# Enhanced Camouflaged Object Detection for Agricultural Pest Management: Insights from Unified Benchmark Dataset Analysis

Syed Muhammad Hussain*
*School of Science and Engineering*
*Habib University*
Karachi, Pakistan
0009-0000-4613-845X

Samiya Ali Zaidi*
*School of Science and Engineering*
*Habib University*
Karachi, Pakistan
0009-0008-1907-1542

Afsah Hyder*
*School of Science and Engineering*
*Habib University*
Karachi, Pakistan
0009-0008-1758-490X

Syed Muhammad Ali Rizvi*
*School of Science and Engineering*
*Habib University*
Karachi, Pakistan
0009-0000-6756-5530

Muhammad Farhan
*School of Science and Engineering*
*Habib University*
Karachi, Pakistan
0000-0002-8244-8313

***Abstract*—In recent decades, the severity of climate change has led to a rise in the frequency of agricultural pest attacks on farms causing significant economic damage and food shortages. Effective management of pests, specifically Camouflaged pests, poses significant challenges in agriculture, requiring accurate automated detection and segmentation. In this study, we leverage state-of-the-art object detection and segmentation models, specifically the single-stage YOLOv8 model, fine-tuned on a large-scale Unified Benchmark Camouflaged Object Detection Dataset (UBCODD) consisting of 52,447 images. Furthermore, we extend our analysis to benchmark agricultural pest datasets such as IP-102 and the Locust-Mini Dataset, showcasing competitive performance metrics. This integrated approach allows us to capture agricultural camouflaged pests with greater detail and accuracy. Our findings lay the groundwork for the advancement of single-stage object detectors and segmentation models in the field of agriculture. Moreover, we contribute to open-source initiatives in agricultural technology by generating bounding box annotations for the entire IP-102 and binary masks for the Agricultural Pests Image Dataset. This research signifies a significant advancement in agricultural pest recognition and segmentation using cutting-edge computer vision technologies.**

***Index Terms*—Camouflaged Pests, Camouflaged Object Detection, Pest Detection, UBCODD, Camouflaged Segmentation**

## I. Introduction

Insect pests pose a relentless threat to agricultural production, jeopardizing both the quality and quantity of crops. This issue leads to significant financial losses for agricultural nations, especially in regions like Asia, Africa, and South America, where undetected infestations can notably impact food security [1]–[7]. Traditional pest detection methods, reliant on manual identification by agricultural experts, are time-consuming, subjective, and costly, particularly for large farms [8]. Hence, there is an need for more efficient, automated, and objective solutions for pest detection.

Advancements in computer vision offer promising alternatives for insect pest recognition systems. These systems have the potential to revolutionize agricultural practices by significantly improving detection efficiency and overcoming limitations associated with manual detection [7], [9], [10]. However, a critical aspect of robust automated pest detection lies in addressing the diverse camouflage capabilities of agricultural pests, which pose challenges in real-world agricultural settings along with variations in lighting, viewpoint, and environmental complexity.

Many agricultural pests have camouflage abilities, which they employ as a survival technique [11]. Agricultural pests often blend seamlessly with their surroundings, making them difficult to identify, as can be seen in figure 1. In such scenarios requiring real-time camouflaged object detection, a lightweight model is highly beneficial. We fine-tuned YOLOv8 detection and segmentation models using the Unified Benchmark Camouflaged Object Detection Dataset (UBCODD). These fine-tuned models were then applied to benchmark agricultural pest datasets, achieving notable results in pest detection and segmentation.

This paper makes the following contributions to the field of agricultural camouflage pest detection and segmentation:

- Creation of a Unified Benchmark Camouflaged Object Detection Dataset (UBCODD): We combined major COD datasets and prepared ground truth labels in YOLO format for detection and segmentation.
- Fine-tuning of YOLOv8 Models: We fine-tuned YOLOv8 models for both detection and segmentation tasks on UBCODD, achieving comparable benchmark accuracies

All* authors contributed equally to this research.

while enabling near real-time inference capabilities for detecting and segmenting camouflaged pests.

- Evaluations on Agricultural Pest Datasets: The optimized models were rigorously evaluated on benchmark agricultural pest datasets, including IP102 [12] and the Locust Mini-Dataset [13], establishing new benchmarks for detection performance in the agricultural sector.
- Open-Source Contributions: In support of collaborative research and innovation, the UBCODD dataset, including ground truth labels and pre-trained model weights, are released. Furthermore, bounding box annotation for IP102 dataset [12] and binary masks for the Agricultural Pests Image dataset [14] were generated to facilitate further research and qualitative analyses in pest detection.

These initiatives aim to fill current research gaps and offer practical, technologically advanced solutions for enhancing agricultural pest management strategies.

Fig. 1: Sample images from the IP-102 dataset depicting pests camouflaging with their surroundings [12].

The paper is organized as follows: Section II reviews literature in the fields of agricultural pest recognition and COD. Section III outlines the process of creating the UBCODD dataset, including label generation, model selection, training optimization, and evaluation methodology on benchmark agricultural datasets. Section IV presents our findings on COD and agricultural datasets. Finally, Section V concludes the paper and discusses our open-source contributions.

## II. Related Works

In this section, we provide an overview of existing research on agricultural pest detection and camouflaged object detection and explore the interrelation between these two domains.

### *A. Agricultural Pest Detection*

Liu et al. proposed EG-PraNet, an improvement on PraNet [13], which utilizes the Group Reverse (GR) module for efficient feature extraction and incorporates data augmentation to enhance model robustness in detecting locusts. Similarly, Zhang et al. [15] addressed low efficiency and unreliable cotton pest detection by employing ECA for complex backgrounds, focal loss for imbalanced samples, and hard swish activation for improved performance. A two-stage pest detection method was presented by Chen et al. [16], with a context feature enhancement module, the region of interest (RoI) feature fusion module, and the non-task separation module, achieving a high mAP of 72.7% on the SimilarPest5 dataset. Additionally, Yang et al. [17] achieved reduced model complexity and faster detection for maize pest identification by replacing network modules with efficient alternatives, demonstrating the potential for lightweight models in agriculture.

Much of the current research in agricultural pest detection targets specific pest categories [13], [15]–[17]. While this approach has yielded valuable advancements, it limits the applicability of these methods to the vast diversity of agricultural pests encountered in real-world scenarios. A crucial, unifying characteristic is the ability of agricultural pests to camouflage themselves within their environment [11]. Recognizing this shared property, several curated pest detection datasets, such as Locust-mini [13], leveraged relevant images from established camouflaged object detection benchmarks like COD10K [18] and CAMO [19]. This reliance on general camouflage object datasets underscores the significance of addressing pest detection through the lens of camouflage.

### *B. Camouflage Object Detection*

Various methods have been explored to achieve accurate camouflaged object segmentation. One approach involves designing novel network modules [19], [20]. Le et al.'s Anabranch Network use a dual-branch approach to improve both object identification and segmentation accuracy by combining classification and segmentation tasks [19]. Zhou et al. proposed the Feature Aggregation and Propagation Network (FAP-Net), incorporating the Boundary Guidance Module (BGM) to enhance segmentation accuracy [20].

Another strategy leverages multi-task learning frameworks by incorporating auxiliary tasks alongside the primary segmentation task [21]–[23]. For instance, Lamdouar et al. introduced a registration module highlighting object boundaries and a segmentation module for detecting moving objects [21]. Another approach involves using an MGL model to effectively leverage high-order relations of graphs to locate camouflaged objects and enhance their boundaries [22]. Additionally, Zhong et al. innovatively applied frequency domain to enhance segmentation accuracy through a frequency enhancement module [23].

Although advancements in segmentation models produce commendable accuracy yet achieving a balance between accuracy and real-time performance is necessary. This prioritization necessitates models optimized for fast inference. For this purpose, we used a lightweight single-staged object detector and segmentation model, YOLOv8. By leveraging a comprehensive and intricate training dataset, we achieved significant results for agricultural pest detection through fine-tuning COD datasets.

## III. METHODOLOGY

This section provides an overview of the methodology used in our research, detailing the systematic approach to dataset unification, model fine-tuning, and adaptations for agricultural pest detection.

### A. UBCODD: a Unification of the Benchmark Datasets

This subsection outlines the creation of the Unified Benchmark Camouflaged Object Detection Dataset (UBCODD), formed by combining five benchmark COD datasets [18], [19], [21], [24], [25]. This integration enhances the diversity and complexity of camouflage scenarios for training and evaluation, which is essential for improving detection accuracy in real-world applications.

TABLE I: Dataset Information.

| Dataset | Number of Images |
|---|---|
| COD10K [18] | 10000 |
| MoCA [21] | 37000 |
| NC4K [24] | 4121 |
| CHAMELEON [25] | 76 |
| CAMO [19] | 1250 |

UBCODD, comprising 52,447 images, represents a comprehensive resource for developing and evaluating camouflage detection and segmentation models. Table I details the datasets integrated into UBCODD.

UBCODD is organized into three main categories: training, validation, and testing, ensuring its suitability for model development and evaluation. Figure 2 represents the process for generating YOLO format labels from binary masks, aiding in both object detection and segmentation tasks. This process streamlines the dataset's use for training and evaluation tasks. UBCODD represents a significant advancement in the field of camouflage object detection, providing researchers with a solid foundation for advancing the SOTA models in this domain.

### B. Fine-Tuning the YOLOv8 Object Detection & Segmentation Models

This subsection evaluates YOLOv8, a highly recognized model known for its efficiency in real-time object detection and segmentation tasks [26]. YOLOv8's lightweight design, precision, and fast inference make it particularly suitable for real-time applications. Its performance is assessed using the MS COCO dataset, a widely recognized benchmark [27].

To adapt YOLOv8 to the specific domain of camouflage object detection and segmentation, it was fine-tuned with the UBCODD dataset, utilizing different variants such as YOLO XLarge, Medium, and Nano. This involved initializing YOLOv8 with pre-trained weights from MS COCO and then training it further on UBCODD. For object detection training, a batch size of 32 was employed over 150 epochs, with testing conducted on the COD10K test dataset. For segmentation, a batch size of 16 was maintained over 100 epochs, with evaluations on the COD10K, CAMO, and CHAMELEON test datasets to compare against other SOTA models.

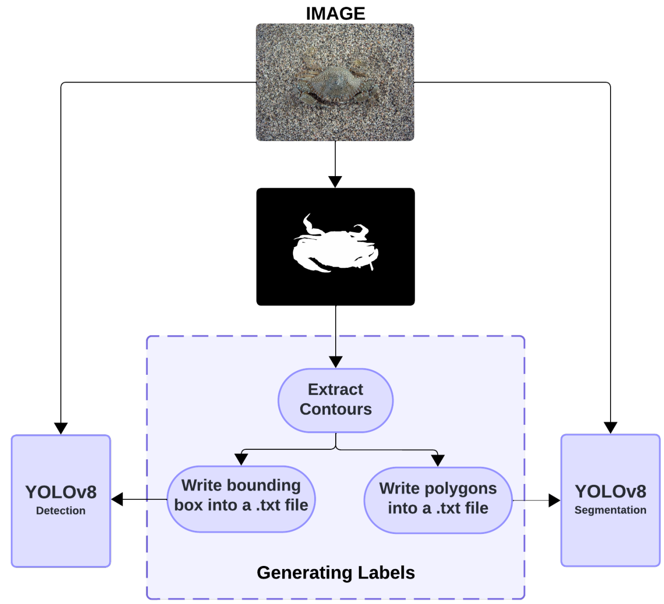


Fig. 2: The process of label creation followed by the subsequent integration of both the labels and the original image into the model for training, for both detection and segmentation.

To ensure model stability and prevent overfitting, a patience level of 25 was implemented. If the validation loss failed to improve for 25 consecutive epochs, the training process was terminated. This approach aimed to enhance YOLOv8's performance in detecting and segmenting camouflaged objects within the UBCODD dataset. Experiments were conducted using an NVIDIA TITAN RTX 24 GB GPU.

The fine-tuning efforts were particularly focused on adapting YOLOv8 for accurate detection and segmentation of agricultural pests, showcasing the model's applicability to challenges in agriculture camouflage pest detection.

### C. Fine-Tuning UBCODD's YOLO Model for Agricultural Pest Detection and Segmentation

Insect pests significantly impact agricultural product yield, necessitating accurate recognition for timely preventive measures to mitigate economic losses. Building upon the effectiveness demonstrated by our YOLOv8 model on the challenging UBCODD dataset, we aimed to assess its applicability to agricultural pest detection and segmentation tasks. To achieve this, we incorporated two additional benchmark datasets: IP102 [12], and Locust-Mini [13].

*1) IP102: A Large-Scale Benchmark Dataset for Insect Pest Recognition:* IP102 dataset [12], encompassing over 75,000 images across 102 pest categories, exhibiting a natural long-tailed distribution. Approximately 19,000 images are annotated with bounding boxes for object detection. IP102 adopts a hierarchical taxonomy, grouping insect pests affecting specific agricultural products into the same upper-level category. Leveraging our UBCODD fine-tuned YOLO detection model, we apply it to detect the provided 19,000 labels on the dataset, further fine-tuning it on the training dataset. We evaluate the model's performance on detection standard metrics and compare the results with other SOTA models on the test dataset.

*2) Locust-Mini Dataset:* Locusts pose a serious threat to global food security, capable of causing widespread damage and economic loss through massive, migratory swarms. Traditional detection methods face challenges due to their adeptness at camouflage.

To address this challenge, we directly evaluate our fine-tuned detection and segmentation models on the Locust-Mini [13] test dataset. As the training dataset is not publicly available, we utilize the test dataset comprising diverse images of locusts, including camouflaged ones sourced from benchmark datasets like COD10K and the internet. We assess the models' performance on the defined evaluation metrics and compare it with other SOTA models for benchmarking purposes in case of segmentation task.

## IV. Results & Analysis

In this section, we define standard evaluation metrics for object detection and segmentation. We then present results from training various YOLOv8 models (xlarge, medium, nano) on the UBCODD dataset and compare them against SOTA models. Finally, we apply these fine-tuned models to an agricultural pest detection and segmentation benchmark dataset to evaluate their generalization capability and performance against SOTA models in this domain.

### A. Evaluation Metrics

*1) Object Detection:* The following metrics are used for evaluating the performance of object detection:

- Precision: Ratio of true positives to the total number of positive predictions.

$$\text{Precision} = \frac{Tp}{Tp + Fp} \tag{1}$$

- Recall: Proportion of true positives correctly identified by the model.

$$\text{Recall} = \frac{Tp}{Tp + Fn} \tag{2}$$

- Mean Average Precision (mAP): Average of the precision values calculated for each class.
- mAP[0.5:0.95]: Mean Average Precision over different Intersection over Union (IoU) thresholds, from 0.50 to 0.95.

*2) Object Segmentation:* Below are the primary metrics used to evaluate the performance of object segmentation:

- Structure Measure ($S_\alpha$) [28]: It aims to gauge the structural similarity between the regional perception ($S_r$) and object perception ($S_o$). It is defined by

$$S_\alpha = \alpha \times S_o + (1 - \alpha) \times S_r \tag{3}$$

where $\alpha \in [0, 1]$ is a trade-off parameter and it is set to 0.5 as default.

- Enhanced-Alignment Measure ($\alpha E$) [29]: Assesses both pixel-level similarity and image-level statistics concurrently, aligning with human visual perception. It is defined by

$$\alpha E = \frac{1}{w \times h} \sum_{x=1}^{w} \sum_{y=1}^{h} \phi_{FM}(x, y) \tag{4}$$

where $\phi_{FM}$ is the enhanced alignment matrix of the foreground map; and $h$ and $w$ are the height and the width of the map, respectively.

- Weighted F-Measure (wF): This represents a harmonic mean of precision and recall, as in:

$$wF = \frac{(1 + \beta^2) \times \text{Precision} \times \text{Recall}}{\beta^2 \text{Precision} + \text{Recall}} \tag{5}$$

The weighted F-measure allocates distinct weights, $\beta$, to precision and recall, a useful feature when one of these metrics is deemed more significant than the other.

- Mean Absolute Error (M): is the average pixel-level relative error between the ground truth ($G$) and the normalized prediction ($S$), defined as:

$$M = \frac{1}{W \times H} \sum_{i=1}^{W} \sum_{j=1}^{H} |S(i, j) - G(i, j)| \tag{6}$$

Both the ground truth and normalized prediction are normalized to $[0, 1]$.

### B. YOLOv8 Model Results: Object Detection and Segmentation on Benchmark Datasets

The YOLOv8 model represents a SOTA single-stage model [26]. In this subsection, we present the results obtained by various versions of the YOLOv8 for both detection (v8x, v8m, v8n) and segmentation (v8Segx, v8Segm, v8Segn) models on benchmark datasets. We compare these results against SOTA models in the segmentation domain.

*1) Detection:* For the task of object detection, we fine-tuned three different YOLOv8 variants – YOLOv8n, YOLOv8m, and YOLOv8x – on the UBCODD dataset. This dataset contains generated bounding box labels for each image, as detailed in the subsection III-A. The model's performance improved steadily with each epoch during training, as depicted in Figure 4. We then tested these models against the benchmark COD10K testing dataset [18]. The results of this experiment are shown in Table II.

TABLE II: Object Detection Performance Metrics on COD10K test dataset.

| Model | Precision | Recall | mAP50 | mAP50-95 |
|---|---|---|---|---|
| YOLOv8x | 0.98 | 0.91 | 0.95 | 0.88 |
| YOLOv8m | 0.94 | 0.83 | 0.89 | 0.77 |
| YOLOv8n | 0.71 | 0.45 | 0.54 | 0.37 |

We were unable to directly compare the performance of our fine-tuned YOLOv8 models against SOTA models because, unlike segmentation tasks where well-established SOTA models exist, the field of camouflaged object detection currently

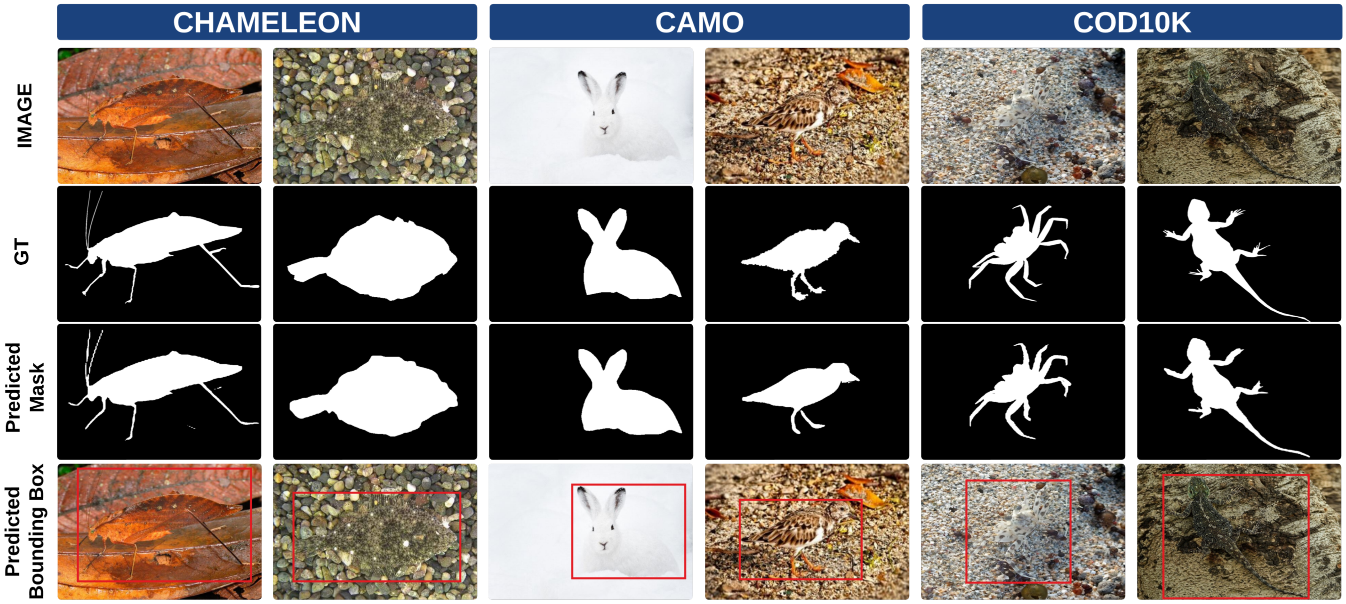


Fig. 3: Visual Performance of the Model. From top to bottom: name of the testing dataset, original images, original ground truths, ground truths generated by our fine-tuned YOLOv8, and bounding boxes produced by our fine-tuned YOLOv8.

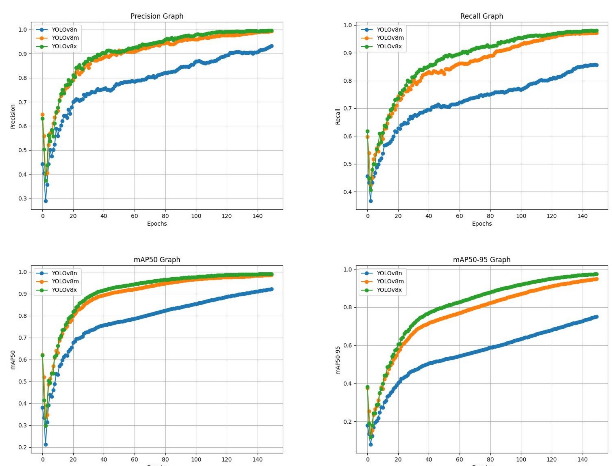


Fig. 4: Improvement in YOLOv8's Performance While Training on Precision, Recall, mAP.

lacks a standard benchmark for evaluating object detection models.

*2) Segmentation:* We also fine-tuned the YOLOv8 segmentation model on the UBCODD dataset. The results were then benchmarked on three standard datasets for camouflaged object detection: CHAMELEON, CAMO, and COD10K [18], [19], [25]. Table III provides a detailed comparison of evaluation metrics for different SOTA segmentation models. The findings demonstrate that our fine-tuned YOLOv8 segmentation models outperformed other SOTA models, especially on the CAMO and COD10K test datasets. Coupled with a rapid inference time, this performance suggests its capability to improve object segmentation tasks across diverse real-world scenarios [26].

The visual performance of our best-performing models, YOLOv8x and YOLOv8segx, across both segmentation and detection tasks, is showcased in Figure 3. This aids in understanding the generalizability and performance of our model across diverse benchmark datasets.

### *C. Using UBCODD-Trained Models for Agricultural Pest Detection and Segmentation*

Building upon the performance of the detection and segmentation tasks on the UBCODD training dataset, this subsection explores the suitability of these models for detecting and segmenting agricultural pests. Following an evaluation of the model's performance on benchmark COD datasets, we aim to tackle the concealed and camouflaged nature of agricultural pests and insects. For this, we utilize two benchmark Pest Detection datasets, namely IP-102 and Locust-Mini, to further validate the generalizability and effectiveness of our models in the agricultural domain.

*1) IP102: A Large-Scale Benchmark Dataset for Insect Pest Recognition:* IP-102, initially designed as a benchmark pest detection dataset primarily for classification tasks, also includes 19,000 images designated for object detection [12]. Leveraging this detection subset, we fine-tuned our top-performing object detection model, YOLOv8x. The results of this fine-tuning experiment, presented in Table IV, demonstrate that our model outperforms other SOTA models on this dataset. Furthermore, visual performance analysis of the YOLOv8x model in Figure 5 showcases its ability to detect camouflaged pests with high confidence scores. However, we couldn't fine-

TABLE III: Evaluation metrics for various segmentation models on benchmark datasets.

| Method | COD10K-Test (2,026 images) | | | | CAMO-Test (250 images) | | | | CHAMELEON-Test (76 images) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Sα ↑ | αE↑ | wF↑ | M↓ | Sα ↑ | αE↑ | wF↑ | M↓ | Sα ↑ | αE↑ | wF↑ | M↓ |
| CPD [30] | 0.752 | 0.820 | 0.557 | 0.049 | 0.712 | 0.813 | 0.561 | 0.108 | 0.860 | 0.908 | 0.753 | 0.044 |
| PraNet [13] | 0.768 | 0.836 | 0.599 | 0.047 | 0.738 | 0.814 | 0.613 | 0.098 | 0.864 | 0.918 | 0.784 | 0.038 |
| MINet-R [31] | 0.759 | 0.832 | 0.580 | 0.045 | 0.749 | 0.835 | 0.635 | 0.090 | 0.844 | 0.919 | 0.746 | 0.040 |
| SINet [18] | 0.771 | 0.807 | 0.565 | 0.048 | 0.742 | 0.834 | 0.601 | 0.101 | 0.869 | 0.903 | 0.749 | 0.041 |
| LSR [24] | 0.767 | 0.861 | 0.611 | 0.045 | 0.712 | 0.791 | 0.583 | 0.104 | 0.846 | 0.913 | 0.767 | 0.046 |
| PFNet [32] | 0.800 | 0.868 | 0.660 | 0.040 | 0.782 | 0.852 | 0.695 | 0.085 | 0.882 | 0.942 | 0.810 | 0.033 |
| $C^2F$-Net [33] | 0.810 | 0.875 | 0.674 | 0.038 | 0.791 | 0.863 | 0.706 | 0.083 | 0.886 | 0.931 | 0.824 | 0.032 |
| MGL [22] | 0.811 | 0.865 | 0.666 | 0.037 | 0.775 | 0.847 | 0.673 | 0.088 | 0.893 | 0.923 | 0.813 | 0.030 |
| SegMaR (Stage-4) [34] | 0.833 | 0.895 | 0.724 | 0.033 | 0.815 | 0.872 | 0.742 | 0.071 | **0.906** | **0.954** | **0.860** | **0.025** |
| YOLOv8Seg (Nano) | 0.771 | 0.852 | 0.652 | 0.058 | 0.796 | 0.844 | 0.739 | 0.084 | 0.818 | 0.883 | 0.744 | 0.045 |
| YOLOv8Seg (Medium) | 0.812 | 0.892 | 0.710 | 0.038 | 0.872 | 0.913 | 0.844 | 0.045 | 0.870 | 0.926 | 0.828 | 0.031 |
| YOLOv8Seg (Xlarge) | **0.834** | **0.906** | **0.746** | **0.033** | **0.902** | **0.948** | **0.886** | **0.031** | 0.885 | 0.947 | 0.854 | 0.028 |

tune and evaluate our segmentation models on this dataset as the curators didn't provide any binary masks.

TABLE IV: Object Detection Results on the IP102 [12] Test Dataset.

| Method | mAP50 | mAP50-95 |
|---|---|---|
| FPN [35] | 0.549 | 0.281 |
| TOOD [36] | 0.439 | 0.265 |
| SSD300 [37] | 0.472 | 0.215 |
| PAA [38] | 0.427 | 0.252 |
| Dynamic R-CNN [39] | 0.507 | 0.294 |
| Sparse R-CNN [40] | 0.332 | 0.211 |
| YOLOv3 [41] | 0.506 | 0.257 |
| YOLOX [42] | 0.521 | 0.311 |
| C3M-YOLO [43] | 0.572 | 0.349 |
| YOLOv8x (Ours) | **0.678** | **0.434** |

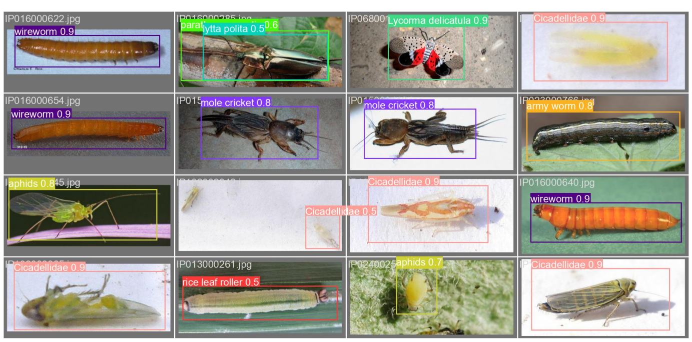


Fig. 5: Bounding Box Predictions of YOLOv8x on the IP-102 Test Dataset.

*2) Locust-Mini:* The Locust-Mini dataset contains diverse images of locusts, including camouflaged ones, gathered from various benchmark datasets like COD10K and the internet [44]. During the evaluation, we directly performed inference on different variants of our UBCODD fine-tuned detection and segmentation models on a test dataset comprising 120 images. Table V showcases the results achieved on the Locust-Mini dataset by our segmentation model compared to other fine-tuned SOTA segmentation models and Table VI showcases the results achieved by our object detection model. We couldn't benchmark the detection model against a SOTA detection model because no such model currently exists for this specific dataset.

We also illustrate the visual performance in Figure 6. These results were obtained without further fine-tuning our existing YOLOv8x detection and YOLOv8Segx segmentation models due to the unavailability of the Locust-Mini training set during the experimental phase of our research. Had the training dataset been accessible, our model could have potentially surpassed the existing results, given its comparable performance even without fine-tuning.

TABLE V: Segmentation Results on the Locust Mini Test Dataset [44].

| Model | Sα ↑ | αE↑ | wF↑ | M↓ |
|---|---|---|---|---|
| UNet [45] | 0.468 | – | 0.127 | 0.120 |
| SINet [18] | 0.887 | – | 0.805 | 0.024 |
| PraNet [13] | 0.812 | – | 0.676 | 0.043 |
| Polyp-PVT [46] | 0.895 | – | 0.850 | 0.018 |
| Improved PraNet [44] | 0.886 | – | 0.811 | 0.023 |
| YOLOv8Seg (Nano) | 0.814 | 0.916 | 0.720 | 0.043 |
| YOLOv8Seg (Medium) | 0.845 | 0.941 | 0.774 | 0.033 |
| YOLOv8Seg (XLarge) | 0.853 | 0.950 | 0.802 | 0.030 |

TABLE VI: Detection Results on the Locust Mini Test Dataset [44].

| Model | Precision | Recall | mAP50 | mAP50-95 |
|---|---|---|---|---|
| YOLOv8x | 0.82 | 0.72 | 0.75 | 0.68 |
| YOLOv8m | 0.80 | 0.68 | 0.72 | 0.61 |
| YOLOv8n | 0.75 | 0.65 | 0.67 | 0.60 |

The results of this subsection validate the performance of our fine-tuned UBCODD detection and segmentation models on both benchmark datasets, IP-102 and Locust-Mini. As a result, we are demonstrating the adaptability and effectiveness of these models in the field of agriculture for both the detection and segmentation of agricultural pests and insects.

## V. Conclusion

This paper presented a comprehensive investigation into the performance of a fine-tuned YOLOv8 model for the detection and segmentation of agricultural pests, particularly those with camouflage characteristics. The results from our


Fig. 6: Visual Performance of the Model on the Locust Mini Dataset. From top to bottom: original images, their corresponding ground truths, ground truths generated by our fine-tuned YOLOv8, and bounding boxes produced by our fine-tuned YOLOv8.

experiments on the Unified Benchmark Camouflaged Object Detection Dataset (UBCODD) demonstrate the effectiveness of fine-tuning YOLOv8 models for real-time inference in challenging agricultural environments. Furthermore, our findings underscore the potential of state-of-the-art computer vision techniques in addressing critical challenges in agricultural pest management.

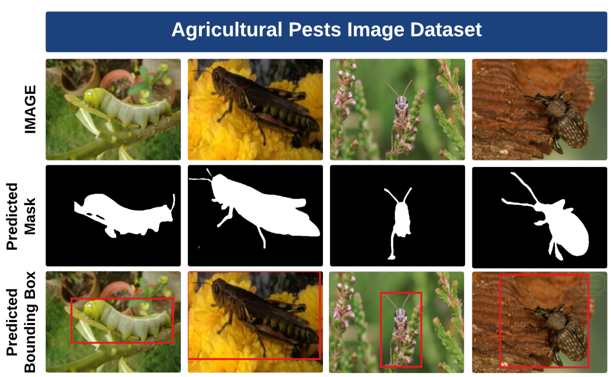


Fig. 7: Predictions on the Agricultural Pests Image Dataset [14].

### A. Open-Sourcing of UBCODD Dataset

As part of our commitment to advancing research in this domain, we plan to open-source the complete UBCODD dataset, along with prepared ground truth labels, code implementation for model training and evaluation, and pre-trained model weights. We believe that sharing these resources with the research community will facilitate collaboration, reproducibility, and further advancements in agricultural computer vision. Figure 7 showcases the compelling results of predicted masks and bounding boxes for Agricultural Pests Image Dataset [14].

## References

[1] CM Oliveira, AM Auad, SM Mendes, and MR Frizzas, "Economic impact of exotic insect pests in brazilian agriculture," *Journal of Applied Entomology*, vol. 137, no. 1-2, pp. 1–15, 2013.

[2] CM Oliveira, AM Auad, SM Mendes, and MR Frizzas, "Crop losses and the economic impact of insect pests on brazilian agriculture," *Crop protection*, vol. 56, pp. 50–54, 2014.

[3] Smriti Sharma, Rubaljot Kooner, and Ramesh Arora, "Insect pests and crop losses," *Breeding insect resistant crops for sustainable agriculture*, pp. 45–66, 2017.

[4] Dilbar Hussain, Muhammad Asrar, Bushra Khalid, Faisal Hafeez, Muhammad Saleem, Muazzma Akhter, Maqsood Ahmed, Imran Ali, and Kanwal Hanif, "Insect pests of economic importance attacking wheat crop (triticum aestivum l.) in punjab, pakistan," *International Journal of Tropical Insect Science*, vol. 42, no. 1, pp. 9–20, 2022.

[5] Kris AG Wyckhuys, Yanhui Lu, Wenwu Zhou, Matthew JW Cock, Steven E Naranjo, Atumurirava Fereti, Frances E Williams, and Michael J Furlong, "Ecological pest control fortifies agricultural growth in asia–pacific economies," *Nature Ecology & Evolution*, vol. 4, no. 11, pp. 1522–1530, 2020.

[6] Denys Yemshanov, Robert G Haight, Frank H Koch, Robert C Venette, Tom Swystun, Ronald E Fournier, Mireille Marcotte, Yongguang Chen, and Jean J Turgeon, "Optimizing surveillance strategies for early detection of invasive alien species," *Ecological Economics*, vol. 162, pp. 87–99, 2019.

[7] Fina Faithpraise, Phil Birch, Rupert Young, Joseph Obu, Bassey Faithpraise, and Chris Chatwin, "Automatic plant pest detection recognition using k-means clustering algorithm correspondence filters," *International Journal of Advanced Biotechnology and Research*, vol. 4, pp. 1052–1062, 02 2013.

[8] Heba Al-Hiary, Sulieman Bani-Ahmad, Mohammad Ryalat, Malik Braik, and Zainab Alrahamneh, "Fast and accurate detection and classification of plant diseases," *International Journal of Computer Applications*, vol. 17, 03 2011.

[9] Matheus Cardim Ferreira Lima, Maria Elisa Damascena de Almeida Leandro, Constantino Valero, Luis Carlos Pereira Coronel, and Clara Oliva Gonçalves Bazzo, "Automatic detection and monitoring of insect pests—a review," *Agriculture*, vol. 10, no. 5, 2020.

[10] Ana Cláudia Teixeira, José Ribeiro, Raul Morais, Joaquim J. Sousa, and António Cunha, "A systematic review on automatic insect detection using deep learning," *Agriculture*, vol. 13, no. 3, 2023.

[11] S. Lev-Yadun, A. Dafni, M. A. Flaishman, M. Inbar, I. Izhaki, G. Katzir, and G. Ne'eman, "Plant coloration undermines herbivorous insect camouflage," *BioEssays*, vol. 26, no. 10, pp. 1126–1130, 2004.

[12] Xiaoping Wu, Chi Zhan, Yu-Kun Lai, Ming-Ming Cheng, and Jufeng Yang, "IP102: A large-scale benchmark dataset for insect pest recognition," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 8787–8796.

[13] Deng-Ping Fan, Ge-Peng Ji, Tao Zhou, Geng Chen, Huazhu Fu, Jianbing Shen, and Ling Shao, "Pranet: Parallel reverse attention network for polyp segmentation," in *International conference on medical image computing and computer-assisted intervention*. Springer, 2020, pp. 263–273.

[14] Marionette, "Agricultural pests image dataset," Mar 2023.

[15] Y. Zhang, B. Ma, Y. Hu, C. Li, and Y. Li, "Accurate cotton diseases and pests detection in complex background based on an improved yolox model," *Computers and Electronics in Agriculture*, vol. 203, pp. 107484, 2022.

[16] H. Chen, R. Wang, J. Du, T. Chen, H. Liu, J. Zhang, R. Li, and G. Zhou, "Feature refinement method based on the two-stage detection framework for similar pest detection in the field," *Insects*, vol. 14, pp. 819, 2023.

[17] S. Yang, Z. Xing, H. Wang, X. Dong, X. Gao, Z. Liu, X. Zhang, S. Li, and Y. Zhao, "Maize-yolo: A new high-precision and real-time method for maize pest detection," *Insects*, vol. 14, pp. 278, 2023.

[18] Deng-Ping Fan, Ge-Peng Ji, Guolei Sun, Ming-Ming Cheng, Jianbing Shen, and Ling Shao, "Camouflaged object detection," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 2777–2787.

[19] Z. Nie M.-T. Tran T.-N. Le, T. V. Nguyen and A. Sugimoto, "Anabranch network for camouflaged object segmentation," *Computer vision and image understanding, vol. 184, pp. 45–56*, 2019.

[20] Tao Zhou, Yi Zhou, C. G. J. Y., and Y. Zhang, "Feature aggregation and propagation network for camouflaged object detection," *IEEE Transactions on Image Processing*, 2022.

[21] W. Xie H. Lamdouar, C. Yang and A. Zisserman, "Betrayed by motion: Camouflaged object discovery via motion segmentation," *Proceedings of the Asian Conference on Computer Vision*, 2020.

[22] Qiang Zhai, Xin Li, Fan Yang, Chenglizhao Chen, Hong Cheng, and Deng-Ping Fan, "Mutual graph learning for camouflaged object detection," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 12997–13007.

[23] Y. Zhong, B. Li, L. Tang, S. Kuang, S. Wu, and S. Ding, "Detecting camouflaged object in frequency domain," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 4504–4513.

[24] Yunqiu Lyu, Jing Zhang, Yuchao Dai, Aixuan Li, Bowen Liu, Nick Barnes, and Deng-Ping Fan, "Simultaneously localize, segment and rank the camouflaged objects," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021.

[25] Przemysław Skurowski and Paweł Kasprowski, "Evaluation of saliency maps in a hard case–images of camouflaged animals," in *2018 IEEE International Conference on Image Processing, Applications and Systems (IPAS)*. IEEE, 2018, pp. 244–249.

[26] Ultralytics, "Ultralytics/ultralytics: New - yolov8 in pytorch," .

[27] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick, "Microsoft COCO: Common Objects in Context," in *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13*. Springer, 2014, pp. 740–755.

[28] Deng-Ping Fan, Ming-Ming Cheng, Yun Liu, Tao Li, and Ali Borji, "Structure-measure: A new way to evaluate foreground maps," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 4548–4557.

[29] Deng-Ping Fan, Cheng Gong, Yang Cao, Bo Ren, Ming-Ming Cheng, and Ali Borji, "Enhanced-alignment measure for binary foreground map evaluation," *arXiv preprint arXiv:1805.10421*, 2018.

[30] Zhe Wu, Li Su, and Qingming Huang, "Cascaded partial decoder for fast and accurate salient object detection," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 3907–3916.

[31] Youwei Pang, Xiaoqi Zhao, Lihe Zhang, and Huchuan Lu, "Multi-scale interactive network for salient object detection," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 9413–9422.

[32] Haiyang Mei, Ge-Peng Ji, Ziqi Wei, Xin Yang, Xiaopeng Wei, and Deng-Ping Fan, "Camouflaged object segmentation with distraction mining," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 8772–8781.

[33] Yujia Sun, Geng Chen, Tao Zhou, Yi Zhang, and Nian Liu, "Context-aware cross-level fusion network for camouflaged object detection," *arXiv preprint arXiv:2105.12555*, 2021.

[34] Qi Jia, Shuilian Yao, Yu Liu, Xin Fan, Risheng Liu, and Zhongxuan Luo, "Segment, magnify and reiterate: Detecting camouflaged objects the hard way," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 4713–4722.

[35] Tsung-Yi Lin, Piotr Dollár, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie, "Feature pyramid networks for object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2117–2125.

[36] Chengjian Feng, Yujie Zhong, Yu Gao, Matthew R Scott, and Weilin Huang, "Tood: Task-aligned one-stage object detection," in *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*. IEEE Computer Society, 2021, pp. 3490–3499.

[37] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C Berg, "Ssd: Single shot multibox detector," in *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*. Springer, 2016, pp. 21–37.

[38] Kang Kim and Hee Seok Lee, "Probabilistic anchor assignment with iou prediction for object detection," in *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXV 16*. Springer, 2020, pp. 355–371.

[39] Hongkai Zhang, Hong Chang, Bingpeng Ma, Naiyan Wang, and Xilin Chen, "Dynamic r-cnn: Towards high quality object detection via dynamic training," in *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XV 16*. Springer, 2020, pp. 260–275.

[40] Peize Sun, Rufeng Zhang, Yi Jiang, Tao Kong, Chenfeng Xu, Wei Zhan, Masayoshi Tomizuka, Lei Li, Zehuan Yuan, Changhu Wang, et al., "Sparse r-cnn: End-to-end object detection with learnable proposals," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 14454–14463.

[41] Joseph Redmon and Ali Farhadi, "Yolov3: An incremental improvement," *arXiv preprint arXiv:1804.02767*, 2018.

[42] Zheng Ge, Songtao Liu, Feng Wang, Zeming Li, and Jian Sun, "Yolox: Exceeding yolo series in 2021," *arXiv preprint arXiv:2107.08430*, 2021.

[43] Lijuan Zhang, Cuixing Zhao, Yuncong Feng, and Dongming Li, "Pests identification of ip102 by yolov5 embedded with the novel lightweight module," *Agronomy*, vol. 13, no. 6, pp. 1583, 2023.

[44] Limei Liu, Meng Liu, Kexin Meng, Linqiang Yang, Min Zhao, and Shuli Mei, "Camouflaged locust segmentation based on pranet," *Computers and Electronics in Agriculture*, vol. 198, pp. 107061, 2022.

[45] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*. Springer, 2015, pp. 234–241.

[46] Dong Bo, Wang Wenhai, Fan Deng-Ping, Li Jinpeng, Fu Huazhu, and Shao Ling, "Polyp-PVT: Polyp segmentation with PyramidVision transformers," 2023.