International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Using Machine Learning in Detecting Ganoderma Disease in Oil Palm Plants

Mardiana Wahyuni*[1], T. Sabrina[2], Mukhlis[2] & Heri Santoso[3]

Submitted: 12/12/2023 **Revised**: 16/01/2024 **Accepted**: 01/02/2024

Abstract: Ganoderma disease is asymptomatic, posing challenges in detection and identification. The utilization of Remote Sensing techniques is expected to provide rapid, accurate, and large-scale detection information. This study aims to identify and classify oil palm plant images recorded by a UAV, analyze digital vegetation indices, and apply RF and SVM algorithms. The UAV equipped with sensors capturing three bands: R, G, NIR. Data processing was conducted using software such as Mission Planner, Agisoft MetaShape, Mapir Camera Control, ArcGIS 10.5, Envi 5.3, and R Studio. The study locations were at Pabatu. Image recording took place on July 3, 2021. The observed parameters included disease incidence, reflectance values, vegetation indices (NDVI, GNDVI, SAVI, SR, CIgreen), with plant classes categorized as Healthy (H) and Infected (I). Infected plants exhibited lower reflectance compared to healthy plants. All vegetation indices, including NDVI, GNDVI, SAVi, SR, and CIgreen, were lower in infected plants compared to healthy plants. The SVM algorithm demonstrated the highest accuracy of 93.55% compared to RF only 84.42%. With Machine Learning algorithm, disease occurrences where imagery has been recorded can be predicted. This map can serve as fundamental information for control strategies, production calculations, and other cultivation activities.

Keywords: *Ganoderma, UAV, reflectance, vegetation indices, Random Forest algorithm, Support Vector Machine algorithm*

1. Introduction

Ganoderma basal stem rot caused by the pathogen *Ganoderma boninense* has become a threat to the sustainability of oil palm plantations in Indonesia. The total area of oil palm plantations in Indonesia is about 14 million hectares, with crude palm oil (CPO) production of 36.5 million tons. Ganoderma disease leads to growth disturbance, decreased production, and even plant death [1], [2]. The potential loss caused by Ganoderma disease in Malaysia reaches up to 45% [3], [4].

The pathogen *Ganoderma boninense* belongs to the white rot fungus, which has a high ability to degrade lignin starting from the basal stem, thus disrupting the absorption of water and nutrients from the soil [5]. One characteristic of Ganoderma disease is its asymptomatic nature; by the time visual symptoms appear, the plant is already severely affected [6], [7].

Early detection of Ganoderma disease incidence is crucial for control strategies and preventing further losses [6]. Manual detection/monitoring by observing symptoms on the stem and canopy of the plant has limitations such as requiring a long time, a large workforce, and high accuracy from the personnel.

Therefore, there is a need for a technology-based method to ensure the sustainability of large-scale oil palm plantations in Indonesia. The use of Remote Sensing (RS) techniques is expected to provide up-to-date and large-scale information for detecting Ganoderma disease incidence [8].

Identification and classification of plant health conditions on a company-scale oil palm plantation are related to managing large amounts of data (big data). The application of Machine Learning (ML) techniques is highly beneficial for forming algorithmic models that can identify Ganoderma disease occurrences within a sample area (ground truth). These models are then applied to rapidly, accurately, and comprehensively predict disease incidence in target areas.

The Random Forest (RF) algorithm can be used for the classification of large-scale data using principles similar to decision tree formation [9]. Another alternative is the Support Vector Machine (SVM) technique, which can be used for classification purposes by recognizing patterns in a dataset and forming classification hyperplanes [10].

The use of Linear Discriminant Analysis (LDA), K-Nearest Neighbor (KNN), and Naive Bayes (NB) algorithms to detect and classify Ganoderma disease based on measured reflectance values using a spectroradiometer resulted in the highest accuracy value of 96% with the NB method [11].

The RF, Decision Tree (DT), and SVM algorithms were also implemented to detect and classify Ganoderma disease outbreaks based on World View satellite imagery [12]. The accuracy values of these algorithms were 53.1% (RF), 53.3% (DT), and 54.1% (SVM). The RF algorithm was further applied to detect Ganoderma disease based on oil palm plant imagery recorded with an Unmanned Aerial Vehicle (UAV) [13]. The accuracy value using the RF algorithm increased to 73.95%.

This study aims to apply the RF and SVM algorithms to classify the incidence of Ganoderma disease in oil palm plants and create a distribution map based on ground truth observations and UAV-recorded imagery. The accuracy of the classification algorithms is crucial as a basis for determining actions for controlling Ganoderma disease in oil palm plants.

[1] Doctoral Program in Agricultural Sciences, Faculty of Agriculture, Universitas Sumatera Utara, Medan 20155 Indonesia
[2] Agrotechnology Study Program, Faculty of Agriculture, Universitas Sumatera Utara, Medan 20155 Indonesia
[3] Indonesia Oil Palm Research Institute, Medan 20178 Indonesia
** Corresponding Author Email: mardiana@itsi.ac.id*

Study Area

The research was conducted in the Pabatu oil palm plantation owned by the state-owned enterprise (BUMN) PT. Perkebunan Nusantara IV, located in the Serdang Bedagai regency, North Sumatra Province, Indonesia (Figure 1). The coordinates of the study area are $3^0 14^I 36^{II}$ North Latitude and $99^0 6^I 50^{II}$ East Longitude, with an elevation of about 200 m above sea level, and the soil type is Inceptisol (*Typic Dystrudepts*). The Pabatu plantation consists of 7 divisions with a total area of 2,998 hectares of oil palm plantations. The research was conducted in divisions III and IV. The study focused on oil palm trees planted in 2004 and 2005 (categorized as mature plants and representing the 3^{rd} generation), consisting of 12 blocks with a total area of 247 hectares. Initially, the oil palm trees were planted at a density of 130 trees per hectare (with a triangular spacing of 9.4 m). The mortality rate of oil palm trees due to Ganoderma disease in the research area varied between 24-52% [14].

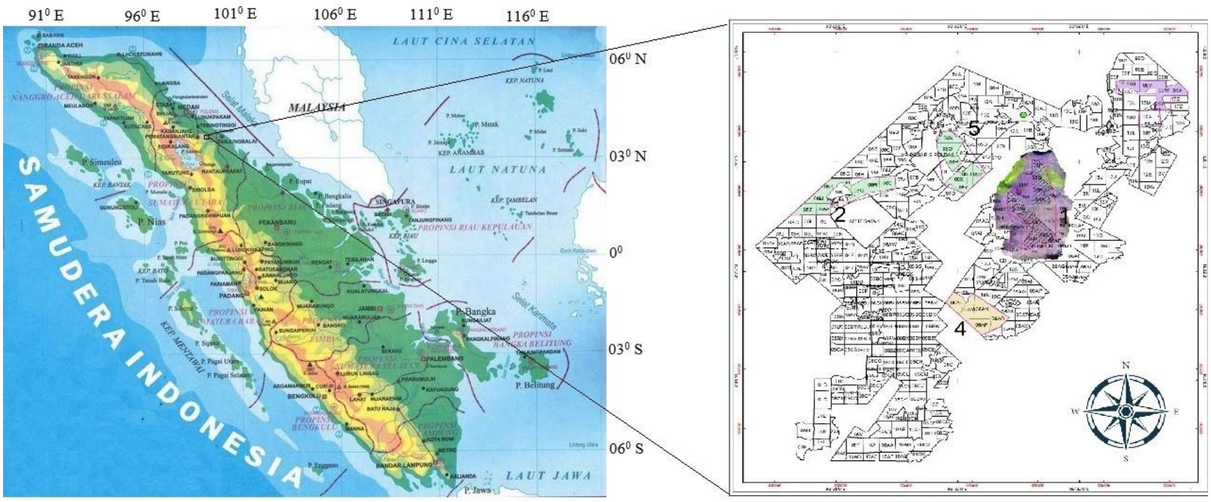

Fig 1. Study Area

2. Research Methodology

Research Flowchart and software utilized are presented in Figure 2.

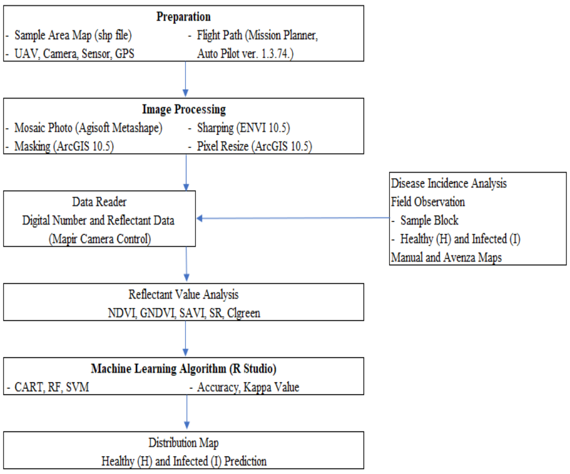

Fig 2. Research Flowchart

3.1. Preparation

The basic map used is in shapefile (.shp) format, was obtained from the Head Office of PT. Perkebunan Nusantara IV in Medan. The unmanned aerial vehicle (UAV) used to capture images of the oil palm trees in the Area of Interest (AOI) was a T Tail Voltron UAV equipped with a MAPIR Survey 3 camera, a Sony Exmax Rimx 117 sensor, and filters sensitive to plant health, namely Red (R), Green (G), and Near Infra Red (NIR) at respective wavelengths of 550 nm, 660 nm, and 850 nm. The UAV was also equipped with a Global Positioning System (GPS), following the method used [12]. The flight path was planned using Mission Planner Auto Pilot software version 1.3.74. The UAV/drone was flown at an altitude of 300-350 m dfhgfsh above ground level, with a photo coverage distance rtyutgegrere of 40 m, a resolution of 9-10 cm, a UAV speed of 10-12 m/s, with an overlap of 75% within the flight path, and an overlap of 80% between flight paths.

3.2. Image Capture

Based on the flight path plan, the image capture was conducted on July 3, 2021, with details provided in Table 1.

Table 1. Actual Image Capture of Oil Palm Plants

Division	Planting Year	Area (ha)	Time (*WIB*)	Minutes	Number of Photos
IV	2004	34	10.30-10.59	29	188
III	2005	213	08.49-09.41	50	175
	Total	247		79	363

The average recording capacity of the UAV in the research area is 3.12 hectares per minute.

3.3. Image Processing

The image processing begins with the mosaic of photos, which involves combining the images using Agisoft Metashape software. Each image is merged based on the continuity of coordinates obtained from GPS readings. Subsequently, masking is performed to separate the Area of Interest (AOI) from other areas captured by the UAV using ArcGIS 10.5 software. Image sharpening is carried out by adjusting the image brightness to improve pixel value readings using ENVI 5.3 software. Pixel resizing is also performed to enhance the visibility of oil palm plant points, resizing them from 0.17 m to 2.0 m [13]. An example of pixel resizing can be seen in Figure 2.

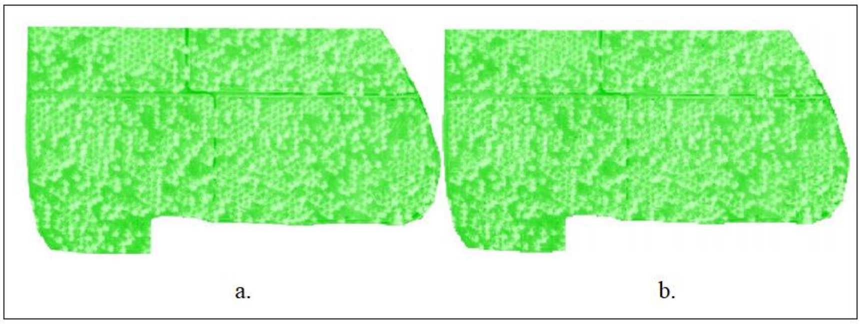

Fig 3. Example of the resized pixel in block 05H; (a) Initial image (0.17 m) and (b) Resized image (2.0 m)

3.4. Reflectance Value Reading

The values recorded by the camera and sensor are in the form of Digital Numbers (DN). DN values represent values within a range/spectrum from white to black, ranging from 2^8 bytes (0-256) and 2^{16} bytes (65,536). Conversion of DN values to reflectance values (ranging from 0 to 1) is performed using Mapir Camera Control (MCC) software, following the procedure outlined: https://www.mapicamera/ pages/calibrating+images+n+mapir+camera-control-application. The reflectance values are read using ENVI 5.3 software. The sample points are derived from the ground truth, comprising approximately 10% of the designated AOI selected through random purposive sampling.

Table 2. Example of Reflectance Value Reading

No.	UTM Coordinate Position		Reflectance	Reflectance	Reflectance
	Latitude	Longitude	R	G	NIR
1	3193748	2611477	0,2097	0,1944	0,3882
2	3189292	2664799	0,1318	0,1216	0,2470
3	3260587	2612992	0,091887	0,084695	0,184311
4	3264389	2704249	0,067913	0,084695	0,129406
5	3271994	2803112	0,185787	0,158663	0,325498
6	3271994	2886764	0,151823	0,121679	0,301967
7	3275797	2985626	0,175798	0,158663	0,325498
8	3325227	2491315	0,01397	0,011919	0,02351
9	3341198	2586756	0,101876	0,084695	0,203921
10	3344240	2668887	0,131844	0,121679	0,247261
11	3345625	2739240	0,249719	0,194455	0,458841

3.5. Field Observation

Field observations of Ganoderma disease incidence were conducted manually based the symptoms on the stems and leaves of oil palm plants, simplified into two categories: Healthy (H) and Infected (I) [7]. Infected plants were identified by single or combined symptoms, including more than 3 unopened spear leaves, leaf chlorosis, frond breakage, stem rot, and the presence of fruiting bodies. The observation results were manually recorded and mapped and transferred to the points on the oil palm plant image based on their spatial similarity using the Avenza Map software.

Observation samples were taken from 5 blocks covering an area of 69 ha ($\pm$ 30% representing the AOI). In block 05K, all plants were observed, while in blocks 05E, 05F, 04A, and 04B, 130 plants were observed in each block (representing a 1 ha area), guided by intervals of every 10^{th} row and every 5^{th} tree within each row, similar to leaf sampling for oil palm plants [15]. Reflectance values in the R, G, and NIR bands were read using ENVI 10.5 software.

3.6. Vegetation Index Analysis

The vegetation index parameter in this study was adjusted based on the bands used, namely R, G, and NIR, using the following formula [13], [16] in Table 3.

Table 3. Indexes and Formulas

Indexes	Formulas
Normalized Difference Vegetation Index (NDVI)	NDVI = (NIR-R) / (NIR+R)
Green Normalized Difference Vegetation Index (GNDVI)	GNDVI = (NIR-G) / (NIR+G)
Soil Adjusted Vegetation Index (SAVI)	SAVI = {(NIR-R)/(NIR+R+L) x (1+L)}
L represents the soil adjustment coefficient, which is commonly set to 0.5 for areas with moderate soil coverage.	
Simple Ratio (SR)	SR = IR / R
Chlorophyll Index (CI)	CI green = (R/G) - 1

3.7. Machine Learning Algorithm

The algorithms used for classification are supervised methods, specifically Random Forest (RF) and Support Vector Machine (SVM). The input variables include the reflectance values in the R, G, and NIR bands, as well as vegetation indices such as NDVI, GNDVI, SAVI, SR, and Cl green, resulting in a total of 8 variables. The classification classes consist of 2 categories: Healthy (H) and Infected (I) plant status.

The algorithmic steps are processed using R software (R Studio), with a 75% separation for training data to build the model and a 25% separation for testing data to measure the accuracy obtained from the Confusion Matrix.

Table 4. R language programming code (Machine Learning) steps in R Studio

No.	Steps	Code (R/python programming language in R Studio)

1.	Adding Packages and Library	#Packages
		install.packages("e1071")
		install.packages("rpart")
		install.packages("rpart.plot")
		install.packages("caret")
		install.packages("caTools")
		#Library
		library(e1071)
		library(rpart)
		library(rpart.plot)
		library(caret)
		library(caTools)

2.	Table naming initialization preparation (Label) ✓ First data container initiation from Ground Truth spreadsheet data	pabatuGT

	✓ Second data container similar to first data container to be a classified purposed data to be separated into training and testing data	pabatuProcessed
	✓ Training data table	pabatuTraining
	✓ Testing data table	pabatuTesting
	✓ Table name contains data formula to create data prediction	pabatuModel
	✓ Table name contains predicted data	pabatuPredict
3.	Pre-process Ground Truth spreadsheet data set into pabatuGT table notes: Ctrl-C at spreadsheet data first	pabatuGT<-read.delim("clipboard")
4.	Create table to contain separeted data into training and testing data	pabatuProcessed<-createDataPartition(pabatuGT$Status,p=0.75,list=FALSE)
	a. calculate training table	pabatuTraining<-pabatuGT[pabatuProcessed,]
	b. calculate testing table	pabatuTesting<-pabatuGT[-pabatuProcessed,]
5.	Create Model contained formula to predict data of independent variable	
	a.apply/create Random Forest based model	pabatuModel=randomForest(as.factor(Status)~.,pabatuTraining,method="class")
	b. apply/create SVM based model	pabatuModel=svm(as.factor(Status)~.,pabatuTraining,method="class")
6.	Classify independent variable/column predicted value ("Status" variable) with the aid of randomForest/SVM	pabatuPredict=predict(pabatuModel,pabatuTesting,type="class")
7.	Measure Confusion Matrix to show difference between "Status" variable of Ground Truth Data (block 05K) compare to Predicted Data	confusionMatrix(pabatuPredict, as.factor(pabatuTesting$Status))
8.	Create Distribution Map of Ground Truth and based on Prediction Data ArcGIS	

9. STOP

The components of the confusion matrix results are shown in Table 4.

Table 5. Explanation of Confusion Matrix Components

No.	Component	Description	Form
1	TP (True Positive)	Plants classified as H and predicted as H	TP
2	FP (False Positive)	Plants classified as H and predicted as I	FP
3	FN (False Negative)	Plants classified as I and predicted as H	FN
4	TN (True Negative)	Plants classified as I and predicted as I	TN
5	Accuracy	Percentage of correct predictions from all data	TP+TN/data
6	Sensitivity/Producers Accuracy +	Percentage of true positive predictions from positive data	TP/(TP+FP)
7	Specificity/Producers Accuracy -	Accuracy in predicting negative classification from negative data	TN/(TN+FN)
8	Positive Predict Value/User Accuracy +	Positive predictions in the positive category	TP/(TP+FN)
9	Negative Predict Value/Mapping Accuracy +	Negative predictions in the negative category	TN/(FP+TN)
10	Prevalence/Use Accuracy -	Percentage of negative predictions from all data	TN/ all data
11	Detection rate/Mapping Accuracy -	Percentage of positive predictions from all data	TP/ all data

3.8. Distribution Map

The distribution map represents the outcome of predictions (Y predict) from the selected algorithm, resulting in symbols representing oil palm plants on the Healthy (H) or Infected (I) class map.

4. Result and Discussion

4.1. Ganoderma Disease Incidence

The field observation results of Ganoderma disease incidence are presented in Table 6.

Table 6. Incidence of Ganoderma Disease in Pabatu Plantation.

Division	Planting Year	Block	Area (ha)	Sample	Healthy (H)		Infected (I)	
III	2005	05K	16	1.436	1110	(77,30%)	326	(22,70%)
	2005	05E	11	130	86	(66,20%)	44	(33,80%)
	2005	05F	8	130	98	(75,40%)	32	(24,60%)
IV	2004	04A	11	130	116	(89,20%)	14	(10,80%)
	2004	04B	23	130	75	(57,70%)	55	(42,30%)
		Total	69	1.956	1485	(75,90%)	471	(24,10%)

The incidence of Ganoderma disease in Pabatu plantations varies, with an average of 24.1%. The high incidence of the disease in mature and old oil palm trees may be attributed to the colonization of pathogens that are already present in the soil. The *Ganoderma boninense* pathogen is soil-borne, and its mycelium, which is resistant to basidiospores and chlamydospores, can persist in the soil for a long time [5], [17]. In general, the soil fertility conditions in Pabatu plantations suffer from several nutrient deficiencies, particularly in potassium (K) and magnesium (Mg) [14]. Low fertility levels in the soil can act as a contributing factor to the increased incidence of Ganoderma disease [6], [18].

4.2. Reflectance Value

Based on a total of 231 observed samples in Pabatu, with 184 (79.6%) classified as Healthy (H) and 47 (20.4%) classified as Infected (I), a descriptive analysis of the reflectance values of oil palm trees is presented in Table 7. Plants strongly absorb light in the visible spectrum, especially in the Red (R) range with wavelengths of 606-725 nm, while in the infrared spectrum (NIR) with wavelengths of 680-780 nm, they reflect more light [16], [19]. The lowest values of reflectance are found in the green (G) spectrum, while the highest values are observed in the infrared (NIR) spectrum. In healthy plants, the reflectance values are low in the R and G spectra, and high in the NIR spectrum. Reflectance patterns have been applied by several researchers to identify diseases in various plants, including oil palm [19], [20], [12]; peanut plants [21], and grapevines [22].

Table 7. Analysis of Observed Reflectance Values

Category	Statistic	Reflectance		
		R	G	NIR
Healthy (H)	Mean	0,2031	0,1702	0,3802
	Minimum	0,0034	0,0045	0,0078
	Maximum	0,5094	0,2835	0,9975
Infected (I)	Mean	0,1151	0,113	0,1697
	Minimum	0,0081	0,0071	0,0088
	Maximum	0,2835	0,3042	0,4843
% I/H	Mean	56,70%	66,40%	44,60%

4.3. Vegetation Indices

It has been suggested that vegetation indices can be used as an indication to quantitatively measure agronomic parameters such as leaf area, biomass, nutrient sufficiency, and plant health [19]. The results of vegetation index calculations for oil palm plants in the Pabatu plantation are presented in Table 8.

Table 8. Vegetation Indices of Oil Palm Plants in the Pabatu Plantation

Vegetation Index	Healthy (H)	Infected (I)	I/H (%)
NDVI	0,29	0,22	75,86%
GNDVI	0,34	0,22	64,71%
SAVI	0,21	0,10	47,62%
SR	1,81	1,72	95,03%
Clgreen	1,06	0,64	60,38%

Generally, the vegetation indices values in Infected (I) plants are lower compared to Healthy (H) plants. The method with the largest difference is SAVI, while the method with the least difference is SR. These differences align with the observations of reflectance values in the R, G, and NIR bands of oil palm plants.
The average NDVI value in Healthy (H) plants falls within the good category range, which is 0.18-0.42 [23]. Visually, there is still little difference in leaf greenness between Healthy (H) and Infected (I) plants. This may be due to good plant management practices, such as the application of compound fertilizer (NPKNg) at a rate of 4-6 kg and dolomite fertilizer at a rate of 1.5-2.0 kg per tree per year.
The higher value of GNDVI compared to NDVI is influenced by the absorption of the Green (G) spectrum, which is highly sensitive to leaf chlorophyll content. It is suggested that plants with chlorophyll types a (blue-green) and b (yellow-green) absorb the R and G spectra, resulting in low reflectance values [22]. The use of SAVI aims to reduce the environmental conditions' influence, including possible soil reflection around the plants [16]. The SR index shows the lowest difference between Healthy (H) and infected (I) plant conditions, while the CIgreen value in healthy plants is 1.06, indicating no significant difference in reflectance. Both the red and green spectra are strongly absorbed in healthy plants, resulting in low reflectance values.

4.4. Machine Learning Classification Algorithm

4.4.1. Dataset

The input dataset for classification variables consists of observations of reflectance values and vegetation indices for a total of 231 plants, with a partial example presented in Table 9.

Table 9. Dataset of 8 variables for the Machine Learning Algorithm process

No.	Latitude	Longitude	R	G	NIR	NDVI	GNDVI	SAVI	SR	Clgreen	Category
1	6326843	4869074	0,204	0,188	0,376	0,296	0,334	0,239	1,842	1,0016	Healthy
2	6350843	4797074	0,170	0,154	0,294	0,268	0,313	0,194	1,732	0,9112	Healthy
3	6358843	4741074	0,227	0,200	0,402	0,278	0,336	0,232	1,769	1,0117	Healthy
4	6286843	4661074	0,369	0,228	0,721	0,323	0,519	0,332	1,955	2,1584	Healthy

5	6278843	4525074	0,427	0,348	0,794	0,301	0,391	0,320	1,860	1,2856	Healthy
6	6270843	4421074	0,188	0,169	0,329	0,274	0,236	0,092	0,166	1,6166	Infected
7	6270843	4293074	0,244	0,224	0,459	0,305	0,345	0,268	1,879	1,0513	Healthy
8	6278843	4189074	0,278	0,247	0,498	0,284	0,337	0,259	1,794	1,0168	Healthy
9	6294843	4093074	0,180	0,161	0,324	0,286	0,335	0,215	1,800	1,0089	Healthy
10	6302843	3989074	0,123	0,109	0,218	0,278	0,437	0,122	0,286	2,5504	Infected
…	…	…	…	…	…	…	…	…	…	…	…
231	6254843	3773074	0,295	0,261	0,539	0,293	0,058	0,295	0,046	1,1238	Infected

Based on the provided data, a Machine Learning algorithm is executed using R studio software.

4.4.2. The Confusion Matrix

The results of the confusion matrix and accuracy values for the RF and SVM algorithms are presented in Table 10.

Table 10. Confusion Matrix Results and Accuracy Values

Model	**Random Forest**			**Support Vector Machine**		
	H	**I**	**Total**	**H**	**I**	**Total**
Healthy (H) - Positif	58	4	62	61	5	66
Infected (I) - Negatif	8	7	15	0	11	11
Total	66	11	77	61	16	77
PA (%)	87,88	63,64	100	66,75		
UA (%)	99,35	85,71	92,42	79,21		
MA (%)	46,67	75,32	100	79,22		
OA (%)	84,42		93,55			
Kappa value/category	0,447 (sufficient)			0,777 (strong)		

The RF algorithm identified 62 (80.5%) Healthy (H) plants and 15 (19.5%) Infected (I) plants. In terms of PA and UA analysis, the RF algorithm showed better capability in identifying healthy plants compared to infected plants, while MA showed an antagonistic behavior by being more effective in mapping infected plants. The overall accuracy (OA) was 84.42%, which falls into the category of excellent. The high accuracy of the algorithm in classifying and mapping Ganoderma disease has been consistently reported [20], [12], [13]. The Kappa coefficient value falls into the category of moderate.

In the SVM algorithm, the number of plants indicated as Healthy (H) was 66 (85.7%), while Infected (I) plants were 11 (14.3%). The PA, UA, and MA values for predicting the Healthy (H) category were very high, with PA and MA reaching 100%. The SVM algorithm also demonstrated good capability in classifying the infected (I) status. It is suggested that the advantage of the SVM algorithm lies in its generality, ability to work with a smaller amount of data, statistical basis, and a good level of suitability [10]. The SVM algorithm has also been used in the field of eye health, with an accuracy of 79.63% [24]. The accuracy value of the SVM algorithm in this study was 93.5%, higher than the RF algorithm. The suitability of the SVM algorithm may be attributed to the pattern of the underlying data, specifically the reflectance variables, which tend to exhibit linearity. It is suggested that the SVM algorithm acts as a linear classifier, grouping data into two categories by forming an ideal or nearest hyperplane/boundary between the two categories [25]. In conditions where the visual distinction between healthy and infected plants is relatively minimal, the SVM algorithm successfully separates/classifies plants into the Healthy (H) and Infected (I) statuses. This is further supported by the Kappa value of 0.777, indicating strong agreement.

4.5. Distribution Map

Based on the algorithm with the highest accuracy, SVM, predictions were made for other blocks whose images were recorded using UAV. Some sample blocks and their predicted plant outcomes are presented in Table 11.

Table 11. Blocks and Predicted Plant Outcomes Using SVM Algorithm

Block	Area (ha)	Number of Plants predicted	Healthy (H)		Infected (I)	
			Total		Total	
04A	23	46	44	-95,70%	2	-4,35%
05B	22	387	252	-65,10%	135	-34,90%
05H	15	2.110	1.801	-85,40%	309	-14,60%
05I	17	684	433	-63,30%	251	-36,70%
05J	15	1.895	1.420	-74,90%	475	-25,10%
05K	16	617	498	-80,70%	119	-19,30%
	108	5.739	4.448	-77,50%	1.291	-22,50%

The prediction results show that the percentage of plants classified as Healthy (H) is 77.50%, while the percentage of Infected (I) plants is 22.50%. These values are close to the ground truth observation of disease incidence, which is 75.9% for healthy plants and 24.1% for infected plants. The identification and classification of plant conditions provide crucial and fundamental information regarding the severity of Ganoderma disease in the research area.

Several objectives of identification and classification, as stated in the decree by PT. Perkebunan Nusantara IV with No. 04.04/SE/OQ/II/2018 dated February 26, 2018, includes (a) obtaining data on productive tree inventory, (b) providing updated reports on Ganoderma attacks from year to year, (c) serving as the basis for production forecasts, (d) calculating the required fertilizer and maintenance costs, and determining the number of seedlings needed for insertion in oil palm plantations with plants aged less than 5 years.

Based on global studies, manual inventories are conducted once every 6 months with a labor cost of 0.3 US dollars per hectare. For example, in the Division III Pabatu location with an area of 657 hectares, a total of 394.2 US dollars or manpower is required for a one-year period with two census rounds. Assuming a tariff of 100,000 Indonesian Rupiah per ***mandays*** in 2022, the funding requirement for inventory would be Rp39,420,000.

It is suggested that the application of remote sensing techniques can help provide rapid, accurate, and large-scale information on plant conditions, thereby preventing the potentially fatal consequences of Ganoderma disease incidence [7], [12], [15].

As an example, the distribution map resulting from the SVM algorithm prediction in the Pabatu plantation is shown in Figures 3(a) and 3(b).

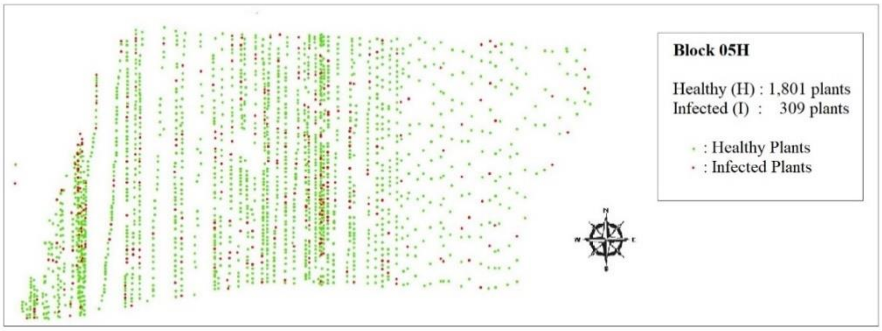

Fig 3.a. Sample of SVM Predicted Distribution Map in Block 05H in the Pabatu estate

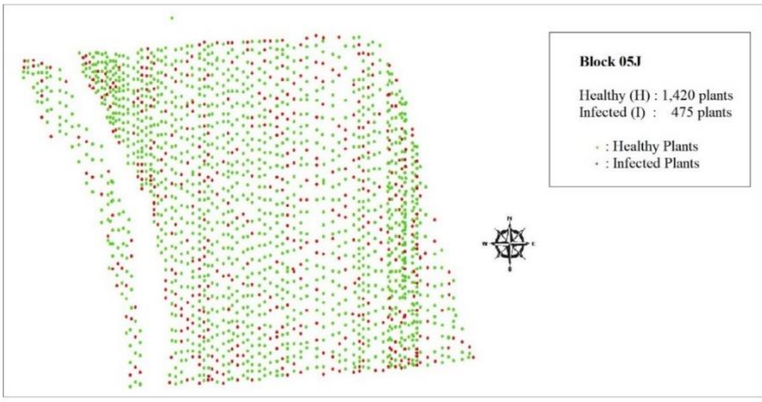

Fig 3.b. Sample of SVM Predicted Distribution Map in Block 05J in the Pabatu estate

Based on the two examples, the percentage of infected plants in block 05H is 14.64%, and in block 05J is 25.07%. It is suggested that the pattern of Ganoderma disease spread is not yet fully understood [5]. Root contact among mature and old oil palm plants is an important factor to consider. With the availability of distribution maps, plantation managers can make informed decisions regarding mechanical and sanitation-based control measures. It has been reported that the construction of isolation ditches to separate healthy and infected plants, as well as mound building, are part of short-term Ganoderma disease control measures [26], [27].

5. Conclusion

The oil palm plantations in the research area have been infected by Ganoderma disease with a percentage of 24.1%. Capturing images of oil palm plants using UAVs and analyzing reflectance and vegetation indices can provide information on the health status of plants categorized as Healthy (H) or Infected (I). The lowest reflectance values were observed in the red band (R), while the highest values were observed in the near-infrared (NIR) band. Infected plants (I) showed lower reflectance compared to healthy plants (H), with percentages of 56.7% for R, 66.4% for G,

and 44.6% for NIR. All vegetation indices, including NDVI, GNDVI, SAVI, SR, and CIgreen, were lower in infected plants compared to healthy ones. The best algorithm for classification was SVM, with an accuracy of 93.55% and a strong Kappa value of 0.777. Using the SVM algorithm, predictions were made on the plant conditions, and distribution maps were generated, providing valuable information for Ganoderma disease control. These maps can be utilized as fundamental information for control strategies, production calculations, and other cultivation activities.

Author contributions

Mardiana Wahyuni: Conceptualization, Methodology, Software, Field Study, Data Analysis, Writing Draft Preparation **T. Sabrina:** Fund, Validation, Supervision **Mukhlis:** Visualization **Heri Santoso:** UAV (Drone), Software, Field Study Reviewing.

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] Direktorat Jendral Perkebunan. 2020. Statistik Perkebunan Kelapa Sawit Indonesia. Direktorat Jendral Perkebunan Jakarta.

[2] J. Flood, L. Keenan, S. Wayne, and Y. Hasan, "Studies on oil palm trunks as sources of infection in the field," Mycopathologia, vol. 159, no. 1, pp. 101–107, 2005, doi: 10.1007/s11046-004-4430-8.

[3] M. S. Arif, A. Roslan, A. S. Idris, and M. Ramle, "Economics of oil palm pests and Ganoderma disease and yield losses," in Proceedings of the Third MPOB-IOPRI International Seminar: Integrated Oil Palm Pests and Diseases Management, Kuala Lumpur Convention Centre Kuala Lumpur, 2011.

[4] R. Hushiarian, N. A. Yusof, and S. W. Dutse, "Detection and control of Ganoderma boninense: Strategies and perspectives," Springerplus, vol. 2, no. 1, pp. 1–12, 2013, doi: 10.1186/2193-1801-2-555.

[5] R. R. M. Paterson, "Ganoderma disease of oil palm—A white rot perspective necessary for integrated control," Crop Protection, vol. 26, no. 9, pp. 1369–1376, Sep. 2007, doi: 10.1016/j.cropro.2006.11.009.

[6] R. H. V Corley and P. B. H. Tinker, The oil palm. John Wiley & Sons, 2008.

[7] C. C. Lelong et al., "Evaluation of Oil-Palm Fungal Disease Infestation with Canopy Hyperspectral Reflectance Data," Sensors, vol. 10, no. 1, pp. 734–747, Jan. 2010, doi: 10.3390/s100100734.

[8] T. M. Azahar, P. Boursier, and I. A. Seman, "Spatial analysis of basal stem rot disease using geographical information system," Map Asia, vol. 1820, 2008.

[9] L. Breiman, "Random forests," Mach Learn, vol. 45, pp. 5–32, 2001.

[10] T. O. Ayodele, "Types of Machine Learning Algorithms," New Advances in Machine Learning, 2010.

S. Liaghat et al., "Early detection of basal stem rot disease (Ganoderma) in oil palms based on hyperspectral reflectance data using pattern recognition algorithms," Int J Remote Sens, vol. 35, no. 10, pp. 3427–3439, 2014.

[12] H. Santoso, H. Tani, X. Wang, A. E. Prasetyo, and R. Sonobe, "Classifying the severity of basal stem rot disease in oil palm plantations using WorldView-3 imagery and machine learning algorithms," Int J Remote Sens, vol. 40, no. 19, pp. 7624–7646, 2019.

[13] H. Santoso, "Pengamatan dan Pemetaan Penyakit Busuk Pangkal Batang di Perkebunan Kelapa Sawit Menggunakan Unmanned Aerial Vehicle (UAV) dan Kamera Multispektral," Jurnal Fitopatologi Indonesia, vol. 16, no. 2, pp. 69–80, Dec. 2020, doi: 10.14692/jfi.16.2.69-80.

[14] PPKS, "Laporan Rekomendasi Pemupukan Tanaman Kelapa Sawit Kebun Tinjowan PT. Perkebunan Nusantara IV pada tahun 2018," 2018.

[15] K. Utami, S. Supriadi, and K. S. Lubis, "Evaluasi Sifat Fisik Tanah Terhadap Laju Infeksi Ganoderma di Perkebunan Kelapa Sawit (Studi Kasus: PT. PD. PATI)," Jurnal Agroekoteknologi Universitas Sumatera Utara, vol. 4, no. 3, p. 108266, 2016.

[16] J. L. Hatfield, A. A. Gitelson, J. S. Schepers, and C. L. Walthall, "Application of Spectral Remote Sensing for Agronomic Decisions," Agron J, vol. 100, no. S3, May 2008, doi: 10.2134/agronj2006.0370c.

[17] A. Susanto, P. S. Sudharto, and R. Y. Purba, "Enhancing biological control of basal stem rot disease (Ganoderma boninense) in oil palm plantations," Mycopathologia, vol. 159, no. 1, pp. 153–157, Jan. 2005, doi: 10.1007/s11046-004-4438-0.

[18] B. H. Prasetyo and danD A. Suriadikarta, "Karakteristik, potensi, dan teknologi pengelolaan tanah ultisol untuk pengembangan pertanian lahan kering di Indonesia," Jurnal Litbang Pertanian, vol. 25, no. 2, pp. 39–46, 2006.

[19] H. Z. M. Shafri and N. Hamdan, "Hyperspectral imagery for mapping disease infection in oil palm plantation using vegetation indices and red edge techniques," Am J Appl Sci, vol. 6, no. 6, 2009, doi: 10.3844/ajassp.2009.1031.1035.

[20] H. Santoso, H. Tani, and X. Wang, “Random Forest classification model of basal stem rot disease caused by Ganoderma boninense in oil palm plantations,” Int J Remote Sens, vol. 38, no. 16, pp. 4683–4699, Aug. 2017, doi: 10.1080/01431161.2017.1331474.

[21] V. Martínez-Martínez, J. Gomez-Gil, M. L. Machado, and F. A. C. Pinto, “Leaf and canopy reflectance spectrometry applied to the estimation of angular leaf spot disease severity of common bean crops,” PLoS One, vol. 13, no. 4, p. e0196072, 2018.

[22] N. S. Lang, L. Mills, R. L. Wample, J. Silbernagel, E. M. Perry, and R. Smithyman, “Remote image and leaf reflectance analysis to evaluate the impact of environmental stress on grape canopy metabolism,” Horttechnology, vol. 10, no. 3, pp. 468–474, 2000.

[23] R. SHOFIYATI1, K. HONDA, N. T. S. WIJESEKERA, and WIDAGDO, “Pemantauan Kekeringan Lahan Pertanian Menggunakan Teknologi Remote Sensing dan SIG di DAS Brantas Hulu,” Jurnal Tanah dan Iklim, no. 20, pp. 24–34, 2002.

[24] W. Li et al., “Outlier detection and removal improves accuracy of machine learning approach to multispectral burn diagnostic imaging,” J Biomed Opt, vol. 20, no. 12, p. 121305, 2015.

[25] Y. L. Pavlov, “Random forests,” Random Forests, pp. 1–122, 2019, doi: 10.1201/9780367816377-11.

[26] A. Susanto, A. Prasetyo, and S. Wening, “Laju Infeksi Ganoderma pada Empat Kelas Tekstur Tanah,” Jurnal Fitopatologi Indonesia, vol. 9, no. 2, pp. 39–46, Apr. 2013, doi: 10.14692/jfi.9.2.39.

[27] A. E. Prasetyo, A. Susanto, and C. Utomo, “Metode penghindaran penyakit busuk pangkal batang kelapa sawit (Ganoderma boninense) dengan sistem lubang tanam besar,” Jurnal Penelitian Kelapa Sawit16, vol. 2, pp. 77–86, 2008.