Salud, Ciencia y Tecnología. 2026; 6:2641
doi: 10.56294/saludcyt20262641

ORIGINAL

Lettuce Plant Disease Recognition Using Android-Based CNN Algorithm Method

Reconocimiento de enfermedades en plantas de lechuga mediante un algoritmo CNN basado en Android

RZ Abdul Aziz[1], MS Hasibuan[1], Nanda Satria Putra[1]

[1]Magister Informatics Engineering, Faculty Computer Science, Institute Informatics and Business Darmajaya, Bandar Lampung, Indonesia.

Cite as: Aziz RA, Hasibuan M, Putra NS. Lettuce Plant Disease Recognition Using Android-Based CNN Algorithm Method. Salud, Ciencia y Tecnología. 2026; 6:2641. https://doi.org/10.56294/saludcyt20262641

Submitted: 17-08-2025 **Revised:** 25-10-2025 **Accepted:** 03-12-2025 **Published:** 01-01-2026

Editor: Prof. Dr. William Castillo-González

Corresponding author: RZ Abdul Aziz

ABSTRACT

Introduction: disease detection in lettuce (Lactuca sativa L.) is crucial to enhance crop yields and prevent losses caused by bacterial, fungal, and weed-related infections. This study aimed to develop an Android-based lettuce disease detection application using a Convolutional Neural Network (CNN) algorithm to assist farmers in identifying plant diseases in real time.
Method: the research used a dataset of 2,320 lettuce leaf images obtained from Kaggle, categorized as healthy, bacterial, fungal, and shepherd's purse weed. The dataset was preprocessed through labeling, normalization, and augmentation to improve model robustness. The CNN architecture comprised four convolution layers followed by max-pooling, dense, and softmax output layers. The model was trained using TensorFlow and deployed through TensorFlow Lite for mobile implementation.
Results: the CNN model achieved 93,67 % training accuracy and 93,99 % validation accuracy, demonstrating good generalization without overfitting. The evaluation using confusion matrix and classification reports showed high performance, particularly in identifying healthy and shepherd's purse weed categories with F1-scores of 0,94 and 0,99, respectively. The Android application successfully detected diseases in real time and provided users with diagnostic results, historical data, and treatment suggestions.
Conclusions: the developed CNN-based Android application proved effective for automatic lettuce disease detection with high accuracy and practical usability for farmers. Future studies could enhance performance through more advanced CNN architectures such as VGG16 or ResNet50 and the use of more detailed datasets for improved disease classification.

Keywords: Plant Disease Detection; Convolution Neural Network; Tensorflow Lite; Android; Lettuce.

RESUMEN

Introducción: la detección de enfermedades en la lechuga (Lactuca sativa L.) es crucial para mejorar el rendimiento de los cultivos y prevenir pérdidas causadas por infecciones bacterianas, fúngicas y de malezas. Este estudio tuvo como objetivo desarrollar una aplicación para Android de detección de enfermedades en lechuga mediante un algoritmo de Red Neuronal Convolucional (CNN) para ayudar a los agricultores a identificar enfermedades de las plantas en tiempo real.
Método: se utilizó un conjunto de datos de 2,320 imágenes de hojas de lechuga obtenidas de Kaggle, categorizadas como sanas, con bacterias, hongos y con la maleza conocida como bolsa de pastor. El conjunto de datos se preprocesó mediante etiquetado, normalización y aumento de datos para mejorar la robustez del modelo. La arquitectura de la CNN consta de cuatro capas convolucionales seguidas de capas de salida de max-pooling, densas y softmax. El modelo se entrenó con TensorFlow y se implementó mediante TensorFlow Lite para su uso en dispositivos móviles.

Resultados: el modelo CNN alcanzó una precisión de entrenamiento del 93,67 % y una precisión de validación del 93,99 %, demostrando una buena generalización sin sobreajuste. La evaluación mediante la matriz de confusión y los informes de clasificación mostró un alto rendimiento, especialmente en la identificación de las categorías de lechuga sana y con la presencia de la maleza conocida como bolsa de pastor, con puntuaciones F1 de 0,94 y 0,99, respectivamente. La aplicación para Android detectó con éxito enfermedades en tiempo real y proporcionó a los usuarios resultados de diagnóstico, datos históricos y sugerencias de tratamiento.
Conclusiones: la aplicación para Android basada en redes neuronales convolucionales (CNN) desarrollada demostró ser eficaz para la detección automática de enfermedades en lechuga, con alta precisión y gran utilidad práctica para los agricultores. Estudios futuros podrían mejorar el rendimiento mediante arquitecturas CNN más avanzadas, como VGG16 o ResNet50, y el uso de conjuntos de datos más detallados para una mejor clasificación de las enfermedades.

Palabras clave: Detección de Enfermedades en Plantas; Red Neuronal Convolucional; Tensorflow Lite; Android; Lechuga.

INTRODUCTION

Lettuce is one of the horticultural crops that has high economic value and is widely consumed as fresh vegetables in Indonesia.[1] However, lettuce production often decreases due to the attack of various diseases, such as bacterial, fungal infections, and disorders caused by shepherd purse weeds. In the region of Sahabat Hidroponik Lampung, farmers often face similar problems that hamper their yields. Early detection of diseases in lettuce plants is an important solution to prevent the wider spread of diseases and increase agricultural productivity.[2]

Several studies have been conducted on lettuce disease detection using image processing and machine learning techniques. For example, Kumar et al. developed a deep learning-based approach using a pre-trained ResNet model to classify Plant diseases with an accuracy of 96,49 %.[3] Similarly, Zhou et al proposed SVM model for detecting hydroponic lettuce: leafroll and brown blotch disease (BBD) with accuracy 93 %.[4] Moreover, Sanida et al. This study developed a hybrid CNN model that achieved a testing accuracy of 99,17 % in classifying tomato leaf diseases.[5] In addition, a study by Zhou et al. demonstrated the effectiveness of using hyperspectral imaging combined with deep learning for early detection of compound in lettuce leaves [6]. However, three key limitations remain:

- many existing models focus on tomato, rice, or other major crops, with limited attention to lettuce;
- studies on lettuce either use heavy computational models or require laboratory-level imaging (e.g., hyperspectral), making them unsuitable for practical field use; and
- very few solutions are optimized for real-time mobile deployment, resulting in a gap between research prototypes and tools accessible to farmers.

The development of artificial intelligence (AI) and Machine Learning technology allows the application of image classification models to automatically detect plant diseases.[7] In this study, an Android-based application was developed[8] that utilizes the Convolutional Neural Network (CNN) algorithm to detect diseases in lettuce leaves.[9] The dataset used consists of 2320 images obtained from Kaggle and includes healthy categories, bacteria, fungi, and shepherd purse weeds, the model is trained using TensorFlow[10] and implemented into TensorFlow Lite to run efficiently on mobile devices.[11]

With this application, farmers can easily upload or take pictures of lettuce leaves using a cell phone camera to get real-time disease detection results.[12,13] This application also provides information related to the detected disease.[14]

Given the rapid progress in lightweight deep learning models and mobile computing, there is a strong need for an efficient, field-deployable system capable of detecting lettuce diseases directly through smartphones. The Convolutional Neural Network (CNN) offers powerful capabilities in image-based classification while remaining suitable for integration with TensorFlow Lite for mobile environments.

METHOD

This study employed an applied experimental research approach designed to develop and evaluate a mobile-based system for automatic lettuce disease detection using deep learning. The investigation was conducted in two primary settings: (1) a computational environment for model development, training, and performance evaluation using Python and TensorFlow, and (2) a practical agricultural context at Sahabat Hidroponik Lampung, where farmers' experiences and real plant conditions were used to validate the relevance of the application. The research workflow consisted of four major stages—data collection, preprocessing, algorithm development, and model training/testing—each aimed at ensuring that the resulting CNN-based Android application could operate reliably under real-world conditions. These stages structured the methodological foundation of the study and guided the development of the disease detection system, as illustrated in figure 1.

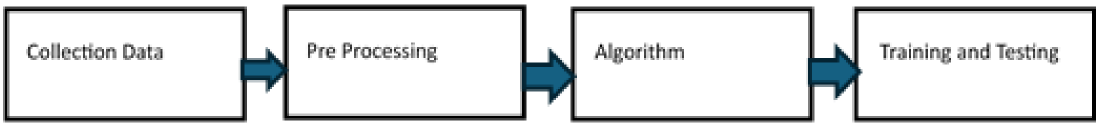

Figure 1. Research Method

Data Collection

Data collection in this study was carried out using the dataset available on the Kaggle platform, which consists of 2320 images of lettuce leaves with healthy, bacterial, fungal, and shepherd's purse weeds categories.(15) In addition, data was also obtained through interviews with hydroponic farmers in Sahabat Hidroponik Lampung to understand the problems they face related to lettuce plant diseases. Direct observation of the condition of infected lettuce leaves was also conducted to ensure the suitability of the data used in model training.(16) The collected data then went through a preprocessing process, including image labeling and augmentation, to increase data variation and improve model performance in detecting lettuce plant diseases more accurately.(17)

Preprocessing Data

The data preprocessing stage includes labeling the images into four categories (healthy, bacterial, fungal, and shepherd's purse weeds), normalizing the pixel values (0-255 to 0-1) as well as augmentation such as rotation, shifting, and flipping to increase data variation.(18,19) The dataset is then divided into 80 % for training and 20 % for testing to ensure the model can learn well and avoid overfitting.(20) With optimal preprocessing, the model is expected to be more accurate in detecting lettuce plant diseases.(21)

Algorithm

This research uses Convolutional Neural Network (CNN) architecture.

Architecture CNN

The Convolutional Neural Network (CNN) architecture in this study consists of four 3x3 convolution layers with stepwise filters (256, 128, 128, 64), followed by a 2x2 MaxPooling layer. After that, the features are processed using a Flatten layer and a Dense layer with 32 neurons and a ReLU activation function. The output layer has 4 neurons with Softmax activation for multi-class classification. The model accepts 256x256x3 input images and is designed to achieve high accuracy with low risk of overfitting.

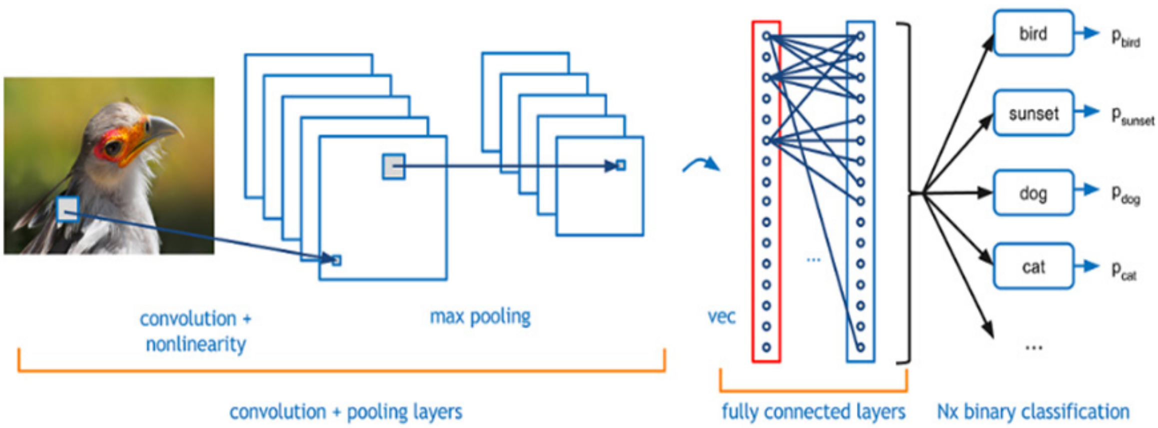

Figure 2. Architecture CNN(22)

Training and Testing

The training and testing procedures are carried out in the following stages:

Data Visualization

Data visualization is used to analyze patterns and trends in the dataset 23. The dataset used consists of 1 856 images for training and 464 images for testing.

https://doi.org/10.56294/saludcyt20262641 ISSN: 2796-9711

Figure 3. Labeling

This figure illustrates the process of labeling and categorizing lettuce leaf images into four classes: healthy, bacterial, fungal, and shepherd's purse weed. Each image in the dataset was assigned a corresponding label to facilitate supervised learning during CNN model training. The labeling ensured accurate mapping between input images and their target outputs, which is crucial for optimizing classification performance

The value of the model training procedure used in this study is as follows:

- Epochs: 50
- Accuracy Training: 93,67 %
- Loss Training: 0,1596
- Accuracy Validation: 93,99 %
- Loss Validation: 0,2625
- Learning Rate: 0,001

This model shows the highest validation accuracy of 93,99 % with a small difference between training and validation accuracy, so it has good generalization ability without signs of overfitting.

Testing Method

After the model is trained, testing is performed using a pre-separated test dataset (20 % of the total data).[24] At this stage, the model is tested to predict data labels that have never been seen before. Evaluation is done using Confusion Matrix, which shows the number of correct and incorrect predictions for each category.[25]

Evaluation

Model evaluation is done by assessing accuracy, loss, and other metrics.[26] Some of the evaluation aspects used in this research include:

- Accuracy Model: Training accuracy reached more than 92 % with validation accuracy of 93,67 %.
- Loss Model: The final validation loss value is about 0,2278, indicating that the model is not overfitting.
- Confusion Matrix: Used to evaluate classification errors and determine metrics such as precision, recall, and F1-score.

RESULTS

Model Training Performance

The CNN model was trained over 50 epochs using 1,856 training images and validated with 464 images.[27] The model achieved a training accuracy of 93,67 % and a validation accuracy of 93,99 %. The corresponding training loss reached 0,1596, while the validation loss stabilized at 0,2625, indicating consistent model convergence.

Figure 4 displays the accuracy trend for both training and validation sets across all epochs, while figure 5 presents the corresponding loss curves.

The classification report produced the following quantitative values:

Table 1. Dates

Class	Precision	Recall	F1-score	Support
Bacterial	0,90	0,86	0,88	102
Fungal	0,82	0,84	0,83	116
Healthy	0,93	0,95	0,94	119
Shepherd's purse weed	1,00	0,98	0,99	127

The overall classification accuracy achieved during testing was 91 %, with a weighted F1-score of 0,91.

https://doi.org/10.56294/saludcyt20262641 ISSN: 2796-9711

Application Testing

Application performance was evaluated using real-time trials on lettuce leaf images captured through the Android interface.

Across multiple validation attempts, accuracy values varied by disease type:

- Shepherd's purse weed detection: up to 100 % accuracy in several trials.
- Fungal disease detection: consistently high accuracy, also reaching 100 % in certain cases.
- Healthy class detection: high but variable accuracy across trials.
- Bacterial disease detection: demonstrated moderate variation, with lower accuracy compared to the other classes.

Figure 4. Graph train and validation accuracy

This figure shows the comparison of training and validation accuracy over 50 epochs. The graph indicates that the model consistently improved its accuracy during training and achieved a validation accuracy of 93,99 %. The close alignment between the two accuracy curves demonstrates that the model generalized well without significant overfitting, confirming the robustness of the CNN architecture in distinguishing lettuce disease categories.

Figure 5. Graph train and validation loss

This figure presents the trend of training and validation loss values across all epochs. The decreasing pattern of both curves signifies that the model successfully minimized prediction errors during training. The final validation loss of approximately 0,26 suggests a stable learning process and balanced performance between training and testing data, supporting the model's reliability for real-time disease detection in lettuce plants.

DISCUSSION

The CNN-based model demonstrated strong performance in classifying lettuce leaf conditions, achieving a validation accuracy of 93,99 % and an overall testing accuracy of 91 %. These results indicate that the proposed model is capable of extracting relevant visual features necessary for distinguishing among bacterial, fungal, healthy, and shepherd's purse weed classes. Comparable studies in the literature report similar or higher accuracies when using deep learning for plant disease detection. For example, Ferentinos achieved up to 99,53 % accuracy using deep CNN architectures for various crop diseases, while Sanida et al. reported 99,17 % accuracy in tomato leaf disease classification using a hybrid CNN model.[28] Although the performance of the current model is slightly lower, its accuracy remains competitive considering it uses a relatively lightweight architecture optimized for mobile deployment.

The highest performance in this study was observed in the healthy and shepherd's purse weed categories (F1-scores of 0,94 and 0,99), suggesting that these classes contain clearer visual distinctions. This pattern aligns with prior findings by Deng et al., who noted that weed-induced symptoms tend to produce distinct texture and shape variations that are easily detectable using machine learning classifiers. In contrast, the lower F1-scores for bacterial and fungal classes (0,88 and 0,83) indicate overlapping visual characteristics. Similar challenges have been identified by Sarkar et al., who reported reduced accuracy when diseases share morphological similarities or exhibit early-stage symptoms that are difficult to distinguish.

The model's performance trends also reflect findings from hyperspectral and imaging-based research. For instance, Ban et al. demonstrated that fungal infections in lettuce present subtle spectral signatures, which may require more advanced architectures or richer datasets for improved classification. The present study, using standard RGB images, thus faces intrinsic limitations in capturing subtle disease cues. This may explain the variability observed in bacterial and fungal detection during application testing.

The Android-based implementation using TensorFlow Lite extends the contribution beyond algorithmic development. While many prior studies focus solely on model accuracy in controlled environments, fewer works translate these models into mobile-ready applications accessible for farmers. Similar mobile-based solutions in other crops—such as AI-enabled tomato disease detection systems—have shown substantial potential for real-world decision support. The present application aligns with this trend by demonstrating functional real-time detection capabilities and providing diagnostic outputs that can support on-site disease management.

Despite the promising results, certain limitations remain. The dataset, although augmented, consists of images primarily sourced from Kaggle, which may not fully represent environmental variations encountered in field conditions. Prior studies emphasize the necessity of diverse training data to improve generalizability. Additionally, deeper architectures such as VGG16, ResNet50, or InceptionV3 have been shown to improve feature extraction in plant pathology tasks, particularly when symptoms are subtle or overlapping. Hence, adopting such architectures could enhance the model's sensitivity to bacterial and fungal diseases.

Overall, the findings demonstrate that a lightweight CNN can achieve strong classification accuracy and be successfully integrated into a mobile system for real-time lettuce disease detection. The results are consistent with international literature highlighting the effectiveness of deep learning in agricultural diagnosis while underscoring the importance of dataset diversity and model complexity for further improvement.

CONCLUSIONS

The overall performance demonstrates that even a lightweight CNN can achieve competitive accuracy in plant disease detection. Similar works in the international literature report higher accuracies using deeper architectures. Ferentinos and Sanida et al. achieved accuracies above 99 % using more complex deep-learning models with millions of parameters. The slightly lower accuracy in this study is therefore expected, given that the model was optimized for mobile inference rather than computational depth.

The higher accuracy in the healthy and shepherd's purse weed classes aligns with findings by Deng et al., who reported that weed-induced disorders and normal leaf patterns have distinctive morphological markers that are easily captured by image-based learning systems. These conditions often exhibit strong texture and shape contrast, resulting in clearer CNN feature extraction.

The lower performance in bacterial and fungal classes is consistent with the observations of Sarkar et al., who found that early-stage diseases display overlapping visual symptoms that are difficult to distinguish using standard RGB imagery. Many bacterial and fungal pathogens affect lettuce with similar necrotic or chlorotic patterns, reducing intra-class separability. Ban et al. further demonstrated that fungal infections show subtle spectral differences in hyperspectral imaging that are not detectable with RGB sensors. This mechanism explains the model's classification variability for these two categories.

The successful integration of the model in an Android application contributes to ongoing global efforts to develop field-ready AI systems for agriculture. While most research focuses on controlled-lab classification accuracy, fewer studies adapt models for real-time mobile usage. The present results support the findings of Yang et al., who documented that lightweight models can be successfully deployed for practical diagnostic use in hydroponic systems.

REFERENCES

1. Nurhidayati, Machfudz M, Basit A. Yield and Nutritional Quality of Green Leafy Lettuce (Lactuca sativa L.) under Soilless Culture System Using Various Composition of Growing Media and Vermicompost Rates. Caraka Tani: Journal of Sustainable Agriculture. 2021;36(2):201-12.

2. Ban S, Tian M, Hu D, et al. Evaluation and Early Detection of Downy Mildew of Lettuce Using Hyperspectral Imagery. Agriculture (Basel). 2025;15(5):444. Available from: https://doi.org/10.3390/agriculture15050444

3. Ferentinos KP. Deep learning models for plant disease detection and diagnosis. Comput Electron Agric. 2018;145:311-8.

4. Deng W, Zhou F, Gong Z, et al. Disease Feature Recognition of Hydroponic Lettuce Images Based on Support Vector Machine. Traitement du Signal. 2022;39(2):617-25.

5. Sanida MV, Sanida T, Sideris A, et al. An Efficient Hybrid CNN Classification Model for Tomato Crop Disease. Technologies (Basel). 2023;11(1):10. Available from: https://doi.org/10.3390/technologies11010010

6. Zhou X, Sun J, Tian Y, et al. Hyperspectral technique combined with deep learning algorithm for detection of compound heavy metals in lettuce. Food Chem. 2020;321:126503. Available from: https://doi.org/10.1016/j.foodchem.2020.126503

7. Orchi H, Sadik M, Khaldoun M. On using artificial intelligence and the internet of things for crop disease detection: A contemporary survey. Agriculture (Basel). 2022;12(1):9. Available from: https://doi.org/10.3390/agriculture12010009

8. Handoyono NA, Rabiman R. Development of android-based learning application in EFI materials for vocational schools. J Phys Conf Ser. 2020;1456:012050. Available from: https://doi.org/10.1088/1742-6596/1456/1/012050

9. Yang R, Wu Z, Fang W, et al. Detection of abnormal hydroponic lettuce leaves based on image processing and machine learning. Inf Process Agric. 2023;10(1):1-10.

10. Pang B, Nijkamp E, Wu YN. Deep Learning With TensorFlow: A Review. J Educ Behav Stat. 2020;45(2):227-48.

11. Huber P, Calatroni A, Rumsch A, et al. Review on Deep Neural Networks Applied to Low-Frequency NILM. Preprints. 2021:2021040421. Available from: https://doi.org/10.20944/preprints202104.0421.v1

12. Singka P, Nakthap N, Bunnjaweht D, et al. An AI-Based Growth Assessment of Green Oak Leaf Lettuce via LINE Chatbot for Amateur Growers: A Preliminary Study. In: 2024 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT & NCON). 2024. p. 275-9.

13. Nashif S, Raihan MR, Islam MR, et al. Heart Disease Detection by Using Machine Learning Algorithms and a Real-Time Cardiovascular Health Monitoring System. World J Eng Technol. 2018;6(4):854-73.

14. Casalinuovo IA, Di Pierro D, Coletta M, et al. Application of Electronic Noses for Disease Diagnosis and Food Spoilage Detection. Sensors (Basel). 2006;6(11):1428-39.

15. Quaranta L, Calefato F, Lanubile F. KGTorrent: A Dataset of Python Jupyter Notebooks from Kaggle. In: 2021 IEEE/ACM 18th International Conference on Mining Software Repositories (MSR). 2021. p. 592-6. Available from: https://doi.org/10.1109/MSR52588.2021.00072

16. Sarkar C, Gupta D, Gupta U, et al. Leaf disease detection using machine learning and deep learning: Review and challenges. Appl Soft Comput. 2023;145:110534. Available from: https://doi.org/10.1016/j.asoc.2023.110534

17. Bauer A, Bostrom AG, Ball J, et al. Combining computer vision and deep learning to enable ultra-scale aerial phenotyping and precision agriculture: A case study of lettuce production. Hortic Res. 2019;6:70. Available from: https://doi.org/10.1038/s41438-019-0151-5

18. Nakhmani A, Tannenbaum A. A new distance measure based on generalized Image Normalized Cross-Correlation for robust video tracking and image recognition. Pattern Recognit Lett. 2013;34(3):315-21.

19. Alomar K, Aysel HI, Cai X. Data Augmentation in Classification and Segmentation: A Survey and New Strategies. J Imaging. 2023;9(2):46. Available from: https://doi.org/10.3390/jimaging9020046

20. Montesinos López OA, Montesinos López A, Crossa J. Overfitting, Model Tuning, and Evaluation of Prediction Performance. In: Multivariate Statistical Machine Learning Methods for Genomic Prediction. Springer International Publishing; 2022. p. 109-39.

21. Deng W, Zhou F, Gong Z, et al. Disease Feature Recognition of Hydroponic Lettuce Images Based on Support Vector Machine. Traitement du Signal. 2022;39(2):617-25.

22. Rizkinaswara L. domain go.id dan desa.id. aptika.kominfo. 2020 Jan 30. Available from: https://aptika.kominfo.go.id/2020/01/8312/

23. Barnett D, Arts I, Penders J. microViz: an R package for microbiome data visualization and statistics. J Open Source Softw. 2021;6(65):3201.

24. Ahmad PN, Guo J, AboElenein NM, et al. Hierarchical graph-based integration network for propaganda detection in textual news articles on social media. Sci Rep. 2025;15(1):74126. Available from: https://doi.org/10.1038/s41598-024-74126-9

25. Salih AA, Abdulazeez AM. Evaluation of Classification Algorithms for Intrusion Detection System: A Review. J Soft Comput Data Min. 2021;2(1):31-40.

26. Maxwell AE, Warner TA, Guillén LA. Accuracy assessment in convolutional neural network-based deep learning remote sensing studies—part 1: Literature review. Remote Sens (Basel). 2021;13(13):2450. Available from: https://doi.org/10.3390/rs13132450

27. Fu Y, Wang F, Shao Z, et al. On the Integration of Spatial-Temporal Knowledge: A Lightweight Approach to Atmospheric Time Series Forecasting. arXiv. 2025:2408.09695. Available from: https://arxiv.org/abs/2408.09695

28. Joseph VR, Vakayil A. SPlit: An Optimal Method for Data Splitting. Technometrics. 2022;64(2):166-76.

FINANCING
The author would like to thank Institute Informatics and Business Darmajaya, the Faculty of Computer Science who have helped and supported this research.

CONFLICT OF INTEREST
Declare potential conflicts of interest.

AUTHORSHIP CONTRIBUTION
Conceptualization: RZ Abdul Aziz.
Data curation: Nanda Satria Putra.
Formal analysis: MS Hasibuan.
Research: MS Hasibuan.
Methodology: RZ Abdul Aziz.
Project management: MS Hasibuan.
Resources: Nanda Satria Putra.
Software: Nanda Satria Putra.

Supervision: MS Hasibuan.
Validation: RZ Abdul Aziz.
Display: Nanda Satria Putra.
Drafting - original draft: MS Hasibuan.
Writing - proofreading and editing: RZ Abdul Aziz.