Jurnal Teknik Informatika (JUTIF) Vol. 6, No. 5, October 2025, Page. 3871-3885
P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5298

Data Augmentation-Driven Predictive Performance Refinement in Multi-Model Convolutional Neural Network for Cocoa Ripeness Prediction

Apriani*[1], I Nyoman Switrayana[2], Rifqi Hammad[3], Pahrul Irfan[4], Gede Yogi Pratama[5]

[1,2,5]Computer Science, Universitas Bumigora, Indonesia
[3]Software Engineering, Universitas Bumigora, Indonesia
[4]Informatics Engineering, Universitas Mataram, Indonesia

Email: [1]apriani@universitasbumigora.ac.id

Received : Sep 8, 2025; Revised : Sep 26, 2025; Accepted : Sep 29, 2025; Published : Oct 23, 2025

Abstract

Timely and accurate prediction of cocoa fruit ripeness is critical for optimizing harvest schedules, improving yield quality, and supporting post-harvest processing. Conventional visual inspection methods are prone to subjectivity and inconsistencies, especially when distinguishing among multiple ripeness levels based on fruit age. This study proposes a deep learning approach that leverages multi-model convolutional neural network transfer learning combined with image data augmentation to classify cocoa fruit into four maturity stages derived from fruit age. An augmented dataset of cocoa fruit images was used to fine-tune five well-established pre-trained models: MobileNetV2, Xception, ResNet50, DenseNet121, and DenseNet169. Data augmentation techniques were employed to increase variability and improve model generalization. Model evaluation was conducted using a standard 80:20 training-to-testing split to ensure sufficient data for learning while preserving a representative test set across all ripeness classes. The results demonstrate that DenseNet169 consistently outperformed other models, achieving the highest average accuracy of 85,05%, followed by DenseNet121 84,06%. Across all models, the use of data augmentation led to notable performance gains, highlighting its importance in enhancing predictive capability and reducing overfitting. The proposed framework shows promising potential for automating ripeness classification in agricultural contexts, offering a robust, scalable, and accurate solution for intelligent cocoa harvest management. This work contributes to the growing application of deep learning in precision agriculture, particularly in addressing fine-grained classification problems using limited but enriched visual data.

Keywords : *Cocoa, Convolutional Neural Network, Data Augmentation, Prediction, Ripeness.*

1. INTRODUCTION

Cocoa (Theobroma cacao L.) is a national strategic commodity that plays an important role in supporting the Indonesian economy and food industry [1], [2]. Data from the Central Statistics Agency (BPS) shows that in 2023, Indonesia's cocoa production reached more than 632,120 tons [3], making it the third largest cocoa producer in the world after Ivory Coast and Ghana [4]. However, the quality of the cocoa harvest is greatly influenced by the timing of the harvesting process [5]–[8]. Harvesting unripe or overripe fruit can reduce the quality of the beans, impacting the taste of processed products such as chocolate [9]–[11], and causes economic losses to farmers and downstream industry players [12]. So far, determining the level of ripeness of cocoa fruit still relies on manual visual methods by farmers [13], [14], which is subjective and prone to inconsistency [15], [16].

The challenge becomes even greater when it is necessary to classify multiple maturity levels that have subtle visual differences [17], [18]. Traditionally, ripeness classification relies on manual inspection based on pod color, texture, and size, procedures that are time-consuming, subjective, and prone to human error. Variability in visual appearance due to environmental factors further complicates the assessment process. This study addresses the problem of automating ripeness prediction using a

Vol. 6, No. 5, October 2025, Page. 3871-3885
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5298

data-driven approach grounded in deep learning, with the goal of minimizing inconsistencies and enhancing decision-making efficiency at the field level. Deep learning technology, particularly Convolutional Neural Network (CNN), has proven effective in solving various image classification problems in the agricultural sector [19]–[25], including for predicting fruit ripeness [25]–[31]. However, most of the previous approaches only use one to three types of CNN architectures [32]–[41] and face performance limitations when the amount and variety of data is limited [42]–[46].

On the other hand, variations in lighting, shooting angles, and differences in fruit age in field images are challenges in themselves that require a more complex and adaptive classification approach [47]–[51]. This study offers an innovative and novel approach, namely by comparing several CNN architectures through a transfer learning-based multi-model CNN strategy using five popular pre-trained models: MobileNetV2, Xception, ResNet50, DenseNet121, and DenseNet169. This multi-model approach allows for comparison and integration of the strengths of each model in extracting visual features in cocoa fruit images, resulting in a more accurate classification for four ripeness levels based on fruit age. The main novelty of this research is the extensive application of image augmentation and processing techniques to enrich the amount and variety of training data.

The augmentation techniques used include rotation, flipping, zooming, shifting, shearing, brightness adjustments, and filling empty pixels using interpolation. All aimed at creating a more realistic and diverse representation of the data. This strategy significantly improves the model's generalization ability and reduces the risk of overfitting, especially in the context of limited real-world data. The research formulates cocoa ripeness classification as a supervised image classification problem, where the objective is to predict one of four discrete maturity classes using RGB images as input. The challenge lies in learning discriminative visual features from relatively subtle color and texture changes across ripeness stages. Performance is measured using classification accuracy on test data, highlighting the potential of this approach to serve as a reliable tool for ripeness monitoring in agricultural applications. With this approach, the research is expected to contribute to the development of an intelligent system for cocoa ripeness classification that is not only accurate and robust, but also efficient and replicable for various other agricultural commodities. The application of this system has significant potential to strengthen precision agriculture practices and data-driven crop management, particularly in tropical cocoa-producing countries like Indonesia.

2. METHOD

The proposed methodology consists of a sequential pipeline designed to predict cocoa fruit ripeness through deep learning-based image classification. As illustrated in Figure 1, the process comprises five primary stages: data collection, data preparation, augmentation, modeling, and evaluation.

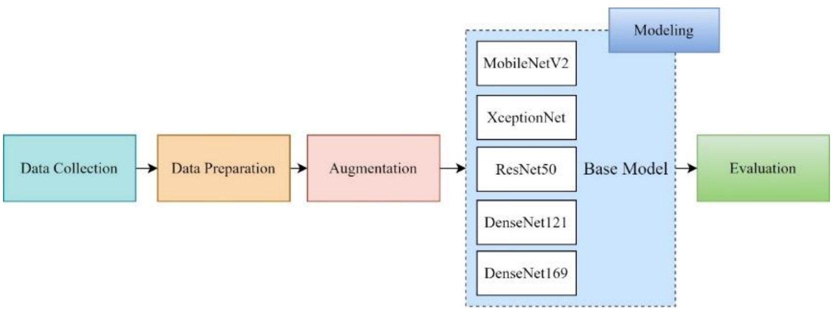


Figure 1. Process Pipeline of Cocoa Ripeness Classification

Figure 1 Cocoa Maturity Classification Process Flow starting from data collection, data preparation, augmentation, modeling and evaluation.

Vol. 6, No. 5, October 2025, Page. 3871-3885
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5298

2.1. Data Collection

The image dataset used in this study consists of cocoa fruit photographs captured under natural lighting conditions, with varied backgrounds and orientations. Each image is labeled based on the age of the fruit, which is grouped into four ripeness stages: stage 1 (0–2 months), stage 2 (2–4 months), stage 3 (4–6 months), and harvest stage (>6 months). These stages were determined based on field measurements and expert annotation using harvest-day metadata, ensuring biological relevance and consistency in maturity classification. To support generalization, image acquisition was conducted over different time periods and environmental conditions. The dataset aims to simulate real-world variability, allowing models to learn features that are robust against external noise and inconsistencies. All collected data were stored in high-resolution format before being processed in subsequent stages.

2.2. Data Preparation

The raw images were first passed through a preprocessing pipeline to ensure uniformity in size and structure. This step began with manual cropping of each image to isolate the cocoa fruit and remove irrelevant background content, improving the signal-to-noise ratio in feature extraction. After cropping, all images were resized to **128×128 pixels**, a resolution chosen to balance computational efficiency and preservation of spatial details relevant to fruit surface texture and shape. Normalization was also applied by scaling pixel values to a range of [0, 1], which helps stabilize training dynamics in convolutional neural networks. Labels were encoded into categorical classes corresponding to the four ripeness levels. This preparation ensured that the models received consistent input formats [52], enabling fair comparison across architectures during the modeling phase.

2.3. Data Augmentation

Table 1. Data Augmentation Technique’s Purpose and Effect

Augmentation	Value	Purpose and Effect
rotation_range	20 degrees	Randomly rotates the image within ±20° to simulate different orientations of fruit.
width_shift_range	0.15	Translates the image horizontally by up to 15% to mimic variation in positioning.
height_shift_range	0.15	Applies vertical translation up to 15% to simulate changes in camera perspective.
zoom_range	0.2	Zooms in/out by up to 20% to account for varying camera distances.
shear_range	0.15	Applies affine transformation to simulate slant or distortion in the image.
brightness_range	[0.8, 1.2]	Adjusts image brightness to replicate different lighting conditions in the field.
horizontal_flip	True	Flip the image horizontally to handle the left-right symmetry of cocoa fruit.
fill_mode	'nearest'	Fills missing pixels after transformation using nearest-neighbor interpolation.

To improve the robustness and generalization of the models, extensive image augmentation was performed. The augmentation process included several randomized transformations. Specifically, the integration of comprehensive image data augmentation techniques including geometric transformations such as rotation, width shift, height shift, zoom, shear, and horizontal flip, as well as photometric

 Vol. 6, No. 5, October 2025, Page. 3871-3885
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5298

transformation through brightness adjustment, was critical for simulating diverse real-world visual conditions encountered during cocoa fruit imaging. ata augmentation acts as a form of regularization by artificially expanding the training set's variability, preventing overfitting and improving model generalization, especially when working with limited datasets in computer vision tasks [53]. By increasing data diversity without requiring additional manual collection, the augmented dataset allows models to learn more invariant and discriminative features, particularly for subtle maturity differences between cocoa ripeness levels. The data augmentation techniques applied in this study are summarized in Table 1. Empirical studies demonstrate that combining multiple augmentation types outperforms single transformations and offers comparable generalization gains to more complex domain generalization approaches [54].

2.4. Modelling

The augmented dataset was used to fine-tune five different pre-trained convolutional neural networks: **MobileNetV2, XceptionNet, ResNet50, DenseNet121,** and **DenseNet169**. These models were selected for their proven performance on image classification benchmarks and varying architectural complexity. Notably, MobileNetV2 has been successfully applied for fruit-related vision tasks, achieving high accuracy [55]. Xception, which utilizes depthwise separable convolutions, has demonstrated strong accuracy and efficiency across multiple image classification benchmarks [52] and [56]. ResNet50 variants have shown high accuracy in distinguishing maturity stages in palm fruit datasets [57]. DenseNet121, characterized by dense connectivity facilitating gradient flow and feature reuse, has demonstrated superior accuracy (exceeding 94 %) in cocoa fruit disease classification tasks when compared directly with MobileNetV2 [58]. DenseNet169 was ranked among the top performing models in recent avocado ripeness classification studies when compared with other DenseNet variants [59]. Each model was initialized with weights pre-trained on ImageNet and adapted to the cocoa ripeness classification task through transfer learning. Model evaluation was conducted using a standard **80:20 train-test split**, where 80% of the data was used for training and 20% for testing. The class distribution was preserved in both sets to ensure a representative evaluation. The proposed model architecture is designed based on transfer learning using pre-trained convolutional neural networks (CNNs) combined with a classification head tailored for cocoa ripeness classification.

2.5. Modelling

The evaluation stage is conducted to measure the model's performance in classifying cocoa pod ripeness based on age categories. To assess the model's ability to distinguish between these four ripeness levels, various evaluation metrics—namely accuracy, precision, recall, and F1-score are employed, as described in [60], and detailed in Equations 1–4. In the context of multiclass classification, True Positives (TP) indicate the number of instances correctly classified into a particular class, True Negatives (TN) represent instances correctly identified as not belonging to that class, False Positives (FP) refer to instances wrongly assigned to that class, and False Negatives (FN) are cases where instances of a given class are misclassified as another. Given the imbalance in the distribution of samples across the ripeness classes, relying solely on accuracy may provide a misleading evaluation. Therefore, precision, recall, and F1-score are included to offer a more comprehensive assessment. These metrics help account for the uneven class representation by evaluating how well the model performs per class, especially in detecting minority classes that may be underrepresented in the dataset.

$$\text{Precision} = \frac{TP}{TP+FP} \tag{1}$$

$$\text{Recall} = \frac{TP}{TP+FN} \tag{2}$$

Vol. 6, No. 5, October 2025, Page. 3871-3885
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5298

$$\text{F1} - \text{Score} = \frac{2\,x\,Precision\,x\,Recall}{Precision+Recall} \quad (3)$$

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \quad (4)$$

3. RESULT AND DISCUSSION

3.1. Data Collection

This study utilizes the RipSetCocoaCNCH12 dataset, a curated collection of cocoa pod images systematically categorized into four ripeness stages: stage 1 (C1), stage 2 (C2), stage 3 (C3), and harvest stage (C4) based on the age of the fruit after pollination [61]. Each image in the dataset is labeled with its corresponding maturity level and was acquired under natural lighting conditions, representing realistic field scenarios. Figure 2 presents the number of samples available for each cocoa pod ripeness category. The dataset includes diverse pod appearances, covering variations in shape, surface texture, and pigmentation, which are critical visual indicators of ripeness progression. The dataset supports research in visual classification by offering well-annotated, high-resolution images, enabling effective training and evaluation of deep learning models. In this work, only the frontal view of each pod was used to maintain consistency in visual features across samples. Prior to model training, all images underwent preprocessing steps, including cropping to isolate the pod from the background and resizing to 128×128 pixels to match model input requirements. The diversity and quality of the dataset make it suitable for benchmarking convolutional neural network performance in ripeness classification tasks and provide a realistic foundation for practical implementation in precision agriculture.

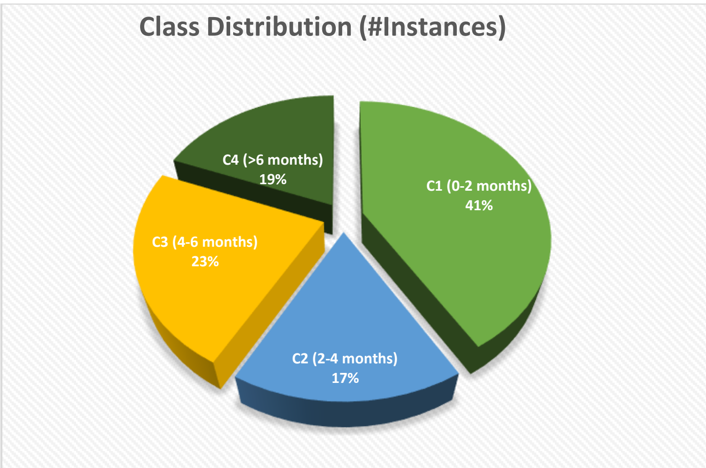


Figure 2. Number of Images in Each Ripeness Stage Category

Figure 2 shows the number of images in each category of chocolate maturity stage starting from C1 (0-2 months) 41%, C2 (2-4 months) 17%, C3 (4-6 months) 23% and C4 (>6 months) 19%.

3.2. Data Preparation

The data preparation process focused on isolating individual cocoa fruits from complex image backgrounds and standardizing their format for classification. The original dataset comprised RGB images accompanied by segmentation masks that highlighted the regions corresponding to cocoa fruits. These masks were utilized to locate the contours or boundaries of each fruit instance within an image. Once the contours were identified, a bounding box was generated around each fruit, and the region inside the box was extracted. As depicted in Figure 3, the segmentation mask is used to extract the

Vol. 6, No. 5, October 2025, Page. 3871-3885
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5298

relevant region from the original image. To enhance precision, only sufficiently large contours were retained to avoid including noise or irrelevant objects. The resulting cropped images contained only the fruit of interest with the background removed, ensuring that the model could focus solely on features pertinent to ripeness classification. An example of this can be seen in Figure 3.

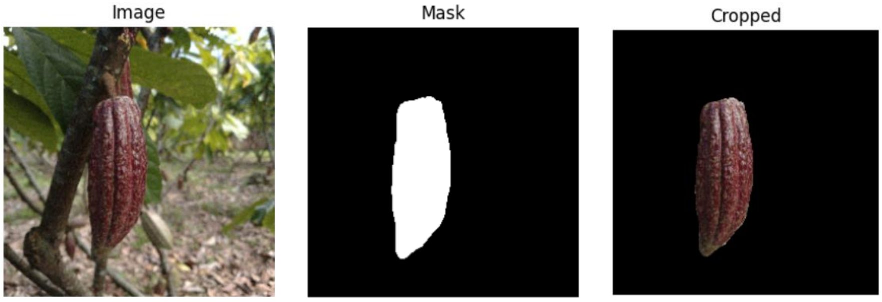


Figure 3. Example of A Cocoa Pod Image, Its Corresponding Segmentation Mask, and The Resulting Cropped Output

Figure 3 shows an example of a cocoa fruit image, its segmentation, and its cropping results. Each isolated fruit image was then converted to a standardized format by resizing it to a fixed resolution of 128 × 128 pixels. Additionally, pixel values were normalized to a [0,1] scale to facilitate stable neural network training [62]. The cropped images were saved with transparent backgrounds to eliminate any residual noise from the surrounding environment. This step not only improved the quality of the training data but also contributed to reducing overfitting and enhancing generalization performance during the learning phase. The combination of instance-level cropping, background removal, and image normalization provided a clean and consistent dataset for training deep learning models with improved accuracy and robustness. Figure 4 illustrates representative samples of cocoa fruit across four ripeness stages, categorized based on chronological age since fruit set.

Figure 4. Sample of Cocoa Ripeness Stages

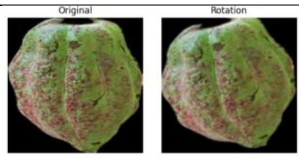
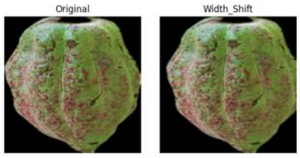
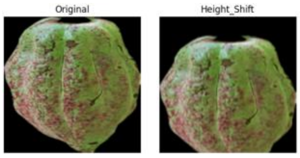
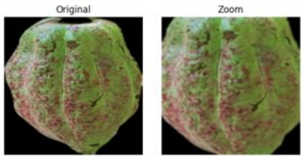
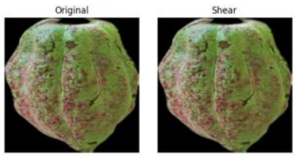
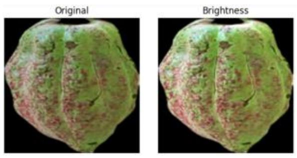
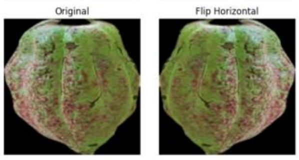
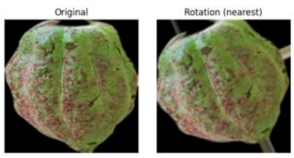
Figure 4 shows an example of cocoa fruit ripeness levels from stages 1 to 4. Stage 1 (0–2 months) is characterized by a predominantly green surface with minimal pigmentation, indicating the early developmental phase where physiological changes are minimal. In Stage 2 (2–4 months), fruits begin to show purplish discoloration or streaks, marking the onset of biochemical maturation processes such as chlorophyll degradation and anthocyanin accumulation. These visual cues provide early indicators of internal physiological changes relevant for harvest timing. As the fruit progresses to Stage 3 (4–6 months), the color transition intensifies toward a reddish-brown hue with more uniform distribution, reflecting the progression of sugar accumulation and pericarp softening. In the final Stage 4 (>6 months), cocoa pods exhibit a dominant yellow coloration, indicating full ripeness and optimal harvest readiness. These color and texture transformations are critical markers for classifying ripeness in image-based analysis systems. Leveraging these visual features through convolutional neural networks enables non-destructive, automated assessment of ripeness stages with high accuracy.

Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 6, No. 5, October 2025, Page. 3871-3885
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5298

3.3. Data Augmentation

To address the issue of limited dataset size and to improve the model's generalization capability, this study applied a comprehensive data augmentation strategy to the training images. Data augmentation artificially increases the diversity and volume of training data by applying a range of transformations that simulate real-world variabilities. This process is especially crucial in image classification tasks involving agricultural products, where external factors like lighting, orientation, and background may influence model performance. The visual outcomes of various data augmentation techniques are shown in Table 2.

Table 2. Results of Data Augmentation

No	Augmentation Method	Original Image Vs Augmented Image
1	rotation_range	
2	width_shift_range	
3	height_shift_range	
4	zoom_range	
5	shear_range	
6	brightness_range	
7	horizontal_flip	
8	fill_mode	

Each augmentation operation was applied randomly during training, ensuring that the model encountered a different transformed version of the image on each epoch. This stochasticity plays a key role in reducing overfitting by preventing the model from memorizing fixed patterns and instead forcing it to learn more robust, abstract features. The inclusion of spatial (rotation, shift, shear), photometric (brightness), and geometric (zoom, flip) transformations simulates real-world variability that might occur in uncontrolled agricultural environments. These augmentations also help the model become invariant to irrelevant factors such as pod alignment, illumination differences, and slight occlusions, which are common in field conditions. Ultimately, this augmentation strategy significantly contributes to increasing the effective size and richness of the training dataset, leading to improved classification performance, as reflected in the evaluation metrics reported in the experimental results.

3.4. Modelling

As illustrated in Figure 5, the pipeline begins with a set of input images, which can be either raw or augmented, depending on the experimental configuration. These images are fed into a base CNN model, which acts as a feature extractor.

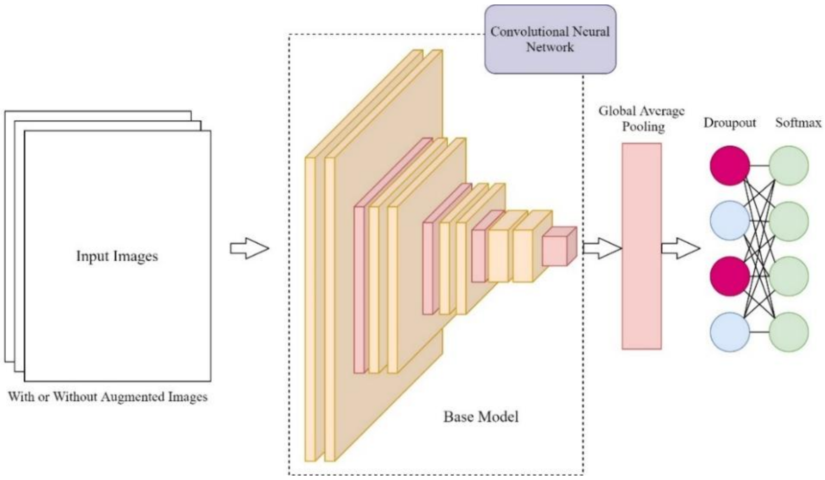


Figure 5. Model Architecture with Transfer Learning

The base models used in this study include MobileNetV2, XceptionNet, ResNet50, DenseNet121, and DenseNet169. Each of these models was pre-trained on the ImageNet dataset and fine-tuned for the cocoa dataset by replacing their original classification layers with a new custom head. The convolutional blocks of the base model act as feature extractors [63], transforming the input image into high-level feature maps that capture spatial patterns and important structures. These feature maps are then passed to a Global Average Pooling (GAP) layer, which compresses each map into a single value by computing the average across all spatial dimensions. This approach significantly reduces the number of parameters and helps prevent overfitting while maintaining spatial invariance. Following the GAP layer, a dropout layer is applied with a dropout rate (e.g., 0.3) to further mitigate overfitting by randomly deactivating a subset of neurons during training. The final layer is a dense softmax classifier with four output neurons, corresponding to the four maturity stages of cocoa fruits. This layer computes the probability distribution over the ripeness classes, and the class with the highest probability is selected as the predicted maturity level.

The model was trained for a maximum of 50 epochs using a batch size of 32. The Adam optimizer with a learning rate of 1e-4 was employed to ensure stable and efficient convergence during training. To prevent overfitting, early stopping was applied with a patience value of 10, allowing the training process to halt automatically if no improvement in validation loss was observed for 10 consecutive epochs. This modular architecture not only leverages the strong representational capabilities of pre-trained CNNs but also introduces lightweight adaptations through global average pooling and dropout, making it highly effective and computationally efficient for the target classification task.

Vol. 6, No. 5, October 2025, Page. 3871-3885
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5298

3.5. Evaluation

To assess the effectiveness of the proposed classification model in identifying cocoa ripeness stages, a series of performance metrics were evaluated. These include accuracy, precision, recall, and F1-score, which provide a comprehensive view of the model's predictive capabilities across all classes. Table 3 summarizes the evaluation metrics obtained from the test dataset.

Table 3. Model Performance Evaluation Results

No	Model	Performance Without Augmentation				Performance With Augmentation			
		Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score	Accuracy
1	MobileNetv2	81,07	81,35	81,09	81,35	82,30	82,28	82,26	82,28
2	XceptionNet	80,02	80,28	79,89	80,28	83,75	83,84	83,72	83,84
3	ResNet50	80,32	80,64	79,94	80,64	83,35	83,49	83,26	83,49
4	DenseNet121	82,04	82,28	82,02	82,28	84,05	84,06	84,04	84,06
5	DenseNet169	84,79	84,91	84,82	84,91	84,94	85,05	84,92	85,05

The performance evaluation results demonstrate that the proposed deep learning-based classification models exhibit significant improvements when augmented training data is applied, confirming the importance of data diversity in image-based agricultural tasks. Furthermore, the superior performance of transfer learning-based models highlights the benefits of leveraging pretrained convolutional neural networks for domain specific classification tasks. Transfer learning has proven particularly useful in agricultural contexts where acquiring large annotated datasets is often difficult. To better illustrate the comparative performance of the proposed CNN models in classifying cocoa pod ripeness, a graphical visualization of the evaluation metrics is presented in Figure 5 and Figure 6. These figures display the precision, recall, F1-score, and accuracy achieved by each model before and after applying data augmentation, respectively. The purpose of these visual comparisons is to highlight not only the relative strength of each architecture but also the positive impact of augmentation strategies on model generalization. The comparison results for each architecture without data augmentation can be seen in Figure 6 and the comparison result for each architecture with data augmentation can be seen in Figure 7.

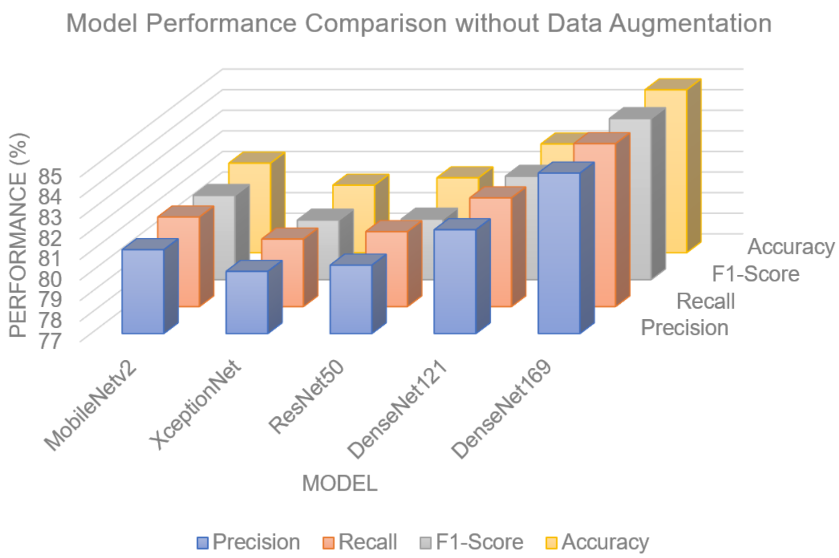


Figure 6. Model Performance Comparison Without Data Augmentation

Vol. 6, No. 5, October 2025, Page. 3871-3885
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5298

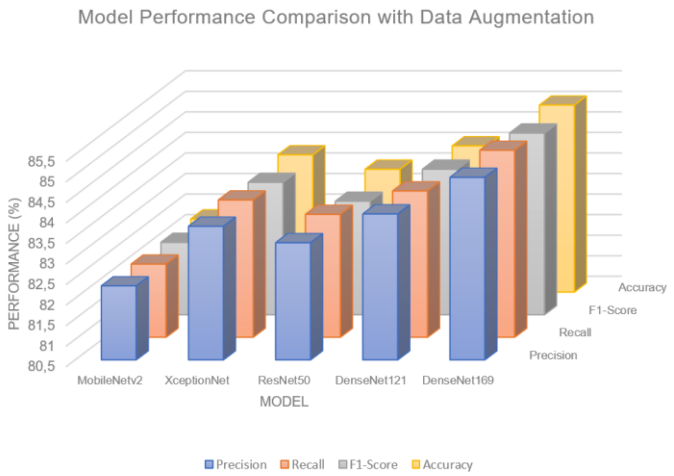


Figure 7. Model Performance Comparison with Data Augmentation

Figure 6 visualizes model performance *without data augmentation*. Here, DenseNet169 clearly outperforms the other models across all metrics, achieving the highest accuracy (84.91%) and F1-score (84.82%). This suggests that its densely connected structure is particularly effective in extracting fine-grained visual features from raw RGB images of cocoa pods. Other models, such as MobileNetV2 and ResNet50, show comparatively lower scores, indicating a limited ability to capture subtle ripeness cues in the absence of enriched training data. In contrast, Figure 7 illustrates the performance *with data augmentation applied*. All models demonstrate noticeable improvements, especially Xception, which experiences a sharp rise in F1-score from 79.89% to 83.72%. This underscores the sensitivity of certain lightweight or depthwise convolution-based models to the diversity and variability of input data. Interestingly, while DenseNet169 maintains its position as the top performing model, the performance gap between it and the other architectures narrows after augmentation, validating the hypothesis that appropriate augmentation can partially compensate for model complexity.

The figures also support the observation that data augmentation enhances not only accuracy but also model balance, as evident from the consistent rise in recall and F1 score across models. This indicates that the models are improving not only in identifying positive cases but also in reducing false negatives, which is an essential aspect for real world agricultural decision support systems where missing a ripe pod can lead to economic loss. Overall, the visualizations in Figure 5 and Figure 6 provide clear empirical support for the combined benefits of proper model selection and data augmentation. The strong baseline performance of DenseNet architectures, along with the performance improvements brought by augmentation techniques, reaffirms the effectiveness of a transfer learning based multi model approach for addressing complex classification challenges in agricultural domains. A comparison of the accuracy of the model performance with and without data augmentation can be seen in Figure 8.

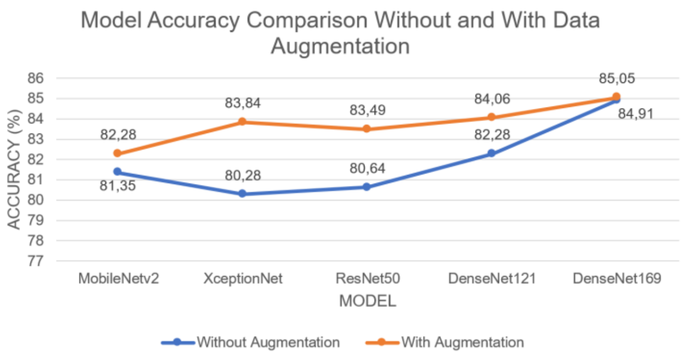


Figure 8. Model Performance Comparison with and Without Data Augmentation

Vol. 6, No. 5, October 2025, Page. 3871-3885
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5298

To further illustrate the impact of data augmentation on model performance, Figure 8 presents a comparative line graph of classification accuracy across five CNN architectures, evaluated with and without augmentation techniques. This visualization provides a clear empirical overview of how each model responds to augmented training data in the context of cocoa ripeness classification. As shown in Figure 8, all models experienced improvements in accuracy when trained with augmented data. The most notable increase was observed in XceptionNet, which improved from 80.28% to 83.84%, reflecting an absolute gain of 4.43%. This substantial enhancement confirms that models utilizing depthwise separable convolutions benefit significantly from expanded and varied training distributions. Similarly, ResNet50 and DenseNet121 improved by 3.53% and 2.16% respectively, while MobileNetV2 followed with a 1.14% gain. This highlights their capacity to generalize more effectively when exposed to a richer set of image conditions, including variations in brightness, orientation, scale, and perspective that were introduced during augmentation.

The DenseNet169 model remained the top performing architecture in both training scenarios. Although the performance gain appears marginal at 0.16%, it is important to note that the model had already achieved a high baseline accuracy without augmentation. This suggests a saturation point where further increases in data diversity yield diminishing returns. It also indicates that dense connectivity architectures are inherently robust, even when trained on limited samples. Overall, Figure 7 visually reinforces the argument that data augmentation is not only beneficial but essential for improving model accuracy, particularly for architectures that are more dependent on diverse and enriched training data. These results also underscore the potential of combining augmentation strategies with transfer learning to develop scalable, robust image classification systems for precision agriculture applications.

3.6. Discussion

Among the tested architectures, DenseNet169 achieved the highest classification performance with an accuracy of 85.05% and 84.91%, both with and without data augmentation. This outcome is consistent with prior studies indicating that densely connected networks enable more efficient feature propagation and reuse, making them highly effective for capturing subtle inter-class differences in agricultural imagery [64]. Meanwhile, Xception demonstrated the most notable performance gain post-augmentation. This suggests that depthwise separable convolutions used in Xception may benefit more from augmented data variability, a finding that aligns with the work of [52]. The consistent performance improvements observed across all models validate the effectiveness of the applied augmentation strategies, including rotation, width and height shifts, zooming, shearing, brightness adjustments, horizontal flipping, and pixel filling, as these collectively simulate diverse real-world field conditions commonly encountered in agricultural imagery. These results support previous findings by, which demonstrate that augmentation techniques substantially enhance the generalizability of deep learning models, especially in domains with limited or imbalanced datasets such as fruit quality assessment [65].

The success of DenseNet169 and DenseNet121 in this study also reinforces the idea that deeper and more complex models, when adequately supported by data augmentation, are better suited for high resolution visual discrimination tasks such as ripeness detection, which typically involve subtle texture and color variations. From an application perspective, the findings suggest that DenseNet169 is the most suitable model for deployment in controlled environments with sufficient computational resources. On the other hand, MobileNetV2 can serve as a lightweight alternative for edge computing scenarios, including mobile based field assessment tools. This differentiation supports the development of tiered systems that aim to balance classification performance with computational efficiency depending on field conditions and available infrastructure. Nevertheless, even though the achieved accuracy exceeds 85 percent, further improvements are still necessary to ensure reliable deployment across diverse

Vol. 6, No. 5, October 2025, Page. 3871-3885
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5298

agroecological zones. Variability in lighting, occlusion from leaves, and differences among cultivar phenotypes remain potential sources of classification errors that need to be addressed in future research.

4. CONCLUSION

This study presents a deep learning-based approach for classifying cocoa fruit maturity levels using transfer learning with multiple convolutional neural network architectures. The method effectively extracts discriminative visual features relevant to four ripe stages derived from the fruit's chronological age. The integration of comprehensive image data augmentation techniques, including geometric and photometric transformations, significantly enhanced the models' generalization capabilities and performance consistency across varying visual conditions. Among the evaluated models, DenseNet169 consistently achieved the highest classification performance, with an accuracy reaching 85.05%, demonstrating its superior feature extraction and representational capacity when combined with data augmentation. The empirical results confirm that applying augmentation strategies not only enriches the training set diversity but also leads to measurable improvements in classification robustness across all tested architectures. The findings underscore the effectiveness of a multi-model transfer learning framework enhanced by image augmentation for supporting automated, non-destructive assessment of cocoa ripeness. This approach offers practical implications for agricultural technology, particularly in optimizing harvest timing and improving post-harvest processing decisions. Future work may explore combining multimodal data such as spectral or textural cues and integrating lightweight architectures for real-time field deployment.

REFERENCES

[1] J. Neilson, A. Dwiartama, N. Fold, and D. Permadi, "Resource-based industrial policy in an era of global production networks: Strategic coupling in the Indonesian cocoa sector," *World Dev.*, vol. 135, no. 105045, pp. 1–12, 2020, doi: https://doi.org/10.1016/j.worlddev.2020.105045.

[2] I. Idawati *et al.*, "Cocoa farmers' characteristics on climate variability and its effects on climate change adaptation strategy," *Glob. J. Environ. Sci. Manag.*, vol. 10, no. 1, 2024, doi: 10.22034/gjesm.2024.01.21.

[3] E. Marsoro *et al.*, *Statistik Kakao Indonesia 2023*, 8th ed. Jakarta: Badan Pusat Statistik Indonesia, 2024.

[4] J. E. Kongor, M. Owusu, and C. Oduro-Yeboah, "Cocoa production in the 2020s: challenges and solutions," *CABI Agric. Bioscinece*, vol. 5, no. 102, 2024, doi: https://doi.org/10.1186/s43170-024-00310-6.

[5] A. M. Calvo, B. L. Botina, M. C. García, W. A. Cardona, A. C. Montenegro, and J. Criollo, "Dynamics of cocoa fermentation and its effect on quality," *Sci. Rep.*, vol. 11, no. 1, 2021, doi: 10.1038/s41598-021-95703-2.

[6] A. P. Romero Vergel, A. V. Camargo Rodriguez, O. D. Ramirez, P. A. Arenas Velilla, and A. M. Gallego, "A Crop Modelling Strategy to Improve Cacao Quality and Productivity," *Plants*, vol. 11, no. 2, 2022, doi: 10.3390/plants11020157.

[7] R. Niikoi Kotey *et al.*, "Effects of Fermentation Periods and Drying Methods on Postharvest Quality of Cocoa (Theobroma Cacao) Beans in Ghana," *J. Food Qual.*, vol. 2022, 2022, doi: 10.1155/2022/7871543.

[8] N. N. Suh and E. L. Molua, "Cocoa production under climate variability and farm management challenges: Some farmers' perspective," *J. Agric. Food Res.*, vol. 8, 2022, doi: 10.1016/j.jafr.2022.100282.

[9] M. Santander *et al.*, "Unravelling Cocoa Drying Technology: A Comprehensive Review of the Influence on Flavor Formation and QualityNo Title," *Foods*, vol. 14, no. 5, 2025, doi: https://doi.org/10.3390/foods14050721.

[10] E. Subroto, M. Djali, R. Indiarto, E. Lembong, and N. Baiti, "Microbiological Activity Affects Post-Harvest Quality of Cocoa (Theobroma cacao L.) Beans," *Horticulturae*, vol. 9, no. 7, 2023, doi: https://doi.org/10.3390/horticulturae9070805.

Vol. 6, No. 5, October 2025, Page. 3871-3885
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5298

[11] L. Goya, J. E. Kongor, and S. de Pascual-Teresa, "From Cocoa to Chocolate: Effect of Processing on Flavanols and Methylxanthines and Their Mechanisms of Action," *Int. J. Mol. Sci.*, vol. 23, no. 22, 2022, doi: https://doi.org/10.3390/ijms232214365.

[12] J. E. Kongor and D. Rahadian Aji Muhammad, "Processing of Cocoa and Development of Chocolate Beverages," 2023.

[13] R. Essah, D. Anand, and S. Singh, "An intelligent cocoa quality testing framework based on deep learning techniques," *Meas. Sensors*, vol. 24, 2022, doi: 10.1016/j.measen.2022.100466.

[14] K. J. Ayikpa, D. Mamadou, P. Gouton, and K. J. Adou, "Classification of Cocoa Pod Maturity Using Similarity Tools on an Image Database: Comparison of Feature Extractors and Color Spaces," *Data*, vol. 8, no. 6, 2023, doi: 10.3390/data8060099.

[15] J. Y. Goh, Y. M. Yunos, and M. S. M. Ali, "Fresh Fruit Bunch Ripeness Classification Methods: A Review," *Food Bioprocess Technol.*, vol. 18, pp. 183–206, 2024, doi: https://doi.org/10.1007/s11947-024-03483-0.

[16] A. Julca-Otiniano *et al.*, "New Races of Hemileia vastatrix Detected in Peruvian Coffee Fields," *Agronomy*, vol. 14, no. 8, p. 1811, 2024, doi: https://doi.org/10.3390/agronomy14081811.

[17] M. Rizzo, M. Marcuzzo, A. Zangari, A. Gasparetto, and A. Albarelli, "Fruit ripeness classification: A survey," *Artificial Intelligence in Agriculture*, vol. 7. 2023, doi: 10.1016/j.aiia.2023.02.004.

[18] W. Chen *et al.*, "MLP-based multimodal tomato detection in complex scenarios: Insights from task-specific analysis of feature fusion architectures," *Agriculture*, vol. 221, 2024, doi: https://doi.org/10.1016/j.compag.2024.108951.

[19] J. Lu, L. Tan, and H. Jiang, "Review on convolutional neural network (CNN) applied to plant leaf disease classification," *Agriculture (Switzerland)*, vol. 11, no. 8. 2021, doi: 10.3390/agriculture11080707.

[20] N. Mamat, M. F. Othman, R. Abdoulghafor, S. B. Belhaouari, N. Mamat, and S. F. Mohd Hussein, "Advanced Technology in Agriculture Industry by Implementing Image Annotation Technique and Deep Learning Approach: A Review," *Agriculture (Switzerland)*, vol. 12, no. 7. 2022, doi: 10.3390/agriculture12071033.

[21] M. Altalak, M. A. Uddin, A. Alajmi, and A. Rizg, "Smart Agriculture Applications Using Deep Learning Technologies: A Survey," *Appl. Sci.*, vol. 12, no. 12, 2022, doi: 10.3390/app12125919.

[22] V. G. Dhanya *et al.*, "Deep learning based computer vision approaches for smart agricultural applications," *Artificial Intelligence in Agriculture*, vol. 6. 2022, doi: 10.1016/j.aiia.2022.09.007.

[23] A. Bouguettaya, H. Zarzour, A. Kechida, and A. M. Taberkit, "Deep learning techniques to classify agricultural crops through UAV imagery: a review," *Neural Computing and Applications*, vol. 34, no. 12. 2022, doi: 10.1007/s00521-022-07104-9.

[24] Y. Yuan, L. Chen, H. Wu, and L. Li, "Advanced agricultural disease image recognition technologies: A review," *Information Processing in Agriculture*, vol. 9, no. 1. 2022, doi: 10.1016/j.inpa.2021.01.003.

[25] A. Khan, A. D. Vibhute, S. Mali, and C. H. Patil, "A systematic review on hyperspectral imaging technology with a machine and deep learning methodology for agricultural applications," *Ecological Informatics*, vol. 69. 2022, doi: 10.1016/j.ecoinf.2022.101678.

[26] S. Geerthik, G. A. Senthil, K. J. Oliviya, and R. Keerthana, "A System and Method for Fruit Ripeness Prediction Using Transfer Learning and CNN," 2024, doi: https://doi.org/10.1109/IC3IoT60841.2024.10550209.

[27] N. Aherwadi, U. Mittal, J. Singla, N. Z. Jhanjhi, A. Yassine, and M. S. Hossain, "Prediction of Fruit Maturity, Quality, and Its Life Using Deep Learning Algorithms," *Electron.*, vol. 11, no. 24, 2022, doi: 10.3390/electronics11244100.

[28] C. Prasad, S. Kumar, and S. D. Rathod, "Intelligent System for Predicting Orange Fruit Ripeness by Integrating Climatic and Imaging Data using CNN Model for High-Accuracy Classification," 2024, doi: https://doi.org/10.1109/ICAIQSA64000.2024.10882459.

[29] J. K. Basak *et al.*, "Prediction of physicochemical properties of strawberry fruits using convolutional neural network-regression models," 2025. doi: https://doi.org/10.1007/s13580-025-00717-8.

[30] D. Minagawa and J. Kim, "Prediction of Harvest Time of Tomato Using Mask R-CNN,"

AgriEngineering, vol. 4, no. 2, 2022, doi: 10.3390/agriengineering4020024.

[31] A. G. Said and B. Joshi, "SmartRipen: LSTM-GRU feature selection& XGBoost-CNN for fruit ripeness detection," *Food Phys.*, vol. 2, 2025, doi: https://doi.org/10.1016/j.foodp.2025.100053.

[32] S. Han, J. Liu, G. Zhou, Y. Jin, M. Zhang, and S. Xu, "InceptionV3-LSTM: A Deep Learning Net for the Intelligent Prediction of Rapeseed Harvest Time," *Agronomy*, vol. 12, no. 12, 2022, doi: 10.3390/agronomy12123046.

[33] Y. Yu, J. Huang, L. Wang, and S. Liang, "A 1D-inception-ResNet based global detection model for thin-skinned multifruit spectral quantitative analysis," *Food Control*, vol. 167, no. 110823, 2025, doi: https://doi.org/10.1016/j.foodcont.2024.110823.

[34] N. Begum and M. K. Hazarika, "Maturity detection of tomatoes using transfer learning," *Meas. Food*, vol. 7, 2022, doi: 10.1016/j.meafoo.2022.100038.

[35] Harsh Mundhada, Sanskriti Sood, Saitejaswi Sanagavarapu, Rina Damdoo, and Kanak Kalyani, "Fruit Detection and Three-Stage Maturity Grading Using CNN," *Int. J. Next-Generation Comput.*, 2023, doi: 10.47164/ijngc.v14i1.1099.

[36] H. Sun, S. Zhang, R. Ren, and L. Su, "Maturity Classification of 'Hupingzao' Jujubes with an Imbalanced Dataset Based on Improved MobileNet V2," *Agric.*, vol. 12, no. 9, 2022, doi: 10.3390/agriculture12091305.

[37] A. A. Abdelhamid, A. A. Alhussan, A.-S. T. Qenawy, A. M. Osman, A. M. Elshewey, and M. Eed, "Potato Harvesting Prediction Using an Improved ResNet-59 Model," *Potato Res*, vol. 68, pp. 1049–1068, 2025, doi: https://doi.org/10.1007/s11540-024-09773-6.

[38] Z. Wang *et al.*, "An improved Faster R-CNN model for multi-object tomato maturity detection in complex scenarios," *Ecol. Inform.*, vol. 72, 2022, doi: 10.1016/j.ecoinf.2022.101886.

[39] K. Sabanci, "Benchmarking of CNN Models and MobileNet-BiLSTM Approach to Classification of Tomato Seed Cultivars," *Sustain.*, vol. 15, no. 5, 2023, doi: 10.3390/su15054443.

[40] E. Gautama, T. K. A. Rahman, and L. Kamelia, "Measurement Of Optimizer Performance On The EfficientNet Architecture In Convolutional Neural Network For Classification Of Matoa Maturity Levels," 2024, doi: https://doi.org/10.1109/ICWT62080.2024.10674685.

[41] P. Nahak, D. K. Pratihar, and A. K. DebView, "Tomato maturity stage prediction based on vision transformer and deep convolution neural networks," *Int. J. Hybrid Intell. Syst.*, vol. 21, no. 1, 2025, doi: https://doi.org/10.3233/HIS-240021.

[42] C. Wang *et al.*, "Application of Convolutional Neural Network-Based Detection Methods in Fresh Fruit Production: A Comprehensive Review," *Frontiers in Plant Science*, vol. 13. 2022, doi: 10.3389/fpls.2022.868745.

[43] F. Xiao, H. Wang, Y. Xu, and R. Zhang, "Fruit Detection and Recognition Based on Deep Learning for Automatic Harvesting: An Overview and Review," *Agronomy*, vol. 13, no. 6. 2023, doi: 10.3390/agronomy13061625.

[44] S. Espinoza, C. Aguilera, L. Rojas, and P. G. Campos, "Analysis of Fruit Images With Deep Learning: A Systematic Literature Review and Future Directions," *IEEE Access*, vol. 12, pp. 3837–3859, 2023, doi: https://doi.org/10.1109/ACCESS.2023.3345789.

[45] C. C. Ukwuoma, Q. Zhiguang, M. B. Bin Heyat, L. Ali, Z. Almaspoor, and H. N. Monday, "Recent Advancements in Fruit Detection and Classification Using Deep Learning Techniques," *Math. Probl. Eng.*, vol. 2022, 2022, doi: 10.1155/2022/9210947.

[46] N. Ismail and O. A. Malik, "Real-time visual inspection system for grading fruits using computer vision and deep learning techniques," *Inf. Process. Agric.*, vol. 9, no. 1, 2022, doi: 10.1016/j.inpa.2021.01.005.

[47] Y. Tang *et al.*, "Optimization strategies of fruit detection to overcome the challenge of unstructured background in field orchard environment: a review," *Precision Agriculture*, vol. 24, no. 4. 2023, doi: 10.1007/s11119-023-10009-9.

[48] X. Liu, N. Li, Y. Huang, X. Lin, and Z. Ren, "A comprehensive review on acquisition of phenotypic information of Prunoideae fruits: Image technology," *Frontiers in Plant Science*, vol. 13. 2023, doi: 10.3389/fpls.2022.1084847.

[49] C. Neupane, M. Pereira, A. Koirala, and K. B. Walsh, "Fruit Sizing in Orchard: A Review from Caliper to Machine Vision with Deep Learning," *Sensors*, vol. 23, no. 8. 2023, doi:

Vol. 6, No. 5, October 2025, Page. 3871-3885
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.5.5298

10.3390/s23083868.
[50] H. Naito, K. Shimomoto, T. Fukatsu, F. Hosoi, and T. Ota, “Interoperability Analysis of Tomato Fruit Detection Models for Images Taken at Different Facilities, Cultivation Methods, and Times of the Day,” *AgriEngineering*, vol. 6, no. 2, pp. 1827–1846, 2024, doi: https://doi.org/10.3390/agriengineering6020106.
[51] A. M. Hayajneh, S. Batayneh, E. Alzoubi, and M. Alwedyan, “TinyML Olive Fruit Variety Classification by Means of Convolutional Neural Networks on IoT Edge Devices,” *AgriEngineering*, vol. 5, no. 4, 2023, doi: 10.3390/agriengineering5040139.
[52] I. N. Switrayana and M. Azwar, “Optimizing Scalability in Spice Identification through Transfer Learning with Convolutional Neural Networks,” vol. 11, no. 1, pp. 73–84, 2025, doi: 10.24014/coreit.v11i1.35453.
[53] T. Kumar, R. Brennan, A. Mileo, and M. Bendechache, “Image Data Augmentation Approaches: A Comprehensive Survey and Future Directions,” *IEEE Access*, vol. 12, pp. 187536–187571, 2024, doi: 10.1109/ACCESS.2024.3470122.
[54] M. Schwonberg, F. El Bouazati, N. M. Schmidt, and H. Gottschalk, “Augmentation-based Domain Generalization for Semantic Segmentation,” *IEEE Intell. Veh. Symp. Proc.*, vol. 2023-June, 2023, doi: 10.1109/IV55152.2023.10186752.
[55] T. B. Shahi, C. Sitaula, A. Neupane, and W. Guo, “Fruit classification using attention-based MobileNetV2 for industrial applications,” *PLoS One*, vol. 17, no. 2 February, pp. 1–21, 2022, doi: 10.1371/journal.pone.0264586.
[56] F. Salim, F. Saeed, S. Basurra, S. N. Qasem, and T. Al-Hadhrami, “DenseNet-201 and Xception Pre-Trained Deep Learning Models for Fruit Recognition,” *Electron.*, vol. 12, no. 14, 2023, doi: 10.3390/electronics12143132.
[57] M. Han and C. Yi, “Deep Convolutional Neural Networks for Palm Fruit Maturity Classification,” pp. 1–9, 2025, [Online]. Available: http://arxiv.org/abs/2502.20223.
[58] K. R. Ariawan, A. A. G. Ekayana, I. P. Y. Indrawan, K. R. Winatha, and I. N. A. F. Setiawan, “Performance Comparasion of DenseNet-121 and MobileNetV2 for Cacao Fruit Disease Image Classification,” *Indones. J. Data Sci.*, vol. 6, no. 1, pp. 30–38, 2025, doi: 10.56705/ijodas.v6i1.233.
[59] S. Nuanmeesri, “Enhanced hybrid attention deep learning for avocado ripeness classification on resource constrained devices,” *Sci. Rep.*, vol. 15, no. 1, pp. 1–15, 2025, doi: 10.1038/s41598-025-87173-7.
[60] I. N. Switrayana, S. Hadi, and N. Sulistianingsih, “A Robust Gender Recognition System using Convolutional Neural Network on Indonesian Speaker,” vol. 13, pp. 1008–1021, 2024.
[61] J. F. Restrepo-Arias, M. I. Salinas-Agudelo, M. I. Hernandez-Pérez, A. Marulanda-Tobón, and M. C. Giraldo-Carvajal, “RipSetCocoaCNCH12: Labeled Dataset for Ripeness Stage Detection, Semantic and Instance Segmentation of Cocoa Pods,” *Data*, vol. 8, no. 6, 2023, doi: 10.3390/data8060112.
[62] X. Pei *et al.*, “Robustness of machine learning to color, size change, normalization, and image enhancement on micrograph datasets with large sample differences,” *Mater. Des.*, vol. 232, p. 112086, 2023, doi: 10.1016/j.matdes.2023.112086.
[63] I. N. Switrayana and N. U. Maulidevi, “Collaborative Convolutional Autoencoder for Scientific Article Recommendation,” *Proc. - 2022 9th Int. Conf. Inf. Technol. Comput. Electr. Eng. ICITACEE 2022*, pp. 96–101, 2022, doi: 10.1109/ICITACEE55701.2022.9924130.
[64] K. Kayaalp, “A deep ensemble learning method for cherry classification,” *Eur. Food Res. Technol.*, vol. 250, no. 5, pp. 1513–1528, 2024, doi: 10.1007/s00217-024-04490-3.
[65] P. T. Huong, L. T. Hien, N. M. Son, H. C. Tuan, and T. Q. Nguyen, “Enhancing deep convolutional neural network models for orange quality classification using MobileNetV2 and data augmentation techniques,” *J. Algorithms Comput. Technol.*, vol. 19, 2025, doi: 10.1177/17483026241309070.