four categories: unripe, ripe, overripe, and rotten. A balancing process was carried out to ensure that each category contained the same number of images. Data preprocessing and augmentation were also performed before training to improve data quality and support feature learning. These procedures were implemented to obtain a more balanced data distribution and improve the model's ability to recognize each ripeness category.

B. *Research Stages*

$$N_{balanced} = k \times n \quad (1)$$

Figure 1 presents the research stages applied in this study. The flowchart illustrates the complete process, starting from dataset preparation and ending with model evaluation using several performance metrics.

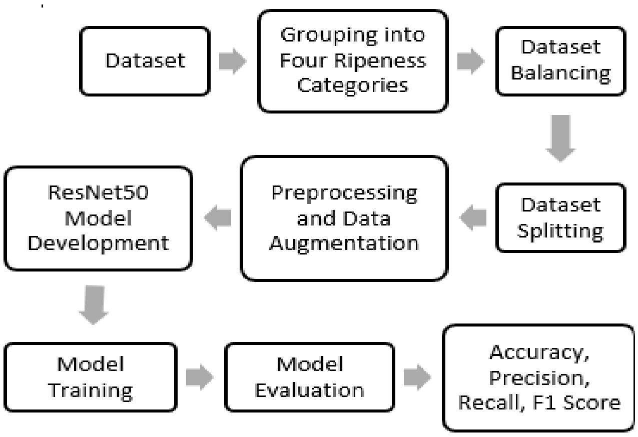

Figure 1. Research Stages

Research began with the preparation of a banana image dataset used for ripeness classification. Images were grouped into four ripeness categories: unripe, ripe, overripe, and rotten according to the available labels. After grouping, dataset balancing was performed to ensure an equal number of images in each category. This step reduced data imbalance that could affect the model's ability to learn characteristics from all classes evenly. The balanced dataset was then divided into training, validation, and testing sets. Next, image preprocessing and data augmentation were applied. Preprocessing adjusted image size and format according to model requirements, while augmentation increased image diversity in training data. Processed images were then used to develop a classification model based on ResNet50 architecture. Training was conducted using training data, whereas validation data were used to monitor learning performance. After training, testing data were used to evaluate classification performance on unseen images. Evaluation employed accuracy, precision, recall, and F1-score metrics.

C. *Banana Dataset*

Dataset used in this study came from BananaRipeness, a public repository developed by Luis Chuquimajo (https://github.com/luischuquim/BananaRipeness). BananaRipeness contains images grouped into four ripeness categories: unripe, ripe, overripe, and rotten. Each image is labeled according to its ripeness level, making it suitable for classification tasks. Previous studies have used ripeness-based image datasets to classify banana maturity levels using Convolutional Neural Network (CNN) and ResNet50 models. These approaches utilize color and texture characteristics of banana peels as primary features for distinguishing ripeness levels [15], [16].

Table I presents sample images from each ripeness category used in this study. Unripe bananas are dominated by green peel color, while ripe bananas show a more uniform yellow appearance. Overripe bananas exhibit brown spots on the peel surface, whereas rotten bananas show dark brown to black discoloration [16]. Based on Table II, image distribution across ripeness categories was not balanced. Unripe contained the largest number of images with 857 samples, while overripe contained the fewest images with 335 samples. Differences in sample size may cause a model to learn patterns from certain categories more frequently than others. Equation (1) calculates the total number of images after the balancing process, where (k) represents the number of ripeness categories and (n) represents the number of images assigned to each category. This study used four ripeness categories ((k = 4)) and 280 images for each category ((n = 280)). As a result, the balanced dataset contained 1,120 images.

TABLE I
BANANA IMAGES IN EACH RIPENESS CATEGORY

Class	**Class A**	**Class B**	**Class C**	**Class D**
Ripeness Category	Unripe	Ripe	Overripe	Rotten
Sample Images				

TABLE II
INITIAL DATASET DISTRIBUTION

Class	Ripeness Category	Number of Images
Class	Unripe	857
Class	Ripe	489
Class	Overripe	335
Class	Rotten	415
Total	All Categories	2,096

The initial dataset consisted of 857 unripe images, 489 ripe images, 335 overripe images, and 415 rotten images. To reduce class imbalance, the number of images in each category was adjusted to 280. Using the same number of images for all categories helped ensure a more balanced data distribution during training and allowed the model to learn representative features from each ripeness category more effectively [16]. Although the balanced dataset consisted of 1,120 images, the use of transfer learning with the pretrained ResNet50 model enabled effective feature learning from a relatively limited number of training images. In addition, data augmentation and regularization techniques were applied during training to increase image diversity and reduce the risk of overfitting.

D. Preprocessing, Data Augmentation, and Data Splitting

This study applied data augmentation to the training dataset through 20° rotation, 15% zoom, 20% horizontal and vertical shifts, and horizontal flipping. These augmentation techniques increased image variation and helped the model learn under different visual conditions. In addition, data augmentation reduced the risk of overfitting during the training process [6], [16]. After preprocessing, the balanced dataset was randomly divided into training, validation, and testing sets with a ratio of 70%, 20%, and 10%, respectively, using a fixed random seed of 42 to ensure reproducibility. Data augmentation was then applied only to the training dataset, while the validation and testing datasets were used without augmentation. The dataset was divided before data augmentation, ensuring that the training, validation, and testing sets remained separate throughout model development. This data split enabled model training, parameter validation, and final performance evaluation using separate image sets. The distribution of images in each ripeness category after balancing and data splitting is presented in Table III.

TABLE III
DATASET DISTRIBUTION AFTER BALANCING AND DATA SPLITTING

Class	Ripeness Category	Training Data (70%)	Validation Data (20%)	Testing Data (10%)	Total
Class A	Unripe	196	56	28	280
Class B	Ripe	196	56	28	280
Class C	Overripe	196	56	28	280
Class D	Rotten	196	56	28	280
Total	All Categories	784	224	112	1,120

Based on Table III, each ripeness category contained 280 images. Each category consisted of 196 training images, 56 validation images, and 28 testing images. In total, 784 images were allocated for training, 224 images for validation, and 112 images for testing. The balanced distribution reduced the impact of class imbalance and allowed each category to contribute equally during training. The testing set consisted of images that were not used during model training, providing a more objective evaluation of the model's performance on unseen test images [17].

E. ResNet50

After preprocessing, data augmentation, and dataset splitting, banana ripeness classification was performed using a Convolutional Neural Network (CNN) based on the ResNet50 architecture. CNN is widely used in image classification tasks because it can automatically learn visual features from images without requiring manual feature extraction. Previous studies have successfully applied CNN-based models to classify banana ripeness using color and texture information obtained from digital images [6], [7], [17].

A CNN architecture generally consists of convolution layers, pooling layers, and fully connected layers. The convolution layer extracts visual features from an input image using kernels or filters. The convolution operation is expressed in Equation (2). During convolution, the model extracts visual information such as color, texture, edges, and object patterns from input images [6], [17]. These visual characteristics are learned automatically through the convolutional filters during feature extraction rather than being extracted manually. Extracted features are then processed using the Rectified Linear Unit (ReLU) activation function shown in Equation (3).

$$S(i,j) = \sum m \sum n I(i-m, j-n) K(m,n) \tag{2}$$
$$f(x) = max\ (0,x) \tag{3}$$

ReLU preserves positive values and converts negative values to zero, allowing the network to learn non-linear patterns more effectively [6], [17]. Feature maps produced by this process are subsequently passed through pooling and fully connected layers before classification is performed. Final classification output is generated using the Softmax function shown in Equation (4). In this study, Softmax calculates the probability of each image belonging to one of four ripeness categories: unripe, ripe, overripe, and rotten. Category with the highest probability is selected as the final classification result [6], [13], [16]. Previous studies reported that CNN-based approaches achieved promising results in fruit ripeness classification. Deep learning, object detection, and image processing techniques have also been applied to improve classification accuracy in fruit maturity assessment [15], [18].

This study employs ResNet50 as the classification model. It was initialized using pretrained ImageNet weights, and a full fine-tuning strategy was applied by allowing all pretrained layers to remain trainable during training. This strategy enabled pretrained visual features to be adapted to the banana ripeness dataset. ResNet50 was developed to overcome the vanishing gradient problem that commonly occurs in deep neural networks. The architecture introduces residual learning through skip connections, allowing information from earlier layers to be passed directly to deeper layers. Residual learning is represented by Equation (5).

$$P(y_i) = \frac{e^{z_i}}{\sum_{j=1}^{n} e^{z_j}} \quad (4)$$

$$H(x) = F(x) + x \quad (5)$$

Within a residual block, input data are added directly to outputs generated by several network layers. This mechanism helps preserve information during training and enables deeper networks to learn more complex visual features. Previous studies reported that ResNet50 achieved high classification performance in fruit ripeness assessment based on visual image characteristics [10], [11], [12], [13], [16].

III. Results and Discussion

A. Model Architecture and Training Configuration

This section presents the architecture and training configuration used for banana ripeness classification. ResNet50 was selected as the base model because its residual learning mechanism helps overcome the vanishing gradient problem commonly encountered in deep neural networks.

- Transfer learning with pretrained ImageNet weights was applied to accelerate the training process and improve feature extraction. Since the available dataset was considerably smaller than ImageNet, pretrained weights provided a useful starting point for learning visual characteristics of banana images.
- The dataset consisted of 1,120 images grouped into four ripeness categories: Unripe (Class A), Ripe (Class B), Overripe (Class C), and Rotten (Class D). Images were divided into training, validation, and testing sets using a ratio of 70%, 20%, and 10%, respectively.
- Before training, all images were resized to 224 × 224 pixels and processed using the preprocessing function associated with ResNet50. These steps ensured that image data matched the input requirements of the network.

GPU acceleration with sufficient memory capacity was used to support parallel computation during model training. Details of model architecture and training parameters are presented in Table IV.

TABLE IV
RESNET50 MODEL ARCHITECTURE

No	Layer	Output Shape	Parameters
1	ResNet50 base	7 × 7 × 2048	23,587,712
2	GlobalAveragePooling2D	2048	0
3	BatchNormalization	2048	8,192
4	Dense (ReLU + L2)	512	1,049,088
5	Dropout	512	0
6	Dense (Softmax)	4	2,052
	Total Parameters		24,647,044

TABLE V
MODEL TRAINING HYPERPARAMETER CONFIGURATION

Hyperparameter	Value / Configuration
Optimization Function	Adam
Learning Rate	0.00001
Max Epochs	100
Mini Batch Size	32
Dropout Rate	0.2
L2 Regularization	0.01
Input Size	224 x 224 pixels
Loss Function	Categorical Crossentropy
Evaluation Metric	Accuracy

Transfer learning was implemented using a full fine-tuning strategy, allowing all pretrained layers to remain trainable during training. As shown in Table IV, feature maps generated by ResNet50 were processed through a GlobalAveragePooling2D layer and a BatchNormalization layer before entering the classification stage. A fully connected layer with 512 neurons was then applied, followed by a dropout layer and a Softmax output layer for four-class classification. Details of training configuration and model parameters are presented in Table V.

According to Table V, the model was trained using the Adam optimizer with a learning rate of 0.00001 and a mini batch size of 32. Training was conducted for 100 epochs. L2 regularization with a value of 0.01 and a dropout rate of 0.2 were applied to reduce overfitting during training. A fully connected layer containing 512 neurons was used in the classification stage before the Softmax output layer.

Classification performance was evaluated using accuracy, precision, recall, F1-score, and confusion matrix analysis, while model optimization was performed using the categorical crossentropy loss function.

B. *Training Configuration*

After the model architecture was defined, training parameters were configured to optimize classification performance. Training was conducted for 100 epochs with a batch size of 32. This batch size was selected to maintain a balance between computational efficiency and training stability. ModelCheckpoint was used to automatically save the model with the highest validation accuracy during training. Consequently, performance evaluation was performed using the best model obtained from the training process. A learning rate of 0.00001 was applied during fine-tuning to support gradual parameter updates. The Adam optimizer was used to optimize model weights throughout training. In addition, L2 regularization and a dropout rate of 0.2 were applied to reduce overfitting and improve model generalization.

C. *Model Performance Evaluation*

This section presents implementation results and performance evaluation of the proposed method. Implementation and testing began with preparation of a computing environment to support efficient and stable model training. Training was conducted on a system equipped with an Intel® Core™ i7-12700H processor and an NVIDIA GeForce RTX 3060 Laptop GPU with 6 GB of VRAM. Python, TensorFlow, and Keras were used to develop and train the model. This hardware and software configuration provided sufficient computational resources to train the ResNet50 architecture with a batch size of 32. After training for 100 epochs, model performance was analyzed using accuracy and loss curves. These metrics illustrate learning behavior and convergence during training. Trends in accuracy and loss are presented in Figure 2. Figure 2 presents training and validation accuracy and loss over 100 epochs. During the early stages of training, both accuracy curves increased rapidly, indicating that the model learned relevant visual features from banana images. Training accuracy continued to improve throughout the training process and reached approximately 99% in the final epochs. Validation accuracy increased substantially during the first 20 epochs and then remained relatively stable between 87% and 89% until the end of training. Loss curves show a gradual decrease for both training and validation data. Training loss consistently declined as the number of epochs increased. A similar trend was observed for validation loss, which continued to decrease throughout the training process. Results show that the model converged during training and maintained stable performance throughout the learning process. Although a performance gap between training and validation accuracy was observed, the use of data augmentation, L2 regularization, and dropout helped reduce overfitting and maintain stable validation performance throughout training.

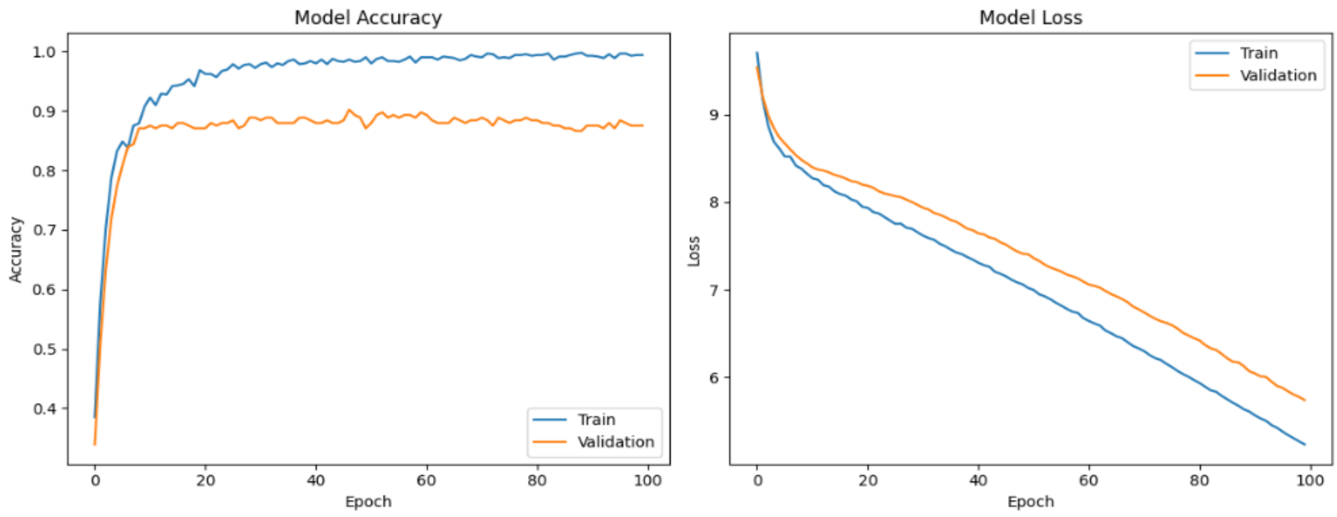

Figure 2. Model Accuracy and Loss Curves

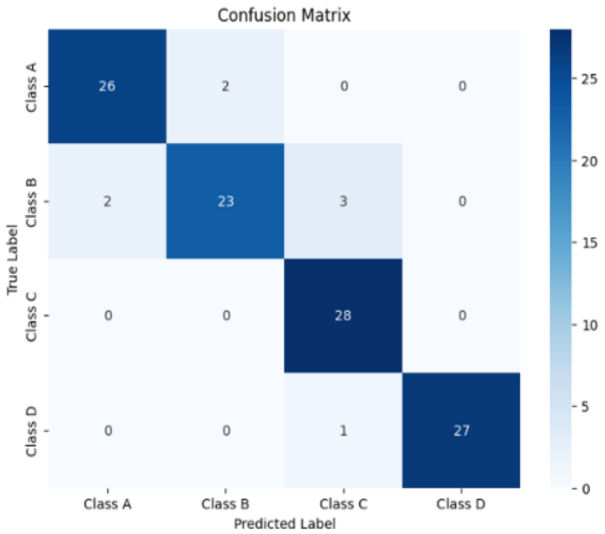

Figure 3. Confusion Matrix of Model Testing Results

Figure 3 presents the confusion matrix obtained from the ResNet50 classification model. Most testing samples were classified correctly, as indicated by the high values along the main diagonal of the matrix. Class C achieved the best result, with all 28 samples classified correctly. Class D also showed strong performance, with 27 of 28 samples correctly identified and only one sample misclassified as Class C. Similarly, Class A recorded 26 correct predictions, while two samples were classified as Class B. These results indicate that the model was able to recognize most ripeness categories accurately.

Most classification errors occurred in Class B. Of the 28 testing samples in this category, 23 were classified correctly, while two samples were predicted as Class A and three samples as Class C. These errors suggest that neighboring ripeness categories may share similar visual characteristics, particularly in terms of peel color and texture. Although several misclassifications were observed, the majority of testing samples were assigned to the correct category. Overall, the confusion matrix indicates that the proposed model achieved high classification performance on the testing dataset.

Table VI presents the classification performance of the ResNet50 model on the testing dataset. The model achieved an overall accuracy of 92.86%, indicating that most testing images were assigned to the correct ripeness category. Precision, recall, and F1-score values were consistently high across all classes, suggesting that the model was able to distinguish banana ripeness levels with good reliability. Class D achieved the highest performance, with a precision of 1.00, a recall of 0.96, and an F1-score of 0.98. Class A and Class C also produced strong results, both obtaining an F1-score of 0.93. Class C achieved a recall of 1.00, indicating that all samples in this category were correctly identified. In contrast, Class B recorded the lowest F1-score of 0.87 and a recall of 0.82. This result is consistent with the confusion matrix presented in Figure 3, where most classification errors occurred in Class B. Similar visual characteristics between neighboring ripeness stages may have contributed to these misclassifications. Overall, F1-score values ranged from 0.87 to 0.98, indicating balanced classification performance across all categories. Combined with the confusion matrix analysis, these results indicate that the proposed model achieved good classification performance on the testing dataset used in this study.

TABLE VI
DETAILED CLASSIFICATION REPORT

Class	Precision	Recall	F1-Score	Support
Class A	0.93	0.93	0.93	28
Class B	0.92	0.82	0.87	28
Class C	0.88	1.00	0.93	28
Class D	1.00	0.96	0.98	28
Accuracy			0.93	112

D. Discussion

Results obtained from the testing dataset show that the proposed ResNet50 model classified banana ripeness into four categories with an overall accuracy of 92.86%. Performance presented in Figure 2, Figure 3, and Table VI indicates that the model was able to distinguish most testing samples correctly. The balanced dataset and image augmentation process may have helped the model learn visual variations found in each ripeness category.

Differences in classification performance were observed among the four classes. Class C and Class D achieved the best results, whereas most classification errors occurred in Class B. Based on the confusion matrix, several samples from Class B were predicted as Class A or Class C. This finding suggests that visual differences between ripe and neighboring ripeness stages were less distinct than those observed in unripe or rotten bananas. Similarities in peel color and texture may have contributed to these errors. During the ripening process, chlorophyll gradually degrades while carotenoid pigments become more visible, resulting in continuous color changes between adjacent ripeness stages. Texture also changes progressively as the fruit matures, further reducing visual differences between adjacent ripeness stages and increasing the possibility of misclassification. A comparable observation was reported in previous studies, where adjacent ripeness categories were more likely to be misclassified because of overlapping visual characteristics [6].

Training and validation curves shown in Figure 2 indicate that the model converged during training. Validation accuracy remained relatively stable after the initial training stage, while validation loss gradually decreased until the final epochs. Although training accuracy was higher than validation accuracy, validation performance remained consistent throughout training. These results indicate that the model maintained stable learning behavior and was able to perform well on testing data.

Compared with conventional visual inspection, the proposed approach provides a more objective method for evaluating banana ripeness. Automated classification using ResNet50 can reduce subjectivity in the assessment process and provide consistent predictions across different ripeness categories. The proposed model has the potential to be integrated into automated fruit sorting systems by combining image acquisition with classification to identify banana ripeness during the sorting process. A similar approach could also be implemented in mobile applications to assist farmers, distributors, and retailers in assessing banana ripeness directly in the field. Practical implementation, however, requires consideration of lighting conditions, camera quality, image acquisition distance, and processing speed because these factors may affect classification performance. Further improvements in model robustness under different operating conditions are needed to support future real-world implementation.

TABLE VII
COMPARISON OF CLASSIFICATION PERFORMANCE

Study	Method	Accuracy (%)
[3]	Dual CNN	90.42
[8]	VGG based CNN	88.67
This study	ResNet50	92.86

Table VII compares the classification performance obtained in this study with results reported in previous research on banana ripeness classification. Study [3] achieved an accuracy of 90.42% using a Dual CNN architecture, while study [8] reported an accuracy of 88.67% using a VGG based CNN model. In the present study, the proposed ResNet50 model achieved an accuracy of 92.86% for classifying four banana ripeness categories. Compared with the studies presented in Table VII, the proposed model produced higher classification accuracy. Differences in performance may be influenced by dataset characteristics, image acquisition conditions, preprocessing techniques, data augmentation strategies, and model configuration. In addition, similarities in color and texture between neighboring ripeness stages may affect classification performance and contribute to prediction errors.

Research on fruit ripeness classification has expanded through the use of image processing, computer vision, and deep learning techniques. Deep learning has been applied to classify fruit maturity levels [19], while image processing combined with fuzzy logic has been used to classify banana size and ripeness [20]. Other studies have explored object detection models such as YOLOv8 for fruit ripeness identification [21]. Multimodal approaches that combine visible-light and hyperspectral imaging have also been reported to improve fruit maturity estimation [22]. An improved banana ripeness detection model was later developed for complex environmental conditions [23]. Comparative studies further showed that deep learning methods generally achieve better performance than conventional machine learning approaches in image-based fruit classification tasks [24]. CNN-based approaches have also been applied to banana ripeness classification. Previous research demonstrated that CNN can be used to classify both banana varieties and ripeness levels [25]. Model robustness under varying illumination conditions was evaluated in [26], while computer vision and deep learning techniques were applied to identify chemically ripened bananas in [27]. Similar approaches have also been used for fruit and vegetable ripeness assessment based on visual image characteristics [28].

Results obtained in this study are consistent with previous findings. Color and texture remain important features for distinguishing different ripeness levels. Most testing samples were classified correctly, although several misclassifications occurred between neighboring categories with similar visual characteristics. This observation is consistent with previous studies that reported classification difficulties when differences in color and texture become less distinct between ripeness stages [16]. These findings were obtained from images collected under specific acquisition conditions and should be interpreted within the scope of this study. Classification performance may vary when images are captured under different lighting conditions, camera specifications, background settings, or banana cultivars. These factors were not evaluated in the present study and may influence model performance in practical applications. Future studies should evaluate more diverse datasets collected under different environmental conditions to further assess model robustness and generalization capability.

IV. CONCLUSION

This study applied the ResNet50 architecture to classify banana ripeness into four categories: unripe, ripe, overripe, and rotten using digital images. The proposed model achieved an accuracy of 92.86% on the testing dataset, demonstrating its ability to learn color and texture characteristics associated with banana ripeness. Most testing samples were classified correctly, although several misclassifications occurred between neighboring categories with similar visual characteristics. The results indicate that ResNet50 achieved good classification performance on the testing dataset used in this study. This approach has potential applications in post-harvest handling, quality control, and agricultural production processes that require rapid and consistent ripeness assessment. Future work may focus on expanding the dataset, increasing image diversity under different environmental conditions, comparing ResNet50 with other deep learning architectures, performing repeated experiments to evaluate model stability, and validating the model using external datasets collected under different acquisition conditions.

REFERENCES

[1] R. Kosasih, Sudaryanto, and A. Fahrurozi, "Classification of six banana ripeness levels based on statistical features on machine learning approach," *International Journal of Advances in Applied Sciences*, vol. 12, no. 4, pp. 317–326, Dec. 2023, doi: 10.11591/ijaas.v12.i4.pp317-326.

[2] W. Widyawati and R. Febriani, "Real-time detection of fruit ripeness using the YOLOv4 algorithm," *Teknika: Jurnal Sains dan Teknologi*, vol. 17, no. 2, p. 205, Nov. 2021, doi: 10.36055/tjst.v17i2.12254.

[3] O. Martínez-Mora *et al.*, "Artificial Vision-Based Dual CNN Classification of Banana Ripeness and Quality Attributes Using RGB Images," *Processes*, vol. 13, no. 7, Jul. 2025, doi: 10.3390/pr13071982.

[4] M. Syifa Aryanta, C. A. Sari, and H. Rachmawanto, "A Banana Disease Detection Using MobileNetV2 Model Based on Adam Optimizer," 2025. [Online]. Available: http://jurnal.polibatam.ac.id/index.php/JAIC

[5] R. E. Saragih and A. W. R. Emanuel, "Banana Ripeness Classification Based on Deep Learning using Convolutional Neural Network."

[6] N. Nafi'iyah, R. Wardhani, and E. Prakasa, "Identification of Banana Ripeness using Convolutional Neural Network Approaches", [Online]. Available: https://orcid.org/0000-0003-4685-6309

[7] Y. A. Nisa, C. A. Sari, E. H. Rachmawanto, and N. Mohd Yaacob, "Ambon Banana Maturity Classification Based On Convolutional Neural Network (CNN)," *sinkron*, vol. 8, no. 4, pp. 2568–2578, Oct. 2023, doi: 10.33395/sinkron.v8i4.12961.

[8] C. Dewi, S. Ueno, and K. Kato, "Artificially Ripeness Identification of Indonesian Banana Cultivar Using Convolution Neural Network," in *ACM International Conference Proceeding Series*, Association for Computing Machinery, Oct. 2023, pp. 107–111. doi: 10.1145/3626641.3626668.

[9] C. A. Sari, L. Hamidah, S. Pradana, and E. H. Rachmawanto, "LogicLink: Journal of Artificial Intelligence and Multimedia in Informatics Identifikasi Citra Jenis Rempah-Rempah Menggunakan Arsitektur RestNet50," vol. 2, no. 1, pp. 27–41, 2025.

[10] E. Prasiwiningrum and Adyanata Lubis, "Classification Of Palm Oil Maturity Using CNN (Convolution Neural Network) Modelling RestNet 50," *Decode: Jurnal Pendidikan Teknologi Informasi*, vol. 4, no. 3, pp. 983–999, Nov. 2024, doi: 10.51454/decode.v4i3.822.

[11] C. Pornpanomchai, "Image based durian (Durie azomethines Linn) sweetness measurement by ResNet50," *Engineering and Applied Science Research*, vol. 51, no. 2, pp. 200–210, Jan. 2024, doi: 10.14456/easr.2024.20.

[12] S. Hardiyanti Putri, "ResNet50-Based Deep Learning Architecture with Focal Loss Optimization for Automated Fruit Ripeness Classification," *Jurnal SISFOKOM (Sistem Informasi dan Komputer)*, 2025, doi: 10.32736/sisfokom.v15i1.2449.

[13] M. Khalis Fikri *et al.*, "Banana Ripeness Classification Using Convolutional Neural Network Based on Resnet-50," *Jurnal Media Elektrik*, vol. 23, no. 2, pp. 157–165, Apr. 2026, doi: 10.59562/metrik.v23i2.11482.

[14] M. M. I. Al-Ghiffary, C. A. Sari, E. H. Rachmawanto, N. R. D. Cahyo, N. M. Yaacob, and R. R. Ali, "Milkfish Freshness Classification Using Convolutional Neural Networks Based on Resnet50 Architecture," *Advance Sustainable Science, Engineering and Technology*, vol. 5, no. 3, Aug. 2023, doi: 10.26877/asset.v5i3.17017.

[15] P. Baglat, A. Hayat, F. Mendonça, A. Gupta, S. S. Mostafa, and F. Morgado-Dias, "Non-Destructive Banana Ripeness Detection Using Shallow and Deep Learning: A Systematic Review," Jan. 01, 2023, *MDPI*. doi: 10.3390/s23020738.

[16] Tjokorda Istri Agung Pandu Yuni Maharani, I Gusti Agung Indrawan, Gede Dana Pramitha, Christina Purnama Yanti, and I Made Marthana Yusa, "Classification of Cavendish Banana Ripeness With CNN Method," *Indonesian Journal of Data and Science*, vol. 6, no. 2, pp. 203–210, Jul. 2025, doi: 10.56705/ijodas.v6i2.259.

[17] S. Anatya, V. C. Mawardi, and J. Hendryli, "Fruit maturity classification using convolutional neural networks method," in *IOP Conference Series: Materials Science and Engineering*, IOP Publishing Ltd, Dec. 2020. doi: 10.1088/1757-899X/1007/1/012149.

[18] R. Erz Saragih, A. Rezki Purnajaya, I. Syafrinal, and Y. Pernando, "Mango and Banana Ripeness Detection based on Lightweight YOLOv8."

[19] M. Winklmair, R. Sekulic, J. Kraus, P. Penava, and R. Buettner, "A deep learning based approach for classifying the maturity of cashew apples," *PLoS One*, vol. 20, no. 6 June, Jun. 2025, doi: 10.1371/journal.pone.0326103.

[20] B. A. Malabag, C. S. Santiago, E. L. Cahapin, J. L. Reyes, and G. S. Legaspi, "Fuzzy Logic-Based Size and Ripeness Classification of Banana using Image Processing Technique," *International Journal of Emerging Technology and Advanced Engineering*, vol. 12, no. 10, pp. 11–18, Oct. 2022, doi: 10.46338/ijetae1022_02.

[21] B. Xiao, M. Nguyen, and W. Q. Yan, "Fruit ripeness identification using YOLOv8 model," *Multimed. Tools Appl.*, vol. 83, no. 9, pp. 28039–28056, Mar. 2024, doi: 10.1007/s11042-023-16570-9.

[22] C. A. Garillos-Manliguez and J. Y. Chiang, "Multimodal deep learning and visible-light and hyperspectral imaging for fruit maturity estimation," *Sensors (Switzerland)*, vol. 21, no. 4, pp. 1–18, Feb. 2021, doi: 10.3390/s21041288.

[23] G. Wang, Y. Gao, F. Xu, W. Sang, Y. Han, and Q. Liu, "A Banana Ripeness Detection Model Based on Improved YOLOv9c Multifactor Complex Scenarios," *Symmetry (Basel).*, vol. 17, no. 2, Feb. 2025, doi: 10.3390/sym17020231.

[24] N. O. M. Salim and A. K. Mohammed, "Comparative Analysis of Classical Machine Learning and Deep Learning Methods for Fruit Image Recognition and Classification," *Traitement du Signal*, vol. 41, no. 3, pp. 1331–1343, Jun. 2024, doi: 10.18280/ts.410322.

[25] L. Jambola, A. R. Darlis, W. Malaha, and D. Aryanta, "Classification system of banana types and ripeness levels based on convolutional neural network," *IAES International Journal of Artificial Intelligence*, vol. 14, no. 6, pp. 4891–4901, Dec. 2025, doi: 10.11591/ijai.v14.i6.pp4891-4901.

[26] L. E. Chuquimarca, B. X. Vintimilla, and S. A. Velastin, "Assessing deep learning model robustness for banana ripeness classification under varying illumination conditions," *Smart Agricultural Technology*, vol. 12, Dec. 2025, doi: 10.1016/j.atech.2025.101333.

[27] A. S and J. Kathirvelan, "Computer vision-based detection and classification of chemically ripened bananas and papayas at vendor site through deep learning AI models using real-time dataset," *Results in Engineering*, vol. 26, Jun. 2025, doi: 10.1016/j.rineng.2025.104730.

[28] E. Tapia-Mendez, I. A. Cruz-Albarran, S. Tovar-Arriaga, and L. A. Morales-Hernandez, "Deep Learning-Based Method for Classification and Ripeness Assessment of Fruits and Vegetables," *Applied Sciences (Switzerland)*, vol. 13, no. 22, Nov. 2023, doi: 10.3390/app132212504.