Techno Nusa Mandiri: Journal of Computing and Information Technology
Vol. 23, No. 2 September 2026 | DOI https://doi.org/10.33480/techno.v23i2.8745

COMPARATIVE ANALYSIS OF THE APPLICATION OF CONVOLUTIONAL NEURAL NETWORKS (CNN) FOR BANANA RIPENESS CLASSIFICATION

Rizqia Fauziah Rachma[1*]; Muhammad Bagus Andra[2]

Computer Science[1,2]
Universitas Nusa Mandiri, Jakarta, Indonesia[1,2]
https://nusamandiri.ac.id/[1,2]
14250002@nusamandiri.ac.id[1*], muhammad.mba@nusamandiri.ac.id[2]

(*) Corresponding Author



The creation is distributed under the Creative Commons Attribution-NonCommercial 4.0 International License.

Abstract— *Banana ripeness classification is an important component of post-harvest agricultural management because ripeness affects product quality, economic value, and marketability. Manual assessment remains subjective, labor-intensive, and susceptible to human error, motivating the development of automated image-based classification systems. This study presents a systematic literature review (SLR) of CNN-based approaches for banana ripeness classification based on 15 selected scholarly studies published between 2023 and 2026. The review synthesizes reported applications of standard CNN, VGG16, VGG19, ResNet-based architectures, and EfficientNet, while examining the influence of preprocessing, data augmentation, class imbalance, cultivar variability, evaluation metrics, model interpretability, and computational efficiency. The reviewed studies generally reported high classification accuracy, but performance varied according to dataset composition, cultivar, image acquisition conditions, preprocessing procedures, and training configurations. Accuracy was the most consistently reported metric, while Precision, Recall, and F1-score provided additional information when available. The synthesis further identified practical challenges related to dataset diversity, model interpretability, and computational requirements for deployment on mobile or edge devices. Based on these findings, a conceptual framework is proposed to integrate data acquisition, preprocessing and imbalance handling, model selection, multi-metric evaluation, explainable AI, and deployment considerations. Future research should prioritize diverse real-world datasets, class-sensitive learning strategies, interpretable models, and computationally efficient architectures for practical agricultural applications.*

Keywords: *Banana Ripeness, Convolutional Neural Network (CNN), Deep Learning, Image Classification, Systematic Literature Review.*

Intisari— Klasifikasi kematangan pisang merupakan salah satu komponen penting dalam manajemen pascapanen karena tingkat kematangan memengaruhi kualitas, nilai ekonomi, dan daya jual produk. Penilaian secara manual masih bersifat subjektif, membutuhkan tenaga, dan rentan terhadap kesalahan manusia, sehingga diperlukan pengembangan sistem klasifikasi otomatis berbasis citra. Penelitian ini menyajikan Systematic Literature Review (SLR) terhadap pendekatan berbasis CNN untuk klasifikasi kematangan pisang berdasarkan 15 artikel ilmiah terpilih yang diterbitkan pada periode 2023–2026. Tinjauan ini menyintesis penerapan CNN standar, VGG16, VGG19, arsitektur berbasis ResNet, dan EfficientNet, serta mengkaji pengaruh preprocessing, augmentasi data, ketidakseimbangan kelas, variasi kultivar, metrik evaluasi, interpretabilitas model, dan efisiensi komputasi. Studi yang ditinjau secara umum melaporkan tingkat akurasi yang tinggi, tetapi kinerja bervariasi berdasarkan komposisi dataset, kultivar, kondisi pengambilan citra, prosedur preprocessing, dan konfigurasi pelatihan. Akurasi merupakan metrik yang paling konsisten dilaporkan, sedangkan Precision, Recall, dan F1-score memberikan informasi tambahan ketika tersedia. Sintesis juga mengidentifikasi tantangan praktis terkait keberagaman dataset, interpretabilitas model, dan kebutuhan komputasi untuk penerapan pada perangkat seluler atau edge. Berdasarkan temuan tersebut, penelitian ini mengusulkan kerangka konseptual yang mengintegrasikan akuisisi data, preprocessing dan penanganan ketidakseimbangan kelas, pemilihan model, evaluasi multi-metrik, Explainable AI, dan*

pertimbangan deployment. Penelitian selanjutnya perlu memprioritaskan dataset dunia nyata yang beragam, strategi pembelajaran yang peka terhadap ketidakseimbangan kelas, model yang dapat diinterpretasikan, serta arsitektur yang efisien secara komputasi untuk penerapan di bidang pertanian.

***Kata Kunci**: Kematangan Pisang, Convolutional Neural Network (CNN), Deep Learning, Klasifikasi Citra, Systematic Literature Review.*

INTRODUCTION

The agricultural sector plays a vital role in the global economy, with bananas (*Musa spp.*) ranking among the fruits with the highest trading volume and consumption worldwide (Baglat et al., 2023; Junior & Astawa, 2024; Thanyacharoen et al., 2025; Yang et al., 2024). The ripeness level of bananas is the primary determinant of quality, economic value, and consumer preference (Baglat et al., 2023; Martínez-Mora et al., 2025; Yang et al., 2024). For diabetic consumers, selecting the appropriate level of ripeness is crucial, as biochemical changes during the ripening process increase simple sugars such as glucose, fructose, and sucrose, which directly affect the glycemic index of the fruit (Maharani et al., 2025; Upadhyay et al., 2023).

Traditionally, banana ripeness assessment in industrial and market settings has relied heavily on manual inspection by experts (Baglat et al., 2023; Dwi Cahyo et al., 2025; Hanifah & Hermawan, 2023; Qoyyum, 2026). This method suffers from significant limitations, including high subjectivity, labor-intensiveness, slow processing speeds, and vulnerability to human error (Baglat et al., 2023; Hanifah & Hermawan, 2023; Nafi'iyah et al., 2023; Soekarta et al., 2024). Consequently, automating ripeness detection systems through computer vision technology has become an urgent necessity to reduce post-harvest losses and improve industrial sorting efficiency (Dwi Cahyo et al., 2025; Halim & Fajar, 2023; Yang et al., 2024).

Recent advancements in artificial intelligence, particularly Deep Learning through the Convolutional Neural Network (CNN) architecture, have offered breakthroughs in non-destructive inspection (Dwi Cahyo et al., 2025; Ramadhan & Pramana, 2026; Yang et al., 2024), conventional machine learning methods that require complex manual feature extraction which often lack accuracy when dealing with complex fruit textures CNN models can automatically learn hierarchical features directly from image data (Dwi Cahyo et al., 2025; Hanifah & Hermawan, 2023; Nafi'iyah et al., 2023; Upadhyay et al., 2023). Various studies have explored diverse CNN architectures, ranging from the VGG (Visual Geometry Group) model, known for its deep feature extraction capabilities, to ResNet (Residual Networks), designed to mitigate vanishing gradient issues in deeper networks (Dewi et al., 2023; Dwi Cahyo et al., 2025; Halim & Fajar, 2023; Qoyyum, 2026; Thanyacharoen et al., 2025).

Despite this, challenges persist in implementing CNN models, such as variations in local cultivar characteristics, lighting conditions in field settings, and the need for models capable of detecting both individual fruit (fingers) and entire bunches (bunches)(Dwi Cahyo et al., 2025; Hanifah & Hermawan, 2023; Ramadhan & Pramana, 2026; Yang et al., 2024). While many studies have reported high accuracy, performance remains heavily dependent on dataset quality, data augmentation techniques, and precise hyperparameter selection (Dwi Cahyo et al., 2025; Junior & Astawa, 2024; Nafi'iyah et al., 2023; Thanyacharoen et al., 2025; Upadhyay et al., 2023).

This study presents a systematic literature review and comparative synthesis of CNN-based approaches for banana ripeness classification. Rather than focusing solely on reported classification accuracy, this review analyzes the methodological and practical factors that influence model performance, including CNN architecture, preprocessing, data augmentation, class imbalance, cultivar variability, evaluation metrics, model interpretability, and computational efficiency. Particular attention is given to the differences in datasets and experimental settings across studies, so that reported performance is interpreted as study-specific evidence rather than as a direct benchmark between architectures.

Based on the synthesis of 15 selected studies, this review develops a conceptual framework that connects multi-cultivar data acquisition, preprocessing and imbalance handling, model selection, multi-metric evaluation, explainable artificial intelligence, and lightweight or edge deployment. This integrated synthesis is intended to identify key research gaps and provide a more practical direction for developing robust and deployable banana ripeness classification systems.

MATERIALS AND METHODS

Research Questions (RQ)

1. RQ1: What are the most effective CNN architectures for banana ripeness classification in terms of accuracy and performance?
2. RQ2: How do data augmentation techniques influence the robustness of deep learning models in post-harvest sorting environments?

3. RQ3: What are the existing research gaps and future directions for the practical implementation of CNN-based systems in agriculture?

Systematic Literature Review Process (PRISMA Flow Diagram)

To ensure a transparent and reproducible review process, this study adopts the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. The literature selection process is structured as follows:

1. Identification: A total of 50 potential studies were identified from electronic databases including IEEE Xplore, ScienceDirect, and Google Scholar based on the predefined keywords covering the publication period of 2023–2026.
2. Screening: Out of the initial 50 records, 20 duplicates were removed, leaving 30 unique records for screening. Subsequently, 15 studies were excluded after an initial screening of titles and abstracts because they did not focus on CNN-based image classification or used non-image sensors.
3. Eligibility: The remaining 15 full-text articles were thoroughly assessed against strict inclusion criteria, such as peer-reviewed status and the provision of clear quantitative performance metrics.
4. Included: Ultimately, 15 relevant studies were selected for final synthesis and qualitative analysis in this review.

Table 1. Summary of Reported Accuracy by CNN Architecture

Architecture	Number of Studies	Average Accuracy (%)	Performance Trend
Standard CNN	3	92.40%	Baseline performance
VGG16/19	6	93.80%	Consistent performance
ResNet 50	6	97.20%	High reported performance

Source: (Research Results, 2026)

Standard CNN models were generally used as baseline approaches for banana ripeness classification and showed moderate to high reported accuracy across the reviewed studies. VGG16 and VGG19 demonstrated relatively consistent performance and were frequently adopted for image-based ripeness classification. ResNet50 also showed high reported accuracy across the studies in which it was applied. However, these accuracy values should be interpreted as descriptive evidence rather than as a direct benchmark among architectures, because the reviewed studies used different datasets, cultivar characteristics, preprocessing procedures, data augmentation strategies, and experimental configurations. Therefore, the observed differences in reported accuracy cannot be attributed solely to the CNN architecture.

Data Summary and Comparative Framework

This research employs a systematic literature review (SLR) design to evaluate and compare the application of Convolutional Neural Networks (CNN) in banana ripeness classification. The research design is structured to identify, select, and synthesize findings from existing scholarly articles, ensuring a transparent and reproducible review process. Data sources for this study were obtained from several reputable academic databases, including IEEE Xplore, ScienceDirect, and Google Scholar, covering literature published between 2023 and 2026. The search strategy utilized a combination of precise keywords such as "Banana Ripeness," "Convolutional Neural Network," "CNN," and "Image Classification" to ensure the comprehensiveness of the gathered literature.

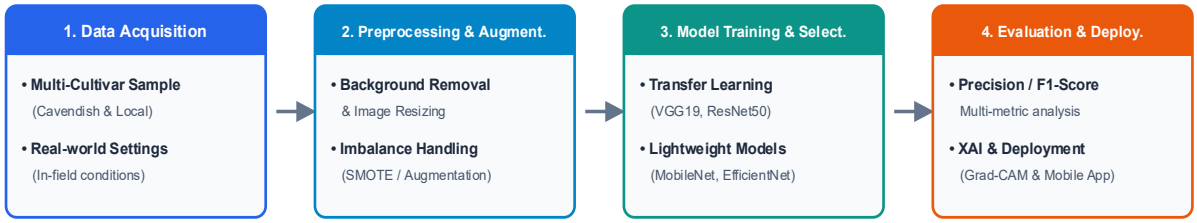
A total of 15 relevant studies were identified based on strict inclusion criteria, which required articles to be peer-reviewed, focus on CNN-based image classification for post-harvest banana ripeness, and provide clear quantitative performance metrics. Studies focusing on banana plant diseases or non-image sensors were excluded to maintain the thematic consistency of the analysis. Data collection techniques involved a multi-stage screening process: initial screening of titles and abstracts, followed by a full-text assessment to extract specific information, including model architecture, dataset size, training configurations, and classification accuracy.

Data analysis was carried out by synthesizing the extracted information into a comparative framework to identify performance trends among different CNN architectures. The analysis examined the influence of hyperparameters, such as the number of epochs, batch sizes, and dataset splits, on model performance. Furthermore, data augmentation techniques used in the selected studies were evaluated to determine their role in enhancing model robustness and mitigating overfitting. The reported classification results were synthesized using comparative performance tables, while confusion matrices were considered where they were available in the reviewed studies to provide additional information on class-specific prediction performance.

Evaluation Metrics and Selection Criteria

The performance of the identified models was synthesized using multiple evaluation metrics reported in the reviewed studies. Accuracy was used as the primary metric to summarize overall classification performance because it was the most consistently reported measure across the selected literature. Precision, Recall, and F1-score were also considered when reported to provide additional information about class-specific prediction reliability, particularly in studies involving imbalanced datasets.

Precision indicates the proportion of predicted samples in a class that are correctly classified, while Recall reflects the model's ability to identify samples belonging to that class. F1-score combines Precision and Recall and is useful for assessing the balance between these two measures. Because the reviewed studies differed in dataset composition, class definitions, and evaluation protocols, these metrics were interpreted descriptively rather than directly aggregated across all studies. The fundamental structure of the CNN model used for banana ripeness classification, which consists of alternating convolutional and pooling layers followed by a fully connected layer, is illustrated in Figure 1.



Source: (Research Results, 2026)

Figure 1. Proposed Conceptual Framework for Banana Ripeness Classification.

Figure 1 presents the conceptual framework developed from the synthesis of the reviewed studies. The framework integrates the main methodological and practical aspects identified in the literature, including multi-cultivar data acquisition, preprocessing and augmentation, CNN model selection, multi-metric evaluation, explainable artificial intelligence, and deployment considerations. These components represent the recurring factors identified throughout the reviewed literature and are integrated to illustrate the key considerations for developing a robust and practically deployable banana ripeness classification system.

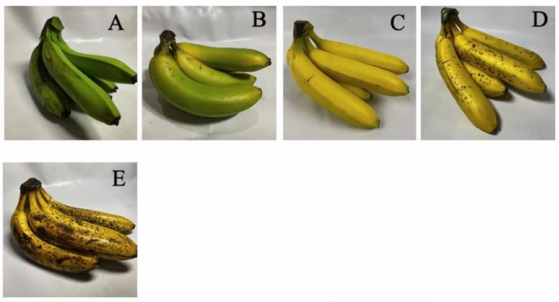
RESULTS AND DISCUSSION

The results of the comparative analysis of the 15 selected studies indicate a clear trend in the performance of Convolutional Neural Network (CNN) architectures for banana ripeness classification. Table 1 summarizes the key configurations and performance metrics extracted from the reviewed literature.

Table 2. Comparative Analysis of CNN Architectures for Banana Ripeness Classification

Study	Model Architecture	Dataset Size (Augmented)	Max Accuracy
Upadhyay et al. (2023)	Standard CNN	458	98.34%
Dewi et al. (2023)	VGG, ResNet18	11,647	88.62%
Nafi'iyah et al. (2023)	Custom CNN	952	97.95%
Thanyacharoen et al. (2025)	VGG16/19, Xception	4,000	93.70%
Yang et al. (2024)	ResNet, VGG	2,520	99.20%
Hanifah & Hermawan (2023)	Standard CNN	216	88.00%
Dwi Cahyo et al. (2025)	VGG19	5,616	98.00%
Halim & Fajar (2023)	VGG16	550	83.53%
Ramadhan & Pramana (2026)	CNN, VGG16, LSTM	600	98.39%
Hastungkoro et al. (2024)	Standard CNN	1,000	95.00%
Junior & Astawa (2024)	VGG16	1,000	97.50%
Soekarta et al. (2024)	EfficientNet	599	90.00%
Maharani et al. (2025)	ResNet50	2,375	99.00%
Martínez-Mora et al. (2025)	CNN, VGG	1,565	90.42%
Riskawati & Djatna (2026)	CNN, VGG16, EfficientNet	500	87.37%

Source: (Research Results, 2026)



Source: (Maharani et al., 2025)

Figure 2. Samples of Banana Ripeness Stages Representing Different Maturity Levels (From Unripe to Overripe).

a. Comparative Performance of CNN Architectures

The comparative synthesis of the 15 selected studies shows that VGG-family architectures, particularly VGG16 and VGG19, were frequently applied to banana ripeness classification and generally reported high classification performance. Their use was associated with their established convolutional feature extraction capabilities, although the resulting performance varied according to input resolution, dataset characteristics, preprocessing procedures, and experimental configurations. ResNet-based architectures, including ResNet50, also reported high accuracy in several studies, particularly when transfer learning and larger datasets were employed. The residual learning mechanism can facilitate the training of deeper networks; however, the observed performance differences across studies cannot be attributed solely to the network architecture because the reviewed studies differed in dataset composition, cultivar, number of classes, image acquisition conditions, and training strategies. Therefore, the reported accuracy values are interpreted as study-specific results and provide evidence of performance trends rather than a direct benchmark among CNN architectures.

b. The Role of Preprocessing and Computational Efficiency

Preprocessing represents an important component of CNN-based banana ripeness classification because the reviewed studies employed different image preparation procedures, including cropping, background removal, resizing, and other image transformations. These differences may affect the visual information presented to the CNN and consequently influence the reported classification performance. Therefore, preprocessing should be considered when interpreting performance differences across studies rather than treating architecture as the only determining factor. Data augmentation was also used in several studies to increase the variability of training images and reduce the risk of overfitting. In addition, the choice of training configuration, including optimizer and hyperparameters, varied across the reviewed studies, indicating that optimization strategies should be considered together with the selected architecture. From a deployment perspective, deeper CNN architectures may require greater computational resources and memory, creating challenges for real-time implementation on mobile or edge devices. This highlights the need to consider computational efficiency alongside classification performance when developing practical banana ripeness classification systems.

c. Challenges in Data Imbalance and Cultivar Variability

Class imbalance is identified as an important concern in banana ripeness classification, particularly when some ripeness categories contain fewer samples than others. An imbalanced distribution may cause models to favor majority classes and may reduce the reliability of predictions for minority categories such as overripe or rotten fruit. The reviewed studies also indicate that cultivar variability is an important consideration for model generalization. Differences in peel color, texture, and ripening characteristics among cultivars may affect the visual patterns learned by CNN models. Therefore, models developed using a limited cultivar or dataset may require additional validation before being applied to other banana varieties. To address these challenges, future studies should consider balanced data collection, data augmentation, and class-sensitive learning strategies such as Synthetic Minority Over-sampling Technique (SMOTE), Focal Loss, or Weighted Cross-Entropy, where appropriate.

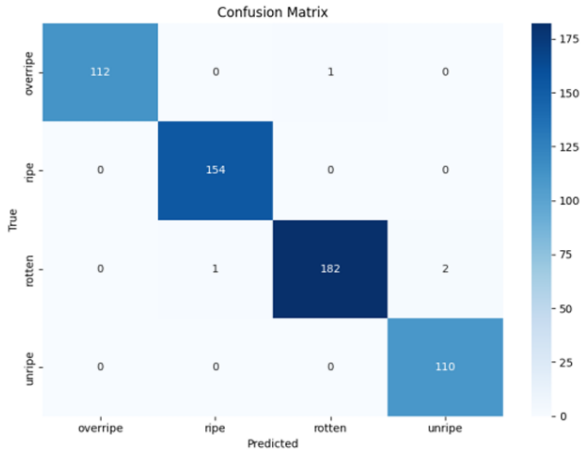
d. Evaluation Metrics and Model Reliability

Beyond overall accuracy, Precision, Recall, and F1-score provide additional perspectives for evaluating CNN-based banana ripeness classification. Accuracy can give a general indication of classification performance, but its interpretation may be limited when the distribution of ripeness classes is imbalanced. Precision is useful for examining the reliability of predictions for a particular class, while Recall indicates the ability of the model to identify samples belonging to that class. F1-score provides a combined measure of Precision and Recall and can therefore be useful when both types of classification error need to be considered. However, the reviewed studies did not report these metrics uniformly. Consequently, Precision, Recall, and F1-score are used in this

review to support the interpretation of reported model reliability where available, rather than as directly comparable measures across all 15 studies.

e. Practical Implementation and Future Directions

The reviewed literature indicates that practical deployment of CNN-based banana ripeness classification requires considerations beyond classification accuracy. Model interpretability is an important aspect for applications in which users need to understand the visual characteristics associated with a classification decision. Explainable AI (XAI) methods, such as Gradient-weighted Class Activation Mapping (Grad-CAM), can be considered to visualize image regions contributing to model predictions. In addition, computational efficiency should be considered when models are intended for mobile or edge-based applications. Lightweight architectures and model optimization techniques may reduce computational and memory requirements while maintaining acceptable classification performance. Broader image collection under different cultivars, backgrounds, and lighting conditions should also be considered to improve the generalizability of future systems in practical agricultural environments.



Source: (Dwi Cahyo et al., 2025)

Figure 3. Example of a Confusion Matrix Reported for a VGG19-Based Banana Ripeness Classification Model. (Dwi Cahyo et al., 2025).

Figure 3 presents an example of a confusion matrix reported by Dwi Cahyo et al. (2025) for the VGG19-based banana ripeness classification model. The matrix provides class-level information that cannot be fully represented by overall accuracy, particularly by showing the distribution of correct and incorrect predictions across ripeness categories. The reported results illustrate the usefulness of confusion matrices for examining class-specific classification behavior. However, this figure represents the results of an individual study and should not be interpreted as representative of all 15 studies included in this review. Differences in datasets, ripeness categories, image acquisition conditions, and experimental configurations may result in different confusion patterns across studies.

Handling Class Imbalance and Cultivar Variability. The synthesis of the reviewed studies indicates that dataset composition is an important factor in the reliability and generalizability of banana ripeness classification. Controlled image acquisition and limited representation of particular cultivars may not fully capture the visual variability encountered in practical agricultural environments. Class imbalance can further affect model learning when some ripeness categories are less represented than others. Future studies should therefore prioritize more balanced and diverse datasets and evaluate model performance across multiple cultivars and environmental conditions. Class-sensitive approaches, including SMOTE, Focal Loss, or Weighted Cross-Entropy, can also be investigated as potential strategies for improving minority-class recognition.

Model Interpretability through Explainable AI (XAI). Another important research gap concerns the interpretability of CNN-based classification models. CNN models may provide accurate predictions while offering limited information about the visual features contributing to those predictions. Explainable AI techniques, particularly Gradient-weighted Class Activation Mapping (Grad-CAM), can be used to visualize image regions that contribute to a model's decision, such as color changes, spots, or texture patterns on the banana peel. Such visualization may help researchers and end-users assess whether the model is responding to relevant visual characteristics rather than irrelevant image regions. Future studies should therefore consider integrating XAI into the classification pipeline and evaluating its usefulness across different cultivars and imaging conditions.

Edge Deployment and Lightweight Architectures. Computational requirements represent another consideration for practical deployment. Deeper architectures such as ResNet50 and VGG19 may require greater memory and computational resources than lightweight CNN architectures, which can limit their suitability for resource-constrained devices. Therefore, future research should investigate lightweight and optimized architectures, such as MobileNet and EfficientNet-Lite, together with techniques such as model quantization and compression. These approaches can be evaluated on edge or mobile platforms, including Raspberry Pi, NVIDIA Jetson, or Android devices, to assess inference latency, memory usage, energy consumption, and

classification performance under practical operating conditions.

Research Gaps and Proposed Future Research Framework

The synthesis of the reviewed studies indicates that several challenges remain in the development of practical CNN-based banana ripeness classification systems. First, dataset diversity remains limited in a number of studies, particularly with respect to cultivar variation, environmental conditions, and image acquisition settings. This limitation may restrict the generalizability of models when they are applied beyond the conditions represented in the training data. Second, practical deployment is not consistently addressed in the reviewed literature, particularly with regard to computational efficiency and implementation on resource-constrained mobile or edge devices. Third, model interpretability remains an important consideration because classification performance alone does not explain which visual characteristics contribute to a model's prediction. These gaps suggest that future research should address dataset diversity, deployment efficiency, and model interpretability alongside classification performance.

To address these gaps, the following directions are proposed for future research:

1. Hybrid Data Strategies: Combine field-collected images with appropriately generated synthetic data to increase dataset diversity across cultivars, ripeness categories, backgrounds, and lighting conditions. The effectiveness of synthetic data should be evaluated against real-world validation data rather than assumed to improve generalization.
2. Lightweight Architecture Optimization: Investigate mobile-oriented and lightweight architectures, such as MobileNet and EfficientNet-Lite, together with model optimization techniques including quantization or compression. Evaluation should include not only classification performance but also inference latency, memory requirements, and energy consumption on target devices.
3. Integration of Explainable AI (XAI): Incorporate techniques such as Grad-CAM to visualize image regions contributing to model predictions. Future studies should evaluate whether the highlighted regions correspond to relevant ripeness characteristics across cultivars and different imaging conditions.

The synthesis of the reviewed studies indicates that reported classification performance varies across CNN architectures and experimental settings. Transfer learning architectures such as VGG19 and ResNet50 frequently reported high accuracy in the studies included in this review; however, these results were obtained using different datasets, cultivars, image acquisition conditions, preprocessing procedures, and training configurations. Therefore, the observed performance should be interpreted as evidence of reported trends rather than as proof of universal architectural superiority. The findings also indicate that pretrained models can be useful when training data are limited, although their effectiveness remains dependent on the characteristics of the target dataset and the experimental design. This suggests that pretrained weights, which capture broad hierarchical visual features from large datasets like ImageNet, are highly advantageous when adapting models to specific agricultural domains where data availability might be limited. In contrast, studies utilizing custom or standard CNN architectures often show high variability in accuracy, ranging from 83% to 98%, depending heavily on the complexity of the dataset and the effectiveness of the preprocessing steps.

The reviewed studies indicate that data augmentation was commonly used to increase the variability of training images and to reduce the risk of overfitting. Several studies employing transformations such as rotation, flipping, or zooming reported favorable classification performance. However, the effect of augmentation cannot be isolated from other experimental factors because the reviewed studies differed in dataset size, preprocessing procedures, model architecture, and training configurations. Therefore, augmentation is considered an important supporting strategy, but its effectiveness should be evaluated within the specific dataset and experimental setting of each study.

When comparing these findings with similar research, the robustness of the VGG16 and ResNet architectures in fruit ripeness detection is consistent with trends in other agricultural sectors, such as palm oil and coffee classification. However, the variation in results specifically the lower accuracy in some VGG16 implementations compared to others suggests that input resolution and batch size configurations are just as critical as the choice of architecture. The results indicate that for real-time mobile applications, lighter models like EfficientNet or optimized shallow CNNs may provide a more practical balance between accuracy and computational speed, even if they sometimes sacrifice marginal gains in total accuracy compared to deep models like ResNet50. Consequently, future

implementations should prioritize a balance between architectural depth and computational constraints based on the specific end-user requirements, such as mobile integration vs. high-throughput industrial sorting.

CONCLUSION

This systematic literature review analyzed 15 selected studies on the application of Convolutional Neural Networks (CNN) for banana ripeness classification published between 2023 and 2026. The synthesis indicates that CNN-based models, including VGG16, VGG19, ResNet-based architectures, and EfficientNet, can achieve high reported classification performance across the reviewed studies. However, the reported results vary according to dataset composition, cultivar, image acquisition conditions, preprocessing procedures, data augmentation, and training configurations. Therefore, accuracy values from different studies should be interpreted as study-specific evidence rather than as a direct benchmark of architectural superiority.

Beyond classification accuracy, the review identifies dataset diversity, class imbalance, model interpretability, and computational efficiency as important considerations for practical deployment. Precision, Recall, and F1-score can provide additional information about model reliability when these metrics are reported. Based on the synthesis, a conceptual framework was developed to connect multi-cultivar data acquisition, preprocessing and imbalance handling, CNN model selection, multi-metric evaluation, Explainable AI, and lightweight deployment. Future research should focus on validating models using more diverse real-world datasets, investigating class-sensitive learning strategies, evaluating interpretability through techniques such as Grad-CAM, and assessing lightweight architectures on mobile or edge devices using measures such as accuracy, inference latency, memory usage, and energy consumption. These directions may help bridge the gap between reported experimental performance and practical agricultural deployment.

REFERENCE

Baglat, P., Hayat, A., Mendonça, F., Gupta, A., Mostafa, S. S., & Morgado-Dias, F. (2023). Non-Destructive Banana Ripeness Detection Using Shallow and Deep Learning: A Systematic Review. In *Sensors* (Vol. 23, Number 2). MDPI. https://doi.org/10.3390/s23020738

Dewi, C., Ueno, S., & Kato, K. (2023). Artificially Ripeness Identification of Indonesian Banana Cultivar Using Convolution Neural Network. *ACM International Conference Proceeding Series*, 107–111. https://doi.org/10.1145/3626641.3626668

Dwi Cahyo, G., Helilintar, R., Nur Farida, I., & Kunci-Akurasi, K. (2025). *Pengolahan Citra Kematangan Pisang Menggunakan Convolutional Neural Network VGG19 1* (Vol. 9). Online. https://doi.org/https://doi.org/10.29407/d079q671

Halim, J., & Fajar, A. N. (2023). *Klasifikasi Pisang Berbasis Algoritma VGG16 Melalui Metode CNN Deep Learning*.

Hanifah, A. I., & Hermawan, A. (2023). Klasifikasi Kematangan Pisang Menggunakan Metode Convolutional Neural Network. *Komputika : Jurnal Sistem Komputer*, *12*(2), 49–56. https://doi.org/10.34010/komputika.v12i2.9999

Junior, J., & Astawa, S. (2024). Pengklasifikasian Kualitas Pisang dengan Deep Learning CNN Arsitektur VGG16. *JNATIA*, *2*(3).

Maharani, T. I. A. P. Y., I Gusti, A. I., Gede, D. P., Christina, P. Y., & I Made, M. Y. (2025). *Classification of Cavendish Banana Maturity with CNN Method ResNet50 architecture*. https://doi.org/https://doi.org/10.56705/ijodas.v6i2.259

Martínez-Mora, O., Capuñay-Uceda, O., Caucha-Morales, L., Sánchez-Ancajima, R., Ramírez-Morales, I., Córdova-Márquez, S., & Cuenca-Mayorga, F. (2025). Artificial Vision-Based Dual CNN Classification of Banana Ripeness and Quality Attributes Using RGB Images. *Processes*, *13*(7). https://doi.org/10.3390/pr13071982

Nafi'iyah, N., Wardhani, R., & Prakasa, E. (2023). *Identification of Banana Ripeness using Convolutional Neural Network Approaches*. https://doi.org/10.1109/IC3INA60834.2023.10285749

Qoyyum, M. (2026). *Klasifikasi Kematangan Buah Pisang Kepok dengan Perbandingan Metode CNN VGG16 dan LSTM* (Vol. 5).

Ramadhan, Z., & Pramana, A. L. (2026). PENGENALAN TINGKAT KEMATANGAN BUAH PISANG BERBASIS CITRA DENGAN METODE CONVOLUTIONAL NEURAL NETWORK. In *Jurnal Mahasiswa Teknik Informatika* (Vol. 10, Number 3).

Soekarta, R., Fadli Hasa, M., & Ode, E. S. (2024). *Klasifikasi Kematangan Buah Pisang Secara Real-Time Menggunakan Convolutional Neural Network Berbasis Android*. *10*(01).

Thanyacharoen, S., Leone, G. R., Jirapattarasakul, S., Akkaralaertsest, T., Joijinda, N., & Yingthawornsuk, T. (2025). *Deep Learning Approaches to Banana Ripeness Detection*.

https://doi.org/10.1109/iEECON64081.2025.10987896

Upadhyay, A., Singh, S., & Kanojiya, S. (2023). Segregation of Ripe and Raw Bananas Using Convolutional Neural Network. *Procedia Computer Science*, *218*, 461–468. https://doi.org/10.1016/j.procs.2023.01.028

Yang, L., Cui, B., Wu, J., Xiao, X., Luo, Y., Peng, Q., & Zhang, Y. (2024). Automatic Detection of Banana Maturity—Application of Image Recognition in Agricultural Production. *Processes*, *12*(4). https://doi.org/10.3390/pr12040799