Article QR

Title:	**Enhancing Agricultural Pest Management with YOLO V5: A Detection and Classification Approach**
Author (s):	Asif Raza[1], M. Kashif Shaikh[1], Osama Ahmed Siddiqui[2], Asher Ali[3], Afshan Khan[3]
Affiliation (s):	[1]Department of Software Engineering, Sir Syed University of Engineering and Technology, Karachi, Pakistan [2]Department of Computer Engineering, Sir Syed University of Engineering and Technology, Karachi, Pakistan [3] Department of Computer Science and Information Technology, Karachi, Pakistan
DOI:	https://doi.org/10.32350.umt-air.32.02
History:	Received: June 3, 2023, Revised: September 22, 2023, Accepted: November 11, 2023, Published: December 2, 2023
Citation:	A. Raza, M. K. Shaikh, O. A. Siddiqui, A. Ali, and A. Khan, "Enhancing agricultural pest management with YOLO V5: A detection and classification approach," *UMT Artif. Intell. Rev.*, vol. 3, no. 2, pp. 00–00, Dec. 2023, doi: https://doi.org/10.32350.umt-air.32.02
Copyright:	© The Authors
Licensing:	This article is open access and is distributed under the terms of Creative Commons Attribution 4.0 International License
Conflict of Interest:	Author(s) declared no conflict of interest

Enhancing Agricultural Pest Management with YOLO V5: A Detection and Classification Approach

Asif Raza[1], M. Kashif Shaikh[3*], Osama Ahmed Siddiqui[2], Asher Ali[2] and Afshan Khan[3]

[1]Department of Software Engineering, Sir Syed University of Engineering and Technology, Karachi, Pakistan

[2]Department of Computer Engineering, Sir Syed University of Engineering and Technology, Karachi, Pakistan

[3]Department of Computer Science and Information Technology, Karachi, Pakistan

ABSTRACT Due to the growing population and the numerous ecological challenges that affect crop yields, the need to modernize the crop production process has become increasingly critical. Swiftly managing potential threats to crops can have a substantial impact on overall crop production. Pests represent a significant menace, capable of causing substantial losses if not effectively controlled in a timely manner. In this study, a deep learning-based method for pest identification is introduced. The approach leverages the YOLO (You Only Look Once) object recognition SSD (single shot detection)algorithm in combination with the pre-trained DARKNET architecture to categorize pests into nine distinct classes. The study utilizes a publicly available dataset sourced from Kaggle, which comprises a total of 7,046 images. The outcomes reveal overall 83% of overall accuracy rate, with a notably low training and validation loss of 0.02%. Moreover, our model exhibits a notable enhancement in localization results, delivering a precision of 0.83, a recall of 0.83, an mAP-0.5 of 0.833, and an mAP-0.5:0.95 of 0.783.

INDEX TERMS learning, yolo, architecture, accuracy, crop, ecological, biodiversity

I. INTRODUCTION

Meeting the growing demand for food as the earth's population expands presents significant challenges for farmers worldwide. Sadly, conventional pest control strategies rely heavily on potentially hazardous synthetic chemicals such as insecticides & fungicides which can have unintended impacts on the environment and public health. To minimize these risks and ensure safe and sustainable food production alternative pest management practices are being developed. For example, crop rotation intercropping natural based insecticides/fungicides and reducing reliance on neonicotinoid insecticides linked with declines in honey bee populations. By adopting these approaches alongside improved farming practices providing healthy and safe food without negative environmental impacts is feasible [1].

Pest control is a critical factor in maximizing agricultural output, and accurate identification of crop-threatening organisms is essential to achieving that goal. Unfortunately, traditional pest recognition approaches can be unreliable due to seasonal fluctuations in pest accumulation and type. Deep learning has consequently provided a practical solution by furnishing the necessary means for

---

* Corresponding Author: mkshaikh@gmail.com

precise detection. This study aimed to evaluate the impact as part of analyses on optimizing proprietary algorithms and models for pest recognition in tomato plants, several optimizations' algorithms were tested including stochastic gradient descent, root mean square propagation, adaptive gradient, and adaptive moment estimation. Following these tests, it became clear that out of these optimization methods adaptive moment estimation produced the best effectiveness levels. This became apparent during testing with our deep convolutional neural networks architecture which was used after training on augmented data., achieving a 93% accuracy rate. This study illustrates the promise of deep learning to increase agricultural efficacy through improved pest identification [2].

Crop damage caused by pests and insects can be detrimental to human health and wealth, thus, the identification and classification of the various species of insects is vital. Traditional methods of manual examination can be time consuming and lack efficacy, whereas, deep learning with a Convolutional Neural Network (CNN) model offers a more effective and reliable approach. In this paper, an ensemble-based model, utilizing transfer learning with VGG16 and VGG19 as pre-trained models, is proposed [3], utilizing a voting classifier ensemble technique with a ResNetv50, the benchmark IP102 dataset was applied for training and assessment of the model's effectiveness. This dataset includes over 75,000 samples from 102 categories. The outcome illustrated that the presented ensemble model achieved a 82.5% accuracy rate, considerably surpassing recent state-of-the-art models developed for the same dataset, and affirming the efficacy of the model [4].

This research captured images of citrus fruit in natural lighting at three points: prior to, initially, and eight days post-infestation. Seventy percent of these pictures were assigned to training, ten percent to validation, and the remaining twenty percent for testing. The experiment also incorporated four pre-trained CNN models: ResNet-50 [5], GoogleNet [6], The Mediterranean fly was successfully identified and classified from healthy fruit by VGG-16 and AlexNet, together with the use of the SGDm, RMSProp, and Adam optimization algorithms. Results showed that VGG-16 achieved the highest detection accuracy and F1 of 98.33% and 98.36% when paired with the SGD algorithm, and AlexNet demonstrated the highest detection accuracy and F1 of 99.33% and 99.34% when working alongside the SGDm algorithm in the third stage. Of the models used, AlexNet with SGDm optimization was the quickest to train, requiring only 323 seconds. The findings of this study confirm that convolutional neural networks and machine vision systems are viable solutions for managing and controlling pests in orchards and other agricultural goods [7].

ViT, the Vision Transformer, is a reliable and expedient means of decreasing the level of subjectivity and improving the accuracy of crop pest detection. By using block partition, the regions exhibiting the most apparent characteristics are chosen with increased precision. Furthermore, the self-attention capacity of the transformer enables a more precise investigation into special solutions which are not readily recognizable in terms of lesions. Through this approach, precise judgment on crop disease and pest classifications is enabled, granting assistance in the optimization of crop disease and pest control procedures, consequently granting macroeconomic

balance and encouraging sustainable growth within the agricultural sector [8].

Fall armyworm (FAW) is an invasive pest that has caused significant economic and environmental losses to many of the world's major maize-producing regions. In Uganda, manual observation of the maize crop has been the primary method of detecting this destructive pest, but its late detection has led to costly delays in administering pest control measures. In order to combat this costly problem, a project was conducted to develop a machine learning classification model that can detect FAW damages in maize using transfer learning. The model was trained using Mobile Net weights and was evaluated using performance metrics. Several metrics, such as accuracy, recall, precision, and F1 score, were used to assess the model's performance. The approach was then put into practice on a platform built for the Raspberry Pi 4 using frameworks, which could be used to detect FAW damage in maize plants at an early stage. This development has the potential to help Ugandan farmers save time and money on the costly pest management processes [9].

Farmers have conventionally depend on on manual methods to detect weeds, diseases, and pests in their crop fields, but this approach is often unreliable and costly. To address these issues, deep learning approaches have been developed to effectively classify various types of plant images. For this purpose, several datasets of plant disease, pest, and weed images were acquired from diverse crops. In order to improve performance, the data were augmented using various deep convolutional neural network (DCNN) architectures, including Mobile Net, DenseNet201, VGG16, Hyper parameter Search, and InceptionV3. Ever since fine-tuning the hyper parameters and layers of each model, the InceptionV3 model achieved the highest accuracy of 87.85%. The Mobile Net and VGG16 models achieved an accuracy of 91.85% and 78.71%, respectively. The Dense-Net model performed the best with an accuracy of 99.62%, while the hyper parameter Search model achieved an accuracy of 71.07%. This model has demonstrated its effectiveness in weed, plant disease, and pest classification, presenting a cost-effective solution for farmers [10].

Insect identification has traditionally relied on manual assessment and traps, but advances in automation and pattern recognition have enabled the development of more effective solutions. Recent work on insect recognition using deep learning has achieved promising results, with the DenseNet201 model demonstrated as the most effective for classification performance. Experiments have shown that the model is capable of classifying 19 different insect classes with a mean accuracy of 87%, even in images with varying degrees of background clutter. This optimized approach for insect pest detection and recognition has the potential to revolutionize insect classification, making the process easier and more accurate for farmers and agricultural organizations [11].

Correct identification of pest is a crucial part of pest control. Unfortunately, traditional pest identification techniques lack accuracy and often misclassify pests, resulting in the improper use of pesticides. To tackle these challenges, the paper presents Deep-PestNet, a unique end-to-end deep learning framework for categorizing pests. Comprised of 11 learnable layers, including 8 convolutional and three fully connected layers, this innovative model uses image rotation and augmentation methods to expand the size

and diversity of the data set. Achieving a perfect accuracy rate, the Deep-PestNet framework was put to the test using the Deng's crop data set. As a further assessment, the framework was tried on the nine types of pests corresponding to the 'Pest Dataset' from Kaggle and produced superb results. Farmers can elevate their crop protection using this model, which has the capacity to transform the way pests are recognized and categorized. Crop yields take an immense hit when afflicted by disease, nutritional shortages, or pests. The employment of deep learning techniques presents a promising prospect for preemptive detection of such troubles, allowing farmers to address the issue in its infancy. An authentic dataset sourced from fields was employed in experimentation, comprising of leaves both undamaged and affected. For detecting disease rhizomes, convolutional neural networks boast an accuracy of 99%. Nutrient deficiency-affected leaves were best detected by ANN with a 96% accuracy. Lastly, VGG-16 models proved most effective in identifying pest pattern leaves with an accuracy of 96%. It is possible that these methods of disease detection could enhance the growth of ginger crops and prompt the development of real-time disease detection applications [12].

The health of a leaf is an important sign of the health of the entire plant, so it is vital to accurately and quickly detect issues such as diseases, pests, and other illnesses in order to effectively grow crops. Deep learning based methods have proved far better than traditional ones, and they are becoming more popular when it comes to digital image processing. This paper presents an AI-based method for identifying and classifying nine kinds of tomato leaf disease, as well as two types of rice leaf disease and pests. A Convolutional Recurrent Neural Network structure, including Gated Recurrent Unit, is used for the tomato leaf dataset, while Transfer Learning was utilized for the rice leaf dataset because of the lack of data and its imbalanced state. The model attained an accuracy of 99.62% for the tomato leaves and 91.63% for the rice leaves, exhibiting the potency of deep learning based techniques in providing timely and accurate recognition of pests and illnesses [13].

This paper proposes an effective and rapid method to identify common tomato pests. To achieve this, a tomato pest image dataset was established and then convolutional neural networks is deployed to recognize eight different classes of pests. Moreover, transfer learning was utilized to minimize the time required for training, while features extracted from pests were paired with three machine learning algorithms, such as discriminant analysis, support vector machine, and k-nearest neighbor approach. Moreover, Bayesian optimization was implemented for optimizing hyper-parameters, along with image augmentation. Of the different models applied, the VGG16 model performed the best with an accuracy of 94.95%, while the ResNet50 combined with discriminant analysis had the highest accuracy at 97.12%. In summary, this paper presents a highly efficient solution for experts and farmers to identify pests, and thereby prevent crop yield and economic losses by blending deep learning and machine learning [14].

This paper categorize the following sections (i)- Introduction, sections (ii)- Related Work, sections (iii) Methodology, sections (iv) - conclusion and future work

II. RELATED WORK

Accurate control of vegetable patch weeds can significantly reduce the amount of weed control needed. In this study, various deep learning object recognition methods were investigated to efficiently detect black pine bast scales from image data captured by pheromone traps. The overall performance and execution time show that the developed model has the potential to identify individual pests accurately and quickly. Further research is needed to improve the model's accuracy and inference speed. In addition, an image stitching method was used and a smartphone application was developed for efficient and convenient analysis. Accurate detection of weeds in vegetable plots is challenging due to the presence of different species of weeds at different growth stages and densities. In this paper, a deep learning-based weed detection method is presented, which classifies all other green objects as weeds while recognizing vegetable crops. Additionally, YOLO-v3's inference time was comparable to Center-Net but much faster than Faster R-CNN [15].

With the desire for better quality jute fiber rising, preventing the spread of jute plant diseases is becoming increasingly important. To combat this issue, YOLO-JD, a deep learning network, was created for detecting diseases from jute plant images. This system integrates SCFEM, DSCFEM, and SPPM - three modules dedicated to efficiently extracting image features - as well as a new image dataset containing ten different jute disease and pest classes. Compared to other similar research, YOLO-JD displayed the best results, achieving an impressive average mAP of 96.63 [16].

As the demand for higher-quality fibers has increased, disease prevention in jute plants has recently become a pressing issue. In order to identify jute diseases in images, this study developed a deep learning network called YOLO-JD. To efficiently extract image features, YOLO-JD incorporates three new modules: the Sand Clock Feature Extraction Module (SCFEM), the Deep Sand Clock Feature Extraction Module (DSCFEM), and the Spatial Pyramid Pooling Module (SPPM). In addition, ten categories of jute pests and diseases were grouped into a new large-scale image dataset. With an average mAP of 96.63, YOLO-JD has the highest detection accuracy when compared to other cutting-edge experiments [17].

The quantity of weed control required can be greatly decreased by effectively controlling weeds in vegetable patches. The presence of various weed species at various growth stages and densities makes it difficult to quickly and accurately identify weeds in vegetable fields. The weed detection method described in this paper uses deep learning to identify vegetable crops while classifying all other green objects as weeds. For YOLO-v3, CenterNet, and Faster R-CNN, the best confidence thresholds are 0.4, 0.6, and 0.4/0.5, respectively, with average precision (AP) above 97 percent on the test dataset. With an F1 score of 0.971, a precision of 0.971, and a recall of 0.970, YOLO-v3 displays the highest accuracy and computational efficiency among the deep learning architectures examined. Additionally, YOLO-v3 has an inference time that is comparable to CenterNet but significantly less than Faster R-CNN [18].

Eggplant (Solanum Melongena L.) is a seasonal vegetable that can be found in a range of sizes, shapes, and hues, including white, green, and purple. Several pests and diseases, such as fruit rot, mealybugs, and leaf beetles, pose a threat to eggplant plants. A camera can be installed to treat and monitor eggplant plants, capturing precise

and quick image data. After being digitally processed, this data can then be used by deep learning techniques for object recognition and evaluation systems in eggplant plants. For solutions and innovations, digital image processing can show how accurately colored fruit stems and leaves are. This study used YOLOv4 to select images of eggplant plants and collect image data on a mobile computing device, such as a Raspberry Pi 4. Based on the images the camera has taken, this image data will then be used to identify plant diseases, with the results being made public either right away or via the Telegram app. We will be able to identify diseases in eggplant plants and take immediate preventative action to increase yield [19].

Crop pest control is crucial for crop yield, but due to the variety of pests and their overlapping physical characteristics, it can be difficult to distinguish between them. Deep-learning based object detection algorithms, such as the YOLO detector, have recently produced excellent results by balancing accuracy and speed. YOLO is effective at detecting objects of average size, but its precision is low when dealing with tiny objects. Dealing with pest datasets, which feature multi-class distinctions and large-scale changes, results in a significant reduction in accuracy. We suggest a detector called YOLO-pest based on YOLOv4 to solve the multi-scale pest detection problem and boost the effectiveness of pest detection. For increased accuracy, this method employs a lightweight but effective mobileNetv3 backbone and a light-weight fusion feature pyramid network. The Croppest12 dataset experiments demonstrate that our improved algorithm outperforms other compared approaches [20].

Precision agriculture should involve early pest management and control because pests can result in significant crop losses. Numerous studies have investigated how to implement automatic because manual sampling and detection used in traditional pest monitoring takes a lot of time and effort. Yet very few research has focused on the automatic tracking of pest flying vegetable insect pests. To bridge this knowledge gap, this study created a method for automatically tracking flying vegetable insect pests. This strategy is supported by two hypotheses: yellow sticky traps can provide reliable information on the population density of flying vegetable insect pests, and a computer-vision-based detector can accurately identify pests in photos. To test these theories, RGB cameras, yellow sticky traps, and a dataset of manually annotated images were put in vegetable fields. The In order to evaluate the "YOLO for Small Insect Pests" (YOLO-SIP) detector, metrics for mean average precision (mAP), average mean absolute error (aMAE), and average mean square error (aMSE) were applied to the dataset. The tests showed that the proposed method successfully achieved a mAP of 84.22 percent, an aMAE of 0.422, and an aMSE of 1.126 when automatically capturing photos of yellow sticky traps. The proposed method seems promising for automatically detecting flying insect pests in vegetables, according to these results [21].

Agriculture production might suffer greatly from pest insects. According to the Food and Agricultural Organization (FAO), insect pests are responsible for 20 to 40% of crop loss annually, which lowers worldwide productivity and poses a substantial challenge to farmers. To transmit the sooty mold disease, these pests feed on the sap of various crop organs, such as leaves fruits, stems, and roots. Due to their rapid action and scalability, pesticides

are frequently used to control these pests, but less frequent use of pesticides is now advised due to environmental pollution and health concerns. Spot spraying could reduce the need for pesticides, but first it's important to identify where the pests are. This highlights the need for innovative agricultural production techniques and systems to address environmental issues, guarantee effectiveness, and encourage sustainability. In order to achieve this, it has become increasingly crucial to develop an object recognition system for the detection and classification of crop-damaging insect pests. This study suggests an automated method for identifying insect pests in digital photos and videos using an IP camera on a smartphone. This method is built on the YOLO object detection architectures, including YOLOv5 (n, s, m, l, and x), YOLOv3, YOLO-Lite, and YOLOR. The object recognition system was trained using a dataset of 7046 images that were captured in the wild under various circumstances. The system was then evaluated and compared using various parameters [22].

Accurate identification of insect pests in the larval stage is crucial for prompt treatment of infected crops, which contributes to a reduction in yield losses in agricultural products. Many image recognition problems in this field have found competitive solutions thanks to convolutional neural network (CNN)-based classification methods. This study created a brand-new pest classification technique called PCNet (Pest Classification Network), focusing on accuracy and small models appropriate for mobile devices. In order to learn the inter-channel pest information and pest positional information of input images, a coordinate attention mechanism (CA) was incorporated into the architecture of PCNet, which was constructed using the EfficientNet V2 backbone. The feature fusion module additionally combined the feature maps produced by the mobile inverted bottleneck (MBConv) and the feature maps produced by average pooling to address the loss of insect pest features in the down sampling processes. To randomly increase data diversity and prevent model overfitting, a stochastic, pipeline-based data augmentation approach was used. The experimental results revealed that the PC-Net model outperformed four lightweight CNN models (Shuffle-Net V2, Mobile-Net V3, Efficient-Net V1 and V2) and three classic CNN models (Alex-Net, VGG16, and ResNet101) in achieving a recognition accuracy of 98.4 percent on a self-built dataset consisting of 30 classes of larvae. Two additional public datasets were used to test the model's robustness [23].

A crucial first step in rice pest control is accurate identification and classification of rice pests. A new algorithm based on fully convolutional networks (FCNs) was put forth for the automatic identification and classification of rice pests in order to address this. This algorithm employed an encoder-decoder in the FCN connected by jump paths combining long jumps and short connections for precise and fine-grained insect boundary detection. One of the novel methods used in this study was the incorporation of a module using a conditional random field (CRF) for the purpose of sharpening insect contours and localizing boundaries. Additionally, a novel Dense-Net framework, with an incorporated attention mechanism (ECA), was proposed for the extraction of insect edge characteristics, which was particularly useful in the effective classification of rice pests. The proposed model underwent rigorous testing using a dataset collected specifically for this study, resulting in an

impressive 98.28% recognition accuracy. The outcomes demonstrated that the proposed model was not only faster and more accurate, but also exhibited excellent robustness in comparison with other models. The study also revealed that the proper segmentation of insect images, before classification, played a crucial role in significantly improving the performance of deep-learning-based classification systems [24].

This study aims to classify pests using the IP-102 dataset using a transfer learning approach with a fine-tuning approach. This technique was chosen for its ability to utilize features and weights from previous training processes, which reduces computation time and improves accuracy. This study tested five pre-trained CNN models: Xception, MobileNetV3L, MobileNetV2, DenseNet-201, and InceptionV3. Both fine-tuning and frozen-slicing techniques are used to further improve the quality of the generated models. Experimental results show that the fine-tuned DenseNet-201 model achieves the highest accuracy rate of 70%, while MobileNetV2 and MobileNetV3L achieve 66% and 68% accuracy, respectively. InceptionV3 and Xception achieved 67% and 69% accuracy. It was concluded that transfer learning methods with fine-tuning methods gave the best results, with the DenseNet 201 model achieving the highest accuracy of 70% [25].

Machine learning is growing in popularity as a method for the quick and accurate classification of plant diseases and pests. To customize deep learning models for the task at hand, transfer learning and deep feature extraction techniques are frequently used. In the experiment, features are extracted using pre-trained deep models, and the extracted features are then classified using RCNN (Region with Convolution Neural Network) and YOLO (You Only Look Once) networks that have been fine-tuned. It uses an adaptation of YOLO, which has shown to have a 95 percent accuracy rate for predicting pests. This study assesses potential challenges in deep learning-based plant disease and pest detection's real-world applications as well as the performance of the current research against other widely used datasets. The accuracy is assessed using data from actual infection and pest images, and the findings indicate a promising future for application [26]. YOLO is a viable strategy and comparatively quick for localization within the custom dataset [27]. Flickr detection will help avoid catastrophic disasters by alarming [28].

III. METHODOLOGY.

A. YOLO V5

The YOLO (You Only Look Once) v5 (version 3) is an object detection algorithm and a significant improvement over its predecessor, YOLO v4. The rationale behind YOLO v5 is to further enhance real-time object detection capabilities by addressing several limitations and improving upon existing features:

1). IMPROVED DETECTION ACCURACY

It is designed to improve object detection accuracy compared to YOLO v4. By introducing a more complex and deeper neural network architecture, it became better at detecting small and closely spaced objects.

2). MULTIPLE DETECTION SCALES

It is featuring pyramid network to detect objects at multiple scales. This allows it to identify objects of various sizes and significantly reduces the problem of objects appearing too small to detect.

3) USE OF DARKNET-53 BACKBONE

It is more robust and deeper convolutional neural network (CNN) called Darknet-53 as its backbone architecture. This contributes to better feature extraction and improved performance.

4) THREE DETECTION HEADS

Three separate detection heads, each responsible for detecting objects at different scales. These detection heads process features from different layers of the network, which further improves detection accuracy.

5) BOUNDING BOX ADJUSTMENTS

It can adjust the bounding boxes in real-time to better fit the shape of the detected objects, reducing false positives and increasing detection precision.

6) CLASS PREDICTION

It predicts multiple classes for each detected object, allowing it to classify objects into different categories.

7) NON-MAXIMUM SUPPRESSION

To reduce duplicate detections and improve precision, YOLO v4 uses non-maximum suppression techniques to filter out redundant bounding boxes.

8) SPEED AND REAL-TIME CAPABILITY

It maintains real-time detection capabilities, making it suitable for applications requiring fast object detection.

9) GENERAL APPLICABILITY

It can be applied to a wide range of object detection tasks, making it versatile for various use cases, including surveillance, autonomous vehicles, and robotics. Overall, the rationale behind YOLO v5 was to create an object detection algorithm that is both accurate and fast, making it well-suited for real-time applications across different domains. It aimed to address the shortcomings of its predecessors and provide state-of-the-art performance in object detection.

Pest detection entails a number of coordinated steps. To begin with, we gathered pest images for the training and assessment of DL models. Then, in order to artificially increase the sample size, we slightly altered the existing images based on specific parameters after preprocessing the entire dataset with annotation and augmentation. Using the IP-23 records as a starting point, we then used YOLO object detection to find pests. Using a split dataset for validation, we verified the fine-tuned model's detection performance and assessed the outcomes. The ideal model for use in agriculture was ultimately chosen.

B. DATASET COLLECTION

We gathered images from the Internet to train and validate our pest detection system. 7046 images across 23 major categories by searching different databases and search engines like Kaggle, Google, Baidu, Iostock, Dream, Flickr, and Bing. A representative sample of the dataset is displayed below. The images were scaled to a uniform size of 640x640. Following that, we separated the images into 8 distinct classes, as seen in the figure.
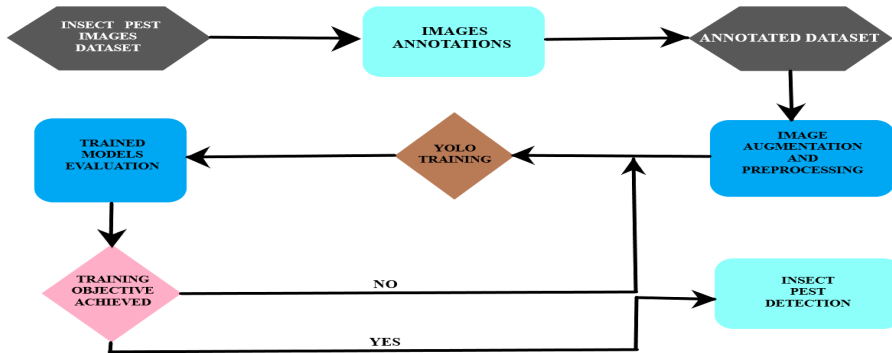
FIGURE 1. Methodology

FIGURE 2. Image scaling and Classes

C. DATA ANNOTATION

Annotating the images is part of the preprocessing done on them prior to training a DL model. During the annotation process, significant traits of various pest species are noted in the images and assigned labels. Due to the fact that feature labeling allows the computer to learn the features from the images, its accuracy is essential to the training model's accuracy. Using the open-source program labeling, which graphically labels images, we localized and saved the positions and classes of pests in the images as text files. Each bounding box's object class, height, and width are listed in the text file on a

separate line with coordinates normalized between 0 and 1. Finally, we processed the training set's data for data refinement in order to expand the sample size and extract the characteristics of pests that fit into various labelled categories. This helped us avoid over-fitting the trained model.

D. DATA AUGMENTATION

The challenge of gathering large amounts of data for training purposes frequently results in the issue of insufficient data volume in data analysis. This may result in the DL model being over fitted, which can be avoided by increasing the training examples. Over fitting can also be avoided

Department of Information Systems

by using data augmentation methods like geometric transformations. This study demonstrated the use of data augmentation techniques in the detection of pests by applying geometric transformations to pest images, including rotation, horizontal flip, hue, blur, and saturation.

The IP-23 dataset includes 7046 samples from 23 different pest classes. We divided the dataset into a training set and a validation set with a ratio of 7:3 at the class level to train the pest detection model. As a result, the detection task generated 2113 validation images and 4933 training images as mentioned in table I.

TABLE I

DATASET DISTRIBUTION TRAINING AND VALIDATION

Total Images	Classes	Training	validation	Split ratio
7046	23	4933	2113	70:30

III. RESULTS AND FINDINGS

The simulation process involves two distinct phases, namely, the cloning of YOLO v5 and the subsequent training phase, which spans an extensive duration. Python commands and environment is set to Google Colab with the following libraries and parameters.

The PyTorch library is a widely acclaimed open-source machine learning framework known for its versatile applicability in a range of deep learning endeavors, encompassing tasks such as the training and deployment of neural networks. Additionally, the inclusion of 'Image' from the IPython.display module serves as a common practice to facilitate image display within coding environments. Furthermore, the command git clone https://github.com/ultralytics/yolov5' signifies an endeavor to replicate the YOLOv5 repository, a prominent and widely recognized model for object detection, from its original source on GitHub. This repository holds significant importance in the field of computer vision and has garnered extensive attention for its capabilities in object detection tasks. torch.cuda.get_arch_list(): This line is attempting to retrieve the list of GPU architectures that are supported by PyTorch on the current system. It can be useful to check the available GPU architectures when dealing with GPU-related computations.

A. TRAINING THE YOLOV5 MODEL

The model is configured with specific settings essential for the training process, including parameters like image size, batch size, and the number of epochs, dataset configuration, initial weights, and the directory for saving results. This setup is driven by a set of hyper parameters employed during training. To initiate the training process for a YOLOv5 model, a Python script named 'train.py' is executed. The input image size for training is set to 416 pixels, which dictates the dimensions to which training images are resized. The batch size for training is established at 30, determining the number of images processed in each iteration during training. The model is trained over 150 epochs, with each epoch representing a complete pass through the training dataset. The data configuration is specified in a data.yaml file, which includes critical dataset information such as the paths to training and validation images and the number of classes. Pre-trained model weights, 'yolov5s.pt', are utilized to initiate training. Caching is enabled using the 'cache' flag, which accelerates the training process by storing preprocessed data. Lastly, the results and logs from training are saved in the project directory '/content/drive,

encompassing critical training metrics, model checkpoints, and other pertinent files. This comprehensive setup forms the basis for the YOLOv5 model's training process.

1)TF LITE MOBILE EMBEDDED DEVICES

 # !python export.py --weights /content/drive/MyDrive/INSECTOUTPUT/ exp4/weights/best.pt --include tflite --img 416 --data /content/yolov5/insect1-2/data.yaml

Command is used to convert pre-trained YOLOv5 model weights into a Tensor Flow Lite model. Tensor Flow Lite is a framework for deploying machine learning models on mobile and embedded devices, and it is often used when you need to run a model on resource-constrained platforms. The specified input image size and data configuration help tailor the TF Lite model for specific use case and dataset. User typically run this command in an environment with the necessary dependencies and access to the pre-trained YOLOv5 weights.

To show the results of our tests more clearly, we arbitrarily chose some pictures from the test set. The YOLOv5x has good recognition performance for the various insect pest categories due to the convolutional neural network algorithm's capacity to boost generalization and eliminate the need for human feature extraction as mentioned in (Figure 17).
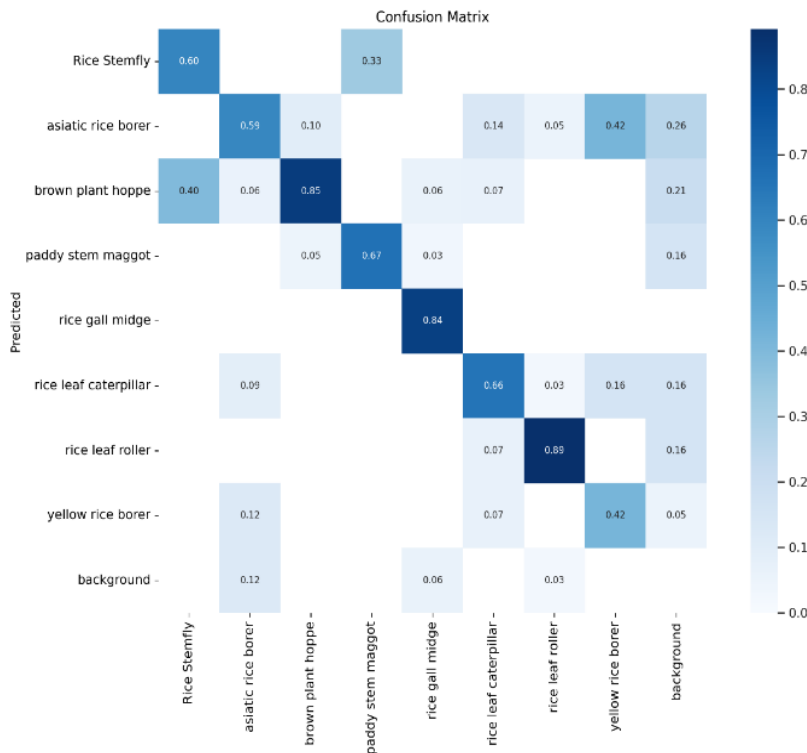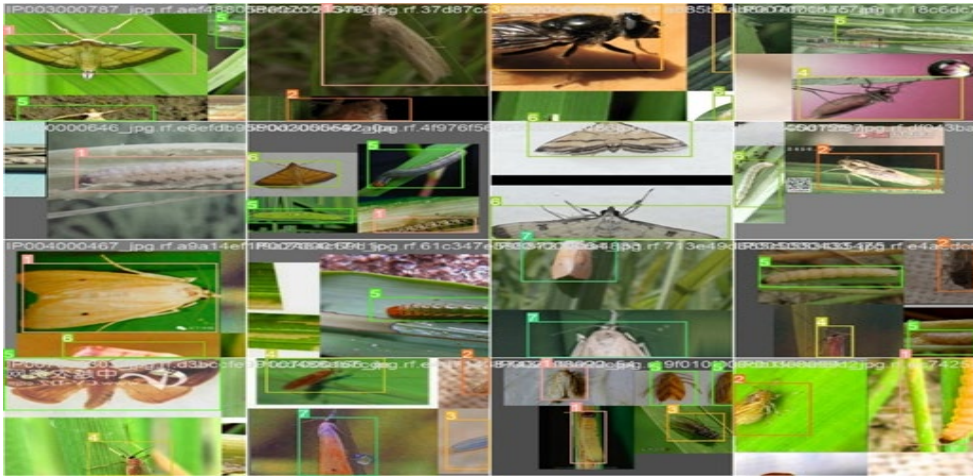

FIGURE 3. Confusion matrix

Department of Information Systems

FIGURE 4. Training batch 1

FIGURE 5. Training batch 1
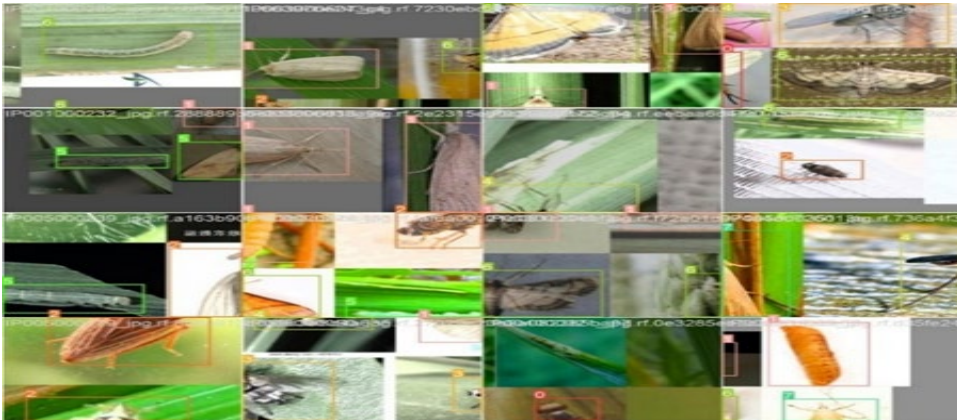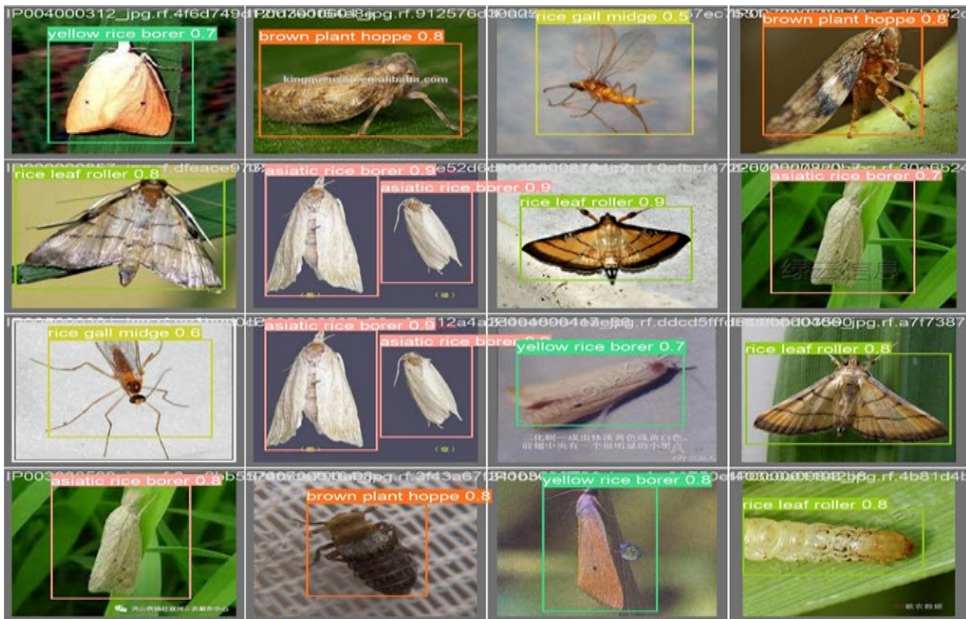

FIGURE 6. Training batch 1

FIGURE 7. Training batch 2

FIGURE 8. Training batch 2

FIGURE 9. Training Batch 2

**FIGURE 10.** Validation batch prediction

Image gathering for pest identification and categorization was the initial stage of this project. For one of the pests, there is no official data set. A system for automatically capturing images was created. Based on search terms and picture backdrops, this method searches and downloads images using the Google image engine. For each of the aforementioned pests, I made a picture. In English, Portuguese, and German, we searched for beetles using their common and scientific names. The expected background is green vegetation; thus, a search with a green background filter was also run.

Data gathering is sped up by automatic download, although some photographs don't correspond to pests. This required manual filtering by an expert after receiving all the photographs for each bug. Images that are not related to or classified as objects of interest are excluded by this filter. Additionally, several photographs were cropped since they contained extraneous data that would have harmed the training algorithm's effectiveness. The final dataset was produced after the filtering procedure. Because we are employing a supervised approach, you should be aware that each image has a label that corresponds to the specified class.

Limited pest species and small specimen image sets collected in a laboratory environment, pest identification Approaches based on manual feature design can give high accuracy rates for pest identification. However, pest images from real fields have complex backgrounds, and there are different pest poses (interspecies difference, interspecies similarity, and many non-target insects). Under the bio-diverse and complex field conditions, Traditional feature-based identification methods show poor robustness and generalization.
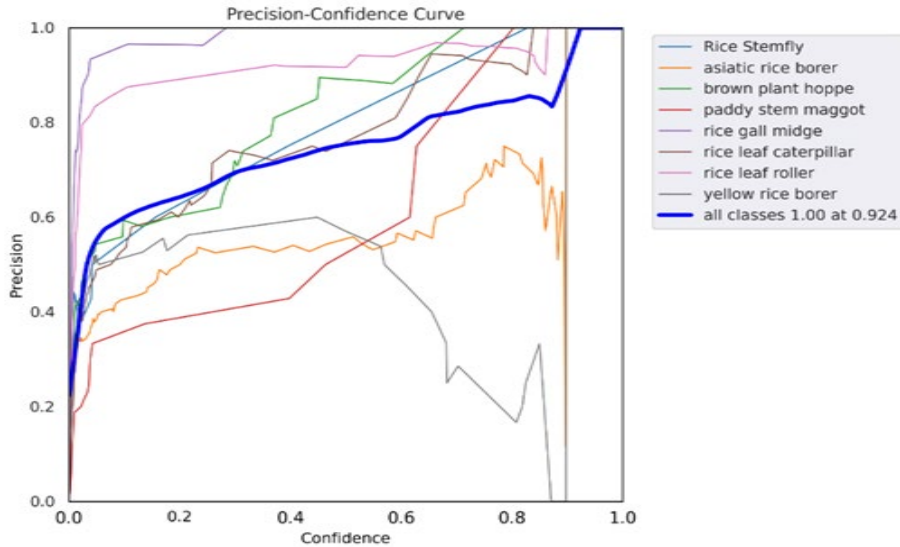
FIGURE 11. Precision confidence curve
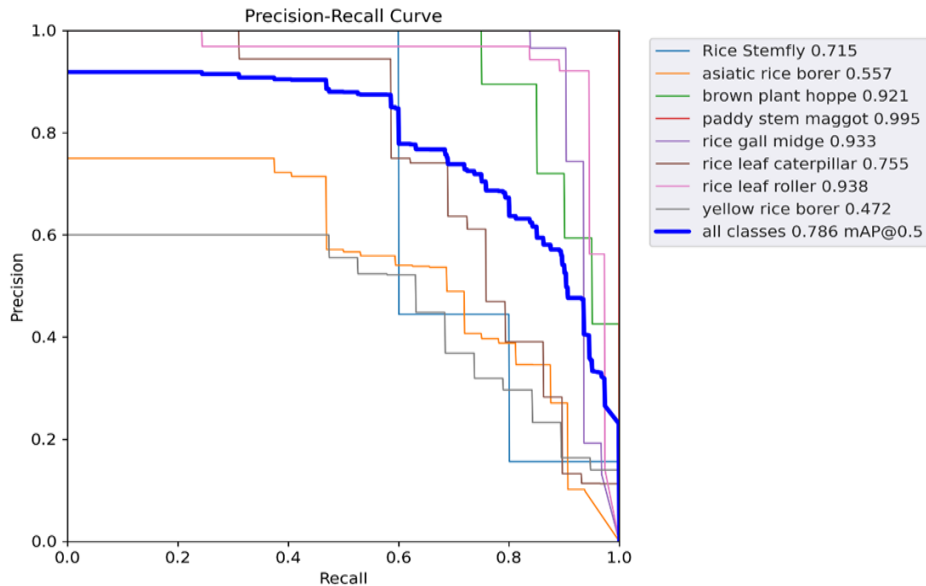
FIGURE 12. Precison recall curve

The Yolo model delivers state-of-the-art performance for a variety of computer vision tasks.

Various supervised DCNN models were implemented to accomplish the task of identifying pests, and the results were compared in the next section. First, a general CNN architecture containing the

same convolutional layers as the retrained VGG-

16 model was used as the basis. We then implemented his YOLO model using his VGG-16 architecture, which is the same as the shared layer.

An attention mechanism was implemented in these models to visualize areas where the algorithm resides. Stay focused during and after training. Finally, we tested a semantic segmentation network (Seg-Net) model to determine location and extent in addition to classifying the defined pests in the image provided. These approaches are shown below.
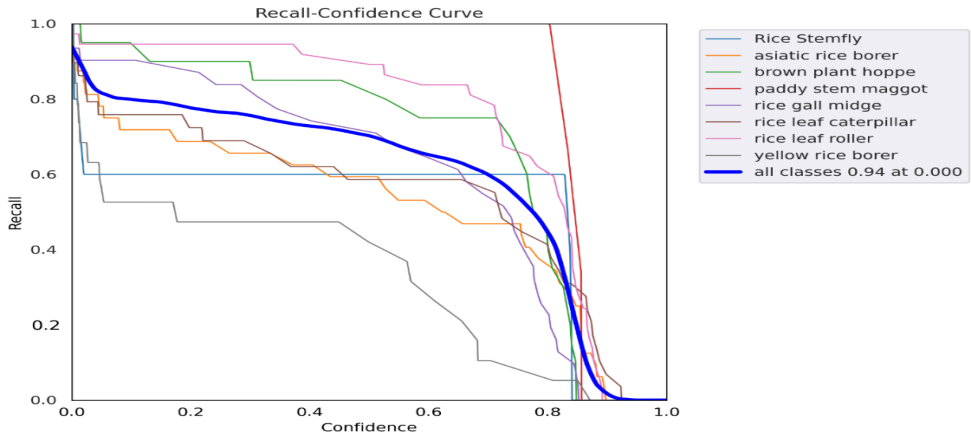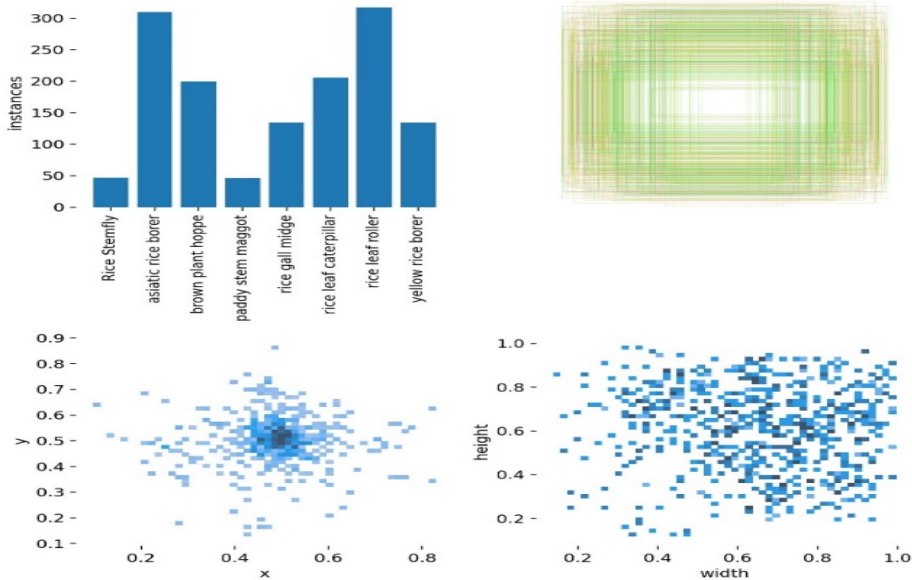

FIGURE 13. Recall confidence curve

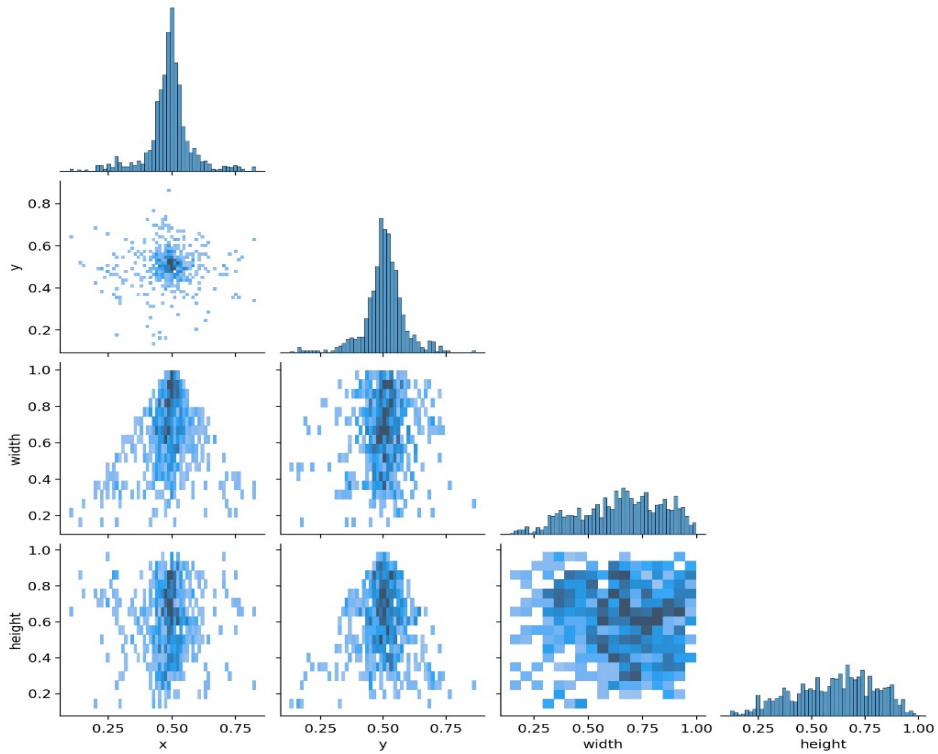**FIGURE 14.** Labels and class interval

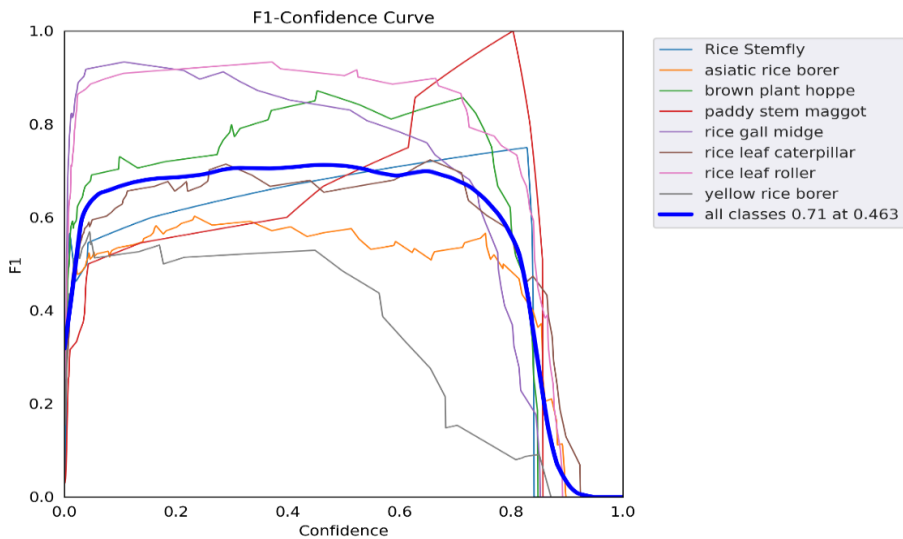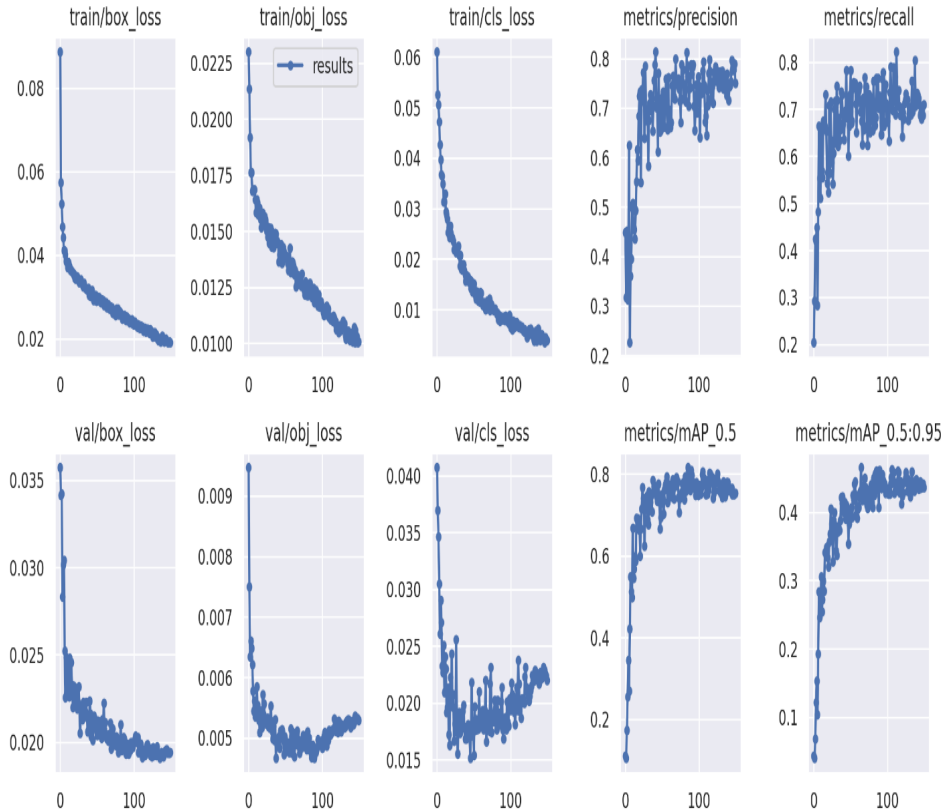**Figure 15.** Labels correlogram

FIGURE 16. F1-confidence Interval

FIGURE 17. Loss of training box

mAP (0.5), mAP (0.95)**: mAP stands for Mean Average Precision, a common metric used in object detection and other tasks. The "(0.5)" and "(0.95)" refer to different thresholds for considering a prediction correct. mAP (0.5) is the mean average precision calculated with an IoU (Intersection over Union) threshold of 0.5, while mAP (0.95) uses a stricter threshold of 0.95. These metrics evaluate the model's ability to make precise and accurate predictions.

IV. CONCLUSION AND FUTURE WORK

In the agricultural domain, pests have been a major concern for farmers worldwide. These pests cause damage in fields, leading to significant losses in crop yield. To mitigate this problem, accurate and early detection of pests is necessary using plant images. Although researchers have been working on accurate detection, there is still room for improvement due to several factors such as noise, illumination, occlusion,

etc. To address this issue, we propose the use of an end-to-end deep learning YOLOv5 model based on different features of pests. Our model provides a significant improvement in localization results,

offering 0.83 Precision, 0.83 Recall, 0.833 mAP-0.5, 0.783 mAP-0.5:0.95 compared to existing methods. In this article, we demonstrate the effectiveness of the YOLOv5 model for pest detection and classification. We believe that our work can be extended for multi-classification of complex medical images in the future. We encourage researchers in the field to consider our proposed model for pest detection in agricultural settings.

REFERENCES.

[1] H. Venthur and J.-J. Zhou, "Odorant receptors and odorant-binding proteins as insect pest control targets: A comparative analysis," *Front. Physiol.*, vol. 9, Art, no. 1163, Aug. 2018, doi: https://doi.org/10.3389/fphys.2018.01163

[2] T. Saranya, C. Deisy, S. Sridevi, K. S. Muthu, and M. K. A. A. Khan, "Performance Analysis of first order optimizers for plant pest detection using deep learning," in *Mach. Learn. Image Process. Net. Secur. Data Sci.*, N. Khare, D. S. Tomar, M. K. Ahirwal, V. B. Semwal, and V. Soni, Eds. Jan. 2023, pp. 37–52. doi: https://doi.org/10.1007/978-3-031-24367-7_4

[3] S. Mascarenhas and M. Agarwal, "A comparison between VGG16, VGG19 and ResNet50 architecture frameworks for image classification," in *Int. Conf. Disrupt. Technol Multi-Discipl. Res. Appl.*, Bengaluru, India, Nov. 2021, pp. 96–99. doi: https://doi.org/10.1109/CENTCON52345.2021.9687944

[4] Z. Anwar and S. Masood, "Exploring deep ensemble model for insect and pest detection from images," *Proc. Comput. Sci.*, vol. 218, pp. 2328–2337, 2023, doi: https://doi.org/10.1016/j.procs.2023.01.208

[5] "Residual networks (ResNet) - deep learning." GeeksforGeeks. https://www.geeksforgeeks.org/residual-networks-resnet-deep-learning/ (accessed Feb. 26, 2023).

[6] R. Alake, "Deep learning: GoogLeNet explained." Medium.com. https://towardsdatascience.com/deep-learning-googlenet-explained-de8861c82765 (accessed Feb. 26, 2023).

[7] R. Hadipour-Rokni, E. A. Asli-Ardeh, A. Jahanbakhshi, I. E. paeen-Afrakoti, and S. Sabzi, "Intelligent detection of citrus fruit pests using machine vision system and convolutional neural network through transfer learning technique," *Comput. Biol. Med.*, vol. 155, Art. no. 106611, Mar. 2023, doi: https://doi.org/10.1016/j.compbiomed.2023.106611

[8] X. Fu *et al.*, "Crop pest image recognition based on the improved ViT method," *Info. Proc. Agricul*. Feb. 2023, doi: https://doi.org/10.1016/j.inpa.2023.02.007

[9] J. K. Lugemwa, "An embedded, machine learning-enabled platform for in-field screening of plant disease and pest damage.," Ph.D. dissertation, Makerere Univ., Kampala, Uganda., 2023. [Online]. Available: http://dissertations.mak.ac.ug/handle/20.500.12281/14870

[10] S. D. Meena, M. Susank, T. Guttula, S. H. Chandana, and J. Sheela, "Crop yield improvement with weeds, pest and disease detection," *Proc. Comput. Sci.*, vol. 218, pp. 2369–2382, 2023,

doi: https://doi.org/10.1016/j.procs.2023.01.212

[11] R. Akter, M. S. Islam, K. Sohan, and M. I. Ahmed, "Insect recognition and classification using optimized densely connected convolutional neural network," in *12th Int. Conf. Info. Sys. Adv. Technol.*, M. R. Laouar, V. E. Balas, B. Lejdel, S. Eom, and M. A. Boudia, Eds. 2023, pp. 251–264. doi: https://doi.org/10.1007/978-3-031-25344-7_23

[12] H. Waheed, N. Zafar, W. Akram, A. Manzoor, A. Gani, and S. Islam, "Deep learning based disease, pest pattern and nutritional deficiency detection system for 'Zingiberaceae' crop," *Agriculture*, vol. 12, no. 6, Art. no. 742, May 2022, doi: https://doi.org/10.3390/agriculture12060742

[13] D. Mondal, K. Roy, D. Pal, and D. K. Kole, "Deep learning-based approach to detect and classify signs of crop leaf diseases and pest damage," *SN Comput. Sci.,* vol. 3, no. 6, Art. no. 433, Aug. 2022, doi: https://doi.org/10.1007/s42979-022-01332-5

[14] M.-L. Huang, T.-C. Chuang, and Y.-C. Liao, "Application of transfer learning and image augmentation technology for tomato pest identification," *Sustain. Comput. Inform. Syst.*, vol. 33, Art. no. 100646, Jan. 2022, doi: https://doi.org/10.1016/j.suscom.2021.100646

[15] W. Yun, J. P. Kumar, S. Lee, D.-S. Kim, and B.-K. Cho, "Deep learning-based system development for black pine bast scale detection," *Sci. Rep.*, vol. 12, no. 1, Art. no. 606, Jan. 2022,

doi: https://doi.org/10.1038/s41598-021-04432-z

[16] D. Li, F. Ahmed, N. Wu, and A. I. Sethi, "YOLO-JD: A deep learning network for jute diseases and pests detection from images," *Plants*, vol. 11, no. 7, Art. no. 937, Mar. 2022, doi: https://doi.org/10.3390/plants11070937

[17] L. Xinmao, L. Yihui, X. Mingl, T. Shuijiaol, and M. Zhandong, "Research on identification of main cotton pests based on deep learning," in *IEEE 2nd Int. Conf. Digital Twins Parallel Intell.*, Boston, MA, USA, Oct. 2022, pp. 1–4. doi: https://doi.org/10.1109/DTPI55838.2022.9998883

[18] X. Jin, Y. Sun, J. Che, M. Bagavathiannan, J. Yu, and Y. Chen, "A novel deep LEARNING-BASED method for detection of weeds in vegetables," *Pest Manag. Sci.*, vol. 78, no. 5, pp. 1861–1869, May 2022, doi: https://doi.org/10.1002/ps.6804

[19] S. W. Nasution and K. Kartika, "Eggplant disease detection using yolo algorithm telegram notified," *Int. J. Eng. Sci. Inf. Technol.*, vol. 2, no. 4, Art. no. 4, Dec. 2022, doi: https://doi.org/10.52088/ijesty.v2i4.383

[20] S. Dong, J. Zhang, F. Wang, and X. Wang, "YOLO-pest: a real-time multi-class crop pest detection model," in I*nt Conf Comput Appl Info Secur.*, Wuhan, China, May 2022, doi: https://doi.org/10.1117/12.2637467

[21] Q. Guo, C. Wang, D. Xiao, and Q. Huang, "Automatic monitoring of flying vegetable insect pests using an RGB camera and YOLO-SIP detector," *Precis. Agric.*, vol. 24, pp.

436–457, Sep. 2022, doi: https://doi.org/10.1007/s11119-022-09952-w

[22] I. Ahmad et al., "Deep learning based detector YOLOv5 for identifying insect pests," *Appl. Sci.*, vol. 12, no. 19, Art. no. 10167, Oct. 2022, doi: https://doi.org/10.3390/app121910167

[23] T. Zheng, X. Yang, J. Lv, M. Li, S. Wang, and W. Li, "An efficient mobile model for insect image classification in the field pest management," *Eng. Sci. Technol. Int. J.*, vol. 39, Art. no. 101335, Mar. 2023, doi: https://doi.org/10.1016/j.jestch.2023.1 01335

[24] H. Gong *et al.*, "Based on FCN and DenseNet framework for the research of rice pest identification methods," *Agronomy*, vol. 13, no. 2, Art. no. 410, Jan. 2023, doi: https://doi.org/10.3390/agronomy130 20410

[25] A. P. Syahputra, A. C. Siregar, and R. W. S. Insani, "Comparison of CNN models with transfer learning in the classification of insect pests," IJCCS Indones. *J. Comput. Cybern. Syst.*, vol. 17, no. 1, Art. no. 103, Feb. 2023, doi: https://doi.org/10.22146/ijccs.80956

[26] M. Sujaritha, M. Kavitha, and S. Roobini, "Pest detection using improvised YOLO architecture," in *Comput. Vision Mach. Intell. Parad. SDGs*, vol. 967, R. J. Kannan, S. M. Thampi, and S.-H. Wang, Eds. Jan. 2023, pp. 59–67. doi: https://doi.org/10.1007/978-981-19-7169-3_6

[27] H. Zaki, M. Shaikh, M. Tahir, M. Naseem, and M. Khan, "Smart surveillance and detection framework using YOLOv3 algorithm", *PakJET*, vol. 5, no. 4, pp. 36–43, Dec. 2022.

[28] M. K. Shaikh, S. Palaniappan, and T. Khodadadi. An AI-driven automotive smart black box for accident and theft prevention. *Int. J. Model. Ident. Cont.*, vol. 39, no. 4, pp. 332–339, June 2020, doi: https://doi.org/10.1504/IJMIC.2021.1 23800