Detecting *Hemileia vastatrix* using Vision AI as Supporting to Food Security for Smallholder Coffee Commodities

Teddy Suparyanto$^{1,2}$, Erick Firmansyah$^1$, Tjeng Wawan Cenggoro$^{2,3}$, Digdo Sudigyo$^2$, Bens Pardamean$^{2,4}$

$^1$Faculty of Agricultural Technology, Institute of Agriculture STIPER Yogyakarta, Indonesia 55282
$^2$Bioinformatics & Data Science Research Center Bina Nusantara University Jakarta, Indonesia 11480
$^3$Computer Science Department, School of Computer Science, Bina Nusantara University, Jakarta, Indonesia 11480
$^4$Computer Science Department, BINUS Graduate Program - Master of Computer Science Program, Bina Nusantara University, Jakarta, Indonesia 11480

Email: teddysup@binus.ac.id

**Abstract.** The loss of food crops during the COVID-19 pandemic threatens food security in Indonesia as one of the world's top coffee producers. This affects the insecurity of coffee commodities which is influenced by several factors other than the pandemic such as pests, plant diseases, and extreme weather. Plant diseases, such as leaf rust, are a significant factor in the insecurity problem in coffee commodities. The fungus *Hemileia vastatrix* B et Br causes leaf rust disease, which is a pest that frequently damages coffee plants. This disease not only interferes with plant growth but also causes a decrease in coffee quality and quantity. This initial research aims to carry out early prevention of these diseases as supporting to food security in smallholder coffee commodities. An AI-based visual detection application is the result of this research. We collected 100 images of coffee leaves from various coffee plants. The image is reshaped to 256 x 256 pixels and randomly trimmed to 224 x 224 pixels to fulfill the size requirements of a standard Deep Learning technique. Each image was classified into two classes by a plant pest and disease specialist. The dataset was divided into training, validation, and testing series with a ratio of 60:20:20 for training procedures. The Convolutional Neural Network (CNN) research method used a variation of the ResNet CNN model with 18 layers. The best model validation was 59%.

1. **Introduction**

The COVID-19 outbreak shown a significant impact on agriculture, including the coffee plantation industry. Almost all sectors of the coffee plantation industry are carried out by farmers who are very vulnerable to a crisis [1]. In addition, the impact of the pandemic has also decreased the production of food crops, especially in Indonesia. In a report from the Bappenas, there was a decline in the food crop subsector in the third quarter of 2020 compared to the previous two years [2]. Stable food security in the agricultural sector in all countries will trigger the diversification of agricultural products, thus helping the coffee industry. So that food security is very influential on the sustainability of the coffee industry for farmers. Other issues, such as insufficient coffee supply and excessive coffee prices, might lead to commodity instability in this market [1]. Extreme climate changes and exacerbated by the attack pests...
organism become a crucial factor. As a result, farmers' communities that produce 70% of the world's coffee supply are the biggest victims of food security problems [3,4].

Indonesia, the world's fourth-largest coffee producer since 1998, tries to meet high requirements in terms of both quality and quantity of coffee goods. Preventing the attack by Pest Organisms Plants (OPT), which affected economic losses for coffee agricultural actors dominated by independent farmers, is part of this effort. The fungus Hemileia vastatrix B et Br causes leaf rust disease, which is a pest that frequently damages coffee plants. This disease not only interferes with plant growth and also results in loss of coffee quality and quantity.

Arabica coffee is more susceptible to leaf rust disease than robusta and liberica coffee [5,6]. The decrease in coffee production due to leaf rust attack can reach more than 25% [6], with the level of plant damage reaching 58% [7]. The effects of leaf rust attacks can be gradual for several years after the attack, although the new-onset can be controlled. As a result, coffee farmers have difficulty in achieving maximum production. Attacks at an advanced stage were difficult to control and had inflicted heavy damage.

Plant diseases can be identified in the field by looking for signs and symptoms. Pathogenic microorganisms that infect the leaves are the hallmark of the cause of this disease. This characteristic can also be observed with the ordinary eye, which is a sign of this disease. For example, mycelia shaped like cotton is one sign of a pathogenic fungus that infects the plant. The symptoms are highly unique depending on the infecting species, so it can be used in the field to identify the type of pathogen infecting leaves [8].

Even though this disease has been developing for more than a century, and inflicts massive losses in Indonesian coffee plantations. There has been no effective treatment for the condition until recently [9]. Therefore, it becomes crucial to evaluate attack intensity with high accuracy and precision, even when the disease occurs in short intervals and with low severity. Early observation and proper identification of the disease will ensure successful control [10]. Effective disease identification and recognition methods can be developed to reduce leaf rust disease and increase Indonesian coffee productivity [11]. Several efforts can be made to prevent pests and disease from spreading in crops, such as early and controlled integrated pest and disease control, which minimizes the risk of crop loss and reduces the demand for pesticides. Early diagnosis of the disease can prevent the emergence of severe attacks and facilitate disease control activities that have the potential to increase coffee productivity [12].

For efficient pest and disease control, symptom severity is a crucial indicator besides plant disease causative agents. Plant stress levels are also diagnosed and quantified, which are two equally important tasks for integrated pest control operations [13]. The severity of plant disease attacks can be measured by calculating symptomatic plant tissue percentage area [14]. Although numerous studies have concentrated on the main problem of plant foliar disease, few have focused on establishing methods capable of evaluating stress severity.

Artificial Intelligence (AI) is now extensively used in numerous sectors, including predicting or simulating biological processes [15], environmental issues [16,17], modern society [18–20], and health issues [21]. The application of AI in agriculture has also begun to be developed by researchers to overcome production problems such as creating various systems for automatic calculations on crop yields [22,23], viewing production quality [24–26], harvest product quality classification [27], etc. However, the detection of plant diseases, especially coffee leaf disease, has not been widely applied in Indonesia.

Therefore, in answering the problem of supporting food security for coffee commodities for farmers, we developed an application by applying vision AI to detect bacterium H. vastatrix on the coffee leaf that causes leaf rust disease.
2. Literature Review

2.1. Hemileia vastatrix

*Hemileia vastatrix* is a fungus that attacks coffee plants, especially on the leaves. These fungi cause leaf rust disease in coffee plants which reducing coffee production by up to 25%. In infected leaves, *H. vastatrix* was found in the form of uredospores (rust fungal spores). However, other forms found can be in the form of uredium and mycelium (a collection of rust fungal hyphae) that infect coffee plants [28,29]. Uredospores become the means of transmission of this disease by infecting the stomata on the leaves. The spread of uredospores can be through agricultural tools, natural air and water media, and physical contact with organisms.

The characteristic colour of yellow-orange spots in this disease is caused by uredospores attached to the upper surface of the leaves. In Indonesia, damage by this disease to coffee plants reaches 58%. The development of leaf rust disease is a significant issue in Indonesian coffee production [30,31]. Symptoms of yellow spots on coffee leaves are the initial diagnosis of *H. vastatrix* infection. The colour change from yellow to white in the uredospore causes the infected leaves to prematurely and slump young leaves. The severity impact of this leaf disease spreads on the leaves and twigs, which makes these parts fall out and die of drought [32].

2.2. Artificial Intelligence in Plant Disease Detection

Artificial implementation in agriculture is widely applied, one of which is to detect plant diseases. ML (Machine Learning) and CV (Computer Vision) technology support humans to improve automation in the detection and recognition of crop diseases. Computer vision, known as Vision AI is an AI system that can recognize and identify an object using a computer device. Deep Learning (DL) is part of vision AI that promises to detect plant and plant diseases that show phenotype or physical characteristics of diseased organisms [33–35].

The application of deep learning in recognizing diseased plants using pixel-wise operations, this method analyzes and classifies the infected (sick) and healthy leaves of the affected plants. Diseased/infected plants show visible patterns in certain parts that are the characteristics of plant disease [35]. Dhakal and Shayka’s research [36], which uses deep learning to detect plant diseases based on images, has an accuracy of 98.59%. These results were also validated by plant pathologists to guide the use of datasets in this technology. Validation from a plant pathologist is crucial to determine the appropriate dataset of disease information to be detected. Plant diseases infect more than one part of the plant, one of which is Cassava disease. This problem affects the accuracy of the image dataset that will be used in detecting the disease. Leaf images are the most easily recognized dataset by Vision AI for detecting plant diseases [37]. The Ramcharan et al. [38] research utilizes CNN (Convolutional Neural Network) to determine Cassava disease with leaf images that showed reliable and high accuracy results. Feature extraction for disease identification is needed to improve accuracy with unorganized dataset images. In the study, Al-bayati and Stündag [39] used feature fusion and extracted the leaf area affected by the disease. Dandawate and Kokare's study [39] converted images from RGB to HSV colour space, which then used Scale Invariant Feature Transform to detect soybean disease based on the shape of the leaves. Colour texture features and statistical classification algorithms were used by Pydipati et al. [40] to identify the citrus disease. They also used the Color Cooccurrence Method to see if colour variables like HIS (Hue, Saturation, and Intensity) could support in identifying diseased leaves. The accuracy of their procedure was higher than 0.95.

3. Methodology

3.1. Coffee Leaves Dataset

The research method AI dataset comprising coffee leaves images and the matching normal/health and leaf rust/fungal infected is required to construct an AI system for *Hemileia vastatrix* detection. As a result, we gathered 100 coffee leaves images from different coffee plants. Figure 1 depicts a selection of images from the dataset. The image is reshaped to 256 x 256 pixels and randomly trimmed to 224 x 224 pixels to fulfill the size requirements of a standard Deep Learning technique. Each image was
classified into two classes by a plant pest and disease specialist, as shown in Table 1. The dataset was divided into training, validation, and testing series with a ratio of 60:20:20 for training procedures [27].

| Classes   | Code | Description                                                        |
|-----------|------|--------------------------------------------------------------------|
| Infected  | I    | The leaf colour has yellow spots on the upper surface and orange uredospores on the lower surface. |
| Normal    | N    | The leaf shows green colour in the whole part as an indication of healthy coffee plants. |

*Figure 1. Samples of coffee leave images in the dataset*

3.2. *AI System for Hemileia vastatrix detection*

The research method Convolutional Neural Network (CNN) [41] was the Deep Learning algorithm employed in the web application that implemented with AI system. In specifically, as shown in Figure 2, we employed a ResNet [42] variation of the CNN model with 18 layers, dubbed ResNet18. ResNet is capable of outperforming humans in picture classification by implementing residual connections, which are depicted in Figure 3 with arched arrows. The residual connection can be expressed mathematically as $g_t (x) = f_t (g_{t-1} (x)) + g_{t-1} (x)$, where $g_t (x)$ is the output of the $t^{th}$ layer block and $f_t (x)$ is the convolutional operation in the $t^{th}$ layer block.

A large dataset is necessary to optimize the outcomes of a Deep Learning model. Unfortunately, our dataset is far smaller than what would be required to train a real Deep Learning model. We used random horizontal flipping as an additional data augmentation procedure in addition to random cropping to solve this problem [27].
4. Results and Discussion

4.1. General Performance
The ResNet18 model was trained using Stochastic Gradient Descent, with a momentum of 0.9. The initial learning rate decreases by a factor of 0.1 every seven epochs. Epoch 20 terminated the training phase. The loss values and accuracies are shown in Figure 3 as the training process progresses. The best model has a validation accuracy of 59%. The graphs show that the model is overfitting due to the vast difference between validation loss and training. This result is almost the same as what happened in the manufacture of palm fruit ripeness detection. The number of data samples does not affect the accuracy results even though it has been reproduced with ten crops [27]. The technique of multiplying data with Ten Crops and random horizontal flipping was able to increase the accuracy value of AI even though it was a little.

![Figure 3. ResNet18's performance was reported in its loss.](image)

4.2. Per-Class Performance
The model's validation F1-score for each class is shown in Table II. In the courses with the incredibly small sample size, N had the highest F1 Score. It should be emphasized, however, that even if only one
sample is misclassified, F1-Score swings in a small sample can be significant. As a result, the F1-Scores of I and N may be misrepresentative.

However, Kamilaris and Prenafeta-Boldu [34] reported in their paper review that many papers related to the detection of plant and leaf diseases and using the DL (Deep Learning) technique reported excellent results with an adequate sample. This result indicates that the accuracy is greater than 0.95 and that the F1 score is greater than 0.92. When compared with the results of this study, the discrepancy F1 score is very different. Therefore, that in future research, additional samples are needed to increase the accuracy of AI in detecting *Hemileia vastatrix* in coffee plants.

| Class | #Training Images | #Val Images | F1-Score |
|-------|-----------------|------------|----------|
| I     | 36              | 9          | 0.53     |
| N     | 32              | 8          | 0.63     |

**Table 3. Confusion matrix**

| Prediction | Actual Data | Precision |
|------------|-------------|-----------|
|            | I           | N         |           |
| I          | 4           | 5         | 0.67      |
| N          | 2           | 6         | 0.55      |

4.3. Web-based Leaf Disease Application

A web-based leaf disease application compiled in python using the Flask library. Figure 4 shows a simple interface for coffee leaf disease detection. In this interface, the user is asked to upload a photo file of coffee leaves. After selecting the file and uploading it, the interface will show the predicted percentage of the photo being sick or healthy leaves.

![Screenshot of application interface](image)

**Figure 4. Screenshot of application interface**

5. Conclusion

*Hemileia vastatrix* is a coffee leaf disease that develops very quickly. This fungus is due to climate change which increases the cases of this disease in coffee plants today. Early detection of this disease helps farmers in tackling *Hemileia vastatrix* attacks. This AI-based application is the initial research, in which the data used is still limited so that the results obtained are still at an accuracy of 59%.

In future work, more datasets will be added to the detecting application to improve the accuracy level. In addition, standardization of sample photos is applied to enhance the validation and accuracy of results. The web-based application will be upgraded using Android, making it easier for coffee farmers to access this application.
References

[1] Kanniah J C 2020 Why Food Security Is a Coffee Industry Issue Perfect Dly. Grind

[2] Bappenas 2020 Laporan Perkembangan Ekonomi Indonesia dan Dunia Triwulan III Tahun 2020 (Jakarta: Kedeputian Bidang Ekonomi Kementerian PPN/Bappenas)

[3] Eakin H, Tucker C and Castellanos E 2006 Responding to the coffee crisis: a pilot study of farmers’ adaptations in Mexico, Guatemala and Honduras Geogr. J. 172 156–71

[4] Eakin H, Winkels A and Sendzimir J 2009 Nested vulnerability: exploring cross-scale linkages and vulnerability teleconnections in Mexican and Vietnamese coffee systems Environ. Sci. & Policy 12 398–412

[5] Sukamto S 1998 Pengelolaan penyakit tanaman kopi Kumpul. Mater. Pelatih. Pengelolaan Org. Penganggu Tanam. Kopi. Pas. Penelit. Kopi dan Kakao. Jember 15

[6] Semangun H 2006 Pengantar Ilmu Penyakit Tumbuhan Pengantar Ilmu Penyakit Tumbuhan (Yogyakarta: Gadjah Mada University Press)

[7] Rosmahani, Rachmawati D, Sarwono, Soleh M and Jumadi 2005 Pengkajian aplikasi PHT untuk meningkatkan produksi dan pengaruhnya terhadap pendapatan petani kopi arabika Agrosains 77–85

[8] Agrios G N 1988 Plant pathology. 3rd (San Diego: Academic Press Inc.)

[9] Esgario J G M, Krohling R A and Ventura J A 2020 Deep learning for classification and severity estimation of coffee leaf biotic stress Comput. Electron. Agric. 169 105162

[10] Abadi A L 2003 Ilmu Penyakit Tumbuhan vol 145 (Malang: Bayu Media)

[11] Defitiri Y 2016 Pengamatan beberapa penyakit yang menyerang tanaman kopi (Coffea sp) di desa Mekar Jaya kecamatan Betara kabupaten Tanjung Jabung Barat J. Media Pertan. 1 78–84

[12] Putri A Y P and Sodik A 2019 Identifikasi Penyakit Tanaman Kopi Arabika dengan Metode K-Nearest Neighbor (K-NN) Pros. Sem. Nas. Sains Teknol. Terp. (SNTEKPAN 2019) vol 1 pp 759–64

[13] Vandermeer J, Perfecto I and Liere H 2009 Evidence for hyperparasitism of coffee rust (Hemileia vastatrix) by the entomogenous fungus, Lecanicillium lecanii, through a complex ecological web Plant Pathol. 58 636–41

[14] Bock C H, Poole G H, Parker P E and Gottwald T R 2010 Plant disease severity estimated visually, by digital photography and image analysis, and by hyperspectral imaging CRC. Crit. Rev. Plant Sci. 29 59–107

[15] Lumbanraja F R, Mahesworo B, Cenggoro T W, Budiarto A and Pardamean B 2019 An evaluation of deep neural network performance on limited protein phosphorylation site prediction data Procedia Comput. Sci. 157 25–30

[16] Caraka R E, Tahmid M, Putra R M, Iskandar A, Mauludin M A, Goldameir N E, Rohayani H, Pardamean B and others 2018 Analysis of plant pattern using water balance and cimogram based on oldeman climate type IOP Conference Series: Earth and Environmental Science vol 195 p 12001

[17] Pardamean B, Cenggoro T W, Chandra B J and Rahutomo R 2020 Data annotation system for intelligent energy conservator in smart building IOP Conference Series: Earth and Environmental Science vol 195 p 12008

[18] Muchtar K, Rahman F, Cenggoro T W, Budiarto A and Pardamean B 2018 An improved version of texture-based foreground segmentation; block-based adaptive segmenter Procedia Comput. Sci. 135 579–86

[19] Cenggoro T W, Tanzil F, Aslamiah A H, Karuppiah E K and Pardamean B 2018 Crowdsourcing annotation system of object counting dataset for deep learning algorithm IOP Conference Series: Earth and Environmental Science vol 195 p 12063

[20] Pardamean B, Muljo H H, Abid F, Susanto A, Cenggoro T W and others 2021 RHC: a dataset for in-room and out-room human counting Procedia Comput. Sci. 179 33–9

[21] Caraka R E, Nugroho N T, Tai S-K, Chen R-C, Toharudin T and Pardamean B 2020 Feature importance of the aortic anatomy on endovascular aneurysm repair (EVAR) using Boruta and Bayesian MCMC Commun. Math. Biol. Neurosci. 2020 Article--ID
[22] Cenggoro T W, Budiarto A, Rahutomo R and Pardamean B 2018 Information system design for deep learning based plant counting automation 2018 Indonesian Association for Pattern Recognition International Conference (INAPR) pp 329–32

[23] Rahutomo R, Perbangsa A S, Lie Y, Cenggoro T W and Pardamean B 2019 Artificial Intelligence Model Implementation in Web-based Application for Pineapple Object Counting Proc. 2019 Int. Conf. Inf. Manag. Technol. ICIMTech 2019 vol 1 pp 525–30

[24] Putra D P, Bimantio M P, Sahfitra A A, Suparyanto T and Pardamean B 2018 Simulation of Availability and Loss of nutrient Elements in Land with Android-based Fertilizing Applications Proc. 2018 Int. Conf. Inf. Manag. Technol. ICIMTech 2020 pp 312–7

[25] Samodro B, Mahesworo B, Suparyanto T, Atmajad B S, Pardamean B and others 2020 Maintaining the quality and aroma of coffee with fuzzy logic coffee roasting machine IOP Conference Series: Earth and Environmental Science vol 426 p 12148

[26] Putra D P, Bimantio M P, Suparyanto T and Pardamean B 2021 Expert System for Oil Palm Leaves Deficiency to Support Precision Agriculture Proc. 2018 Int. Conf. Inf. Manag. Technol. ICIMTech 2021 vol 1 pp 33–6

[27] Harsawardana, Rahutomo R, Mahesworo B, Cenggoro T W, Budiarto A, Suparyanto T, Surya Atmajad B, Samoedro B and Pardamean B 2020 AI-based ripeness grading for oil palm fresh fruit bunch in smart crane grabber IOP Conf. Ser. Earth Environ. Sci. 426 12147

[28] Semangun H 2000 Penyakit-Penyakit Tanaman Perkebunan di Indonesia (Yogyakarta: Gadjah Mada University Press)

[29] Mahfud M C 2012 Teknologi dan strategi pengendalian penyakit karat daun untuk meningkatkan produksi kopi nasional Pengemb. Inov. Pertan. 5 44–57

[30] Kushalappa A C and Eskes A B 1989 Coffee Rust: Epidemiology, Resistance, and Management (Florida: CRC Press, Inc)

[31] Rosmahani L, Mahfud M, Handoko, Rahmawati D, Sarwono, Soleh M and Subagio H Uji penerapan teknologi PHT tingkat petani oleh petani pada kopi arabika rakyat di dataran tinggi Prosiding Seminar dan Ekspose Teknologi BPTP Jatim pp 441–53

[32] Agrios G N 2005 Plant Pathology. 5th Edition (San Diego: Academic Press Inc.)

[33] Wang G, Sun Y and Wang J 2017 Automatic image-based plant disease severity estimation using deep learning Comput. Intell. Neurosci. 2017

[34] Kamilaris A and Prenafeta-Boldú F X 2018 Deep learning in agriculture: A survey Comput. Electron. Agric. 147

[35] Khan R U, Khan K, Albattah W and Qamar A M 2021 Image-based detection of plant diseases: from classical machine learning to deep learning journey Wirel. Commun. Mob. Comput. 2021

[36] Dhakal A and Shakya S 2018 Image-based plant disease detection with deep learning Int. J. Comput. Trends Technol. 61 26–9

[37] Verma S, Chug A, Singh A P, Sharma S and Rajvanshi P 2019 Deep learning-based mobile application for plant disease diagnosis: A proof of concept with a case study on tomato plant Applications of Image Processing and Soft Computing Systems in Agriculture (IGI Global) pp 242–71

[38] Ramcharan A, Baranowski K, McCloskey P, Ahmed B, Legg J and Hughes D P 2017 Deep learning for image-based cassava disease detection Front. Plant Sci. 8 1852

[39] Al-bayati J S H and Üstündag B B 2020 Evolutionary feature optimization for plant leaf disease detection by deep neural networks Int. J. Comput. Intell. Syst. 13 12–23

[40] Pydipati R, Burks T F and Lee W S 2006 Identification of citrus disease using color texture features and discriminant analysis Comput. Electron. Agric. 52 49–59

[41] LeCun Y, Boser B, Denker J S, Henderson D, Howard R E, Hubbard W and Jackel L D 1989 Backpropagation applied to handwritten zip code recognition Neural Comput. 1 541–51

[42] He K, Zhang X, Ren S and Sun J 2016 Deep Residual Learning for Image Recognition Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. CVPR 2016 pp 770–8