Journal of Information Systems Engineering and Management
2025, 10(2)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

An Improved Yolov8 for Enhanced Disease Detection and Localization in Cotton Plants

Jesal Desai [1*], Dr. Amit Ganatra [2]
[1]Research Scholar, U & P U Patel Department of Computer Engineering, Chandubhai S. Patel Institute of Technology, Faculty of Technology and Engineering, Charotar University of Science and Technology, Changa, India
[2]Provost, Parul University (PU), Waghodia 391760, Vadodara, India
* Corresponding author's Email:jaddesai654@gmail.com

ARTICLE INFO
Received: 17 Oct 2024
Revised: 19 Dec 2024
Accepted: 30 Dec 2024

ABSTRACT
Enhancing disease detection and localization in cotton plants using an upgraded version of YOLOv8, a cutting-edge object detection model noted for its efficiency and accuracy. Cotton plants are subject to a variety of diseases, including bacterial blight, armyworm, and powdery mildew, which reduce agricultural production. The study focuses on using advanced deep learning architectures, notably ResNet50 combined with YOLOv8 and Faster R-CNN, to improve disease identification and localization in cotton plants. CVAT was used to annotate a rigorously curated dataset of 2100 photos with bounding boxes, allowing for model training. To guarantee accurate performance assessment, the dataset was divided into three sets: training, validation, and test. ResNet50 was incorporated into the YOLOv8 and Faster R-CNN architectures, which were specifically designed to identify and localize bacterial blight, armyworm, and powdery mildew, as well as healthy cotton plants. Fine-tuning hyperparameters and optimizing model configurations were used in the experiments to obtain high illness classification accuracy and exact localization. The results obtained were commendable, with both models achieving an object detection IoU of approximately 0.95. It is worth noting that ResNet50-YOLOv8 exhibited superior classification accuracy at 98%, while ResNet50-FRCNN achieved a respectable 90%, thereby illustrating nuanced performance variations. In terms of precision, recall, and F1-score, ResNet50-YOLOv8 presented impressive values of 0.989, 0.986, and 0.988, respectively. Contrastingly, ResNet50-FRCNN displayed values of 0.858, 0.905, and 0.88, indicating variations in precision and recall

Keywords: Cotton Plant Disease Detection, Localization, Hybrid models, ResNet50, YOLOv8, Faster R-CNN

1. Introduction

According to the "Indian Food and Agriculture Organization" (FAO), agricultural plants have the potential to prevent a global decline in production of approximately 20-40% [1]. Plant disease is defined as any condition that hinders a plant from reaching its maximum productivity or potential [2]. Plant diseases pose a significant obstacle to both global food security and the future sustainability of agriculture [3]. Hence, it is essential to pay attention to crops and detect diseases at an early stage to reduce crop losses. Traditional methods to detect diseases in plants include visual inspection [4], symptom recognition, field surveys, sampling and laboratory analysis, expert consultation and observations of crops, and classification of diseases based on changes in the color, appearance, or texture of leaves. Laboratory-based methods, including hyperspectral techniques, gas chromatography, mass spectrometry, Thermography, and polymerase chain reaction have been used to identify diseases [5]. However, these methods suffer from drawbacks such as long detection times, the high level of expertise required, limited accuracy, the inability to detect at early stages, limited scalability, and a lack of real-time monitoring. Artificial Intelligence presents immense opportunities in agriculture [6]. Deep learning and computer vision have revolutionized plant disease detection by automating the process of identifying and diagnosing diseases in crops. Leveraging neural network architectures, such as "Convolutional Neural Networks" (CNNs), can be used to analyze large datasets of plant images and learn intricate patterns and features associated with various diseases. This allows for the accurate classification of healthy and infected plants based on visual cues, eliminating the need for manual inspection. Deploying such models in real-time applications, whether through mobile devices or remote sensing technologies, facilitates the rapid and scalable monitoring of agricultural fields. The integration of deep learning and computer vision not only expedites the detection process but also opens avenues for proactive disease management strategies, ultimately contributing to more sustainable and productive agriculture.

Existing solutions in the field have explored various deep-learning methods for disease detection in plants. For instance, VGGNet pre-trained on ImageNet [7], and an enhanced convolutional neural network (CNN) [8] is used for rice plant analysis, while basic CNN architectures such as GoogLeNet, AlexNet, and ResNet have been employed for identifying tomato leaf diseases [9]. The DenseNet structure has been shown to achieve almost the same accuracy level as ResNet with fewer parameters [10]. In another method, model training is accelerated and semantic information is extracted from diseased leaf images using transfer learning with an adapted CNN, integrating deep and handcrafted features while enhancing discriminative power with a center loss constraint [11]. Also, methods like Resnet-50 with a modified Red Deer optimization algorithm are used for salient feature extraction before using the Deep Learning Convolutional neural network for the plant village dataset and the rice plant dataset [12]. In [13], Residual Network (ResNet34) was trained to perform the task of classification on 15200 crop images. Multimodal, with SVM classifier as pre-extraction along with Res net -18, Res net-34, and VGG-16, respectively, are used for disease detection in apple leaf [14]. Lightweight, improved object detection models like the Yolov5-based algorithm are implemented to achieve real-time localization and ripeness detection in tomato fruits [15]. In [16], A multilayered convolutional neural network is suggested for the categorization of mango leaves that are affected by anthracnose sickness.

ResNet50, a deep neural network architecture, was applied to plant disease detection to harness its advanced feature extraction capabilities. The ResNet model is trained using an open-source data set [17]. With 50 layers and residual connections, ResNet50 overcomes the challenges of training deep networks, enabling the effective learning of intricate disease-related patterns from plant images. Its architecture aids in capturing varied features that are important for the accurate classification of healthy and diseased plants. ResNet50 excels in plant disease detection tasks and offers heightened accuracy. By leveraging the benefits of ResNet50, the model can efficiently analyze large datasets, providing a robust tool for the automated, precise, and scalable monitoring of plant health and contributing to early disease detection and effective agricultural management. The residual block of the ResNet is measured using Equation.

$$o = G(z, U + z) \qquad (1)$$

Here, the input and output layers represent variables like o and z. The remaining maps are denoted by the function G [18].

You Only Look Once (YOLO), originally proposed by Joseph Redmon, Ross Girshick, Ali Farhadi, and Santosh Divvala, in 2016, is widely regarded as one of the most efficient detection algorithms due to its unique methodology [19]. Through one-stage detection, the category and location of the object can be obtained [20]. YOLOv8 "You Only Look Once version 8", a state-of-the-art object detection model, is employed in plant disease detection owing to its efficiency and accuracy. With a well-organized architecture, YOLOv8 processes entire images in one forward pass, thereby facilitating real-time disease localization. Their multiscale detection capabilities are useful for identifying various diseases in diverse plant species. YOLOv8 ensures the accurate localization and classification of diseases within plants, enabling precise intervention. The speed and accuracy of the model make it a preferred choice for large-scale agricultural monitoring. YOLOv8 enhances the efficiency of plant disease detection, contributing to timely and targeted agricultural management strategies.

The Fast “Region-based Convolutional Neural Network” (R-CNN) is a type of detection network that takes in region suggestions as input and returns the objects of those suggestions together with the offsets of the bounding boxes [21]. This method uses an instrument known as a Region Proposal Network (RPN), which generates prospective object regions before classifying and refining them [22]. Fast RCNN is more efficient than RCNN since it eliminates the need to input 2000 region proposals to the CNN, resulting in reduced processing time [23]. Faster R-CNN is applied to plant disease detection and localization due to its robust object detection capabilities. Utilizing a two-stage process involving region proposal and classification, Faster R-CNN accurately identifies and localizes diseases within plant images. Its Region Proposal Network (RPN) efficiently generates potential bounding boxes, enhancing localization precision. Faster R-CNN excels at handling overlapping instances, which are a common occurrence in plant disease scenarios. With a balanced trade-off between accuracy and speed, faster R-CNN is suitable for real-time detection and localization in agriculture. The versatility of the model allows it to adapt to various plant species and disease types, making it a valuable tool for comprehensive plant health monitoring, aiding in timely intervention and effective disease management strategies in agricultural settings.

In this study, two hybrid models were developed. These hybrid models were specifically crafted for the detection and localization of diseases in cotton plants. Three distinct disease categories, namely, bacterial blight, armyworm, and powdery mildew, were considered, along with a fourth class denoting a healthy status. The primary objective of these models was to discern whether a given cotton plant exhibited signs of disease, and if so, to identify the specific disease category. Furthermore, the models were designed to pinpoint the affected region of the leaf, providing a comprehensive analysis of plant health for effective agricultural management. These hybrid models represent a powerful synergy that leverages the strengths of both architectures, thereby leading to enhanced plant disease detection and localization. ResNet50, renowned for its depth and feature extraction process, captures intricate patterns in plant images and learns rich representations of various diseases. By serving as the backbone, ResNet50 provides a robust foundation for subsequent detection models. For YOLOv8, ResNet50 enriches the model with detailed features, enabling YOLOv8 to make precise predictions at different scales. YOLOv8’s ability to process images holistically aligns seamlessly with ResNet50’s feature-rich representations, enhancing the real-time detection efficiency. Similarly, when paired with Faster R-CNN, ResNet50 contributes to accurate region proposals through its feature maps. The faster R-CNN’s two-stage process for region proposal and classification is empowered by the detailed features extracted by ResNet50, resulting in refined disease localization. By combining ResNet50 with both YOLOv8 and Faster R-CNN, the fused models benefited from a holistic understanding of plant images at various scales, leading to improved accuracy in disease detection and localization. This fusion of capabilities addresses the complexity of diverse plant diseases, making this combined approach a potent solution for advancing automated plant health monitoring and enabling more effective agricultural management strategies.

In a comparative study, the ResNet50-YOLOv8 integration exhibited superior performance with a classification accuracy of 97.50% and an average IOU of 0.95 across four classes. In contrast, the ResNet50-Faster R-CNN combination achieved a classification accuracy of 90%, with a slightly higher average IOU of 0.953. These results highlight the effectiveness of combining ResNet50 with YOLOv8 for precise plant disease detection and localization.

The main contributions of the study are described below:

1. To curate this dataset, approximately 2100 images were sourced from Kaggle. Bounding boxes were annotated across all images using the Computer Vision Annotation Tool (CVAT).
2. In this paper, we have created and compared two hybrid models by integrating ResNet50 into Faster R-CNN and ResNet50 into Yolov8, which involved substituting the original backbone with a pre-trained ResNet50 that has 50 layers in 5 blocks.

3. Evaluation measures such as training accuracy, testing accuracy, precision, recall, and F1-score were used to evaluate classification results, while the IoU metric was used to assess object detection performance.
4. Finally, an overall comparison of both models was conducted. Both models performed almost at par, with Yolov8-ResNet50 showing slightly better performance.

2. Review of literature

Dang et al. [24] completed an extensive study on seven iterations of the state-of-the-art YOLO object detector to accurately identify the location of weeds in cotton fields. The dataset used in this study was the CottonWeedDet12, consisting of about 5648 pictures representing 12 distinct weed categories. The Monte Carlo cross-validation, conducted with 5 replications, revealed that the detection accuracy, as assessed by mAP@50, ranged from 88.14% for YOLOv3-tiny to 95.22% for YOLOv4.Two distinct experiments were carried out, one including augmented data and the other without. The inclusion of augmented data resulted in expedited model training. Furthermore, the mean average precision at 50 (mAP@50) had a boost to 95.63%. Wang et al., (2023) [25] presented a novel intelligent recognition technique that utilizes an enhanced version of the YOLOv8 model. The suggested approach achieves exceptional recognition accuracy and speed. In the Backbone network, the Global Attention Mechanism (GAM) is used to assign weights to relevant feature information, resulting in enhanced model accuracy. Upon verification of the rice and cotton datasets, the mean average precision (mAP) accuracy indicator achieves 71.27% and 82.91% in the respective datasets. The upgraded model's detection accuracy is shown to be effective and superior when compared to the Faster R-CNN, YOLOv7, and the original YOLOv8 models. Zheng et al., (2024) [26] presented a YOLOv8-DMAS model that utilized the YOLOv8 detection technique to accurately identify cotton weeds in challenging situations. In order to improve the model's capacity to detect various weed types at different scales, the BottleNeck components in the C2f network are substituted with the Dilation-wise Residual Module (DWR). Additionally, the final layer of the backbone now includes the Multi-Scale module (MSBlock). SoftNMS is implemented as a replacement for the original “Non-Maximum Suppression” (NMS) approach in order to enhance the accuracy of dense weed identification. The findings demonstrate that the enhanced model can effectively identify cotton weeds in intricate field settings in real time, hence offering valuable assistance for intelligent weeding investigations.

Zhao et al. (2022) [27] The MC-YOLOv4 model was implemented by replacing the CSPDarkNet53 component in the YOLOv4 network with the MobileNet v3 network, resulting in improved efficiency. Additionally, an attention mechanism was included to improve the capability of extracting features for the detection of weeds in potato fields. As a consequence, there was a 3.2% enhancement in the mean average precision (mAP) value of YOLOv4. Chavan et al. [28] combined AlexNet with VGGNet and used the Plant seedling dataset to build the AgroAVNET model. In the case of single-object identification, when each picture would only include a single weed, the model attained a classification accuracy of 93.64%. Abdalla et al. (2019) [29] examined the efficacy of identifying weeds in dense vegetation using semantic segmentation and reached a 96% accuracy rate using a pre-trained VGG16 network. Ferreira et al. (2019) [30] evaluated the use of clustered Convolutional Neural Network (CNN) algorithms using both labeled and unlabeled data in order to decrease the amount of manual labeling required. The method demonstrated high efficacy in semi-automatic weed detection using annotated data. Zhao et al. [31] implemented an enhanced DenseNet algorithm capable of reliably detecting companion weeds in uncultivated maize seedlings. By incorporating DropBlock regularization and "Efficient Channel Attention" (ECA), the model was able to attain an average classification accuracy of 98.63%. Han et al. (2023) [**32**] demonstrated a streamlined crop recognition model for identifying maize seedlings in challenging field conditions. Integrated the ExG index algorithm with the improved Otsu approach to efficiently extract weed information, taking into account the distribution features of the weeds. Liao et al. (2024) [**33**] introduced the SC-Net model, which utilizes a banded convolutional network and the banded multiscale convolution technique to expand the receptive field of the convolutional layer. Additionally, the model employs an attention-based feature fusion approach to successfully combine low- and high-level information for efficient identification of rice and weeds.

Ren et al., (2024) [34] suggested a new and better weed identification system, based on YOLOv8. By incorporating multi-scale dilation convolution and large kernel convolution, the Dilated Feature Integration Block enhances feature extraction in the backbone network by making use of data from many levels and scales. While the number of model parameters is just 6.62 M, extensive studies on two publicly available datasets demonstrate that the proposed model surpasses the benchmark model. Specifically, mAP50 improves by 4.7%, mAP75 by 5.0%, and 5.3% by 3.3%. Gao et al., (2024) [35]. Developed and tested a smart detection model for cotton pests and illnesses using deep learning. By combining state-of-the-art Transformer technology with knowledge graphs, the model is able to improve the accuracy of pest and disease feature detection. Mobile systems can now process data and analyze inferences more efficiently thanks to edge computing technologies. A rate of 0.94 for accuracy, 0.95 for mean average precision (mAP), and 49.7.90 for frames per second (FPS) were all reached by the suggested strategy in the experiments. Zayani et al., (2024) [36] presented a YOLOv8 algorithm-based deep-learning strategy for automated tomato disease identification. The model improved upon an earlier Roboflow dataset and reached a total accuracy of 66.67%. But there is a wide range of class-specific performance, which shows how difficult it is to distinguish between certain disorders. Additional investigation into data balance, investigating other designs, and implementing disease-specific metrics is recommended. Wang et al. (2022) [37], presented a detector that takes contextual information into account in order to overcome the difficulty of dealing with complicated backdrops in remote sensing photos. It also made improvements to the RCNN region proposal network. Furthermore, authors in paper [38] designed an MPFPN to retrieve minor object characteristics that had been lost in the deep semantic data. However, these approaches aren't practicable for low-power edge image processing devices since they need a lot of memory and computing power.

3. Methodology

2.1 Dataset

The dataset employed in this study focuses on cotton plants and encompasses three distinct diseases bacterial blight, armyworm, and powdery mildew—along with a category denoting the healthy state. To curate this dataset, approximately 2100 images were sourced from Kaggle https://www.kaggle.com/datasets/shaikmahmamadrafi/cotton-plant-disease-detection-datasets.

However, the downloaded images lacked bounding boxes, a critical component for model training. To address this, bounding boxes were meticulously annotated across all images using the Computer Vision Annotation Tool (CVAT). This comprehensive annotation process ensured the accurate delineation of disease-affected regions within each image.

Subsequently, the dataset was partitioned into training, validation, and test sets. The training set comprised 1750 images, providing a substantial volume for the models to learn and generalize patterns. The validation set, consisting of 250 images, served as a crucial component for hyperparameter tuning and performance validation during training. The test set, comprising approximately 100 images, remained untouched during training and validation, facilitating an unbiased assessment of model generalization to unseen data.

The dataset was then prepared for compatibility with the distinct architectures of ResNet50 combined with Faster R-CNN and ResNet50 combined with YOLOv8. For the former, the dataset was converted to the JSON format, facilitating seamless integration with the ResNet50 and Faster R-CNN model pipeline. For the latter, the dataset underwent conversion to the XML and TXT format, aligning with the requirements of the ResNet50 and YOLOv8 model fusion.

This meticulous dataset preparation process ensures the models are equipped with a diverse and representative collection of images, encompassing the nuanced variations in cotton plant diseases. The strategic partitioning of the dataset into training, validation, and test sets, combined with precise bounding box annotations, establishes a robust foundation for the subsequent evaluation of ResNet50 with Faster R-CNN and ResNet50 with YOLOv8 models. These efforts contribute to the reliability and generalizability of the models in accurately detecting and localizing diseases on cotton plants.

2.2 Resnet50 and Yolov8

The YOLOv8 architecture consists of three main components: Backbone, Neck, and Head [39]. The backbone is a modified variant of the CSPDarknet53 architecture, and the neck network, also known as the mid-section or transition layers, follows the backbone. It is responsible for aggregating and refining features extracted by the backbone. The design, techniques like "Feature Pyramid Networks" (FPN) or "Path Aggregation Networks" (PANet) are frequently employed, whereas the head has several convolutional layers followed by a sequence of fully connected layers. The primary function of the head is to make predictions regarding the boundaries of the boxes, objectness ratings, and class probabilities for the items that have been spotted inside an image.

Several parameters are integral to YOLOv8, including a default input size of 640 × 640 and a default number of layers set to 53. For the BBox loss, YOLOv8 employs three distinct loss functions: 1. Complete Intersection over Union (CIoU) Loss: Utilized for bounding box regression, CIoU loss is designed to improve localization accuracy. DFL (Distribution Focal Loss) Loss: This loss function contributes to better handling of challenging scenarios, enhancing the model's ability to detect objects, especially those of smaller sizes. BCE (Binary Cross-Entropy) loss for classification Loss: Applied to classification tasks, BCE loss is instrumental in optimizing the accuracy of class predictions. Empirical evidence has demonstrated that these losses significantly enhance the ability to identify items, particularly when dealing with minuscule objects.

In this approach, the backbone of Yolov8 was modified from CSPDarknet to Resnet50. To achieve this, adjustments were made to the model. Yaml file in the Ultralytics YOLOv8 repository, cloned from GitHub. After cloning the repository, direct navigation into the paralytics/models/v8 directory was performed. Subsequently, relevant sections in the model yaml file were updated using the ResNet50 architecture, accompanied by the necessary configuration changes. An additional update was made to the cfg file to specify the path of the modified model.

The loading of the pre-trained ResNet50 weights was achieved by specifying the path to the weights file in the weight parameter. The adjustment involved changing weights from None to path_to_resnet50_weights.pt. The code in tasks.py and trainer.py handles the loading of weights, ensuring the expected functionality.

Considering the differences between the YOLOv8 and ResNet50 architectures, changing the backbone to ResNet50 necessitated further adjustments for compatibility and optimal model performance. Experiments with various configurations were performed, and the results were systematically evaluated to determine the optimal setup for disease detection in cotton plants.

In the YOLOv8 model, the backbone comprises initial layers that process the input image before passing it through detection layers. This backbone extracts essential features from an image and is crucial for subsequent object detection.

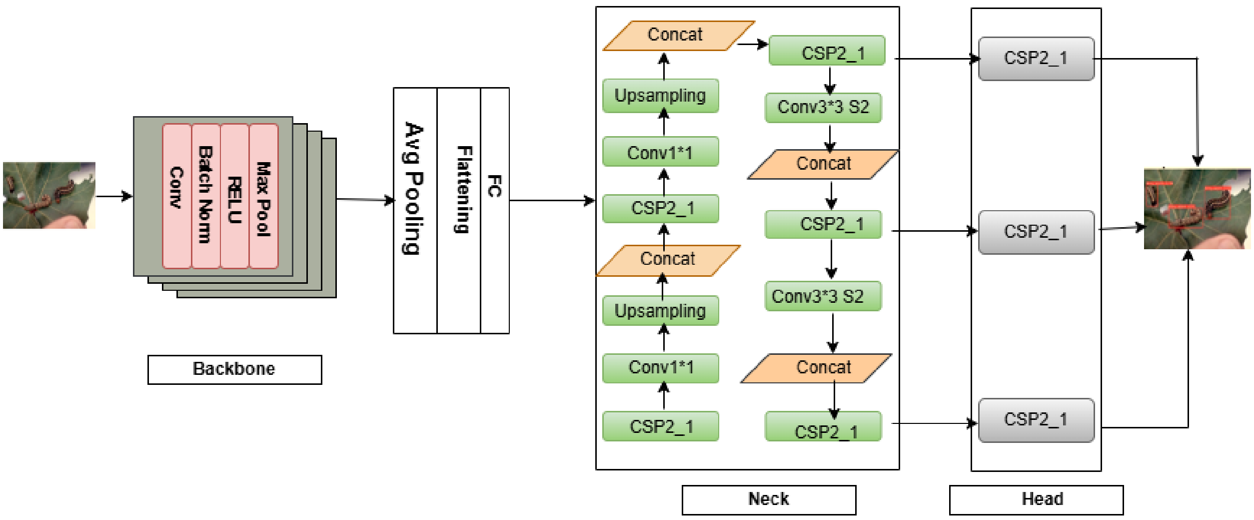


Figure 1: **ResNet50 with YOLOv8 Architecture.**

2.3 Resnet50 and FRCNN

The Faster R-CNN architecture encompasses several key components, each playing a crucial role in the ability of the model to detect and localize objects within an image.

2.3.1 Backbone - Convolutional Feature Extraction:

A conventional backbone network is tasked with processing the input image and extracting hierarchical features that are essential for subsequent stages of object detection. These features serve as the foundation for accurate localization and classification.

2.3.2 Region Proposal Network (RPN):

The RPN, an integral part of the Faster R-CNN architecture, is responsible for generating region proposals or candidate bounding boxes. Operating on the feature maps produced by the ResNet50 backbone, the RPN identifies potential objects within the image and enhances the localization efficiency of the model.

2.3.3 Region of Interest (RoI) pooling:

Pooling of the Regions of Interest (RoI) occurs after the RPN for the suggested areas. In order to maintain input size consistency for the following layers, this step entails extracting feature maps from RoIs with varied sizes. RoI pooling enhances the adaptability of the model to objects of diverse scales.

2.3.4 Fully Connected Layers (FCNs):

The RoI-pooled features progress to Fully Connected Layers (FCNs), where intricate processing occurs. These layers are pivotal for refining bounding box coordinates and predicting class probabilities, particularly in the context of disease categories for plant disease detection.

2.3.5 Bounding-box Regression and Classification:

The final output of the model encompasses predictions for the bounding box coordinates (regression) and class probabilities (classification). Bounding box regression is typically represented as offsets from the anchor boxes proposed by the RPN. This comprehensive output encapsulates the model's understanding of object locations and categories within a given image.

In our approach, the integration of ResNet50 into Faster R-CNN involves substituting the original backbone with pre-trained ResNet50, which has 50 layers in 5 blocks. This strategic substitution allows the model to leverage ResNet50's powerful feature extraction capabilities, enhancing its ability to discern complex patterns in data. Importantly, this modification maintained the consistency of the overall architecture, encompassing the RPN, RoI pooling, and final classification layers. This ensures seamless integration of ResNet50 into the general Faster R-CNN framework, resulting in improved performance, particularly tailored for the nuanced domain of plant disease detection and localization.

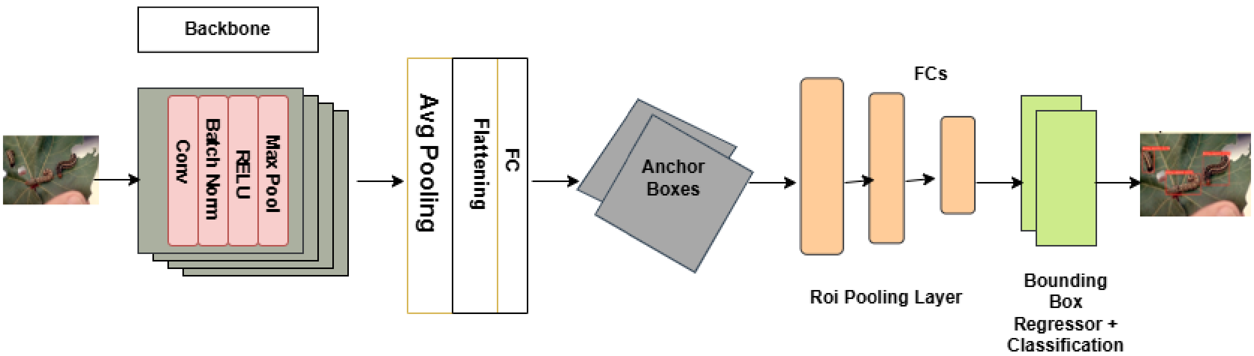


Figure 2: **ResNet50 with FRCNN Architecture.**

2.4 Hyperparameters

2.4.1 Resnet50 – Yolov8

The hybrid ResNet50 with YOLOv8 model for plant disease detection and localization was configured mostly with the default hyperparameters of Yolov8. Input images of size 640 × 640 pixels, strike a balance between computational efficiency and spatial information retention. A batch size of 4 was used instead of the default, which was 64 for optimizing the training process. The Adam optimizer facilitated adaptive learning rate adjustments, set at 0.01, with a momentum of 0.937 for efficient optimization. A weight decay of 0.0005 addresses overfitting concerns. The training spanned 300 epochs, implementing early stopping criteria to conclude at the 270th epoch.

2.4.2 Resnet50-FRCNN

The image size chosen for the experiment was 1333 × 800 pixels. The dataset was organized into three folders: training, validation, and test. To maintain the integrity of the actual test data and ensure its complete novelty, the images in the validation folder were duplicated in the test folder.

The neck type selected for the architecture was a Feature Pyramid Network. The optimization algorithm employed was "Stochastic Gradient Descent" (SGD) with a learning rate of 0.02 and a momentum of 0.9. Additionally, the weight decay was set to 0.0001 to mitigate overfitting during training.

The loss function utilized for training was the Cross-Entropy Loss. Non-maximum suppression (NMS) was applied to the model's output to filter redundant bounding boxes. In this process, an Intersection over Union (IoU) threshold of 0.7 was specified to determine the overlap required for NMS to suppress redundant detections.

The warm-up parameter was configured as "Linear." This strategy facilitated a smooth transition in the learning rate from an initial low value to the desired learning rate. This approach is particularly beneficial in the initial stages of training, when the model acquires fundamental features and representations, contributing to a more stable and controlled training process. Finally, the hybrid model of Resnet50-FRCNN was trained for 300 epochs.

Overall, these experimental configurations were designed to enhance the performance of the model by carefully selecting parameters and strategies in the training pipeline.

4. Benefits

ResNet50 with YOLOv8

Combining ResNet50 with YOLOv8 enhances cotton plant disease detection and localization. YOLOv8's real-time detection and ResNet50 feature extraction synergize, enabling the swift and accurate identification of diseases. The model excels in scenarios demanding rapid multi-instance detection, making it a valuable tool for efficient and timely plant health monitoring in cotton fields.

ResNet50 with Faster R-CNN (FRCNN)

Integrating ResNet50 with Faster R-CNN refined cotton plant disease localization. ResNet50's robust feature extraction, coupled with Faster R-CNN's
region-based approach, ensures the precise identification and localization of diseases. This hybrid model excels in scenarios where a detailed understanding and accurate localization of disease regions are critical, offering valuable insights into cotton plant pathology.

5. Results and Implementation

4.1 Training

Additionally, graphical representations were employed to comprehensively assess the training progress and performance stability of both hybrid models. The training graphs depict the evolution of the average precision, recall, and F1-score across epochs for ResNet50-YOLOv8 and ResNet50-FRCNN.

The graphs show convergence and fluctuations in these metrics, offering valuable insights into the learning trajectories of the models. Notably, trends in precision, recall, and F1-score can be observed, aiding in the identification of optimal epochs where the models achieve a balance between accuracy and generalization.

Precision, Recall, and F1-score in case ResNet50-YOLOv8 reached saturation at the 270th epoch, giving values of 0.989, 0.986, and 0.988, respectively. Contrastingly, ResNet50-FRCNN displayed values of 0.858, 0.905, and 0.88, at the 300th epoch, indicating variations in precision and recall.

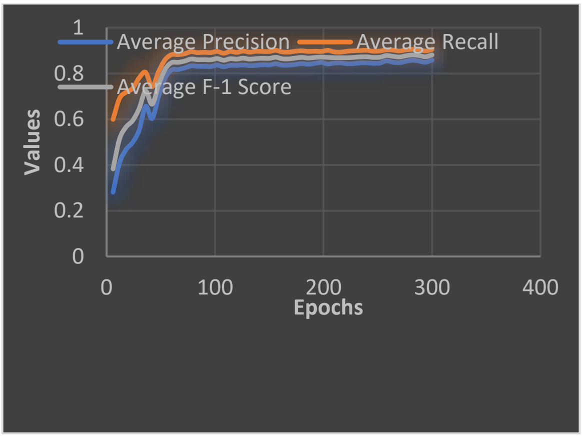


Figure 3.: Resnet50 + FRCNN Training Graph

Figure 3 displays the model's performance metrics Average Precision, Average Recall, and Average F-1 Score across 400 training epochs. The Average Precision starts around 0.2 and rapidly increases to approximately 0.85 by the 50th epoch. It then gradually improves, reaching slightly above 0.9 by the 400th epoch. The Average Recall begins at around 0.5 and climbs steeply to about 0.9 by the 50th epoch, maintaining stability with minor fluctuations around 0.9 for the remaining epochs. The Average F-1 Score starts at about 0.3 and quickly rises to approximately 0.85 by the 50th epoch, continuing to improve gradually and ending just under 0.9 by the 400th epoch. The model's performance significantly improves within the first 50 epochs and then maintains high values close to 0.9 with minimal variations thereafter.

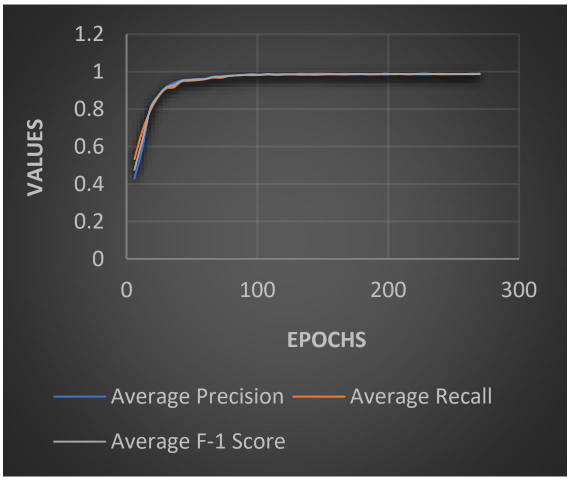


Figure 4: Resnet50 + YoloV8 Training Graph

Figure 4 illustrates the model's performance metrics Average Precision, Average Recall, and Average F-1 Score over 300 training epochs. The Average Precision begins low but quickly rises to approximately 0.95 by the 30th epoch, after which it stabilizes and maintains this value throughout the remaining epochs. Similarly, the Average Recall starts low, sharply increases to about 0.95 by the 30th epoch, and then remains consistent at this level for the rest of the training period. The Average F-1 Score follows the same trend, starting low, rapidly increasing to around 0.95 by the 30th epoch, and stabilizing at this value for the duration of the epochs. The model quickly achieves high performance within the first 30 epochs and maintains these high values close to 0.95 with minimal variation thereafter.

4.2Testing

100 unseen images were used for testing, with 25 images randomly selected from each class of cotton plants: Bacterial Blight, Powdery Mildew, Army Worm, and Healthy Plants. These images were consistent across both models for a fair comparison. The results of the hybrid model are summarized below.

Table 1: Resnet50 with Yolov8 Result

Disease Name	Total Images	Correctly Detected	Incorrect Detected
Army Worm	25	24	1
Bacterial Blight	25	25	0
Healthy	25	25	0
Powdery Mildew	25	24	1
Total	**100**	**98**	**2**
	Classification Accuracy		**98.00%**
	Average IoU		**0.95**

Table 1 shows a model's effectiveness in detecting several plant diseases across 100 images, grouped into four categories: Army Worm, Bacterial Blight, Healthy, and Powdery Mildew. Out of 25 images in the Army Worm category, the model successfully spotted 24 examples while incorrectly detecting one. For Bacterial Blight, the model accurately recognized 25 instances out of 25 images while incorrectly detecting 0. For the Healthy, the model successfully spotted 25 instances out of 25 photos while incorrectly detecting 0. The algorithm accurately recognized 24 incidences of Powdery Mildew out of 25 photos, while incorrectly detecting one. Overall, the model correctly recognized 98 of 100 photos, with a classification accuracy of 98.00%. Furthermore, the model's Average Intersection over Union (IoU) score is 0.95, showing that its detections are very accurate.

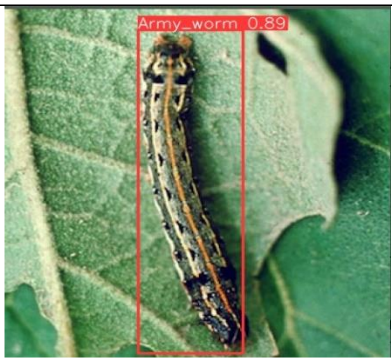
Table 2: Resnet50 with FRCNN Result

Disease Name	Total Images	Correctly Detected	Incorrect Detected/Not detected
Army Worm	25	22	3
Bacterial Blight	25	21	4
Healthy	25	25	0
Powdery Mildew	25	22	3
Total	**100**	**90**	**10**
	Classification Accuracy		**90.00%**
	Average IoU		**0.955**

Table 2 presents the performance of a model in detecting various plant diseases across a total of 100 images, divided into four categories: Army Worm, Bacterial Blight, Healthy, and Powdery Mildew. For the Army Worm category, out of 25 images, the model correctly detected 22 cases and incorrectly detected 3 cases. For the Bacterial Blight, out of 25 images, the model correctly detected 21 cases and 4 incorrectly detected. For the Healthy, out of 25 images, the model correctly detected 25 cases and 0 were incorrectly detected. For the Powdery Mildew, out of 25 images, the model correctly detected 22 cases and 3 incorrectly detected. Overall, the model correctly identified 98 out of 100 images, resulting in a classification accuracy of 90.00%. Additionally, the model's Average Intersection over Union (IoU) score is 0.955, indicating high precision in its detections.

4.3Model Output Comparison

Table 3: Model Output Result Comparison

Disease	Resnet50 + Yolov8	Resnet50 + Faster RCNN	Comments
Army Worm			Correctly classified by both models with different IoU values
Bacterial Blight			Resnet50 +Yolov8 - Correctly classified Resnet50 + FRCNN - Incorrectly classified
Healthy			Similar results

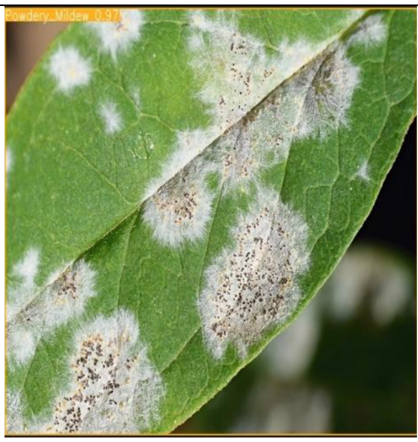
Powdery Mildew			Same classification results, different IoU results

Table 3 compares the performance of two model combinations for identifying plant diseases: Resnet50 + Yolov8 and Resnet50 + Faster RCNN. Both models properly diagnosed the disease "Army Worm," but with varying Intersection over Union (IoU) values. For "Bacterial Blight," the Resnet50 + Yolov8 model properly categorized the disease, but the Resnet50 + Faster RCNN model wrongly classified it. The results for "Healthy" plants were similar across both models. Finally, for "Powdery Mildew," both models produced the same classification results but with different IoU values, showing differences in the precision of their predictions.

5 Conclusion

In this research endeavor, we implemented two hybrid models, namely, ResNet50 combined with YOLOv8 and ResNet50 paired with FRCNN, to detect and localize diseases in cotton plants. Our efforts started with the careful preparation of datasets through image collection and annotation. The results obtained were commendable, with both models achieving an object detection IoU of approximately 0.95. It is worth noting that ResNet50-YOLOv8 exhibited superior classification accuracy at 98%, while ResNet50-FRCNN achieved a respectable 90%, thereby illustrating nuanced performance variations.

In terms of precision, recall, and F1-score, ResNet50-YOLOv8 presented impressive values of 0.989, 0.986, and 0.988, respectively. Contrastingly, ResNet50-FRCNN displayed values of 0.858, 0.905, and 0.88, indicating variations in precision and recall.

Undertakings in the future can focus on experimenting with various backbones and exploring alternative object detection models to improve architectures. Additionally, expanding the models to encompass a wider range of plant species and diseases and deploying them at the edge for real-time detection are crucial directions for practical implementation. These kinds of studies aim to refine existing models and broaden their applicability, ultimately making them valuable tools for precision agriculture and disease management. In conclusion, this research establishes a foundation for advancements in plant disease detection, highlighting the importance of ongoing exploration and adaptation in diverse agricultural contexts. The proposed metrics provide valuable insights into the intricate performance of each hybrid model, guiding future enhancements and practical applications.

Funding Information

The authors have not taken funding from anywhere.

Author's Contributions

Author Jesal Desai is responsible for doing a literature survey, implementing models, and writing the paper.

Dr.Amit Ganatra has guided author Jesal Desai in finalizing the topic and writing the paper.

Conflict of Interest

"The authors declare that there are no conflict.

References

[1] P. Deepika and B. Arthi, "Prediction of plant pest detection using improved mask FRCNN in cloud environment," Measurement: Sensors, vol. 24, Dec. 2022, doi: 10.1016/j.measen.2022.100549.

[2] K. R. Chairma Lakshmi, B. Praveena, G. Sahaana, J. Nithya Jenev, T. Gnanasekaran, and M. Hashim, “Yolo for Detecting Plant Diseases,” in Proceedings of the 3rd International Conference on Artificial Intelligence and Smart Energy, ICAIS 2023, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 1029–1034. doi: 10.1109/ICAIS56108.2023.10073875.

[3] P. Sajitha, A. D. Andrushia, N. Anand, and M. Z. Naser, “A review on machine learning and deep learning image-based plant disease classification for industrial farming systems,” 2024, doi: 10.1016/j.jii.2024.100572.

[4] Y. Toda and F. Okura, “How convolutional neural networks diagnose plant disease,” Plant Phenomics, vol. 2019, 2019, doi: 10.34133/2019/9237136.

[5] S. Ramesh et al., “Plant Disease Detection Using Machine Learning,” in 2018 International Conference on Design Innovations for 3Cs Compute Communicate Control (ICDI3C), IEEE, Apr. 2018, pp. 41–45. doi: 10.1109/ICDI3C.2018.00017.

[6] V. Dharmaraj and C. Vijayanand, “Artificial Intelligence (AI) in Agriculture,” Int.J.Curr.Microbiol.App.Sci, vol. 7, no. 12, pp. 2122–2128, 2018, doi: 10.20546/ijcmas.2018.712.241.

[7] J. Chen, J. Chen, D. Zhang, Y. Sun, and Y. A. Nanehkaran, “Using deep transfer learning for image-based plant disease identification,” 2020, doi: 10.1016/j.compag.2020.105393.

[8] V. Balaji et al., “Deep Transfer Learning Technique for Multimodal Disease Classification in Plant Images,” 2023, doi: 10.1155/2023/5644727.

[9] M. H. Saleem, J. Potgieter, and K. M. Arif, “plants Plant Disease Detection and Classification by Deep Learning”, doi: 10.3390/plants8110468.

[10] R. I. Hasan, S. M. Yusuf, and L. Alzubaidi, "Plants Review of the State of the Art of Deep Learning for Plant Diseases: A Broad Analysis and Discussion", doi: 10.3390/plants9101302.

[11] X. Fan, P. Luo, Y. Mu, R. Zhou, T. Tjahjadi, and Y. Ren, "Leaf image-based plant disease identification using transfer learning and feature fusion," Comput Electron Agric, vol. 196, May 2022, doi: 10.1016/j.compag.2022.106892.

[12] S. R. G. Reddy, G. P. Saradhi Varma, and R. L. Davuluri, “Resnet-based modified red deer optimization with DLCNN classifier for plant disease identification and classification,” 2022, doi: 10.1016/j.compeleceng.2022.108492.

[13] V. Kumar, H. Arora, Harsh, and J. Sisodia, “ResNet-based approach for Detection and Classification of Plant Leaf Diseases,” in 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), IEEE, Jul. 2020, pp. 495–502. doi: 10.1109/ICESC48915.2020.9155585.

[14] X. Li and L. Rai, “Apple Leaf Disease Identification and Classification using ResNet Models,” in ICEICT 2020 - IEEE 3rd International Conference on Electronic Information and Communication Technology, Institute of Electrical and Electronics Engineers Inc., Nov. 2020, pp. 738–742. doi: 10.1109/ICEICT51264.2020.9334214.

[15] T. Zeng, S. Li, Q. Song, F. Zhong, and X. Wei, “Lightweight tomato real-time detection method based on improved YOLO and mobile deployment,” Comput Electron Agric, vol. 205, Feb. 2023, doi: 10.1016/j.compag.2023.107625.

[16] U. P. Singh, S. S. Chouhan, S. Jain, and S. Jain, “Multilayer Convolution Neural Network for the Classification of Mango Leaves Infected by Anthracnose Disease,” IEEE Access, vol. 7, pp. 43721–43729, 2019, doi: 10.1109/ACCESS.2019.2907383.

[17] P. Sirenjeevi, J. M. Karthick, K. Agalya, R. Srikanth, T. Elangovan, and R. Nareshkumar, “Leaf Disease Identification using ResNet,” in Proceedings of the International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering, ICECONF 2023, Institute of Electrical and Electronics Engineers Inc., 2023. doi: 10.1109/ICECONF57129.2023.10083963.

[18] G. Ganesan and J. Chinnappan, “Hybridization of ResNet with YOLO classifier for automated paddy leaf disease recognition: An optimized model,” J Field Robot, vol. 39, no. 7, pp. 1087–1111, Oct. 2022, doi: 10.1002/rob.22089.

[19] A. Mohandas, M. S. Anjali, and U. Rahul Varma, “Real-Time Detection and Identification of Plant Leaf Diseases using YOLOv4-tiny,” in 2021 12th International Conference on Computing Communication and Networking Technologies, ICCCNT 2021, Institute of Electrical and Electronics Engineers Inc., 2021. doi: 10.1109/ICCCNT51525.2021.9579783.

[20] W. Bao, Z. Zhu, G. Hu, X. Zhou, D. Zhang, and X. Yang, “UAV remote sensing detection of tea leaf blight based on DDMA-YOLO,” Comput Electron Agric, vol. 205, Feb. 2023, doi: 10.1016/j.compag.2023.107637.

[21] K. Ashokkumar, S. Parthasarathy, S. Nandhini, and K. Ananthajothi, "Prediction of grape leaf through digital image using FRCNN," Measurement: Sensors, vol. 24, Dec. 2022, doi: 10.1016/j.measen.2022.100447.

[22] R. Rajamohanan and B. C. Latha, "An Optimized YOLO v5 Model for Tomato Leaf Disease Classification with Field Dataset," Engineering, Technology and Applied Science Research, vol. 13, no. 6, pp. 12033–12038, Dec. 2023, doi: 10.48084/etasr.6377.

[23] H. Rehana, M. Ibrahim, and Md. H. Ali, "Plant Disease Detection using Region-Based Convolutional Neural Network," Mar. 2023, [Online]. Available: http://arxiv.org/abs/2303.09063

[24] Dang, Fengying, Dong Chen, Yuzhen Lu, and Zhaojian Li. "YOLOWeeds: A novel benchmark of YOLO object detectors for multi-class weed detection in cotton production systems." *Computers and Electronics in Agriculture* 205 (2023): 107655.

[25] Wang, Yuchun, Cancan Yi, Tao Huang, and Jun Liu. "Research on Intelligent Recognition for Plant Pests and Diseases Based on Improved YOLOv8 Model." *Applied Sciences* 14, no. 12 (2024): 5353.

[26] Zheng, Lu, Junchao Yi, Pengcheng He, Jun Tie, Yibo Zhang, Weibo Wu, and Lyujia Long. "Improvement of the YOLOv8 Model in the Optimization of the Weed Recognition Algorithm in Cotton Field." *Plants* 13, no. 13 (2024): 1843.

[27] Zhao, J.; Tian, G.; Qiu, C.; Gu, B.; Zheng, K.; Liu, Q. Weed Detection in Potato Fields Based on Improved YOLOv4: Optimal Speed and Accuracy of Weed Detection in Potato Fields. *Electronics* **2022**, *11*, 3709.

[28] Chavan, Trupti R., and Abhijeet V. Nandedkar. "AgroAVNET for crops and weeds classification: A step forward in automatic farming." *Computers and electronics in agriculture* 154 (2018): 361-372.

[29] Abdalla, Alwaseela, Haiyan Cen, Liang Wan, Reem Rashid, Haiyong Weng, Weijun Zhou, and Yong He. "Fine-tuning convolutional neural network with transfer learning for semantic segmentation of ground-level oilseed rape images in a field with high weed pressure." *Computers and electronics in agriculture* 167 (2019): 105091.

[30] dos Santos Ferreira, Alessandro, Daniel Matte Freitas, Gercina Gonçalves da Silva, Hemerson Pistori, and Marcelo Theophilo Folhes. "Unsupervised deep learning and semi-automatic data labeling in weed discrimination." *Computers and Electronics in Agriculture* 165 (2019): 104963

[31] Zheng, Lu, Junchao Yi, Pengcheng He, Jun Tie, Yibo Zhang, Weibo Wu, and Lyujia Long. "Improvement of the YOLOv8 Model in the Optimization of the Weed Recognition Algorithm in Cotton Field." *Plants* 13, no. 13 (2024): 1843.

[32] Han, Xiaowu, Han Wang, Ting Yuan, Kunlin Zou, Qianfeng Liao, Kai Deng, Zhiqin Zhang, Chunlong Zhang, and Wei Li. "A rapid segmentation method for weed based on CDM and ExG index." *Crop Protection* 172 (2023): 106321.

[33] Liao, Juan, Minhui Chen, Kai Zhang, Huiyu Zhou, Yu Zou, Wei Xiong, Shun Zhang, Fuming Kuang, and Dequan Zhu. "SC-Net: A new strip convolutional network model for rice seedling and weed segmentation in paddy field." *Computers and Electronics in Agriculture* 220 (2024): 108862.

[34] Ren, Doudou, Wenzhong Yang, Zhifeng Lu, Danny Chen, Wenxuan Su, and Yihang Li. "A Lightweight and Dynamic Feature Aggregation Method for Cotton Field Weed Detection Based on Enhanced YOLOv8." *Electronics* 13, no. 11 (2024): 2105.

[35] Gao, Ruicheng, Zhancai Dong, Yuqi Wang, Zhuowen Cui, Muyang Ye, Bowen Dong, Yuchun Lu, Xuaner Wang, Yihong Song, and Shuo Yan. "Intelligent cotton Pest and disease detection: Edge computing solutions with transformer technology and knowledge graphs." *Agriculture* 14, no. 2 (2024): 247.

[36] Zayani, Hafedh Mahmoud, Ikhlass Ammar, Refka Ghodhbani, Albia Maqbool, Taoufik Saidani, Jihane Ben Slimane, Amani Kachoukh et al. "Deep Learning for Tomato Disease Detection with YOLOv8." *Engineering, Technology & Applied Science Research* 14, no. 2 (2024): 13584-13591.

[37] Wang, Yan, Chaofei Xu, Cuiwei Liu, and Zhaokui Li. "Context information refinement for few-shot object detection in remote sensing images." *Remote Sensing* 14, no. 14 (2022): 3255.

[38] Liu, Yingjie, Fengbao Yang, and Peng Hu. "Small-object detection in UAV-captured images via multi-branch parallel feature pyramid networks." *IEEE access* 8 (2020): 145740-145750.

[39] F. Aziz, F. Ernawan, M. Fakhreldin, and P. W. Adi, "YOLO Network-Based for Detection of Rice Leaf Disease," in 2023 International Conference on Information Technology Research and Innovation, ICITRI 2023, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 65–69. doi: 10.1109/ICITRI59340.2023.10249843.