

Micro dataset segmentation for vanilla disease detection

Radityo H. J. Notonegoro, Dewi Agushinta R., Diana Ikasari, Rifki Kosasih

Department of Information Technology, Faculty of Computer Science and Information Technology, Universitas Gunadarma,
Depok, Indonesia

Article Info

Article history:

Received Aug 15, 2025

Revised Jul 20, 2026

Accepted Aug 12, 2026

Keywords:

Disease

Image

Preprocessing

Segmentation

Vanilla

ABSTRACT

Vanilla cultivation in Indonesia is often affected by various plant diseases, which pose significant challenges to productivity. However, images of vanilla disease data available on the internet are very limited for public access; the easily accessible disease images are mostly found in books related to vanilla or vanilla gardening. In this study, data collection involved 18 images of vanilla plants infected with basal stem rot caused by the fungus *Sclerotium rolfsii* and 20 images of healthy green vanilla plants sourced from websites. A pretrained U-Net model from “ImageNet” with a mobile network version 2 (MobileNetV2) backbone was employed, using parameters of 100 epochs, batch size 16, image size 256×256 , and learning rate 0.0001. The model was trained using four different color preprocessing methods: RGB, CIE Lab, HSV, and YCbCr. Experimental results showed that the YCbCr color representations consistently achieved the best segmentation performance on both training and validation datasets. Color representations such as CIE Lab and YCbCr can better capture contrast and lighting variations compared to RGB, resulting in a more robust and generalizable model.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Radityo H. J. Notonegoro

Department of Information Technology, Faculty of Computer Science and Information Technology
Universitas Gunadarma

Margonda Raya St. No. 100, Pondok Cina, Depok 16424, West Java, Indonesia

Email: raditheo@staff.gundarma.ac.id

1. INTRODUCTION

Vanilla cultivation in Indonesia is often affected by various plant diseases, which contribute to vanilla low quantity and quality. Diseases that affecting vanilla plant such as, stem rot caused by *Fusarium oxysporum*, the base stem rot sclerotium, the pods rot, vanilla shoots tor, and anthracnose caused by *Colletotrichum gloeosporioides* [1]. Diseases that infect vanilla plants affect the availability of vanilla commodities in meeting market demand. Early detection of these diseases helps prevent the spread of pathogens to other parts of the plant or between individual vanilla plants. Artificial intelligence (AI) can assist in detecting the presence of vanilla diseases by capturing images and collecting them, after which the data is prepared before being processed by a model to identify the type of disease affecting the plant.

Many image segmentations implemented in detection for crop diseases, early and late blight in potato using U-Net for semantic segmentation achieved 89.91% on mean intersection over union (MIoU) [2], tea coal, target spot, and brown blight in tea plantations with mask region-based convolutional neural network (R-CNN), wavelet transform, and four-channeled residual network (F-RNet), the Mask R-CNN model was able to detect 98.7% and F-RNet model achieved an accuracy of 88% [3]; apple disease identification using dice focal loss-U-Net + convolutional block attention module (DFL-U-Net + CBAM), in the evaluation of lesion segmentation and disease identification, the model achieved a MIoU of 91.07%, a

mean pixel accuracy (MPA) of 95.58%, and an F1-score of 95.16% [4], convolution neural network (CNN), support vector machine (SVM), k-nearest neighbors (KNN) model with GoogleNet obtained 98.50% [5], Mask R-CNN with residual network 50 layers (ResNet50) backbone provides 80.5% accuracy [6]; tomato leaf with modified U-Net obtained 98.66%, 98.5% MIoU score, and 98.73% on dice coefficient [7]; and other model using SVM obtained 100% accuracy for healthy leaf, 90% septoria spot, 95% early blight, and 85% late blight [8]; pear leaf with mask RCNN showed experimental result MPA 91.49% and MIoU 88.55% [9]; maize plant disease using region of interest attention network (RA-Net) framework obtained 43.10% on Dice intersection over union (IoU) comparison with transformer and reverse attention based method [10]; disease leaves with small inception model obtained 99.75% [11]; guava leaf disease using hybrid approach model U-Net-mobile network version 2 (MobileNetv2) obtained 92.41% accuracy on average, the segmentation model got 83.40%, and the detection model based on you only look once version 5 (YOLOv5) gained 71% on mean average precision (mAP) [12]; and peach disease using Mask R-CNN with ResNet101 as backbone obtained 46.50% on mAP segmentation [13].

Images of vanilla plant diseases on the internet are rarely available for public access. Most accessible disease images are found in books related to vanilla or vanilla gardening. However, images of vanilla plants themselves are quite widely available, such as on the global biodiversity information facility (GBIF), website, GitHub, and other opensource sites. There are also datasets that include mixed content with vanilla products (such as dried vanilla pods) available on Kaggle.

2. METHOD

In this study, data collection involved obtaining 20 images for each vanilla plants conditions such as healthy plants, infected with basal stem rot, stem root, and anthracnose caused by the fungus *Sclerotium rolfsii*, *Fusarium sp.*, *Colletotrichum sp.* Then masks were then applied using LabelMe, marking the infected areas (foreground) and the healthy or background areas (background). All images were in RGB format (.jpg), and the masks were in PNG format (.png), organized based on disease category. This limited dataset needs to be handled properly to achieve balance, in order to produce a good-fit model, as shown in Figure 1.

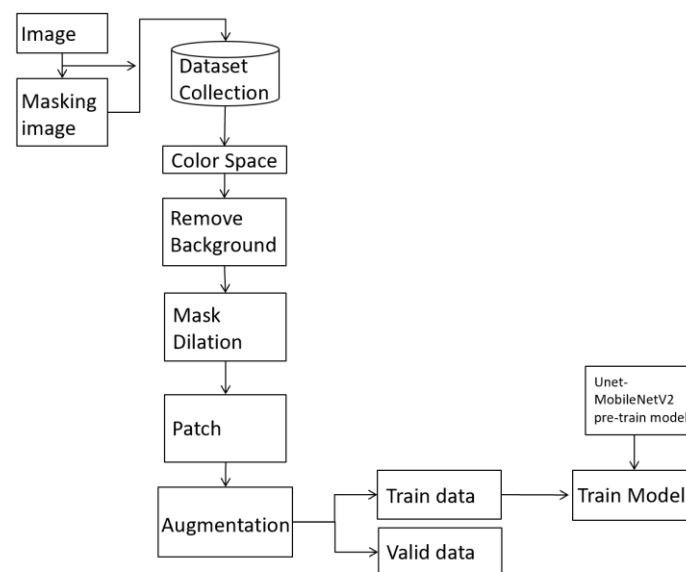


Figure 1. Dataset limited handled workflow

2.1. Convert to color space

Various color space models—such as RGB, HSI, CMYK, Lab, and HSV—can be utilized for analysis. Typically, the acquired images are initially processed within the RGB color space, and color transformations are applied to the images to standardize the color representation [14]. The mean and standard deviation of each channel in an RGB image are also calculated. A contour is defined as a line that traces all the pixels along the boundaries of an element [15]. The YCbCr color space separates luminance information (Y) from chrominance information (Cb and Cr), making it more stable under varying lighting conditions. The

converted YCbCr images are saved as intermediate outputs for further visual analysis. This step aims to enhance the model's robustness against lighting variations and differences in disease coloration.

2.2. Remove background

Each pixel in the original image with a mask value of 0 (background) is set to zero (black). This is intended to emphasize the foreground area (e.g., diseased regions) and reduce noise from irrelevant background information. The initial images were processed to eliminate background interference through segmentation [16].

2.3. Mask dilatation

The disease regions in the mask images are sometimes thin or fragmented, which may lead to insufficient representation during patch extraction. To address this, morphological dilatation was applied to the masks using a 5×5 structuring element, it helps ensure that the lesion area is completely included [17]. This operation enlarges the foreground areas, ensuring that a greater portion of the diseased region is included within extracted patches. Consequently, this step improves the quality of training data by increasing the representation of relevant disease features, facilitating better model learning.

2.4. Patch

The cropped images and masks were divided into non-overlapping patches of 256×256 pixels. This patching approach enabled accurate disease classification, images need to be split into smaller patches or segments. These segmented portions are then passed to a classifier to identify any signs of disease [18]. Each patch was evaluated, and only those containing at least 1% foreground pixels were retained for further processing. This filtering step aims to prevent the model from training on patches without meaningful information (i.e., empty or background-only regions). The model is able to improve its accuracy in classifying and detecting plant diseases by learning from a broader and more varied collection of samples [19].

2.5. Augmentation

Datasets commonly consist of both raw and synthetically augmented images. When the available original data is limited, augmentation becomes essential to support the training process. This step generates extra samples within the dataset, which helps reduce generalization errors in the CNN and enhances the model's ability to handle unseen data more effectively [20].

To reduce the risk of overfitting—where the model overly adapts to the training set and performs poorly on unseen data—data augmentation techniques are frequently employed. Two widely used approaches are image manipulation and image erasing [21]–[23]. The former involves altering the image either spatially (e.g., flipping or rotating) or by modifying pixel-level characteristics such as brightness and contrast. These operations are generally divided into geometric augmentations, which adjust the spatial configuration of the image, and non-geometric augmentations, which focus on changing pixel values or introducing noise. The selected patches are then augmented using both geometric and non-geometric augmentation steps to increase data variability and enhance model robustness against real-world conditions.

2.5.1. Geometric

Geometric manipulation techniques involve transforming the spatial structure of image data by altering orientation, position, or aspect ratio. These transformations modify the arrangement of pixels through operations such as rotation, translation, shearing, and flipping. Rotation entails rotating the image by an angle ranging from 0° to 360° , enabling the model to recognize objects regardless of their orientation. Flipping reverses the image either horizontally or vertically, enhancing the model's ability to generalize across mirrored variations. These geometric augmentations contribute to improved model robustness by introducing spatial diversity within the training dataset. Each patch undergoes five transformations with specific probabilities: horizontal flipping ($p = 0.5$), random rotation within $\pm 15^\circ$, and resizing back to 256×256 pixels. Each augmentation produces five new variants from a single patch, as illustrated in Figure 2.



Figure 2. Geometric augmentation

2.5.2. Non-geometric

Non-geometric augmentation techniques involve altering the visual characteristics of an image without modifying its spatial structure. These methods include cropping and resizing, noise injection, and color space transformation. Cropping and resizing consist of selecting a smaller region of the image and scaling it to the original dimensions. This approach is commonly applied during the preprocessing stage to improve input consistency and enhance model generalization [21], [24]. Noise injection adds random noise to the image to simulate real-world imperfections, thereby increasing the robustness of neural networks against noisy or low-quality data. Color space transformations adjust individual color channel values independently, helping to reduce biases caused by specific lighting conditions and improving model performance across diverse environments. Jittering randomly modifies image properties such as brightness, contrast, saturation, and hue, controlled by hyperparameters defining allowable variation ranges. Additionally, operations such as blurring, brightness adjustment, and further noise injection are applied to simulate conditions like out-of-focus imagery, varying illumination, and data imperfections. These augmentations improve a model's ability to handle novel or unpredictable real-world scenarios.

2.6. Split

After the augmentation process is completed, all generated image and mask pairs are prepared for model training. However, before being used for training, the dataset needs to be split into three distinct subsets: training, validation, and testing. The splitting is performed in two stages using the following approach: the dataset is first divided into 85% for training and 15% validation.

2.7. Train

Pretrained U-Net segmentation model using the ImageNet dataset, with MobileNetV2 as the backbone, sigmoid activation, early stopping, and learning rate reduction used during train. MobileNet greatly lowers computational expenses and model complexity compared to standard convolutional methods. This design enables MobileNet to maintain an efficient structure while still delivering acceptable accuracy and improves efficiency by using depthwise separable convolutions, reverse residual connections for deeper feature learning, and linear bottlenecks to preserve low-dimensional information [25], [26]. The model trained using four different color preprocessing methods: RGB, CIE Lab, HSV, and YCbCr with parameters were set as follows: 100 epochs, batch size of 16, image size of 256×256 pixels, and a learning rate of 0.0001.

2.8. Evaluation

The evaluation results based on accuracy, Dice, also CrossEntropyloss and Dixeloss for the four models are presented in Tables 1 and 2. The accuracy defined as the number of correct predictions, including both correct positive and negative predictions, compared to the total number of predictions [27]. Accuracy can be calculated using (1). Dice is used most widely used metric for assessing segmentation accuracy. It can be expressed in terms of per-voxel classification of true positives (TP), false positives (FP), and false negatives (FN), as shown in (2). A weighted combination of the Dice similarity coefficient (DSC) and a modified version of the cross-entropy loss is used as the combined loss function, which is defined in (3), where $\alpha \in [0, 1]$ determines the relative importance of the Dice and cross-entropy components in the loss function. Additionally, β adjusts the weight given to false positives and false negatives, with a β value greater than 0.5 placing a higher penalty on false negatives compared to false positives [28]. The modified cross-entropy loss used in the combined loss function is defined in (4).

$$Accuracy = \frac{TP+TN}{TP+TN+FN+FP} \quad (1)$$

$$DSC = \frac{2TP}{2TP+FP+FN} \quad (2)$$

$$L_{combo} = \alpha(L_{mCE}) - (1 - \alpha) \cdot DSC \quad (3)$$

$$L_{mCE} = -\frac{1}{N} \sum_{i=1}^N \beta(y_i \ln(p_i)) + (1 - \beta)[(1 - y_i) \ln(1 - p_i)] \quad (4)$$

Table 1. Evaluation train model

Color models	Accuracy	Dice	Loss
RGB	0.9828	0.9699	0.0740
CIE Lab	0.9874	0.9792	0.0529
YCbCr	0.9877	0.9801	0.0511
HSV	0.9778	0.9652	0.0931

Table 2. Evaluation valid model

Color models	Accuracy	Dice	Loss
RGB	0.9684	0.9790	0.1224
CIE Lab	0.9654	0.9821	0.1417
YCbCr	0.9674	0.9831	0.1386
HSV	0.9617	0.9780	0.1370

3. RESULTS AND DISCUSSION

The model performance on the training dataset using four different color spaces: RGB, CIE Lab, YCbCr, and HSV. Among these, the model trained with the YCbCr color space achieved the highest accuracy (0.9877) and highest dice coefficient (0.9801), while also attaining the lowest loss value (0.0511). This indicates that YCbCr effectively enhances feature representation, enabling the model to better learn object characteristics during training. The CIE Lab model also demonstrated competitive performance with accuracy and Dice values of 0.9874 and 0.9792, respectively. Meanwhile, the RGB-based model, although commonly used, showed slightly lower results (accuracy of 0.9828 and Dice of 0.9699). The HSV color space resulted in the lowest training performance, particularly with the highest loss (0.0931), suggesting that the hue-based representation in HSV might be less suitable for feature learning in this context.

Evaluation results on the validation dataset. Unlike the training phase, the RGB color space achieved the highest accuracy (0.9684), indicating better generalization. However, in terms of the Dice coefficient—which is a more relevant metric for segmentation tasks—the YCbCr model still performed best (Dice = 0.9831), followed closely by CIE Lab (0.9821) and RGB (0.9790). Interestingly, all models exhibited increased loss values on the validation data, indicating a slight overfitting. The RGB model had the lowest validation loss (0.1224), while CIE Lab and YCbCr showed slightly higher losses (0.1417 and 0.1386, respectively). Despite higher loss, their Dice scores remained superior, which suggests that they maintain better spatial segmentation quality. The HSV model consistently ranked lowest across metrics on both training and validation sets, indicating its unsuitability for this specific segmentation task.

The superior performance of the YCbCr color space is consistent with previous studies demonstrating that separating luminance from chrominance improves robustness to illumination variations and enhances lesion feature representation for plant disease segmentation [6], [14]. Although the RGB model achieved the highest validation accuracy, the YCbCr model obtained the highest Dice coefficient, indicating better spatial agreement between the predicted masks and the ground truth. This observation supports the findings of Müller *et al.* [27], who reported that the Dice coefficient is a more reliable metric than pixel accuracy for evaluating semantic segmentation quality. Furthermore, the slight increase in validation loss across all color spaces suggests mild overfitting, which is common in deep learning models trained on limited agricultural image datasets. Similar findings have been reported [21], [25], who emphasized that preprocessing and data augmentation are essential for improving model generalization.

The comparative analysis reveals that YCbCr and CIE Lab color spaces enhance segmentation performance more effectively than RGB and HSV, particularly in terms of the Dice coefficient. These color spaces separate luminance from chrominance, which may improve the model's ability to learn object boundaries and spatial features. However, the discrepancy between training and validation loss across all models suggests the need for additional regularization techniques or data augmentation strategies to reduce overfitting and improve model generalization. Overall, the results demonstrate that transforming images into YCbCr or CIE Lab color spaces before model training can lead to better segmentation outcomes, especially in tasks where spatial accuracy is critical. The identification result from patch shows in Figure 3, where Figure 3(a) shows the original, Figure 3(b) shows the mask prediction, and Figure 3(c) shows the overlay.

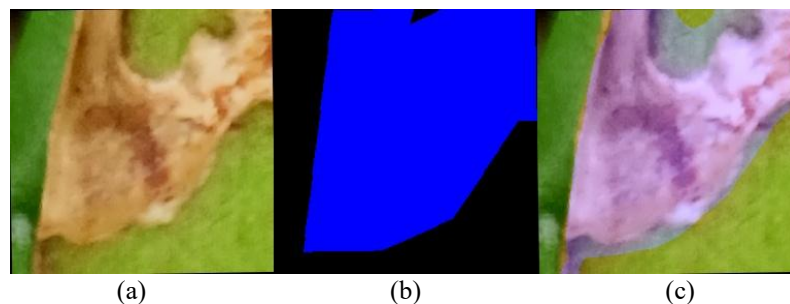


Figure 3. Vanilla disease identification from patch (a) original, (b) mask prediction, and (c) overlay

4. CONCLUSION

This study investigated the impact of different color space representations—RGB, CIE Lab, YCbCr, and HSV—on the performance of a deep learning model for image segmentation. Based on the experimental results, it was found that the YCbCr color space consistently produced the best performance, particularly in terms of the Dice coefficient, which is crucial for segmentation tasks. The CIE Lab color space also demonstrated strong performance, outperforming the standard RGB representation in both training and validation phases. Although the RGB model achieved the highest validation accuracy and lowest validation loss, the YCbCr and CIE Lab models maintained superior segmentation quality, as evidenced by their higher Dice scores. In contrast, the HSV color space resulted in the lowest performance across all metrics, indicating its limited effectiveness in this application. The findings suggest that converting images to YCbCr or CIE Lab color spaces can improve model accuracy and segmentation quality. Future work should explore data augmentation and regularization strategies to address the overfitting observed during validation and further improve model generalization.

FUNDING INFORMATION

This work is partially supported by the Directorate of Research and Community Service, Ministry of Higher Education, Science, and Technology Republic of Indonesia-Grant No. 0419/C3/DT.05.00/2025, on 22 May 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
Radityo H. J. Notonegoro	✓	✓	✓		✓	✓	✓	✓	✓	✓				
Dewi Agushinta R.		✓		✓	✓			✓		✓	✓	✓	✓	
Diana Ikasari	✓									✓	✓	✓	✓	✓
Rifki Kosasih	✓		✓	✓						✓		✓		

C : Conceptualization	I : Investigation	Vi : Visualization
M : Methodology	R : Resources	Su : Supervision
So : Software	D : Data Curation	P : Project administration
Va : Validation	O : Writing - Original Draft	Fu : Funding acquisition
Fo : Formal analysis	E : Writing - Review & Editing	

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

INFORMED CONSENT

This section not applicable. This study did not involve human participants, human data, or any personally identifiable information. All data used were either publicly available, fully anonymized, or derived from non-human sources, and therefore no informed consent was required from individuals.

ETHICAL APPROVAL

This section not applicable. This research did not involve human subjects, human biological materials, or experimental procedures on animals. The work was conducted solely on computational models, publicly available datasets, or non-sensitive data that did not require intervention with living organisms. Therefore, ethical approval from an institutional review board or animal ethics committee was not necessary for this study.

DATA AVAILABILITY

The dataset used in this study was collected and manually annotated by the authors. Because it is part of an ongoing research project, the dataset is not publicly available but can be obtained from the corresponding author, [RHJN], upon reasonable request.

REFERENCES

- [1] Y. Astuti, R. Siagian, C. Daniati, and N. Isnaini, *Pengenalan dan pengendalian opt pada tanaman vanili*. Directorate of Plantation Protection, Directorate General of Estates, Ministry of Agriculture, Jakarta, 2021.
- [2] X. Li, Y. Zhou, J. Liu, L. Wang, J. Zhang, and X. Fan, "The detection method of potato foliage diseases in complex background based on instance segmentation and semantic segmentation," *Frontiers in Plant Science*, vol. 13, Jul. 2022, doi: 10.3389/fpls.2022.899754.
- [3] H. Li *et al.*, "Symptom recognition of disease and insect damage based on Mask R-CNN, wavelet transform, and F-RNet," *Frontiers in Plant Science*, vol. 13, Jul. 2022, doi: 10.3389/fpls.2022.922797.
- [4] X. Zhang, D. Li, X. Liu, T. Sun, X. Lin, and Z. Ren, "Research of segmentation recognition of small disease spots on apple leaves based on hybrid loss function and CBAM," *Frontiers in Plant Science*, vol. 14, Jun. 2023, doi: 10.3389/fpls.2023.1175027.
- [5] S. Alqethami, B. Almtanni, W. Alzhirani, and M. Alghamdi, "Disease detection in apple leaves using image processing techniques," *Engineering, Technology and Applied Science Research*, vol. 12, no. 2, pp. 8335–8341, Apr. 2022, doi: 10.48084/etasr.4721.
- [6] G. Storey, Q. Meng, and B. Li, "Leaf disease segmentation and detection in apple orchards for precise smart spraying in sustainable agriculture," *Sustainability*, vol. 14, no. 3, Jan. 2022, doi: 10.3390/su14031458.
- [7] M. Shoaib *et al.*, "Deep learning-based segmentation and classification of leaf images for detection of tomato plant disease," *Frontiers in Plant Science*, vol. 13, Oct. 2022, doi: 10.3389/fpls.2022.1031748.
- [8] S. U. Rahman, F. Alam, N. Ahmad, and S. Arshad, "Image processing based system for the detection, identification and treatment of tomato leaf diseases," *Multimedia Tools and Applications*, vol. 82, no. 6, pp. 9431–9445, Mar. 2023, doi: 10.1007/s11042-022-13715-0.
- [9] W. Mu, Z. Jia, Y. Liu, W. Xu, and Y. Liu, "Image segmentation model of pear leaf diseases based on mask R-CNN," in *2022 International Conference on Image Processing and Media Computing (ICIPMC)*, 2022, pp. 41–45, doi: 10.1109/ICIPMC55686.2022.00016.
- [10] G. Y. Moon and J. O. Kim, "RoI-attention network for small disease segmentation in crop images," *IEEE Access*, vol. 12, pp. 63725–63734, 2024, doi: 10.1109/ACCESS.2024.3393301.
- [11] Z. Zainab and R. Mahum, "Plant disease detection using deep learning techniques," *ICCK Journal of Image Analysis and Processing*, vol. 1, no. 1, pp. 36–44, Mar. 2025, doi: 10.62762/JIAP.2025.227089.
- [12] J. Rashid, I. Khan, G. Ali, S. ur Rehman, F. Alturise, and T. Alkhalifah, "Real-time multiple guava leaf disease detection from a single leaf using hybrid deep learning technique," *Computers, Materials and Continua*, vol. 74, no. 1, pp. 1235–1257, 2023, doi: 10.32604/cmc.2023.032005.
- [13] N. Yao, F. Ni, M. Wu, H. Wang, G. Li, and W.-K. Sung, "Deep learning-based segmentation of peach diseases using convolutional neural network," *Frontiers in Plant Science*, vol. 13, May 2022, doi: 10.3389/fpls.2022.876357.
- [14] V. G. Krishnan, J. Deepa, P. V. Rao, V. Divya, and S. Kaviarasan, "An automated segmentation and classification model for banana leaf disease detection," *Journal of Applied Biology and Biotechnology*, Oct. 2022, doi: 10.7324/JABB.2021.100126.
- [15] F. P. P. Rajeena, S. U. Aswathy, M. A. Moustafa, and M. A. S. Ali, "Detecting plant disease in corn leaf using EfficientNet architecture—an analytical approach," *Electronics*, vol. 12, no. 8, Apr. 2023, doi: 10.3390/electronics12081938.
- [16] J. F. R.-Arias, J. W. B.-Bedoya, and G. Awad, "Plant disease detection strategy based on image texture and Bayesian optimization with small neural networks," *Agriculture*, vol. 12, no. 11, Nov. 2022, doi: 10.3390/agriculture12111964.
- [17] L. C. Ngugi, M. Abdelwahab, and M. A.-Zahhad, "A new approach to learning and recognizing leaf diseases from individual lesions using convolutional neural networks," *Information Processing in Agriculture*, vol. 10, no. 1, pp. 11–27, Mar. 2023, doi: 10.1016/j.inpa.2021.10.004.
- [18] L. W. Kuswidiyanto, H.-H. Noh, and X. Han, "Plant disease diagnosis using deep learning based on aerial hyperspectral images: a review," *Remote Sensing*, vol. 14, no. 23, 2022, doi: 10.3390/rs14236031.
- [19] S. Abinaya, K. U. Kumar, and A. S. Alphonse, "Cascading autoencoder with attention residual U-Net for multi-class plant leaf disease segmentation and classification," *IEEE Access*, vol. 11, pp. 98153–98170, 2023, doi: 10.1109/ACCESS.2023.3312718.
- [20] T. Ilyas, H. Jin, M. I. Siddique, S. J. Lee, H. Kim, and L. Chua, "DIANA: a deep learning-based paprika plant disease and pest phenotyping system with disease severity analysis," *Frontiers in Plant Science*, vol. 13, Oct. 2022, doi: 10.3389/fpls.2022.983625.
- [21] A. A. AlArfaj *et al.*, "Multi-step preprocessing with U-Net segmentation and transfer learning model for pepper bell leaf disease detection," *IEEE Access*, vol. 11, pp. 132254–132267, 2023, doi: 10.1109/ACCESS.2023.3334428.
- [22] H. Wang, S. Qiu, H. Ye, and X. Liao, "A plant disease classification algorithm based on attention MobileNet V2," *Algorithms*, vol. 16, no. 9, Sep. 2023, doi: 10.3390/a16090442.
- [23] P. Oza, P. Sharma, S. Patel, F. Adedoyin, and A. Bruno, "Image augmentation techniques for mammogram analysis," *Journal of Imaging*, vol. 8, no. 5, May 2022, doi: 10.3390/jimaging8050141.
- [24] T. Kumar, R. Brennan, A. Mileo, and M. Bendechache, "Image data augmentation approaches: a comprehensive survey and future directions," *IEEE Access*, vol. 12, pp. 187536–187571, Mar. 2023, doi: 10.1109/ACCESS.2024.3470122.
- [25] K. Alomar, H. I. Aysel, and X. Cai, "Data augmentation in classification and segmentation: a survey and new strategies," *Journal of Imaging*, vol. 9, no. 2, Feb. 2023, doi: 10.3390/jimaging9020046.
- [26] K. Maharana, S. Mondal, and B. Nemade, "A review: data pre-processing and data augmentation techniques," *Global Transitions Proceedings*, vol. 3, no. 1, pp. 91–99, Jun. 2022, doi: 10.1016/j.gltp.2022.04.020.
- [27] D. Müller, I. S.-Rey, and F. Kramer, "Towards a guideline for evaluation metrics in medical image segmentation," *BMC Research Notes*, vol. 15, no. 1, Dec. 2022, doi: 10.1186/s13104-022-06096-y.
- [28] M. Yeung, E. Sala, C. B. Schönlieb, and L. Rundo, "Unified focal loss: generalising dice and cross entropy-based losses to handle class imbalanced medical image segmentation," *Computerized Medical Imaging and Graphics*, vol. 95, 2022, doi: 10.1016/j.compmedimag.2021.102026.

BIOGRAPHIES OF AUTHORS

Radityo H. J. Notonegoro    holds a bachelor of computers (S.Kom.) in Information Systems, and is currently pursuing a Master of Information Systems Management (M.M.S.I.) in Information Systems Software in addition to several certificates and professional skills. He is currently staff at Universitas Gunadarma. His research areas of interest include artificial intelligence, image processing, and data science. He can be contacted at email: radithjn@staff.gunadarma.ac.id.



Dewi Agushinta R.    is a professor of Information Technology at the post-doctoral program of Information Technology, Universitas Gunadarma. She had her bachelor degree in 1993 and doctoral degree from Universitas Gunadarma in 2007. Her master degree was from Asian Institute Technology, Bangkok, Thailand. She is giving some lectures in the Department of Information System at Faculty of Computer Science and Information Technology and other department in the same university. Her interest research is in human-computer interaction, image processing, and information technology. She can be contacted at email: dewiar@staff.gunadarma.ac.id or dewiar@gmail.com.



Diana Ikasari    is a lecturer at Universitas Gunadarma in the field of Computer Science and Information Technology. She earned her doctorate in Information Technology from Universitas Gunadarma in 2018. Her scientific research focuses on topics related to information systems technology. Her areas of interest include databases and natural language programming. She is the author of the book 'Information retrieval' and has compiled a teaching book titled 'Introduction to geographic information systems'. She also contributed a book chapter entitled 'Geographic information systems in the application of straight-line distance calculation to public high schools (SMAN) in Depok'. She can be contacted at email: d_ikasari@staff.gunadarma.ac.id.



Rifki Kosasih    received the B.S. degree in Mathematics from University of Indonesia, Indonesia, in 2009, the M.S. degree in Mathematics from University of Indonesia in 2012, and the Ph.D. degree from Universitas Gunadarma, Indonesia, in 2015. He is a lecturer in Department of Information Technology, Universitas Gunadarma. His research interests are image processing, object recognition, data science, manifold learning, machine learning, and deep learning. He can be contacted at email: rifki_kosasih@staff.gunadarma.ac.id.