Journal of Theoretical and Applied Information Technology
15th April 2024. Vol.102. No 7

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

REVOLUTIONIZING POTATO LATE BLIGHT SURVEILLANCE: UAV-DRIVEN OBJECT DETECTION INNOVATIONS

YASSINE ZARROUK[1], MIMOUN YANDOUZI[2], MOUNIR GRARI[3], MOHAMMED BOURHALEB[1], MOHAMMED RAHMOUNE[1], KHALID HACHAMI[1]

[1]LaRSA Lab., ENSAO, Mohammed First University, Oujda, Morocco

[2]LSI Lab., ENSAO, Mohammed First University, Oujda, Morocco

[3]MATSI Lab., ESTO, Mohammed First University, Oujda, Morocco

E-mail: m.yandouzi@ump.ac.ma

ABSTRACT

The ongoing integration of cutting-edge technologies is profoundly transforming agricultural oversight, where drones emerge as pivotal instruments for precise crop monitoring, early disease detection, and efficient land management. The harmonious synergy between drones and AI, specifically deep learning, is revolutionizing the surveillance of plant diseases, facilitating accurate realtime detection. This innovative approach not only promises enhanced effectiveness but also fosters sustainable agricultural management, steering the course of modern farming towards intelligent and environmentally conscious practices. This article undertakes a thorough comparative exploration of recent advancements in deep learning-based object detection. It investigates two model families - the single-pass YOLO (You Only Look Once) and the two-pass RCNN (Region-based Convolutional Neural Network) - along with their respective variations, with a particular focus on their potential use in drone-based agricultural surveillance, specifically targeting the detection of Potato Late Blight. The conducted experiments unveil promising results across various metrics, affirming the invaluable role of this tool in the detection and monitoring of agricultural diseases. This research not only contributes to advancing our understanding of deep learning in agricultural contexts but also underscores the significance of integrating cutting-edge technologies for sustainable and efficient farming practices.

Keywords: *Potato Late Blight, Deep Learning, Computer Vision, Drones, UAV, Object Detection*

1. INTRODUCTION

Plant diseases pose a significant threat to food security and agriculture. "Potato Late Blight" serves as a devastating example among these afflictions. The major issue lies in the rapid spread of these diseases, resulting in massive crop losses and compromising food availability. Faced with this threat, it becomes imperative to implement effective preventive measures and surveillance solutions. A robust monitoring system can play a crucial role by enabling early detection of disease signs, providing the opportunity for swift intervention to contain the spread and minimize damage [1]. The need for proactive action in the development of surveillance systems is thus a critical element to ensure crop resilience and secure long-term food safety [2].

Various techniques are employed to monitor and track plant diseases such as "Potato Late Blight." Remote sensing technologies, including satellite and drone imagery, offer a non-invasive means of assessing crop health by detecting subtle changes in vegetation that may indicate the presence of disease. Ground-based sensors equipped with advanced imaging systems and data analytics provide real-time monitoring at the field level. Additionally, molecular diagnostic tools enable the identification of specific pathogens responsible for diseases, facilitating targeted interventions. Traditional scouting, involving manual field inspections, remains a valuable technique for on-the-ground observation of symptoms [3].

The use of drones offers several notable advantages in the field of agricultural surveillance. These devices allow for rapid and precise data collection on a large scale, facilitating early identification of issues such as plant diseases. With their flexible operational capabilities, drones can be deployed at various stages of the agricultural season

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

to monitor crop growth, assess irrigation efficiency, and detect temperature variations. By automating data collection, farmers can make informed decisions and respond quickly to potential challenges, contributing to a more efficient and sustainable agriculture [4].

The fusion of drones and deep learning has revolutionized agricultural surveillance, particularly in object detection. This combination allows drones to capture high-resolution data over fields, while deep learning algorithms rapidly and accurately analyze this information. Automated detection of objects, such as early signs of plant diseases or growth variations, becomes achievable. This approach transforms agricultural monitoring by enabling proactive issue identification and more efficient resource management, paving the way for a sustainable increase in agricultural productivity [5].

The limitations mentioned above have prompted the suggestion of an automated system designed to aid in the detection of potential Potato Late Blight. This system leverages object detection algorithms, notably the YOLO and Faster R-CNN versions.

The subsequent sections of this document are structured as follows: Section two delves into the contextual foundation of our research, offering insights into the background. Section three provides an in-depth examination of relevant literature and related works. In section four, our proposed methodology is outlined and explained. The outcomes and discussions related to the proposed system are presented in section five. Finally, section six offers a conclusion along with perspectives for future considerations.

2. BACKGROUNDS

2.1 Computer Vision

Computer vision is a branch of artificial intelligence focused on imparting machines with the ability to visually understand and interpret the world around them. Its goal is to enable computer systems to perceive, analyze, and make decisions based on visual information extracted from images or videos. In essence, computer vision seeks to replicate the human capacity to visually interpret and comprehend its environment [6].

The application domains of computer vision are broad and continually expanding. Among its common applications are object recognition, face detection, image segmentation, text recognition, video surveillance, augmented reality, autonomous driving, medical imaging, virtual reality, and more. This technology finds use in sectors such as healthcare, security, industry, research, transportation, and significantly contributes to the advancement of various fields by automating and enhancing the visual understanding of computer systems [7].

2.2 Deep Learning

Deep learning, drawing inspiration from the human brain, empowers a computer to autonomously acquire knowledge. Although it is a recent development, it has already had a notable influence, particularly in identifying visual content, understanding spoken language, and processing natural language [8].

While powerful, deep learning faces challenges, notably the requirement for substantial amounts of data for effective learning. Despite these challenges, this captivating field of machine learning yields remarkable results [9].

2.3 Object Detection

Object detection is a crucial aspect of computer vision, aiming to locate and classify objects within an image or video. Various applications, ranging from video surveillance to automatic license plate recognition and medical imaging, heavily rely on this technology [10]. Deep learning has revolutionized object detection by replacing traditional methods with deep neural networks. These models automatically learn complex features, allowing for better generalization and improved performance across diverse scenarios [11].

Two predominant approaches dominate the object detection landscape (Figure 1): single-shot models, exemplified by YOLO, and two-shot models, represented by Faster RCNN. Single-shot models are known for their speed, making them suitable for real-time applications like urban surveillance. On the other hand, two-shot models, though more complex, provide superior precision, making them well-suited for tasks requiring meticulous object detection. The rapid evolution of object detection, fueled by deep learning, promises ongoing advancements. Researchers continuously explore new architectures to enhance accuracy, speed, and model adaptability. This creative dynamic paves the way for even more sophisticated applications, from autonomous driving to early disease detection, shaping the future of computer vision [12].

In object detection research, the current focus is on enhancing model interpretability and resilience against adversarial attacks. This improves their applicability in critical areas such as healthcare and autonomous vehicles.

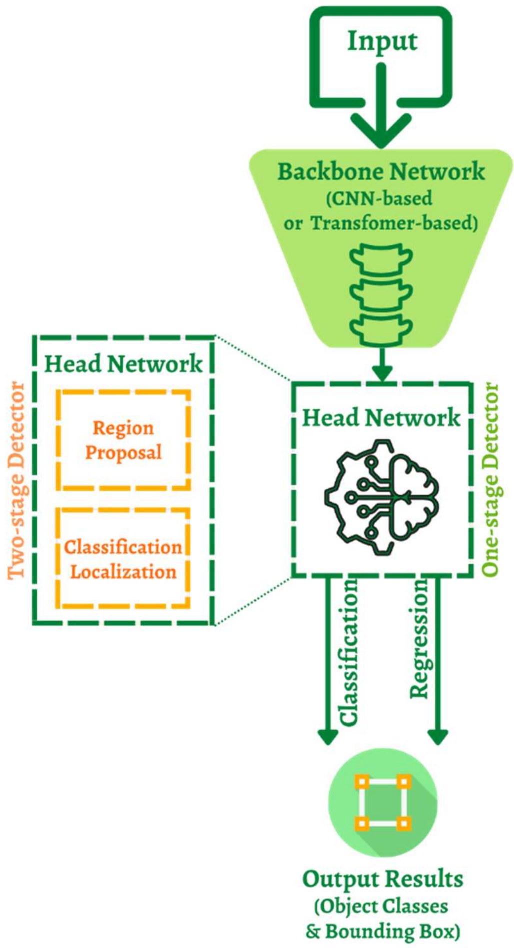

Fig. 1. One-Stage vs. Two-Stage Object Detection Architectures

2.3.1 R-CNN and its variants

R-CNN, Fast R-CNN, and Faster R-CNN mark successive advancements in the object detection field. Each model distinguishes itself through its capacity to enhance the effectiveness and precision of object detection in images. The core concept underlying R-CNN involves segmenting the image into regions of interest (RoI) and subsequently applying a convolutional neural network to each region for feature extraction [13].

Fast R-CNN [14] introduces a noteworthy enhancement by seamlessly integrating the region proposal process directly into the network. In contrast, Faster R-CNN [15] elevates this optimization to a higher level by introducing a dedicated network known as the "RPN" (Region Proposal Network). This specialized network accelerates and refines the generation of region proposals, contributing to a faster and more accurate object detection process.

Consequently, these successive models exemplify a substantial leap forward in the efficiency of object detection. This progress is achieved through the incorporation of refinements such as computational resource sharing, integration of dedicated networks, and the streamlining of the region proposal process [16].

2.3.2 YOLO

YOLO is a widely used convolutional neural network architecture in the realm of real-time object detection. What sets YOLO apart is its innovative ability to execute object detection in a single pass through the network, unlike conventional methods that require multiple steps. The first version, YOLOv1, introduced in 2016, demonstrated high efficiency but had limitations in precision, especially for small objects. Subsequent versions, YOLOv2 (or YOLO9000), YOLOv3, and YOLOv4, brought significant improvements in precision and processing speed. YOLOv2 introduced multi-scale detection and the ability to detect a large number of object classes, while YOLOv3 optimized the architecture for enhanced precision [17], [18].

In 2021, YOLOv5 was unveiled, marking a significant milestone in the evolution of the YOLO series. Subsequent releases, namely YOLO versions v6 [19] and v7 [20] (both launched in 2022), along with the most recent version, v8 [21] introduced in 2023, have continued to elevate the architecture's performance. These newer iterations bring forth notable enhancements in terms of both exactness and processing speed, showcasing the ongoing commitment to refining the YOLO framework. Additionally, the introduction of a novel segmentation pipeline in the latest versions demonstrates a forward-looking approach, expanding the capabilities of YOLO beyond object detection.

3. RELATED WORKS

Arshad et al. [22] aims to enhance agricultural productivity by providing precise and rapid solutions for disease detection. The key goal of this study is to formulate a hybrid deep learning model, PLDPNet, that integrates advanced technologies to effectively predict potato leaf diseases. Researchers created PLDPNet by combining features derived from two well-established deep learning models, VGG19 and Inception-V3, with the addition of vision transformers. This hybrid approach allows the model to leverage the strengths of each component, leading

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

to more accurate and reliable predictions. The dataset used to train and evaluate the PLDPNet model is notable, comprising images of potato leaves classified into three categories: early blight, late blight, and healthy leaves. To validate the model's universality and robustness, researchers tested PLDPNet on additional datasets, including those of apple and tomato leaves. This research highlights the potential of hybrid deep learning in the precise and rapid diagnosis of plant diseases, paving the way for practical applications that could significantly transform agricultural practices.

Anim-Ayeko et al. [23] employed the ResNet-9 model, a deep convolutional neural network, to classify potato and tomato leaf images from the PlantVillage dataset, which contains 6652 images, encompassing healthy leaves and those affected by early and late blight. Initially trained on a subset of 3990 images, ResNet-9 underwent testing on 1331 images after data augmentation to balance class distribution. Model optimization involved fine-tuning hyperparameters such as learning rate and epochs to enhance performance. The results demonstrated exceptional accuracy: a test accuracy of 99.25%, an overall accuracy of 99.67%, a recall of 99.33%, and an F1 score of 99.33%. In addition to quantitative evaluation, the authors utilized saliency maps to provide visual explanations of the model's predictions. These maps highlight regions within leaf images deemed most important for classification, thereby enhancing transparency and understanding of the model's internal workings.

The main objective of the study of Shi et al. [24] was to develop CropdocNet, a deep learning model capable of efficiently processing complex hyperspectral data obtained through aerial imaging. CropdocNet stands out for its ability to integrate and analyze spectral and spatial features to accurately identify cases of late blight, overcoming challenges related to terrain variability and the complexity of disease symptoms. The training and validation dataset included high-resolution hyperspectral images collected using a DJI S1000 drone equipped with a UHD-185 imaging spectrometer from Cubert GmbH. These hyperspectral images cover a wavelength range from 450 nm to 950 nm, comprising 125 spectral bands. In total, 23 hyperspectral images were mosaic-ed for the first experimental site (16382 × 8762 pixels), and 16 hyperspectral images for the second site, reflecting diverse terrain conditions and disease development stages, providing a realistic framework for model evaluation. CropdocNet achieved an average accuracy of 95.75%.

The study of Gao et al. [25] focuses on a Convolutional Neural Network model based on SegNet, specifically tailored for the semantic segmentation of late blight lesions. Field-collected images for training and testing capture the diversity of potato genotypes and disease severity levels. The dataset comprises approximately 500 RGB images from the field, spanning disease severity from 0% to 70%, resulting in 2100 cropped images. For training, 1600 of these cropped images were used, with 250 randomly selected for validation and testing. The results are significant, with an Intersection over Union (IoU) of 0.996 for the background and 0.386 for disease lesions in the test dataset. Furthermore, a linear relationship was established between manually assessed late blight visual scores and the number of lesions detected by deep learning at the canopy level. The study also examined the impact of class weight balancing on segmentation performance, underscoring the importance of class balancing in training deep learning models for agricultural applications.

4. PROPOSED METHOD

The innovative architecture proposed herein represents a groundbreaking leap in the realm of agricultural monitoring systems, designed to significantly augment efficiency. The core strategy involves seamlessly integrating cutting-edge interfaces that harness the power of drone imagery and deploy sophisticated deep learning techniques for disease detection, with a specific emphasis on combatting "Potato Late Blight". This ambitious endeavor hinges on the imperative implementation of deep learning algorithms meticulously crafted to discern and identify the characteristic symptoms of this particular affliction within the vast array of images captured by the aerial drones. These drones, equipped with state-of-the-art high-resolution cameras, ensure the precision of visual data acquisition, capturing intricate details vital for early disease detection. Following the data collection phase, a meticulously crafted pipeline will kick into action, employing pre-trained object detection models to automate the identification of nascent signs of "Potato Late Blight", as elucidated in Figure 2. The outcome of this comprehensive effort is poised to revolutionize agricultural surveillance, proactively mitigating disease risks, and fine-tuning crop management strategies. Ultimately, this initiative promises to usher in a new era of sustainable and high-quality agricultural production, propelling the industry towards greater resilience and productivity.

Journal of Theoretical and Applied Information Technology
15th April 2024. Vol.102. No 7

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Fig. 2. Proposed Architecture

The suggested approach involves five pivotal steps, detailed in Figure 3.

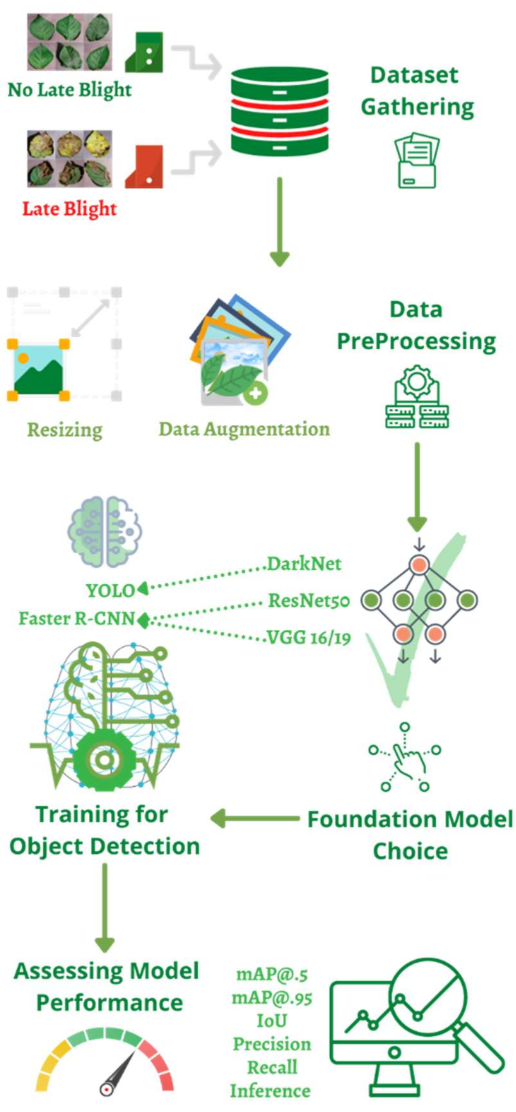

Fig. 3. Proposed Methodology

4.1 Dataset Gathering

The first step in our proposed methodology entails collecting a comprehensive array of images, including late blight-infected potatoes and those devoid of any disease symptoms (Figure 4). Following the implementation of data augmentation techniques, our dataset expands to include a substantial total of 2280 images, all meticulously labeled with the designation "Late Blight". These visuals were ac-quired utilizing ground-based cameras and airborne drones. Ground cameras were utilized to secure detailed imagery, ensuring precise portrayals of the ailment. Conversely, aerial drones were utilized to provide a broader perspective, capturing expansive areas within potato cultivation. The combination of ground and aerial approaches has yielded a comprehensive set of images, ranging from meticulous details to a panoramic overview.

Fig. 4. Sampling Images with and without Late Blight from the Database

We drew upon a diverse array of sources to compile this dataset, with our primary source being photographs of multiple potato-growing plots within the Berkane agricultural zone. These images were personally captured by our team of researchers during on-site visits, providing a detailed and up-to-date perspective on the local agricultural reality. We utilized both our smartphones and video footage captured by a drone. Furthermore, we gathered photographs from datasets available to the public, including online image repositories and platforms. These images were employed to broaden the dataset, thereby furnishing a more diverse array of data for the model to assimilate and learn from.

4.2 Data Preprocessing

The second step involves preprocessing image data, which is a critical stage. Its main aim is to improve the quality and consistency of data by removing undesirable elements such as noise and artifacts, which could adversely affect the model's performance. This involves operations like intensity normalization, distortion correction, and image scaling. By removing these disturbances, preprocessing ensures a clean and uniform input for the model, thereby facilitating convergence during the learning phase. Another critical aspect of the preprocessing process is image data augmentation. This step aims to diversify the dataset by applying various transformations to existing images, such as rotations, flips, zooms, and other geometric modifications with random parameters. The primary objective of augmentation is to enrich the variability of the training data, enabling the model to learn more robust and generalized patterns. By introducing controlled variations in the dataset, data augmentation helps improve the model's ability to handle real-world scenarios and strengthen its resilience to variable conditions. By combining these two steps, the overall preprocessing of image data creates an optimized training set, fostering maximum model performance.

4.3 Foundation Model Choice

As part of our investigation into enhancing the detection of "Late Blight" in images taken by agricultural drones designed for monitoring potato crops, we will explore two distinct approaches: single-pass models and double-pass models. This marks the third stage of our study. For single-pass models, we will explore the performance of YOLOv6, YOLOv7, and YOLOv8 architectures, all utilizing the Darknet backbone. On the other hand, for double-pass models, we will scrutinize the outcomes achieved with Faster R-CNN, employing backbones such as ResNet, VGG16, and VGG19. The selection of these convolutional neural network architectures as backbones is grounded in prior studies demonstrating their efficacy in similar contexts [26]. This comparative approach will enable us to assess and select the most suitable model for our specific task of "Late Blight" detection in agricultural environments.

ResNet. an abbreviation for Residual Network, represents a groundbreaking development in deep learning architecture pioneered by Microsoft Research in 2015. It ingeniously tackles the challenge of training exceptionally deep neural networks by introducing the concept of residual learning. The key breakthrough involves the incorporation of shortcut connections, or skip connections, which allow information to circumvent certain layers during forward propagation [27]. This ingenious design mitigates the vanishing gradient problem, facilitating the training of highly complex networks with enhanced accuracy. Renowned for its efficacy in computer vision tasks like image recognition and object detection, ResNet has become a pivotal model, celebrated for its ability to leverage the benefits of deep neural networks while surmounting com-mon training obstacles [28].

VGG. or Visual Geometry Group, is a family of influential convolutional neural networks (CNNs) known for their straightforward architecture. VGG16 and VGG19, with 16 and 19 layers respectively, are key members. Developed by the University of Oxford, these models feature a sequence of convolutional layers and densely connected layers. Despite their simplicity, VGG16 and VGG19 have proven highly effective in image classification tasks, showcasing their enduring impact on deep learning in computer vision [29].

DarkNet. a nimble and potent neural network framework developed by Joseph Redmon, creator of the YOLO algorithm, stands out for its efficiency in computer vision applications. This lightweight architecture, ideal for real-time object detection, supports both CPU and GPU computations, achieving a favorable balance between speed and accuracy. Darknet's open-source and modular design has propelled its widespread adoption in the deep learning community, owing to its adaptability and seamless integration into diverse projects [30].

4.4 Training for Object Detection

Refining our object detection models through adjustment and fine-tuning, utilizing the preprocessed and augmented dataset, represents the fourth step in our proposed approach. We adopt a dataset split of 70% for training, 20% for validation, and 10% for testing purposes. Our model repertoire includes YOLOv6, YOLOv7, and YOLOv8, as well as Faster R-CNN variants (ResNet50, VGG16 and VGG19). These models leverage a foundational backbone model.

Our training procedure involves utilizing labeled data, typically featuring bounding boxes around objects within images, accompanied by class information for the enclosed objects. The labeling process employs Open-Source Data Labeling software [31]. The input data format varies, with Faster R-CNN utilizing TensorFlow record (TFRecord) files, while YOLO employs TXT annotations and YAML config files. The ultimate goal of this step is to craft a highly accurate and dependable model tailored for the detection of Late Blight.

4.5 Assessing Model Performance

As a final step, we assess the models' performance by analyzing the entire dataset. This assessment utilizes 10% of the test set to gauge the average accuracy and inference speed of each model.

The testing stage is pivotal by allowing us to gauge model performance on novel data, aiding in determining the overall effectiveness of the models in detecting Potato Late Blight, and the potential for real-time use by drone.

5. FINDING RESULTS

In this section, we will begin by outlining the hardware specifications utilized in our study on object detection using a drone-mounted camera. Following this exposition, we will introduce the evaluation metrics employed to gauge the accuracy of the developed system. To assess the object detection results, a comparative analysis will be conducted against a dataset from real potato fields. Lastly, we will conclude this section by presenting and thoroughly discussing the outcomes achieved through the proposed method.

5.1 Technical Specifications

In the context of this study, the experimental setup included a Mavic Air drone by DJI fitted with a camera of high resolution [32]. For training and testing object detection models, we utilized a DELL PowerEdge R740 server featuring an Intel Xeon Silver 4210 2.2G processor and 80GB of RAM. This server was additionally equipped with two NVIDIA RTX A5000 GPUs, each with 24GB of graphics memory.

5.2 Evaluation Metrics

Numerous metrics exist to assess the effectiveness of an object detection algorithm, including:

5.2.1 Precision, recall and F1 score

Precision denotes the number of instances correctly identified as "Potato Late Blight" by the model, divided by the total number of instances detected as "Potato Late Blight," including false positives. Recall measures the model's ability to identify all actual occurrences of "Potato Late Blight" among all real instances and is calculated by dividing the number of instances correctly identified by the model by the total number of actual occurrences of "Potato Late Blight." The F1 score combines precision and recall into a single metric, offering a balanced assessment. It is calculated as the harmonic mean of precision and recall. These metrics are essential for a holistic evaluation of a model's effectiveness in accurately detecting the "Potato Late Blight" object.

$$F1\ score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (1)$$

5.2.2 Average Precision

The Average Precision (AP) is a performance metric that assesses the average precision of a detector across all classes, including the detection of the "Potato Late Blight" object. While it provides a comprehensive overview of the overall effectiveness of the detector, it may lack granularity in evaluating the performance for specific classes such as "Potato Late Blight." To calculate AP, one constructs the precision-recall curve for a given set of detections and then computes the average precision by averaging precision values at regularly spaced recall levels. In the specific case of "Potato Late Blight" detection, the formula for AP is applied to the set of detections related to this particular class [33]. The formula for AP is:

$$AP = \sum_{k=0}^{k=n-1}[Recalls(k) - Recalls(k+1) * Precisions(k)] \quad (2)$$

Where Recalls(n)=0, Precisions(n)=1, and n=Number of thresholds

5.2.3 Mean Average Precision

The mean Average Precision (mAP) stands as a sophisticated metric designed to assess the performance of object detection models. Unlike simpler measures, it considers precision and recall for each class, providing a detailed perspective. Calculated at various confidence thresholds, such as 0.5 and 0.95, mAP offers a thorough assessment of model robustness. Although its complexity can make interpretation challenging, mAP enhances evaluation by providing detailed insights into overall model effectiveness. The mean Average Precision formula encapsulates these considerations, making it a valuable tool for assessing and comparing object detection algorithms.

$$mAP = \frac{1}{N}\sum_{i=1}^{N} APi \quad (3)$$

5.2.4 Intersection over Union

Intersection over Union (IoU) is a metric used in object detection to assess the overlap between a detected object and its ground truth. It ensures the accuracy of object localization. Despite its complex formula involving rectangle coordinates, IoU provides detailed insights into the model's ability to precisely align detected objects with their actual references. IoU compares the intersection area of two rectangles to their union area. While metrics like Average Precision (AP) and mean Average Precision (mAP) are commonly used to evaluate object detection models, IoU complements these by

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

offering specific information on the model's accuracy in localizing objects.

$$IoU = \frac{(Area\ of\ Intersection)}{(Area\ of\ Union)} \quad (4)$$

$$Area\ of\ Intersection = (min(x2, x4) - max(x1, x3)) * (min(y2, y4) - max(y1, y3)) \quad (5)$$

$$Area\ of\ Union = (x2 - x1) * (y2 - y1) + (x4 - x3) * (y4 - y3) - Area\ of\ Intersection \quad (6)$$

5.2.5 Inference Time

The inference time of a deep learning model, particularly in object detection, refers to the duration it takes for the model to analyze an input image and generate predictions. This critical metric is influenced by factors like model architecture and available computational resources, impacting the model's real-time usability. The goal is to optimize inference time while maintaining sufficient accuracy, especially in applications that require a swift response.

5.3 Results Discussion

This study aims to assess the performance of two families of object detection models in identifying Potato Late Blight for the purpose of integrating this drone-based monitoring capability into a comprehensive supervision platform. We evaluated double-stage Faster-RCNN models, each employing different backbone networks, alongside single-stage YOLO models (v6, v7, and v8). These models were trained on a dataset comprising images sourced from the internet and others captured directly in potato fields in Morocco's eastern region. We trained each model for 100 epochs to ensure convergence across all metrics. The models underwent training and evaluation using various measures, including mAP at IoU thresholds of 0.5 and 0.95, recall, precision, and F1 score. Additionally, we measured the inference time (milliseconds/frame) using two Nvidia RTX A5000 GPUs. The F1 score, a metric that combines precision and recall in object detection models, provides a valuable balance between avoiding false detections and effectively capturing real objects. We prioritize the field of plant diseases with an emphasis on mAP@0.5. In critical domains like medicine, precision is paramount, leading to a preference for mAP@0.95. This higher confidence threshold ensures exceptionally reliable results, significantly reducing false positives.

Table 1 and Figure 5 comprehensively present the results obtained by various models during both the entire testing phase and the training process, respectively. An initial observation indicates that all models successfully detect potato late blight with precision, and notably, YOLO models require less training time. Among these models, Faster-RCNN stood out by showcasing the highest performance levels. Specifically, the Faster-RCNN model with ResNet-50 backbones (RS50) achieved an impressive F1 score of 93.97%, along with a mAP@0.5 of 95.32% and a mAP@0.95 of 81.12%. These results clearly underscore the robustness and efficiency of models adopting this particular configuration. This model also stands out for having the lowest inference time among Faster-RCNN models, displaying an average of 71.89 milliseconds per image. On one hand, while YOLO models demonstrated slightly lower performance compared to Faster-RCNN models, with YOLOv8 achieving a mAP@0.5 of 91.45% and a mAP@0.95 of 79.31%, they significantly excel in inference time. Specifically, YOLOv8 boasts an average of only 1.43 milliseconds per image. It is noteworthy that YOLO v6, v7, and v8 are not sequential versions, meaning that one is not necessarily newer than the other. Instead, they represent results from distinct research endeavors. This positions YOLO, especially YOLOv8, as an optimal choice for real-time applications, such as drone data collection, where fast processing speed is crucial. While the Fast-RCNN model (RS50) also yielded satisfactory results, its higher processing speed compared to YOLO models makes it the preferred choice in an architecture where the drone sends images to a ground station responsible for Potato Late Blight object recognition.

Table 1: Outcomes Achieved by the Implemented Models (On the Test Set)

DL Model	Inference Time (ms / frame)	Precision %	Recall %	F1 Score %	IoU %	mAP@0.5 %	mAP@0.95 %
Faster-RCNN (RS50)	~71.89	93.92	94.01	93.97	95.36	95.32	81.12
Faster-RCNN (VGG19)	~97.89	91.15	90.47	90.81	93.44	92.75	80.65
Faster-RCNN (VGG16)	~87.75	91.96	91.47	91.71	93.23	93.22	80.52
YOLOv6	~2.47	84.59	84.80	84.70	88.06	89.12	78.93
YOLOv7	~3.51	84.78	84.99	84.89	89.17	90.29	79.09
YOLOv8	~1.43	87.18	88.34	87.76	91.36	91.45	79.31

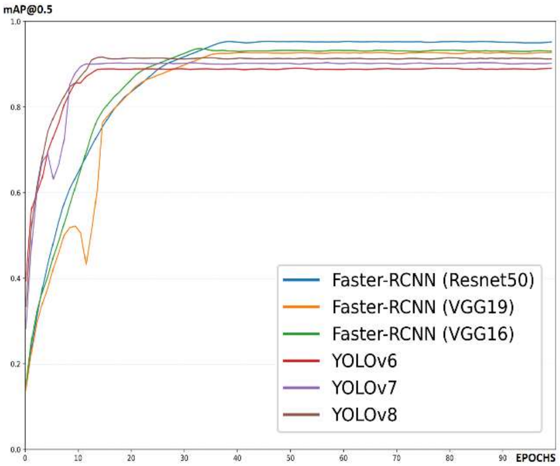

Fig. 5. Evolution of mAP@0.5 Across Epochs for Analyzed Models (Validation Set)

In summary, the decision between YOLO and Faster RCNN models for Potato Late Blight detection depends on the preferred balance between precision and processing speed. For applications prioritizing high accuracy, faster RCNN models, particularly the faster RCNN model (RS50), would be the preferred choice. Conversely, for real-time applications requiring fast processing speeds, YOLO models, especially YOLOv8, emerge as the optimal option (see Figure 6).

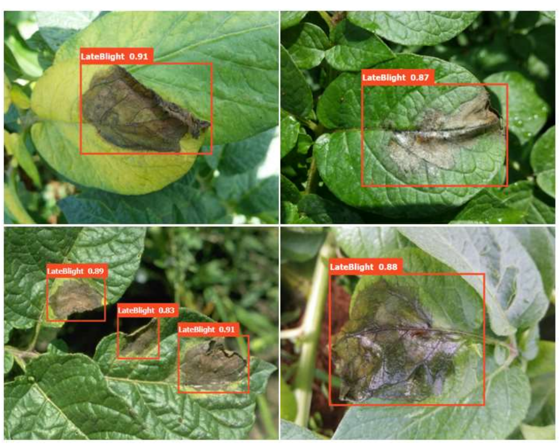

Fig. 6. Potato Late Blight detection examples using YOLOv8

In our research, we've pioneered an innovative methodology by harnessing the synergy of drone technology and object detection to pinpoint plant diseases. This approach stands out from prior efforts by facilitating extensive crop surveillance with heightened precision and efficiency in disease identification. Through the seamless integration of drone flight capabilities with advanced object detection algorithms, we achieve comprehensive coverage of agricultural fields, empowering proactive and swift disease monitoring. This groundbreaking approach holds immense promise in transforming agricultural monitoring practices, ushering in early disease detection that enables targeted interventions and mitigates crop losses.

6. CONCLUSION

New technologies are revolutionizing agricultural supervision, enabling precise crop management, real-time monitoring of environmental conditions, and optimization of yields. This sets the stage for more sustainable and efficient farming practices. The combination of drones and AI, particularly leveraging deep learning, is transforming agricultural monitoring by swiftly and accurately detecting plant diseases. This proactive approach allows for targeted management, reducing crop losses and promoting more efficient, sustainable agriculture. This article presents a comparative study of deep learning models for potential use in detecting "Potato Late Blight" using drones. The research aims to evaluate and compare various approaches to identify the most effective method for early and accurate detection of this disease, paving the way for innovative solutions in agricultural monitoring. We explored two families of object detection models in our study, namely the single-pass YOLO models (versions v6, v7, and v8) and the two-stage Faster R-CNN models with three backbone variants: Res-Net50, VGG16, and VGG19. The obtained results show promise for disease detection and monitoring, positioning this tool as a potentially valuable asset in this field. The studied models exhibit promising results across various metrics, positioning them-selves as a valuable tool for disease detection and monitoring. To select the most suitable model for drone images, achieving equilibrium between precision and speed of processing is crucial. For enhanced precision, it is advisable to use the faster RCNN model. Alternatively, for real-time applications emphasizing speed, the YOLO model, especially YOLOv8, emerges as the optimal choice with a mAP@0.5 of 91.45%, mAP@0.95 of 79.31%, and an inference time of nearly 1.43 milliseconds per image.

This research not only sheds light on the potential of deep learning models for early detection of "Potato Late Blight" using drones but also underscores the critical role of innovative technological integration in advancing agricultural monitoring. By evaluating and comparing various deep learning approaches, this study contributes to the growing body of knowledge aimed at enhancing

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

the precision and efficiency of disease detection in precision agriculture.

Future work could explore alternative datasets, refine model architectures, or investigate the integration of other advanced technologies to enhance the precision and efficiency of disease detection in precision agriculture, with a particular focus on considering and optimizing for varying weather conditions and drone flight altitudes. Furthermore, there is a need to delve into the development of robust algorithms capable of adapting to dynamic environmental factors, such as fluctuating weather patterns and varying terrain, to ensure consistent and reliable performance across diverse agricultural settings. Additionally, exploring the potential synergies between different sensing modalities and data fusion techniques could further enhance the accuracy and utility of disease detection systems in precision agriculture.

REFERENCES:

[1] M. Grari et al., "IoT-Based Approach for Wildfire Monitoring and Detection," pp. 205–213, 2023, doi: 10.1007/978-3-031-35251-5_19.

[2] J. Feng et al., "Research and Validation of Potato Late Blight Detection Method Based on Deep Learning," Agronomy, vol. 13, no. 6, p. 1659, Jun. 2023, doi: 10.3390/agronomy13061659.

[3] Y. Kumar, R. Singh, M. R. Moudgil, and Kamini, "A Systematic Review of Different Categories of Plant Disease Detection Using Deep Learning-Based Approaches," Arch. Comput. Methods Eng., vol. 30, no. 8, pp. 4757–4779, Nov. 2023, doi: 10.1007/s11831-023-09958-1.

[4] S. Lakshmi Priya and R. Subhashini, "Different Disease Detection of Paddy Crops using Drones - A Survey," Proc. 4th Int. Conf. Smart Electron. Commun. ICOSEC 2023, pp. 1221–1227, Jan. 2023, doi: 10.1109/ICOSEC58147.2023.10275809.

[5] N. Bharathiraja, K. Pradeepa, I. J. Sheela, G. Sudhakar, M. V. Kumar, and G. Kaur, "Prediction of Plant Leaf Diseases using Drone and Image Processing Techniques," Proc. - 5th Int. Conf. Smart Syst. Inven. Technol. ICSSIT 2023, pp. 1723–1727, Jan. 2023, doi: 10.1109/ICSSIT55814.2023.10061094.

[6] M. Yandouzi et al., "Review on forest fires detection and prediction using deep learning and drones," J. Theor. Appl. Inf. Technol., vol. 100, no. 12, pp. 4565–4576, 2022.

[7] M. Berrahal et al., "Investigating the effectiveness of deep learning approaches for deep fake detection," Bull. Electr. Eng. Informatics, vol. 12, no. 6, pp. 3853–3860, Dec. 2023, doi: 10.11591/EEI.V12I6.6221.

[8] M. Grari et al., "Using IoT and ML for Forest Fire Detection, Monitoring, and Prediction: a Literature Review," J. Theor. Appl. Inf. Technol., vol. 100, pp. 5445–5461, 2022.

[9] M. Yandouzi et al., "A Lightweight Deep Learning Model for Forest Fires Detection and Monitoring," pp. 697–705, 2023, doi: 10.1007/978-981-19-6223-3_71.

[10] V. Czymmek, L. O. Harders, F. J. Knoll, and S. Hussmann, "Vision-based deep learning approach for real-time detection of weeds in organic farming," I2MTC 2019 - 2019 IEEE Int. Instrum. Meas. Technol. Conf. Proc., vol. 2019-May, May 2019, doi: 10.1109/I2MTC.2019.8826921.

[11] M. Buzzy, V. Thesma, M. Davoodi, and J. M. Velni, "Real-Time Plant Leaf Counting Using Deep Object Detection Networks," doi: 10.3390/s20236896.

[12] M. Yandouzi et al., "Investigation of Combining Deep Learning Object Recognition with Drones for Forest Fire Detection and Monitoring," Int. J. Adv. Comput. Sci. Appl., vol. 14, no. 3, pp. 377–384, 2023, doi: 10.14569/IJACSA.2023.0140342.

[13] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "R-CNN: Regions with CNN features," Proc. ieee Conf. Comput. Vis. pattern Recognit., 2014, [Online]. Available: http://gwylab.com/pdf/rcnn_chs.pdf.

[14] "R-CNN, Fast R-CNN, Faster R-CNN, YOLO — Object Detection Algorithms | by Rohith Gandhi | Towards Data Science." https://towardsdatascience.com/r-cnn-fast-r-cnn-faster-r-cnn-yolo-object-detection-algorithms-36d53571365e (accessed Jan. 16, 2024).

[15] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 6, pp. 1137–1149, Jun. 2015, doi: 10.48550/arxiv.1506.01497.

[16] M. Grari et al., "Forest Fire Detection and Monitoring through Environment Sound Spectrum using Deep Learning," J. Theor. Appl. Inf. Technol., vol. 15, p. 19, 2023, Accessed: Jan. 17, 2024. [Online]. Available: www.jatit.org.

[17] "YOLO: Real-Time Object Detection." https://pjreddie.com/darknet/yolo/ (accessed Sep. 21, 2023).

[18] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," arXiv:1804.02767v1, 2018.
[19] C. Li et al., "YOLOv6: A Single-Stage Object Detection Framework for Industrial Applications," Sep. 2022, Accessed: Mar. 22, 2023. [Online]. Available: https://arxiv.org/abs/2209.02976v1.
[20] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," arXiv Prepr. arXiv2207.02696, 2022.
[21] G. Jocher, A. Chaurasia, and J. Qiu, "YOLOv8," 2023. https://github.com/ultralytics/ultralytics (accessed Mar. 30, 2023).
[22] F. Arshad et al., "PLDPNet: End-to-end hybrid deep learning framework for potato leaf disease prediction," Alexandria Eng. J., vol. 78, pp. 406–418, Sep. 2023, doi: 10.1016/J.AEJ.2023.07.076.
[23] A. O. Anim-Ayeko, C. Schillaci, and A. Lipani, "Automatic blight disease detection in potato (Solanum tuberosum L.) and tomato (Solanum lycopersicum, L. 1753) plants using deep learning," Smart Agric. Technol., vol. 4, Aug. 2023, doi: 10.1016/J.ATECH.2023.100178.
[24] Y. Shi, L. Han, A. Kleerekoper, S. Chang, and T. Hu, "Novel CropdocNet Model for Automated Potato Late Blight Disease Detection from Unmanned Aerial Vehicle-Based Hyperspectral Imagery," Remote Sens., vol. 14, no. 2, Jan. 2022, doi: 10.3390/RS14020396.
[25] J. Gao, J. C. Westergaard, E. H. R. Sundmark, M. Bagge, E. Liljeroth, and E. Alexandersson, "Automatic late blight lesion recognition and severity quantification based on field imagery of diverse potato genotypes by deep learning," Knowledge-Based Syst., vol. 214, Feb. 2021, doi: 10.1016/J.KNOSYS.2020.106723.
[26] M. Yandouzi et al., "Forest Fires Detection using Deep Transfer Learning," Int. J. Adv. Comput. Sci. Appl., vol. 13, no. 8, pp. 268–275, Oct. 2022, doi: 10.14569/IJACSA.2022.0130832.
[27] M. Berrahal et al., "Enhancing the medical diagnosis of COVID-19 with learning-based decision support systems," Bull. Electr. Eng. Informatics, vol. 13, no. 2, pp. 1316–1324, Apr. 2024, doi: 10.11591/eei.v13i2.6293.
[28] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 2016-Decem, pp. 770–778, Dec. 2015, doi: 10.1109/CVPR.2016.90.
[29] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2015.
[30] "Darknet: Open Source Neural Networks in C." https://pjreddie.com/darknet/ (accessed Jan. 02, 2024).
[31] "Open Source Data Labeling | Label Studio." https://labelstud.io/ (accessed Jan. 01, 2024).
[32] "Mavic Air - Product Information - DJI." https://www.dji.com/global/mavic-air/info#specs (accessed Jan. 18, 2024).
[33] "Average Precision | CloudFactory Computer Vision Wiki." https://wiki.cloudfactory.com/docs/mp-wiki/metrics/average-precision (accessed Jan. 05, 2024).