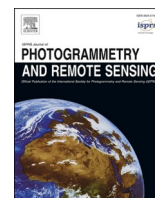




Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

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

Michael Gomez Selvaraj^{a,*}, Alejandro Vergara^a, Frank Montenegro^a, Henry Alonso Ruiz^b, Nancy Safari^c, Dries Raymaekers^d, Walter Ocimati^e, Jules Ntamwira^c, Laurent Tits^d, Aman Bonaventure Omondi^f, Guy Blomme^g

^a International Center for Tropical Agriculture (CIAT), A.A. 6713, Cali, Colombia

^b Department of Soil and Crop Sciences, Texas A&M University, College Station 77843, TX, USA

^c Bioversity International, Bukavu, South Kivu, Congo

^d VITO Remote Sensing, Mol, Belgium

^e Bioversity International, P.O. Box 24384, Kampala, Uganda

^f Bioversity International, c/o IITA – Abomey Calavi, 08 BP 0932 Cotonou, Benin

^g Bioversity International, c/o ILRI, P.O. Box 5689, Addis Ababa, Ethiopia

ARTICLE INFO

Keywords:

Artificial Intelligence
Banana detection
Deep learning
Disease detection
High-resolution satellite image
Surveillance
UAV images

ABSTRACT

Front-line remote sensing tools, coupled with machine learning (ML), have a significant role in crop monitoring and disease surveillance. Crop type classification and a disease early warning system are some of these remote sensing applications that provide precise, timely, and cost-effective information at different spatial, temporal, and spectral resolutions. To our knowledge, most disease surveillance systems focus on a single-sensor based solutions and lagging the integration of multiple information sources. Moreover, monitoring larger landscapes using unmanned aerial vehicles (UAV) are challenging, and, therefore combining high resolution satellite imagery data with advanced machine learning (ML) models through the use of mobile apps could help detect and classify banana plants and provide more information on its overall health status. In this study, we classified banana under mixed-complex African landscapes through pixel-based classifications and ML models derived from multi-level satellite images (Sentinel 2, PlanetScope and WorldView-2) and UAV (MicaSense RedEdge) platforms. Our pixel-based classification from random forest (RF) model using combined features of vegetation indices (VIs) and principal component analysis (PCA) showed up to 97% overall accuracy (OA) with less than

Abbreviations: AI, Artificial Intelligence; ARVI, Atmospherically Resistant Vegetation Index; BBTB, Banana Bunchy Top Virus; BXW, Xanthomonas Wilt of Banana; CAM, Class Activation Map; CE, Commission Error; CIAT, International Center for Tropical Agriculture; CIED, Chlorophyll Index Red Edge; CIG, Chlorophyll Index Green; CIG-A, Chlorophyll Index Green With NIR2; CGIAR, Consultative Group for International Agricultural Research; CM, Confusion Matrix; CNN, Convolutional Neural Network; DL, Deep Learning; ECA, East and Central Africa; EVI, Enhanced Vegetation Index; Fps, False Positives; GEE, Google Earth Engine; GCP, Ground Control Points; GPU, Graphics Processing Unit; Iou, Intersection Over Union; IPVI, Infrared Percentage Vegetation Index; Kappa, Kappa coefficient; Lics, Low-Income Countries; ML, Machine Learning; MLC, Maximum Likelihood Classification; MS, Multispectral; NIR, Near Infrared; NDVI, Normalized Difference Vegetation Index; OA, Overall Accuracy; OE, Omission Error; OSAVI, Optimized Soil-Adjusted Vegetation Index; P4P, Phantom 4 Pro; PA, Producer Accuracy; PCA, Principal Component Analysis; PCR, Polymerase Chain Reaction; PS, PlanetScope; RGB, Red, Green and Blue; RTB, Root Tubers Banana; RVI-B, Ratio Vegetation Index Blue; RVI-BA, Ratio Vegetation Index Blue with NIR2; RVI-Bre1, Ratio Vegetation Index Blue with Rededge1; RVI-Bre2, Ratio Vegetation Index Blue with Rededge2; RVI-Bre3, Ratio Vegetation Index Blue with Rededge3; RVI-G, Ratio Vegetation Index Green; RVI-GA, Ratio Vegetation Index Green with NIR2; RVI-Gre1, Ratio Vegetation Index Green with Rededge1; RVI-Gre2, Ratio Vegetation Index Green with Rededge2; RVI-Gre3, Ratio Vegetation Index Green with Rededge3; RVI-R, Ratio Vegetation Index Red; RVI-RA, Ratio Vegetation Index Red with NIR2; RVI-Rre1, Ratio Vegetation Index Red with Rededge1; RVI-Rre2, Ratio Vegetation Index Red with Rededge2; RVI-Rre3, Ratio Vegetation Index Red with Rededge3; RF, Random Forest; S2, Sentinel 2; SR, Simple Ratio; SVM, Support Vector Machine; TGI, Triangular Greenness Index; Tps, True Positives; UA, User Accuracy; UAV, Unmanned Aerial Vehicle; VIs, Vegetation Indices; VGG, Visual Geometry Group; WV-2, WorldView-2.

* Corresponding author.

E-mail addresses: m.selvaraj@cgiar.org (M. Gomez Selvaraj), j.vergara@cgiar.org (A. Vergara), f.montenegro@cgiar.org (F. Montenegro), henry.ruiz@tamu.edu (H. Alonso Ruiz), dries.raymaekers@vito.be (D. Raymaekers), w.ocimati@cgiar.org (W. Ocimati), laurent.tits@vito.be (L. Tits), b.a.omondi@cgiar.org (A.B. Omondi), g.blomme@cgiar.org (G. Blomme).

<https://doi.org/10.1016/j.isprsjprs.2020.08.025>

Received 30 April 2020; Received in revised form 20 August 2020; Accepted 31 August 2020

Available online 18 September 2020

0924-2716/© 2020 The Author(s). Published by Elsevier B.V. on behalf of International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). This is an

open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

10% omission and commission errors (OE and CE) and Kappa coefficient of 0.96 in high resolution multispectral images. We used UAV-RGB aerial images from DR Congo and Republic of Benin fields to develop a mixed-model system combining object detection model (RetinaNet) and a custom classifier for simultaneous banana localization and disease classification. Their accuracies were tested using different performance metrics. Our UAV-RGB mixed-model revealed that the developed object detection and classification model successfully classified healthy and diseased plants with 99.4%, 92.8%, 93.3% and 90.8% accuracy for the four classes: banana bunchy top disease (BBTD), Xanthomonas Wilt of Banana (BXW), healthy banana cluster and individual banana plants, respectively. These approaches of aerial image-based ML models have high potential to provide a decision support system for major banana diseases in Africa.

1. Introduction

Bananas and plantains serve as a primary food source for around 20 million people in East Africa and 70 million people in West and Central Africa, and given that Sub-Saharan Africa is highly dependent on banana/plantain cultivation for food, income and job security (De Buck and Swennen, 2016). In East and Central Africa (ECA), bananas are mainly grown on small-scale farms with a size ranging from 0.5 to 4 ha (FAOSTAT, 2017). Biotic stresses (bacterial, viral and fungal diseases, and pests) continue to affect the banana production systems in this region, causing partial damage or even the destruction of entire fields (De Buck and Swennen, 2016). The arrival and spread of major devastating banana diseases such as bunchy top disease (BBTD) and Xanthomonas wilt of banana (BXW) pose a considerable threat to food security in the ECA region (Blomme et al., 2017; Niyongere et al., 2012; Niyongere et al., 2015; Ocimati et al., 2019).

Disease surveillance for BXW and BBTD carried out through field visits, or human scouting are often complemented by diagnostic tools such as using growth media and polymerase chain reaction (PCR) (Blomme et al., 2014; Bouwmeester et al., 2016; Ocimati et al., 2019). The inspection of early disease detection is assessed as prevalence (present or not) at plot/field, farm, village, or landscape level, and the procedure by itself at a large scale is challenging and time-consuming (Johansen et al., 2014). These limitations have led scientists to investigate on advanced and novel techniques that could obtain the crop health information rapidly and economically (Heim et al., 2019; Johansen et al., 2014; Steward et al., 2019).

UAVs and satellites with the capability of capturing a large number of high-quality spectral-temporal aerial images are becoming an ultimate technology for assessing yield, health, and economic valuation of the crop (Ji et al., 2018). Conventional classification methods such as support vector machine (SVM), K-nearest neighbor (KNN), maximum likelihood classification (MLC) are successfully applied for crop type classification (Blanzieri and Melgani, 2008; Murthy et al., 2003). However, these methods are slower and involve supervised labeling to get precise data (Ji et al., 2018). Therefore, it is essential to develop crop classification technologies that are pragmatic and can be widely applied in extensive operational settings. Crop classification by remote sensing is an intricate task, especially if different crops have a similar spectral response and growth pattern. In such cases, classification tasks could be enhanced by combining object-based image analysis and advanced machine learning methods (Peña et al., 2014). In the last few years, high resolution UAV images coupled with machine learning (ML) classification techniques are emerging for crop classification (Steward et al., 2019). Recent research (Pourazar et al., 2019; Steward et al., 2019) also demonstrated the potential of deep learning-based instance segmentation approach on field-based disease classification using UAVs, but the efforts towards banana remain limited. Neupane et al. (2019) reported deep learning-based banana classification and height estimation in well-planned commercial fields using UAV-RGB images. However, these models may be more complicated in an African scenario, where mixed-complex agriculture systems bound the area with trees, intercrops, and bananas. Therefore, an automatic classification method to locate or detect banana plants/mats in the mixed-complex landscapes is in urgent

need. Since 2010, UAV-based imaging sensors have been used for providing high operational flexibility in crop monitoring (Sankaran et al. 2015; Shi et al. 2016) but comes with their restrictions in terms of limited aerial coverage, operational skills, country regulations, and permissions, especially in developing countries. Another important advancement in crop monitoring is the availability of high-resolution images from satellites with sub-meter accuracy. With the high spatial and temporal resolution of satellite imagery, multispectral (MS) data can be used for crop classification similar to UAV images with the added advantage of larger area coverage.

Remote sensing permits the non-invasive measurement of biophysical and biochemical parameters of crops, and thus for the nondestructive monitoring of plant health status (Lu et al., 2015; Mahlein et al., 2010; Ramcharan et al., 2017). The rapid technology development in UAV systems and the initiation of low-cost UAVs carrying visible and multi-spectral sensors, provide the opening to capture high spatial and spectral resolution data, especially for disease detection for production fields. The ability to capture the crop phenotypic differences in this complex multi-dimensional system is necessary to understand the host-pathogen interactions better and develop disease-resistant varieties (Steward et al., 2019). In wheat, the mobile and real-time plant disease (MARPLE) integrated diagnostics system already proved to be an early warning system in Ethiopia to directly inform diseases risk forecasting (Carvajal-Yepes et al., 2019). However, the benefits of integrated remote sensing-based diagnostic technologies have not been assessed in low-income countries, where emerging diseases in banana can be particularly devastating (e.g. *Fusarium* TR4 outbreak in Colombia; BXW in eastern DR Congo).

The combination of aerial image information and AI-based approaches have the potential to provide an accurate, high-throughput method for plant disease detection under real-world conditions (Boulent et al., 2019). More recently, Selvaraj et al., (2019) demonstrated the deep learning potential on banana disease detection, which led to the release of a smartphone-based AI app (Tumaini), which could detect six major banana diseases (<https://play.google.com/store/apps/details?id=ciat.cgiar.org.tumaini&hl=en>). The use of these AI apps for disease surveillance on a system-level approach to integrate multi-level sensing systems (satellites, UAVs AI apps, and ground-truthing) is necessary to monitor crop health at different scales. Continuous efforts to leverage advanced technologies for crop disease surveillance and management in low-income countries (LICs) must occur to effectively limit the impact of crop diseases spreading locally, and globally (Carvajal-Yepes et al., 2019). In this framework, the paper aims to investigate the potential of aerial imaging and ML techniques on banana classification and disease detection in the African mixed-landscape. To this aim, we used MS, UAVs (MicaSense RedEdge) and multi-level satellite images (Sentinel 2, PlanetScope and WorldView-2) (1) to classify and localize banana under mixed-complex African landscapes through pixel-based classifications and ML models. (2) Also, developed a object based deep learning mixed models for simultaneous banana detection and classification of two of their major diseases (BXW and BBTD), using UAV-RGB images.

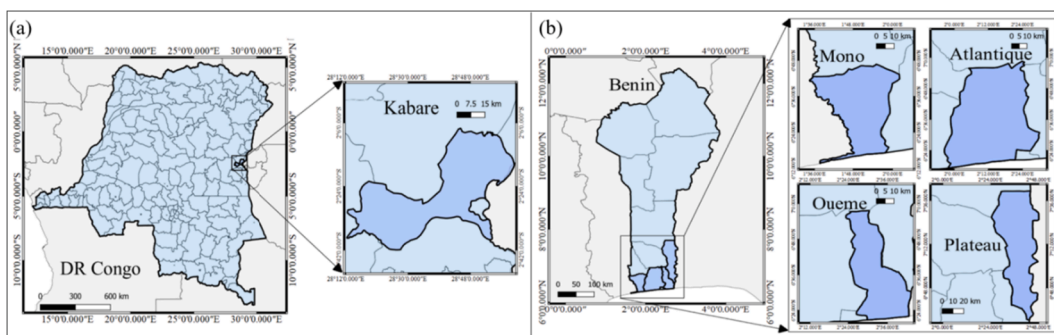


Fig. 1. Location of the study regions a) Kabare district of DR Congo. b) Oueme, Plateau, Mono, and Atlantique departments of Republic of Benin.

2. Materials and methods

2.1. Study area, image data collection, and ground-truthing

Typical banana growing regions were selected from Kabare district, South Kivu province of the Eastern Democratic Republic of Congo (DR Congo), and the Oueme, Plateau, Mono, and Atlantique departments of the southern Republic of Benin were selected as specific study regions (Fig. 1). The Eastern DR Congolese highlands (1,400–1,800 m asl) are dominated by the East African highland banana (*Musa* AAA-EAH genome) types. In contrast, plantain (*Musa* AAB genome) is the most widely grown *Musa* type in the Southern Republic of Benin. Both the regions are characterized by small-scale subsistence type of agriculture with mixed-complex landscapes where banana plants are mixed with shrubs, trees, annual food crops, pastures, buildings, roads, etc. Kabare district, DR Congo, is characterized by a mean annual rainfall of 1656 ± 235 mm (2015–2018) distributed over two rainy seasons (February to May and September to December) and a mean annual temperature of 18°C . The study area in southern Republic of Benin is a coastal lowland area with an average altitude of 11 m in the coastal Atlantique to 250 m

in the Oueme region. The mean annual temperature in the southern Republic of Benin is 28°C , while annual rainfall ranges from 1,000 to 1,300 mm.

In DR Congo, UAV images of BXW infected fields were collected from 15 different locations of the Kabare district, eastern DR Congo during the years of 2017, 2018, and 2019 (Supplementary Table 1). BXW-affected landscapes were selected based on the presence of apparent BXW symptoms in banana fields. Besides, pixel-based banana classification datasets of multispectral UAV and satellite images were collected from three different locations in the Kabare district of DR Congo (Supplementary Table 1). UAV-RGB images of BBTB infected banana fields were collected in Republic of Benin (Oueme Plateau, Mono, and Atlantique departments), West Africa, in 2019. The detailed location descriptions of two experimental sites are shown in Supplementary Table 1.

Ground truth data of disease information was obtained during field surveys where experts walk through banana fields and assess the presence and incidence of the diseases. The specific size for each sample point covered one banana plant, and these samples are classified into two categories: healthy and diseased plants (BXW and BBTB), as

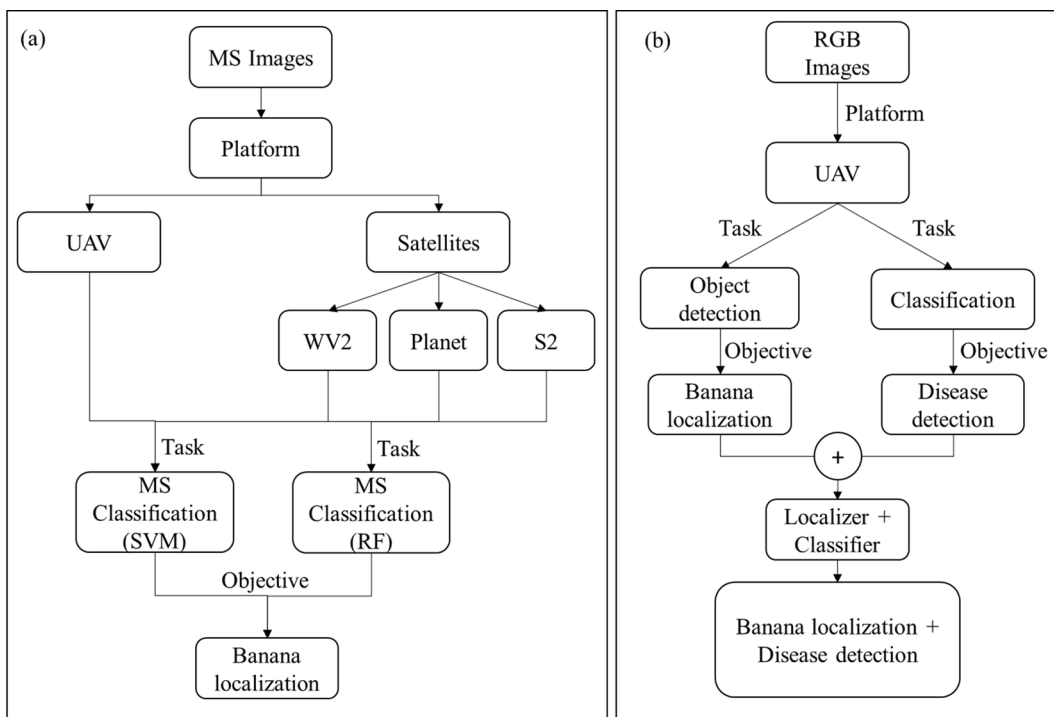


Fig. 2. Banana and their major disease monitoring system per sensor source: (a) Pixel-based banana classification system by MS images and (b) Object-based banana localization (classifying/locating banana under mixed farming system) and disease detection (detecting plant features by deep learning to classify diseased plants from the healthy) by RGB images.

Table 1
Overview of the imaging systems used in this study.

Platform name	Platform type	Sensor type	Resolution	Task	Disease classification (C)	Object detection (Od)	Od + C	Banana classification
3DR SOLO	UAV	Sony-QX1	3.35 cm/px @120 m	VGG-16, Custom model	–	–	–	–
Phantom 4 Pro	UAV	DJI-FC6310	3.54 cm/px @120 m	VGG-16, Custom model	RetinaNet	RetinaNet + Custom model	–	–
Phantom 4 Pro	UAV	MicaSense-RedEdge	8 cm/px @120 m	–	–	–	–	RF, SVM
S2	Satellite	12 bands	10 m/px	–	–	–	–	RF, SVM
Planet	Satellite	4 bands	3 m/px	–	–	–	–	RF, SVM
WV-2	Satellite	8 bands	31 cm/px	–	–	–	–	RF, SVM

reflected by the typical external characteristics and the symptoms expressed by the infected plants. Based on these ground truth information, the UAV-based aerial images were annotated by plant pathologists on the presence of BXW and BBTD (Supplementary Figure 1). The symptoms used to pinpoint a plant affected by BXW are leaf yellowing and wilting, whereas a bunchy top/rosette appearance was the typical symptom of BBTD-affected fields. Using large sets of annotated images, the datasets were prepared to localize and detect banana and their major diseases using machine learning models. The sensors and ML system used in this study are described below.

2.1.1. Sensors and ML system description

Our ML detection system (banana and their major diseases) is based on multispectral and RGB images captured from satellites and UAVs, as presented in Fig. 2. In this study, we used multispectral (MS) satellite and UAV images for pixel-based banana classification in the field landscape (Fig. 2a) and high-resolution UAV-RGB images for both object-based banana localization and disease detection (BXW and BBTD) purposes (Fig. 2b). An overview of the sensor and ML-based banana classification and disease detection system are shown in Table 1. The software and hardware used in this study are presented in Supplementary Table 2.

2.2. Pixel-based banana classification

UAV-MS images were taken (Fig. 2a) with a MicaSense RedEdge multispectral camera attached to a multicopter UAV Phantom 4 Pro (P4P) with automatic flight plan using Pix4Dcapture app (<https://www.pix4d.com/es/producto/pix4dcapture>). In order to collect the aerial multispectral imaging data, the drone flew autonomously along a preset flight path at a speed of 2.5 m/s and a height of 100 m above ground level, achieving ~8.0 cm ground sample distance. The Micasense RedEdge can acquire a 16-bit raw image in five narrow bands from blue, green, red, red edge to near-infrared (NIR), and more details are listed in the Supplementary Table 3.

MS orthomosaics (converting large numbers of individual images into a single geo-referenced mosaicked image) for each flight were constructed using Agisoft Metashape software (Agisoft, 2020) with the protocol described in (Selvaraj et al., 2020). We ended up with three high-quality MS orthomosaics from DR Congo (see Supplementary Table 1) where drone and satellite image dates coincide. We used three different satellite information that acquires low to high spatial resolution reflectance images (geometrically and radiometrically corrected products): WorldView2 (WV-2), (Digital Globe, 2016); PlanetScope (PS), (Planet, 2017); Sentinel 2(S2),(ESA, 2015) (Table 1). The descriptions of collected locations, spectral and spatial characteristics of each satellite are listed in Supplementary Tables 1 and 3.

2.2.1. Data extraction and feature selection

To extract UAV-MS features, we annotated all banana plants (individual and clusters, i.e., mats) located in three MicaSense-derived high-

Table 2
Ground data sources for MS images (UAV and Satellites).

Features	MicaSense*	WV-2**	PS**	S2**
Banana	6430	2897	410	100
Pastures	6522	2510	133	39
Trees	8084	2944	292	62
Non banana	–	–	425	101
Total	21,036	8351	835	201

* UAV classes derived from three different orthomosaics.
** Satellite classes derived from three different types of sensors (WV-2, PS & S2), which coincides with UAV dates.

quality MS orthomosaics from the Kabare district (Supplementary Table 1). To extract reflectance data, we drew 30 polygons, and a total of 21,036 reference points for the three classes (banana, pastures, and trees) were obtained. The data description is shown in Table 2. Simple Ratio (SR) and non-normalized Vegetation Indices (VIs) (Supplementary Table 4) were extracted from the data points to compare the structural part of the plant (referred to NIR band) over the visible region of chlorophyll absorption (Chuvieco, 1991).

To extract satellite image features, the same technique of polygon feature extraction of banana plantations, as described in the above section, was used. Since the three satellite sensors (WV-2, PS, and S2) have different spatial resolutions, the chosen locations were not enough to extract the data points and train the models effectively. Therefore, we selected other banana plantations in the same UAV flown region and derived the same three classes (banana, pastures, and trees), however for PS and S2, we created a fourth class with non-banana features (i.e., the sum of pastures and trees), and this was created to test the performance of the model (Table 2). The amount of reference samples for each class type are shown in Table 2.

2.2.2. Banana classification models & accuracy metrics

We used the classification package from google earth engine (GEE) (Gorelick et al., 2017) that contains traditional ML algorithms for random forest (RF) (Breiman, 2001) and support vector machine (SVM) classifications (Hsu et al., 2003).

The workflow of the classifications is shown below in Fig. 3, and the steps are described as follows:

- 1) VIs calculation, masking of buildings, and soil: The thresholding mask using G-R, and VIs was created.
- 2) Principal component analysis (PCA) of bands and VIs: This technique is used to extract the best information from all the data and test if the model has higher precision, and avoid the autocorrelation of the features.
- 3) Training data collection and labeling: The reference data collected from all the sensors were alienated into 70% and 30% for training and testing, respectively. The feature collection function from GEE was used to store the features.

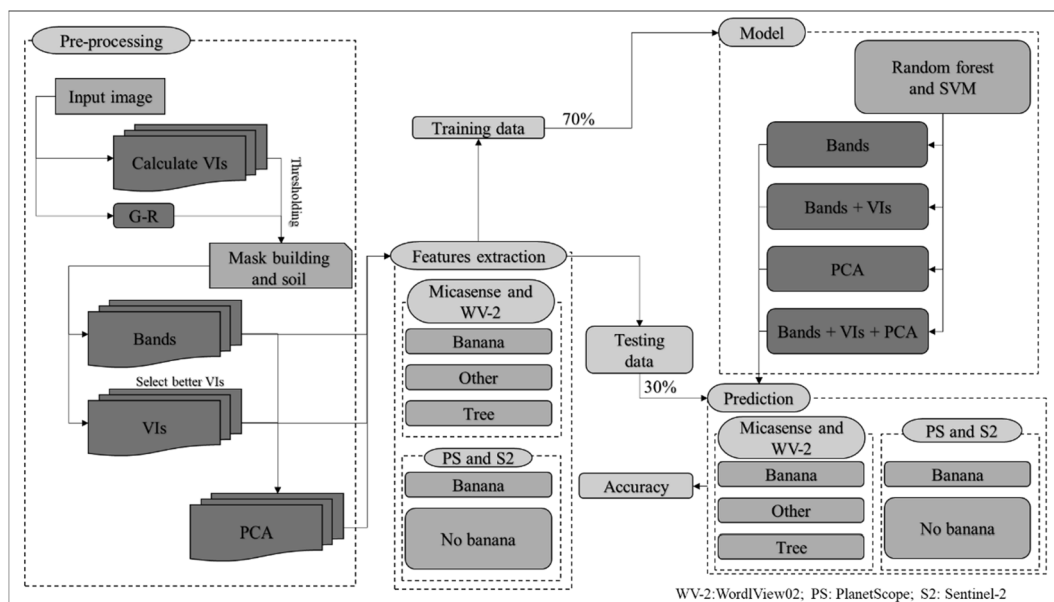


Fig. 3. The workflow of multispectral image-based banana classification. 'B' is Bands, B + VIs is Bands plus Vegetation indices, PCA is Principal component analysis, B + VIs + PCA is Bands plus Vegetation indices plus Principal component analysis.

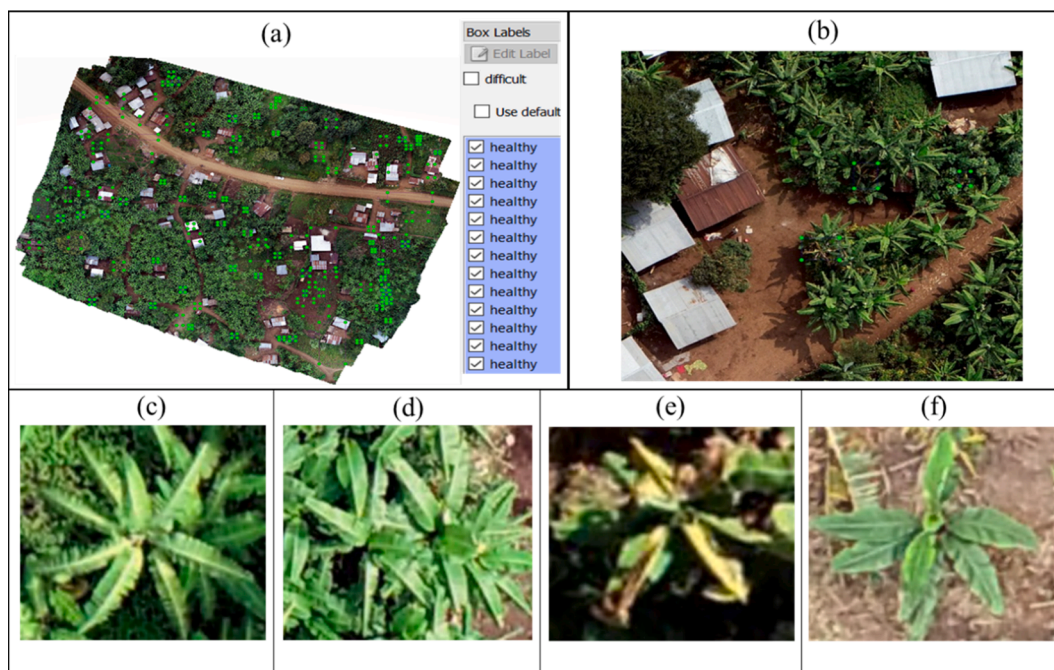


Fig. 4. Data labeling using Labelling software and banana disease detection classes. (a) Software overview (Orthomosaic and labeling) (b) Bounding boxes resulted from labeling process (c) healthy individual plant (d) healthy banana cluster (e) BXW infected plant (yellowing and wilting) (f) BBTB infected plants (rosette appearance).

- 4) Classifier parameter selection and initialization: RF models were trained with 53 trees, minLeafPopulation:3, bagFraction: 0.5. SVM model was trained with decision procedure: 'Voting,' svmType: 'C_SVC,' kernel type: 'LINEAR' and shrinking: TRUE as suggested by Oliphant et al., (2019) for initial values.
- 5) Classifier training: Each classification model was trained with a specific group of features
- 6) Image classification: Image classification for each model was obtained
- 7) Performance metrics: The confusion matrix (CM) is obtained by comparing the classified pixels, and the test data (Stehman, 1997).

Also, the following accuracy metrics were computed: Overall Accuracy (OA), Kappa coefficient (Kappa), per-class User's (UA) and Producer's (PA) Accuracy, as well as Commission Error (CE) and Omission Error (OE) (Campbell and Wynne, 2011; Lucas et al., 1994).

2.3. Object-based banana localization and disease detection by UAV-RGB images

2.3.1. UAV image collection, labeling, and splitting

Initially, UAV-RGB images of BXW and BBTB were taken using a 3DR

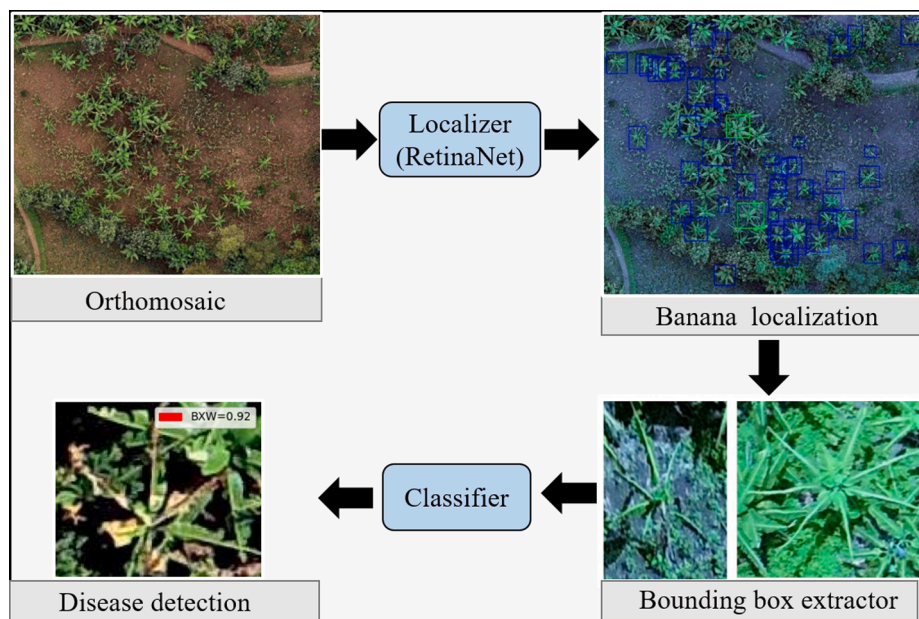


Fig. 5. The workflow of UAV-RGB based mixed model pipeline for instantaneous banana detection and disease classification.

SOLO multicopter UAV with a Sony QX1 RGB camera and subsequently using a DJI Phantom 4 Pro (P4P) UAV with an integrated RGB camera. The flight paths of 3DR Solo UAV and P4P was programmed using Tower app software (Tower, 2017) and Pix4Dcapture software (Pix4D, 2018). To check the sensitivity of UAV captured images at spatial resolution, the flying heights were set at the altitudes of 50 and 100 m above the soil level. The location details are listed in Supplementary Table 1. The undulating terrain and tall trees (e.g., Ficus and Eucalyptus trees) made the UAVs challenging to fly at a low altitude; however, these flying heights were translated to a spatial resolution of 1 and 3 cm based on UAVs altitudes. These UAV-RGB images were also used to develop a DL-based banana detection system (Fig. 2 a, b).

To detect all possible banana plants in the field (either healthy or diseased), we trained the object detection model to represent only one general class called banana. The annotation labels were drawn on the RGB images by plant pathologists using the Labellmg software (Tzutalin, 2015) (Fig. 4 a, b) in the PASCAL VOC format implemented in the ImageNet database (Deng et al., 2009) ending up with 6195 annotations from the RGB orthomosaics. To avoid the colossal processing and memory usage, orthomosaics having more than 10,000 pixels were split into several tiles, based on the original image size, with an overlap of 50 pixels between them. It was considered ideal for reconstructing the orthomosaics after prediction, therefore not to lose samples in the splitting process.

To classify the banana diseases, we used four classes, healthy individual plant, healthy banana cluster, BXW infected, and BBTB infected plants (Fig. 4 c, d, e, f). Using the Labellmg software (Tzutalin, 2015), the total number of 2753 annotations (filtered from the 6195 initial annotations) in the real field landscapes were categorized as 599, 705, 922, and 583 for healthy plants, healthy banana clusters, BXW infected plants and BBTB infected plants, respectively (Supplementary Fig. 1). The healthy individual plant was annotated at different ages, especially that are similar to the age of BBTB infected plants. This is performed to differentiate the BBTB infected plants from healthy young plants, and all the annotated images were confirmed by the plant pathologists. A python script was used to extract the images inside the annotation for each class and create a new dataset with desired classes based on our region of interest.

2.3.2. DL models description

As mentioned earlier, a disease classification model was placed on

top of an object (banana) detection algorithm to classify healthy and diseases classes (Fig. 4). We used two architectures to evaluate the disease classification model; namely, a Visual Geometry Group 16 (VGG-16) pre-trained model and a custom model (Fig. 5), which are mentioned below.

For banana localization, we used RetinaNet (Lin et al., 2017) with a backbone of ResNet50 developed by Fizr (Fizyr, 2019) to localize and detect banana plants. To train the model, we implemented a transfer learning approach with the pre-trained COCO (Common objects in context) data set (Lin et al., 2017), which is available in Tensor Flow object detection API model Zoo (Huang et al., 2017). Finally, we used a mixed model (localizer + classifier) to detect plant status (healthy or diseased) (Fig. 5).

For diseases (BXW and BBTB) recognition, we used two architectures to evaluate the disease classification model; namely, a visual geometry group 16 (VGG-16) pre-trained model and a custom model briefly mentioned below.

VGG-16: VGG-16 was developed by the visual geometry group (VGG) of Oxford University. Due to its decent generalization performance, VGG-16 can increase the classification accuracy by using its pre-trained model on the ImageNet dataset (Deng et al., 2009). To adapt VGG-based architecture used in this study, we altered the model adding a dropout layer of 2% before the soft-max to reduce the overfitting.

Custom Model: As observed in Supplementary Fig. 2, our proposed architecture starts with a fixed size input layer of established image dimensions (64x64x3), followed by four feature extraction blocks with two convolutional layers and one max-pooling layer, which extracts the image features. Then, to keep only the relevant features by class and reduce overfitting issues, a dropout layer of 0.2 was installed. When the image passes through the model, the resulting feature map is transferred to the next fully connected layer, where the inference is made.

2.3.3. Data set training and evaluation metrics for object detection models Training

We trained RetinaNet with RetinaNet50 backbone to develop an object detection (banana) model. The minimum object size allowed is 32 pixels. We used the original architecture without any modifications and divided the dataset into 90% and 10% for training (6287 annotations) and testing (726 annotations), respectively. This proportion was mainly selected to have as many training samples as possible to improve the accuracy. Moreover, this chosen proportion will not affect the model

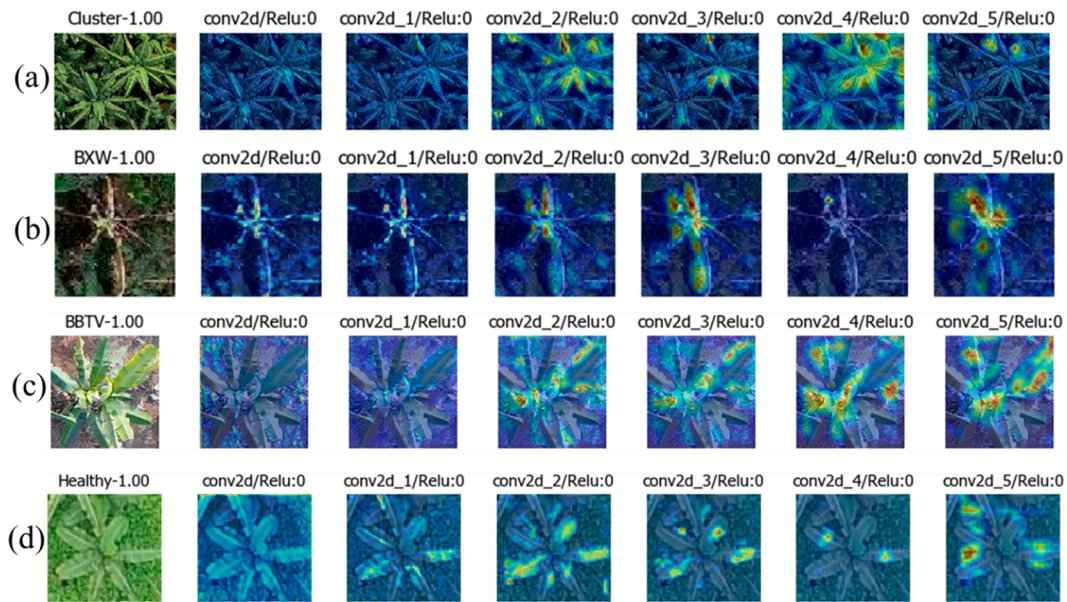


Fig. 6. Diseased and healthy class feature detection by CAM, the highlighted regions indicate the region of interest where the model is detecting special features for prediction. (a) Banana cluster, (b) BXW infection in the leaves (c) BBTVD symptom (d) Healthy individual plant.

behavior since the single class model and cross-validation were not implemented based on the time consumption to train this algorithm. To train the RetinaNet model, we used 50 epochs and a batch size of 8 and 785 steps to cover the entire dataset. Then, the classification model that recognizes the diseases on the bounding boxes produced by the VGG-16 (model 1) was re-trained by 100 epochs, using the 'categorical_crossentropy,' and 'Adam' such as loss function and optimizer. On the other hand, the custom model (Model 2) was trained using the same number of epochs, and parameters as that of the VGG model. Still, for the optimization function, the Stochastic gradient descent with a learning

rate of 0.001, and decay of 0.02 was employed, Fig. 5 shows the details of the custom architecture. The models were fitted using 70% of training (2312) images and 30% of testing (988) images.

Class activation maps (CAM) were generated to understand the predictions of the classification model. As various visualization techniques, CAM creates heat maps to highlight class /specific regions of images such as, (BXW, BBTVD, healthy banana). The following figures exhibit the areas highlighted by the model along the internal layers to infer if an image corresponds to a healthy or a diseased plant. The red color indicates strong influence, while the blue indicates weak

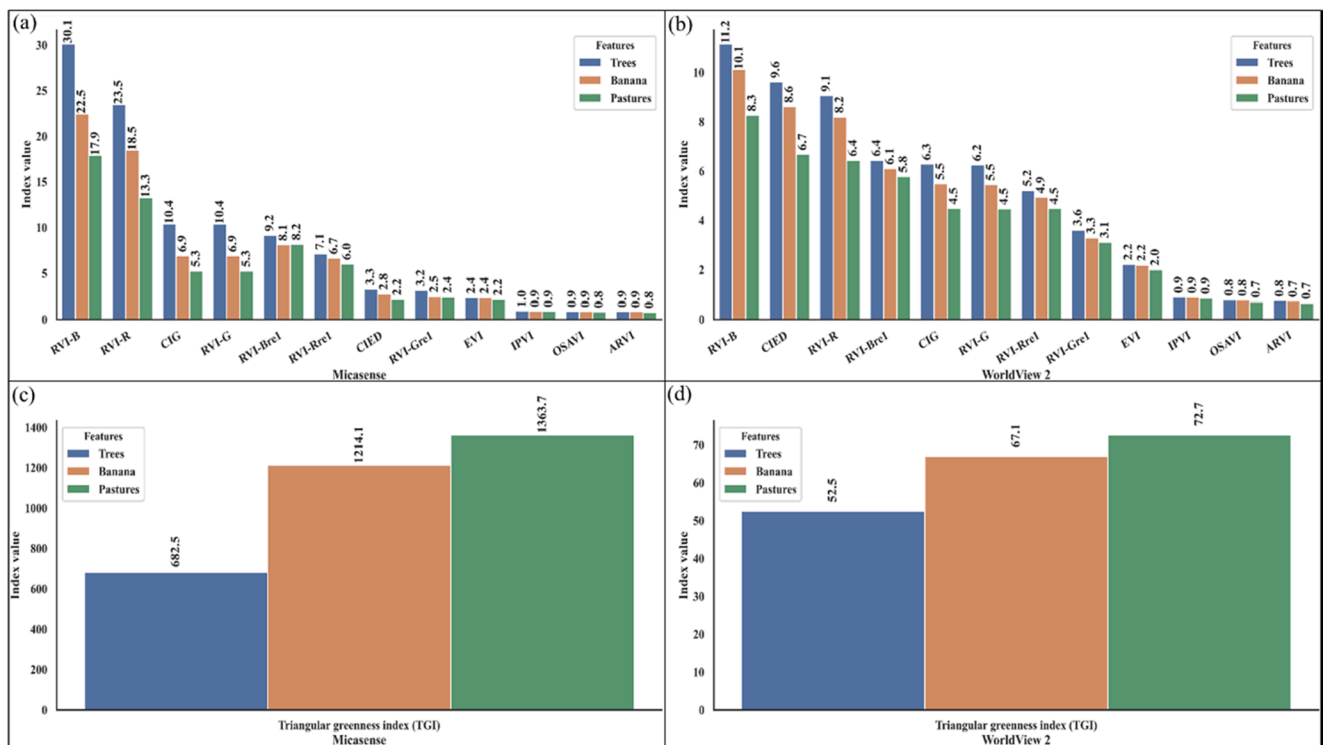


Fig. 7. Spectral signature comparison of high-resolution multispectral images (a) Vegetation index MicaSense; (b) Vegetation index WV-2; (c) TGI in MicaSense and (d) TGI in WV-2.

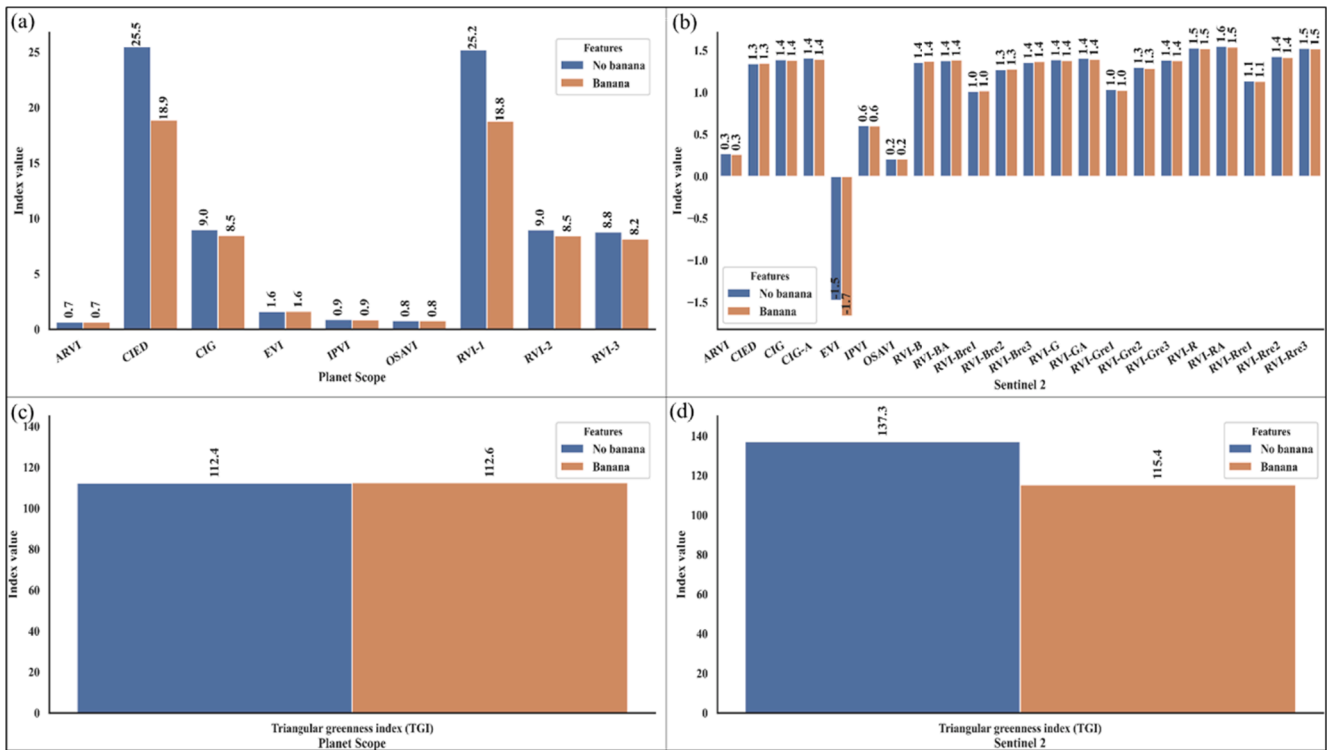


Fig. 8. Spectral signature comparison for medium and low-resolution multispectral images (a) Vegetation index of PS; (b) Vegetation index of S2; (c) TGI in PS and (d) TGI in S2.

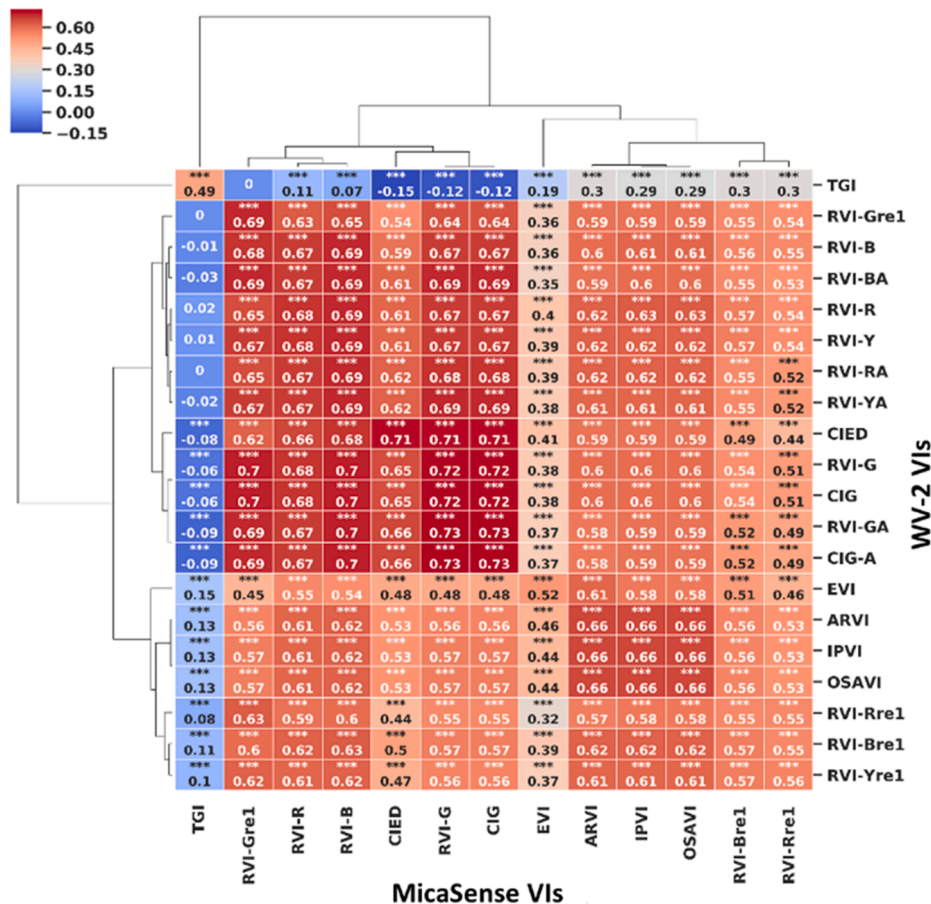


Fig. 9. Correlation between MicaSense and WV-2 VI metrics.

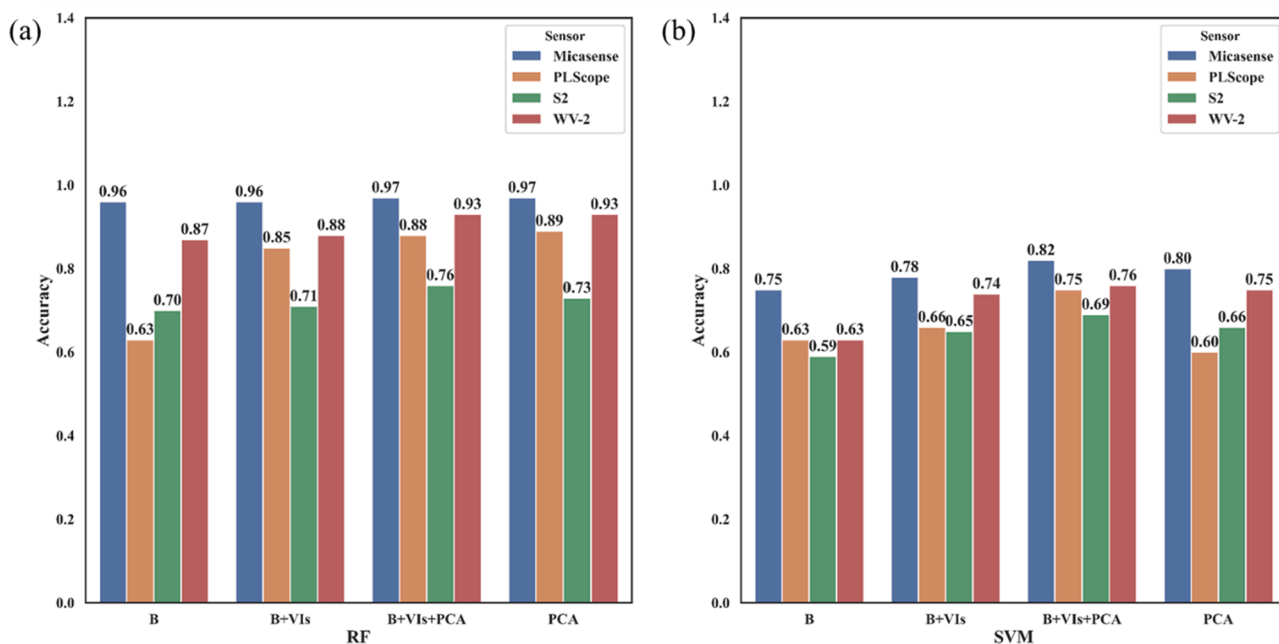


Fig. 10. (a) Overall accuracy of the random forest; (b) Overall accuracy of SVM where 'B' is Bands, B + VIs is bands plus vegetation indices, PCA is principal component analysis, B + VIs + PCA is bands plus vegetation indices plus principal component analysis.

importance on features (Fig. 6).

2.3.4. Evaluation metrics

In this paper, we applied different metrics such as loss function, IoU, precision and recall to evaluate the performance of the banana localization. The loss function provides a statistical representation of the

training process and diagnoses a good-fit of the model. Loss function data was extracted using (Mané, 2015) the TensorFlow’s visualization toolkit (TensorBoard), which helps to track and visualize metrics such as loss and accuracy in ML process. Once we train the model, the Intersection over Union (IoU) was implemented to extract true positives, if B_1 is the ground truth bounding box and B_2 the predicted bounding box,

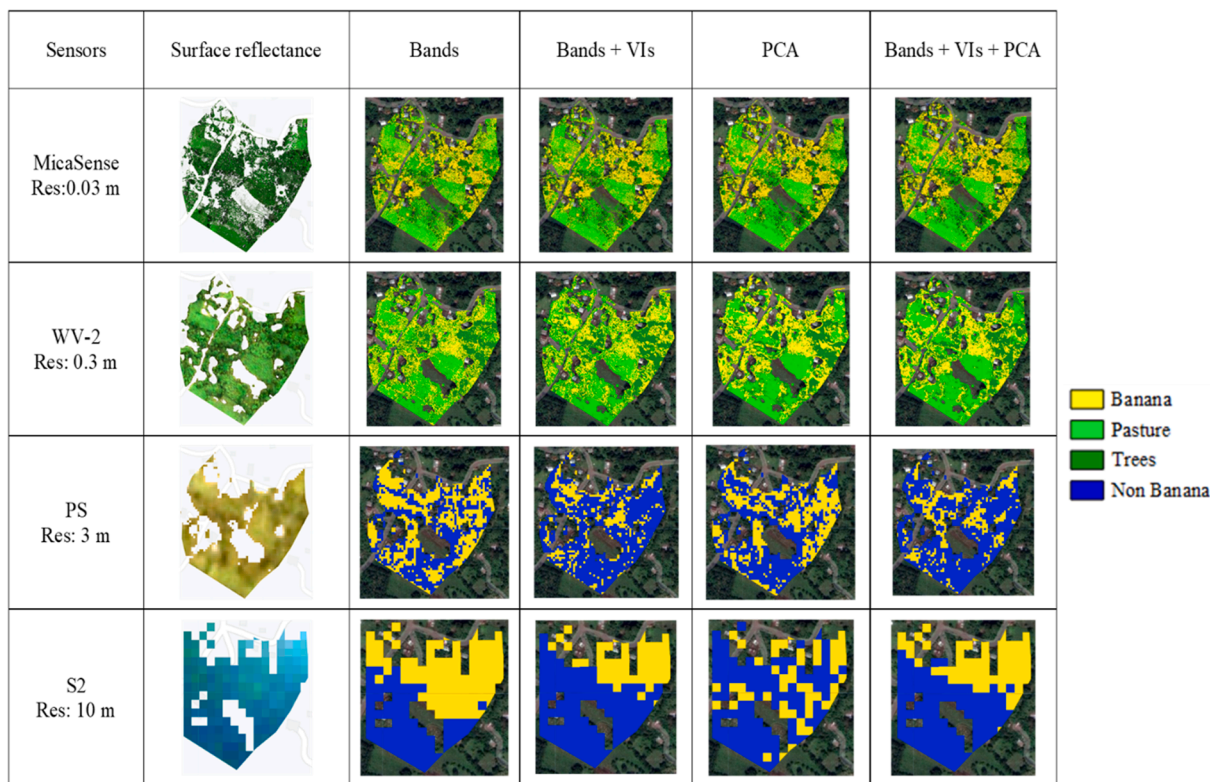


Fig. 11. Classified satellite and UAV-MS images on banana classification map using RF model with different features of the studied banana region in the Muhanh, Kabare district of eastern DR Congo, 22 March 2019 where UAV and satellite image dates coincides. Yellow segmentation indicates banana localization in different levels of resolution from low to high.

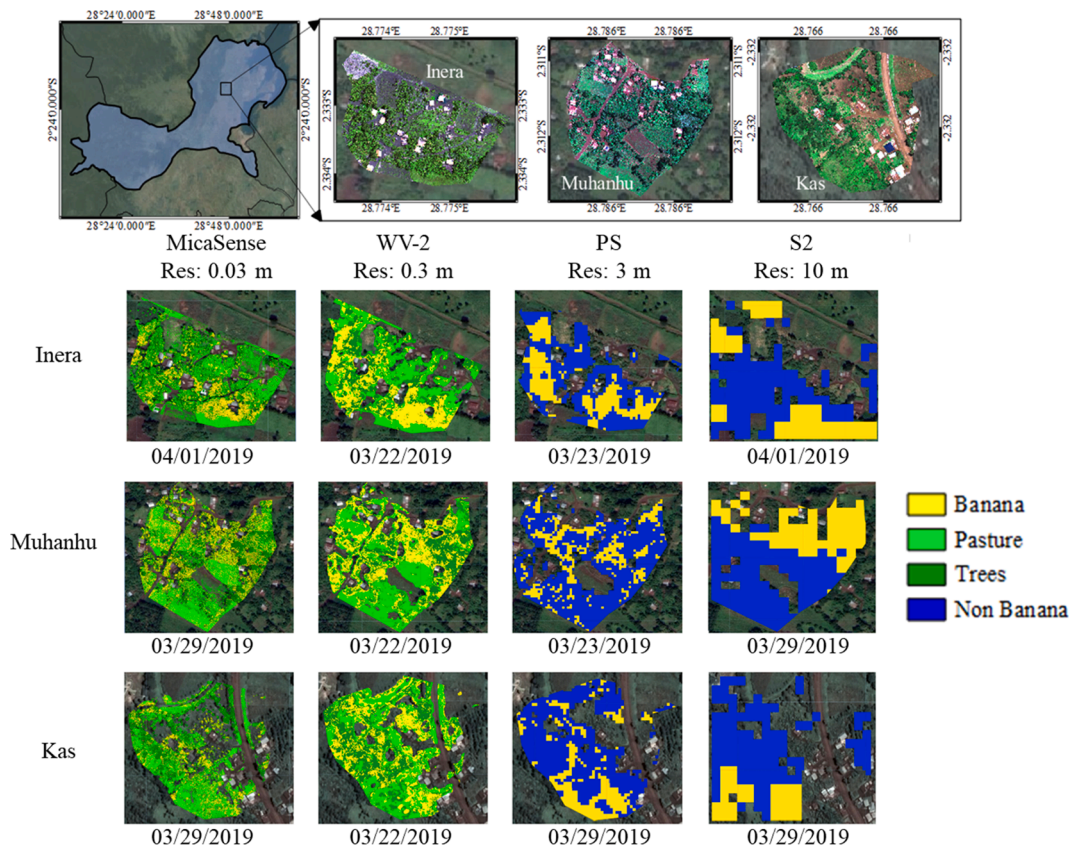


Fig. 12. Visual comparison of RF model (B + VIs + PCA) based banana classification map results for different sensor resolutions in Inera, Muhanhu, and Kas from Kabare district of eastern DR Congo.

IoU is calculated as in the equation (1). In this case, to determine if a prediction is a real positive (TP), we used a value of 0.3 as an acceptable box overlap, so if IoU greater than 0.3, then the current prediction is a TP.

$$IoU(B1, B2) = \frac{B_1 \cap B_2}{B_1 \cup B_2} \quad (1)$$

For disease classification, precision, recall, and F1-score metrics were computed to evaluate the performance of the classification model. Recall, precision, and F1-scores were computed using the same formula reported by Mao et al. (2017). Additionally, we also computed the confusion matrix to evaluate the accuracy per class.

3. Results

3.1. Pixel-based banana classification using MS images

3.1.1. MS signature comparison and feature selection

The normalized VIs are sensitive to environmental factors and causes a high level of saturation (Yang et al., 2008) (Supplementary Fig. 3) and distortion (Azar et al., 2016). To address these limitations, several other derivatives and alternatives to NDVI have been reported in the scientific literature (Yang et al., 2008). We exploited non-normalized VIs for our banana classification task, which is listed in Supplementary Table 4.

A variety of bands and VIs are produced from each sensor (Supplementary Table 4), and the different vegetation classes are shown in Figs. 7 and 8. It was observed that CIG, RVI-B, RVI-G, RVI-R, TGI indices of MicaSense (Fig. 7 a), CIG, CIG-A, RVI-B, RVI-BA, RVI-G, RVI-GA, RVI-R, RVI-RA, TGI indices of WV-2 (Fig. 7 b), CIG, CIED, RVI-B, RVI-G, RVI-R and TGI indices of PS (Fig. 8), were showed the difference between the banana and non-banana classes (Fig. 8). However, in the case of S2, EVI and TGI are the only VIs that differentiate banana and non-banana

classes. In contrast, other VIs was not much useful to separate those classes (Fig. 8 d).

Fig. 9 clearly shows the correlation of VIs derived from MicaSense (UAV) and WV-2, where, a highly significant correlation was achieved in VIs of CIG and RVI-G ($r = 0.73$). Also, ARVI, IPVI, and OSAVI exhibited a high correlation of $r = 0.66$. Our correlation analysis between other sensors like PS vs. UAV, S2 vs. UAV combination is listed in Supplementary Figs. 4 and 5. As expected, the correlation between VIs derived from UAV and other satellite sensors (PS and S2) also exhibited higher positive and significant correlation (Supplementary Figs. 4 and 5), indicating the potential of those sensors on banana classification in mixed-complex African landscapes.

3.1.2. ML models for pixel-wise banana classification

MicaSense versus WV-2 based ML model :

To evaluate the banana classification performance of the ML models (RF and SVM), we used the confusion matrix. The overall results of RF classifier model with combined (Bands + VIs + PCA) (Fig. 11) features demonstrated a maximum OA of 97% and 93% with $\leq 9\%$ OE and CE, kappa of 0.96 and 0.90 for MicaSense and WV-2, respectively (Fig. 10; Supplementary Table 5). Likewise, the SVM model showed 82% and 76% with $< 28\%$ OE and CE, kappa of 0.73 and 0.64 for MicaSense and WV-2, respectively (Supplementary Table 6). In both ML models, combined (Bands + VIs + PCA) features were identified as best performers compared to other features (Fig. 10).

PS versus S2 based ML model :

The models (RF and SVM) trained with the three classes (banana, pasture, and tree) showed low accuracy in the PS, S2 sensors. The results of confusion matrix between the ground-truth and the corresponding classification results of ML model are shown in Supplementary Tables 5 and 6. However, the models trained with the banana classes and non-banana (using PS and S2) showed similar good results compared to

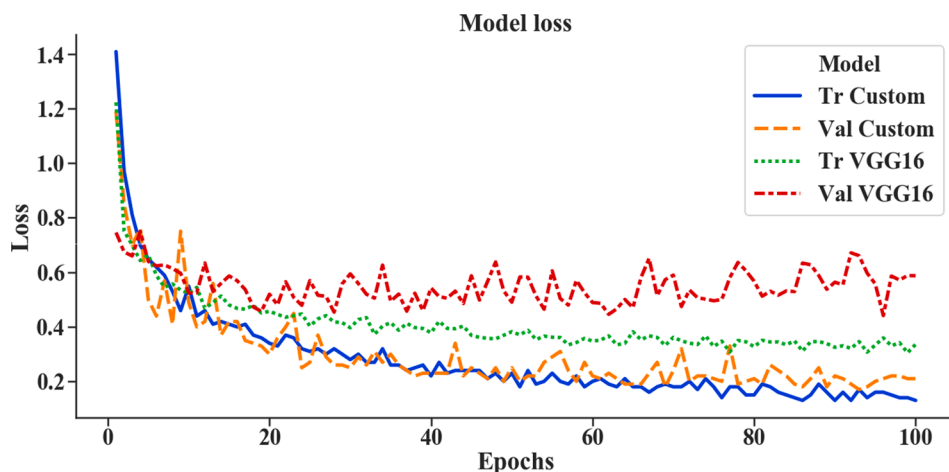


Fig. 13. Loss function of the disease classifier. Tr and Val indicates the training and validation data sets.

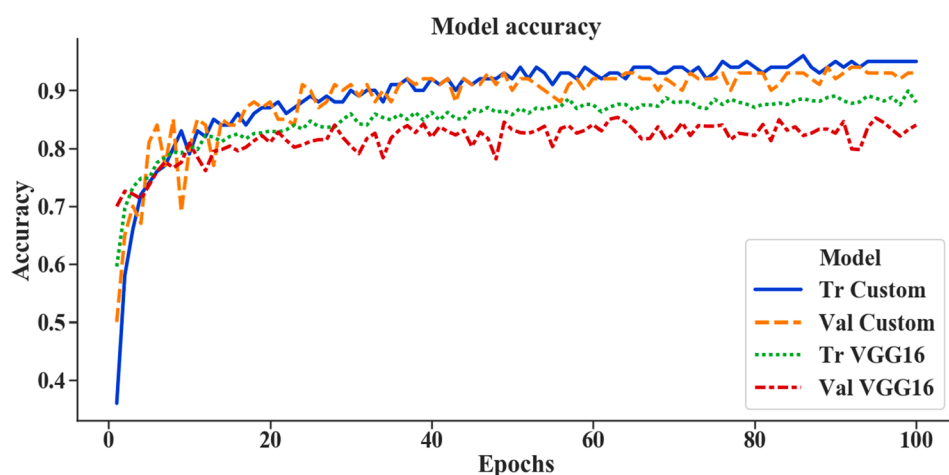


Fig. 14. Accuracy of the disease classifier. Tr and Val indicates the training and validation data sets.

MicaSense and WV-2 models, in which the features extracted from Bands + VIs + PCA were identified as the best to train both models. The RF classifier model exhibited an OA of 88% and 76% with $\leq 25\%$ of OE and CE, kappa of 0.81 and 0.53 for PS and S2, respectively (See Supplementary Table 5). Similarly, the SVM showed an OA of 75% and 69% with $\leq 32\%$ errors, kappa of 0.50 and 0.37 for PS and S2, respectively (Fig. 10; See Supplementary Table 6).

3.1.3. RF based pixelwise banana classification map

Fig. 11 shows the visual comparison of classification map results of Muhanhu Kabare district of eastern DR Congo generated using the RF model with different features for sensors (Fig. 11). It's evident that UAV MicaSense and WV-2 derived images showed higher accuracy and more detailed classification map results obtained with combined (B + VIs + PCA) features (Fig. 11). Also, we derived banana classification maps of two other regions (Inera and Kas) of Kabare district in addition to Mahanhu using RF model with best (B + VIs + PCA) features (Fig. 12). Banana classification maps of three regions of Kabare district using the best studied RF model are depicted in Fig. 12.

3.2. Banana localization and disease detection using UAV-RGB images

3.2.1. Performance metrics of diseases detection (BXW and BBTD)

To recognize the robust model to classify the major banana diseases (BXW, BBTD, we tested two models (VGG and Custom) and healthy plants (individual and banana clusters / mats). Reviewing the learning

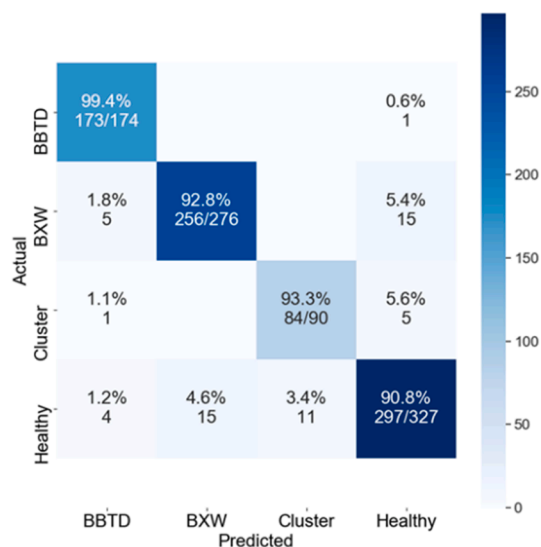


Fig. 15. Confusion matrix of custom model for the major banana disease classifier.

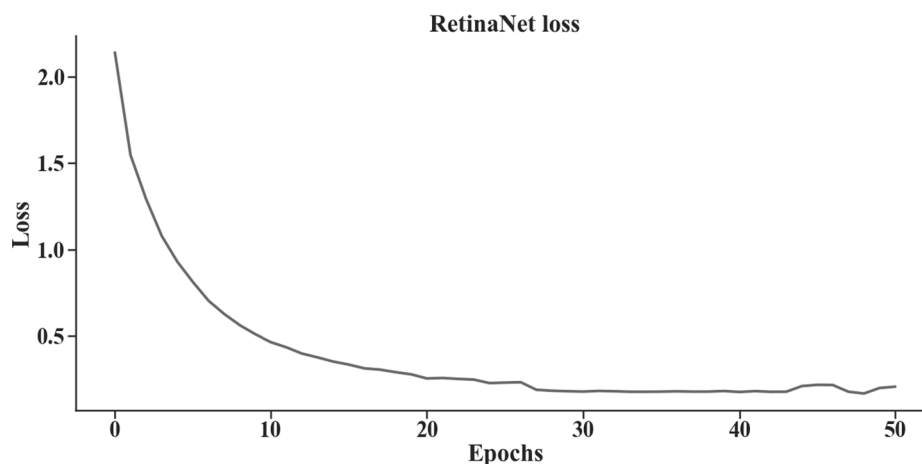


Fig. 16. The loss function for the object (banana) detection model.

Table 3

Evaluation metrics for the object (banana) detection model.

Metrics	Training dataset	Testing dataset
True positives	5564	825
False positives	1478	405
Precision	0.79	0.67
Recall	0.89	0.74

curves (Figs. 13 and 14) of the models during training, a continuous improvement was observed, showing that the developed models may benefit from further training epochs. However, the VGG model showed more variation in the loss function of the data validation set (Fig. 13), and the same was also noticed in the accuracy trend (Fig. 14). Our analysis revealed that the custom model was the best with higher accuracy (0.92) than VGG (0.85); therefore, the custom model can be the right candidate for banana disease classification using UAV-RGB images (Fig. 14). Our results on the trained models suggested that both dropout and data augmentation have positive effects on disease classification. The further combinations showed that our trained model has good control over the rate of training without overfitting, confirming that further improvements such as regularization, aggressive dropouts in later layers, the addition of weight decay could help improve the model. So far, we have not tuned the hyperparameters of the training algorithm, such as the learning rate, perhaps the most essential factor to achieve more accuracy.

The confusion matrix (CM) of the custom model revealed that the classification model developed from this study exhibited more than 90% accuracy (Fig. 15) in all the classes studied. High classification accuracy was observed in the disease classes of BXW (92.8%) and BBTD (99.4%).

In addition to CM, we have also computed the performance of other metrics, such as recall, precision, and F1 score. Recall was found around 0.99, 0.95, 0.92, 0.90, for the BBTD, BXW, healthy plant and cluster/mat classes, respectively. Precision metrics was found 0.98 for the BBTD class, 0.87 for healthy plant cluster/mat, 0.93 for healthy individual plant class, and 0.96 for BXW. The F1-score reported was: 0.92 for the BBTV class, 0.89 for cluster, 0.94 for healthy, and 0.90 for BXW.

3.2.2. Performance metrics for banana localization model

Monitoring the 50 epochs with the loss function shown in (Fig. 16), we found that after 30 epoch, the loss was stable and achieved a minimum value of 0.16 (Fig. 16). Recent literature also reported that training on higher accuracy allowed RetinaNet to detect pneumonia and road scenes analysis around 15 to 100 epochs (Blin et al., 2019; Liu et al., 2019). These results explain that the training model using more epochs will not necessarily decrease the loss significantly.

We extracted the precision and recall data using IoU to pull out the true positives (TPs) and false positives (FPs), and the results indicate that our model can detect 74% of the testing dataset (Table 3). This is an essential task when considering the predictions to the classifier for disease detection. In both the training and testing datasets, FPs are found to be higher than TPs, and this might be due to the reason, that it is nearly impossible to label all the individual and clustered banana in the orthomosaics, that the machine also recognized the non-labelled predictions as datasets and eventually resulted in high FPs.

4. Discussion

4.1. Banana classification using remote sensing & ML models

Banana production worldwide is profoundly affected by many pests and diseases. Rapid and early disease diagnosis are crucial for precision crop management, allowing targeted interventions. Timely and precise diagnosis is a critical initial step in mitigating losses caused by crop diseases. Before performing disease surveillance at a regional or country level, the first step is to detect/locate the target crop remotely with higher accuracy. Once the target crop is classified accurately, then deep learning disease detection models can be applied for disease surveillance.

In this study, the potential use of different VIs derived from UAV and different resolution satellite imagery (S2, PS, and WV-2) on pixel-based banana classification under mixed-complex landscape as well as the result of each VIs on classification accuracy were explored. For medium resolution sensors (S2), EVI and TGI are the only VIs that differentiate banana and non-banana classes. The VIs of CIG, and RVI-G derived from high (PS, WV-2), and very high resolution (MicaSense) sensors are found to be more promising and these VIs use NIR and Green bands to calculate the total chlorophyll content in the leaves (Ahamed et al., 2011). Also, other promising VIs such as ARVI, IPVI, and OSAVI exhibited a high correlation, which might be due to the nature of these VIs having intrinsic atmospheric and soil corrections (Huete et al., 1999). Most of the other VIs also showed an adequate correlation (more than 50%) between MicaSense and WV-2 (Fig. 9). These average correlations are mainly due to the spectral band difference between both sensors, atmospheric correction, spatial, and temporal resolution. It is expected that this correlation could be significantly improved in the future by establishing ground control points (GCPs) in the field to permit overlap between sensors.

The results of the ML-based banana classification model revealed the maximum overall accuracy (OA) with fewer errors (OE and CE), achieved by RF classifier using scenarios of B + VIs + PCA reached a level of 76% to 97% and kappa ranged between 0.53 and 0.96 based on the

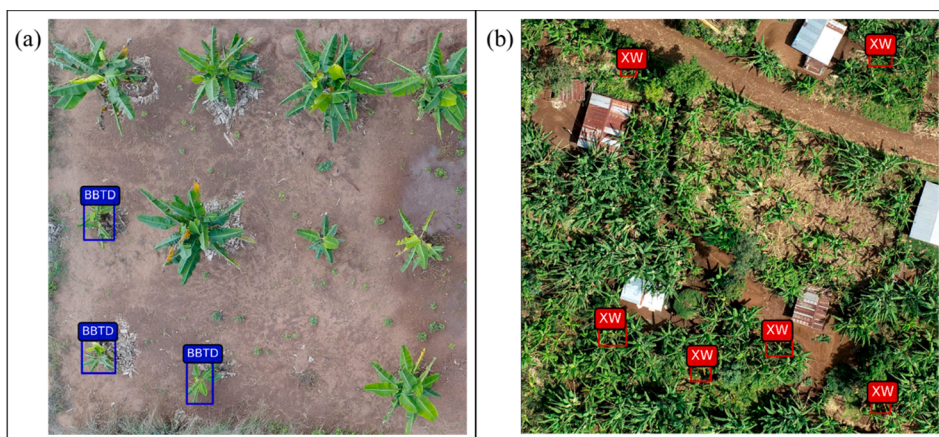


Fig. 17. Banana diseases detection using UAV-RGB based custom model (a) BBTB (blue bounding boxes), (b) BXW (red bounding boxes).

resolution of the sensors used (Fig. 10 a; Supplementary Table 5). The SVM classifier achieved lower OA (69% to 82%) and kappa (0.37 to 0.73) using scenarios of B + Vis + PCA but was lower than the RF classifier (Fig. 10b; Supplementary Table 6). RF proved to be a suitable model for multi-temporal crop classification due to its capacity to generate multiple paths with different variables to classify the same class (Zhong et al., 2019). RF model for crop classification with time-series MODIS data attained an accuracy of 88%, which confirms that the RF algorithm is suitable for selecting features and classifying crops when large volumes of data are used (Hao et al., 2015). The PCA trained model showed a better performance than the trained models with raw bands, and this is because the PCA values have no linear correlation between the components (Wold et al., 1987). When all features (bands, Vis, and PCA) were combined, the overall accuracy improved, showing that the combined information gives better features to classify banana and other classes accurately, even with low and medium resolution satellite images (Fig. 10). The best performed RF model with combined features (Bands + Vis + PCA) showed higher accuracy was used to map banana under mixed-complex African landscapes (Fig. 11&12). Simple Ratio (SR) Vis showed differentiable features and have been proven to be high potential to achieve a useful classification (Chuvieco, 1991). Crop classification from Sentinel-2-derived Vis using ensemble learning and RF model exhibited more than 90% OA (Sonobe et al., 2018). In this study, S2-based banana mapping does not show high accuracy (76% OA), which is mainly due to small size of the data points and mixed-complex setting of African landscapes and. Since African landscapes have mostly small-scale banana farms and often mixed with other classes (e.g., buildings, pastures, trees), the use of low and medium resolution satellite images to classify banana is highly challenging. The low-resolution pixels combined with sparse banana plantations would provide only one spectral signature from different classes resulting in similar Vis responses and bad classification models. The high-resolution sensor-based RF models (UAVs, WV-2 and PS) are found to be more promising and accurate to map banana in mixed African landscapes. Merging very high-resolution data (UAVs & WV-2) with open-source medium resolution satellite information to increment the effectiveness in classification and decision-making seems to be the way forward (Johansen et al., 2020). Improving the spatial resolution of satellites could help replace skill and regulation oriented UAV technologies and monitoring of larger areas with high spatial and temporal resolution (Mayes et al., 2016) in developing countries. Moreover, the use of UAVs to monitor larger landscapes are challenging, so satellite-based ML models could help to classify banana and UAVs and smartphone-based sensors can be utilized to detect more definite problems, such as the disease and health status.

4.2. Mixed model pipeline for banana detection and disease classification

The robust object detection (banana) and custom disease classification mixed model (Fig. 5) developed from this study demonstrate the capability of low-cost UAV-RGB images, and deep learning approaches of disease surveillance tool in Africa. To detect diseases accurately in mixed-complex African landscapes, it is necessary to detect banana plants irrespective of its phenological stages (young to mature) in an individual or cluster banana planting distribution. As showed in Fig. 5, our RetinaNet object detection model accurately detected banana plants regardless of their age and size automatically through orthomosaics, then the developed custom model classified the diseases (BXW and BBTB) using the detector (Fig. 5; Fig. 17). To monitor larger landscapes with good accuracy, we identified the images range from 50 to 100 m height to be sufficient and the orthomosaics are well georeferenced to classify the diseased plants in the mixed landscape without any problems. However, training a RetinaNet model needs intensive labeling to annotate huge orthomosaics, which is time consuming and it took around one month for 50 epochs. Currently, we are in the process of training the developed RetinaNet model with more epochs to improve the model performance.

To check the robustness of the developed custom model, metrics such as loss function, accuracy, and confusion matrix were computed (Figs. 13–15). Even though CM of the custom classification model exhibited more than 90% accuracy (Fig. 15) in all classes studied, the major confusion was noticed between healthy individual banana, banana clusters, and BXW (Fig. 15). This is because the big individual healthy banana and small healthy banana clusters in the mixed landscape looked similar and the model confused little bit (5.6%) and healthy individual banana classes are confused with BXW classes (4.6%) since normal yellow streaks confused with BXW symptoms, but the percentage of confusion was very low. These inter-class confusions can be avoided by collecting and training more distinct features of those classes. As expected, the confusion between BBTB and BXW was very low since these two symptoms are very contrasting with each other (Fig. 4). However, the major challenge of this developed RGB based custom model is to differentiate BXW and closely related wilt such as fusarium wilt. Since the UAV imagery obtained by the RGB sensor has only visual bands, the differences in spectral characteristics between healthy and diseased plants cannot be explained well. So, multi- and hyperspectral data should be further investigated to study the sensitivity of certain bands of BXW. Besides, changes in the spectral characteristic between BXW, Fusarium wilt, and other physiological yellowing phenomena caused by different abiotic stresses such as heat, drought, and nutrition disorders should also be scrutinized. Even though this study discussed different laborious DL approaches, the mixed model framework developed in this study have a simple front-end such a way the user

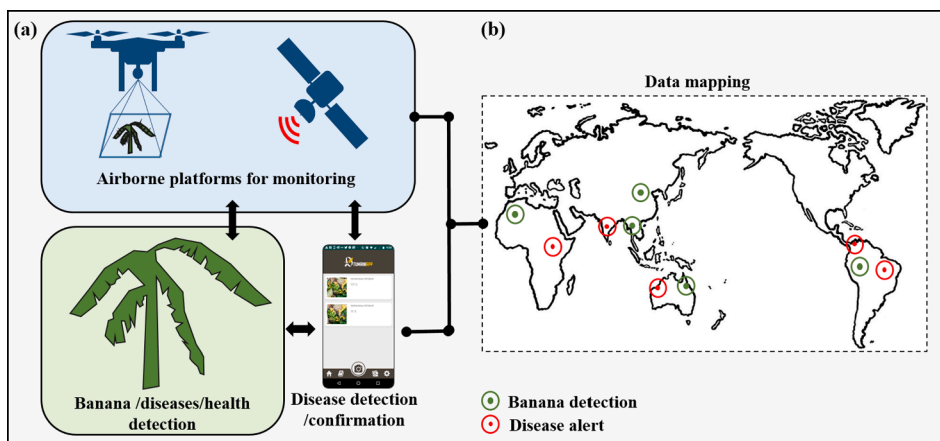


Fig. 18. Envisaged system-level AI-powered disease monitoring system comprising of (a) data extraction and analysis, and (b) data mapping and alerts/early warning.

can upload orthomosaics and receive the classified result with a bounding labels for the major banana diseases (Fig. 17).

4.3. Towards AI-Powered disease surveillance system in Africa

The use of artificial intelligence (AI) to predict crop health and related environmental impacts are developing rapidly. Here we propose a system-level AI-based banana disease alert system (Fig. 18) by integrating high & low-resolution aerial (UAV and satellite) imagery coupled with advanced computer vision algorithms (Smart phone based AI apps). Additional ground-truth data can be achieved through expert field observations and the use of validated AI-powered disease symptom assessment apps (e.g., the Tumaini app). The validated Tumaini AI-powered mobile app (https://play.google.com/store/apps/details?id=ciat.cgiar.org.tumaini&hl=es_419) can also serve as the ground-truth for the BXW and BBTB classification model since these diseases could be easily detected (more than 90% accuracy) under real field conditions (Selvaraj et al., 2019). Moreover, GPS tagged mobile app images will also help to reconfirm RF-based banana classification models developed from this study (Fig. 18). It is obvious that a combination of robust AI detection models with medium to high spatio-temporal aerial image data, and our validated AI-powered mobile app (Tumaini), together with local weather data, can lead to a robust early warning alert systems across banana production landscapes in Africa (Fig. 18). Outputs and models developed from this study can also be integrated into the banana mapper (<http://www.crop-mapper.org/banana/index.html>) and PestDisPlace (<https://pestdisplace.org/>) platforms developed by the Alliance of Bioversity and CIAT to enhance banana mapping and diseases surveillance at a global scale (Fig. 18b).

5. Conclusions and future directions

Our pixel-based banana classification from random forest (RF) model using combined features of vegetation indices (VIs) along with principal component analysis (PCA) showed promising option to map banana under mixed-complex African landscapes. The high resolution sensors (UAVs, PS and WV-2) used in this study were found to be more accurate to map banana than medium resolution satellite images (S2). Accurate banana mapping using open-source medium resolution satellites (S2) under mixed complex system is still challenging, but integrating informations from different (medium and high resolution) sensor sources is way forward. Through UAV-based RGB imagery systems, we were able to detect banana and their major diseases with higher accuracy and less errors. The low-cost UAV-RGB based mixed-model disease classification pipeline developed from this study is feasible and could be further extended to other crop diseases. This mixed-model pipeline based

banana disease classification system developed from this study can be further strengthened by training the CNN model using a larger dataset with multiple crops/diseases, which will be our impending work. The work is underway to collect more ground data points and ground-truth measurements from our global banana partners across Latin America, Africa, and India to enhance existing datasets and validate the developed ML algorithms. The output of this research paper are being integrated into other banana disease surveillance platforms of the CGIAR Research Program on Roots, Tubers and Bananas (RTB) to enhance the digital disease monitoring system at a global scale.

Declaration of Competing Interest

The authors declared that there is no conflict of interest.

Acknowledgments

The authors would like to thank the Alliance of Bioversity International and International Center for Tropical Agriculture (CIAT) Information Technology unit for providing facilities and logistics support. The authors would like to thank Joe Tohme, from the International Center for Tropical Agriculture (CIAT) for his support in this research. The authors would also like to acknowledge Milton Valencia, Manuel Valderrama, Jorge Casas, and Maria Montoya for their help in image annotation. Thanks to Nancy Safari, Jules Ntamwira and Jean-Pierre Mafuta of the Bioversity Bakavu, DR Congo office Innocent Nduwimana of Bioversity Burundi, and Deo Kantungeko of IITA, Burundi for their immense support to collect smartphone images. Thanks to Maxar globe Inc. for providing high resolution images through the CGIAR-Big-data collaboration. We would also like to thank Dr. Sindhuja Sankaran of Washington State University for providing PS satellite imagery, as well as Angela Fernando, CIAT consultant for formatting and technical editing of the MS. Bioversity International provided funding for field UAV image collection in the framework of the RTB-CC3.1. and RTB-BA3.3, RTB-BA 3.4 clusters; Bioversity International and crop nutrition and health research area, CIAT to carry out the image processing and Machine learning model development work (AGBIO1). We thank the RTB Program Management Unit that supported this study and the CGIAR Fund Donors who support RTB (www.cgiar.org/who-we-are/cgiar-fund/fund-donors-2).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.isprsjprs.2020.08.025>.

References

- Agisoft, 2020. Agisoft Metashape. URL <https://www.agisoft.com/> (accessed 4.13.20).
- Ahamed, T., Tian, L., Zhang, Y., Ting, K.C., 2011. A review of remote sensing methods for biomass feedstock production. *Biomass Bioenergy* 35, 2455–2469. <https://doi.org/10.1016/j.biombioe.2011.02.028>.
- Azar, R., Villa, P., Stroppiana, D., Crema, A., Boschetti, M., Brivio, P.A., 2016. Assessing in-season crop classification performance using satellite data: A test case in Northern Italy. *Eur. J. Remote Sens.* 49, 361–380. <https://doi.org/10.5721/EuJRS20164920>.
- Blanzieri, E., Melgani, F., 2008. Nearest neighbor classification of remote sensing images with the maximal margin principle. *IEEE Trans. Geosci. Remote Sens.* 46, 1804–1811. <https://doi.org/10.1109/TGRS.2008.916090>.
- Blin, R., Ainouz, S., Canu, S., Meriaudeau, F., 2019. Road scenes analysis in adverse weather conditions by polarization-encoded images and adapted deep learning: in: 2019 IEEE Intelligent Transportation Systems Conference, ITSC 2019. pp. 27–32. <https://doi.org/10.1109/ITSC.2019.8916853>.
- Blomme, G., Dita, M., Jacobsen, K.S., Vicente, L.P., Molina, A., Ocimati, W., Poussier, S., Prior, P., 2017. Bacterial diseases of bananas and enset: Current state of knowledge and integrated approaches toward sustainable management. *Front. Plant Sci.* <https://doi.org/10.3389/fpls.2017.01290>.
- Blomme, G., Jacobsen, K., Ocimati, W., Beed, F., Ntamwira, J., Sivirihauma, C., Ssekigwoko, F., Nakato, V., Kubiriba, J., Tripathi, L., Tinjaara, W., Mbolela, F., Lutete, L., Karamura, E., 2014. Fine-tuning banana *Xanthomonas* wilt control options over the past decade in East and Central Africa. *Eur. J. Plant Pathol.* 139, 265–281. <https://doi.org/10.1007/s10658-014-0402-0>.
- Boullet, J., Beaulieu, M., St-Charles, P.-L., Théau, J., Foucher, S., 2019. Deep learning for in-field image-based grapevine downy mildew identification, in: Precision Agriculture 2019. Wageningen Academic Publishers, pp. 141–148. <https://doi.org/10.3920/978-90-8686-888-9.16>.
- Bouwmeester, H., Heuvelink, G.B.M., Stoorvogel, J.J., 2016. Mapping crop diseases using survey data: The case of bacterial wilt in bananas in the East African highlands. *Eur. J. Agron.* 74, 173–184. <https://doi.org/10.1016/j.eja.2015.12.013>.
- Breiman, L., 2001. *Machine Learning* 45, 5–32. <https://doi.org/10.1023/A:1010933404324>.
- Campbell, J.B., Wynne, R.H., 2011. Introduction to Remote Sensing, The Guilford press. <https://doi.org/10.1007/s13398-014-0173-2>.
- Carvajal-Yepes, M., Cardwell, K., Nelson, A., Garrett, K.A., Giovani, B., Saunders, D.G.O., Kamoun, S., Legg, J.P., Verdier, V., Lessel, J., 2019. A global surveillance system for crop diseases. *Science* (80-). 364, 1237–1239. <https://doi.org/10.1126/science.aaw1572>.
- Chuvieco, E., 1991. Fundamentos de teledetección espacial. *Estud. Geogr.* 52, 371.
- De Buck, S., Swennen, R., 2016. Bananas, the green gold of the South. *VIB fact Ser.* 1, 1–54.
- Digital Globe, M., 2016. *WorldView-2 data sheet*.
- ESA, 2015. *Sentinel-2 User Handbook*. European Space Agency (ESA).
- FAOSTAT, F., 2017. Food and Agriculture Organization of the United Nations (FAOSTAT). Data of crop production. URL <http://www.fao.org/faostat/en/#data/QC>.
- Fizyr, 2019. Keras RetinaNet. GitHub Repos. <https://doi.org/10.5281/zenodo.3250670>.
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* 202, 18–27. <https://doi.org/10.1016/j.rse.2017.06.031>.
- Hao, P., Zhan, Y., Wang, L., Niu, Z., Shakir, M., 2015. Feature selection of time series MODIS data for early crop classification using random forest: A case study in Kansas, USA. *Remote Sens.* 7, 5347–5369. <https://doi.org/10.3390/rs70505347>.
- Heim, R.H.J., Wright, L.J., Allen, A.P., Geedicke, I., Oldeland, J., 2019. Developing a spectral disease index for myrtle rust (*Austropuccinia psidii*). *Plant. Pathol.* 68, 738–745. <https://doi.org/10.1111/ppa.12996>.
- Hsu, C.-W., Chang, C.-C., Lin, C.-J., 2003. A practical guide to support vector classification. Taipei.
- Huang, J., Rathod, V., Chow, D., Sun, C., Zhu, M., Fathi, A., Lu, Z., 2017. Tensorflow object detection api.
- Huete, A., Justice, C., van Leeuwen, W., 1999. *Modis Vegetation Index (MOD13). Algorithm Theor. basis Doc. 3*.
- Ji, S., Zhang, C., Xu, A., Shi, Y., Duan, Y., 2018. 3D convolutional neural networks for crop classification with multi-temporal remote sensing images. *Remote Sens.* 10, 75. <https://doi.org/10.3390/rs10010075>.
- Jia Deng, Wei Dong, Socher, R., Li-Jia Li, Kai Li, Li Fei-Fei, 2009. ImageNet: A large-scale hierarchical image database, in: IEEE Conference on Computer Vision and Pattern Recognition. <https://doi.org/10.1109/CVPR.2009.5206848>.
- Johansen, K., Duan, Q., Tu, Y.H., Searle, C., Wu, D., Phinn, S., Robson, A., McCabe, M.F., 2020. Mapping the condition of macadamia tree crops using multi-spectral UAV and WorldView-3 imagery. *ISPRS J. Photogramm. Remote Sens.* 165, 28–40. <https://doi.org/10.1016/j.isprs.2020.04.017>.
- Johansen, K., Sohlbach, M., Sullivan, B., Stringer, S., Peasley, D., Phinn, S., 2014. Mapping banana plants from high spatial resolution orthophotos to facilitate plant health assessment. *Remote Sens.* 6, 8261–8286. <https://doi.org/10.3390/rs6098261>.
- Lin, T.-Y., Goyal, P., Girshick, R., He, K., Dollár, P., 2017. Focal Loss for Dense Object Detection. Proceedings of the IEEE International Conference on Computer Vision 2999–3007. <https://doi.org/10.1109/ICCV.2017.324>.
- Liu, M., Tan, Y., Chen, L., 2019. Pneumonia detection based on deep neural network Retinanet. 2019 International Conference on Image and Video Processing, and Artificial Intelligence. International Society for Optics and Photonics 113210F. <https://doi.org/10.1117/12.2539633>.
- Lu, J., Miao, Y., Huang, Y., Shi, W., Hu, X., Wang, X., Wan, J., 2015. Evaluating an unmanned aerial vehicle-based remote sensing system for estimation of rice nitrogen status, in: 2015 Fourth International Conference on Agro-Geoinformatics (Agro-Geoinformatics). IEEE, pp. 198–203. <https://doi.org/10.1109/Agro-Geoinformatics.2015.7248117>.
- Lucas, I.F.J., Janssen, F., van der Wel, F.J., 1994. Accuracy assessment of satellite derived landcover data: A review. *Photogramm. Eng. Remote Sens.* 60, 426–479.
- Mahlein, A.-K., Steiner, U., Dehne, H.-W., Oerke, E.-C., 2010. Spectral signatures of sugar beet leaves for the detection and differentiation of diseases. *Precis. Agric.* 11, 413–431. <https://doi.org/10.1007/s11119-010-9180-7>.
- Mané, D., 2015. TensorBoard: TensorFlow's visualization toolkit.
- Mao, J., Tian, W., Li, P., Wei, T., Liang, Z., 2017. Phishing-alarm: robust and efficient phishing detection via page component similarity. *IEEE Access* 5, 17020–17030. <https://doi.org/10.1109/ACCESS.2017.2743528>.
- Mayes, M.T., Estes, L.D., Gago, X., Debats, S.R., Caylor, K.K., Manfreda, S., Oudemans, P., Ciraolo, G., Maltese, A., Nadal, M., 2016. Using Small Drone (UAS) Imagery to Bridge the Gap Between Field-and Satellite-Based Measurements of Vegetation Structure and Change. *AGU Fall Meeting Abstracts*. B53J-B105.
- Murthy, C.S., Raju, P.V., Badrinath, K.V.S., 2003. Classification of wheat crop with multi-temporal images: Performance of maximum likelihood and artificial neural networks. *Int. J. Remote Sens.* 24, 4871–4890. <https://doi.org/10.1080/0143116031000070490>.
- Neupane, B., Horanont, T., Hung, N.D., 2019. Deep learning based banana plant detection and counting using high-resolution red-green-blue (RGB) images collected from unmanned aerial vehicle (UAV). *PLoS ONE* 14.
- Niyongere, C., Losenge, T., Ateka, E., Nkezahizi, D., Blomme, G., Lepoint, P., 2012. Occurrence and distribution of banana bunchy top disease in the Great Lakes Region of Africa. *Tree For. Sci. Biotechnol.* 6, 102–107.
- Niyongere, C., Omondi, A.B., Blomme, G., 2015. The Banana bunchy top disease. In: Tennant, P., Fermin, G. (Eds.), *Virus Diseases of Tropical and Subtropical Crops*. CAB International Plant Protection Series, Wallingford, United Kingdom, pp. 17–26.
- Ocimati, W., Bouwmeester, H., Groot, J.C.J., Tittonell, P., Brown, D., Blomme, G., 2019. The risk posed by *Xanthomonas* wilt disease of banana: Mapping of disease hotspots, fronts and vulnerable landscapes. *PLoS ONE* 14, 1–19. <https://doi.org/10.1371/journal.pone.0213691>.
- Olyphant, A.J., Thenkabail, P.S., Telugunta, P., Xiong, J., Gumma, M.K., Congalton, R.G., Yadav, K., 2019. Mapping cropland extent of Southeast and Northeast Asia using multi-year time-series Landsat 30-m data using a random forest classifier on the Google Earth Engine Cloud. *Int. J. Appl. Earth Obs. Geoinf.* 81, 110–124. <https://doi.org/10.1016/j.jag.2018.11.014>.
- Peña, J.M., Gutiérrez, P.A., Hervás-Martínez, C., Six, J., Plant, R.E., López-Granados, F., 2014. Object-based image classification of summer crops with machine learning methods. *Remote Sens.* 6, 5019–5041. <https://doi.org/10.3390/rs6065019>.
- Pix4D, 2018. Pix4Dcapture. URL <https://play.google.com/store/apps/details?id=com.pix4d.pix4dmapper> (accessed 4.13.20).
- Planet, 2017. Planet Imagery : Product Specification, Planet.
- Pourazar, H., Samadzadegan, F., Dadrass Javan, F., 2019. Aerial multispectral imagery for plant disease detection: radiometric calibration necessity assessment. *Eur. J. Remote Sens.* 52, 17–31. <https://doi.org/10.1080/22797254.2019.1642143>.
- Ramcharan, A., Baranowski, K., McCloskey, P., Ahmed, B., Legg, J., Hughes, D.P., 2017. Deep learning for image-based cassava disease detection. *Front. Plant Sci.* 8, 1852. <https://doi.org/10.3389/fpls.2017.01852>.
- Sankaran, S., Khot, L.R., Carter, A.H., 2015. Field-based crop phenotyping: Multispectral aerial imaging for evaluation of winter wheat emergence and spring stand. *Comput. Electron. Agric.* 118, 372–379. <https://doi.org/10.1016/j.compag.2015.09.001>.
- Selvaraj, M.G., Valderrama, M., Guzman, D., Valencia, M.O., Ruiz, H., Acharjee, A., 2020. Machine learning for high-throughput field phenotyping and image processing provides insight into the association of above and below-ground traits in cassava (*Manihot esculenta* Crantz). *Plant Methods*. <https://doi.org/10.21203/rs.2.24148/v1>.
- Selvaraj, M.G., Vergara, A., Ruiz, H., Safari, N., Elayabalan, S., Ocimati, W., Blomme, G., 2019. AI-powered banana diseases and pest detection. *Plant Methods* 15, 92. <https://doi.org/10.1186/s13007-019-0475-z>.
- Shi, Y., Thomasson, J.A., Murray, S.C., Pugh, N.A., Rooney, W.L., Shafian, S., Rajan, N., Rouze, G., Morgan, C.L.S., Neely, H.L., 2016. Unmanned aerial vehicles for high-throughput phenotyping and agronomic research. *PLoS ONE* 11, e0159781. <https://doi.org/10.1371/journal.pone.0159781>.
- Sonobe, R., Yamaya, Y., Tani, H., Wang, X., Kobayashi, N., Mochizuki, K., 2018. Crop classification from Sentinel-2-derived vegetation indices using ensemble learning. *J. Appl. Remote Sens.* 12, 026019. <https://doi.org/10.1117/1.jrs.12.026019>.
- Stehman, S.V., 1997. Selecting and interpreting measures of thematic classification accuracy. *Remote Sens. Environ.* 62, 77–89. [https://doi.org/10.1016/S0034-4257\(97\)00083-7](https://doi.org/10.1016/S0034-4257(97)00083-7).
- Steward, B.L., Gai, J., Tang, L., 2019. *The use of agricultural robots in weed management and control*, in: *Agricultural and Biosystems Engineering Publications*.
- Tower, D., 2017. Ground Control Station for Android Devices. URL <https://github.com/DroidPlanner/Tower> (accessed 4.13.20).
- Tzutalin, 2015. *LabelImg* Git code.
- Wold, S., Esbensen, K., Geladi, P., 1987. Principal component analysis. *Chemom. Intell. Lab. Syst.* 2, 37–52. [https://doi.org/10.1016/0169-7439\(87\)80084-9](https://doi.org/10.1016/0169-7439(87)80084-9).
- Yang, Z., Willis, P., Mueller, R., 2008. Impact of Band-Ratio Enhanced Aerial Image To Crop Classification Accuracy. *Pecora* 17 – Futur. L. Imaging...Going Oper. 11.
- Zhong, L., Hu, L., Zhou, H., 2019. Deep learning based multi-temporal crop classification. *Remote Sens. Environ.* 221, 430–443. <https://doi.org/10.1016/j.rse.2018.11.032>.