UNIVERSITÄT AUGSBURG
Universitätsbibliothek

OPEN ACCESS
OPUS AUGSBURG

From pixels to plant health: accurate detection of banana Xanthomonas wilt in complex African landscapes using high-resolution UAV images and deep learning

Juan Jose Mora, Michael Gomez Selvaraj, César Iván Álvarez, Nancy Safari, Guy Blomme

Angaben zur Veröffentlichung / Publication details:

Mora, Juan Jose, Michael Gomez Selvaraj, César Iván Álvarez, Nancy Safari, and Guy Blomme. 2024. "From pixels to plant health: accurate detection of banana Xanthomonas wilt in complex African landscapes using high-resolution UAV images and deep learning." *Discover Applied Sciences* 6: 377. https://doi.org/10.1007/s42452-024-06073-z.

Nutzungsbedingungen / Terms of use:

CC BY 4.0

Dieses Dokument wird unter folgenden Bedingungen zur Verfügung gestellt: / This document is made available under these conditions:
CC-BY 4.0: Creative Commons: Namensnennung
Weitere Informationen finden Sie unter: / For more information see:
https://creativecommons.org/licenses/by/4.0/deed.de

Discover Applied Sciences

Research

From pixels to plant health: accurate detection of banana Xanthomonas wilt in complex African landscapes using high-resolution UAV images and deep learning

Juan Jose Mora[1] · Michael Gomez Selvaraj[1] · Cesar Ivan Alvarez[2] · Nancy Safari[3] · Guy Blomme[4]

Received: 8 April 2024 / Accepted: 1 July 2024
Published online: 11 July 2024
© The Author(s) 2024 OPEN

Abstract

Bananas and plantains are vital for food security and smallholder livelihoods in Africa, but diseases pose a significant threat. Traditional disease surveillance methods, like field visits, lack accuracy, especially for specific diseases like Xanthomonas wilt of banana (BXW). To address this, the present study develops a Deep-Learning system to detect BXW-affected stems in mixed-complex landscapes within the Eastern Democratic Republic of Congo. RGB (Red, Green, Blue) and multispectral (MS) images from unmanned aerial vehicles UAVs were utilized using pansharpening algorithms for improved data fusion. Using transfer learning, two deep-learning model architectures were used and compared in our study to determine which offers better detection capabilities. A single-stage model, Yolo-V8, and the second, a two-stage model, Faster R-CNN, were both employed. The developed system achieves remarkable precision, recall, and F1 scores ranging between 75 and 99% for detecting healthy and BXW-infected stems. Notably, the RGB and PAN UAV images perform exceptionally well, while MS images suffer due to the lower spatial resolution. Nevertheless, specific vegetation indexes showed promising performance detecting healthy banana stems across larger areas. This research underscores the potential of UAV images and Deep Learning models for crop health assessment, specifically for BXW in complex African systems. This cutting-edge deep-learning approach can revolutionize agricultural practices, bolster African food security, and help farmers with early disease management. The study's novelty lies in its Deep-Learning algorithm development, approach with recent architectures (Yolo-V8, 2023), and assessment using real-world data, further advancing crop-health assessment through UAV imagery and deep-learning techniques.

Article Highlights

- Deep-learning models are developed using drone-captured images to detect BXW-affected banana stems in complex landscapes.
- Combinations of vegetation indexes and pansharpening techniques enhance low-resolution multispectral images.
- The system achieves a 75% to 99% precision, recall, and F1 score for identifying healthy and BXW-infected stems in RGB, MS, and pansharpening images.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s42452-024-06073-z.

✉ Michael Gomez Selvaraj, m.selvaraj@cgiar.org; Juan Jose Mora, j.mora@cgiar.org; Cesar Ivan Alvarez, calvarezm@ups.edu.ec; Nancy Safari, nancysafari21@gmail.com; Guy Blomme, g.blomme@cgiar.org | [1]Alliance of Bioversity International and International Center for Tropical Agriculture (CIAT), A.A. 6713, Cali, Colombia. [2]Environmental Research Group for Sustainable Development (GIADES), Salesian Polytechnic University, Rumichaca y Moran Valverde, Quito, EC 170702, Ecuador. [3]Bioversity International, Bukavu, South Kivu, Democratic Republic of Congo. [4]Bioversity International, c/ o ILRI, P.O. Box 5689, Addis Ababa, Ethiopia.

Keywords Xanthomonas wilt · Unmanned aerial vehicles (UAV) · Faster R-CNN · Yolo-V8 · RGB · Multispectral · Panchromatic fusion

Abbreviations

AI	Artificial intelligence
BNDVI	Blue normalized difference vegetation index
BXW	Xanthomonas wilt of banana
CARI	Chlorophyll absorption ratio index
CGIAR	Consultative group for international agricultural research
CIAT	International center for tropical agriculture
CM	Confusion matrix
CNN	Convolutional neural network
DA	Data augmentation
DL	Deep learning
DR Congo	Democratic Republic of Congo
DT	Dataset
ECA	East and Central Africa
Faster RCNN	Faster Region-based Convolutional Neural Network
GNDVI	Green normalized difference vegetation index
GPU	Graphics processing unit
GRVI	Green red vegetation index
GTD	Ground-truth data
HIS	Hue intensity saturation
IOU	Intersection over union
MGTE	Missed ground truth error
ML	Machine learning
MS	Multispectral
NDVI	Normalized difference vegetation index
NDRE	Normalized difference red edge index
NIR	Near infrared
NPCI	Normalized pigment chlorophyll ratio index
P4P	Phantom 4 pro
PAN	Pansharpening
RGB	Red, green blue
ROIs	Regions of interest
RPN	Region proposal network
SSD	Single shot MultiBox detector
UAV	Unmanned aerial vehicle
Vis	Vegetation indices
Yolo-V8	You only look once version 8

1 Introduction

Bananas and plantains are essential staple food crops in sub-Saharan Africa, making significant contributions to household food and nutrition security and the income of rural populations. Sub-Saharan Africa alone accounts for approximately one-third of global banana production, and these crops fulfill over 25% of the food energy requirements for more than 100 million people [1]. Within Africa, East and Central Africa (ECA) stands out as the largest region for both banana production and consumption. Around 20 million people in East Africa and 70 million in West and Central Africa rely on bananas, emphasizing the critical role of these crops as primary food sources [2]. In addition, bananas and plantains play a crucial role in smallholder communities' economic stability and employment opportunities, contributing to their sustainable livelihoods and rural development.

Regrettably, banana production systems in the region face significant challenges posed by various biotic stresses, including bacterial, viral, and fungal diseases and pests. These factors often lead to substantial crop losses, causing severe impacts on the livelihoods of small-scale farmers [3]. Xanthomonas wilt of banana (BXW) is one of the most devastating diseases affecting bananas. It emerged as a major threat in 2001 and has since spread to most banana production zones in East and Central Africa. This disease has had a detrimental effect on the livelihoods of smallholder communities. BXW primarily spreads through contaminated garden tools, insect vectors, and infected planting materials; if uncontrolled, it can result in complete crop losses. Common symptoms of BXW include yellowing and wilting leaves, premature bunch ripening, fruit pulp discoloration, and yellow ooze in pseudostems and leaf vascular tissues [4, 5].

Early detection of crop diseases is crucial for timely intervention and minimizing potential yield losses, leading to maximum benefits. Unfortunately, traditional approaches in disease surveillance, such as scouting and monitoring for crop diseases, have been reliant on visual inspection during fieldwork visits, which is tedious, subjective, and often results in erroneous observations [6]. In Africa, disease surveillance for BXW commonly combines field campaigns and laboratory analysis employing diagnostic tools like growth media and polymerase chain reaction (PCR) [7, 8]. Nevertheless, this conventional approach lacks the ability to provide precise distribution data on BXW infections, which hinders effective disease management strategies in the complex landscapes where bananas are grown. To address these challenges and improve disease surveillance, innovative and technology-driven solutions are necessary to safeguard the region's banana production and the well-being of its farmers. Nevertheless, the conventional approach of assessing Banana Xanthomonas Wilt (BXW) through visual observation is characterized by its consumption of time and potential ineffectiveness, particularly when confronted with expansive areas necessitating vigilant oversight [9]. The incorporation of cutting-edge methodologies like remote sensing, digital imaging, and data analytics exhibits significant potential to fundamentally transform the landscape of disease identification and monitoring within the realm of agriculture [9].

Agricultural drones or unmanned aerial vehicles (UAVs) have emerged as an ideal technology for evaluating crop yields, assessing the health status of crops, and conducting economic evaluations [10]. The rapid advancement of UAV systems, along with the availability of affordable UAVs equipped with visible (RGB) and multispectral (MS) sensors, has unlocked new possibilities for capturing high-resolution spatiotemporal data, enabling the early identification of disease symptoms, even before they become visible to human eye [11, 12]. Based on their captured bands, UAV images can be categorized into RGB and multispectral (MS) images. While RGB cameras have been successful in identifying crop diseases [9, 13, 14], they only measure three bands of the electromagnetic spectrum (red, green, and blue) [15]. In contrast, multispectral images, captured using dedicated sensors, record energy reflected in multiple specific electromagnetic spectrum bands (475–717 nm) [16]. These images provide additional information and enable the calculation of vegetation indices (VIs), which are crucial for the early identification of crop diseases [17]. Consequently, not all wavelengths are suitable for accurate crop disease detection [15]. Therefore, vegetation indexes and the integration of pansharpening methods have enhanced the utility of UAV images by combining RGB and MS images, enabling higher spatial and spectral resolution and overcoming the limitations of each individual image type or multispectral image [18].

Data analytics in disease surveillance empowers farmers with timely and accurate information, enabling them to make informed decisions and safeguard their crops and livelihoods. The images analyzed using deep learning (DL) have revolutionized crop disease classification, achieving remarkable accuracy in image processing and object detection [19]. Two-stage and one-stage DL networks, such as Faster R-CNN and Yolo-V8, respectively, have been developed for image detection, where two-stage models have been found to offer higher accuracy by first identifying the region of interest and then performing classification and bounding box regression. In contrast, one-stage models directly predict the class and bounding boxes of objects in one single pass [9, 20–22]. Faster R-CNN architecture has been applied in several other agricultural artificial-intelligence (AI) projects, detecting diseases and pests in crops like rice, roses, sweet peppers, and bananas [20, 23–26]. On the other hand, Yolo-V8, while being a recently released architecture, has been used for phenotype calculation of rapeseed pods based and small object identification in unmanned aerial vehicle images [26, 27]. The practical application of DL in agriculture, particularly for crop disease detection, has highlighted promising research value and market prospects [28].

Additional studies have focused on various deep learning methods for detecting and classifying foliar diseases in different crops. Mishra et al. [29] discuss the identification of weed density in soya bean fields using the Inception V4 deep convolutional neural network architecture. By employing RGB images and advanced vegetation segmentation techniques, the study achieved precision, recall, and F1 scores of 97, 99, and 98%, respectively. Similarly, Nagaraju et al. [30] address the challenge of overfitting in deep learning models used for agricultural image classification by introducing two novel algorithms: IPTA (an adaptive supervised learning approach for image transformation) and IMHSA (an unsupervised RGB image segmentation method), significantly improving the model's validation accuracy

from 65 to 73%. Subsequently, Kaur et al. [31] focus on plant disease identification in grapevines, utilizing transfer learning with EfficientNet B7 and logistic regression to detect four types of diseases. The proposed method achieved a constant accuracy of 98.7% after 92 epochs, showcasing its effectiveness in comprehensive disease detection. Although DL-based instance segmentation for field-based disease classification using UAVs has seen advancements [12, 19, 32, 33]; however, efforts specifically targeting banana diseases have been limited. Neupane et al. [34] proposed a machine-learning-based approach for banana disease classification and height estimation in well-planned commercial/mono-cropped banana fields using UAV-RGB images. However, implementing these models in African agricultural systems, characterized by mixed-complex landscapes, presents challenges [9]. While previous studies highlight the use of deep learning and convolutional neural networks for disease detection at the leaf level, as well as image segmentation and data augmentation, this paper distinguishes itself by employing aerial imagery, incorporating multiple sensors for image input, and utilizing vegetation index manipulation and pansharpening strategies. It is important to highlight the pansharpening process because, in the case of satellite remote sensing, we have advantages compared to using only RGB or MS bands for object detection. Additionally, it establishes comparisons between one-stage and two-stage models, emphasizing its unique approach in leveraging UAV imagery and deep learning techniques for crop health assessment. This paper addresses the urgent need for a robust deep learning-based disease detection model to accurately identify healthy and diseased plants in mixed-complex, banana-based landscapes in real-life scenarios.

Validation of unseen data is essential for assessing the effectiveness of deep learning models and bolstering their resilience and precision in real-world scenarios [35]. Nevertheless, many studies in deep learning tend to underestimate the significance of appraising models using new and unfamiliar data, underscoring the necessity to employ this methodology [36]. Through iterative refinement based on unseen data analysis, models can improve their performance and instill greater confidence in their practical use [35].

Accurate disease classification is pursued by examining individual stems, each consisting of a single pseudostem and associated leaves. This approach is pertinent as banana plants typically comprise a corm mat with multiple stems. An object-based Deep Learning (DL) pipeline is developed, utilizing RGB, multispectral, and pansharpened images to classify BXW (Banana Xanthomonas Wilt) and healthy stems. The model is subsequently evaluated using unseen ground-truth data. This proposed research will address a significant research gap by offering a comprehensive approach to disease classification in the mixed-complex agricultural landscapes of Africa using advanced DL techniques.

The contributions of this work is to:

1. Curate a dataset of expert-annotated aerial images depicting banana Xanthomonas wilt (BXW) from an African hotspot d. This dataset is valuable for training and testing DL models, benefiting researchers working on similar problems.
2. Overcome the limitation of insufficient samples, data augmentation techniques were employed to enhance the network's detection and identification capabilities. By increasing the number of training images through augmentation, the robustness of the DL model is improved.
3. Developed an enhanced object detection model that accurately detects and identifies BXW diseases in complex African landscapes using low-cost UAV imagery (RGB, MS, and pansharpening). The model combines the Faster R-CNN algorithm with a transfer learning ResNet-50 DL architecture, ensuring reliable performance. Parallel to this, Yolo-V8 also provides adequate performance under specific scenarios.
4. The performance of the Faster R-CNN and Yolo-V8 model was validated using images collected from two types of sensors (RGB and multispectral) and at two different heights (100 and 60 m). This validation study represents a pioneering effort to assess disease detection accuracy in African mixed complex systems using previously unseen data.
5. The proposed methodology lays the foundation for an early warning system and an automated disease management system for BXW. This approach facilitates the implementation of the Tumaini AI-powered mobile app database, enabling proactive measures for disease control and management.

In conclusion, this research aims to develop a BXW disease detection system that enhances efficient and effective disease management at the farm level.

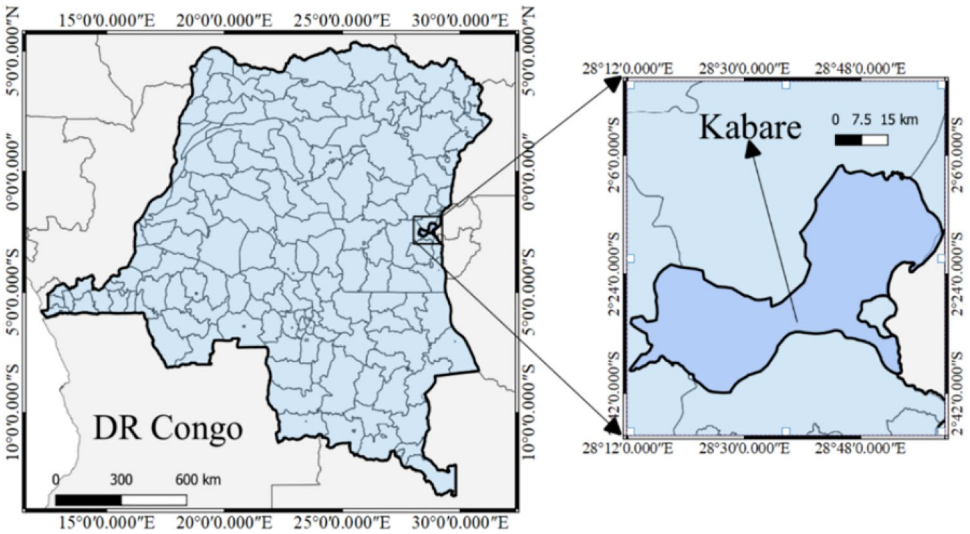

Fig. 1 Location of the study region, Kabare district of Eastern Democratic Republic of Congo

2 Materials and methods

2.1 Study area

The study was conducted in the northeastern Kabare district region, as shown in Fig. 1, in the South Kivu province of the Eastern Democratic Republic of Congo (DR Congo). The research spanned from 2019 to 2022 and covered 21 small banana plots. The region experiences an average annual rainfall of 1656 ± 235 mm (2015–2018), with two distinct rainy seasons from February to May and September to December. The mean annual temperature in this region is 18 °C. Each of the 24 plots exhibits a small-scale subsistence agriculture system, featuring mixed-complex landscapes of banana plants, shrubs, trees, annual food crops, pastures, buildings, and roads. Supplementary Table 1 contains information regarding each location's date, altitude, latitude, and longitude.

2.2 Ground-truth data (GTD)

BXW-affected landscapes were selected based on visible symptoms such as yellowing and wilting leaves, premature bunch ripening, and yellow ooze in pseudostems and leaf petioles. To obtain ground-truth data (GTD) on the disease, field surveys were conducted across the 21 banana fields (2019–2022). The presence and incidence of the disease were assessed, as shown in Fig. 2, focusing on a single unhealthy stem at each sample point. Based on their external characteristics and symptoms, the stems were categorized as healthy or diseased (with BXW presence). These annotations were represented in a point shapefile using QGIS software.

2.3 UAV data

2.3.1 UAV image collection

As shown in Fig. 3, a MicaSense-RedEdge MS camera and a DJI-FC6310 RGB camera attached to a DJI Phantom 4 Pro (P4P) drone were used to capture the necessary field images. The drone flights were planned through the Pix4D-capture app for automatic flight planning, ensuring smooth and efficient flights. This app plans the flight in the area based on coordinates. The application is free to use and designed for Pix4D (https://www.pix4d.com/es/producto/pix4dcapture/).

Fig. 2 Symptoms of Banana Xanthomonas Wilt (Unhealthy banana stems)

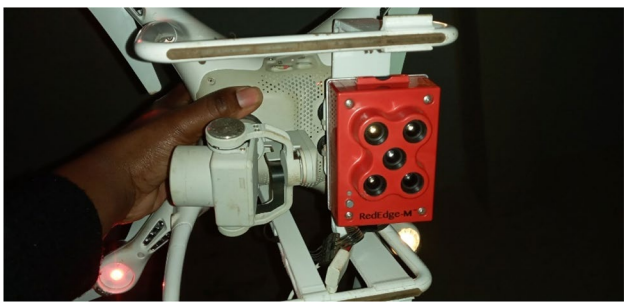

Fig. 3 MicaSense-RedEdge MS Camera mounted in P4P drone

Table 1 UAV camera sensors and flight characteristics

Camera	Sensor	Radiometric Resolution	Spectral Resolution—Bands	Flight height (m)	Spatial resolution (cm)
DJI-FC6310	RGB—Visible	8 bits	3 bands Red (R) Green (G) Blue (B)	60 100	1.36 2.59
MicaSense RedEdge	MS—Multispectral	16 bits	5 bands R, G, B Red Edge (RE) Near-infrared (NIR)	60 100	4.17 6.53

A constant speed of 2.5 m/s was maintained for all the different flight conditions, and the height above ground level was adjusted to 60 m and 100 m. For more information about each camera and the flight characteristics, please refer to Table 1.

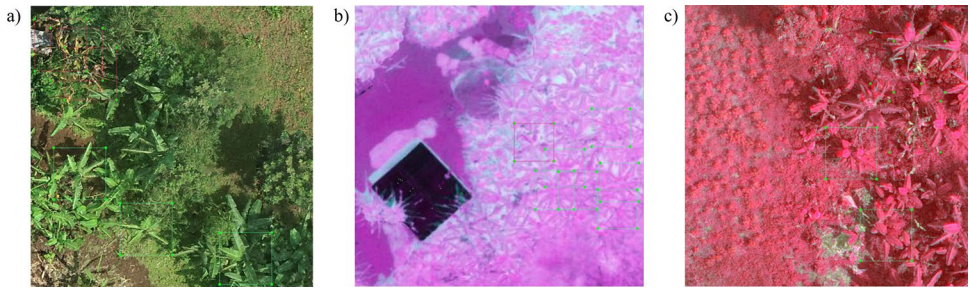

Fig. 4 Examples of Orthomosaics focused on banana fields: **a** RGB, **b** MS, and **c** PAN

Table 2 Description of each Vegetation indices

Name	Abbreviation	Formula
Normalized Difference vegetation index	NDVI	$\frac{NIR-RED}{NIR+RED}$
Normalized Difference Red Edge Index	NDRE	$\frac{NIR-REDEDGE}{NIR+REDEDGE}$
Chlorophyll Absorption Ratio Index	CARI	$\frac{REDEDGE}{GREEN-1}$
Normalized pigment chlorophyll ratio index	NPCI	$\frac{RED-BLUE}{RED+BLUE}$
Green–Red Vegetation Index	GRVI	$\frac{RED-GREEN}{RED+GREEN}$
Green Normalized Difference Vegetation Index	GNDVI	$\frac{NIR-GREEN}{NIR+GREEN}$
Blue Normalized Difference Vegetation Index	BNDVI	$\frac{NIR-BLUE}{NIR+BLUE}$

Table 3 Vegetation indices combination to form orthomosaics of three bands

Dataset	3-Band Image Combination
DT1	RGB (RED, GREEN, BLUE)
DT2	NDVI, NDRE, CARI
DT3	NDRE, NDVI, NPCI
DT4	GRVI, NDVI, GNDVI
DT5	NPCI, BNDVI, GRVI
DT6	NPCI, NDVI, BNDVI
DT7	NPCI, NDRE, BNDVI
DT8	PAN (NIR, RED, GREEN)

The image collection process faced various challenges due to the low drone flight altitude, undulating terrain, and tall trees like Ficus and Eucalyptus. These factors made it more difficult to accurately capture the images of banana stems. For a comprehensive understanding of the image locations and the capture dates, please see Supplementary Table 1.

2.3.2 Orthomosaic generation (RGB—MS and PAN)

The construction of RGB and MS orthomosaics for each flight is performed using Agisoft Professional Metashape software [37]. This involves converting numerous individual images into a unified georeferenced stitched image. For the PAN image, the software ArcGIS Pro was utilized. The process began with an initial georeferencing stage, where RGB and MS images were aligned using the same ground control points. This was followed by a fusion step utilizing the Hue-Intensity-Saturation (HIS) pansharpening method [38]. The objective of the PAN image was to blend the strengths of RGB and MS images, resulting in a fused image with enhanced spatial and spectral resolution (Fig. 4).

2.3.3 Band-based image and vegetation indices combinations as model inputs

To compare the accuracy of DL models in identifying the presence of BXW, an input dataset (DT) consisting of 3-band-based images was used. These images contain three bands of information, such as Red, Green, and Blue (RGB) or similar multi-band representations. For RGB orthomosaics, the visible bands of dataset 1 (DT1) were used. For MS images, seven vegetation indices (VIs) were computed using the Pheno-I protocol [10, 39] to enhance vegetation health (See Table 2). To create the 3-band images as an input, the VIs bands from DT2 to DT7 were combined, as detailed in Table 3. The last input dataset (DT8) is the PAN orthomosaic, which uses the 3-band combination NIR, R, and G. Next, the orthomosaics were split into smaller images using the GTD point shapefile. Finally, a polygon shapefile was created using a Python script to identify a boundary box of the healthy or unhealthy stem, converting them to XML files suitable for DL training (Fig. 5).

2.4 Object detection using deep learning

2.4.1 Data augmentation

A challenge arises in multiple agricultural deep-learning projects due to the limited instances representing the minor class, resulting in unbalanced datasets [40]. This imbalance implies an unequal data distribution across different classes when training the DL model. Specifically, most training data for infected crops contain more healthy crop instances than unhealthy ones, making it difficult to accurately identify the minority class using DL techniques [41].

In this context, the project employed data augmentation (DA) as an artificial alternative to increase elements in the minority class [42]. The project used six DA techniques that were combined and applied randomly to create more images of the minority class. The techniques used were:

(i) Image scale variation: Images are made to seem closer or farther away. (ii) Image rotation: Images are rotated to a random number of degrees between 0 and 60 to the left or right. (iii) Image translation: Images are moved along the horizontal and vertical axis. (iv) Image flipping: Images are flipped over their vertical axis, creating a mirror effect. (v) Pixel-brightness variation: Pixel brightness is multiplied by a random number, making the image darker or lighter. (vi) Noise addition: Random Gaussian noise or Gaussian blur is added to the image.

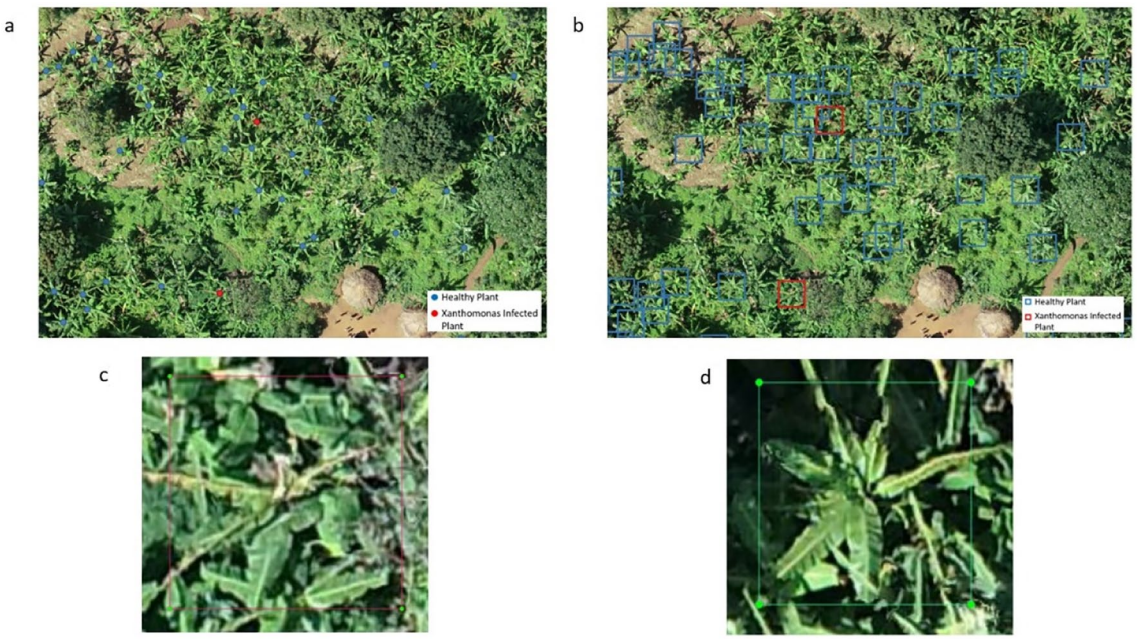

Fig. 5 Data labeling using QGIS software and banana disease detection classes: **a** Healthy and Sick stem labeling using vector points, **b** Polygon generation on each stem, **c** BXW infected banana stem **d** healthy banana stem

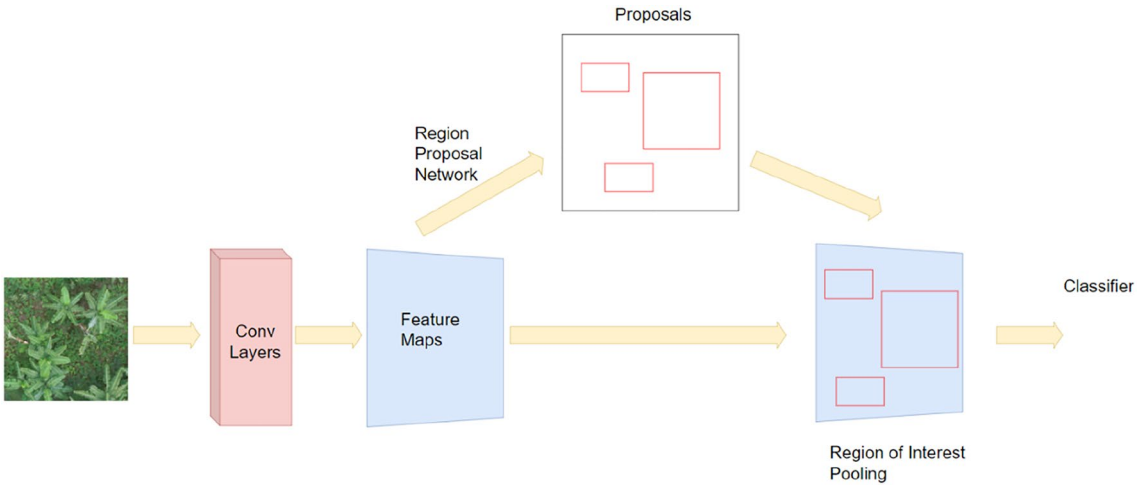

Fig. 6 Faster R-CNN architecture overview

2.4.2 Faster R-CNN overview

Faster R-CNN is an object-detection system that combines two modules to achieve accurate and efficient detection [43]. The first module is a deep convolutional network, acting as a region proposal network (RPN). In the Region Proposal Network, anchors characterized by predetermined ratios and scales are produced through a sliding window mechanism applied to the upper feature maps of a Convolutional Neural Network (CNN) [44]. This network learns anchor coordinate offsets and objectness scores (distinguishing foreground from background) directly from these feature maps. Subsequently, proposals are formulated by modifying these anchors with the learned coordinate offsets [44]. These proposals are then ranked based on foreground scores and subjected to a selection process through non-maximum suppression. This method effectively filters out redundant proposals, ensuring that only the most relevant proposals are retained for further processing [43].

The second module is the Fast R-CNN detector, which uses the proposed regions and their associated features for precise object classification and bounding box regression. Both modules share convolutional layers, reducing redundant computations and improving computational efficiency. The unified network is trained using alternating training or approximate joint-training approaches, where the RPN is initially trained, and its proposals are used to train the Fast R-CNN. Figure 6 displays a general overview of the architecture.

2.4.3 Yolo-V8 overview

Joseph Redmon introduced the YOLO (You Only Look Once) architecture [45, 46] in 2016, gaining prominence for its efficient feature fusion and high-accuracy detection capabilities. The latest iteration, YOLOv8, was unveiled in 2023. This architecture is structurally composed of three primary segments: the backbone, neck, and head, as seen in Fig. 7 [47].

The backbone of YOLOv8 resembles that of its predecessor, YOLOv5, employing a modified CSPDarknet53 backbone. However, it introduces the C2f module as a replacement for the CSPLayer used in YOLOv5 [46, 47]. YOLOv8 integrates a Spatial Pyramid Pooling Fast (SPPF) layer to enhance computational efficiency. This layer consolidates features into a uniform map, effectively circumventing issues like image distortion from scaling image regions and cropping. This improvement also notably augments the model's generalization capacity [48].

The neck component of YOLOv8 adopts the combined feature fusion approach of the Path Aggregation Network [46, 49] and the principles of the Feature Pyramid Network (FPN) [46, 50]. In a significant design evolution, the head module of YOLOv8 incorporates a mainstream Decoupled Head structure. This structure distinctly separates tasks related to target detection, splitting the classification and detection heads to independently extract target location and category information. By conducting fusion post-learning through different network branches, the model mitigates the additional latency typically associated with convolution in the decoupled head. This strategic design choice not only enhances the generalization capability but also bolsters the robustness of the model [46, 47].

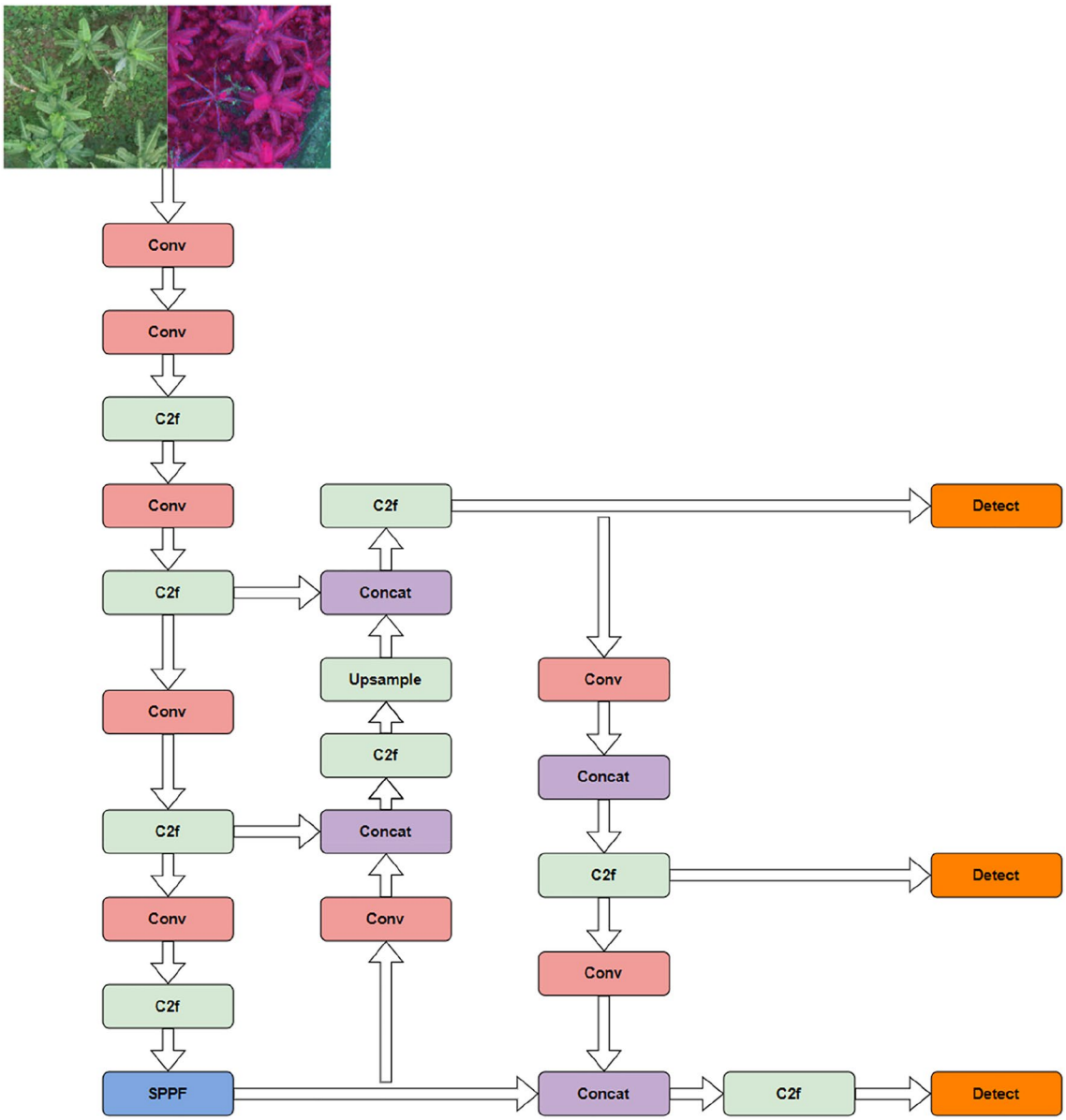

Fig. 7 Yolo-V8 architecture overview

2.4.4 Model training conditions

For model training, a transfer-learning approach was used, utilizing pre-trained models from the COCO (Common Objects in Context) dataset [51] through TensorFlow object-detection API zoo models with Falcon CV 2.0 [52] for the Faster R-CNN model and directly using Python and the ultralytics library for the Yolo-V8 model.

The model development process was executed on a remote server with the following specifications: NVIDIA Tesla M60 graphics card, Linux Ubuntu operative system, and Python 3.7 (see Table 4 for more details).

2.4.5 Model validation

The evaluation of the DL models involved assessing various performance metrics, including precision, recall, loss function, F1 score, and MissedGTError. The Intersection over Union (IoU) method was used to determine the accuracy of the model's predictions. The IoU measures the overlap between the predicted bounding box (B2) and the

Table 4 Hardware and software used in this study

Hardware and software	Specifications
Image acquisition hardware	
Quadcopter UAV	3DR Solo
Quadcopter UAV	Phantom 4 pro
RGB Camera	FC6310
Phantom 4 RGB camera	Sony Qx1
Multispectral camera	MicaSense RedEdge
Image processing software	
Orthomosaic Generation	Agisoft MetaShape
Vegetation Indices extraction	Pheno-i
DL Modeling hardware	
RAM Memory	128 GB
Processor	Intel Xeon E5-2667 v4 @ 3.20 GHz × 16
GPU	NVIDIA Tesla M60
DL Modeling software	
OS	Ubuntu Linux
Programming language	Python 3.7
Labelling software	QGIS
Deep learning libraries	FalconCV with TensorFlow 2.0 and TensorFlow object detection API

ground-truth bounding box (B1) using Eq. 1. For this evaluation, a threshold of 0.5 was set. The prediction is correct if the IoU is greater than 0.5 and the predicted label matches the ground-truth label.

$$IoU = \frac{B_1 \cap B_2}{B_1 \cup B_2} \tag{1}$$

In detection models, precision and recall are two important metrics used to evaluate model performance in identifying positive and negative instances (Healthy and BXW in this case).

Precision refers to the ratio between true-positive instances and the total number of instances the model identifies as positive. In other words, precision measures the model's accuracy in identifying positive instances [53]. Conversely, Recall measures the fraction of true-positive detections among all actual positive instances in the dataset [53]. The F1 score combines precision and recall and provides a more general metric of the model performance. A high F1 score indicates that a model has both high precision and high recall and, therefore, is a good overall performer [53].

$$precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \tag{2}$$

$$recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \tag{3}$$

$$f1\ score = \frac{2 \times precision \times recall}{recall + precision} \tag{4}$$

The loss function serves as a statistical representation of the training process and helps diagnose how well the model fits the data. We extracted the loss function data using [54, 55] and the TensorFlow's visualization toolkit (TensorBoard) to track and visualize metrics such as loss during the DL training process.

MissedGTError (MGTE) is a performance metric used to evaluate the accuracy of object detection models. It calculates the ratio of ground-truth images that were not detected and the total number of ground-truth images in the dataset [56]. While other metrics such as the confusion matrix, precision, and recall provide information on the accuracy of correctly detected ground-truth labels per class, MGTE focuses on the number of ground-truth instances that the model missed

detecting. This metric helps distinguish between two types of errors: misclassifications and missed detections. Misclassifications occur when an object is detected but assigned to the wrong class (e.g., identifying a healthy stem infected with BXW). On the other hand, missed detections happen when ground-truth objects are not detected at all, regardless of their class (healthy or infected). A 0% score indicates that all ground-truth objects were detected, while a score of 100% indicates that none were detected, representing the worst-case scenario.

$$MGTE = \frac{Non\ detected\ Annotations}{Total\ number\ of\ ground\ truth\ Annotations} 100\% \tag{5}$$

A workflow shows the methodology pipeline as a summary in Fig. 8.

3 Results

3.1 UAV dataset collection and preprocessing

Collecting reliable and accurate images from disease hotspots and precise data annotation is crucial in developing any DL detection model. This project created MS and RGB image datasets using UAV (Unmanned Aerial Vehicle) images collected in Eastern Congo. The RGB dataset comprises approximately 16,195 annotations, with 86% healthy stem annotations and 14% sick annotations. The MS distribution is roughly 75% healthy and 25% sick stems. Both datasets are imbalanced, which is common in agriculture due to the smaller number of BXW-infected stems (minority class) compared to healthy stems in the field. This imbalance can sometimes lead to difficulties in identifying the minority class in DL models [41]. Data augmentation techniques are employed to address this and improve model performance to increase the number of instances in the minority class [42].

Among the available six augmentation techniques, two were randomly selected and applied to each annotation of BXW-infected stems as seen in Fig. 9. This procedure generated 2281, 2804, and 1809 new annotations for sick stems in the RGB, MS, and PAN datasets, respectively. As a result, the distribution of healthy to sick stem annotations in the RGB dataset changed from 86% healthy and 14% sick to 75% healthy and 25% sick. In contrast, both the MS and PAN datasets witnessed an increase in BXW-infected stem annotations from 25 to 35% of the entire dataset. The distributions of the original and new datasets are presented in Table 5 for visual comparison.

3.2 Model training

The transfer learning model training was conducted over 100,000 steps for the Faster R-CNN model and 200 steps for the Yolo-V8 model. The data was divided into an 80% training and 20% training validation split. Table 5 provides the exact number of annotations for each dataset.

An additional dataset of unseen stem annotations was used to evaluate the model's performance. This dataset included annotations from images taken at 100 m and 60 m heights. For the 100 m height images, there were 2026 RGB annotations (1546 Healthy–480 BXW), 3,310 MS annotations (2485 Healthy–825 BXW), and 890 PAN annotations (581 Healthy–309 BXW). Similarly, for the 60 m height images, there were 1,581 annotations (1084 Healthy–497 BXW) for both RGB and MS. Additionally, for 60 m PAN, 1179 annotations (815 Healthy–364 BXW) were considered. These unseen annotations provided a comprehensive test dataset to evaluate the model's performance.

3.3 DI-based detection models for healthy and BXW-infected banana stems

Eight distinct models were created using the Faster R-CNN ResNet-50 architecture and another eight using the Yolo-V8 architecture, all aimed at distinguishing between healthy and BXW-infected banana stems within complex mixed systems. These models include one RGB model (DT1), one PAN model (DT8), and six MS models (DT2 to DT7). During the training phase, a portion of the data, specifically 20%, was used for validation to enable the model to assess its learning progress. This validation process involves various loss functions to evaluate the model's performance. These functions include classification loss and localization loss, which measure the model's confidence in making accurate predictions and its ability to accurately determine object positions within the input data [57]. A comprehensive total loss function was derived by

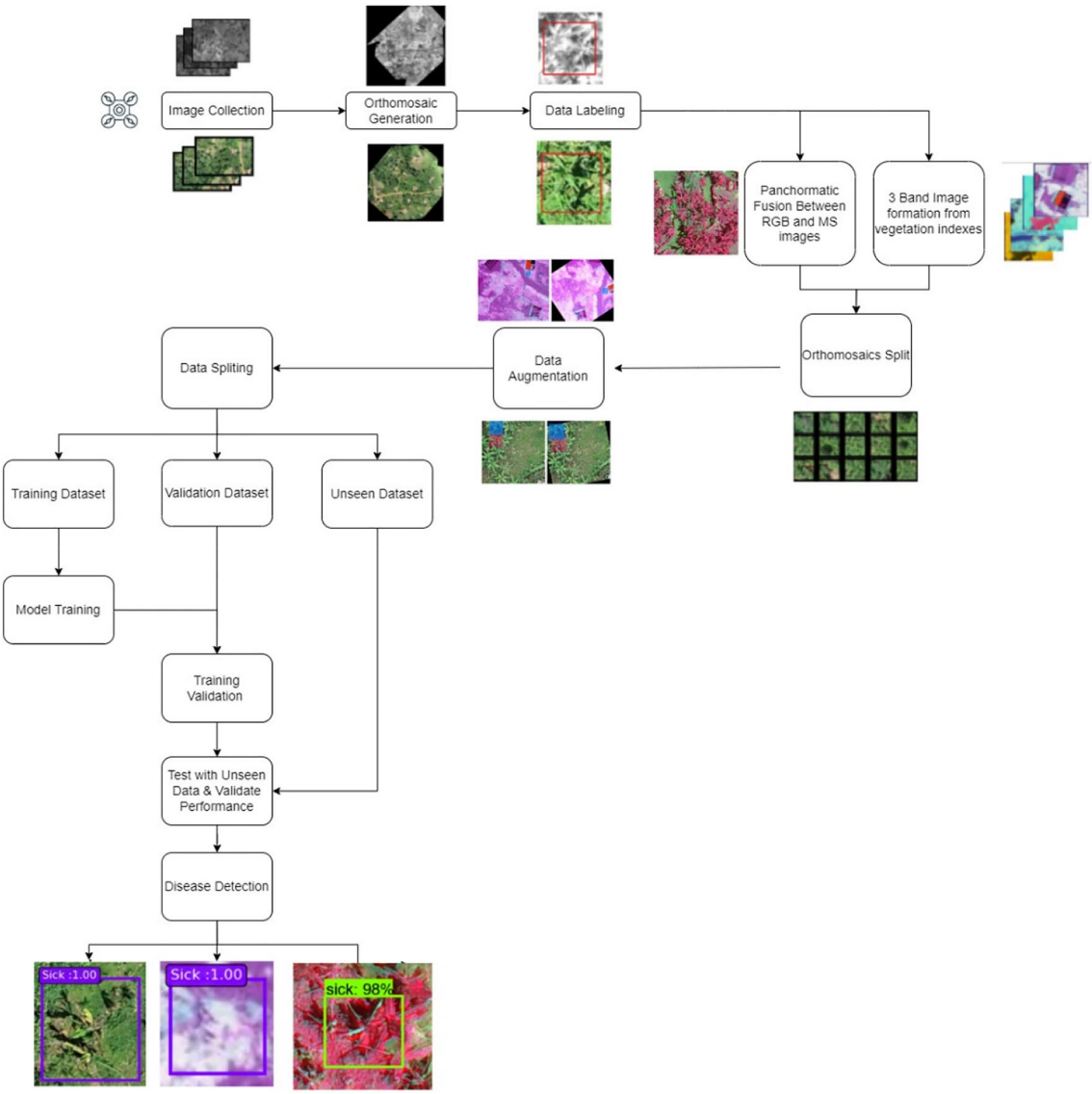

Fig. 8 Methodology pipeline

combining the classification and localization losses, offering a broader perspective of the training process and helping monitor the model's overall performance. As shown in Fig. 10, the loss curves of the models during training consistently decrease with each step. This suggests that the models can benefit from further training epochs, enhancing performance.

3.4 Performance metrics for dl model testing with unseen data at 100 m

After developing and validating the DL models with suitable datasets, the models were tested using real-world images. These images were selected carefully to ensure they differed from those used in training, effectively representing unfamiliar or "unseen" data for the models. This testing phase allowed us to assess how well the models perform and generalize under real-world conditions, providing valuable insights beyond the training context.

Tables 6 and 7 detail the MissedGTError (MGTE), precision, recall, and F1 score values for the 100 m data for the Faster R-CNN and Yolo-V8 architectures, respectively. Among the various models for Faster R-CNN seen in Table 6, the RGB model showcases the best performance, with an MGTE of 11.2%, followed by PAN with 18.7% and DT3 with 20.33%. However, DT7 and DT5 exhibited higher MGTE rates at 33 and 29%, respectively. These rates imply that, on average, approximately 1 out of every 10 annotations in the RGB model goes undetected, while DT5 and DT7 fail to detect approximately 1 out of every 3 detections. Additionally, the PAN model fails to detect 1 out of every 5 annotations. Figure 11 illustrates the

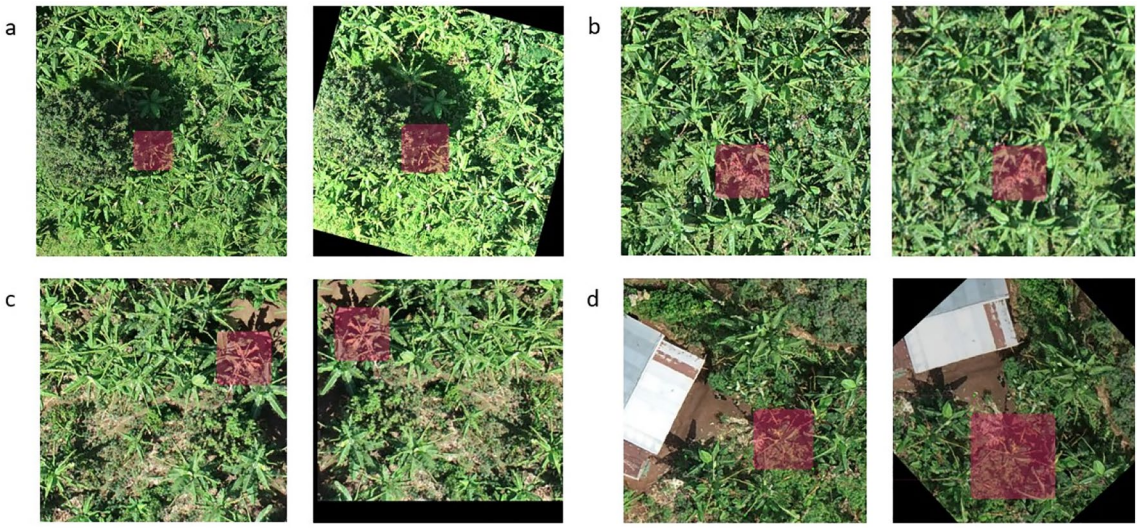

Fig. 9 Data Augmentation on RGB orthomosaic split images: (**a**) Original image (Left), rotation and augmented brightness (Right), (**b**) Original image (Left), mirroring and noise addition (Right), (**c**) original image (Left), mirroring and translation (Right), (**d**) Original image (Left), image rotation and decreased pixel brightness (Right)

Table 5 Dataset class distributions before and after data augmentations

Label	Vegetation indexes	Before data augmentation			After data augmentation		
		% Healthy annotations	% Sick annotations	Total annotations	% Healthy annotations	% Sick annotations	Total annotations
DT1	RGB	86	14	16195	75	25	18476
DT2	NDVI, NDRE, CARI	78	22	13766	65	35	16570
DT3	NDRE, NDVI, NPCI	78	22	13766	65	35	16570
DT4	GRVI, NDVI, GNDVI	78	22	13766	65	35	16570
DT5	NPCI, BNDVI, GRVI	78	22	13766	65	35	16570
DT6	NPCI, NDVI, BNDVI	78	22	13766	65	35	16570
DT7	NPCI, NDRE, BNDVI	78	22	13766	65	35	16570
DT8	PAN (NIR, RED, GREEN)	78	22	8283	64	36	10092

confusion matrix for the unseen data. It reveals a significant count of false positives for the MS models, indicating that they identify many healthy stems as having BXW symptoms. This suboptimal performance of the MS models is also reflected in the metrics shown in Table 6. For instance, F1 scores for DT5 and DT7 vary from 0.40 to 0.31 for healthy stems and 0.04 to 0.06 for BXW-infected stems. Similar patterns emerge for the other MS models. On the other hand, the RGB and PAN models (DT1 and DT8) demonstrated good performance in detecting healthy and BXW-infected stems. Their F1 scores are 0.88 and 0.91 for healthy detections and 0.74 and 0.82 for BXW annotations, respectively. This indicates that, overall, these models successfully detect each stem corresponding to its ground truth annotation.

For the Yolo-V8 models shown in Table 7, the MGTE of DT1 and DT8 present lower values than the Faster R-CNN model, with values of 7.8% and 9.8%, respectively, while all the other MS models present values from 49 to 79% showcasing difficulties to identify both healthy and sick stems for the MS images. Regarding precision, recall, and F1-Score, most metrics present high values, mostly over 0.9, except for the DT6 recall. These results confirm Yolo-V8 as the best-performing model for the 100 m RGB and PAN datasets.

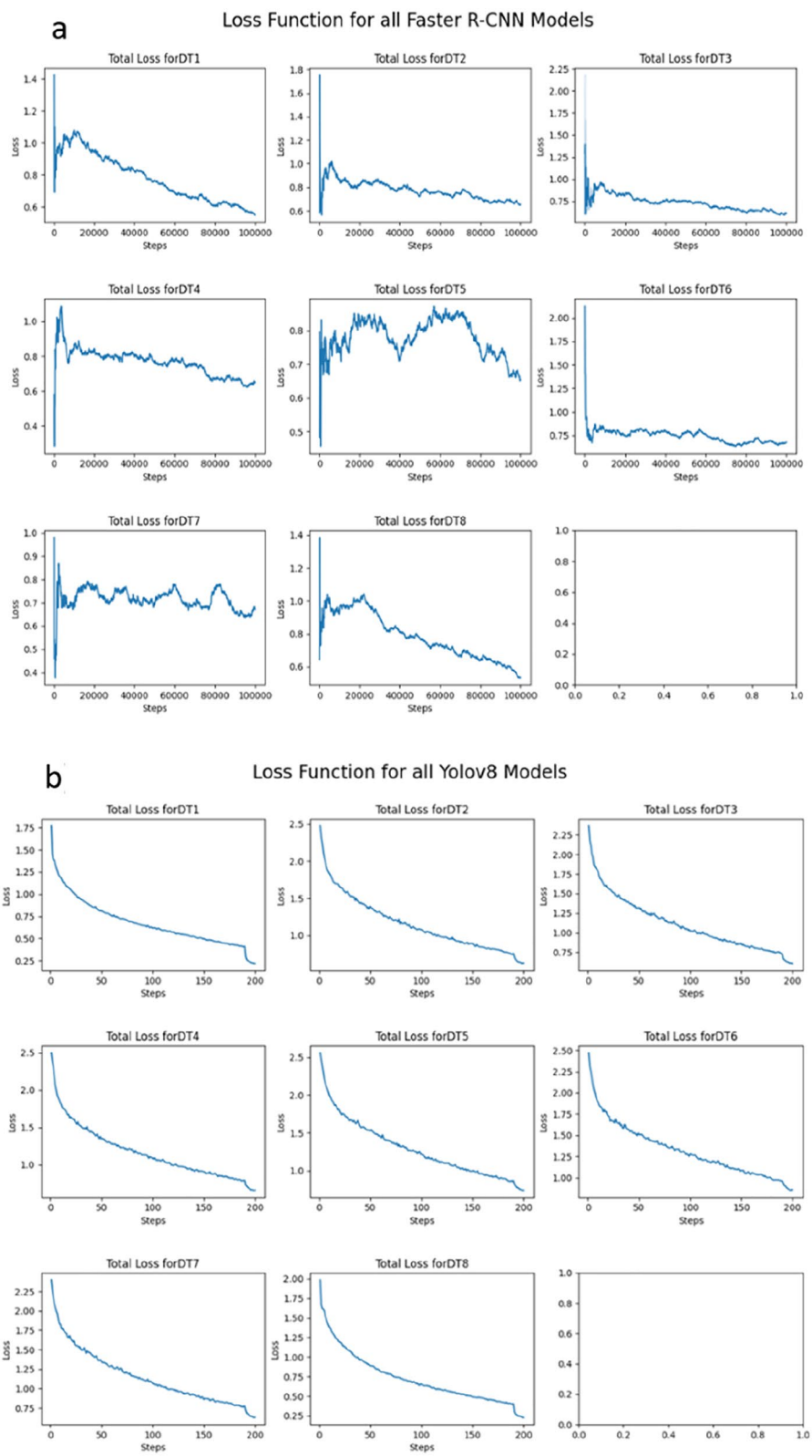

Fig. 10 Total loss function for RGB (DT1), MS (DT2-DT7), and PAN models (DT8) for (**a**) Faster R-CNN, (**b**) Yolo-V8

Table 6 Undetected Instances by FasterRCNN model from unseen 100 m Data

Label	Vegetation indexes	Undetected BXW annotations	Undetected healthy annotations	Total number of annotations undetected	Total number of annotations	MGTE (%)	Healthy precision	Healthy recall	Healthy F1-score	Sick precision	Sick recall	Sick F1-score
DT1	RGB	63	165	228	2026	11.2	0.917	0.847	0.881	0.686	0.813	0.744
DT2	NDVI, NDRE,CARI	351	362	713	3310	21.5	0.525	0.174	0.261	0.085	0.33	0.136
DT3	NDRE, NDVI, NPCI	310	363	673	3310	20.33	0.559	0.183	0.275	0.105	0.4	0.167
DT4	GRVI, NDVI, GNDVI	340	389	729	3310	22.0	0.6	0.269	0.372	0.64	0.217	0.099
DT5	NPCI, BNDVI, GRVI	399	561	960	3310	29.0	0.604	0.302	0.403	0.026	0.085	0.04
DT6	NPCI, NDVI, BNDVI	335	469	804	3310	24.2	0.504	0.149	0.23	0.1	0.391	0.159
DT7	NPCI, NDRE, BNDVI	436	658	1094	3310	33.0	0.548	0.219	0.313	0.038	0.146	0.061
DT8	PAN(NIR, RED, GREEN)	27	141	168	890	18.7	0.976	0.86	0.915	0.731	0.948	0.825

Table 7 Undetected Instances by Yolo-V8 model from unseen 100 m Data

Label	Vegetation indexes	Undetected BXW annotations	Undetected healthy annotations	Total number of annotations undetected	Total number of annotations	MGTE (%)	Healthy precision	Healthy recall	Healthy F1-score	Sick precision	Sick recall	Sick F1-score
DT1	RGB	89	69	158	2026	7.80	0.991	0.968	0.98	0.93	0.992	0.96
DT2	NDVI, NDRE,CARI	393	1430	1823	3310	55.07	0.966	0.95	0.958	0.93	0.99	0.96
DT3	NDRE, NDVI, NPCI	335	1875	2210	3310	66.76	0.951	0.927	0.939	0.93	0.992	0.96
DT4	GRVI, NDVI, GNDVI	389	1707	2096	3310	63.30	0.955	0.968	0.962	0.93	0.992	0.96
DT5	NPCI, BNDVI, GRVI	325	2024	2349	3310	70.1	0.953	0.946	0.949	0.93	0.992	0.96
DT6	NPCI, NDVI, BNDVI	558	2082	2640	3310	79.70	0.688	1	0.815	0.93	0.992	0.96
DT7	NPCI, NDRE, BNDVI	340	1300	1640	3310	49.0	0.968	0.95	0.959	0.93	0.992	0.96
DT8	PAN(NIR, RED, GREEN)	50	37	87	890	9.8	0.995	0.958	0.976	0.93	0.992	0.96

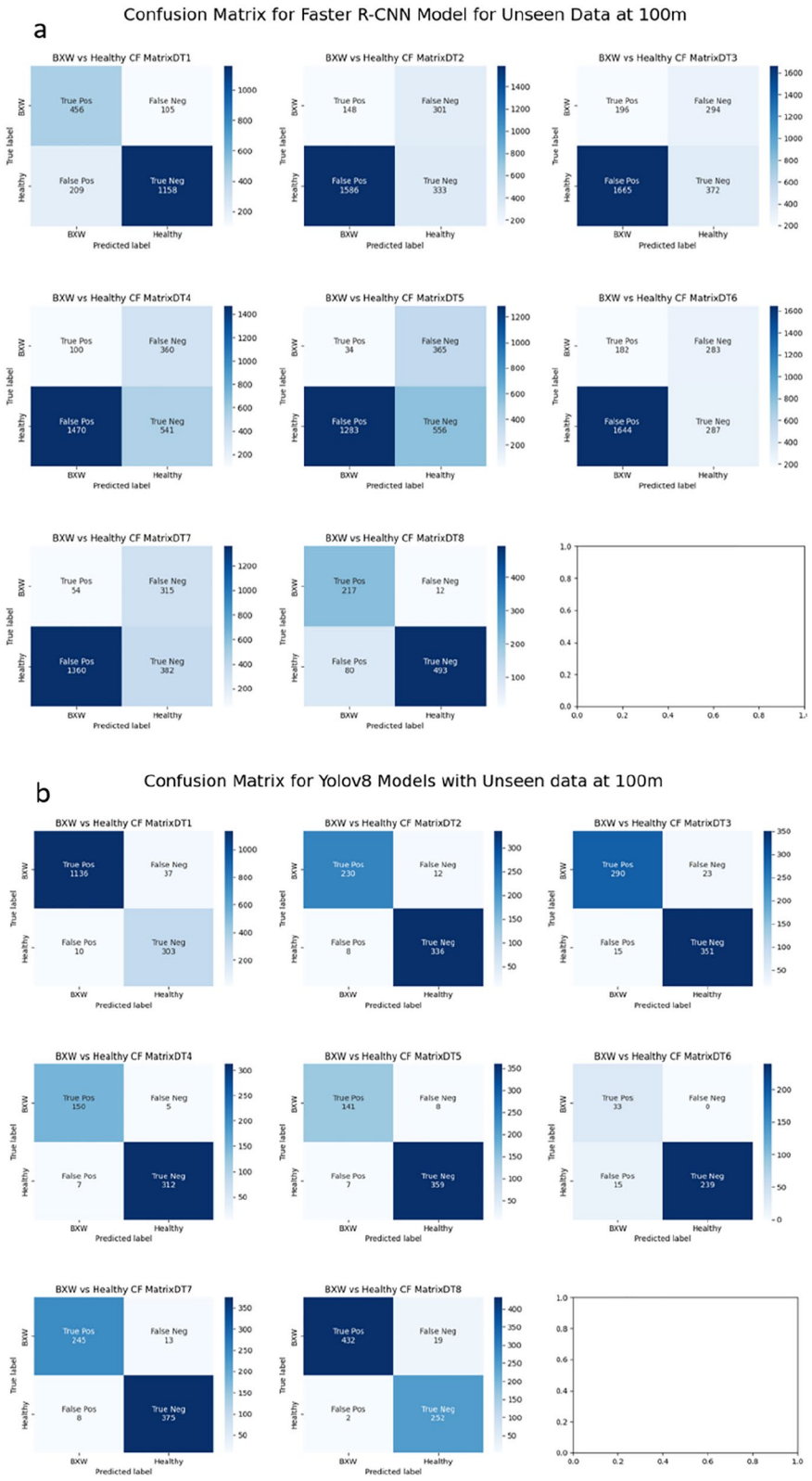

Fig. 11 Confusion Matrix for RGB (DT1), MS (DT2-DT7), and PAN models (DT8) with unseen 100 m data for (**a**) Faster R-CNN, (**b**) Yolo-V8

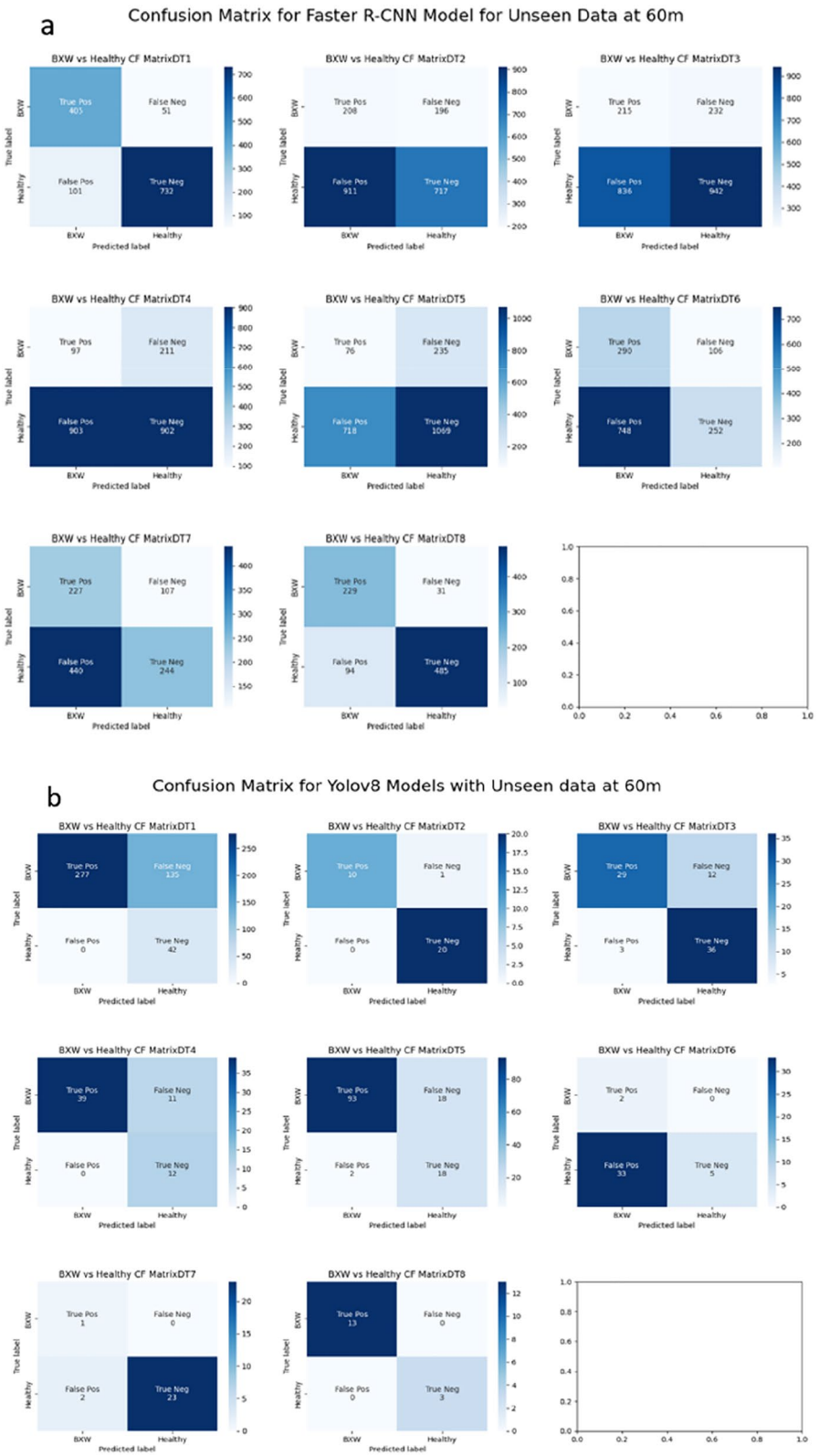

Fig. 12 Confusion Matrix for RGB (DT1), MS (DT2-DT7), and PAN models (DT8) with unseen 60 m data for (**a**) Faster R-CNN, (**b**) Yolo-V8

Table 8 Undetected Instances by FasterRCNN model from unseen 60 m Data

Label	Vegetation indexes	Undetected BXW annotations	Undetected healthy annotations	Total number of annotations undetected	Total Number of annotations	MGTE (%)	Healthy precision	Healthy recall	Healthy F1-score	Sick precision	Sick recall	Sick F1-score
DT1	RGB	89	69	158	2026	13.5	0.935	0.879	0.906	0.8	0.888	0.842
DT2	NDVI, NDRE,CARI	393	1430	1823	3310	15.8	0.785	0.44	0.564	0.186	0.515	0.273
DT3	NDRE, NDVI, NPCI	335	1875	2210	3310	17.0	0.802	0.53	0.638	0.205	0.481	0.287
DT4	GRVI, NDVI, GNDVI	389	1707	2096	3310	17.5	0.81	0.5	0.618	0.097	0.315	0.148
DT5	NPCI, BNDVI, GRVI	325	2024	2349	3310	18.8	0.82	0.598	0.692	0.096	0.244	0.138
DT6	NPCI, NDVI, BNDVI	558	2082	2640	3310	16.1	0.704	0.252	0.371	0.279	0.732	0.404
DT7	NPCI, NDRE, BNDVI	340	1300	1640	3310	18.6	0.695	0.357	0.471	0.34	0.68	0.454
DT8	PAN(NIR, RED, GREEN)	50	37	87	890	16.7	0.94	0.838	0.886	0.709	0.881	0.786

3.5 Performance metrics for Dl model testing with unseen data at 60 m height

The images collected by drone at 60 m altitude were used to validate the models previously trained and assess the impact of the newly attained resolution. Figure 12 and Table 8 present this scenario's confusion matrix and performance metrics. When considering precision, recall, and F1 score, the models exhibit improved performance for the 60 m images compared with the 100 m images for Faster R-CNN. This improvement applies to the MS and RGB models, while the PAN metrics slightly decrease. In the case of DT1, the F1 score experiences a slight boost for healthy stem detection and witnesses an increase of nearly 10% for sick stem F1 score. The MS models show significant improvement in F1 score for detecting healthy stems, where recall and precision values range from 0.5 to 0.8. PAN achieves an F1 score between 0.78 and 0.88 for BXW and healthy stems, respectively.

When assessing undetected instances for the Faster R-CNN models in the 60 m images, as shown in Table 8, the RGB and PAN models exhibit similar MGTE values for 100 and 60 m data. In the case of the MS models, there is an improvement in MGTE, particularly noticeable for DT5 and DT7, where the non-detection rate decreased by around 12% each. Similarly, DT2, DT3, and DT4 demonstrate lower MGTE, indicating that these models successfully detect more MS images.

Contrary to the improvement from Faster-RCNN, Yolo-V8 models perform poorly with all the 60 m images. As seen in Table 9, the lowest MGTE is DT1, which has increased from 7.8 to 52.2%; the other MGTE values for the rest of the models are between 89 and 99%, showing that the model is not detecting most images. The rest of the metrics present high values, especially the precision, recall, and f1 score for the sick stems, which are all perfect. On the other hand, the Healthy stem F1 score ranges from 0.5 to 1, depending on the model. However, due to the high number of missed detections for models DT2 to DT7, the detection metrics are irrelevant since these models cannot detect the vast majority of the ground truth images.

4 Discussion

4.1 BXW detection using UAV remote sensing & DL models

Banana cultivation, particularly in Africa, confronts significant challenges due to pests and diseases, notably the destructive Xanthomonas Wilt of bananas (BXW). Timely and accurate disease detection is crucial for minimizing crop losses. Traditional methods relying on visual evaluations and manual indicators lack consistency and rigor. This study captures images of banana stems affected by BXW using UAV-mounted RGB and MS cameras within mixed-complex production landscapes.

As a result, models based on Faster R-CNN and YOLO-V8 were trained to distinguish between healthy and infected stems using the collected UAV images. Testing these models with real-world unseen data yielded promising results. Notably, for detecting healthy stem at 60 m height, specific multispectral VI combinations, such as DT3 (NDRE, NDVI, NPCI), DT4 (GRVI, NDVI, GNDVI), and DT5 (NPCI, BNDVI, GRVI), showed better performance. These combinations focus on chlorophyll content and its variations, important indicators of stem health [58–60]. However, the MS models exhibited limited effectiveness in detecting BXW-infected stems, attributed to inadequate spatial resolution. Conversely, RGB and PAN models better identified healthy and BXW-infected stems. The RGB model detected annotated stems at a rate of 86.5%, while PAN achieved 84.3%. The challenges of incomplete field labeling (Fig. 13) and stem density variations in mixed complex landscapes (Fig. 14) impacted model training and detection accuracy.

Various DL models have encountered challenges with partially labeled images, especially when dealing with images that contain multiple objects, making annotation a difficult and time-consuming task. This challenge is particularly prominent in microscopical and pathological datasets [61, 62]. Nonetheless, model calibration techniques like positive-unlabeled learning [62] and probability-based image similarity methods [63], have shown promising results in improving DL-model performance. Additionally, adopting strategies like complete labeling and considering the characteristics of complex landscapes can further enhance model performance in such scenarios.

The performance of Faster R-CNN models in this project, utilizing PAN and RGB, has demonstrated accurate detection of both healthy and diseased stems (Fig. 15). This observation aligns with previous studies that have highlighted the effectiveness of Faster R-CNN in accurately detecting diseases across various crops. For example, in a study by Zhang et al. [64], a similar approach addressed the challenge of limited datasets for soybean leaf-disease detection by creating synthetic-image datasets, achieving an average mean accuracy of 83.34% through a fusion of multiple Faster R-CNN

Table 9 Undetected Instances by Yolo-V8 model from unseen 60 m Data

Label	Vegetation indexes	Undetected BXW annotations	Undetected healthy annotations	Total number of annotations undetected	Total number of annotations	MGTE (%)	Healthy precision	Healthy recall	Healthy F1-score	Sick precision	Sick recall	Sick F1-score
DT1	RGB	89	69	158	2026	52.2	1	0.672	0.804	1	1	1
DT2	NDVI, NDRE, CARI	393	1430	1823	3310	97.6	1	0.909	0.952	1	1	1
DT3	NDRE, NDVI, NPCI	335	1875	2210	3310	89.3	0.906	0.707	0.795	1	1	1
DT4	GRVI, NDVI, GNDVI	389	1707	2096	3310	92.5	1	0.78	0.876	1	1	1
DT5	NPCI, BNDVI, GRVI	325	2024	2349	3310	90.3	0.979	0.838	0.903	1	1	1
DT6	NPCI, NDVI, BNDVI	558	2082	2640	3310	95.8	0.057	1	0.108	1	1	1
DT7	NPCI, NDRE, BNDVI	340	1300	1640	3310	97.7	0.333	1	0.5	1	1	1
DT8	PAN(NIR, RED, GREEN)	50	37	87	890	98.6	1	1	1	1	1	1

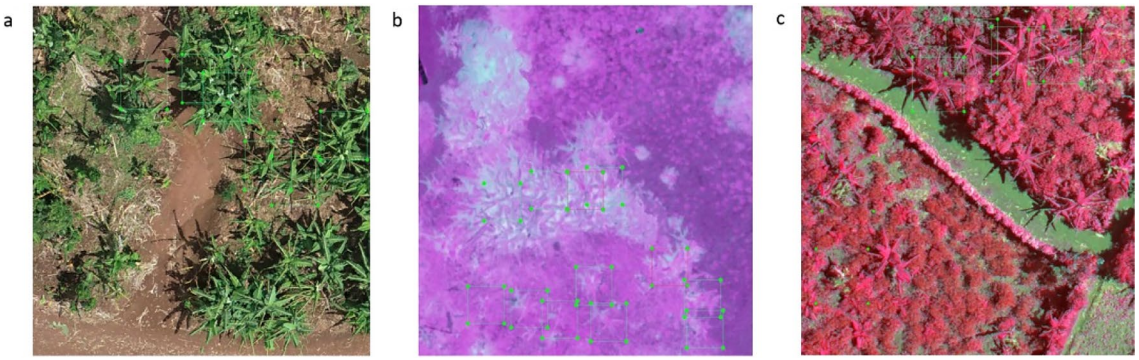

Fig. 13 Partially Annotated Banana Fields: **a** RGB, **b** multi-spectral, **c** PAN

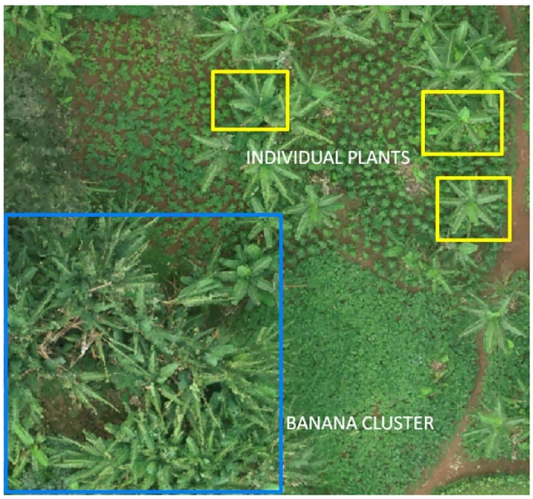

Fig. 14 Banana Plantation Density in Mixed Complex Landscapes

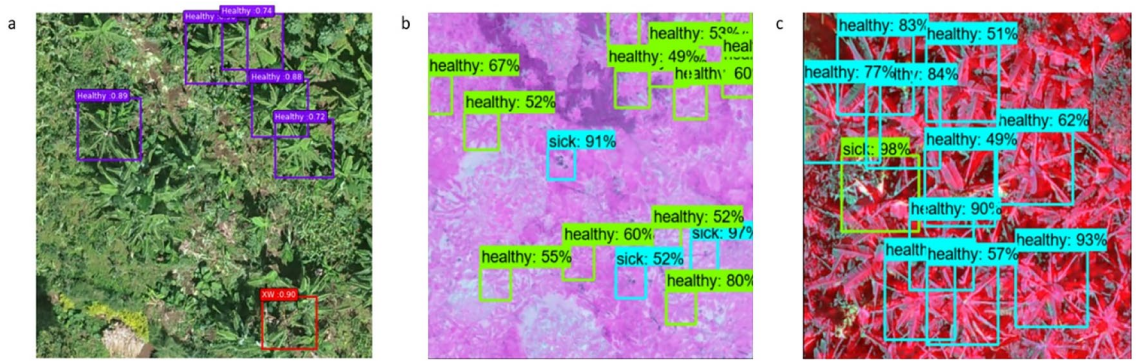

Fig. 15 Banana BXW detection using UAV-DL based model: **a** RGB, **b** multi-spectral, **c** PAN

functions. In another study by Parvathi and Selvi [65], proposed an improved Faster R-CNN algorithm with the ResNet-50 network for detecting maturity stages in coconuts, outperforming other detection methods. The diagnosis of rice leaf diseases has also been accomplished using the Faster R-CNN framework, showcasing high accuracy (ranging from 98 to 99%) in identifying common rice leaf diseases and healthy leaves [66]. Furthermore, Ghoury utilized transfer learning with two pre-trained DL models, namely Single Shot MultiBox Detector (SSD) with MobileNet v1 and Faster R-CNN with Inception v2, to differentiate between healthy and diseased grapes and grape leaves [67]. The Faster R-CNN Inception

v2 model achieved a classification accuracy of 95.57% for all testing images, surpassing the SSD_MobileNet v1 model's accuracy of 59.29% despite having a longer processing time.

Faster R-CNN, adopts a distinct two-step approach as a two-stage detection algorithm, compared to the one-stage detectors. Regions of interest (ROIs) are first identified in the two-stage framework using selective search algorithms. These regions are then processed by a DL model to generate bounding box predictions for objects. In contrast, one-stage detectors predict bounding boxes directly in a single pass of the DL model. Recent studies in remote sensing have primarily focused on using one-stage detectors, such as YOLO, RetinaNet, and CenterNet [19]. Nonetheless, there are still instances where two-stage detectors, such as Faster R-CNN, still have their applications [43, 68]. The choice between detectors depends on the specific needs of the task and the desired balance between accuracy and efficiency. It is crucial to recognize that both two-stage and one-stage detectors possess unique strengths and limitations.

The results obtained in this research exhibit both scenarios in which Faster R-CNN and Yolo-V8 perform well. Yolo-V8 tends to detect healthy and sick stems better than Faster R-CNN in scenarios with high spatial resolution, and the images are very similar to those used for training (100 m height). On the other hand, Faster R-CNN tends to perform better with MS images and is more robust when facing RGB and PAN images at different heights. This differential performance between the two models can be attributed to the inherent design and processing characteristics of one-stage and two-stage detectors.

Although one-stage detectors like Yolo-V8 are known for their efficient design, faster speeds, and reasonably good accuracy [69], they have faced accuracy challenges compared to two-stage detectors like Faster R-CNN. Two-stage detectors, by design, operate on a smaller number of proposals, allowing for a larger head network dedicated to proposal classification and regression. This setup facilitates more comprehensive feature extraction. Moreover, two-stage detectors utilize the RoIAlign operation to ensure high-quality feature extraction for each proposal and maintain location consistency [70]. In contrast, one-stage detectors may encounter feature misalignment issues when different region proposals share the same feature or high feature variations due to their coarse and spatially implicit representation of proposals [70]. Additionally, two-stage detectors perform object location regression twice, once in each stage, leading to refined bounding boxes and improved precision in object localization [70]. As observed in our study, these advantages collectively contribute to the superior accuracy of two-stage detectors across various object detection tasks.

The detection of healthy and BXW-infected banana stems in this project benefits from the accuracy provided by a two-stage model such as Faster R-CNN and the performance of a one-stage model under specific conditions. Nevertheless, differentiating BXW from other banana diseases or biotic stresses with similar symptoms remains challenging for model-based identification. Enhancements in MS models can aid in identifying spectral differences between BXW and other conditions. Additionally, MS models can help distinguish BXW from yellowing caused by abiotic stresses like heat, drought, and nutrition disorders. Addressing these challenges requires advancements in data collection, annotation techniques, and spectral analysis. Overcoming these obstacles will improve model performance, enable effective disease management strategies, and boost banana production in mixed-complex African systems. To ensure greater accuracy in specific disease detection, such as BXW, models should be validated using more targeted diagnostic methods [69, 70].

4.2 DI model performance with higher-resolution images

The performance of the MS models in detecting healthy and infected stems is less effective when using images captured at 100 m altitude for the Faster R-CNN model. This could be attributed to the limited spatial resolution of the images resulting from high-altitude flights and the specific sensor used [71]. A combination of physical adjustments and post-processing techniques can be employed to improve detection performance. Lowering the flight altitude is an effective way to enhance spatial resolution. Operating the flights at lower altitudes allow for closer proximity to the target, resulting in clearer and more detailed images [72]. In this study, flights were conducted at 60 m altitude, and the resulting higher-resolution images showed improved performance metrics for the MS Faster R-CNN models compared to the 100 m images.

Furthermore, panchromatic fusion techniques by integrating RGB and MS sensor data can provide an additional boost to spatial resolution. Panchromatic fusion combines high-resolution panchromatic imagery from the RGB sensor with multispectral data from the MS sensor, resulting in improved spatial details and overall image resolution. The incorporation of PAN images in training a new model (DT8) demonstrated good detection performance for both healthy and BXW-infected stems, surpassing the performance of RGB model at 100 m and maintaining good metrics at 60 m for the two-stage model. PAN sharpening techniques contribute to the improved detection of both healthy and diseased stems in image analysis and DL applications, offering more accurate insights for informed decision-making in agriculture.

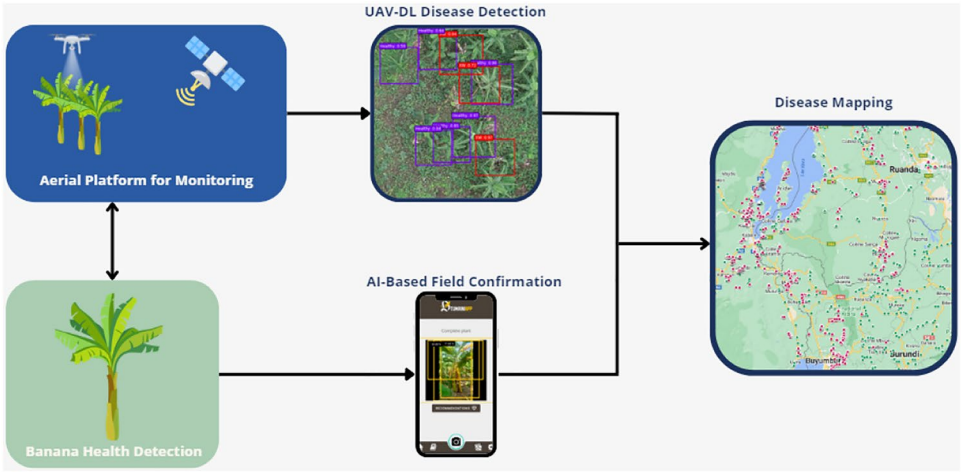

Fig. 16 Multiplatform Early Warning Alert System for Banana Production

While higher-resolution images offer advantages in BXW detection, there are associated tradeoffs to carefully consider. Lower altitude flights, although improving spatial resolution, may result in longer flight times and increased processing duration due to the larger number of data points collected [72]. Additionally, post-processing techniques like PAN sharpening require higher computational power. Therefore, the optimal orthomosaic resolution should be determined by a balanced consideration of flight and processing costs. Our study underscores that improving spatial resolution through strategies like lower altitude flights and panchromatic fusion techniques significantly enhances the performance of DL models in detecting healthy and infected stems. These advancements directly contribute to more precise identification and classification of various land-cover types, vegetation indices, and specific stem traits, ultimately supporting more effective decision-making in agricultural practices.

4.3 BXW surveillance system in Africa

Developing UAV-based DL models to detect BXW in Central Africa can potentially bring about significant positive changes in BXW management and prevention within the region. Utilizing cost-effective UAV-RGB/MS tools and DL models, this detection system offers effective and practical support for monitoring BXW across diverse landscapes.

The impact of this technology manifests in several notable ways. Firstly, the developed DL models exhibit remarkable proficiency in accurately detecting healthy and BXW-infected banana stems. This proficiency facilitates early detection of BXW outbreaks, enabling swift intervention and targeted management strategies to curtail the disease's propagation. This approach minimizes crop losses and plays a crucial role in safeguarding food security in the region.

Secondly, using UAV-RGB/MS tools enables the acquisition of high-resolution aerial images, facilitating comprehensive monitoring of banana plantations across extensive areas. This technique provides a holistic understanding of BXW's distribution and severity, facilitating the identification of hotspots and areas that require immediate attention. Moreover, the developed airborne platform can be complemented with in-field scouting, sampling, and PCR testing [68, 73], as well as integration with other technological resources like the AI-driven banana app, Tumaini, [20] and GPS mapping of healthy and diseased banana crops (Fig. 16). These additional tools further contribute to establishing robust early warning alert systems across African banana production landscapes. Overall, integrating UAV-based DL models and related technologies holds immense potential for improving BXW management and prevention efforts. This approach can enhance disease control, minimize crop losses, and promote sustainable banana production in Central Africa by enabling early detection, precise monitoring, and complementary tools.

Regarding our objectives, our study achieved a professional annotation dataset from BXW at UAV images in Africa, the first time this kind of dataset has been obtained for future training or studies in the region. Secondly, this dataset uses different combinations of vegetation indices, bands, and fusion data (pansharpening) to allow us to understand the importance of spatial and spectral resolution to object detection, focusing on banana diseases. For example, an important characteristic to enhance is the flight height to get data because we obtain a better spatial resolution at a lower flight height. The influence of the deep learning model is not the most important in most cases; an important part involves the analysis of input data, in this case, images with their characteristics. Our study involves the analysis of input data to

continue with DL modeling using advanced techniques (Faster R-CNN and Yolo-V8); they were compared with the same dataset, having better results with Yolo-V8 in some cases.

In summary, the Faster-RCNN models performed better with the 60 m height images, demonstrating improved precision, recall, and F1 score for both MS and RGB models. Notably, the MS models significantly improve the detection of healthy stems, while the PAN model maintains relatively stable performance. These findings suggest that two-stage models capitalize on the improved resolution provided by the 60 m images, leading to enhanced disease-detection capabilities. However, Yolo-V8 models present top performance for detecting healthy and infected stems for RGB and PAN images at 100 m heights. However, images at lower altitudes confuse the model, drastically affecting the model's detecting capabilities. Additionally, this model presents difficulties detecting stems for all MS images regardless of the height.

The technical analysis is one of the significant steps in our project; however, our best contribution is to have a methodological section for an early warning system and an automated disease management system for BXW in DR Congo and other African countries with similar conditions and limitations. It will give us a background to facilitate the implementation of the Tumaini AI-powered mobile app database, enabling proactive measures for disease control and management.

5 Conclusions and future directions

In summary, this research has greatly improved the detection of banana diseases, specifically BXW in diverse African systems. A precise model for symptomatic stem detection was developed, and expert-annotated datasets were compiled for future studies. Advanced DL algorithms have increased detection accuracy, equipping farmers with efficient tools for early disease identification and management.

Despite the promising outcomes, numerous domains stand to gain from additional in-depth investigation and refinement. Collecting images at lower altitudes or utilizing equipment with resolutions capable of achieving similar resolutions to RGB is crucial to improving the performance of MS-DL models or using pansharpening techniques. Additionally, further research should focus on expanding dataset expansion and comprehensive field annotations to enhance the accuracy of diseased stem identification.

By integrating the UAV-RGB pipeline with other digital tools, such as satellite imagery, AI-powered smartphone apps like Tumaini (https://play.google.com/store/apps/details?id=ciat.cgiar.org.tumaini&hl=e), along with linked disease mapping databases, advanced early warning systems can continue to be further developed across banana production landscapes.

This research sets the stage for future crop-disease detection and management advancements within mixed-complex African systems. By leveraging UAV remote sensing, DL models, and digital tools, agricultural practices can be further enhanced, and food security in Africa and beyond can be strengthened.

Acknowledgements The authors thank the Alliance of Bioversity International and the International Center for Tropical Agriculture (CIAT) Information Technology unit for providing facilities and logistics support. The authors thank Dr. Joe Tohme from the International Center for Tropical Agriculture (CIAT) for supporting this research. The authors would also like to acknowledge Henry Ruiz and John Valderrama for their contributions to developing the DL pipeline. Thanks to Nancy Safari of The Alliance for her immense support in collecting RGB and MS drone images in the South Kivu Province of eastern DR Congo. Thanks to Angela Fernando, CIAT consultant, for formatting and technical editing of the manuscript. Thanks also to Vincent Johnson editorial consultant to the Alliance science writing service, for final editing work. Bioversity International provided funding for field UAV image collection in the framework of the Roots, Tubers and Bananas program. We thank the RTB Program Management Unit and the Plant Health Initiative that supported this study and the CGIAR Fund Donors (www.cgiar.org/who-we-are/cgiar-fund/fund-donors-2).

Author contributions Michael Gomez Selvaraj, Juan Mora, and Guy Blomme designed the study, performed the experiments, and are the main contributing authors of the paper. Juan Mora, Cesar I. Alvarez, and Michael Gomez Selvaraj conducted data analysis, trained algorithms, and analyzed the data. Guy Blomme and Nancy Safari collected drone images of the disease, confirmed the symptoms, and pre-screened all the images collected in Africa. Michael Gomez Selvaraj, Juan Mora, and Cesar I. Alvarez wrote the manuscript, and all the authors checked it.

Funding Bioversity International, in the framework of the RTB-CC3.1 cluster, provided funding for field image collection and the CIAT Agrobiodiversity Research Area to carry out the image processing work.

Data availability The remotely sensed and field sampling data utilized in this study are accessible upon reasonable request from the corresponding author. The FALCONCV platform for DL model development, created by our lab, is accessible at https://github.com/haruiz/CvStudio. Additionally, the Banana AI app, a product of our lab's work, is available for free download on Google Play through Tumaini—Apps on Google Play at https://play.google.com/store/apps/details?id=ciat.cgiar.org.tumaini&hl=en_US&gl=US.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication All authors agreed to publish this manuscript.

Competing interests The authors declare that they have no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Tripathi L, Mwangi M, Abele S, Aritua V, Tushemereirwe WK, Bandyopadhyay R. Xanthomonas wilt: a threat to banana production in East and Central Africa. Plant Dis. 2009;93(5):440–51.
2. FAOSTAT. Crop Production Database. http://www.fao.org/faostat/en/#data/QC. Accessed on Apr 19 2020. http://www.fao.org/faostat/en/#data/QC
3. Swennen De Buck SR. "Bananas, the green gold of the South. VIB, Facts Ser. 2016;1:1–54.
4. Blomme G, et al. Bacterial diseases of bananas and enset: current state of knowledge and integrated approaches toward sustainable management. Front Plant Sci. 2017. https://doi.org/10.3389/fpls.2017.01290.
5. Blomme G, et al. A control package revolving around the removal of single diseased banana stems is effective for the restoration of Xanthomonas wilt infected fields. Eur J Plant Pathol. 2017;149(2):385–400. https://doi.org/10.1007/s10658-017-1189-6.
6. Barbedo J. A review on the use of unmanned aerial vehicles and imaging sensors for monitoring and assessing plant stresses. Drones. 2019;3(2):40. https://doi.org/10.3390/drones3020040.
7. Blomme G, et al. Fine-tuning banana Xanthomonas wilt control options over the past decade in East and Central Africa. Eur J Plant Pathol. 2014;139(2):271–87. https://doi.org/10.1007/s10658-014-0402-0.
8. Ocimati W, Bouwmeester H, Groot JCJ, Tittonell P, Brown D, Blomme G. The risk posed by Xanthomonas wilt disease of banana: mapping of disease hotspots, fronts and vulnerable landscapes. PLoS ONE. 2019;14(4):e0213691. https://doi.org/10.1371/journal.pone.0213691.
9. Gomez Selvaraj M, 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;169:110–24. https://doi.org/10.1016/j.isprsjprs.2020.08.025.
10. Alvarez-Mendoza CI, et al. Predictive modeling of above-ground biomass in brachiaria pastures from satellite and UAV imagery using machine learning approaches. Remote Sens. 2022;14(22):5870. https://doi.org/10.3390/rs14225870.
11. Ji S, Zhang C, Xu A, Shi Y, Duan Y. 3D convolutional neural networks for crop classification with multi-temporal remote sensing images. Remote Sens. 2018;10(2):75. https://doi.org/10.3390/rs10010075.
12. Neupane K, Baysal-Gurel F. Automatic identification and monitoring of plant diseases using unmanned aerial vehicles: a review. Remote Sens. 2021;13(19):3841. https://doi.org/10.3390/rs13193841.
13. Sugiura R, et al. Field phenotyping system for the assessment of potato late blight resistance using RGB imagery from an unmanned aerial vehicle. Biosyst Eng. 2016;148:1–10. https://doi.org/10.1016/j.biosystemseng.2016.04.010.
14. Özgüven MM. Determination of sugar beet leaf spot disease level (Cercospora Beticola Sacc.) with image processing technique by using drone. Curr Investig Agric Curr Res. 2018. https://doi.org/10.3247/CIACR.2018.05.000214.
15. Zhang D, Zhou X, Zhang J, Lan Y, Xu C, Liang D. Detection of rice sheath blight using an unmanned aerial system with high-resolution color and multispectral imaging. PLoS ONE. 2018;13(5):e0187470. https://doi.org/10.1371/journal.pone.0187470.
16. Hagen N, Kudenov MW. Review of snapshot spectral imaging technologies. Opt Eng. 2013;52(9):090901. https://doi.org/10.1117/1.OE.52.9.090901.
17. 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. https://doi.org/10.1016/j.compag.2020.105446.
18. Raimundo J, Lopez-Cuervo Medina S, Prieto JF, de Aguirre Mata J. Super resolution infrared thermal imaging using pansharpening algorithms: quantitative assessment and application to UAV thermal imaging. Sensors. 2021;21(4):1265. https://doi.org/10.3390/s21041265.
19. Shahi TB, Xu C-Y, Neupane A, Guo W. Recent advances in crop disease detection using UAV and deep learning techniques. Remote Sens. 2023;15(9):2450. https://doi.org/10.3390/rs15092450.
20. Selvaraj MG, et al. AI-powered banana diseases and pest detection. Plant Methods. 2019;15(1):92. https://doi.org/10.1186/s13007-019-0475-z.
21. Velumani K, et al. Estimates of maize plant density from UAV RGB images using faster-RCNN detection model: impact of the spatial resolution. Plant Phenomics. 2021. https://doi.org/10.3413/2021/9824843.
22. Nawaz M, et al. A robust deep learning approach for tomato plant leaf disease localization and classification. Sci Rep. 2022;12(1):18568. https://doi.org/10.1038/s41598-022-21498-5.

23. Cynthia ST, Md Shahrukh Hossain K, Md Hasan N, Asaduzzaman Md, Das AK. Automated Detection of Plant Diseases Using Image Processing and Faster R-CNN Algorithm. In: 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI). IEEE; 2019. pp. 1–5. https://doi.org/10.1109/STI47673.2019.9068092.
24. AT Nieuwenhuizen, J Hemming, HK. Suh. Detection and classification of insects on stick-traps in a tomato crop using Faster R-CNN. In: Proceedings of the Netherlands Conference on Computer Vision, Eindhoven, September 26–27, 2018 (pp. 1–5). Agro Field Technology Innovations, GTB Teelt & Gewasfysiologie A, WPR GTB AgroFoodRobotics. Eindhoven: Agro Field Technology Innovations; 2018. pp. 1–5.
25. Peng J, et al. Wild animal survey using UAS imagery and deep learning: modified Faster R-CNN for kiang detection in Tibetan plateau. ISPRS J Photogramm Remote Sens. 2020;169:364–76. https://doi.org/10.1016/j.isprsjprs.2020.08.026.
26. Lin T-L, Chang H-Y, Chen K-H. The pest and disease identification in the growth of sweet peppers using faster R-CNN and mask R-CNN. J Internet Technol. 2020;21(2):605–14.
27. Huangfu Z, Li S. Lightweight you only look once v8: an upgraded you only look once v8 algorithm for small object identification in unmanned aerial vehicle images. Appl Sci. 2023;13(22):12369. https://doi.org/10.3390/app132212369.
28. Zhu N, et al. Deep learning for smart agriculture: concepts, tools, applications, and opportunities. Int J Agric Biol Eng. 2018;11(4):21–8. https://doi.org/10.25165/j.ijabe.20181104.4475.
29. Mishra AM, Harnal S, Gautam V, Tiwari R, Upadhyay S. Weed density estimation in soya bean crop using deep convolutional neural networks in smart agriculture. J Plant Dis Prot. 2022;129(3):593–604. https://doi.org/10.1007/s41348-022-00595-7.
30. Nagaraju M, Chawla P, Upadhyay S, Tiwari R. Convolution network model based leaf disease detection using augmentation techniques. Expert Syst. 2022. https://doi.org/10.1111/exsy.12885.
31. Kaur P, et al. Recognition of leaf disease using hybrid convolutional neural network by applying feature reduction. Sensors. 2022;22(2):575. https://doi.org/10.3390/s22020575.
32. Pourazar H, Samadzadegan F, Dadrass Javan F. Aerial multispectral imagery for plant disease detection: radiometric calibration necessity assessment. Eur J Remote Sens. 2019;52(3):17–31. https://doi.org/10.1080/22797254.2019.1642143.
33. Steward B, Gai J, Tang L. The use of agricultural robots in weed management and control. In: Billingsley J, editor. Robotics and automation for improving agriculture. London: Burleigh Dodds Science Publishing; 2019. p. 161–86.
34. Neupane B, Horanont T, Hung ND. Deep learning based banana plant detection and counting using high-resolution red-green-blue (RGB) images collected from unmanned aerial vehicle (UAV). PLoS ONE. 2019;14(10):e0223906. https://doi.org/10.1371/journal.pone.0223906.
35. W Wu, H Xu, S Zhong, MR Lyu, I King. Deep validation: Toward detecting real-world corner cases for deep neural networks. In: 2019 49th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN); 2019. pp. 125–137.
36. Ghojogh B, Crowley M. The theory behind overfitting, cross validation, regularization, bagging, and boosting: tutorial. arXiv preprint. 2019. https://doi.org/10.4855/arXiv.1905.12787.
37. Agisoft. Agisoft Metashape. https://www.agisoft.com. Accessed 23 Jun 2020. https://www.agisoft.com
38. Sulaiman AG, Elashmawi WH, Eltaweel GhS. IHS-based pan-sharpening technique for visual quality improvement using KPCA and enhanced SML in the NSCT domain. Int J Remote Sens. 2021;42(2):537–66. https://doi.org/10.1080/01431161.2020.1811913.
39. Selvaraj MG, Valderrama M, Guzman D, Valencia M, Ruiz Guzman H, Acharjee A. 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. 2020. https://doi.org/10.2120/rs.2.24148/v3.
40. Bhatt N, Varma S. Classify-imbalance data sets in IoT framework of agriculture field with multivariate sensors using centroid-based oversampling method. Natl Acad Sci Lett. 2023. https://doi.org/10.1007/s40009-023-01249-4.
41. Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. J Big Data. 2019;6(1):60. https://doi.org/10.1186/s40537-019-0197-0.
42. dos Tanaka FHKS, Aranha C. Data augmentation using GANs. Proc Mach Learn Res. 2019;1:1–16.
43. Ren S, He K, Girshick R, Sun J. Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans Pattern Anal Mach Intell. 2017;39(6):1137–49. https://doi.org/10.1109/TPAMI.2016.2577031.
44. X Mai, H Zhang, and MQ-H Meng. Faster R-CNN with Classifier Fusion for Small Fruit Detection. In 2018 IEEE International Conference on Robotics and Automation (ICRA). IEEE; 2018. pp. 7166–7172. https://doi.org/10.1109/ICRA.2018.8461130.
45. J Redmon, S Divvala, R Girshick, A Farhadi. You Only Look Once: Unified, Real-Time Object Detection. Proceedings of the IEEE conference on computer vision and pattern recognition; 2016. pp. 779–786
46. Shi J, Bai Y, Zhou J, Zhang B. Multi-crop navigation line extraction based on improved YOLO-v8 and threshold-DBSCAN under complex agricultural environments. Agriculture. 2023;14(1):45. https://doi.org/10.3390/agriculture14010045.
47. Terven J, Cordova-Esparza D. A comprehensive review of YOLO architectures in computer vision: from YOLOv1 to YOLOv8 and YOLO-NAS. Mach Learn Knowl Extr. 2023. https://doi.org/10.3390/make5040083.
48. He K, Zhang X, Ren S, Sun J. Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Trans Pattern Anal Mach Intell. 2015;37(9):1904–16. https://doi.org/10.1109/TPAMI.2015.2389824.
49. S Liu, L Qi, H Qin, J Shi, J Jia. Path Aggregation Network for Instance Segmentation. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE; 2018. pp. 8759–8768. https://doi.org/10.1109/CVPR.2018.00913.
50. T-Y Lin, P Dollar, R Girshick, K He, B Hariharan, S Belongie. Feature Pyramid Networks for Object Detection. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE. 2017. pp. 936–944. https://doi.org/10.1109/CVPR.2017.106.
51. Lin T-Y, et al. Microsoft COCO: common objects in context. In: Fleet D, Pajdla T, Schiele B, Tuytelaars T, editors., et al., Lecture notes in computer science. Cham: Springer International Publishing; 2014. p. 740–55.
52. Henry Ruiz. FalconCV. https://github.com/haruiz/FalconCV. Accessed 03 Jun 2022. https://github.com/haruiz/FalconCV
53. Mao J, Tian W, Li P, Wei T, Liang Z. Phishing-alarm: robust and efficient phishing detection via page component similarity. IEEE Access. 2017;5:17020–30. https://doi.org/10.1109/ACCESS.2017.2743528.
54. D Mané. TensorBoard: TensorFlow's visualization toolkit. https://github.com/tensorflow/tensorboard. Accessed 28 Aug 2020. https://github.com/tensorflow/tensorboard

55. Tensorflow. TensorBoard: TensorFlow's visualization toolkit. https://www.tensorflow.org/tensorboard?hl=en. Accessed 19 Jul 2022. https://github.com/tensorflow/tensorboard
56. D Bolya, S Foley, J Hays, and J Hoffman. Tide: A general toolbox for identifying object detection errors. In Computer Vision–ECCV 2020: 16th European Conference. Glasgow: Proceedings, Part III 16; 2020, pp. 558–573. 2020
57. Liu W, et al. SSD: single shot MultiBox detector. In: Leibe B, Matas J, Sebe N, Welling M, editors., et al., Computer vision—ECCV 2016. ECCV 2016. New York: Springer; 2016. p. 21–37.
58. Clevers JGPW, et al. MERIS and the red-edge position. Int J Appl Earth Obs Geoinf. 2001;3(4):313–20. https://doi.org/10.1016/S0303-2434(01)85038-8.
59. Hancock DW, Dougherty CT. Relationships between blue- and red-based vegetation indices and leaf area and yield of alfalfa. Crop Sci. 2007;47(6):2547–56. https://doi.org/10.2135/cropsci2007.01.0031.
60. Pavlovic D, Nikolic B, Djurovic S, Waisi H, Andjelkovic A, Marisavljevic D. Chlorophyll as a measure of plant health: agroecological aspects. Pesticidi i fitomedicina. 2014;29(1):21–34. https://doi.org/10.2298/PIF1401021P.
61. H. Li et al. A Novel Loss Calibration Strategy for Object Detection Networks Training on Sparsely Annotated Pathological Datasets. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2020. MICCAI 202 Lecture Notes in Computer Science. Springer; 2020, pp. 320–329. https://doi.org/10.1007/978-3-030-59722-1_31
62. Zhao Z, Pang F, Liu Z, Ye C. Positive-unlabeled learning for cell detection in histopathology images with incomplete annotations. In: de Bruijne M, Cattin PC, Cotin S, Padoy N, Speidel S, Zheng Y, Essert C, editors. Lecture notes in computer science. Cham: Springer; 2021. p. 509–18.
63. S Wu, S Wang, B Pan, Q Ji. Deep Facial Action Unit Recognition from Partially Labeled Data. In: 2017 IEEE International Conference on Computer Vision (ICCV). IEEE; 2017. pp. 3971–3979. https://doi.org/10.1109/ICCV.2017.426.
64. Zhang K, Wu Q, Chen Y. Detecting soybean leaf disease from synthetic image using multi-feature fusion faster R-CNN. Comput Electron Agric. 2021;183:106064. https://doi.org/10.1016/j.compag.2021.106064.
65. Parvathi S, Tamil Selvi S. Detection of maturity stages of coconuts in complex background using Faster R-CNN model. Biosyst Eng. 2021;202:119–32. https://doi.org/10.1016/j.biosystemseng.2020.12.002.
66. Bari BS, et al. A real-time approach of diagnosing rice leaf disease using deep learning-based faster R-CNN framework. PeerJ Comput Sci. 2021;7:e432. https://doi.org/10.7717/peerj-cs.432.
67. S Ghoury, C Sungur, A Durdu. Real-Time Diseases Detection of Grape and Grape Leaves using Faster R-CNN and SSD MobileNet Architectures. In International Conference on Advanced Technologies, Computer Engineering and Science (ICATCES 2019). Alanya; 2019. pp. 39–44.
68. Amarasingam N, Gonzalez F, Salgadoe ASA, Sandino J, Powell K. Detection of white leaf disease in sugarcane crops using UAV-derived RGB imagery with existing deep learning models. Remote Sens. 2022;14(23):6137. https://doi.org/10.3390/rs14236137.
69. Li Z, Peng C, Yu G, Zhang X, Deng Y, Sun J. Light-head R-CNN: in defense of two-stage object detector. arXiv preprint. 2017. https://doi.org/10.4855/arXiv.1711.07264.
70. Lu X, Li Q, Li B, Yan J. MimicDet: bridging the gap between one-stage and two-stage object detection. In: Vedaldi A, Bischof H, Brox T, Frahm J-M, editors. Computer vision—ECCV 2020. ECCV 2020. New York: Springer; 2020. p. 541–57.
71. Sabottke CF, Spieler BM. The effect of image resolution on deep learning in radiography. Radiol Artif Intell. 2020;2(1):e190015. https://doi.org/10.1148/ryai.2019190015.
72. Seifert E, et al. Influence of drone altitude, image overlap, and optical sensor resolution on multi-view reconstruction of forest images. Remote Sens. 2019;11(10):1252. https://doi.org/10.3390/rs11101252.
73. Sandino J, Gonzalez F, Mengersen K, Gaston KJ. UAVs and machine learning revolutionising invasive grass and vegetation surveys in remote arid lands. Sensors. 2018;18(2):605. https://doi.org/10.3390/s18020605.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.