

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/363174564>

# Detection of White Leaf Disease in Sugarcane Using Machine Learning Techniques over UAV Multispectral Images

Article in *Drones* · September 2022

DOI: 10.3390/drones6090230

CITATIONS

0

READS

2

4 authors, including:



**Amarasingam Narmilan**  
South Eastern University of Sri Lanka

38 PUBLICATIONS 28 CITATIONS

[SEE PROFILE](#)



**Surantha Salgadoe**  
University of New England (Australia)

19 PUBLICATIONS 80 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



International Roundtable on The Impacts of Extreme Natural Events: Science and Technology for Mitigation (IRENE2017) [View project](#)



Assessment of Smart Farming Technologies in Batticaloa District, Sri Lanka [View project](#)

## Article

# Detection of White Leaf Disease in Sugarcane Using Machine Learning Techniques over UAV Multispectral Images

Amarasingam Narmilan <sup>1,2,\*</sup> , Felipe Gonzalez <sup>1</sup> , Arachchige Surantha Ashan Salgadoe <sup>3</sup>  and Kevin Powell <sup>4</sup> 

<sup>1</sup> School of Electrical Engineering and Robotics, Faculty of Engineering, Queensland University of Technology (QUT), 2 George Street, Brisbane City, QLD 4000, Australia

<sup>2</sup> Department of Biosystems Technology, Faculty of Technology, South Eastern University of Sri Lanka, University Park, Oluvil 32360, Sri Lanka

<sup>3</sup> Department of Horticulture and Landscape Gardening, Wayamba University of Sri Lanka, Makandura, Gonawila 60170, Sri Lanka

<sup>4</sup> Sugar Research Australia, P.O. Box 122, Gordonvale, QLD 4865, Australia

\* Correspondence: narmilan.amarasingam@hdr.qut.edu.au

**Abstract:** *Sugarcane white leaf phytoplasma* (white leaf disease) in sugarcane crops is caused by a phytoplasma transmitted by leafhopper vectors. White leaf disease (WLD) occurs predominantly in some Asian countries and is a devastating global threat to sugarcane industries, especially Sri Lanka. Therefore, a feasible and an effective approach to precisely monitoring WLD infection is important, especially at the early pre-visual stage. This work presents the first approach on the preliminary detection of sugarcane WLD by using high-resolution multispectral sensors mounted on small unmanned aerial vehicles (UAVs) and supervised machine learning classifiers. The detection pipeline discussed in this paper was validated in a sugarcane field located in Gal-Oya Plantation, Hingurana, Sri Lanka. The pixelwise segmented samples were classified as ground, shadow, healthy plant, early symptom, and severe symptom. Four ML algorithms, namely XGBoost (XGB), random forest (RF), decision tree (DT), and K-nearest neighbors (KNN), were implemented along with different python libraries, vegetation indices (VIs), and five spectral bands to detect the WLD in the sugarcane field. The accuracy rate of 94% was attained in the XGB, RF, and KNN to detect WLD in the field. The top three vegetation indices (VIs) for separating healthy and infected sugarcane crops are modified soil-adjusted vegetation index (MSAVI), normalized difference vegetation index (NDVI), and excess green (ExG) in XGB, RF, and DT, while the best spectral band is red in XGB and RF and green in DT. The results revealed that this technology provides a dependable, more direct, cost-effective, and quick method for detecting WLD.

**Keywords:** white leaf disease; sugarcane; UAV multispectral images; machine learning; precision agriculture



**Citation:** Narmilan, A.; Gonzalez, F.; Salgadoe, A.S.A.; Powell, K.

Detection of White Leaf Disease in Sugarcane Using Machine Learning Techniques over UAV Multispectral Images. *Drones* **2022**, *6*, 230.

<https://doi.org/10.3390/drones6090230>

Academic Editor: Diego González-Aguilera

Received: 21 August 2022

Accepted: 24 August 2022

Published: 1 September 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Sugarcane (*Saccharum officinarum*) is a tropical plant, and it is the most important sugar extracting crop in Sri Lanka [1,2]. Sugarcane white leaf disease (WLD) is one of the most economically important diseases in Sri Lanka's sugarcane industry [3], and WLD severely progresses in ratoon sugarcane, which ultimately affects yield [4]. WLD is caused by a phytoplasma, an obligate plant parasite that attacks plant phloem tissue. It is transmitted through leafhopper insect vectors [3–5]. Cream-white stripes are developed parallel to the midrib of sugarcane leaves, eventually covering the entire leaf in the infected crops. Other symptoms of WLD include stunted stalks, the absence of lateral shoots on the upper portion of infected stalks, and eventual plant death. Currently, there are no sugarcane varieties found to be resistant to WLD in Sri Lanka [4]. As a preventive approach, growers still follow traditional scouting methods all over the field, monitoring disease symptoms with human eyes and burning infected crops on the spot. However, this method requires

a significant amount of time to watch the entire field to identify infected areas in large field sugarcane plantations. Thus, precision agriculture technologies aided with modern computational machine learning approaches may provide an effective way of detecting sugarcane WLD on-field, an alternative to human-based methods.

Precision agriculture is a smart farming method that uses current technologies to examine and manage changes within an agricultural field to maximize cost-effectiveness, sustainability, and environmental protection [6–8]. Precision agriculture is crucial to seeking low-input, high-efficiency, and sustainable methods in agricultural industries [9]. Recent improvements in the application of UAV-based remote sensing in crop production have proved crucial in improving crop productivity [10]. Remote sensing for precision agriculture is based on the indirect detection of soil and crop reflected radiation in an agricultural field [11]. This approach is well suited for monitoring plant stress and disease since it provides multitemporal and multispectral data. UAVs are increasingly used for agriculture to collect high-resolution images and videos for post-processing. Artificial intelligent (AI) approaches are used to process these UAV images for planning, navigation, and georeferencing, as well as for a variety of agricultural applications [12]. UAVs and advanced computational ML techniques are increasingly used to forecast and improve yield in various farming industries, including sugarcane [10].

León-Rueda et al. [13] examined the use of multispectral cameras mounted on UAVs to classify commercial potato vascular wilt using supervised random forest classification. Su et al. [14] investigated the yellow rust disease in winter wheat using a multispectral camera by selecting spectral bands and SVI with a high discriminating capability. Albetis et al. [15] assessed the possibility of distinguishing Flavescence dorée symptoms using UAV multispectral imaging. Gomez Selvaraj et al. [16] examined the potential of aerial imagery and machine learning approaches for disease identification in bananas by classifying and localizing bananas in mixed-complex African environments using pixel-based classifications and machine learning models. Lan et al. [17] assessed the feasibility of large-area identification of citrus Huanglongbing using remote sensing and committed to improving the accuracy of detection using numerous ML techniques, including support vector machine (SVM), K-nearest neighbor (KNN), and logistic regression (LR). Table 1 represents the application of UAVs for disease management in precision agriculture, and Table 2 shows the use of UAVs for pest and disease control in the sugarcane sector.

ML algorithms have been used to monitor the crop status in many remote sensing applications in agriculture [30–33]. ML methods attempt to establish a relationship between crop parameters to forecast crop production [34]. Artificial neural networks (ANN), random forests (RF), SVM, and decision trees (DT) are relevant algorithms in remote sensing applications [35].

Saini and Ghosh [36] utilized XGBoost (XGB), stochastic gradient boosting (SGB), RF, and SVM for rice mapping crops in India to evaluate the efficacy of ensemble methods. Huang et al. [37] used VIs generated from canopy level hyperspectral scans to examine the utility of the RF technique in combination with the XGB approach for detecting wheat stripe rust early and mid-term. Compared to typical machine learning approaches, the XGB, as a unique ML methodology, can reduce model overfitting and computing effort [37]. Tageldin et al. [38] used the XGB method to predict the occurrence of cotton leaf miner infestation with an accuracy of 84 percent, which was greater than the findings obtained using algorithms, such as RF and logistic regression. The RF non-parametric classifier is an ensemble-based machine learning technique that combines the predictions of many decision tree classifiers using a voting strategy [39]. Santoso et al. [40] assessed the RF model's potential for predicting BSR disease in oil palm fields and produced BSR disease distribution maps. With the cascade parallel random forest (CPRF) algorithm and a 20-year examination of pertinent data, Zhang [41] identified the pattern of rice diseases. Samajpati and Degadwala [42] experimented with identifying apple scab, apple rot, and apple blotch utilizing the RF algorithm. Some of the researchers suggested a model that employs a decision tree to identify and categorize leaf disease and boosts its detection

accuracy while reducing detection time compared to the current system using DT models [43–45]. K-nearest neighbor (KNN) is a prevalent machine learning algorithm that performs well in supervised learning scenarios and simple recognition issues [46]. Vaishnave et al. [47] developed the ML model by KNN algorithm to detect the groundnut leaf disease, and Krithika and Grace [48] used a KNN classifier to identify the grape leaf diseases. Kapil et al. [49] developed a system for recognizing cotton leaf disease by the KNN algorithm.

**Table 1.** Application of UAVs for disease management in precision agriculture.

No	Crop	Disease	Location	UAV Sensor	Reference
01	Citrus	Citrus greening	Iran	Micasense RedEdge camera	[18]
02	Cotton	Leaf Blight Disease	Brazil	Multispectral TetraCam ADC camera	[19]
03	Maize	Maize streak virus disease	Zimbabwe	Parrot Sequoia multispectral camera	[20]
04	Vineyard	Vine disease	France	Survey2 sensor	[21]
05	Cotton	Root rot disease	USA	Micasense RedEdge camera	[22]
06	Vineyards	Grapevine disease	France	Micasense RedEdge camera	[15]
07	Wheat	Helminthosporium leaf blotch (HLB)	China	Phantom 4 RGB camera	[23]
08	Soybean	Soybean leaf diseases	Brazil	Phantom 3 Sony EXMOR sensor	[24]
09	Vine	Esca disease	France	RGB camera	[25]
10	Wheat	Fusarium Head Blight	China	Hyperspectral camera	[26]

**Table 2.** Use of UAVs for pest and disease control in the sugarcane sector.

No	Purpose	UAV Type/Sensor Type/Sprayer Type	Location	Reference
01	Detection of WLD	Six rotors, VESPA HEX 650	MicaSense Multispectral camera Thailand	[5]
02	Pesticide application	Quadrotor: Jimu 3WWDZ-1013 with Centrifugal mist sprayers	China	[2]
03		Multi-rotor UAV (four-rotor electronic UAV, 3WWDZ-10A)	China	[27]
04		A TY-800 single-rotor UAV	China	[28]
05	Monitoring if Mosaic Virus	SX8 multirotor UAS with eight propellers	Hyperspectral camera model DT-0014 Brazil	[29]

Vegetation indices are numerical metrics used in remote sensing applications to assess the differentiation of vegetation cover, vigor, and growth dynamics. A sum, difference, ratio, or other linear combination of reflectance factor or radiance measurements from two or more wavelength intervals normally constitutes the vegetation index. It is utilized to increase the reliability of regional and temporal comparisons of terrestrial photosynthetic activity and canopy structure variation by enhancing the contribution of vegetation features [50]. A VI's ability to detect WLD-infected sugarcane via image processing from a multispectral camera placed on a UAV was examined by Sanseechan et al. [5]. Moriya et al. [29] developed a method for accurately identifying and mapping mosaic virus

in sugarcane using aerial surveys conducted with a UAV equipped with a hyperspectral camera. A few research studies have been undertaken using ML techniques over UAV multispectral images to identify the other sugarcane diseases, and no research studies have been undertaken related to detection of WLD using ML models and high-resolution UAV imagery in sugarcane crops. Therefore, this study proposes developing a method for identifying sugarcane WLD by combining UAV technology with high-resolution multispectral cameras and multiple machine learning classification algorithms. There were two sub-goals: (1) to correlate the VIs with the fluctuation in severity level of WLD in the sugarcane field; and (2) to evaluate the detection performance in WLD severity levels using various ML approaches.

UAV-based remote sensing can assist farmers in analyzing crop health and management in precision agriculture. Early detection of WLD in Sri Lankan sugarcane fields will be used to implement effective management measures throughout the crop’s early phases. This method will aid in disease management in sugarcane farms by eliminating the requirement for conventional methods [51]. Ultimately, it will help farmers and the cane industry in Sri Lanka recover economically. However, the commercial application of UAVs and artificial intelligence algorithms in sugarcane sectors has been limited due to various variables, including technology, UAV legislation, and cost [51].

## 2. Methodology

### 2.1. Process Pipeline

As depicted in Figure 1, a process pipeline with four key components was developed: acquisition, preprocessing, training, and prediction. Images are downloaded, orthorectified, mosaicked, and preprocessed to extract samples with crucial features and then we labelled them. The data were then supplied to supervised machine learning classifiers, trained, and optimized for detection. The complete orthorectified data were then analyzed to determine where WLD crops would grow in the field. Images were collected, orthorectified, and preprocessed to extract samples with essential characteristics and then we labelled them. The data were subsequently sent to supervised machine learning classifiers that had been trained and optimized for detection.

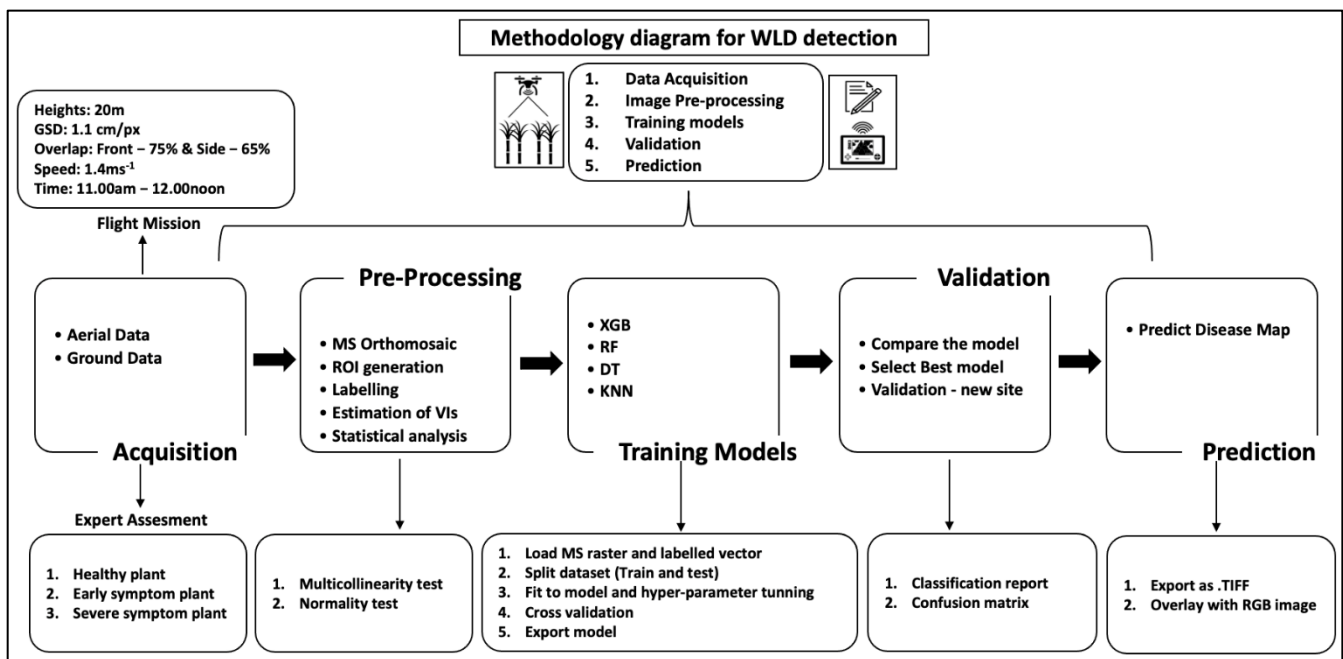


Figure 1. Main steps of the proposed methodology for a single flight campaign.

## 2.2. Study Site

The study was conducted in a 1.24-hectare sugarcane field in Gal-Oya Plantation, Hingurana, Sri Lanka ( $7^{\circ}16'42.94''$  N,  $81^{\circ}42'25.53''$  E) during the sugarcane growing season of October 2021 (Figure 2). For this experiment, two-month-old sugarcane plants with an average height of 1.2 m were chosen.



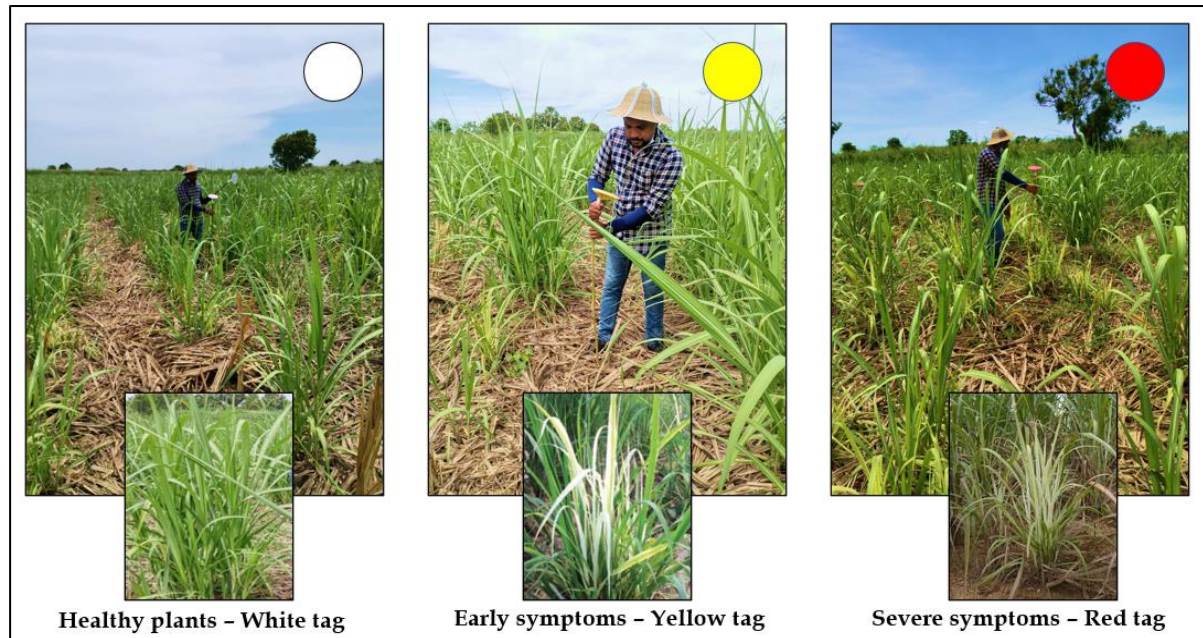
**Figure 2.** Study area.

Disease plants were randomly sampled throughout the field for the levels of disease severities followed by the natural disease occurrence pattern in the field. Field agronomists confirmed the following during this experiment: (1) Ridges and furrows irrigation method was used in the field, and there was no water stress to the plants; (2) the entire experimental site had a uniform soil type (sandy to clay loam soils); (3) fertilizers were applied in the recommended level to the entire experimental field, and there was no fertilizer stress to the plants; and (4) WLD disease was transmitted by insect vector and was not associated with soil or water, and this symptom was developed only by WLD. Due to the above four reasons, it was not necessary to design the experiment for block design in this site.

## 2.3. Ground Truth Data Collection

Experts visually inspected and labelled diseased and healthy plants as ground truth before image acquisition to train and test the classifier [39]. The sugarcane plants (a total of 150 plants) were classified into three types, healthy plants (50 plants), early symptoms plants (50 plants), and severe symptoms plants (50 plants) by using different color tags such as white tag, yellow tag, and red tag, respectively, that were installed in the training site manually as shown in Figure 3. A total of 90 plants were classified into three types, healthy plants (30 plants), early symptoms plants (30 plants), and severe symptoms plants (30 plants), by using color tags in the testing site for validation. Early symptom plants

were characterized by the youngest leaves appearing white with older leaves remaining green. Pure white leaves classify the severe plant symptoms in most leaves with stunted growth [52].



**Figure 3.** Ground truth classification of sugarcane crops in the field.

#### 2.4. UAV Platform

DJI P4 multispectral UAV was used to conduct the experiment in the sugarcane field. DJI P4 Multispectral is a fully integrated UAV platform, and it can complete the data collection task independently without the help of other aircraft. It has a take-off weight of 1487 g, and the average flight time is 27 min. The P4 Multispectral imaging system contains six cameras with 1/2.9-inch CMOS sensors, including an RGB camera that produces images in the JPEG format and a multispectral camera array containing five cameras (Figure 4b) that produce multispectral images in the TIFF format. It uses a global shutter to ensure performance. The five cameras in the multispectral camera array can capture photos in the following imaging bands: Blue (B): 450 nm  $\pm$  16 nm; green (G): 560 nm  $\pm$  16 nm; red (R): 650 nm  $\pm$  16 nm; red edge (RE): 730 nm  $\pm$  16 nm; and near-infrared (NIR):

840 nm  $\pm$  26 nm [53] without zoomable. Table 3 shows the information on central wavelength and wavelength width for DJI P4 multispectral camera [53].



**Figure 4.** UAV and camera used in this study: (a) DJI P4 with remote controller and (b) RGB sensor for visible light imaging and five monochrome sensors (Source [54]).

**Table 3.** Spectral band information for the DJI P4 Multispectral.

Band	Central Wavelength (nm)	Wavelength Width (nm)
Blue	450	32
Green	560	32
Red	650	32
Red edge	730	32
Near-infrared	840	32

Table 4 shows the camera specifications of the DJI P4 Multispectral. The remote controller features, as shown in the Figure 4a, of DJI's long-range transmission technology can control the aircraft and the gimbal cameras at a maximum transmission range of 4.3 mi (7 km). It is possible to connect an iPad to the remote controller via the USB port to use the DJI GS Pro app to plan and perform missions. It can also be used to export the captured images for analysis and create multispectral maps [53].

**Table 4.** Camera specification of the DJI P4 Multispectral.

Camera Component	Specifications
Sensors	Six 1/2.9" CMOS, including one RGB sensor for visible light imaging and five monochrome sensors for multispectral imaging. Each Sensor: Effective pixels 2.08 MP (2.12 MP in total)
Filters	Blue (B): 450 nm $\pm$ 16 nm; Green (G): 560 nm $\pm$ 16 nm; Red (R): 650 nm $\pm$ 16 nm; Red edge (RE): 730 nm $\pm$ 16 nm; Near-infrared (NIR): 840 nm $\pm$ 26 nm
Lenses	FOV (Field of View): 62.7° Focal Length: 5.74 mm (35 mm format equivalent: 40 mm), autofocus set at $\infty$ Aperture: f/2.2
Monochrome Sensor Gain	1–8x
Electronic Global Shutter	1/100–1/20,000 s (visible light imaging); 1/100–1/10,000 s (multispectral imaging)

The RTK module is integrated directly into the Phantom 4 RTK, providing real-time, centimeter-level positioning data for improved absolute accuracy on image metadata. The GSD for the P4 multispectral is (H/18.9) cm/pixel. Height can be calculated based on the accuracy needed for flight mission.

### 2.5. Collection of Multispectral UAV Images

A UAV flying operation was undertaken during the growing season utilizing a DJI P4 multispectral system on a sunny day between 11:00 a.m. and 12:00 p.m. The visible-to-near-infrared spectral range of the DJI P4 multispectral camera comprises five bands with wavelengths of 450.0 nm, 560.0 nm, 650.0 nm, 730.0 nm, and 840.0 nm, respectively (blue, green, red, red edge, and near-infrared). The flying altitude was 20 m and maintained by a barometer present in the UAV. DJI P4 multispectral UAV uses barometer to maintain the altitude. It uses mean barometer to measure air pressure and establish and maintain a stable altitude during flying. Barometer can rapidly measure changes in atmospheric pressure to help ensure the UAV is flying at the appropriate elevation. Additionally, the experiment site is in the level surface. Therefore, it was easy to maintain the height in the same altitude. The size of pixels in terms of real-world dimensions for this experiment was 1.1 cm/pixel.

The UAV was flown at different heights, 10 m, 15 m, 20 m, and 25 m, before conducting the flight mission to select the suitable height needed for labelling the WLD over the multispectral orthomosaic image. The flight campaign at 15 m was selected as captured data, which provided the best outcomes in terms of WLD detection, UAV endurance, and battery capacity. The speed of the UAV and front and side overlap of images were 1.4 m per second and 75% and 65%, respectively. The experiment was conducted between 11:00 a.m. and 12:00 p.m. because plant leaves are erect and at maximum transpiration



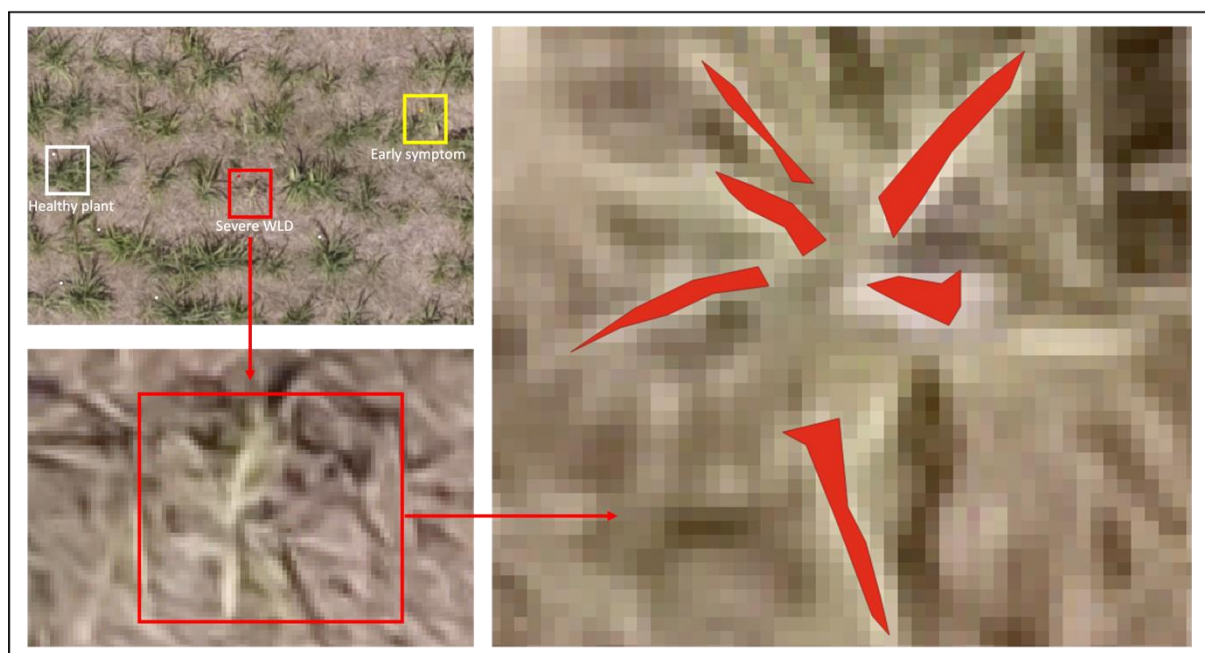
time (active time for plants) at the time of image capture. Early morning and late afternoon or evening are not suitable for conducting this experiment due to dew on the plants in the early morning and dropping of leaves in the late afternoon or evening. Additionally, this will have not an effect on the VIs values. Therefore, the time of image capture is very important for developing the accurate WLD detection models.

## 2.6. Software and Python Libraries

This research was conducted using several software tools and python libraries. Agisoft Metashape (Version 1.6.6; Agisoft LLC, Petersburg, Russia) was used to process, filter, and orthorectify 5600 raw photos for multispectral image analysis. A set of images from cropped regions was extracted and then labelled using QGIS (Version 3.2.0; Open-Source Geospatial Foundation, Chicago, IL, USA). Visual Studio Code (VS Code) 1.70.0 was used as source-code editor to develop the different ML algorithms using the Python 3.8.10 programming language. Several libraries were used for data manipulation and machine learning, including Geospatial Data Abstraction Library (GDAL) 3.0.2, eXtreme Gradient Boosting (XGBoost) 1.5.0, Scikit-learn 0.24.2, OpenCV 4.6.0.66, and Matplotlib 3.4.3.

## 2.7. Data Labelling

A mask for each image was generated by assigning integer values for every highlighted pixel to perform image labelling. The integer values were set as follows: 1 = ground cover; 2 = shadow; 3 = healthy; 4 = early symptoms; and 5 = severe WLD by using QGIS. Each bright colored pixel was filtered from an orthomosaic image. A new shapefile was created to draw the polygons on the multispectral orthomosaic image to label each class by using toggle editing and adding polygon tools in the QGIS. In total, 471,748 pixels were labelled from all the classes based on the ground truth information by observing the different color tags in the orthomosaic image as shown in Figure 3. The edges of the plant leaves were not labelled to prevent the misclassification of mixed pixels. All the selected 150 plants' leaves were labelled by using the polygon tool in QGIS as shown in Figure 5. Ground truth shape files (.shp) were exported for training the different ML models. Shape region in the shape file was converted into labelled pixels using translation techniques before training the data.



**Figure 5.** Pixelwise labelling of severe symptom of WLD plant using QGIS tool.

### 2.8. Statistical Analysis for Algorithm Development

Statistical analysis was conducted using multicollinearity testing and normality testing to select the best fit ML models before tuning them with labelled data. From an initial list of twenty VIs, only six of them were chosen to train the models via multicollinearity testing via variable inflation factors (*VIF*) to avoid model overfitting. Finally, eleven input features (five bands and six VIs) were used to develop the ML models to detect WLD. Variance inflation factor was used to measure how much the variance of the estimated regression coefficient is inflated if the independent variables are correlated [54]. *VIF* is calculated as shown in Equation (1).

$$VIF = \frac{1}{1 - R_i^2} = \frac{1}{Tolerance} \quad (1)$$

where  $R_i^2$  represents the unadjusted coefficient of determination for regressing the *i*-th independent variable on the remaining ones, and tolerance is simply the inverse of the *VIF*. The lower the tolerance, the more likely is the multicollinearity among the variables. The value of *VIF* = 1 indicates that the independent variables are not correlated to each other. If the value of *VIF* is  $1 < VIF < 5$ , it specifies that the variables are moderately correlated to each other. The challenging value of *VIF* is between 5 to 10 as it specifies the highly correlated variables. If  $VIF \geq 5$  to 10, there will be multicollinearity among the predictors in the regression model, and  $VIF > 10$  indicates the regression coefficients are feebly estimated with the presence of multicollinearity [54].

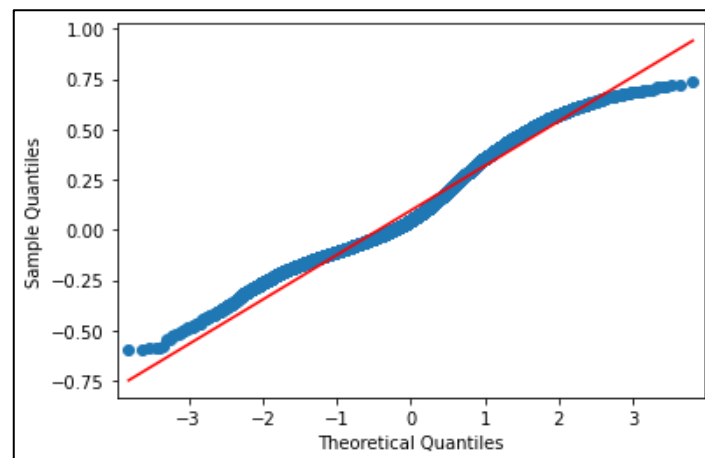
Based on the literature review [54–56], input features, which are not correlated among all input features and moderately correlated among all input features, such as blue, green, red, red edge, NIR, normalized difference vegetation index (NDVI), green normalized difference vegetation index (GNDVI), normalized difference red edge index (NDRE), green chlorophyll index (GCI), modified soil-adjusted vegetation index (MSAVI), and excess green (ExG), were selected to train the models, as shown in the Table 5. Highly correlated input variables, such as leaf chlorophyll index (LCI), difference vegetation index (DVI), ratio vegetation index (RVI), enhanced vegetation index (EVI), triangular vegetation index (TVI), green difference vegetation index (GDVI), normalized green red difference index (NGRDI), atmospherically resistant vegetation index (ARVI), structure insensitive pigment index (SIPI), green optimized soil adjusted vegetation index (GOSAVI), excess red (ExR), excess green red (ExGR), normalized difference index (NDI), and simple ratio index (SRI), were not selected to train the ML models due to higher *VIF* that range from around 7 to 22 [54].

**Table 5.** *VIF* values for selected VIs.

Id	Input Variables	<i>VIF</i>
0	Blue	1.2271
1	Green	1.7911
2	Red	1.1987
3	Red Edge	1.1243
4	NIR	1.5213
5	NDVI	4.0372
6	GNDVI	3.4279
7	NDRE	1.0976
8	GCI	4.4612
9	MSAVI	1.0121
10	ExG	3.0231

A second statistical experiment of normality test was conducted to determine whether sample data have been drawn from a normally distributed population for the development of ML models. Different normality tests, namely quantile-quantile (Q-Q) plot, were conducted to confirm the normal distribution of features. Figure 6 shows the Q-Q plot confirming that the data were adequately close to the theoretical reference line, representing a sound

model fit. The python libraries, such as matplotlib, numpy, statsmodels.graphics.gofplots, and scipy.stats, were used to develop the Q-Q plot.



**Figure 6.** Normal Q-Q plot for the observed sample against theoretical quantiles.

### 2.9. Development of Classification Algorithms and Prediction

The development of algorithms includes multiple steps to load, preprocess, fit the classifier to the data, and prediction. The processing phase converts the read data into a collection of features, which are then analyzed by the classifier as shown in Table 6. An orthomosaic multispectral raster was loaded into the algorithm to calculate spectral indexes and improve the detection rates as mentioned in step 5. For this approach, the VIs, such as ExG, GCI, MSAVI, GNDVI, NDRE, and NDVI, are estimated (step 6) as shown in Table 7. All five bands in the multispectral raster, as well as in the estimated vegetation spectral indexes, are denominated as input features (step 7).

**Table 6.** Steps in algorithms (XGB, RF, DT, and KNN) development-detection and segmentation of WLD using multispectral imagery.

- 
- Step 1**-import required modules and libraries.
  - Step 2**-Load input file (Multispectral images as .tiff) and ground truth file (Ground truth shape file as .shp).
  - Step 3**-Extract the bands (features) from input file (blue, green, red, red edge, and NIR) through GDAL library.
  - Step 4**-Define the wavelength of each band based on the DJI P4 camera (wavelengths = [475.0, 560.0, 668.0, 717.0, 840.0]).
  - Step 5**-Store the bands in the variable (V) as five input features.
  - Step 6**-Estimation of selected VIs (Additional input features-six).
  - Step 7**-Append the VIs and five bands and store in the same variable (V)-Total-11 input features.
  - Step 8**-Search for the number of classes of the labelled data.
  - Step 9**-Filter unlabelled data from the source image and store their values in the 'X' features variable and store in the array (x\_array).
  - Step 10**-Select only labelled data from the labelled image and store their values in the 'y' labels variable and store in the array (y\_array).
  - Step 11**-Splitting the dataset into the 'Training' set (75%) and 'Test' set (25%)
  - Step 12**-Data normalization (Feature Scaling) of the 'X' features matrix for Euclidean distance
  - Step 13**-Fitting Classifier (XGB, RF, DT, and KNN) to the training set (11 input features)
  - Step 14**-manual hyper-parameter tuning based on the algorithms as shown in Table 7.
  - Step 15**-Applying k-fold Cross Validation
  - Step 16**-Export and save the model.
  - Step 17**-Predict the values for each sample in x-array
  - Step 18**-Export the output file as tagged image file (TIF) format.
-

**Table 7.** Estimation of VIs.

No	Vegetation Indices	Formula	References
01	Normalized Difference Vegetation Index (NDVI)	$NDVI = \frac{NIR - R}{NIR + R}$	[57]
02	Green Normalized Difference Vegetation Index (GNDVI)	$GNDVI = \frac{NIR - G}{NIR + G}$	[58–60]
03	Normalized Difference Red Edge Index (NDRE)	$NDRE = \frac{NIR - Red\ Edge}{NIR + Red\ Edge}$	[61–63]
04	Green Chlorophyll Index (GCI)	$GCI = \frac{NIR}{G} - 1$	[64,65]
05	Modified Soil-Adjusted Vegetation Index (MSAVI)	$MASVI = \frac{(2 * NIR + 1 - \sqrt{(2 * NIR + 1)^2 - 8 * (NIR - R)})}{2}$	[66]
06	Excess Green (ExG)	$ExG = \frac{2G - R - B}{R + G + B}$	[25]

The labelled regions from the ground-based assessments are exported from QGIS and loaded into an array (y\_array) (step 10). In all, 471,748 pixelwise samples were filtered and randomly divided into a training array (75%) and a testing array (25%) (step 11). In step 13, data are processed into different ML classifiers. This study employed four (04) machine learning regression methods, XGB, RF, DT, and KNN, to detect the sugarcane WLD from multispectral UAV images. Finally, the fitted model is validated using k-fold cross-validation (step 15). In the prediction stage, unlabelled pixels are processed in the optimized classifier, and their values are displayed in the same 2D spatial image from the orthorectified multispectral raster (step 17). Each image's identified pixels are then colored differently and exported in TIF format, which can be read with geographic information system (GIS) platforms (step 18). The best performant model for identifying WLD in the sugarcane field was selected by comparing performance metrics, such as precision, recall, f1 score, and accuracy. Further details on the calculation of these metrics can be found in Section 3.4.

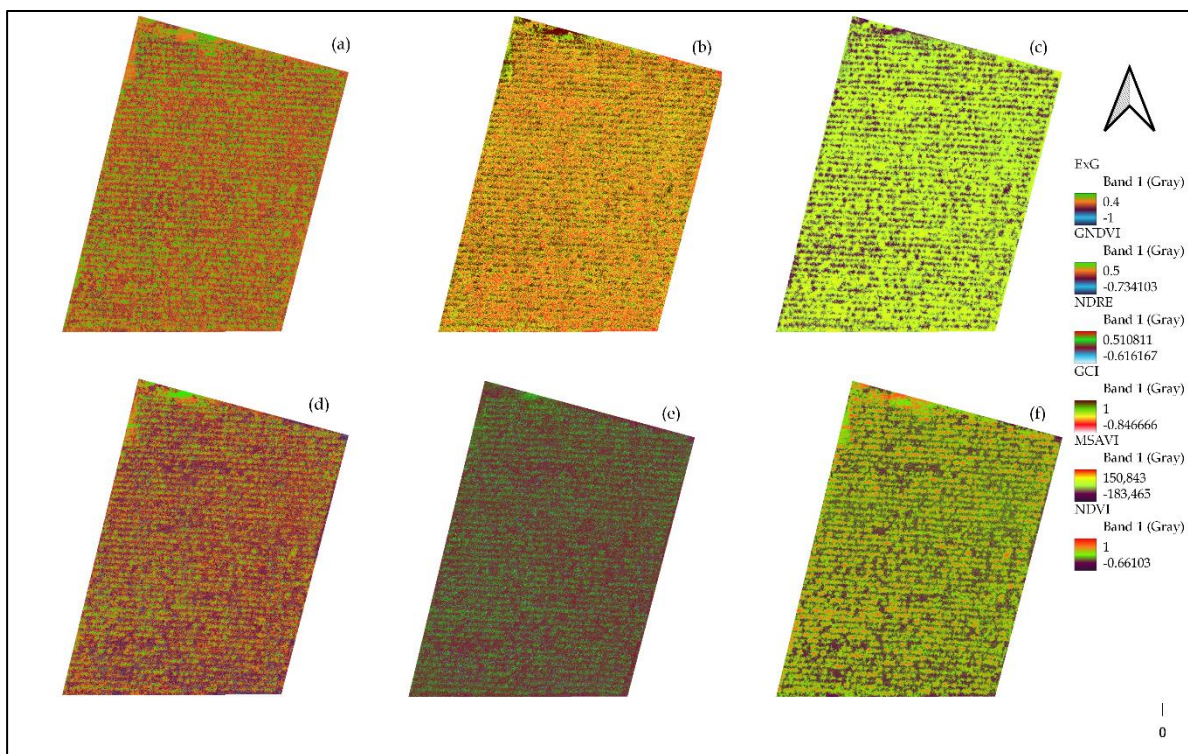
### 2.10. Validation

For validation, 90 sugarcane plants were classified into three types, healthy plants (30 plants), early symptoms plants (30 plants), and severe symptoms plants (30 plants), by using different color tags, such as white tag, yellow tag, and red tag, respectively, as shown in Figure 3, in the testing site. The labelling was performed the same as in the training site that is mentioned in the Section 2.6. Then, a python script was developed to validate the validation accuracy in the testing site. Finally, an input file (multispectral images as .tiff for testing site), a ground truth file (ground truth shape file as .shp for testing site), and a best model file (as .json exported from training) were loaded into different algorithms for estimating the validation accuracy.

## 3. Results

### 3.1. Estimation of Vegetation Indices

Leaf pigments' absorption characteristics govern spectral reflectance. Therefore, any variation in pigment concentrations correlates closely with the health and production of the plant [20]. Six (06) VIs were selected based on the results from multicollinearity testing and variable optimization techniques [51,67] to develop the different ML models to detect the WLD, as shown in Figure 7 and Table 8. To construct the various VIs, reflectance values in multispectral bands corresponding to blue (B): 450 nm ± 16 nm; green (G): 560 nm ± 16 nm; red (R): 650 nm ± 16 nm; red edge (RE): 730 nm ± 16 nm; near-infrared (NIR): 840 nm ± 26 nm were utilized in this study.



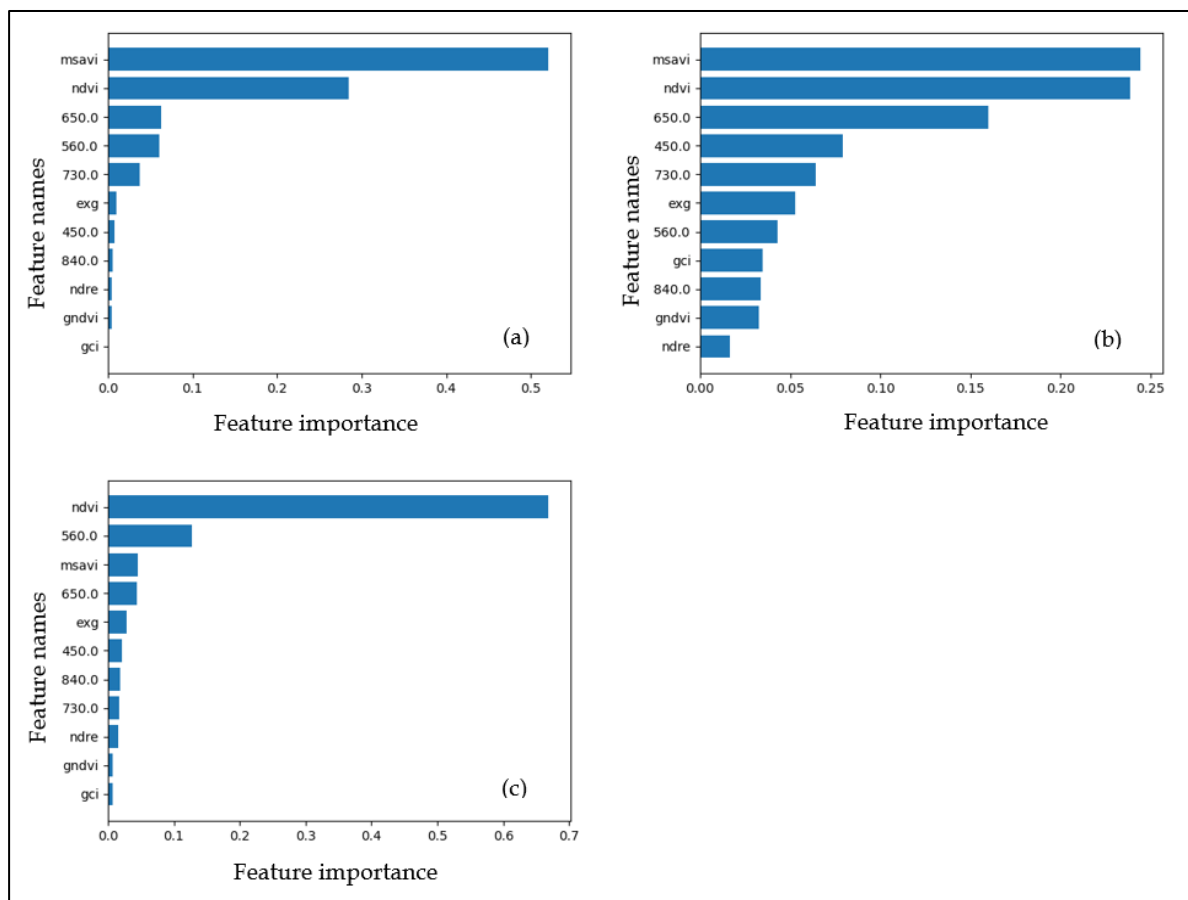
**Figure 7.** Spectral vegetation indices used in this study: (a) ExG, (b) GCI, (c) MSAVI, (d) GNDVI, (e) NDRE, and (f) NDVI.

**Table 8.** Algorithms and their hyper-parameters and specific libraries.

Algorithms	Hyper-Parameters	Specific Libraries
XGB	estimators = 100 learning rate = 0.1 maximum depth = 3	xgboost (import XGBClassifier)
RF	max_depth = 20 random_state = 10 n_estimators: 100 min_samples_split = 2 min_samples_leaf = 1	sklearn.ensemble (import Random Forest Classifier)
DT	max_depth = 20 random_state = 10 min_samples_split = 2 min_samples_leaf = 1	sklearn.tree (import Decision Tree Classifier)
KNN	n_neighbors = 5 p = 2 leaf_size = 30	sklearn.neighbors (import KNeighbors Classifier)

### 3.2. Ranking of Feature Importance

As shown in Figure 8, individual five bands and selected VIs were ranked using feature importance techniques with python programming during the model development. The top five important features in XGB models are MSAVI, NDVI, red, green, and NIR. Moreover, MSAVI, NDVI, red, blue, and NIR were ranked as the top five features in the RF model, while NDVI, green, MSAVI, red, and ExG were the top five in the DT model. However, GCI shows the lowest rank in XGB and DT models while NDRE shows the lowest ranking in the RF model to detect the WLD in the sugarcane field.



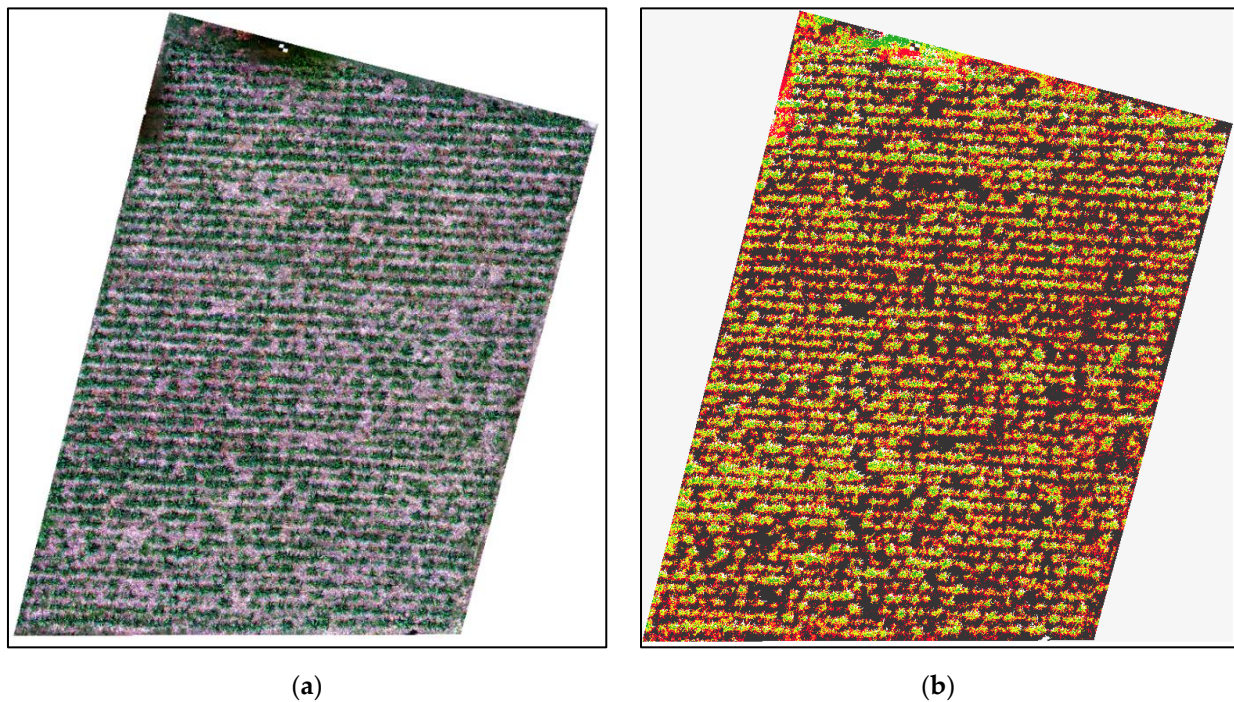
**Figure 8.** Ranking the features. (a): XGB, (b) RF, and (c) DT.

### 3.3. Segmentation Results of the Proposed Approaches

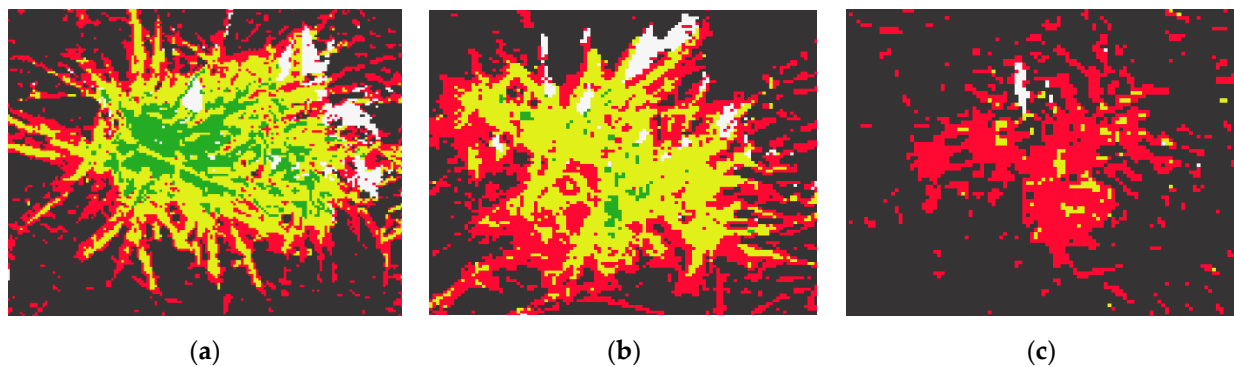
Figure 9a represents the multispectral orthomosaic image generated from the UAV raw images, while Figure 9b shows the WLD spatial map developed by XGB, which is an optimized model among other ML models. The severe WLD plant shows a red color in almost all the canopy areas, as shown in Figure 10c. In early symptom plants, most of the canopy region shows yellow, as shown in Figure 10b, while healthy plants show green color in most of the canopy region (Figure 10a). However, the margin of the canopy shows red color in all the classifications due to dead leaves presented in each sugarcane crop. The spatial distribution of the severity of the WLD of sugarcane is plotted in Figure 9 using different prediction models, such as (a) XGB, (b) RF, (c) DT, and (d) KNN. Segmented images for photo interpretation and accuracy indicators were implemented for validation purposes. In total, 117,937 labelled pixels were evaluated from the test to assess the algorithms. Figure 11 represents the segmentation results of healthy, early symptoms, and severe symptoms of WLD in sugarcane plants for different ML models.

### 3.4. Confusion Matrix and Classification Report

The training performance of various machine learning models, such as XGB, RF, DT, and KNN, was compared over consecutive runs by overall accuracy, F1 score, precision, and recall. The results indicate that all machine learning models performed similarly well in the suggested pipeline for detecting WLD. The classification results indicated that all models achieved high accuracy. The confusion matrix of each model as represented in Table 9 and the classification reports for each model were as shown in Table 10.



**Figure 9.** Segmentation results of the proposed approach: (a) multispectral image and (b) segmentation result of XGB model.



**Figure 10.** Segmented classification of sugarcane plants: (a) healthy plant, (b) early symptom plant, and (c) severe diseased plant.

The descriptors true positive ( $TP$ ), false positive ( $FP$ ), true negative ( $TN$ ), and false-negative ( $FN$ ) were utilized to construct the confusion matrix (Equation (2)) and subsequently calculate the overall accuracy (Equation (3)), precision (Equation (4)), recall (Equation (5)), and F-score (Equation (6)) [68].

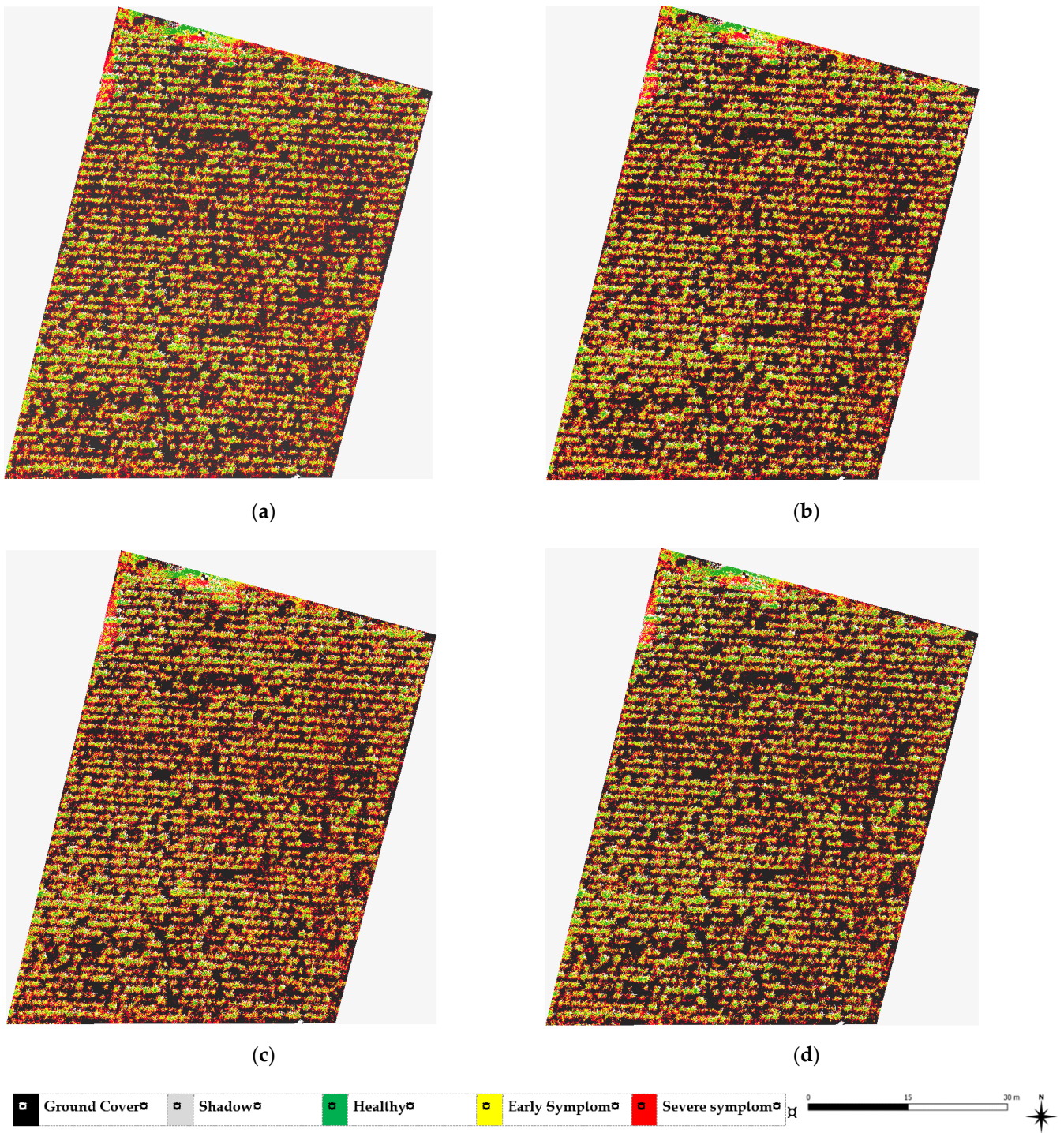
$$\text{Confusion Matrix} = \begin{bmatrix} TP & FP \\ FN & TN \end{bmatrix} \quad (2)$$

$$\text{Overall Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (4)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (5)$$

$$\text{F score} = \frac{2TP}{FP + 2TP + FN} \quad (6)$$



**Figure 11.** Classification and segmentation outputs of all the trained models: (a) XGB, (b) RF, (c) DT, and (d) KNN.



**Table 9.** Confusion matrix of different classifiers in the training site.

XGB	Ground Cover	Shadow	Healthy	Early Symptom	Severe Symptom
Ground Cover	19,248	1	0	2	73
Shadow	5	2029	59	21	21
Healthy	0	39	2552	20	0
Early Symptom	8	26	61	1975	569
Severe symptom	145	37	1	555	2011
RF					
Ground Cover	26,993	5	0	1	107
Shadow	10	2846	86	26	12
Healthy	0	56	3582	30	0
Early Symptom	10	34	84	2747	840
Severe symptom	184	52	1	753	2819
DT					
Ground Cover	26,884	11	0	13	198
Shadow	13	2754	107	46	60
Healthy	0	107	3488	72	1
Early Symptom	13	57	86	2531	1028
Severe symptom	203	43	0	940	2623
KNN					
Ground Cover	19,292	1	0	2	65
Shadow	7	2018	74	18	18
Healthy	0	62	2532	17	0
Early Symptom	16	35	88	1860	631
Severe symptom	171	42	2	609	1925

**Table 10.** Classification report for different ML models in the training site.

XGB	Precision	Recall	F <sub>1</sub> -Score	Accuracy
Ground cover	0.99	1	0.99	
Shadow	0.95	0.95	0.95	
Healthy	0.95	0.98	0.97	0.94
Early symptom	0.77	0.75	0.76	
Severe symptom	0.75	0.73	0.74	
RF				
Ground cover	0.99	1	0.99	
Shadow	0.95	0.96	0.95	
Healthy	0.95	0.98	0.97	0.94
Early symptom	0.77	0.74	0.76	
Severe symptom	0.75	0.74	0.74	
DT				
Ground cover	0.99	0.99	0.99	
Shadow	0.93	0.92	0.93	
Healthy	0.95	0.95	0.95	0.93
Early symptom	0.70	0.68	0.69	
Severe symptom	0.67	0.69	0.68	
KNN				
Ground cover	0.99	1	0.99	
Shadow	0.94	0.95	0.94	
Healthy	0.94	0.97	0.95	0.94
Early symptom	0.74	0.71	0.72	
Severe symptom	0.73	0.7	0.71	

Precision: Ratio between true positives and the sum of true positives and false positives; Recall: Ratio between true positives and the sum of true positives and false negatives.

The results show that 94% of overall accuracy was attained in the XGB, RF, and KNN to detect WLD in the field, even though the DT model also shows good overall accuracy of 93%. Among five classes, ground cover, shadow, and healthy plants were classified with more than 93% of precision, recall, and F<sub>1</sub> scores in all the models. In contrast, early

and severe symptom crops were classified with more than 75% accuracy in XGB and RF models. However, the DT model obtained the lowest precision, recall, and F1 scores of 67%, 69%, and 68% to classify the severe symptom crops. According to the previous studies, Sandino et al. [69] detected healthy and infected trees in the forest with exotic pathogens using the XGboost algorithms with a 97% classification accuracy. Santoso et al. [40] identified healthy and unhealthy oil palms with an overall accuracy of 91% using the RF classifier model. Sandika et al. [70] presented a classification scheme for three grape diseases: anthracnose, powdery mildew, and downy mildew by RF. The proposed system achieved a classification accuracy of 86%. Suresha et al. [71] proposed a method for identifying blast and brown spot diseases in rice with a KNN classifier with an accuracy of 76.59%. Abdulridha et al. [46] developed the KNN algorithms with an overall classification accuracy of 94%, 95%, and 96% to detect citrus canker on tree canopies in the orchard. Zhang et al. [50] built optimal BFW classification models with higher overall accuracy (OA) of 97.28% by RF based on the five multispectral bands.

### 3.5. Testing and Validation in a Different Field at Gal-Oya Plantation

The K-fold cross-validation technique was used to develop the best trained models of XGB, RF, DT, and KNN. Finally, the best models were used to detect the WLD in the testing site located in the same region (Figure 2). In addition to that, validation was performed by observing and labelling the color tags representing the testing site as explained in Section 2.3. Finally, another classification report was developed as shown in Table 11. The results show that 92% of overall accuracy was attained in the XGB, RF, and KNN to detect WLD in the different fields by using the same ML models while 91% of accuracy was obtained by the DT model.

**Table 11.** Classification report for different ML models in the testing site.

	XGB	Precision	Recall	F <sub>1</sub> -Score	Accuracy
	Healthy	0.93	0.97	0.95	0.92
	Early symptom	0.75	0.74	0.73	
	Severe symptom	0.72	0.72	0.71	
	RF				0.92
	Healthy	0.91	0.93	0.96	
	Early symptom	0.74	0.71	0.74	
	Severe symptom	0.71	0.72	0.71	
	DT				0.91
	Healthy	0.93	0.92	0.93	
	Early symptom	0.68	0.66	0.68	
	Severe symptom	0.69	0.65	0.67	
	KNN				0.92
	Healthy	0.89	0.94	0.94	
	Early symptom	0.73	0.70	0.70	
	Severe symptom	0.71	0.67	0.69	

### 3.6. Model Training Time at the Training Site

The training times for each approach under the computer capacity of the 11th Gen Intel(R) Core (TM) i7-1185G7 @ 3.00GHz, 1805 MHz, 4 Core(s), 8 Logical Processor(s), and 16.0 GB RAM in Microsoft Windows 10 Enterprise are listed in Table 12. The most accurate team, XGB, also had the smallest training time, nine minutes. KNN's training took the most time, 29 min, but had the same total accuracy as XGB. RF and DT had the same overall accuracy of 94% and 93%, respectively, with 15 and 18 min of training duration.

**Table 12.** The training time of the XGB, RF, DT, and KNN.

Classifier	Training Time (Minutes)
XGB	9
RF	15
DT	18
KNN	29

#### 4. Discussion

The current study demonstrates a viable strategy for detecting WLD in sugarcane fields by UAVs and machine learning-based classification models. This methodology will give a realistic, accurate, and efficient method for determining the presence of WLD in vast sugarcane fields. VIs are crucial for developing the best classification algorithms because diseases cause changes in the color, water content, and cell structure of the leaves, which are reflected in the spectrum [72]. Pigment changes cause visible spectral responses, while changes in cell structure cause near-infrared spectral responses. Initially, twenty VIs were selected, and only six VIs were chosen via multicollinearity testing and feature selection techniques to minimize the training time and resource requirements of the computer because training time is crucial for model evaluation and to avoid model overfitting. However, the UAV-derived spectral bands and indices used in this work are not disease-specific; hence, they can only measure different infestation levels or damage when a single disease impacts the crop, as they cannot differentiate between their different types of diseases.

Feature selection is important to attain a higher classification accuracy with less training time. However, it isn't easy to obtain the best time and accuracy, and hence a balance must be established based on users' requirements. Different color tags were used for ground truth measurements for post image processing of labelling. However, a handheld GPS meter with higher accuracy can locate each class due to the unavailability of high accuracy GPS meters. However, it is a good method for validating the prediction results in the segmented images. Since conducting ground truth investigations into plant diseases needs professional competence and is time- and labor-intensive, most of the research relied heavily on sampling surveys, as did the evaluation outcomes [50]. Two-month-old sugarcane plants were selected in this study because young plants are highly affected by WLD in the sugarcane industries. Early detection of illnesses is crucial for successful mitigation actions [46]. However, this study should continue further in various sugarcane crop stages. Additionally, flight missions should be conducted in different climatic seasons in different sugarcane varieties to find the severity level of this incident.

Variable optimization was implemented, and just six variables were deemed essential for developing the various ML prediction models. During the optimization process, all ML models eliminated less significant variables. Other studies have indicated that excluding insignificant factors improves the classification performance of machine learning. When variable relevance is very low, the variable is either unimportant or substantially collinear with another variable or variables. Based on the five-band pictures and VIs, the selected ML models, such as XGB, RF, DT, and KNN, produced distribution maps with comparable results. In addition, the overall classification accuracy of this study employing multispectral VIs produced from UAVs is equivalent to similar studies described in the preceding sections. RF has great precision, excellent outlier tolerance, and parameter selection. León-Rueda et al. [13] also used the RF classifier for the classification process. Lan et al. [17] evaluated the feasibility of monitoring citrus Huanglongbing (HLB) by using multispectral images, VIs, and KNN algorithm because KNN is one of the simplest classification algorithms available. It may be used to solve classification and regression predicting problems with extremely competitive results [46]. The DT algorithm tends to have more numerical features in the classification results for data within consistent sample sizes in each category [73]. However, as a result, XGB was chosen as the ideal technology for monitoring WLD in the sugarcane field because it is highly flexible and works well in small to medium datasets.

Therefore, the best prediction model was developed with high accuracy within a short training time.

In the segmentation results, the margin of all the crops was shown in red color due to the dead leaves. Therefore, precision, recall, and F1 score for early and severe symptoms were reduced during the training process. It is a limitation of this study. However, severely diseased plants can be detected easily if the segmented crop canopy is covered completely with red color. Therefore, further research should be conducted to determine the usefulness of deep learning algorithms for detecting WLD in sugarcane fields. Only four ML algorithms were selected in this study based on the previous studies mentioned in Section 1. However, other ML models, such as SVM and LR, can be developed to detect the WLD while comparing with existing models. In addition to these research gaps, high-resolution hyperspectral cameras can improve accuracy, and disease-specific VIs should be developed to detect the specific disease in the sugarcane field.

## 5. Conclusions

This research utilized multispectral UAV images and machine learning methods to detect WLD in a sugarcane field. High-resolution multispectral images and pixel-by-pixel classification answered the need for precise and efficient detection and segmentation approaches for WLD monitoring. The classification performance of four machine learning (ML) methods (XGB, RF, DT, and KNN) was comprehensively evaluated from multiple perspectives, including classification accuracies based on pixel scale and plant scale, the degree of agreement with the ground truth density maps, and the identified areas of infection. The total accuracy of all ML models for five-band pictures was greater than 93%. The experimental results reveal that both XGB and RF performed well in classification. DT, however, demonstrated the lowest classification performance. The five-multispectral-band XGB model with a higher OA of 94% and a faster running duration of nine minutes was deemed the best-supervised model. This study's findings could guide sugarcane plantation management for disease identification by pinpointing the precise location of infected areas in sugarcane fields.

**Author Contributions:** As the corresponding author, A.N. conducted the UAV flight mission and analysis and prepared the work for final submission. F.G. provided general oversight and contributed to the writing and editing of the document. A.S.A.S. gives technical direction for the UAV flight mission and research design and feedback on the publication draft. K.P. contributed to the manuscript's editing. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Acknowledgments:** The authors thank the Gal-Oya Plantation in Sri Lanka for allowing the UAV flight operation and collecting ground truth data. In addition, the authors are extremely appreciative to the Centre for Agriculture and the Bioeconomy (CAB), QUT, SEUSL, AHEAD, and the World Bank for providing us with a scholarship to complete this experiment successfully. Lastly, the authors would like to thank friends and co-workers for their assistance throughout the experiment. We are indebted to the anonymous reviewers and editors for their insightful remarks on our article.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Braithwaite, K.S.; Croft, B.J.; Magarey, R.C. Progress in Identifying the Cause of Ramu Stunt Disease of Sugarcane. *Proc. Aust. Soc. Sugar Cane Technol.* **2007**, *29*, 235–241.
2. Wang, X.; Zhang, R.; Shan, H.; Fan, Y.; Xu, H.; Huang, P.; Li, Z.; Duan, T.; Kang, H.; Huang, Y.; et al. UAV control of major sugarcane disease and pest. *Agric. Biotechnol.* **2019**, *8*, 48–51.

3. Chanchala, K.M.G.; Dayasena, Y.A.P.K.; Wanasinghe, V.K.A.S.M.; Hemachandra, K.S.; Nugaliyadde, L.; Witharama, W.R.G. Viruliferous Nature of the Sugarcane White Leaf Disease Vector; *Deltocephalus Menoni* (Hemiptera: Cicadellidae). In Proceedings of the Seventh Symposium on Plantation Crop Research—Towards Achieving Sustainable Development Goals in the Plantation Sector, Colombo, Sri Lanka, 4–6 November 2019; Rubber Research Institute of Sri Lanka: Agalawatta, Sri Lanka, 2019; pp. 1583–1590.
4. Wickramasinghe, K.P.; Wijesuriya, A.; Ariyawansa, B.D.S.K.; Perera, A.M.M.S.; Chanchala, K.M.G.; Manel, D.; Chandana, R.A.M. Performance of Sugarcane Varieties in a White Leaf Disease (WLD)—Prone Environment at Pelwatte. Available online: <http://sugarres.lk/wp-content/uploads/2020/05/Best-Paper-Award----Seventh-Symposium-on-Plantation-Crop-Research-2019.pdf> (accessed on 8 December 2021).
5. Sansechan, P.; Saengprachathanarug, K.; Posom, J.; Wongpichet, S.; Chea, C.; Wongphati, M. Use of vegetation indices in monitoring sugarcane white leaf disease symptoms in sugarcane field using multispectral UAV aerial imagery. *IOP Conf. Ser. Earth Environ. Sci.* **2019**, *301*, 012025. [[CrossRef](#)]
6. Narmilan, A.; Puvanitha, N. Mitigation Techniques for Agricultural Pollution by Precision Technologies with a Focus on the Internet of Things (IoTs): A Review. *Agric. Rev.* **2020**, *41*, 279–284. [[CrossRef](#)]
7. Narmilan, A.; Niroash, G.; Sumangala, K. Assessment on Consequences and Benefits of the Smart Farming Techniques in Batticaloa District, Sri Lanka. *Int. J. Res. Publ.* **2020**, *61*, 14–20. [[CrossRef](#)]
8. Narmilan, A. E-Agricultural Concepts for Improving Productivity: A Review. *Sch. J. Eng. Technol.* **2017**, *5*, 10–17. [[CrossRef](#)]
9. Mazzia, V.; Comba, L.; Khaliq, A.; Chiaberge, M.; Gay, P. UAV and Machine Learning Based Refinement of a Satellite-Driven Vegetation Index for Precision Agriculture. *Sensors* **2020**, *20*, 2530. [[CrossRef](#)]
10. Amarasingam, N.; Salgadoe, A.S.A.; Powell, K.; Gonzalez, L.F.; Natarajan, S. A review of UAV platforms, sensors, and applications for monitoring of sugarcane crops. *Remote Sens. Appl. Soc. Environ.* **2022**, *26*, 100712. [[CrossRef](#)]
11. Kim, H.; Kim, W.; Kim, S. Damage Assessment of Rice Crop after Toluene Exposure Based on the Vegetation Index (VI) and UAV Multispectral Imagery. *Remote Sens.* **2020**, *13*, 25. [[CrossRef](#)]
12. García, L.; Parra, L.; Jimenez, J.; Lloret, J.; Mauri, P.; Lorenz, P. DronAway: A Proposal on the Use of Remote Sensing Drones as Mobile Gateway for WSN in Precision Agriculture. *Appl. Sci.* **2020**, *10*, 6668. [[CrossRef](#)]
13. León-Rueda, W.A.; León, C.; Caro, S.G.; Ramírez-Gil, J.G. Identification of diseases and physiological disorders in potato via multispectral drone imagery using machine learning tools. *Trop. Plant Pathol.* **2021**, *47*, 152–167. [[CrossRef](#)]
14. Su, J.; Liu, C.; Coombes, M.; Hu, X.; Wang, C.; Xu, X.; Li, Q.; Guo, L.; Chen, W.-H. Wheat yellow rust monitoring by learning from multispectral UAV aerial imagery. *Comput. Electron. Agric.* **2018**, *155*, 157–166. [[CrossRef](#)]
15. Albetis, J.; Duthoit, S.; Guttler, F.; Jacquin, A.; Goulard, M.; Poilvé, H.; Féret, J.-B.; Dedieu, G. Detection of Flavescente dorée Grapevine Disease Using Unmanned Aerial Vehicle (UAV) Multispectral Imagery. *Remote Sens.* **2017**, *9*, 308. [[CrossRef](#)]
16. Selvaraj, M.G.; Vergara, A.; Montenegro, F.; Ruiz, H.A.; Safari, N.; Raymaekers, D.; Ocimati, W.; Ntamwira, J.; Tits, L.; Omondi, A.B.; et al. Detection of banana plants and their major diseases through aerial images and machine learning methods: A case study in DR Congo and Republic of Benin. *ISPRS J. Photogramm. Remote Sens.* **2020**, *169*, 110–124. [[CrossRef](#)]
17. Lan, Y.; Huang, Z.; Deng, X.; Zhu, Z.; Huang, H.; Zheng, Z.; Lian, B.; Zeng, G.; Tong, Z. Comparison of machine learning methods for citrus greening detection on UAV multispectral images. *Comput. Electron. Agric.* **2020**, *171*, 105234. [[CrossRef](#)]
18. DadrasJavan, F.; Samadzadegan, F.; Pourazar, S.H.S.; Fazeli, H. UAV-based multispectral imagery for fast Citrus Greening detection. *J. Plant Dis. Prot.* **2019**, *126*, 307–318. [[CrossRef](#)]
19. Xavir, T.W.F.; Souto, R.N.V.; Statella, T.; Galbieri, R.; Santos, E.S.; Suli, G.S.; Zeilhofer, P. Identification of ramularia leaf blight cotton disease infection levels by multispectral, multiscale uav imagery. *Drones* **2019**, *3*, 33. [[CrossRef](#)]
20. Chivasa, W.; Mutanga, O.; Biradar, C. UAV-based multispectral phenotyping for disease resistance to accelerate crop improvement under changing climate conditions. *Remote Sens.* **2020**, *12*, 2445. [[CrossRef](#)]
21. Kerkech, M.; Hafiane, A.; Canals, R. Vine disease detection in UAV multispectral images using optimized image registration and deep learning segmentation approach. *Comput. Electron. Agric.* **2020**, *174*, 105446. [[CrossRef](#)]
22. Wang, T.; Thomasson, J.A.; Yang, C.; Isakeit, T.; Nichols, R.L. Automatic classification of cotton root rot disease based on uav remote sensing. *Remote Sens.* **2020**, *12*, 1310. [[CrossRef](#)]
23. Huang, H.; Deng, J.; Lan, Y.; Yang, A.; Zhang, L.; Wen, S.; Zhang, H.; Zhang, Y.; Deng, Y. Detection of helminthosporium leaf blotch disease based on UAV Imagery. *Appl. Sci.* **2019**, *9*, 558. [[CrossRef](#)]
24. Tetila, E.C.; Machado, B.B.; Menezes, G.K.; Oliveira, A.D.S.; Alvarez, M.; Amorim, W.P.; Belete, N.A.; Da Silva, G.G.; Pistori, H. Automatic Recognition of Soybean Leaf Diseases Using UAV Images and Deep Convolutional Neural Networks. *IEEE Geosci. Remote Sens. Lett.* **2019**, *17*, 903–907. [[CrossRef](#)]
25. Kerkech, M.; Hafiane, A.; Canals, R. Deep leaning approach with colorimetric spaces and vegetation indices for vine diseases detection in UAV images. *Comput. Electron. Agric.* **2018**, *155*, 237–243. [[CrossRef](#)]
26. Xiao, Y.; Dong, Y.; Huang, W.; Liu, L.; Ma, P. Wheat fusarium head blight detection using uav-based spectral and texture features in optimal window size. *Remote Sens.* **2021**, *13*, 2437. [[CrossRef](#)]
27. Zhang, X.-Q.; Liang, Y.-J.; Qin, Z.-Q.; Li, D.-W.; Wei, C.-Y.; Wei, J.-J.; Li, Y.-R.; Song, X.-P. Application of Multi-rotor Unmanned Aerial Vehicle Application in Management of Stem Borer (Lepidoptera) in Sugarcane. *Sugar Tech* **2019**, *21*, 847–852. [[CrossRef](#)]
28. Zhang, P.; Zhang, W.; Sun, H.; Fu, H.; Liu, J. Effect of the downwash flow field of a single-rotor uav on droplet velocity in sugarcane plant protection. *Engenharia Agrícola* **2021**, *41*, 235–244. [[CrossRef](#)]

29. Moriya, E.A.S.; Imai, N.N.; Tommaselli, A.M.G.; Miyoshi, G.T. Mapping Mosaic Virus in Sugarcane Based on Hyperspectral Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *10*, 740–748. [CrossRef]
30. Tsouros, D.C.; Bibi, S.; Sarigiannidis, P.G. A review on UAV-based applications for precision agriculture. *Information* **2019**, *10*, 349. [CrossRef]
31. Marin, D.B.; Ferraz, G.A.E.S.; Santana, L.S.; Barbosa, B.D.S.; Barata, R.A.P.; Osco, L.P.; Ramos, A.P.M.; Guimarães, P.H.S. Detecting coffee leaf rust with UAV-based vegetation indices and decision tree machine learning models. *Comput. Electron. Agric.* **2021**, *190*, 106476. [CrossRef]
32. De Rosa, D.; Basso, B.; Fasiolo, M.; Friedl, J.; Fulkerson, B.; Grace, P.R.; Rowlings, D.W. Predicting pasture biomass using a statistical model and machine learning algorithm implemented with remotely sensed imagery. *Comput. Electron. Agric.* **2020**, *180*, 105880. [CrossRef]
33. Puig Garcia, E.; Gonzalez, F.; Hamilton, G.; Grundy, P. Assessment of Crop Insect Damage Using Unmanned Aerial Systems: A Machine Learning Approach. In Proceedings of the MODSIM 2015, 21st International Congress on Modelling and Simulation, Gold Coast, Australia, 24 November–4 December 2015; Available online: <http://www.mssanz.org.au/modsim2015/F12/puig.pdf> (accessed on 14 January 2022).
34. Feng, L.; Zhang, Z.; Ma, Y.; Du, Q.; Williams, P.; Drewry, J.; Luck, B. Alfalfa Yield prediction using uav-based hyperspectral imagery and ensemble learning. *Remote Sens.* **2020**, *12*, 2028. [CrossRef]
35. Osco, L.P.; Ramos, A.P.M.; Pereira, D.R.; Moriya, A.S.; Imai, N.N.; Matsubara, E.T.; Estrabis, N.; de Souza, M.; Junior, J.M.; Gonçalves, W.N.; et al. Predicting canopy nitrogen content in citrus-trees using random forest algorithm associated to spectral vegetation indices from Uav-imagery. *Remote Sens.* **2019**, *11*, 2925. [CrossRef]
36. Saini, R.; Ghosh, S.K. Crop classification in a heterogeneous agricultural environment using ensemble classifiers and single-date Sentinel-2A imagery. *Geocarto Int.* **2019**, *36*, 2141–2159. [CrossRef]
37. Huang, L.; Liu, Y.; Huang, W.; Dong, Y.; Ma, H.; Wu, K.; Guo, A. Combining Random Forest and XGBoost Methods in Detecting Early and Mid-Term Winter Wheat Stripe Rust Using Canopy Level Hyperspectral Measurements. *Agriculture* **2022**, *12*, 74. [CrossRef]
38. Tageldin, A.; Adly, D.; Mostafa, H.; Mohammed, H.S. Applying Machine Learning Technology in the Prediction of Crop Infestation with Cotton Leafworm in Greenhouse. *Biorxiv* **2020**, 1–26. [CrossRef]
39. Pourazar, H.; Samadzadegan, F.; Javan, F.D. Aerial multispectral imagery for plant disease detection: Radiometric calibration necessity assessment. *Eur. J. Remote Sens.* **2019**, *52*, 17–31. [CrossRef]
40. Santoso, H.; Tani, H.; Wang, X. Random Forest classification model of basal stem rot disease caused by *Ganoderma boninense* in oil palm plantations. *Int. J. Remote Sens.* **2017**, *38*, 4683–4699. [CrossRef]
41. Zhang, L.; Xie, L.; Wang, Z.; Huang, C. Cascade Parallel Random Forest Algorithm for Predicting Rice Diseases in Big Data Analysis. *Electronics* **2022**, *11*, 1079. [CrossRef]
42. Samajpati, B.J.; Degadwala, S.D. Hybrid Approach for Apple Fruit Diseases Detection and Classification Using Random Forest Classifier. In Proceedings of the 2016 International Conference on Communication and Signal Processing (ICCSPP), Melmaruvathur, India, 6–8 April 2016; pp. 1015–1019. [CrossRef]
43. Rajesh, B.; Vishnu Sai Vardhan, M.; Sujihelen, L. Leaf Disease Detection and Classification by Decision Tree. In Proceedings of the 4th International Conference on Trends in Electronics and Informatics (ICOEI 2020), Tirunelveli, India, 15–17 June 2020; pp. 705–708.
44. Bhatia, G.G.; Singh, A.; Chug, A.; Singh, A.; Bhatia, A.; Singh, A.P. Refactoring and its effects on Maintainability View project Advance Shadow Edge Detection and Removal (ASEDR) View project plant disease detection for high dimensional imbalanced dataset using an enhanced decision tree approach. *Int. J. Future Gener. Commun. Netw.* **2020**, *13*, 71–78. [CrossRef]
45. Sabrol, H.; Kumar, S. Intensity based feature extraction for tomato plant disease recognition by classification using decision tree. *Int. J. Comput. Sci. Inf. Secur.* **2016**, *14*, 622–626. Available online: <https://sites.google.com/site/ijcsis> (accessed on 28 April 2022).
46. Abdulridha, J.; Batuman, O.; Ampatzidis, Y. UAV-Based remote sensing technique to detect citrus canker disease utilizing hyperspectral imaging and machine learning. *Remote Sens.* **2019**, *11*, 1373. [CrossRef]
47. Vaishnave, M.P.; Srinivasan, P.; Suganya Dev, K.; ArutPerumJothi, G. Detection and Classification of Groundnut Leaf Diseases Using KNN Classifier. In Proceedings of the 2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN), Pondicherry, India, 29–30 March 2019.
48. Krithika, N.; Grace Selvarani, A. An Individual Grape Leaf Disease Identification Using Leaf Skeletons and KNN Classification. In Proceedings of the 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), Coimbatore, India, 17–18 March 2017.
49. Prashar, K.; Talwar, R.; Kant, C. CNN based on Overlapping Pooling Method and Multi-layered Learning with SVM & KNN for American Cotton Leaf Disease Recognition. In Proceedings of the International Conference on Automation, Computational and Technology Management (ICACTM), Coimbatore, India, 24–25 April 2019; pp. 330–333.
50. Zhang, S.; Li, X.; Ba, Y.; Lyu, X.; Zhang, M.; Li, M. Banana Fusarium Wilt Disease Detection by Supervised and Unsupervised Methods from UAV-Based Multispectral Imagery. *Remote Sens.* **2022**, *14*, 1231. [CrossRef]
51. Narmilan, A.; Gonzalez, F.; Salgadoe, A.S.A.; Kumarasiri, U.W.L.M.; Weerasinghe, H.A.S.; Kulasekara, B.R. Predicting Canopy Chlorophyll Content in Sugarcane Crops Using Machine Learning Algorithms and Spectral Vegetation Indices Derived from UAV Multispectral Imagery. *Remote Sens.* **2022**, *14*, 1140. [CrossRef]

52. Sugar Research Australia (SRA). WLD Information Sheet. Available online: [sugarresearch.com.au](http://sugarresearch.com.au) (accessed on 13 April 2022).
53. P4 Multispectral—Specifications—DJI. Available online: <https://www.dji.com/au/p4-multispectral/specs> (accessed on 10 August 2022).
54. Shrestha, N. Detecting Multicollinearity in Regression Analysis. *Am. J. Appl. Math. Stat.* **2020**, *8*, 39–42. [[CrossRef](#)]
55. Daoud, J.I. Multicollinearity and Regression Analysis. *J. Phys. Conf. Ser.* **2017**, *949*, 012009. [[CrossRef](#)]
56. Senaviratna, N.A.M.R.; Cooray, T.M.J.A. Diagnosing Multicollinearity of Logistic Regression Model. *Asian J. Probab. Stat.* **2019**, *5*, 1–9. [[CrossRef](#)]
57. Imran, A.; Khan, K.; Ali, N.; Ahmad, N.; Ali, A.; Shah, K. Narrow band based and broadband derived vegetation indices using Sentinel-2 Imagery to estimate vegetation biomass. *Glob. J. Environ. Sci. Manag.* **2020**, *6*, 97–108. [[CrossRef](#)]
58. Marcial-Pablo, M.D.J.; Gonzalez-Sanchez, A.; Jimenez-Jimenez, S.I.; Ontiveros-Capurata, R.E.; Ojeda-Bustamante, W. Estimation of vegetation fraction using RGB and multispectral images from UAV. *Int. J. Remote Sens.* **2018**, *40*, 420–438. [[CrossRef](#)]
59. Scher, C.L.; Karimi, N.; Glasenhardt, M.; Tuffin, A.; Cannon, C.H.; Scharenbroch, B.C.; Hipp, A.L. Application of remote sensing technology to estimate productivity and assess phylogenetic heritability. *Appl. Plant Sci.* **2020**, *8*, e11401. [[CrossRef](#)]
60. Avola, G.; Di Gennaro, S.F.; Cantini, C.; Riggi, E.; Muratore, F.; Tornambè, C.; Matese, A. Remotely sensed vegetation indices to discriminate field-grown olive cultivars. *Remote Sens.* **2019**, *11*, 1242. [[CrossRef](#)]
61. Boiarskii, B. Comparison of NDVI and NDRE Indices to Detect Differences in Vegetation and Chlorophyll Content. *J. Mech. Contin. Math. Sci.* **2019**, *4*, 20–29. [[CrossRef](#)]
62. Eitel, J.U.H.; Vierling, L.A.; Litvak, M.E.; Long, D.S.; Schulthess, U.; Ager, A.A.; Krofcheck, D.J.; Stoscheck, L. Broadband, red-edge information from satellites improves early stress detection in a New Mexico conifer woodland. *Remote Sens. Environ.* **2011**, *115*, 3640–3646. [[CrossRef](#)]
63. Zhang, J.; Wang, C.; Yang, C.; Xie, T.; Jiang, Z.; Hu, T.; Luo, Z.; Zhou, G.; Xie, J. Assessing the effect of real spatial resolution of in situ UAV multispectral images on seedling rapeseed growth monitoring. *Remote Sens.* **2020**, *12*, 1207. [[CrossRef](#)]
64. Kumar, V.; Sharma, A.; Bhardwaj, R.; Thukral, A.K. Comparison of different reflectance indices for vegetation analysis using Landsat-TM data. *Remote Sens. Appl. Soc. Environ.* **2018**, *12*, 70–77. [[CrossRef](#)]
65. Gitelson, A.A.; Viña, A.; Arkebauer, T.J.; Rundquist, D.C.; Keydan, G.; Leavitt, B. Remote estimation of leaf area index and green leaf biomass in maize canopies. *Geophys. Res. Lett.* **2003**, *30*, 1248. [[CrossRef](#)]
66. Xue, J.; Su, B. Significant remote sensing vegetation indices: A review of developments and applications. *J. Sens.* **2017**, *2017*, 1353691. [[CrossRef](#)]
67. Chang, A.; Yeom, J.; Jung, J.; Landivar, J. Comparison of canopy shape and vegetation indices of citrus trees derived from UAV multispectral images for characterization of citrus greening disease. *Remote Sens.* **2020**, *12*, 4122. [[CrossRef](#)]
68. Ampatzidis, Y.; Partel, V. UAV-based high throughput phenotyping in citrus utilizing multispectral imaging and artificial intelligence. *Remote Sens.* **2019**, *11*, 410. [[CrossRef](#)]
69. Sandino, J.; Pegg, G.; Gonzalez, F.; Smith, G. Aerial Mapping of Forests Affected by Pathogens Using UAVs, Hyperspectral Sensors, and Artificial Intelligence. *Sensors* **2018**, *18*, 944. [[CrossRef](#)]
70. Sandika, B.; Avil, S.; Sanat, S.; Srinivasu, P. Random Forest Based Classification of Diseases in Grapes from Images Captured in Uncontrolled Environments. In Proceedings of the 2016 IEEE 13th International Conference on Signal Processing (ICSP), Chengdu, China, 6–10 November 2016; pp. 1775–1780. [[CrossRef](#)]
71. Suresha, M.; Shreekanth, K.N.; Thirumalesh, B.V. Recognition of Diseases in Paddy Leaves Using kNN Classifier. In Proceedings of the 2017 2nd International Conference for Convergence in Technology (I2CT), Mumbai, India, 7–9 April 2017; pp. 663–666.
72. Guo, A.T.; Huang, W.J.; Dong, Y.Y.; Ye, H.C.; Ma, H.Q.; Liu, B.; Wu, W.B.; Ren, Y.; Ruan, C.; Geng, Y. Wheat Yellow Rust Detection Using UAV-Based Hyperspectral Technology. *Remote Sens.* **2021**, *13*, 123. [[CrossRef](#)]
73. Lei, S.; Luo, J.; Tao, X.; Qiu, Z. Remote Sensing Detecting of Yellow Leaf Disease of Arecanut Based on UAV Multisource Sensors. *Remote Sens.* **2021**, *13*, 4562. [[CrossRef](#)]