Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v10i3.16055 p-ISSN : 2541-044X

CLASSIFICATION OF COFFEE FRUIT DRYING USING VGG16

Annisa Diyan Novitasari[1),]Yufis Azhar[2)]
[1, 2)]Faculty of Engineering, Informatics Study Program, Universitas Muhammadiyah Malang, Malang, Indonesia
[1)]annisadiyannumm@webmail.umm.ac.id, [2)*] yufis@umm.ac.id

Submitted :April 6, 2026 | **Accepted** : April 27, 2026 | **Published** : July 5, 2026

Abstract: The drying process is a crucial stage in coffee post-harvest handling that directly affects the final product quality, especially in the specialty coffee segment. Assessment of the coffee fruit drying level in the field is still largely carried out visually and subjectively, which can potentially lead to inconsistent quality. This study aims to develop an automatic classification system for coffee fruit drying levels based on digital images using a deep learning method with the Convolutional Neural Network (CNN) VGG16 architecture. The dataset used consists of 561 coffee fruit images classified into three classes: Wet, Medium, and Dry. The preprocess sing stages include background removal, auto-cropping, and image standardization. Two models were developed: a baseline model without data augmentation and a model with data augmentation and selective fine-tuning on the final layers of VGG16. The evaluation results show that the baseline model achieved a validation accuracy of 83%, while the model with augmentation and fine-tuning improved the accuracy to 94%, accompanied by significant increases in precision, recall, and F1-score values. The proposed model also demonstrates a high and stable level of prediction confidence. These results prove that the VGG16 approach is effective for classifying coffee fruit drying levels and has the potential to be applied as an objective post-harvest quality control support system.

Keywords: Coffee Fruit; Drying; Image Classification; Deep Learning; VGG16

INTRODUCTION

Coffee is a commodity with high economic value, but the sustainability of its production faces technical challenges at the post-harvest stage. An optimal drying process determines whether coffee fruit can meet the desired quality standards, especially for the specialty coffee category, which has a high premium value. Unfortunately, the assessment of drying levels in the field still relies on subjective, manual visual evaluations that are prone to human error. As a result, the variability in production quality is high and often fails to meet consistent quality standards, particularly for the specialty coffee market, which is highly sensitive to quality.(Silva et al., 2022). Digital image processing technology offers an objective solution for quality control in the agroindustry sector (Sutisna et al., 2020). This technology has been successfully applied for detecting fruit ripeness levels, as well as classifying plant diseases(Folla & Bulan, 2023). Color and texture are key visual indicators of the drying level that can be observed through digital images (Raysyah et al., 2021). However, conventional image processing methods face significant technical challenges, such as variations in lighting and the complexity of extracting visual features that are relevant for accurately distinguishing drying levels (Khairudin & Wahyu Saputro, 2022).

Deep learning, particularly Convolutional Neural Networks (CNNs), offers an optimal solution with automatic feature extraction capabilities (Plested & Gedeon, 2022). CNN has proven to excel in agricultural applications such as plant disease detection and product quality identification without the need for manual feature definitions (Kusuma et al., 2025). In the context of coffee, previous research has focused more on coffee diseases, using architectures such as GoogLeNet and ResNet, MobileNetV2, as well as VGG-19 (Abuhayi & Mossa, 2023). However, the use of deep learning specifically for classifying drying levels is still limited due to differences in visual characteristics; disease detection focuses on local features (spots), whereas drying involves global morphological changes that are more difficult to distinguish(Sucia et al., 2023). The VGG16 architecture was chosen in this study because of its advantage in capturing texture-heavy features compared to lighter models(Rani et al., 2025). VGG16 also has excellent transfer learning capabilities, which have been shown to significantly improve classification accuracy on small agricultural datasets. The application of transfer learning with VGG16 has also been successfully used in various agricultural applications, including plant disease classification, with satisfactory results(Al Sahili & Awad, 2022).

*name of corresponding author

e-ISSN : 2541-2019
p-ISSN : 2541-044X

In the context of deep learning applications for the coffee industry, previous research has focused on the detection and classification of coffee diseases such as Coffee Leaf Rust and Coffee Berry Disease (Montalbo & Hernandez, 2020). Various CNN architectures and transfer learning techniques have been used with very promising results(Plested & Gedeon, 2022). CNNs with feature concatenation based on GoogLeNet and ResNet achieved 99.08% accuracy, DenseNet achieved 99.57%, and MobileNetV2 even 99.93% for classifying robusta coffee leaf diseases(Abuhayi & Mossa, 2023; Aufar & Kaloka, 2022). VGG-19 with image preprocessing and contours has also been successfully used with an F1-Score of 90%(Montalbo & Hernandez, 2020). However, the use of deep learning to classify and measure the drying level of coffee fruit based on color and texture is still very limited(Ungkawa & Hakim, 2023). This gap arises due to differences in visual characteristics, where disease detection focuses on local features such as spots, while drying involves gradual global morphological changes that are more difficult to distinguish(Sucia et al., 2023). The VGG16 architecture has been proven superior in image classification tasks and is considered a reliable baseline model for modern computer vision applications(Plested & Gedeon, 2022). VGG16 has a simple architecture with layered 3×3 convolutions that allow for effective hierarchical feature extraction. This simplicity facilitates understanding of the model's mechanisms and interpretability of the results. The selection of VGG16 in this study is based on its ability to capture rich texture-heavy features compared to lightweight models such as MobileNet (Rani et al., 2025). Efficient models often perform aggressive dimensionality reduction that can remove fine spatial details such as coffee fruit surface wrinkles, even though these details are key indicators of drying levels. VGG16 also has excellent transfer learning capabilities, where weights trained on large datasets like ImageNet can be reused for limited agricultural datasets. The use of transfer learning with VGG16 has been shown to improve accuracy on small-sized datasets compared to training a model from scratch. This approach has been widely applied in agricultural domains, including fruit quality detection, plant nutrient deficiency identification, and plant disease classification (Al Sahili & Awad, 2022).

Although deep learning and CNN have rapidly advanced in agriculture, there remains a research gap, particularly in classifying the drying levels of coffee fruit (Ungkawa & Hakim, 2023). Most studies focus on detecting coffee diseases or classifying ripe fruits, rather than assessing drying conditions(Raysyah et al., 2021). Furthermore, few studies have simultaneously combined color and texture feature extraction using VGG16 to classify drying levels into practical field categories. To address this need, this study defines output classes based on observable visual changes at three distinct stages of the sun-drying process: (1) Wet, representing coffee fruit at the beginning of the drying day with bright red or purple color and glossy surfaces; (2) Medium, representing fruit at the midpoint of drying with transitional grayish-brown color and more complex texture; and (3) Dry, representing fruit at the end of the drying process with dark brown color and wrinkled surface texture**.** These class definitions were established in collaboration with experienced coffee farmers who verified that each acquisition time corresponded to a visually distinct and practically recognizable drying stage, thereby reducing subjectivity in the ground truth annotation. This gap presents opportunities to develop automated systems that can enhance the consistency and objectivity of coffee drying assessments(Silva et al., 2022).

LITERATURE REVIEW

Digital Image Processing in Agroindustry

Digital image processing has become one of the important approaches in improving objectivity and efficiency in quality control processes in the agroindustry sector. This technology enables the identification of visual characteristics such as color, texture, and shape automatically, so that it can replace subjective manual methods (Sutisna et al., 2020). In practice, image processing has been successfully applied to detect fruit ripeness levels as well as identify plant diseases accurately (Folla & Bulan, 2023). In addition, research by (Raysyah et al., 2021) shows that color and texture features are the main indicators in determining the physical condition of an agricultural product, including in the drying process. However, conventional image processing methods still have limitations, especially in dealing with lighting variations and the complexity of extracting relevant features. This can reduce system accuracy under uncontrolled environmental conditions (Khairudin & Wahyu Saputro, 2022).

Deep Learning and Convolutional Neural Network (CNN)

The development of deep learning, especially Convolutional Neural Network (CNN), has brought significant progress in the field of computer vision. CNN has the ability to perform automatic feature extraction without requiring manual feature engineering, making it more adaptive to variations in visual data(Plested & Gedeon, 2022). In the agricultural sector, CNN has been widely used for various applications such as plant disease detection and product quality classification. (Kusuma et al., 2025) that CNN is capable of producing high performance in image-based classification due to its ability to capture complex patterns in visual data. Several popular CNN architectures such as GoogLeNet, ResNet, DenseNet, and MobileNetV2 have shown excellent performance, even achieving accuracy above 99% in coffee leaf disease classification (Abuhayi & Mossa, 2023).

*name of corresponding author

SinkrOn
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16055
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Application of Deep Learning in Coffee

Research related to coffee has mostly focused on detecting diseases such as Coffee Leaf Rust and Coffee Berry Disease (Montalbo & Hernandez, 2020). Various deep learning approaches have been used and have produced very high performance in classifying these diseases. For example, CNN based on feature concatenation with GoogLeNet and ResNet architectures is able to achieve accuracy up to 99.08%, while DenseNet achieves 99.57% and MobileNetV2 reaches 99.93% (Abuhayi & Mossa, 2023). In addition, the use of VGG-19 with certain preprocessing techniques also shows good performance with an F1-Score of 90% (Montalbo & Hernandez, 2020). However, most of these studies focus on detecting local features such as spots or damage on leaves, making them less relevant for drying cases that involve global changes in the structure and color of coffee fruit (Sucia et al., 2023).

VGG16 Architecture and Transfer Learning

VGG16 is one of the CNN architectures that is widely used in various image classification tasks. This model has a simple structure with small-sized convolution layers (3×3) arranged deeply to capture hierarchical features effectively (Plested & Gedeon, 2022). The main advantage of VGG16 lies in its ability to capture complex texture features, making it suitable for cases involving subtle visual changes such as drying levels (Rani et al., 2025). Compared to lightweight models such as MobileNet, VGG16 tends to retain richer spatial details. In addition, VGG16 supports a transfer learning approach, where weights that have been trained on large datasets such as ImageNet can be reused for smaller datasets. This approach has been proven to significantly improve model accuracy in various agricultural applications (Al Sahili & Awad, 2022).

Research Gap

Although deep learning technology has rapidly developed in the agricultural field, there is still a research gap in classifying coffee drying levels. Most studies focus more on detecting diseases or classifying fruit ripeness, rather than assessing drying conditions (Ungkawa & Hakim, 2023). In addition, studies that combine color and texture features simultaneously to classify drying levels are still very limited. In fact, the drying process involves more complex global morphological changes compared to disease detection which is local in nature (Raysyah et al., 2021). Therefore, it is necessary to develop an automatic classification system that is able to identify coffee drying levels based on digital images using a deep learning approach. This study attempts to fill this gap by using the VGG16 architecture to classify drying levels into Wet, Medium, and Dry categories, so that it can improve the objectivity and consistency of coffee quality assessment.

METHOD

This study uses a dataset of 561 coffee fruit images collected directly from coffee farmers through an image acquisition process in the field. All images depict the real condition of coffee fruit at various stages of the drying process, thus providing an accurate visual representation of color and texture changes at each drying phase (as shown in Figure 1). The dataset is divided into three main classes, namely Wet, Medium, and Dry, with the distribution of the number of samples per class presented in Table 1.

Table 1. Dataset Class Distribution

Fruit Condition	Wet	Medium	Dry
Collection Time	Dataset collection conducted at the beginning of the drying day	Dataset collection conducted in the middle of the drying day	Dataset collection conducted at the end of the drying process
Number of Images	183	198	180

Figure 1. Dataset Samples

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16055
e-ISSN : 2541-2019
p-ISSN : 2541-044X

The relatively balanced data distribution among the three classes helps ensure that the developed classification model does not experience bias toward any particular class. The class labels were determined based on the time of image acquisition during the natural sun-drying process: images collected at the beginning of the drying day were labeled Wet, images collected at the midpoint of the drying day were labeled Medium, and images collected at the end of the drying process were labeled Dry. This time-staged labeling approach reflects the progressive visual transformation of coffee fruit throughout drying, characterized by changes in color (from bright red/purple to dark brown) and surface texture (from glossy to wrinkled). Each image in the dataset carries visual information in the form of color, texture, and surface structure attributes of the coffee fruit, which are then used as the basis for automatically determining the drying stage through deep learning–based image analysis.(Raysyah et al., 2021). This dataset, with a realistic representation of field conditions, becomes an important foundation in the development of an image-based drying quality analysis system, as it serves as the primary source for the training and validation process of the coffee fruit drying level identification model (Silva et al., 2022).

Research Stages

Based on the dataset of 561 coffee fruit images that have been collected, the stages of the research methodology begin with image preprocessing to improve the quality of visual input, followed by the development of a VGG16-based deep learning model, the implementation of different training strategies for two model variants, and finally a comprehensive evaluation of the classification performance of the three drying level classes in accordance with the research framework shown in Figure 2.

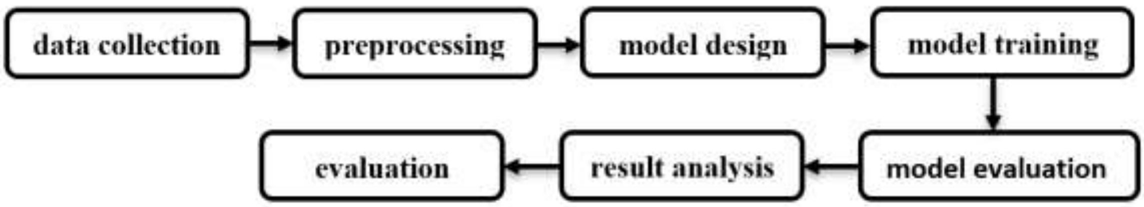

Figure 2. Research Framework

Image Preprocessing

The preprocessing stage was carried out through three main steps to improve the quality of image input (Sutisna et al., 2020). First, background removal was performed using the rembg library to isolate the coffee fruit and eliminate visual noise from the surrounding environment. Second, the images were converted to PNG format to maintain transparency and ensure quality without data loss. Third, the auto-cropping process using the getbbox() function removed empty areas around the object so that the coffee fruit consistently filled the image frame(Ikhsanudin et al., 2023). This preprocessing stage ensures that the VGG16 model receives high-quality, cleaned, and standardized input. As a result, feature extraction accuracy improves and the model can more effectively distinguish the three drying level classes (Wet, Medium, Dry).

Dataset Split

The dataset of 561 coffee fruit images was divided into three subsets using a hold-out split strategy. The training set consists of 392 images (69.9%) used to train the model, the validation set consists of 84 images (15.0%) used to monitor model performance and prevent overfitting during training, and the test set consists of 85 images (15.2%) used for final evaluation on unseen data. This three-way split ensures that model performance assessment is conducted objectively, as the test set is completely isolated from the training and validation process. The split was performed with a fixed random seed to ensure reproducibility of the experimental results.

VGG16 Model Design

The model design uses the VGG16 architecture (Visual Geometry Group with 16 layers) as the backbone for visual feature extraction, with two variants to compare the effectiveness of data augmentation (Akil, 2025). The baseline model was developed by freezing all VGG16 layers (trainable=False) so that it functions as a fixed feature extractor, and only adding classification layers consisting of GlobalAveragePooling2D for spatial feature aggregation, a Dense layer with 128 units and ReLU activation, and an output layer with softmax for three-class classification (Wet, Medium, Dry)(Montalbo & Hernandez, 2020).

The second model applies data augmentation using random flip (horizontal and vertical), random rotation (0.3 radians), random zoom (0.3 scale), random contrast (0.2), and custom sharpening augmentation of 20% to strengthen coffee fruit texture details (Yanto et al., 2021). In the augmented model, fine-tuning was performed by

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16055
e-ISSN : 2541-2019
p-ISSN : 2541-044X

activating the last four layers of VGG16 (block5 layers) while freezing the preceding layers, allowing the model to adjust specific weights to the characteristics of the coffee fruit dataset (Al Sahili & Awad, 2022). Both models were compiled using the Adam optimizer, categorical_crossentropy loss function, and accuracy metric, with different learning rates (0.001 for baseline, 0.0001 for augmented) to maintain training stability (Shao et al., 2024). This architectural design enables an objective comparison between pure transfer learning and a combination of transfer learning with fine-tuning and data augmentation (Plested & Gedeon, 2022).

Model Performance Evaluation

Model performance evaluation was conducted using standard multi-class classification metrics, including accuracy, precision, recall, and F1-score for each class (Wet, Medium, Dry) through a classification report (Rifa'i et al., 2025). In addition, a confusion matrix was used to provide a detailed overview of the model's classification patterns, showing the number of correct and incorrect predictions for each class and identifying which classes are frequently confused or misclassified (Kusuma et al., 2025). The confusion matrix was visualized in the form of a heatmap to facilitate interpretation of classification error patterns.

Evaluation was performed on a validation dataset separate from the training set to ensure objectivity in performance assessment. The baseline model showed high training accuracy but lower validation accuracy, confirming the occurrence of overfitting during training and indicating poor model generalization to previously unseen data (Xiao et al., 2021). In contrast, the VGG16 model with data augmentation and fine-tuning showed significant improvements across all evaluation metrics (accuracy, precision, recall, F1-score) with minimal differences between training and validation metrics, indicating a more robust model capable of recognizing visual variations of coffee fruit images more accurately. The comparison of these two models validates the effectiveness of data augmentation and fine-tuning strategies in improving model performance for the classification of coffee fruit drying levels(Shao et al., 2024).

RESULT

The research results show a significant performance comparison between the two VGG16 model variants in classifying coffee fruit drying levels based on comprehensive evaluation metrics.

a. Baseline Model Without Data Augmentation

The baseline model without data augmentation achieved an overall accuracy of 83% with good precision, recall, and F1-score for all three classes. At the per-class level, the baseline model showed the best performance on the Wet class with precision and recall of 0.96 and 0.96, respectively, followed by the Dry class with precision of 0.80 and recall of 0.74, while the Medium class showed the lowest performance with precision of 0.75 and recall of 0.80, as presented in Table 2.

Table 2. Results Without Augmentation

Class	Precision	Recall	F1-Score	Support
Wet	0.96	0.96	0.96	27
Dry	0.8	0.74	0.77	27
Medium	0.75	0.8	0.77	30
Accuracy			0.83	84
Macro Avg	0.84	0.83	0.84	84

The baseline model confusion matrix shows that most samples in the Wet class were successfully classified correctly, with 26 out of 27 samples predicted according to their class and only 1 sample misclassified as the Medium class. In the Dry class, the model produced 20 correct predictions out of 27 samples, with 7 samples incorrectly predicted as the Medium class. Meanwhile, the Medium class had 24 correct predictions out of 30 samples, with 1 misclassification to the Wet class and 5 misclassifications to the Dry class, as shown in Figure 3. This pattern indicates that the baseline model has difficulty distinguishing between the Dry and Medium classes, which have relatively similar visual characteristics.

*name of corresponding author

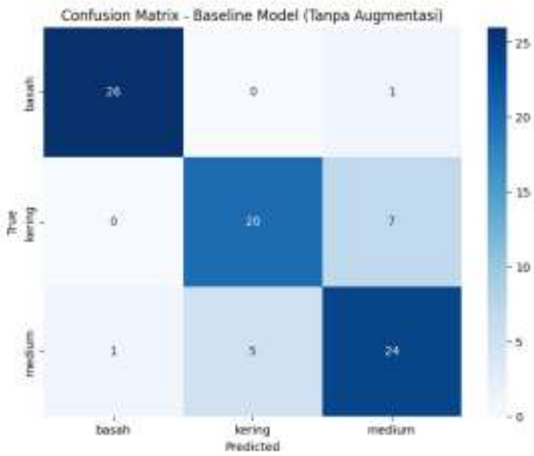

Figure 3. CM Results Without Augmentation

As illustrated in Figure 4, the baseline model accuracy plot shows a relatively smooth training pattern with a gradual increase. Training accuracy increases consistently until reaching approximately 92%, while validation accuracy increases more gradually and tends to stabilize in the range of 83–85% in the final epochs. This pattern indicates a mild overfitting tendency, characterized by a gap between training and validation accuracy of approximately 7–9%. From the loss perspective, the curves show that training loss continues to decrease consistently to a value of around 0.25, whereas validation loss decreases more slowly and tends to plateau around 0.33, indicating limitations in the model's generalization capability. This condition shows that the baseline VGG16 model with all layers fully frozen functions as a stable feature extractor but has limitations in adapting high-level features to the specific characteristics of the coffee fruit dataset, causing validation performance to plateau at an accuracy level of around 83%.

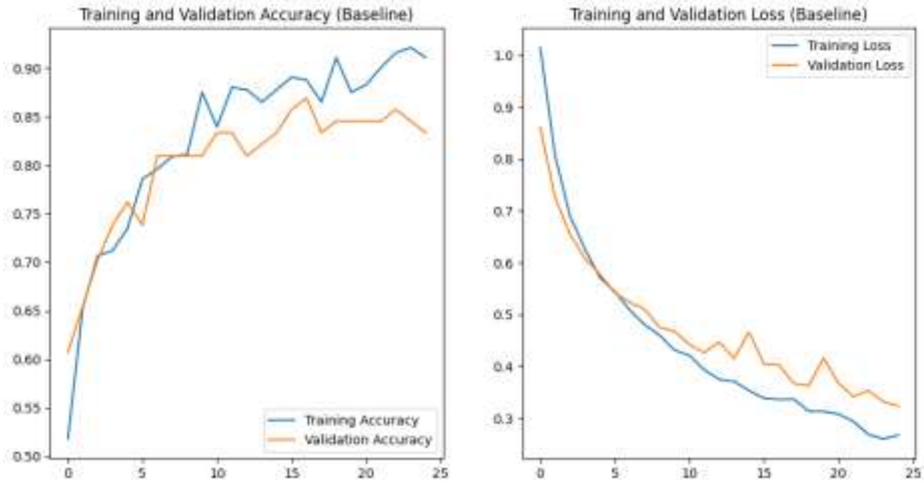

Figure 4. Training and Validation Without Augmentation

b. Model with Data Augmentation and Fine-Tuning

In contrast, the model with data augmentation and fine-tuning shows consistent performance improvements across all metrics. Overall accuracy increases to 94%, indicating a significant improvement compared to the baseline model. At the class level, the Wet class achieves the highest performance with precision of 1.00 and recall of 1.00 (F1-score 1.00), the Dry class shows improvement with precision of 1.00 and recall of 0.81 (F1-score 0.90), while the Medium class experiences the most significant improvement with precision of 0.86 and recall of 1.00 (F1-score 0.92), as presented in Table 3. In particular, the greatest improvement occurs in the Medium class, especially in the recall aspect, which reaches the maximum value, indicating the model's success in recognizing all samples of that class.

Class	Precision	Recall	F1-Score	Support
Wet	1.00	1.00	1.00	27
Dry	1.00	0.81	0.9	27
Medium	0.86	1.00	0.92	30
Accuracy			0.94	84
Macro Avg	0.95	0.94	0.94	84

Table 3. Results With Augmentation

The confusion matrix of the model with augmentation shows a more balanced and accurate classification pattern. All 27 samples of the Wet class were successfully classified correctly without any errors. In the Dry class, the model produced 22 correct predictions out of 27 samples, with 5 samples misclassified as the Medium class. Meanwhile, all 30 samples of the Medium class were successfully predicted correctly without any misclassification to other classes, as illustrated in Figure 5. This pattern indicates that the application of augmentation and fine-tuning significantly reduces ambiguity between the Dry and Medium classes, particularly by improving the model's ability to consistently recognize the Medium class.

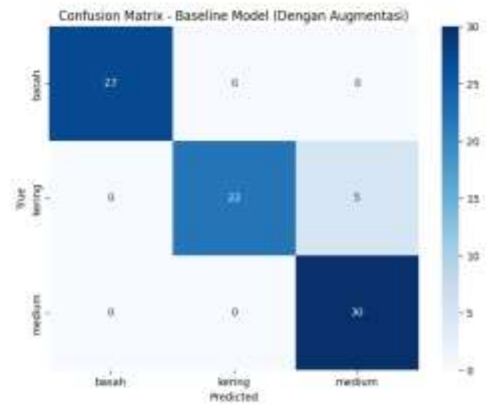

Figure 5. Confusion Matrix Results With Augmentation

As illustrated in Figure 6, the accuracy and loss plots show a more stable and convergent training pattern. Training accuracy gradually increases to reach the range of 93–95%, while validation accuracy reaches approximately 94% in the final epoch. The gap between training and validation accuracy is relatively small and stable, in the range of 1–3%, indicating reduced overfitting compared to the baseline model. From the loss perspective, training loss decreases consistently to around 0.18, while validation loss also decreases and stabilizes in the range of 0.17–0.19, showing a healthy convergence pattern. Overall, these curves indicate that the application of data augmentation and selective fine-tuning on the final layers of VGG16 successfully improves the model's generalization ability, resulting in more robust and stable performance compared to the baseline model.

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16055
e-ISSN : 2541-2019
p-ISSN : 2541-044X

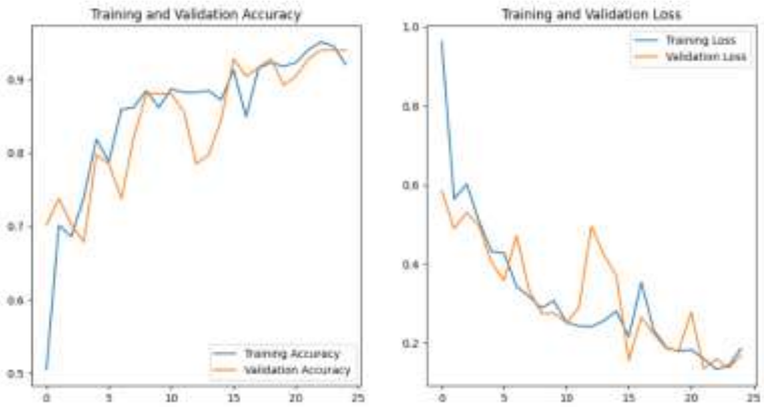

Figure 6. Training and Validation With Augmentation

Analysis of Confidence Score and Prediction

a. Confidence Score Distribution

The confidence score analysis on the validation set shows that the model with augmentation has a high and relatively consistent level of confidence. The average confidence score of the augmented model reaches 0.9436 (94.36%), which is higher than that of the baseline model. The confidence score distribution of the augmented model shows that the majority of predictions are concentrated in the range of 0.95–1.00, with most samples having very high confidence values close to 1.00, indicating that the model has strong confidence in its prediction results, as shown in Figure 7. This distribution pattern appears sharper and more centralized, in contrast to the baseline model, which shows a wider spread of confidence scores extending to lower values. This condition indicates that the application of augmentation and fine-tuning not only improves classification accuracy but also enhances the stability and confidence of the model in generating predictions on validation data.

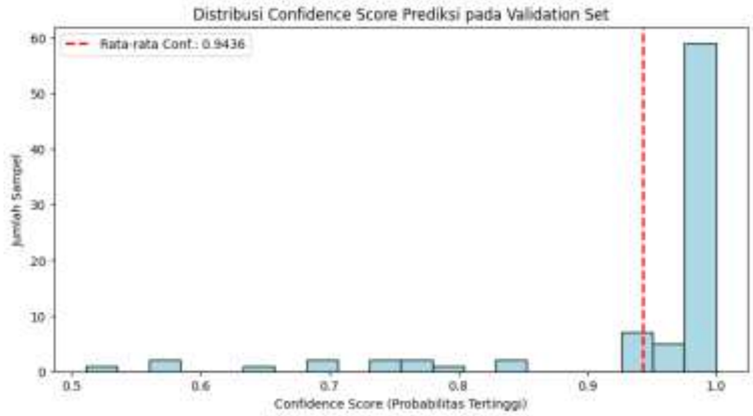

Figure 7. Confidence Score Distribution

b. Visual Analysis of Model Predictions

The visual analysis of the model's predictions shows that the model with augmentation is able to accurately classify most coffee fruit samples across the three drying levels. Examples of correct predictions show that the model can recognize distinctive visual characteristics for each class: Wet coffee fruits with light red/purple color and glossy surfaces, Medium coffee fruits with grayish-brown color and more complex texture, and Dry coffee fruits with dark brown color and more wrinkled surface texture. Nevertheless, there are several cases of incorrect predictions, indicating that the model has difficulty distinguishing samples that lie between classes (boundary cases), especially in the transition between the Medium and Dry classes, which are visually similar, as shown in Figure 8. Examples of misclassification cases indicate that coffee fruits with ambiguous visual characteristics (e.g., fruits in the middle of the drying process) are often predicted into adjacent classes.

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16055
e-ISSN : 2541-2019
p-ISSN : 2541-044X

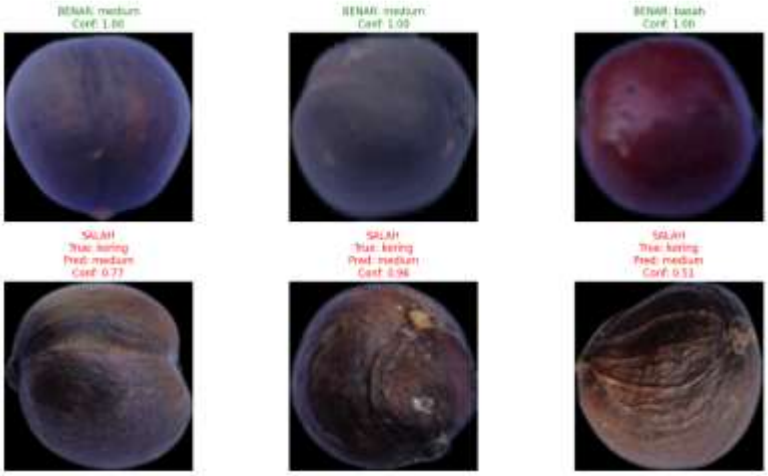

Figure 8. Prediction Analysis

c. Summary of Results

Based on the confidence score analysis on the validation set, the model with augmentation and fine-tuning shows a high and stable level of prediction confidence, with an average confidence score of 94.36%. Most predictions are concentrated in the range of 0.95–1.00, indicating that the model is not only accurate but also consistent in providing predictions with a high level of confidence. The sharp and centralized distribution of confidence scores indicates an improvement in prediction quality compared to the baseline model, which tends to have a wider spread of confidence values.

The visual analysis of predictions reinforces these quantitative findings. The model is able to recognize the main visual characteristics of the three coffee fruit drying level classes, namely Wet, Medium, and Dry, with high accuracy. The remaining prediction errors generally occur in samples with ambiguous visual characteristics, particularly at the transition boundary between the Medium and Dry classes, which are visually highly similar. These findings indicate that although the model already has good generalization capability, the main challenge still lies in classifying samples with intermediate conditions, which require more discriminative feature representations.

DISCUSSIONS

The results of this study show that the VGG16 architecture with a transfer learning approach is able to achieve an accuracy of 83% in the baseline model and increases to 94% in the model with data augmentation and fine-tuning for the task of classifying coffee fruit drying levels. This improvement validates that low-level features such as edges, textures, and colors learned from the ImageNet dataset can be effectively combined with high-level domain-specific features, resulting in good classification performance even though the dataset used is relatively small and originates from the agricultural domain. In the per-class evaluation, the model shows the best performance in the Wet class, with a precision value of 0.96, recall of 0.96, and F1-score of 0.96. This indicates that the visual characteristics of wet coffee fruits, such as bright red or purple color and relatively glossy surfaces, are highly distinctive and easily recognized by the model. In contrast, the greatest challenge arises in classifying the Medium and Dry classes, particularly for samples located in the transition zone. Misclassifications in these two classes decrease from 12 to 5 samples after the application of augmentation, but still indicate that the gradual transition in color and surface texture between the mid-drying and end-of-drying stages does not have a clearly defined visual boundary. This phenomenon is consistent with coffee drying theory, where the Medium and Dry phases have overlapping visual characteristics, unlike the Wet phase, which is more easily distinguishable.

The application of custom data augmentation in the form of random flip, rotation, zoom, contrast, as well as 20% sharpening, combined with selective fine-tuning on the last four layers of VGG16, is proven to provide a significant reduction in classification errors. The reduction in errors reaches 27% in the Dry class and 18% in the Medium class, indicating that this strategy is effective in strengthening the model's ability to distinguish subtle texture and color characteristics. The overall accuracy improvement of 11 percentage points (from 83% to 94%) carries substantial practical significance because it directly reduces the misclassification of coffee fruit in the final drying stage, which is crucial for objective assessment of post-harvest drying completion in the specialty coffee segment. The sharpening augmentation is specifically designed to highlight surface texture details that serve as the main distinguishing features among drying phases.

The selective fine-tuning strategy is also proven to be more efficient than full retraining in terms of resource requirements and training time, and more effective than the full freeze approach because it still allows domain

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v10i3.16055 p-ISSN : 2541-044X

adaptation to occur. The use of a lower learning rate (0.0001) plays an important role in maintaining training stability and preventing overly aggressive weight updates in the pretrained layers.

The confidence score distribution of the augmented model shows an average value of 94.36%, with 56.7% of predictions falling within the range of 0.99–1.00, indicating a very high level of prediction confidence. Analysis of incorrect predictions shows that errors generally occur in lower confidence ranges (0.48–0.74), indicating that the model has good calibration awareness. This condition opens opportunities for implementing a hybrid human–AI system with a confidence threshold of 0.80. Under this threshold, only about 5–10% of ambiguous samples require manual inspection, while the remaining 90–95% can be processed automatically, increasing system reliability for field applications. In addition, the application of domain-specific preprocessing in the form of background removal using rembg, auto-cropping with the getbbox() function, and conversion of image formats to PNG contributes significantly to improving model performance. These stages are effective in eliminating visual noise from the surrounding environment and maximizing the signal-to-noise ratio, allowing the model to focus more on the main visual characteristics of coffee fruits. These results confirm the fundamental principle that preprocessing quality has a direct influence on the quality of deep learning model outputs.

When compared to previous studies that generally focus on coffee disease detection such as Coffee Leaf Rust or Coffee Berry Disease using CNNs, this study has a fundamental difference. Disease detection tends to rely on local features with clear visual boundaries, whereas drying level classification involves gradual and continuous global morphological changes. Thus, this study fills a research gap by developing a classification system specifically for coffee fruit drying levels, rather than merely disease detection, and demonstrates that deep learning approaches remain effective even when facing more complex visual challenges.

Overall, the results of this study indicate that the application of data augmentation and selective fine-tuning on the VGG16 model is able to consistently improve the model's generalization capability compared to the baseline model without augmentation. The augmented model not only achieves higher accuracy, but also shows better training stability, reduced misclassification among overlapping classes, and a high level of prediction confidence, resulting in a more robust and reliable coffee fruit drying level classification system.

CONCLUSION

This study demonstrates that a deep learning approach based on a Convolutional Neural Network (CNN) with a VGG16 architecture is capable of effectively classifying the drying levels of coffee fruit based on visual characteristics of color and texture. The baseline model implementing transfer learning without data augmentation achieved a validation accuracy of 83%. However, the model still exhibited limitations in generalization capability, as indicated by discrepancies between training and validation performance and a high rate of misclassification in classes with overlapping visual characteristics, particularly between the Medium and Dry classes.

The application of data augmentation and selective fine-tuning on the final layers of the VGG16 architecture was proven to significantly improve model performance. The augmented model achieved a validation accuracy of 94%, accompanied by consistent improvements in precision, recall, and F1-score across all classes. The most significant improvement occurred in the Medium class, which previously had the highest misclassification rate, with recall reaching 100%. These results indicate that data augmentation and fine-tuning strategies are effective in enhancing the model's ability to recognize visual variations in the gradual transition phase of the drying process.

Analysis of the confusion matrix and confidence scores shows that the augmented model is not only more accurate but also more stable and consistent in producing predictions. The average confidence score of 94.36% reflects a high level of prediction certainty. The remaining misclassifications generally occurred in samples with ambiguous visual characteristics, particularly at the transition boundary between the Medium and Dry classes. These findings are consistent with the continuous nature of the coffee drying process, in which visual boundaries between classes are not always clearly defined.

Overall, this study contributes to filling the research gap related to the application of deep learning for classifying the drying levels of coffee fruit, which has previously focused more on disease detection or fruit maturity assessment. The combination of domain-based image preprocessing, data augmentation, and selective fine-tuning of the VGG16 architecture has been shown to produce a more robust, objective, and reliable classification system. The proposed system has the potential to be implemented as an automated post-harvest quality assessment tool to improve quality consistency, reduce subjectivity in manual evaluations, and support the competitiveness of the coffee industry, particularly in the specialty coffee segment.

REFERENCES

Abuhayi, B. M., & Mossa, A. A. (2023). Coffee disease classification using Convolutional Neural Network based on feature concatenation. *Informatics in Medicine Unlocked*, *39*, 101245. https://doi.org/10.1016/j.imu.2023.101245

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16055
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Akil, I. (2025). Comparison of Coffee Bean Roasting Level Classification Using ResNet50 and VGG16. *Jutisi : Jurnal Ilmiah Teknik Informatika Dan Sistem Informasi*, *14*(2), 1124. https://doi.org/10.35889/jutisi.v14i2.3122

Al Sahili, Z., & Awad, M. (2022). The power of transfer learning in agricultural applications: AgriNet. *Frontiers in Plant Science*, *13*, 992700. https://doi.org/10.3389/fpls.2022.992700

Aufar, Y., & Kaloka, T. P. (2022). Robusta coffee leaf diseases detection based on MobileNetV2 model. *International Journal of Electrical and Computer Engineering (IJECE)*, *12*(6), 6675. https://doi.org/10.11591/ijece.v12i6.pp6675-6683

Folla, M. M., & Bulan, S. J. (2023). Penerapan Metode Gray Level Co-Occurrence Matrix dalam Mengklasifikasi Tingkat Kematangan Buah Naga Berbasis Citra. *Jurnal Komputer dan Informatika*, *11*(2), 210–116. https://doi.org/10.35508/jicon.v11i2.11847

Ikhsanudin, M. A., Setyati, E., & Junaedi, H. (2023). DETEKSI TINGKAT KEMANISAN BUAH SEMANGKA (CITRULLUS LANATUS) BERDASARKAN CIRI KULIT BUAH DENGAN MENGGUNAKAN METODE CNN (CONVOLUTIONAL NEURAL NETWORK). *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, *8*(4), 1501–1513. https://doi.org/10.29100/jipi.v8i4.4118

Khairudin, N. A. & Wahyu Saputro. (2022). KLASIFIKASI KUALITAS MUTU BUAH DELIMA DENGAN MENGGUNAKAN EKSTRAKSI GRAY LEVEL CO-OCCURRENCE MATRIX (GLCM) DAN K-NEAREST NEIGHBOR (KNN). *Jurnal Informatika Teknologi dan Sains*, *4*(3), 273–278. https://doi.org/10.51401/jinteks.v4i3.1990

Kusuma, B., Hermanto, T. I., & Lestari, C. D. (2025). KLASIFIKASI JENIS PENYAKIT PADA TANAMAN PADI MENGGUNAKAN ALGORITMA CONVOLUTIONAL NEURAL NETWORK. *JIKO (Jurnal Informatika dan Komputer)*, *9*(1), 40. https://doi.org/10.26798/jiko.v9i1.1395

Montalbo, F. J. P., & Hernandez, A. A. (2020). Classifying Barako coffee leaf diseases using deep convolutional models. *International Journal of Advances in Intelligent Informatics*, *6*(2), 197. https://doi.org/10.26555/ijain.v6i2.495

Plested, J., & Gedeon, T. (2022). *Deep transfer learning for image classification: A survey* (arXiv:2205.09904). arXiv. https://doi.org/10.48550/arXiv.2205.09904

Rani, R., Sahoo, J., Bellamkonda, S., & Kumar, S. (2025). Attention-enhanced corn disease diagnosis using few-shot learning and VGG16. *MethodsX*, *14*, 103172. https://doi.org/10.1016/j.mex.2025.103172

Raysyah, S. R., Veri Arinal, & Dadang Iskandar Mulyana. (2021). KLASIFIKASI TINGKAT KEMATANGAN BUAH KOPI BERDASARKAN DETEKSI WARNA MENGGUNAKAN METODE KNN DAN PCA. *JSiI (Jurnal Sistem Informasi)*, *8*(2), 88–95. https://doi.org/10.30656/jsii.v8i2.3638

Rifa'i, W. R. N., Erwanto, D., & Yanuartanti, I. (2025). Evaluasi Kinerja CNN dengan Optimizer RMSprop, Adam dan SGD dalam Klasifikasi Penyakit Daun Anggur. *Jurnal Teknik Elektro*, *10*(1).

Shao, Y., Yang, J., Zhou, W., Sun, H., Xing, L., Zhao, Q., & Zhang, L. (2024). An Improvement of Adam Based on a Cyclic Exponential Decay Learning Rate and Gradient Norm Constraints. *Electronics*, *13*(9), 1778. https://doi.org/10.3390/electronics13091778

Silva, C. S. D., Coelho, A. P. D. F., Lisboa, C. F., Vieira, G., & Teles, M. C. D. A. (2022). Post-harvest of coffee: Factors that influence the final quality of the beverage. *Revista Engenharia Na Agricultura - REVENG*, *30*, 49–62. https://doi.org/10.13083/reveng.v30i1.12639

Sucia, D., ShintyaLarasabi, A. T., Azhar, Y., & Sari, Z. (2023). Classification of Coffee Leaves Diseases Using CNN. *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*. https://doi.org/10.22219/kinetik.v8i3.1745

Sutisna, S. P., Waluyo, R., Aldiansyah, F., & Rahmat, M. (2020). APLIKASI PENGOLAHAN CITRA UNTUK PROSES SORTASI BUAH MANGGA BERDASARKAN DIMENSI DAN BOBOT. *Jurnal Ilmiah Rekayasa Pertanian dan Biosistem*, *8*(1), 12–19. https://doi.org/10.29303/jrpb.v8i1.151

Ungkawa, U., & Hakim, G. A. (2023). Klasifikasi Warna pada Kematangan Buah Kopi Kuning menggunakan Metode CNN Inception V3. *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, *11*(3), 731. https://doi.org/10.26760/elkomika.v11i3.731

Xiao, M., Wu, Y., Zuo, G., Fan, S., Yu, H., Shaikh, Z. A., & Wen, Z. (2021). Addressing Overfitting Problem in Deep Learning-Based Solutions for Next Generation Data-Driven Networks. *Wireless Communications and Mobile Computing*, *2021*(1), 8493795. https://doi.org/10.1155/2021/8493795

Yanto, B., Fimawahib, L., Supriyanto, A., Hayadi, B. H., & Pratama, R. R. (2021). Klasifikasi Tekstur Kematangan Buah Jeruk Manis Berdasarkan Tingkat Kecerahan Warna dengan Metode Deep Learning Convolutional Neural Network. *INOVTEK Polbeng - Seri Informatika*, *6*(2), 259. https://doi.org/10.35314/isi.v6i2.2104

*name of corresponding author