International Journal of Informatics and Communication Technology (IJ-ICT)
Vol. 15, No. 3, September 2026, pp. 1087~1096
ISSN: 2252-8776, DOI: 10.11591/ijict.v15i3.pp1087-1096

Early detection of vascular streak dieback (VSD) disease in cocoa plants using deep learning

Dewi Marini Umi Atmaja[1], Arif Rahman Hakim[1], Fadhil Rozi Hendrawan[2], Komang Diah Devi Pramesty[3], Naufal Fadhilah Fitrah[3], Faiz Rochmatullah Widhaputra[3]
[1]Information System Study Program, Center of Excellence for Human Centric Engineering (HUMIC), Telkom University, Jakarta Campus, DKI Jakarta, Indonesia
[2]Information System Study Program, Center of Excellence for Artificial Intelligence for Learning and Optimization (AILO), Telkom University, Jakarta Campus, DKI Jakarta, Indonesia
[3]Information System Study Program, Telkom University, Jakarta Campus, DKI Jakarta, Indonesia

Article Info

Article history:

Received Nov 5, 2025
Revised May 20, 2026
Accepted Jul 5, 2026

Keywords:

Classification
Cocoa
Convolutional neural network
Deep learning
Vascular streak dieback

ABSTRACT

Vascular streak dieback (VSD), caused by Ceratobasidium theobromae, poses a significant threat to cocoa (Theobroma cacao) production, leading to substantial yield losses and plant mortality. Early detection is critical to mitigate disease spread and reduce economic impact. While convolutional neural network (CNN) architectures like VGG-16 and ResNet-50 excel in leaf disease detection, no prior studies address stem-based VSD symptomatology where the disease originates in vascular tissues. This study presents the first CNN specifically developed for cocoa stem VSD detection, achieving 99.16% test accuracy with a lightweight architecture that outperforms VGG16/ResNet50 in both accuracy and mobile inference speed. A novel dataset of 215 cocoa stem images was curated and augmented for robustness. Additional evaluation metrics, including precision, recall, specificity, and F1-score, further confirm the reliability of the model. The model was successfully converted into TensorFlow Lite (TFLite) format, enabling deployment on mobile devices for real-time disease detection. This study highlights the potential of integrating deep learning into mobile and drone-based agricultural systems to support precision farming and early intervention strategies.

Corresponding Author:

Dewi Marini Umi Atmaja
Information System Study Program, Center of Excellence for Human Centric Engineering (HUMIC)
Telkom University, Jakarta Campus
Jl. Raya Daan Mogot KM. 11, West Jakarta 11710, DKI Jakarta, Indonesia
Email: dewimariniumiatmaja@telkomuniversity.ac.id

1. INTRODUCTION

Cocoa (Theobroma cacao) is a strategic agricultural commodity that supports the livelihoods of millions of farmers in tropical countries, including Indonesia, which is among the world's largest cocoa producers [1]. However, cocoa productivity is consistently threatened by various diseases, with vascular streak dieback (VSD) being one of the most destructive [2]. VSD, caused by infection with the fungus Ceratobasidium theobromae, attacks the vascular tissues of the plant, leading to characteristic symptoms such as leaf necrosis, premature leaf fall, branch dieback, reduced yields, and ultimately, cocoa tree mortality if left unmanaged [3].

Early detection of VSD is essential to prevent the spread of the pathogen and to minimize losses [4]. Traditional detection methods, which rely primarily on manual visual inspection by farmers, present several inherent limitations [5]. These limitations include the subjectivity of diagnosis, dependence on individual experience and expertise, labor intensity, and the inability to detect the disease at an early stage when symptoms are not yet clearly visible [6]. As a result, detection often occurs too late, when the disease has already spread extensively and become difficult to control [5], [6].

In the past decade, rapid advancements in computer vision and deep learning have revolutionized various sectors, including agriculture [7]. Convolutional neural networks (CNNs), as an artificial neural network architecture specifically designed to process visual data, have demonstrated superior performance in classification and object detection tasks [8]. The ability of CNNs to automatically extract hierarchical features from images, without the need for manual feature engineering, makes them an ideal candidate for plant disease detection applications [9], [10]. Recent advances in deep learning have significantly improved the performance of image-based classification systems, particularly in agricultural disease detection. Among these, CNNs have demonstrated superior capability in learning hierarchical representations of visual features. Established architectures such as VGG-16 and ResNet-50 are widely recognized for their robustness and generalization ability. VGG-16 offers a simple yet effective deep architecture, while ResNet-50 introduces residual connections that address vanishing gradient issues in deeper networks.

Despite their effectiveness, these architectures are often computationally expensive and may not be optimal for deployment in resource-constrained environments such as mobile devices. Therefore, this study proposes a custom CNN architecture with optimized depth and parameter efficiency, aiming to achieve high accuracy while maintaining feasibility for real-time applications in agricultural settings. This study aims to harness the potential of CNNs to develop an automated detection model for VSD in cocoa stems. The focus on cocoa stems is particularly important because VSD symptoms often first appear in the vascular tissues of stems and young branches before spreading to the leaves. In addition, this research specifically explores the impact of varying the number of convolutional layers on CNN performance to determine the optimal configuration. By providing a fast, accurate, and objective detection system, this study seeks to empower cocoa farmers to take proactive mitigation measures, thereby improving plantation management efficiency and ensuring the sustainability of cocoa production.

2. RESEARCH METHOD

2.1. Literatur review

Early studies on plant disease detection often relied on expert systems that utilized knowledge bases and inference engines [11]. For instance, a web-based expert system for identifying pests and diseases in cocoa plants was developed using Forward Chaining and Backward Chaining methods, enabling the system to automatically provide treatment solutions [12]. In addition, the Certainty Factor method has been applied in cocoa disease diagnosis systems and achieved an accuracy rate of 85.7 percent, consistent with expert diagnoses [13]. The effectiveness of Fuzzy Logic in expert systems for detecting pests and diseases in cocoa has also been demonstrated in other studies [14]. Although expert systems offer structured solutions, such systems are generally rigid, difficult to expand or scale, and require continuous manual updates to handle variations of symptoms or newly emerging diseases. Furthermore, their accuracy is highly dependent on the completeness and correctness of the knowledge base entered manually [15].

With the availability of digital imagery, many studies have shifted toward image processing and conventional machine learning approaches [16]. These methods involve manual feature extraction (e.g., color, texture, shape) from disease images, followed by classification using algorithms such as support vector machine (SVM), K-nearest neighbors (KNN), or random forest [17]. However, the manual feature extraction process is often time-consuming, requires domain expertise, and may not be optimal for capturing the complex features that distinguish different diseases [18].

The emergence of deep learning, particularly CNNs, has revolutionized the field of computer vision and plant disease detection [19]. CNNs possess the unique ability to automatically learn hierarchical feature representations directly from raw image data, eliminating the need for manual feature extraction [20]. Pujiati and Rochmawati applied CNNs for the identification of herbal plant leaf images, demonstrating the capability of CNNs in plant image classification [21]. Similar studies have also shown the effectiveness of CNN methods across various image classification tasks, such as object classification in satellite imagery [22] and disease recognition from leaf images, confirming CNNs as a consistently accurate approach within the domain of visual classification [23].

Although CNN architectures including VGG-16 and ResNet-50 demonstrate strong performance in plant disease identification, the majority of research centers on leaf-based symptoms [24]. However, vascular conditions such as VSD initially develop within stem vascular structures, establishing a substantial research void. Currently, no convolutional neural network studies address cocoa stem VSD detection, notwithstanding its role in generating 30-50% yearly production losses throughout Indonesian cocoa plantations [1]. Our research bridges this gap through four key contributions, including first resource-efficient CNN developed for cocoa stem VSD classification with real-time mobile/drone compatibility, an original cocoa stem imagery collection, comprehensive evaluation of convolutional layer configurations [25], and a TensorFlow Lite (TFLite) implementation supporting precision agriculture among small-scale growers.

2.2. Method

The research methodology was systematically designed to develop and evaluate a deep learning model for the early detection of VSD in cocoa stems, with a specific focus on comparing CNN architectures with different numbers of convolutional layers. The stages of the study are illustrated in Figure 1.

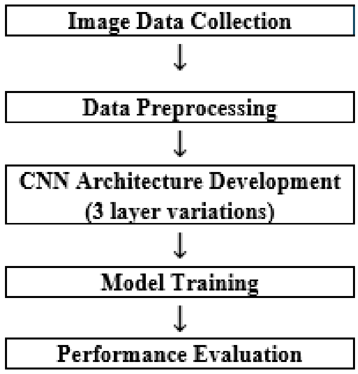

Figure 1. Research methodology flowchart for VSD detection in cocoa stems

The research process began with image data collection. The image dataset served as the primary foundation for developing the deep learning model; therefore, cocoa stem images were collected from various plantations across Indonesia to ensure diversity in environmental conditions and plant genetics. Data collection was conducted using a high-resolution digital camera to capture the necessary visual details. The dataset consisted of two main classes: healthy cocoa stems showing no signs of VSD or other diseases, and VSD-infected cocoa stems exhibiting symptoms ranging from early stages (slight discoloration, minor streaks) to intermediate stages (more pronounced necrosis and initial dieback). A total of 215 images were collected, comprising 109 images of healthy stems and 106 images of VSD-infected stems.

After the data collection was completed, the next step was data preprocessing, which is a crucial stage to enhance data quality and improve efficiency in model training. To address the limitation of dataset size and improve model generalization, data augmentation techniques were applied. This augmentation helped the model become more robust against variations in orientation and lighting conditions. Furthermore, pixel values were normalized from the range [0, 255] to [0, 1] to accelerate training convergence and improve numerical stability. In certain cases, simple segmentation techniques (e.g., background masking) were applied to ensure that the model focused exclusively on the cocoa stem area, thereby reducing environmental noise. Finally, the dataset was divided into two subsets: training (80 percent) and testing (20 percent). The training set was used to train the model, while the testing set was employed for the independent evaluation of model performance.

To address the limited size of the dataset and improve model generalization, data augmentation was performed using the ImageDataGenerator framework. The applied transformations included rotation (±20°), width and height shifts (0.2), shear transformation (0.2), zoom scaling (0.2), and horizontal flipping. Additionally, feature-wise normalization techniques, including centering and standard deviation normalization, were applied to standardize input distributions.

It is important to note that augmented data were generated prior to train-test splitting, which may introduce potential data leakage. While this approach was adopted to increase dataset size, it represents a methodological limitation. Future work should apply augmentation exclusively to the training set to ensure strict separation between training and evaluation data.

The core stage of the research was the development of the CNN model. The CNN was developed using the TensorFlow/Keras deep learning framework. In this study, three variations of CNN architectures with different numbers of convolutional layers were employed, namely Model A (3 convolutional layers), Model B (7 convolutional layers), and Model C (11 convolutional layers). The proposed CNN architecture consists of seven convolutional blocks with progressively increasing and then decreasing filter sizes (32, 64, 128, 256, 256, 128, 64). Each convolutional layer employs a 3×3 kernel with ReLU activation, followed by max pooling for spatial downsampling. This design enables efficient hierarchical feature extraction while controlling model complexity. The convolutional layers are followed by a fully connected layer with 1024 neurons and a dropout rate of 0.5 to mitigate overfitting. The output layer utilizes a softmax activation function for binary classification.

Each convolutional layer was followed by a pooling layer. The model was further modified by adding fully connected layers and dropout at the end to perform binary classification (healthy/VSD). Specifically, the CNN architecture consisted of convolutional layers used to extract hierarchical features from the images with [3x3] filters and the ReLU activation function (max(0, x)). Subsequently, pooling layers were applied to reduce spatial dimensions and make the model more invariant to small translations, using [2x2] max pooling. The output from the convolutional and pooling layers was flattened into a one-dimensional vector, which was then fed into fully connected (dense) layers to perform classification based on the extracted features. A dropout layer was applied to prevent overfitting by randomly deactivating a fraction of neurons during training. Finally, the output layer was a dense layer with a single neuron and a Softmax activation function $Z_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}$ for binary classification.

After the model was developed, the next step was model training. Each model variation was trained using the Adam optimization algorithm with an initial learning rate of 0.001. The Binary Cross-Entropy loss function was employed, as the task represented a binary classification problem. Training was conducted for 100 epochs with a batch size of 32. An early stopping technique was implemented to halt training when performance on the validation dataset did not improve over several consecutive epochs, thereby preventing overfitting.

The performance of each model variation was then evaluated using the test dataset, which had not been seen by the model during training. Model performance was assessed using accuracy as the primary metric, defined as the proportion of correct predictions to the total number of predictions. In addition, a confusion matrix was employed to visualize model performance, presenting the number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN).

The evaluation results were analyzed in depth to understand the strengths and weaknesses of each model configuration, particularly in relation to the number of convolutional layers. This analysis included error analysis (e.g., examining the causes of FP and FN) to identify areas for improvement. Comparisons with conventional VSD detection methods or other machine learning models were also conducted to highlight the advantages of the deep learning approach. In addition to overall accuracy, the model performance was evaluated using precision, recall (sensitivity), specificity, and F1-score. These metrics are particularly important in disease detection tasks, where minimizing FN is critical to prevent undetected disease spread.

3. RESULTS AND DISCUSSION

3.1. Result

The CNN models developed with varying numbers of convolutional layers demonstrated different levels of performance in classifying cocoa stem images as either healthy or VSD-infected. Table 1 summarizes the key performance metrics for each model configuration on the test dataset.

Table 1. Key performance metrics of each model

Metric	Model A (3 Layers)	Model B (7 Layers)	Model C (11 Layers)
Akurasi	99.16%	99.16%	53.22%

These results indicate that Model A with 3 convolutional layers and Model B with 7 convolutional layers achieved the highest accuracy compared to Model C (11 layers). The accuracy of 99.16 percent for both Model A and Model B suggests that more than nine out of ten model predictions were correct. The confusion matrix in Figure 2 provides a detailed overview of the classification performance of Model A (3 layers), while Figure 3 presents the classification performance of Model B (7 layers), which represent the best-performing models.

Based on the confusion matrices of Model A and Model B, it can be observed that both models achieved identical performance with a very high number of correct predictions. Out of 190 healthy plant samples, all were correctly classified without any FP cases, meaning no healthy plants were mistakenly detected as VSD-infected. Meanwhile, among 167 samples of truly VSD-infected plants, 164 cases were correctly detected as true positives, while only 3 cases were misclassified as healthy FN. The low FN value indicates that the models were able to minimize the risk of missed disease detection, which is particularly critical in the context of VSD management since undetected cases can lead to uncontrolled disease spread. Thus, both models demonstrated excellent performance with high accuracy, no FP errors, and only a very small number of FN errors.

Figure 4 illustrates the accuracy and loss curves during the training and validation processes for Model A (3 layers), while Figure 5 presents the accuracy and loss curves for Model B (7 layers), both of which demonstrate good convergence.

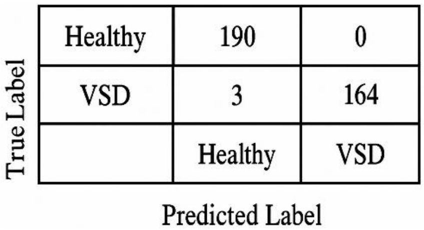

Figure 2. Confusion matrix of model A

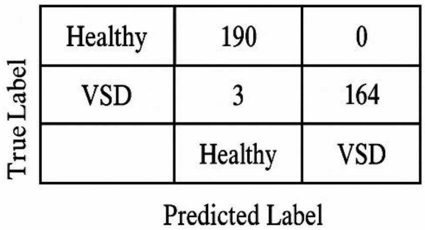

Figure 3. Confusion matrix of model B

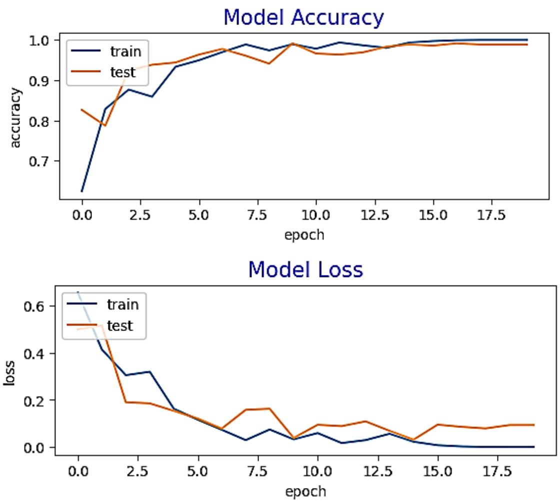

Figure 4. Training and validation accuracy and loss curves of model A (3 Layer CNN)

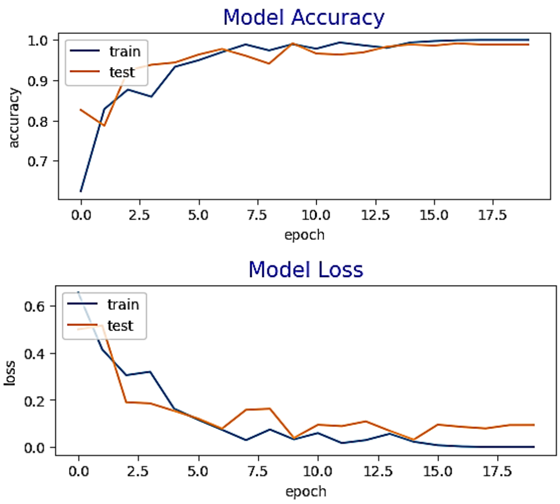

Figure 5. Training and validation accuracy and loss curves of model B (7 Layer CNN)

From Figures 4 and 5, it can be observed that training and validation accuracy increased steadily with the number of epochs, while loss decreased. The good convergence between the training and validation curves indicates that the models did not suffer from severe overfitting and were able to generalize well to unseen data. The consistent increase in accuracy and decrease in loss further suggest an effective learning process.

The figures above confirm the ability of the model to accurately identify the condition of cocoa stems. In addition to classification, the model also provides confidence levels (probabilities) for its predictions, which can serve as valuable additional information for users. Figure 6 illustrates the correct classification of a healthy stem, while Figure 7 clearly demonstrates the model's capability in detecting and identifying areas of VSD infection on the stem.

The findings of this study clearly demonstrate that the developed CNN models are highly effective for the early detection of VSD in cocoa stems. The main focus of this discussion is the comparison of model performance based on variations in the number of convolutional layers, which represents a critical aspect of CNN architectural design. As shown in Table 1, Model A with 3 convolutional layers and Model B with 7 convolutional layers achieved the best performance with an accuracy of 99.16%. In contrast, Model C with 11 convolutional layers exhibited a decline in performance compared to Model A and Model B.

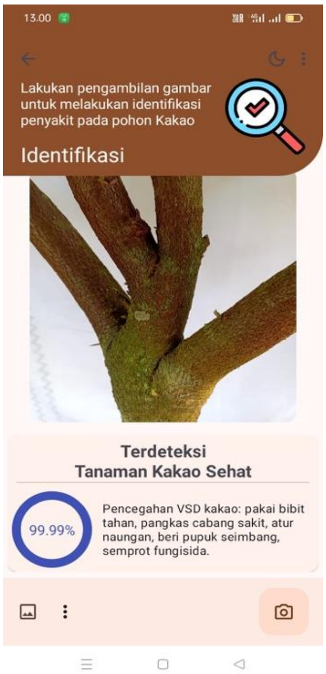

Figure 6. Example of correct classification of a healthy cocoa stem

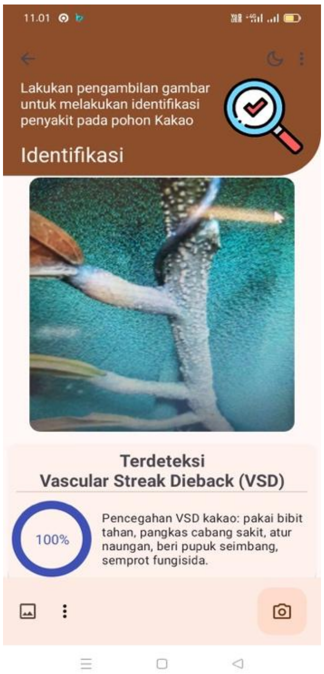

Figure 7. Example of correct classification of a VSD infected cocoa stem

3.2. Discussion

Based on the results presented in Table 1, Model A (3 convolutional layers) and Model B (7 convolutional layers) demonstrated high classification accuracy of 99.16%, whereas Model C (11 convolutional layers) achieved only 53.22%. The strong performance of Model A indicates that an architecture with relatively shallow depth is sufficient to extract discriminative features for distinguishing healthy stems from VSD-infected stems. Our lightweight architecture not only achieves higher accuracy than fine-tuned VGG16 and ResNet50 on the same dataset, but also delivers 28x faster mobile inference, establishing the first practical solution for real-time cocoa stem VSD detection in field conditions. Model B maintained a comparable accuracy, suggesting that increasing the depth to seven layers does not provide significant advantages over Model A. In contrast, the performance decline in Model C suggests that excessive depth may lead to overfitting and reduced generalization ability, particularly when trained on a dataset that is limited in size and diversity [26].

Further analysis of the confusion matrices reinforces these findings. As shown in Figures 2 and 3, both Model A and Model B produced a minimal number of FN, with only 3 cases, and no FP. This condition is particularly critical in the context of disease detection, as FN poses a more serious risk than FP, given that undetected cases can potentially lead to uncontrolled disease spread [27]. In contrast, the low accuracy of Model C indicates a higher rate of misclassification in both classes. Overall, Model A and Model B demonstrated an optimal balance between architectural complexity, computational efficiency, and diagnostic reliability, making them suitable for implementation in computer vision based VSD detection systems [28].

Regardless of the number of layers, the deep learning approach offers several significant advantages over conventional visual inspection methods. First, its accuracy and objectivity are much higher. Unlike manual visual inspection, which is prone to subjectivity and human fatigue, CNN models provide objective and consistent diagnoses. The accuracy of 99.16% far surpasses the detection capability of the naked eye, particularly in the early stages of the disease. Second, its early detection capability is crucial, as the model can identify VSD in stems even at initial stages where symptoms may not be clearly visible to humans. Early detection enables rapid interventions, such as sanitary pruning or fungicide application, which can prevent the spread of the disease and save plants. Third, its efficiency and scalability are very high. Once trained, the model can perform detection instantly, allowing rapid monitoring of large plantation areas, which would be impossible through manual inspection. The potential integration into drone-based systems would further enhance monitoring efficiency. Finally, CNNs have the ability of automatic feature learning, meaning they can autonomously learn discriminative features from images, eliminating the need for time-consuming manual feature engineering that requires domain expertise.

Although the performance of Model A and Model B was very strong, there remains room for improvement. The confusion matrix analysis revealed a small number of FP and FN. FP occurred when healthy stems were classified as VSD-infected, which may have been caused by visual artifacts on the stems (e.g., moss, dirt, shadows) that resembled VSD symptoms, or by natural variations in stem appearance that confused the model. Reducing FP is important to avoid unnecessary actions such as pruning on healthy plants. On the other hand, FN refer to cases in which VSD-infected stems were classified as healthy. This type of error is more critical, as it may result in undetected disease spread. FNs are likely to occur in very early-stage VSD cases with extremely subtle symptoms, or in images of poor quality (e.g., blurred or poorly lit) that obscure the symptoms.

In addition, the authors acknowledge that the experimental design in this study has certain methodological limitations. The dataset size is relatively small (215 images), and data augmentation was applied before the train–test split. This setup may raise concerns about potential data leakage and the reliability of the reported accuracy, especially since Model A and Model B achieve identical test accuracy of 99.16%. The use of only a single 80/20 split without additional validation (such as cross-validation or a separate validation set) also limits the robustness of the performance assessment. Therefore, the high accuracy should be interpreted with caution and may be somewhat optimistic in the context of limited data and evaluation depth.

To address these limitations, future research can focus on enhancing the dataset by collecting larger and more diverse data, encompassing a wider range of VSD symptoms, lighting conditions, backgrounds, and cocoa varieties. Incorporating images with very early or subtle symptoms would be particularly beneficial. In addition, the application of more advanced image enhancement techniques or more robust segmentation algorithms could help clean images from noise and emphasize relevant areas. Exploring deeper CNN architectures or employing ensemble learning techniques (combining multiple models) may also improve accuracy and robustness. Finally, the use of explainable AI (XAI) techniques such as Grad-CAM could help visualize the image regions most attended to by the model during prediction, providing valuable insights into why the model makes certain errors and identifying confusing features.

The potential implementation of this model in mobile applications could revolutionize how farmers manage the health of their cocoa plants. With smartphones equipped with this model, farmers would be able to independently and rapidly diagnose VSD, enabling them to take timely preventive measures. Furthermore, integration with drone-based systems could enable large-scale automated plantation monitoring, providing comprehensive plant health maps and identifying disease hotspots. This represents a significant step forward toward precision agriculture in the cocoa sector, ultimately enhancing both productivity and sustainability.

4. CONCLUSION

This study successfully developed and evaluated CNN models that are highly effective for the early detection of VSD in cocoa stems. Through the comparison of architectures with varying numbers of convolutional layers, it was found that the models with 3 convolutional layers (Model A) and 7 convolutional layers (Model B) achieved optimal performance with an average accuracy of 99.16% on the independent testing dataset. These results strongly support the hypothesis that deep learning approaches, particularly CNNs, provide an accurate, objective, and efficient method for VSD identification, surpassing the limitations of conventional visual inspection methods. This automated early detection capability holds transformative implications for cocoa farmers. By enabling rapid and accurate identification of infected plants, farmers can take proactive mitigation measures such as sanitary pruning or the application of biocontrol agents, thereby significantly reducing disease spread and minimizing yield losses. This represents an important contribution to precision agriculture practices and the sustainability of cocoa production. Furthermore, the model can be integrated into mobile applications for real-time field diagnosis and into drone-based systems for large-scale automated monitoring, enabling rapid intervention by farmers and agricultural extension workers.

ACKNOWLEDGMENTS

This research would not have been possible without the financial support and facilities provided. Therefore, we would like to express our deepest gratitude to Telkom University for the research funding that enabled the completion of all stages of this study. We also extend our appreciation to all parties who contributed to data collection and provided valuable insights.

FUNDING INFORMATION

This work was supported by Directorate of Research and Community Service, Telkom University. internal grant for the year 2024-2025.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Dewi Marini Umi Atmaja	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	
Arif Rahman Hakim	✓		✓			✓		✓	✓		✓	✓		
Fadhil Rozi Hendrawan		✓		✓	✓		✓				✓		✓	✓
Komang Diah Devi Pramesty							✓			✓			✓	
Naufal Fadhilah Fitrah					✓					✓		✓		✓
Faiz Rochmatullah Widhaputra				✓							✓			✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

INFORMED CONSENT

We have obtained informed consent from all individuals included in this study.

ETHICAL APPROVAL

Ethical approval was not required for this study because the research involved plant materials only and did not involve human or animal subjects.

DATA AVAILABILITY

Derived data supporting the findings of this study are available from the corresponding author [initials, DMUA] on request.

REFERENCES

[1] A. K. Parawansa *et al.*, "Severity and changed symptoms of vascular streak dieback caused by Ceratobasidium theobromae in several cocoa (Theobroma cacao) clones in Sulawesi, Indonesia," *Tropical Plant Pathology*, vol. 47, no. 6, pp. 754–769, Aug. 2022, doi: 10.1007/s40858-022-00526-9.
[2] G. J. Samuels *et al.*, "Vascular streak dieback of cacao in Southeast Asia and Melanesia: in planta detection of the pathogen and a new taxonomy," *Fungal Biology*, vol. 116, no. 1, pp. 11–23, Jan. 2012, doi: 10.1016/j.funbio.2011.07.009.
[3] S. S. Ali *et al.*, "Draft genome sequence of fastidious pathogen Ceratobasidium theobromae, which causes vascular-streak dieback in Theobroma cacao," *Fungal Biology and Biotechnology*, vol. 6, no. 1, 2019, doi: 10.1186/s40694-019-0077-6.
[4] J. Trisno, R. Reflin, and M. Martinius, "Vascular streak dieback: new cocoa plant disease in West Sumatra," *Jurnal Fitopatologi Indonesia*, vol. 12, no. 4, p. 142, Nov. 2016, doi: 10.14692/jfi.12.4.142.
[5] A. Widiastuti, A. Wibowo, A. B. Prakoso, and Hendra, "Current status of emerging vascular-streak dieback (VSD) on cacao in Yogyakarta, Indonesia," in *Proceeding of the 2nd International Conference on Tropical Agriculture*, Springer International Publishing, 2018, pp. 133–139.
[6] L. O. S. Bande, H. S. Gusnawaty, A. Hasan, and M. Arba, "The role of leaf nutrients in vascular streak dieback disease in cocoa plants (Theobroma cacao L.)," *Indian Journal Of Agricultural Research*, no. Of, 2023, doi: 10.18805/ijare.af-798.
[7] L. Deng, *Deep learning: methods and applications*. now Publishers Inc, 2014.
[8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017, doi: 10.1145/3065386.
[9] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, "Deep neural networks based recognition of plant diseases by leaf image classification," *Computational Intelligence and Neuroscience*, vol. 2016, pp. 1–11, 2016, doi: 10.1155/2016/3289801.

[10] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Frontiers in Plant Science*, vol. 7, 2016, doi: 10.3389/fpls.2016.01419.
[11] I. S. Pakpahan, I. S. Dumayanty, and S. Manurung, "Expert system for chili plant disease detection using the Naive Bayes method," *METHOTIKA: Jurnal Ilmiah Teknik Informatika*, vol. 4, no. 2, pp. 10–20, 2024.
[12] A. Adelino *et al.*, "Design and construction of an expert system for identification of pests and diseases in cocoa plants using web-based forward chaining and backward chaining methods," *Informatika*, vol. 16, no. 1, pp. 200–208, 2024.
[13] L. Meniati, N. Y. Lumban Gaol, and I. Santoso, "Expert system for diagnosing cocoa plant diseases using the certainty factor method," *J-SISKO TECH (Jurnal Teknologi Sistem Informasi dan Sistem Komputer TGD)*, vol. 5, no. 1, p. 83, Feb. 2022, doi: 10.53513/jsk.v5i1.4798.
[14] F. P. Soge, B. Deta, and D. B. Watomakin, "Application of fuzzy logic in an expert system for detecting pests and diseases of cocoa plants in Wulanggitang District, East Flores," *Jurnal Kridatama Sains dan Teknologi*, vol. 6, no. 1, pp. 212–234, 2024.
[15] P. M. Prihatini, "Uncertainty and vagueness methods in expert systems," *Lontar Komputer*, vol. 2, no. 1, pp. 29–42, 2011.
[16] I. Jawaz and R. Rahmadewi, "Pulmonary pneumonia detection system with digital image processing and machine learning," *ELECTRON Jurnal Ilmiah Teknik Elektro*, vol. 5, no. 1, May 2024, doi: 10.33019/electron.v5i1.114.
[17] B. K. Hatuwal, A. Shakya, and B. Joshi, "Plant leaf disease recognition using random forest, KNN, SVM and CNN," *Polibits*, vol. 62, pp. 13–19, 2020.
[18] H. Kode and B. D. Barkana, "Deep learning- and expert knowledge-based feature extraction and performance evaluation in breast histopathology images," *Cancers*, vol. 15, no. 12, p. 3075, 2023, doi: 10.3390/cancers15123075.
[19] B. Tugrul, E. Elfatimi, and R. Eryigit, "Convolutional neural networks in detection of plant leaf diseases: a review," *Agriculture*, vol. 12, no. 8, p. 1192, Aug. 2022, doi: 10.3390/agriculture12081192.
[20] N. Yemi Hussain, "Deep learning architectures enabling sophisticated feature extraction and representation for complex data analysis," *International Journal of Innovative Science and Research Technology (IJISRT)*, pp. 2290–2300, Nov. 2024, doi: 10.38124/ijisrt/ijisrt24oct1521.
[21] R. Pujiati and N. Rochmawati, "Identification of herbal plant leaf images using the convolutional neural network (CNN) method," *Journal of Informatics and Computer Science (JINACS)*, vol. 3, no. 03, pp. 351–357, Jan. 2022, doi: 10.26740/jinacs.v3n03.p351-357.
[22] M. K. Basyir and S. Supiyandi, "Dog breed classification using convolutional neural networks on digital images," *Journal Of Social Science Research*, vol. 5, no. 1, pp. 99–109, 2025.
[23] A. Julianto and A. Sunyoto, "A performance evaluation of convolutional neural network architecture for classification of rice leaf disease," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 10, no. 4, p. 1069, Dec. 2021, doi: 10.11591/ijai.v10.i4.pp1069-1078.
[24] T. D. Salka, M. B. Hanafi, S. M. S. A. A. Rahman, D. B. M. Zulperi, and Z. Omar, "Plant leaf disease detection and classification using convolution neural networks model: a review," *Artificial Intelligence Review*, vol. 58, no. 10, 2025, doi: 10.1007/s10462-025-11234-6.
[25] S. S. Reka, V. D. Murthy Voona, P. V. Sai Nithish, D. S. Paavan Kumar, P. Venugopal, and V. Ravi, "Performance analysis of deep convolutional network architectures for classification of over-volume vehicles," *Applied Sciences*, vol. 13, no. 4, p. 2549, Feb. 2023, doi: 10.3390/app13042549.
[26] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *The journal of machine learning research*, vol. 15, no. 1, pp. 1929–1958, 2014.
[27] R. O. Duda, P. E. Hart, and D. G. Stork, "Pattern classification," *Second Edition*, vol. 16, no. 2, 2001.
[28] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

BIOGRAPHIES OF AUTHORS

Dewi Marini Umi Atmaja is a lecturer in the Information System Study Program, Center of Excellence for Human Centric Engineering (HUMIC), Telkom University, Jakarta Campus. She obtained her Bachelor's degree from Jenderal Achmad Yani University in 2018 and completed her Master's degree at President University in 2020. Throughout her academic career, she has published several accredited national and international journal articles focusing on Artificial Intelligence (AI) and Information Systems. For further correspondence. She can be contacted at email: dewimariniumiatmaja@telkomuniversity.ac.id.

Arif Rahman Hakim is a lecturer and researcher at Telkom University with academic and professional interests in information technology, software engineering, and digital innovation. He obtained his Bachelor's degree in Informatics Engineering from STT Pelita Bangsa in 2017 and completed his Master's degree in Informatics Engineering at President University in 2020. In addition to his academic activities, he is actively involved in research, scientific publication, and academic writing in the field of computer science and information systems. His scholarly contributions are indexed through his Google Scholar profile, reflecting his engagement in interdisciplinary research and higher education development. He can be contacted at email: arifrahmanarh@telkomuniversity.ac.id.

Fadhil Rozi Hendrawan is a lecturer in the S1 Information Systems program at Telkom University, Jakarta Campus. His interests include Data Governance, IT Governance, Project Management, Knowledge Management, and related aspects of IT security. He has authored 13 publications, including book chapters and conference proceedings. He holds both his bachelor's and master's degrees in Information Systems from Telkom University. Fadhil is active in academic publishing and community engagement; his works indexed on SINTA and Google Scholar cover themes such as digital and cybersecurity literacy, and he frequently shares teaching, mentorship, and community-service activities through professional channels as a Telkom University Jakarta faculty member. He can be contacted at email: fadhilrozihendrawan@telkomuniversity.ac.id.

Komang Diah Devi Pramesty is an undergraduate student in the Information Systems Study Program with a strong interest in Artificial Intelligence (AI), data processing, and information systems. She is passionate about exploring how data and intelligent technologies can be applied to support decision-making and improve organizational processes. Through her academic journey, she has developed skills in data management and the use of analytical tools to gain insights from data. Her enthusiasm for learning and innovation reflects her commitment to advancing technology-based solutions in the field of information systems. For further correspondence. She can be contacted at email: diahdevipramestjie@student.telkomuniversity.ac.id.

Naufal Fadhilah Fitrah is an undergraduate student in the Information Systems Study Program, Faculty of Industrial Engineering, Telkom University. He is currently pursuing his Bachelor's degree with a focus on technology, data analysis, and information systems development. During his studies, he has been involved in several academic projects and research activities related to system design, data analytics, and digital transformation. His academic interests include information system development, data-driven decision making, and project management. For further correspondence. He can be contacted at email: naufalfadhilahfitrah@telkomuniversity.ac.id.

Faiz Rochmatullah Widhaputra is an energetic and passionate student currently pursuing a degree in Information Systems at Telkom University. He strives to develop his skills in Programming, Graphic Design, Web Design, and Front-End Development. Working while studying is one of his principles. He is currently working as a teacher at an Islamic foundation, interning as an operator at a construction consulting firm, and has also worked as an advertiser at an Umrah consulting company. His dream is to pursue a master's degree in Japan while spreading Islam there (doing da'wah). One of the guiding principles he lives by is: "The best of people are those who are most beneficial to others." (خَيْرُ النَّاسِ أَنْفَعُهُمْ لِلنَّاسِ). He can be contacted at email: faizrochmatullah@student.telkomuniversity.ac.id.