ILKOM Jurnal Ilmiah Vol. 18, No. 1, April 2026, pp. 85-96
Accredited 2nd by KEMDIKTISAINTEK No. 10/C/C3/DT.05.00/2025; E-ISSN 2548-7779 | P-ISSN 2087-1716

Comparative Performance Analysis of Modified VGG16 and Slim-CNN for Arabica Coffee Bean Defect Classification

Yusriel Ardian [a,1,*]; I Nyoman Gede Arya Astawa [b,2]; Novta Danyel Irawan [a,3]; I Putu Bagus Arya Pradnyana [b,4]; Agung Sulistyo [a,5]
[a] *Politeknik Unisma Malang, Jl. MT. Haryono 193, Malang, 65145, Indonesia*
[b] *Politeknik Negeri Bali, Bukit Jimbaran, Kuta Selatan, Badung, 80361, Indonesia*
[1] *ardian.yusriel@gmail.com;* [2] *arya_kmg@pnb.ac.id;* [3] *novta@polisma.ac.id;* [4] *bagusarya12@pnb.ac.id;* [5] *agung@polisma.ac.id*
** Corresponding author*

Article history: Received February 03, 2026; Revised February 11, 2026; Accepted March 06, 2026; Available online April 20, 2026

Abstract
Defect detection in Arabica coffee beans is a critical aspect of quality control, particularly for export-oriented commodities that require consistent visual standards and uniform quality across production batches. Black and partial-black defects are known to significantly affect market value, quality perception, and sensory characteristics. Meanwhile, manual inspection processes remain vulnerable to evaluator subjectivity and inter-operator inconsistency.This study aims to conduct a comparative analysis between a Modified VGG16 architecture and Slim-CNN for detecting these two defect categories using a deep learning-based Convolutional Neural Network (CNN) approach. The dataset consists of 4,080 high-resolution images of Arabica green coffee beans captured using a 24.2 MP macro camera under controlled lighting conditions to minimize shadows and visual distortion. To preserve the natural characteristics of the defects, minimal data augmentation was applied using cropping and 15-degree rotation techniques. The Modified VGG16 architecture was simplified by reducing the complexity of the fully connected layers, integrating batch normalization, and applying dropout to enhance training stability and computational efficiency. Slim-CNN was employed as a lightweight comparative model with fewer parameters and lower memory requirements, making it suitable for resource-constrained deployment scenarios. Four training schemes were evaluated using variations in learning rate and epoch number to assess configuration impacts on performance. Experimental results show that Modified VGG16 achieved the highest test accuracy of 86.7% at a learning rate of 0.001 with 3 epochs, demonstrating a strong balance between training and validation accuracy. Slim-CNN exhibited shorter training time and lower computational complexity, although with slightly lower classification accuracy compared to Modified VGG16. These findings highlight a trade-off between classification performance and computational efficiency in selecting CNN architectures for coffee bean defect detection. Although the results demonstrate potential for industrial automatic classification systems, further validation using larger datasets and more comprehensive evaluation schemes is required to improve model generalization. This study contributes to the development of a more measurable, adaptive, and efficient deep learning-based coffee quality inspection system to support agro-export industry requirements.

Keywords: Arabica Coffee Beans; Defect Detection; Modified VGG16; Slim-CNN; Image Classification

Introduction

Defect detection in Arabica coffee beans plays a crucial role in maintaining product quality and ensuring compliance with international export standards. Beyond quality assurance, this aspect contributes directly to the economic stability of coffee-producing regions. Visual defects, particularly black and partial-black beans, significantly affect flavor characteristics and reduce cupping scores. This decline in quality ultimately weakens the market value of coffee in the global specialty coffee market [1], [2], [3]. Export-grade coffee classification primarily depends on visual inspection. Therefore, the development of technology capable of objectively detecting defects with high precision has become increasingly essential. Such systems help producers meet stricter quality standards while minimizing human subjectivity inherent in manual inspection processes [4], [5], [6].

Recent advancements in deep learning, especially Convolutional Neural Networks (CNNs), have demonstrated promising performance across various agricultural image classification tasks, ranging from plant disease detection to surface defect identification in agricultural and industrial products [7], [8], [9], [10]. CNNs excel at learning complex hierarchical visual patterns, making them highly suitable for automated quality assessment systems.

Among various CNN architectures, VGG16 is widely recognized for its strong hierarchical feature extraction capability and is frequently utilized as a feature extractor for complex and high-resolution images [11], [6], [12], [13]. However, this performance advantage comes at the cost of substantial computational requirements, which limits real-time implementation in resource-constrained industrial environments [14], [11]. Several studies have attempted to address this limitation through architectural modifications such as reducing convolutional depth, incorporating batch normalization, and applying optimized training strategies to enhance generalization while reducing computational load [15], [16], [17]. Additionally, lightweight CNN architectures have gained increasing attention as efficient alternatives designed for low-latency and resource-constrained deployment scenarios without significantly sacrificing performance [9], [18].

Despite these advancements, comparative studies specifically evaluating modified VGG16 architectures against lightweight CNN models remain limited, particularly in the context of detecting black and partial-black defects in Arabica coffee beans. [6], [5]. Moreover, previous research often relies on extensive and complex augmentation pipelines, while systematic evaluation of minimal augmentation strategies—such as controlled cropping and rotation—remains relatively underexplored in coffee defect classification [19], [20], [21].

This study addresses these gaps through a comprehensive comparative analysis using 4,080 high-resolution images of Arabica green coffee beans captured with a 24.2 MP macro setup. The dataset represents variations in shape, texture, and defect intensity. Data augmentation was intentionally restricted to minimal transformations—cropping and controlled 15-degree rotation—to preserve the natural characteristics of defects and avoid excessive synthetic distortion [22]. The VGG16 architecture was modified by reducing convolutional depth and simplifying fully connected layers. Batch normalization and dropout regularization were incorporated to reduce model complexity and mitigate overfitting during training [23], [6], [24]. Meanwhile, Slim-CNN was implemented as a lightweight, resource-efficient alternative designed for scalable industrial deployment [25], [26].

Through controlled experiments and performance evaluation, this study analyzes training accuracy, validation accuracy, test accuracy, and cross-entropy loss to provide a comprehensive understanding of model efficiency, classification reliability, and practical deployment feasibility in automated coffee quality inspection systems [27], [28]. The findings contribute to advancing automation in the agricultural sector, particularly in improving coffee sorting and quality control processes that traditionally rely on manual methods [29], [30], [31], [32] .

Method

A. Data Collection

The dataset used in this study consists of 4,080 high-resolution images of Arabica green coffee beans representing two primary defect categories: black defects and partial-black defects. All images were captured using a 24.2 MP macro camera under controlled lighting conditions to ensure uniform illumination and minimize shadows [6], [31], [33], [4].

Each coffee bean was photographed from multiple angles to capture variations in surface texture and defect distribution comprehensively. The dataset collection process was conducted manually to ensure visual consistency and to accurately represent real-world defect conditions. This approach is particularly important in deep learning-based classification tasks, as accurate visual documentation of black defects directly influences quality assessment and export eligibility standards.

B. Data Preprocessing

Before the training process begins, all images are first processed to ensure consistency and compatibility with the CNN architecture requirements. Each image is resized to $224 \times 224 \times 3$, and subsequently normalized to a pixel value range commonly used in deep learning frameworks. The data are then systematically organized into three primary subsets—training, validation, and testing datasets—allowing the model evaluation process to be conducted in a structured and systematic manner [34], [12], [16].

The dataset used in this study consists of 4,080 images evenly distributed across two classes, comprising 2,040 images of black defects and 2,040 images of partial-black defects. To ensure fair and reproducible evaluation, the dataset was partitioned using a stratified random sampling method with a composition of 70% training data (2,856 images), 15% validation data (612 images), and 15% testing data (612 images). The stratification approach was applied to maintain consistent class proportions within each subset. A fixed random seed was employed during the splitting process to ensure experimental reproducibility.

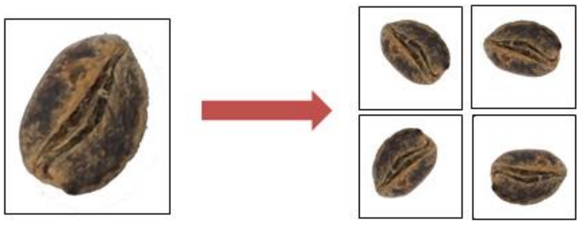


Figure 1. Preprocessing of Arabica coffee bean images, illustrating the comparison between the original images captured by the camera and the images after being resized to 224 × 224 × 3 as input to the CNN architecture.

Figure 1 presents the images of Arabica coffee beans before and after the resizing process. The original images captured using a macro camera exhibit varying resolutions and aspect ratios. Therefore, all images were resized to 224 × 224 × 3 to comply with the input specifications of the CNN architecture employed in this study. This preprocessing stage plays a critical role in reducing noise while normalizing inter-sample variations. With more uniform data conditions, the training process becomes more stable, and the model's generalization capability can be improved more effectively. [12], [6].

C. Data Augmentation

To enhance model robustness without compromising the natural characteristics of coffee bean defects, this study adopts a minimalist data augmentation strategy. This approach was deliberately selected and differs from many other defect detection studies that typically employ extensive augmentation schemes involving various geometric and photometric transformations. [35], [36].

Minimalist Augmentation Approach:

- ***Cropping***: applied to emphasize localized defect patterns on the surface of the coffee bean. This approach aligns with simple geometric augmentation practices, such as cropping and rotation, which are recognized as effective in increasing data diversity without altering the semantic meaning of class labels in the classification process [35].
- ***Rotation***: a 15-degree rotation was applied to represent the orientation variations of coffee beans commonly encountered during field sorting processes. This approach is consistent with the use of small-angle rotations in defect detection tasks within industrial and agricultural sectors, aiming to improve model robustness without distorting the natural structure of the object [35], [28].

Unlike many defect detection studies in the industrial and agricultural sectors, this research does not rely on an extensive augmentation pipeline. In general, other studies employ combinations of flipping, rescaling, brightness or contrast adjustment, and even advanced techniques such as mosaic augmentation and GAN-based approaches to compensate for data limitations [35], [21], [4], [37]. In this study, the augmentation strategy was intentionally kept highly minimal. The primary focus was directed toward evaluating its impact on model accuracy. This approach also addresses a gap in the defect detection literature, as the effect of limited augmentation on classification performance has rarely been systematically examined, particularly in the context of coffee bean defects [28], [37].

D. Modified VGG16 Architecture

The primary model in this study adopts a modified VGG16 architecture tailored to meet the requirements of coffee bean defect classification.

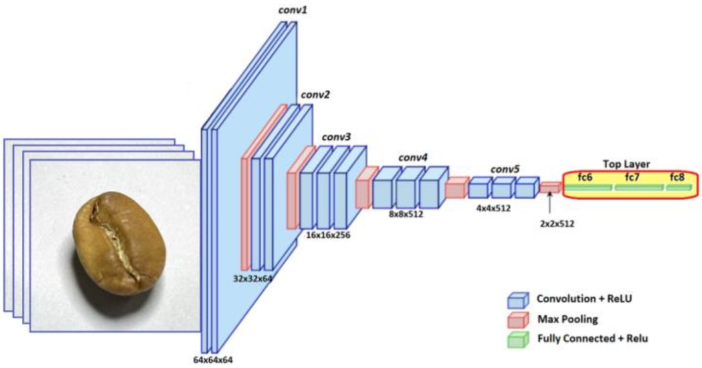


Figure 2. VGG16 architecture for detecting defects in Arabica coffee beans

The modifications were implemented to reduce computational complexity while preserving the required feature extraction capability. Such approaches, including reducing convolutional blocks, incorporating batch normalization, and simplifying the fully connected layers, have previously been shown to be effective in rice seed defect detection and various other classification tasks [13], [23].

- **Three active convolutional blocks** are used as the feature extraction backbone. This design directly reduces the number of parameters and computational load, aligning with the lightweight VGG16 concept previously applied in rice seed recognition and industrial image analysis tasks [13], [23].
- ***Batch normalization*** is applied to each convolutional layer to maintain activation distribution stability and accelerate convergence during training. This approach is consistent with its implementation in VGG16 development for manufacturing defect detection and peanut variety classification, where it has been shown to improve overall model performance [38], [23].
- In the **fully connected layer**, a dropout mechanism with a rate of 0.5 is applied to mitigate the risk of overfitting. This technique is commonly adopted in modified VGG16 architectures across agricultural and industrial applications to maintain stable model performance [6], [38], [23].
- The **dense layer size** is reduced to 1024 neurons, significantly more compact than the standard VGG16 configuration, which utilizes 4096 neurons. This adjustment aligns with several studies that reduce dense layer dimensions to accelerate inference without causing significant accuracy degradation [13], [23].
- In the output layer, a **softmax** function is employed to map probabilities into two defect classes. This approach is consistent with VGG16 implementations in various defect classification tasks within agricultural and industrial sectors, for both binary and multiclass schemes [6], [38], [11], [15].

With these modifications, the proposed VGG16 architecture is expected to maintain optimal classification performance while achieving greater efficiency when deployed in coffee bean defect detection systems. The selection of VGG16 is based on its consistent track record in agricultural defect classification research and its strong capability in extracting features from high-resolution images. However, the depth of the standard architecture often imposes considerable computational burden. Therefore, this study proposes a modified version designed to balance classification accuracy with training time efficiency.

E. Slim-CNN Architecture

Figure 3 illustrates the main stages of image processing in the CNN model used in this study. The coffee bean image, as input, is processed through a series of convolutional layers to extract essential visual features, such as texture, surface patterns, and defect characteristics [6]. Slim-CNN is implemented as a lightweight alternative CNN architecture designed to accelerate the training process while facilitating deployment in resource-constrained environments [25], [26]. This condition is particularly relevant to on-site coffee bean sorting systems, where computational efficiency and fast response time are critical requirements for practical implementation. The Slim-CNN architecture employed in this study is constructed with two convolutional blocks serving as the core feature extraction components. Each block consists of a convolutional layer followed by batch normalization and a ReLU activation function, and is concluded with a max pooling layer to reduce feature dimensionality.

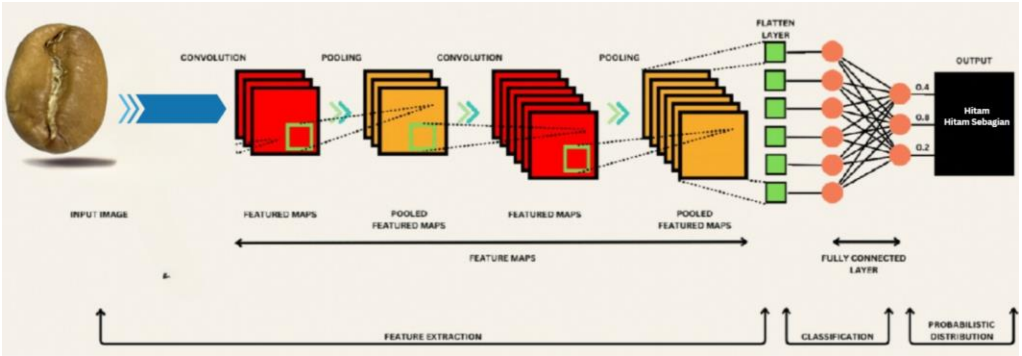


Figure 3. CNN workflow for detecting defects in Arabica coffee beans

After the feature extraction stage, the resulting feature representation is forwarded to a fully connected layer with 256 neurons. A dropout rate of 0.5 is applied at this stage to mitigate the risk of overfitting during training. Final classification is performed using a softmax activation function to map the output into two defect classes. Previous studies have demonstrated that lightweight CNN architectures can be effectively applied across various defect detection domains. These findings suggest advantages in terms of system scalability and real-time processing capability. Therefore, Slim-CNN is utilized as a comparative model to evaluate the balance between computational efficiency and classification accuracy in the task of Arabica coffee bean defect detection [6], [4], [39].

F. Training Options

Three types of optimizers were applied to evaluate the efficiency of the model training process:

1. Adam Optimizer
   - Learning rate: 0.001
   - Epochs: 10
   - Batch size: 32
   - L2 Regularization: 0.001
2. SGD Optimizer
   - Learning rate: 0.001
   - Epochs: 10
   - Batch size: 64
   - L2 Regularization: 0.001
3. RMSprop Optimizer
   - Learning rate: 0.0001
   - Epochs: 15
   - Batch size: 32
   - L2 Regularization: 0.001

The training configurations in this study were selected based on findings from previous research that have demonstrated their effectiveness, particularly in defect detection and agricultural image processing tasks [18], [40], [14].

G. Training Procedure

The model training process consists of the following stages:

1. Dataset Loading

The augmented dataset is loaded into the model training pipeline.

2. Model Initialization

The Modified VGG16 and Slim-CNN architectures are initialized with the predefined layer configurations.

3. **Hyperparameter Configuration**

   Training parameters such as the optimizer, learning rate, batch size, and regularization settings are defined prior to the commencement of training.

4. **Model Training**

   The training process is conducted with validation monitoring performed every 10 iterations to detect potential overfitting.

5. **Model Evaluation**

   The model is evaluated using previously unseen test data to assess its generalization capability.

During training, model weights are updated iteratively and progressively based on the cross-entropy loss value. At each iteration, the optimizer adjusts the model parameters according to its respective update strategy, enabling continuous improvement in model performance [14], [11].

H. Evaluation Metrics

- Model performance is evaluated using:
- Training Accuracy
- Validation Accuracy
- Testing Accuracy
- Cross-Entropy Loss
- Confusion Matrix
- Comparative Performance Analysis

These metrics allow comprehensive assessment of classification capability and training efficiency, enabling fair comparison between the modified VGG16 and Slim-CNN models [18], [20].

I. Experimental Setup

All experiments in this study were implemented using MATLAB with support from the Deep Learning Toolbox. Both the Modified VGG16 and Slim-CNN architectures were trained using the same data splitting scheme and preprocessing stages, ensuring that their performance comparison was conducted objectively and under equivalent conditions.

Based on the experimental results, the optimal configuration was obtained using the Stochastic Gradient Descent with Momentum (SGDM) optimizer, with a learning rate of 0.001, 3 epochs, and a mini-batch size of 64. The L2 regularization factor was set to 0.001 to help control model complexity. Validation was performed every 10 iterations during training, and cross-entropy loss was used as the loss function.

Results and Discussion

1. Overview of Experimental Results

The experimental evaluation was conducted using a dataset consisting of 4,080 images of Arabica coffee beans divided into two defect categories: black defects and partial-black defects. Four training scenarios were designed by varying the primary hyperparameters, namely the learning rate and the number of epochs. In all scenarios, the Modified VGG16 architecture was consistently employed. Meanwhile, the data augmentation strategy was limited to cropping and 15-degree rotation to enable the model to capture natural defect variations without inducing overfitting.

The evaluation protocol utilized the full dataset of 4,080 images, which was divided into three subsets: 70% for training, 15% for validation, and 15% for testing. The split was performed using a stratified approach to ensure balanced class distribution across all subsets. Specifically, the test set consisted of 120 images evenly distributed into 60 black defect images and 60 partial-black defect images. During training and evaluation, cross-entropy was computed as the average loss per mini-batch to ensure consistent and standardized performance measurement.

A summary of model performance, including training accuracy, validation accuracy, test accuracy, and cross-entropy loss, is presented in **Table 1**.

Table 1. Ringkasan hasil *training* dan pengujian

Experiment	LR	Epochs	Validation Accuracy	Training Accuracy	Test Accuracy	Test Loss
Exp 1	0.01	6	96.3%	98.7%	85.0%	2.75
Exp 2	0.001	2	93.9%	94.8%	81.7%	3.41
Exp 3	0.001	3	96.2%	98.1%	86.7%	2.73
Exp 4	0.001	10	98.4%	100.0%	85.0%	2.75

2. Training Performance Analysis

Across all experimental scenarios, the training curves exhibited a rapid increase in accuracy from the first epoch, followed by stabilization at levels above 90%.

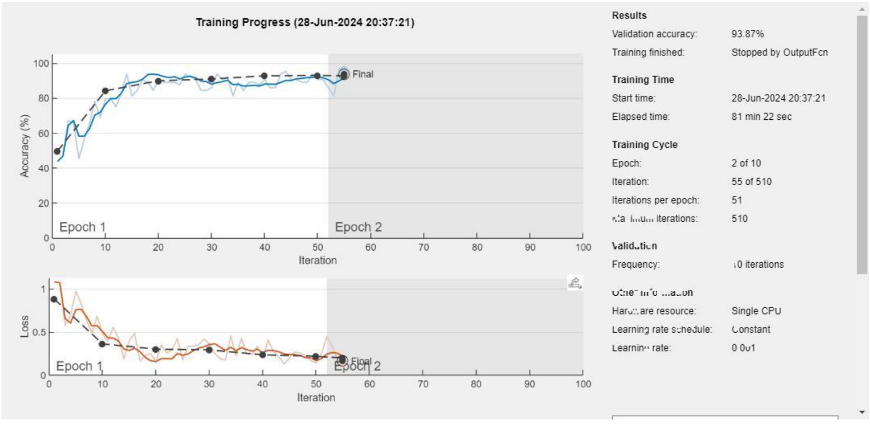
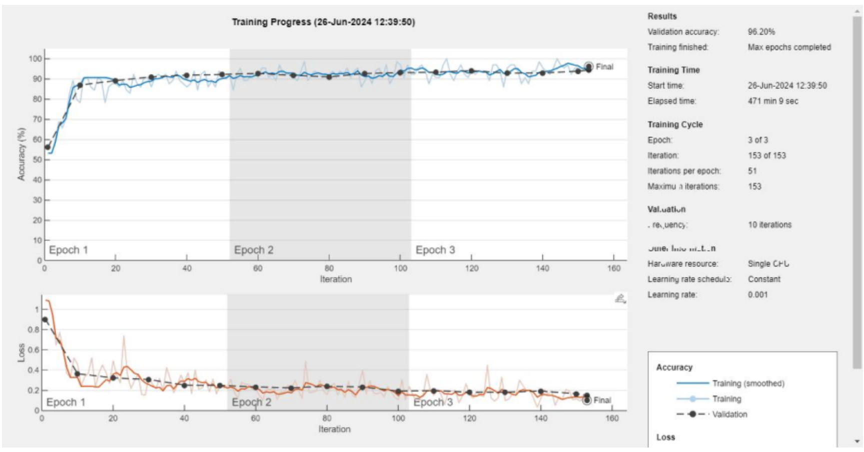


Figure 4. Training performance in Experiment 1

The fourth experiment achieved the highest training accuracy, reaching 100%, with validation accuracy attaining 98.4%. This indicates the model's strong capability in effectively learning feature representations through a longer training cycle. However, an extended training duration began to show signs of mild overfitting. This is reflected in the test accuracy, which plateaued at 85.0% despite the training accuracy reaching its maximum value.



Gambar 5. Training performance in the final experiment

In contrast, the final experiment demonstrated a more balanced performance. The training accuracy of 98.1% and validation accuracy of 96.2% were aligned with the highest test accuracy, reaching 86.7%. These findings indicate that a more moderate number of epochs—specifically 3 epochs with a learning rate of 0.001—resulted in a more optimal learning scheme compared to excessively prolonged training.

3. Testing Performance Analysis

Test accuracy ranged from 81.7% to 87.1%, with Experiment 3 achieving the best performance among all scenarios.

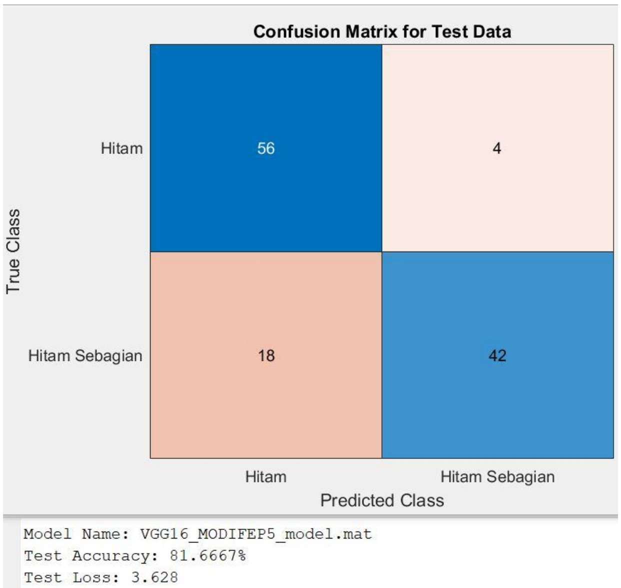


Figure 6. Testing performance in the initial experiment

The lower performance observed in Experiment 2 is associated with an insufficient number of training iterations—only 2 epochs—preventing the model from adequately learning data patterns and leading to underfitting. Meanwhile, Experiment 1 and Experiment 4 both recorded a test accuracy of 85.0%, which can be considered competitive. However, Experiment 3 achieved a more optimal balance between training duration and model generalization capability.

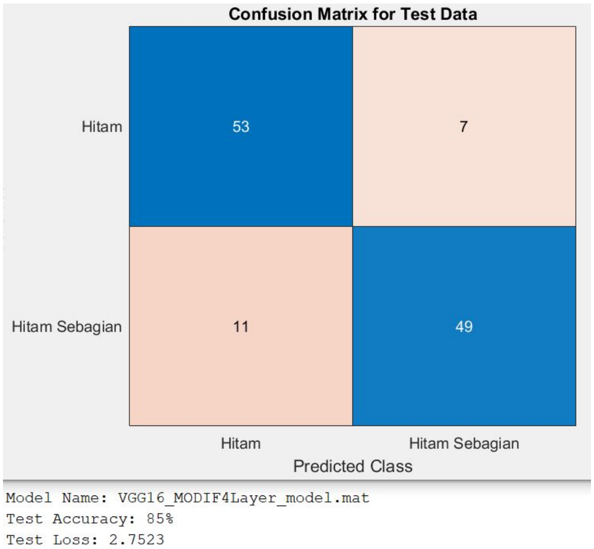


Figure 7. Testing performance in the final experiment

The loss values reported in this study represent the accumulated loss throughout the training process rather than the final per-sample loss on the test data. Therefore, these values reflect the overall learning dynamics of the model during training. To provide a more representative evaluation of classification performance, the interpretation of results is primarily based on test accuracy metrics and confusion matrix analysis, which illustrate the distribution of correct and incorrect predictions for each class.

4. Confusion Matrix Interpretation

The resulting confusion matrix provides a detailed overview of the model's classification behavior for each defect category, as follows:

Confusion matrix Experiment3:

- Black defect coffee beans correctly classified: 56

- Black defect coffee beans misclassified as partial-black defects: 4
- Partial-black defect coffee beans correctly classified: 48
- Partial-black defect coffee beans misclassified: 12

Based on these values:

- Detection accuracy for black defects: 93.3%
- Detection accuracy for partial-black defects: 80.0%

These results indicate that the model is more effective in recognizing the dark texture characteristics of black defects compared to the subtler color gradations found in partial-black defects. This finding is consistent with previous studies in agricultural defect detection, which suggest that stronger color contrast facilitates more accurate feature extraction by CNN models.

5. Comparative Discussion Across All Experiments

The comparison across experiments indicates that model performance is highly influenced by the learning rate configuration and the number of training epochs, as detailed below:

- **A high learning rate (0.01)** accelerates the convergence process but carries the risk of less stable parameter updates. This is evident in the validation curve of Experiment 1, which tends to stagnate.
- **A lower learning rate (0.001)** results in a more stable learning process, as observed in Experiments 2 through 4.
- **A longer training cycle (10 epochs)** is capable of increasing training accuracy; however, this improvement does not necessarily translate into better generalization performance.
- **A moderate training depth (3 epochs)** provides the most optimal balance between accuracy and generalization, as reflected in the highest test accuracy achieved in Experiment 3.
- These findings are consistent with previous research emphasizing that CNN optimization requires a balance between training duration and model generalization performance [6].

6. Performance Comparison between Modified VGG16 and Slim-CNN

This study compares the performance of Modified VGG16 and Slim-CNN in detecting black and partial-black defects in Arabica coffee beans. The evaluation is based on test accuracy, training stability, and computational efficiency. Modified VGG16 achieved the highest accuracy of 86.7% at a learning rate of 0.001 with 3 epochs, while also demonstrating consistency between training and validation accuracy, indicating good generalization capability.

In contrast, Slim-CNN offers shorter training time and lower memory requirements, but with lower classification accuracy. These results indicate a trade-off between computational efficiency and classification performance, suggesting that model selection should be aligned with specific system implementation requirements.

7. Comparison with Previous Studies

The accuracy of 86.7% achieved in this study demonstrates competitive performance when compared to various previous deep learning approaches for agricultural product defect detection. Several studies employing VGG16-based architectures, particularly in plant disease detection tasks, have also reported high accuracy when applying targeted data augmentation strategies and architectural optimization [41], [42]. On the other hand, some studies report accuracy levels exceeding 95%; however, these results are generally supported by substantially larger datasets, more aggressive augmentation strategies, or more complex fine-tuning processes [42].

The distinctive contribution of this study lies in the finding that a minimal augmentation strategy can still yield strong classification performance. This approach preserves the natural texture characteristics of coffee bean defects while avoiding excessive synthetic variation. Consequently, the feature learning process becomes more representative of real-world conditions [6], [4].

8. Practical Implications

This study demonstrates that the modified VGG16 model, despite being supported only by minimal data augmentation, is still capable of effectively classifying defects in Arabica coffee beans. These findings open opportunities for future implementation in automated quality control systems. With further optimization—such as GPU acceleration, deployment of lightweight models like Slim-CNN, or the development of a system prototype—this approach has the potential to assist the coffee processing industry in maintaining quality consistency in accordance with export standards while reducing the risk of human-related errors.

9. Research Limitations

This study has several limitations that should be considered when interpreting the results. The dataset includes only two defect categories—black and partial-black defects—therefore, the model's generalization capability to other defect types cannot yet be ensured. Additionally, the size of the test dataset is relatively limited. Although the achieved performance is competitive, validation using a larger dataset is still required to improve the reliability of the findings. Furthermore, cross-validation techniques were not applied, meaning that performance variation due to the data splitting scheme may still occur.

Moreover, the developed model has not yet been tested directly in an industrial environment. Consequently, its readiness for real-time coffee bean sorting systems still requires further validation. These limitations provide an important foundation for future research to expand dataset coverage, strengthen evaluation strategies, and test the model under more realistic operational scenarios.

Conclusion

This study developed an Arabica coffee bean defect detection model based on a Modified VGG16 architecture with a minimal augmentation strategy. The experimental results indicate that a configuration with a learning rate of 0.001 and 3 epochs achieved the best performance, as evidenced by a test accuracy of 86.7%. Compared to Slim-CNN, Modified VGG16 provided higher classification accuracy, whereas Slim-CNN demonstrated advantages in computational efficiency and resource requirements.

These findings highlight the balance between classification accuracy and system efficiency in selecting an appropriate architecture. Although the model shows potential for implementation in automated defect detection systems, its generalization capability still requires further validation using a larger dataset and more comprehensive evaluation schemes. Implementation in real-time industrial environments also requires additional testing before operational deployment can be achieved.

References

[1] C. N. Cain *et al.*, "Investigating sensory-classified roasted arabica coffee with GC × GC-TOFMS and chemometrics to understand potato taste defect," *Microchem. J.*, 2023, doi: 10.1016/j.microc.2023.109578.

[2] P. Wang, H.-W. Tseng, T.-C. Chen, and C.-H. Hsia, "Deep Convolutional Neural Network for Coffee Bean Inspection," *Sensors Mater.*, 2021, doi: 10.18494/sam.2021.3277.

[3] R. Rusinek *et al.*, "How to Identify Roast Defects in Coffee Beans Based on the Volatile Compound Profile," *Molecules*, vol. 27, 2022, doi: 10.3390/molecules27238530.

[4] S. Chang and C. Huang, "Deep Learning Model for the Inspection of Coffee Bean Defects," *Appl. Sci.*, vol. 11, no. 17, p. 8226, 2021, doi: 10.3390/app11178226.

[5] M. Jumarlis, M. Mirfan, and A. R. Manga, "Classification of Coffee Bean Defects Using Gray-Level Co-Occurrence Matrix and K-Nearest Neighbor," *Ilk. J. Ilm.*, vol. 14, no. 1, pp. 1–9, Apr. 2022, doi: 10.33096/ilkom.v14i1.910.1-9.

[6] Y. Ardian, N. D. Irawan, S. Sutoko, I. N. G. A. Astawa, I. B. I. Purnama, and F. A. Dwiyanto, "A Novel Approach to Defect Detection in Arabica Coffee Beans Using Deep Learning: Investigating Data Augmentation and Model Optimization," *Knowl. Eng. Data Sci.*, vol. 7, no. 1, p. 117, Nov. 2024, doi: 10.17977/um018v7i12024p117-127.

[7] A. J, J. R, D. Popescu, M. Chowdary, and J. D., "Deep Learning-Based Leaf Disease Detection in Crops Using Images for Agricultural Applications," *Agronomy*, 2022, doi: 10.3390/agronomy12102395.

[8] B. Baiju *et al.*, "Robust CRW crops leaf disease detection and classification in agriculture using hybrid deep learning models," *Plant Methods*, vol. 21, 2025, doi: 10.1186/s13007-025-01332-5.

[9] A. Ashurov *et al.*, “Enhancing plant disease detection through deep learning: a Depthwise CNN with squeeze and excitation integration and residual skip connections,” *Front. Plant Sci.*, vol. 15, 2025, doi: 10.3389/fpls.2024.1505857.

[10] W. B. Demilie, “Plant disease detection and classification techniques: a comparative study of the performances,” *J. Big Data*, vol. 11, pp. 1–43, 2024, doi: 10.1186/s40537-023-00863-9.

[11] R. Klangbunrueang, P. Pookduang, W. Chansanam, and T. Lunrasri, “AI-Powered Lung Cancer Detection: Assessing VGG16 and CNN Architectures for CT Scan Image Classification,” *Informatics*, vol. 12, p. 18, 2025, doi: 10.3390/informatics12010018.

[12] N. Zhou, “Image recognition in depth: comparative study of CNN and Pre-trained VGG16 architecture for classification tasks,” vol. 13075, pp. 130751Y-130751Y–7, 2024, doi: 10.1117/12.3026829.

[13] N. Zakaria and Y. M. M. Hassim, “A Review Study oftheVisual Geometry Group Approachesfor Image Classification,” *J. Appl. Sci. Technol. Comput.*, 2024, doi: 10.30880/jastec.2024.01.01.003.

[14] I. N. G. A. Astawa, M. L. Radhitya, I. W. R. Ardana, and F. A. Dwiyanto, “Face Images Classification using VGG-CNN,” *Knowl. Eng. Data Sci.*, vol. 4, no. 1, p. 49, 2021, doi: 10.17977/um018v4i12021p49-54.

[15] I. N. G. A. Astawa and Y. Ardian, “Sistem Deteksi Cacat Biji Kopi Arabika Berbasis Deep Learning CNN VGG-16 dengan Dataset Augmentasi.” 2025.

[16] S. Cong and Y. Zhou, “A review of convolutional neural network architectures and their optimizations,” *Artif. Intell. Rev.*, vol. 56, pp. 1905–1969, 2022, doi: 10.1007/s10462-022-10213-5.

[17] M. A. Widneh, A. A. Alemu, and D. D. Getie, “Exploring Batch Normalization’s Impact on Dense Layers of Multiclass and Multilabel Classifiers,” *Int. J. Intell. Syst.*, vol. 2025, 2025, doi: 10.1155/int/1466655.

[18] P. Peyal and Others, “CNN and SVM Hybrid Model for Corn Leaf Disease Detection,” *Comput. Agric.*, 2024, doi: 10.1016/j.compag.2024.108755.

[19] Y. Chou *et al.*, “Deep Learning Model for Defective Bean Inspection with GAN-Structured Automated Labeled Data Augmentation in Coffee Industry,” *Appl. Sci.*, vol. 9, no. 19, p. 4166, 2019, doi: 10.3390/app9194166.

[20] S. Arwatchananukul, D. Xu, P. Charoenkwan, S. A. Moon, and R. Saengrayap, “Implementing a Deep Learning Model for Defect Classification in Thai Arabica Green Coffee Beans,” *Smart Agric. Technol.*, vol. 9, p. 100680, 2024, doi: 10.1016/j.atech.2024.100680.

[21] J. J. Bird, C. Barnes, L. J. Manso, A. Ek’art, and D. Faria, “Fruit Quality and Defect Image Classification with Conditional GAN Data Augmentation,” *ArXiv*, vol. abs/2104.05647, 2021, doi: 10.1016/j.scienta.2021.110684.

[22] A. Nasiri, A. Taheri-Garavand, D. Fanourakis, Y. Zhang, and N. Nikoloudakis, “Automated Grapevine Cultivar Identification via Leaf Imaging and Deep Convolutional Neural Networks: A Proof-of-Concept Study Employing Primary Iranian Varieties,” *Plants*, vol. 10, 2021, doi: 10.3390/plants10081628.

[23] N. Zakaria and Y. M. M. Hassim, “Improved Image Classification Task Using Enhanced Visual Geometry Group of Convolution Neural Networks,” *JOIV Int. J. Informatics Vis.*, 2023, doi: 10.30630/joiv.7.4.1752.

[24] S. Kazi, B. Palkar, and D. Mishra, “VCNet: Optimized Deep Learning framework with deep feature extraction and genetic algorithm for multiclass rice crop disease detection,” *MethodsX*, vol. 15, 2025, doi: 10.1016/j.mex.2025.103551.

[25] D. Zhang *et al.*, “An efficient lightweight convolutional neural network for industrial surface defect detection,” *Artif. Intell. Rev.*, vol. 56, pp. 10651–10677, 2023, doi: 10.1007/s10462-023-10438-y.

[26] H. Ma, T. Çelik, and H. Li, “Lightweight attention convolutional neural network through network slimming for robust facial expression recognition,” *Signal, Image Video Process.*, vol. 15, pp. 1507–1515, 2021, doi: 10.1007/s11760-021-01883-9.

[27] K. Muchtar *et al.*, "Edge AI-Based Detection for Defective Coffee Beans Using Deep Learning and Streamlit Framework," *IEEE Access*, vol. 13, pp. 67977–67992, 2025, doi: 10.1109/access.2025.3561189.

[28] H. L. Gope, H. Fukai, F. M. Ruhad, and S. Barman, "Comparative analysis of YOLO models for green coffee bean detection and defect classification," *Sci. Rep.*, vol. 14, 2024, doi: 10.1038/s41598-024-78598-7.

[29] T. Thongnop, T. Perpaman, P. Kansiri, W. Nuchda, and S. Peungsungwan, "Quality Sorting of Green Coffee Beans from Wet Processing by Using The Principle of Machine Learning," *ASEAN J. Sci. Eng.*, 2021, doi: 10.17509/ajse.v1i2.41078.

[30] G. Susanibar, J. Ramírez, J. Sánchez, R. Ramirez, and D. Arroyo, "Development of an Automated Machine for Green Coffee Beans Classification by Size and Defects," *J. Adv. Agric. Technol.*, 2021, doi: 10.18178/joaat.8.1.17-24.

[31] S.-Y. Chen, M.-F. Chiu, and X.-W. Zou, "Real-time defect inspection of green coffee beans using NIR snapshot hyperspectral imaging," *Comput. Electron. Agric.*, vol. 197, p. 106970, Jun. 2022, doi: 10.1016/j.compag.2022.106970.

[32] A. J. N. Lualhati, J. B. Mariano, A. E. L. Torres, and S. D. Fenol, "Development and Testing of Green Coffee Bean Quality Sorter using Image Processing and Artificial Neural Network," *Mindanao J. Sci. Technol.*, 2022, doi: 10.61310/mndjsteect.1048.22.

[33] H. Thai, H. Ko, and J. Huh, "Coffee Bean Defects Automatic Classification Realtime Application Adopting Deep Learning," *IEEE Access*, 2024.

[34] R. A. Lashaki, Z. Raeisi, N. Razavi, M. Goodarzi, and H. Najafzadeh, "Optimized classification of dental implants using convolutional neural networks and pre-trained models with preprocessed data," *BMC Oral Health*, vol. 25, 2025, doi: 10.1186/s12903-025-05704-0.

[35] K. Alomar, H. I. Aysel, and X. Cai, "Data Augmentation in Classification and Segmentation: A Survey and New Strategies," *J. Imaging*, vol. 9, 2023, doi: 10.3390/jimaging9020046.

[36] J. Zhang, G. Cosma, and J. Watkins, "Image Enhanced Mask R-CNN: A Deep Learning Pipeline with New Evaluation Measures for Wind Turbine Blade Defect Detection and Classification," *J. Imaging*, vol. 7, 2021, doi: 10.3390/jimaging7030046.

[37] T. Nguyen and others, "Real-Time Coffee Bean Defect Detection Using YOLOv4," *J. Imaging Sci. Technol.*, vol. 67, no. 2, p. 20501, 2023, doi: 10.2352/J.ImagingSci.Technol.2023.67.2.020501.

[38] A. Baiganova, Z. Ubayeva, Z. Taskalyeva, L. Kaparova, R. Nurzhaubaeva, and B. Umirzakova, "Automated Defect Detection in Manufacturing Using Enhanced VGG16 Convolutional Neural Networks," *Int. J. Adv. Comput. Sci. Appl.*, 2025, doi: 10.14569/ijacsa.2025.0160487.

[39] R. V Murthy, V. P. N., P. Nidoni, N. M., and K. Kumar, "Green Coffee Beans Defects Detection Using Convolution Neural Networks," in *International Advanced Research Journal in Science, Engineering and Technology (IARJSET)*, 2022, pp. 138–143. doi: 10.17148/IARJSET.2022.9725.

[40] V. Mulik and V. Patil, "Smart Agriculture: A CNN-Based Approach for Crop Disease Classification and Detection," *2025 12th Int. Conf. Comput. Sustain. Glob. Dev.*, pp. 1–7, 2025, doi: 10.23919/indiacom66777.2025.11115495.

[41] R. Rani, S. Bharany, D. H. Elkamchouchi, A. U. Rehman, R. Singh, and S. Hussen, "VGG-EffAttnNet: Hybrid Deep Learning Model for Automated Chili Plant Disease Classification Using VGG16 and EfficientNetB0 With Attention Mechanism," *Food Sci. Nutr.*, vol. 13, 2025, doi: 10.1002/fsn3.70653.

[42] R. H. Prince *et al.*, "CSXAI: a lightweight 2D CNN-SVM model for detection and classification of various crop diseases with explainable AI visualization," *Front. Plant Sci.*, vol. 15, 2024, doi: 10.3389/fpls.2024.1412988.