# The rapid development of a computer vision model for detecting crop pests for integrated pest management: a case study in the black vine weevil.


## Abstract

Insect pests pose a huge threat to global agriculture, inflicting widespread crop damage and lowering output. Traditional pesticide usage includes downsides such as insect resistance, damage to non-target species, and pesticide residues in agricultural goods, notwithstanding its effectiveness. This work suggests the quick construction of a computer vision model for identifying crop pests, utilizing the black vine weevil as a case study.
The goals of this study are to develop a reproducible methodology for constructing a baseline dataset for pest insect classification, to create a comparative framework for baseline classification and detection models, to test an experiment methodology for rapid training and fine-tuning of pest detection models, and to validate the model's accuracy using real pest images in practical scenarios.
The study gathered and annotated a collection of weevil and earwig photos using the YOLOv8 model for object recognition. To improve performance, several model configurations, such as batch size and picture size, were tested.
At various IOU thresholds, the findings show that the YOLOv8n and YOLOv8s models consistently beat other models in terms of accuracy, recall, and mean average precision. This demonstrates their aptitude for correctly and efficiently detecting weevils and earwigs. Farmers may receive pest information fast and precisely by automating pest identification with computer vision, enabling focused pest management tactics. This adheres to the principles of integrated pest control, lowering dependency on chemicals and minimizing environmental implications.
Finally, the YOLOv8-based computer vision model provides an efficient and resilient approach for identifying weevils and earwigs, hence assisting precision pest management in agriculture. This technology's effective adoption has the potential to improve farming practices, increase efficiency, and reduce environmental concerns related with pest management.


## Introduction

Insect pests are a major cause of crop damage globally, resulting in significant losses in agriculture (Nanni, Maguolo and Pancino, 2020). Crop pests can interfere with metabolic processes and reduce yield output and quality(Muralidharan and Pasalu, 2006). Regular pesticide use is often habitual, which while sometimes effectively contributes to the accumulation of pest resistance and potentially negatively impacting non-pest species, while contributing to business costs(Hazarika, Bhuyan and Hazarika, 2009). Indeed, there is evidence routine pesticide use negatively impacts the ecological environment and leaves pesticide residues in agricultural products(Miller and Spoolman, 2014). Pest population information cannot be easily and accurately obtained, further contributing to the practice of pesticide overuse. On the other hand, if accurate pest population information were easily and efficiently accessible, Adequate pest control approaches, such as the careful use of insecticides, and proper prevention measures may be achievable(Zhang and Swinton, 2009). As a result, information on insect types and quantities is critical and essential. Because of the fast rise of digital image technology, there is a growing tendency towards adopting machine vision technology to address these issues with promising outcomes in the field of agricultural research.(Cho *et al.,* 2007).

Traditional manual methods of insect pest monitoring, relying on human experts, are labour-intensive, subjective, and unsuitable for large-scale applications (Ding and Taylor, 2016).

Workers do this by comparing morphological features such as colour, texture, and other attributes. However, this often requires labour and expertise which are limiting and prone to mistakes (Cho *et al.,* 2007)(Ding and Taylor, 2016). A potential solution is to automate pest identification. This need has given rise to solutions using computer vision (Cho *et al.,* 2007).

Even though certain pertinent studies have achieved significant advancements, their study still mostly focuses on theoretical issues and pays little attention to real-world application scenarios for two key reasons: First, methods other than deep learning are frequently used to measure and identify insects, yet the precision varies (Wang *et al.,* 2012). Furthermore, most recent studies employ insect photos obtained in an ideal lab setting rather than in the field (Gassoumi, Prasad and Ellington, 2000) (Maharlooei *et al.,* 2017); (Larios *et al.,* 2010)(Kang, Cho, and Lee, 2014). Although only a few studies employ photos of insects taken in the wild, these high-resolution photographs must be transferred to a server where the counting and identification tasks are accomplished. To overcome the obstacles Real-time object identification is an important task in computer vision and is commonly used in computer vision systems. An object detector employs an object detection approach to estimate bounding boxes and class probabilities for each item in the input picture when conducting image recognition tasks.

In the framework of comprehensive pest management (IPM), pest control that combines and integrates biological and chemical control is defined (Stern *et al.,* 1959). IPM tries to eliminate insect populations below the economic injury level which results in less effects on environment and human begins. The promise of computer vision models for automated, instantaneous detection of crop pests has gained significant attention. Such a model would facilitate the implementation of precision pest control measures, thereby reducing crop losses and remove the limitations of cost, time, and expertise. Among the pests Otiorhynchus sulcatus, sometimes known as the black vine weevil (BVW), is a significant pest of Horticultural crops, ornamentals used in landscaping. The black vine weevil harms host plants by marginally notching the leaves of broadleaved evergreens and other host plants. Adults prefer to consume plant leaves over subterranean tissues or fruits, resulting in noticeable scratching along the leaf margin(Moorhouse, Charnley and Gillespie, 1992). Although it is well acknowledged that such notching has minimal influence on overall plant health, its presence can significantly reduce the economic value of ornamental plants that are less forgiving of aesthetic defects. (California Agriculture published descriptions of this univoltine species (having a single generation per year) in March–April 1984, January–February 1985, and May–June 1989.The overall aim of this project is to rapidly develop a computer vision model that accurately detects the black vine weevil.

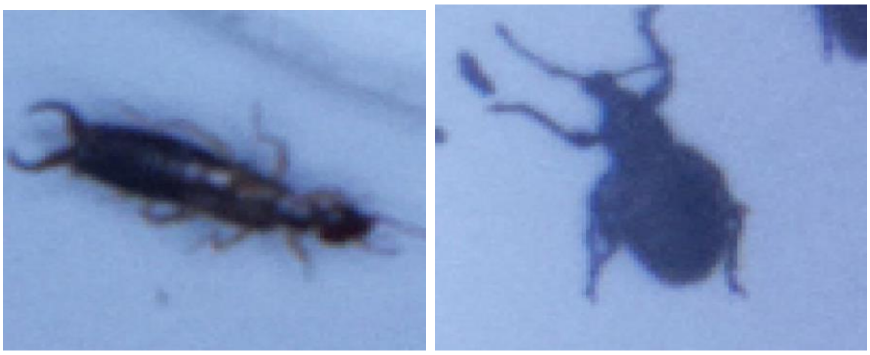
**Objectives**

My specific objectives are: 1)Create a reproducible methodology to rapidly create a baseline dataset for training, testing and validation for pest insect classification; 2) create a reproducible baseline classification and detection model using a comparative framework (different model sizes in a recent YOLO model framework); 3) create and test a reproducible experiment methodology to rapidly train and tune pest detection models for accuracy and speed; and 4) validate an example model using real pest images in a biologically relevant scenario. I will discuss my results in the context of opportunities of computer vision applications for solving integrated pest management problems.

**Dataset**

I have carefully gathered pictures of weevils and earwigs under numerous intricate backdrops for this work because there is not a publicly accessible dataset. Two classes—weevils and earwigs—make up the dataset utilised for the experiment. The input data set had 400 annotated photographs that were separated into 8:1:1 training, validation, and test groups. In the format of Roboflow's YOLOV8 dataset (Dwyer, B., Nelson, J. (2022), Solawetz, J., et al. Roboflow (Version 1.0) [Software]).

a) b)



Two types of classes: a) Earwig b) Weevil

**Proposed method for object detection**

YOLO (You Only Look Once) is a popular object identification system with great accuracy and performance (Redmon *et al.,* 2016) Initially announced YOLO in the publication "You Only Look Once: Unified, Real-Time Object Detection," which is known as a single-stage detector since it accomplishes everything in a single step (Jiang *et al.,* 2022)YOLOv8, the most recent version of this framework developed by Ultralytics (Ultralytics, Maryland, USA), works by partitioning a picture into smaller sections and then predicting a bounding box and class probabilities for each object present in each zone. The Darknet-53 architecture is used by the YOLOv8 algorithm to increase feature extraction, resulting in more accurate object recognition. DarkNet-53 is a 53-layer convolutional neural network that can categorise photos into 1,000 item categories. This network is broken into smaller stages, which are then partially connected, allowing for greater feature reuse and gradient propagation. The YOLOv8 model was chosen for its ability to detect objects quickly while retaining high accuracy. The model's architecture (Figure 1) was set up with DarkNet-53 for improved feature extraction and the Pseudo Ensemble (PS) approach for improved prediction robustness. The model's output supplied bounding boxes and class probabilities for each recognised Weevil and Earwig, which was useful for further investigation.

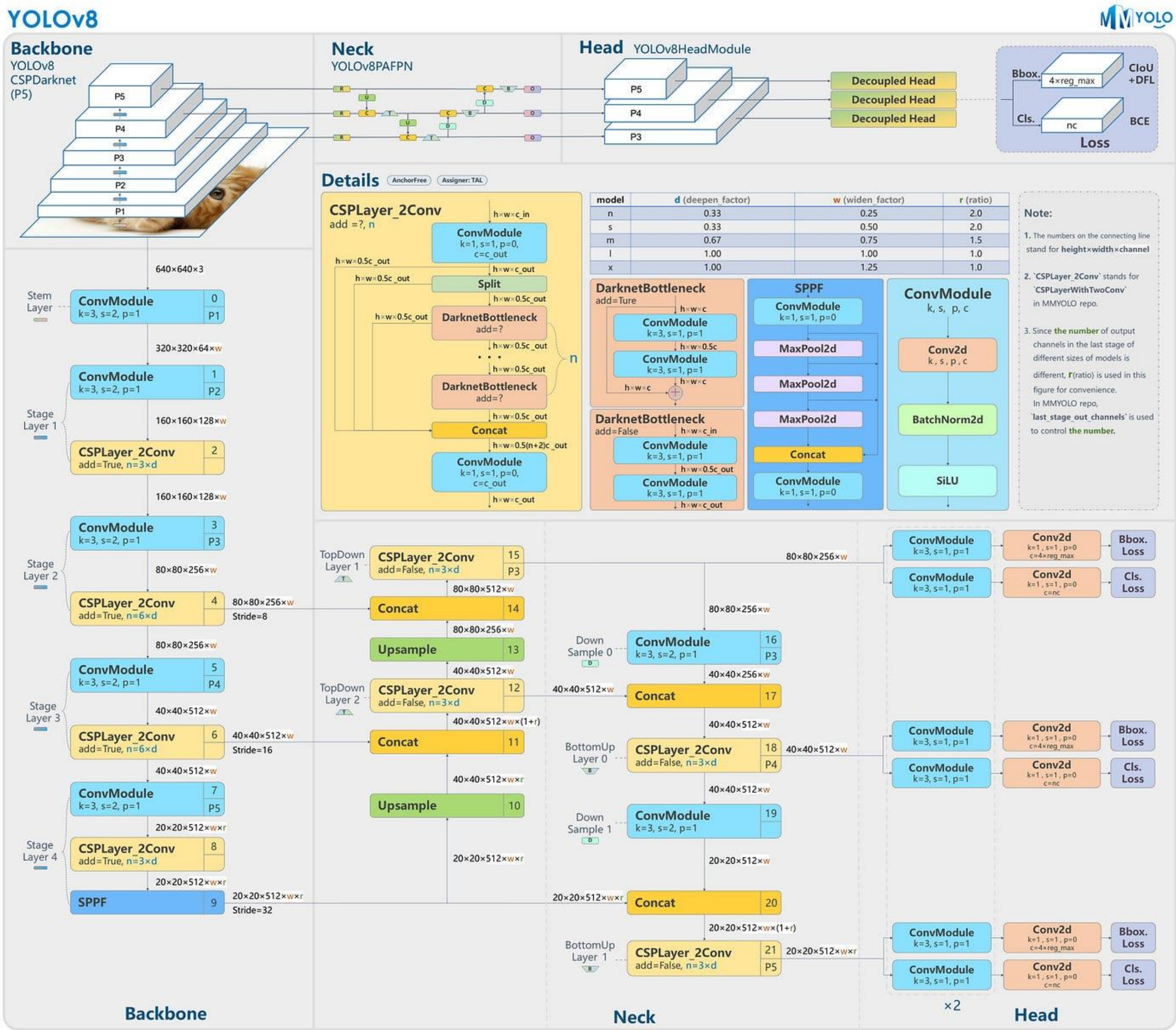



Figure 1: Schematic representation of YOLOV8 model (Ultralytics (Version 8.0.0) [Computer Software] https://github.com/ultralytics/ultralytics

To optimise the model's performance, the specified parameters, such as epochs, batch size, and image size, were carefully adjusted through many training and debugging attempts. The hyperparameters were chosen to provide a good balance of model accuracy and computational efficiency. There are five models (Figure 2) in yolov8 with different speed, accuracy, and size. We have trained our dataset into twenty times to find out the best model for accuracy. While training we have adjusted many parameters specific to model for example image size 100 & 128 and batch size 8 & 16 and epochs 100 which is constant, likewise we have trained each models into four times.

| Model | size (pixels) | $mAP^{val}$ 50-95 | Speed CPU ONNX (ms) | Speed A100 TensorRT (ms) | params (M) | FLOPs (B) |
|---|---|---|---|---|---|---|
| YOLOv8n | 640 | 37.3 | 80.4 | 0.99 | 3.2 | 8.7 |
| YOLOv8s | 640 | 44.9 | 128.4 | 1.20 | 11.2 | 28.6 |
| YOLOv8m | 640 | 50.2 | 234.7 | 1.83 | 25.9 | 78.9 |
| YOLOv8l | 640 | 52.9 | 375.2 | 2.39 | 43.7 | 165.2 |
| YOLOv8x | 640 | 53.9 | 479.1 | 3.53 | 68.2 | 257.8 |

Figure 2: YOLOV8 models (Ultralytics (Version 8.0.0) [Computer Software] https://github.com/ultralytics/ultralytics

**Metrics of performance**

The performance of the object identification models was evaluated in the current work using commonly used measures like as mean average precision (mAP), precision (P), recall (R), IoU, F-1 score, mean average precision (mAP), and so on(Ferri, Hernández-Orallo and Modroiu, 2009) The test results can be interpreted as true positive (TP), false positive (FP), false negative (FN), or true negative (TN) based on the confusion matrix created from the evaluation procedure. During binary classification, the classified item can be identified as TP for IoU 0.5. If the IoU is less than 0.5, it might be classified as FP.

**Precision:** Precision is the fraction of real positive predictions (objects properly recognised) out of all projected positive cases. Higher accuracy values suggest that the models have a low false positive rate, indicating that they are accurate in detecting the important items (Earwig and Weevil) in pest data.

$$P = \frac{TP}{(TP + FP)}$$

**Recall:** Recall, also known as sensitivity, is the fraction of genuine positive predictions in the dataset out of all real positive cases. Higher recall values indicate that the models effectively captured a considerable amount of the real Earwig and Weevil cases, lowering the frequency of false negatives.

$$R = \frac{TP}{(TP + FN)}$$

P and R numbers that are higher signify better detecting abilities. As opposed to this, the F-1 score is the P and R's arithmetic mean, which is expressed as:

$$F1 - SCORE = 2\left[\frac{P \times R}{P + R}\right]$$

**mAP50:** At 50%, the mean average precision is 50%. The Intersection over Union (IOU) criteria measures the average accuracy of the models when the projected bounding boxes and the ground truth overlap by at least 50%. A higher mAP50 score shows that the models are effective in reliably localising the commodities, especially when there is reasonable overlap with the ground truth.

**mAP50-90:** Mean Average Precision from 50% to 90% IOU threshold evaluates the models' performance across a broader range of IOU thresholds. A higher mAP50-90 score indicates that the models can retain consistent accuracy in localising Weevil and Earwig objects even as the overlap gets stricter (IOU rises).

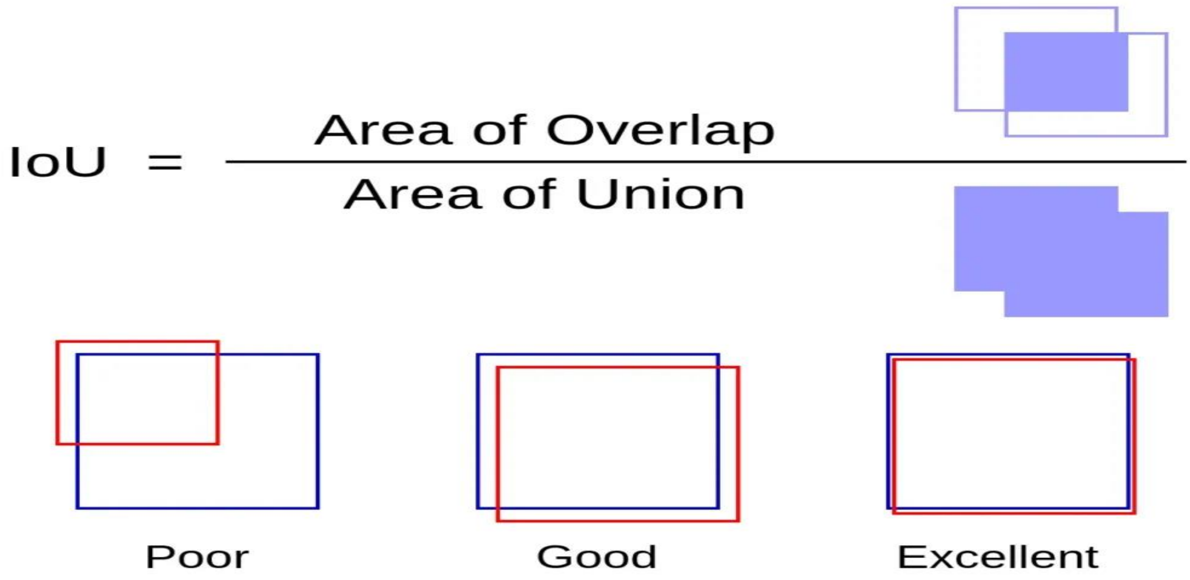



Figure 3: Schematic representation of IOU

https://idiotdeveloper.com/what-is-intersection-over-union-iou/

**Results**

Figure 3 shows instances of Weevil and Earwig recognition produced with the YOLOv8-based model. These findings indicate the model's ability to effectively recognise and delineate the pest detection within the photos. The model's visual outputs not only illustrate the approach's potential in the context of the research, but also provide insight into its practical use for pest detection in Agriculture settings. Tables 1,2,3, and 4 illustrate the model's Precision, recall, mAP50 and mAP50-90 for detecting Weevils and Earwigs, respectively.

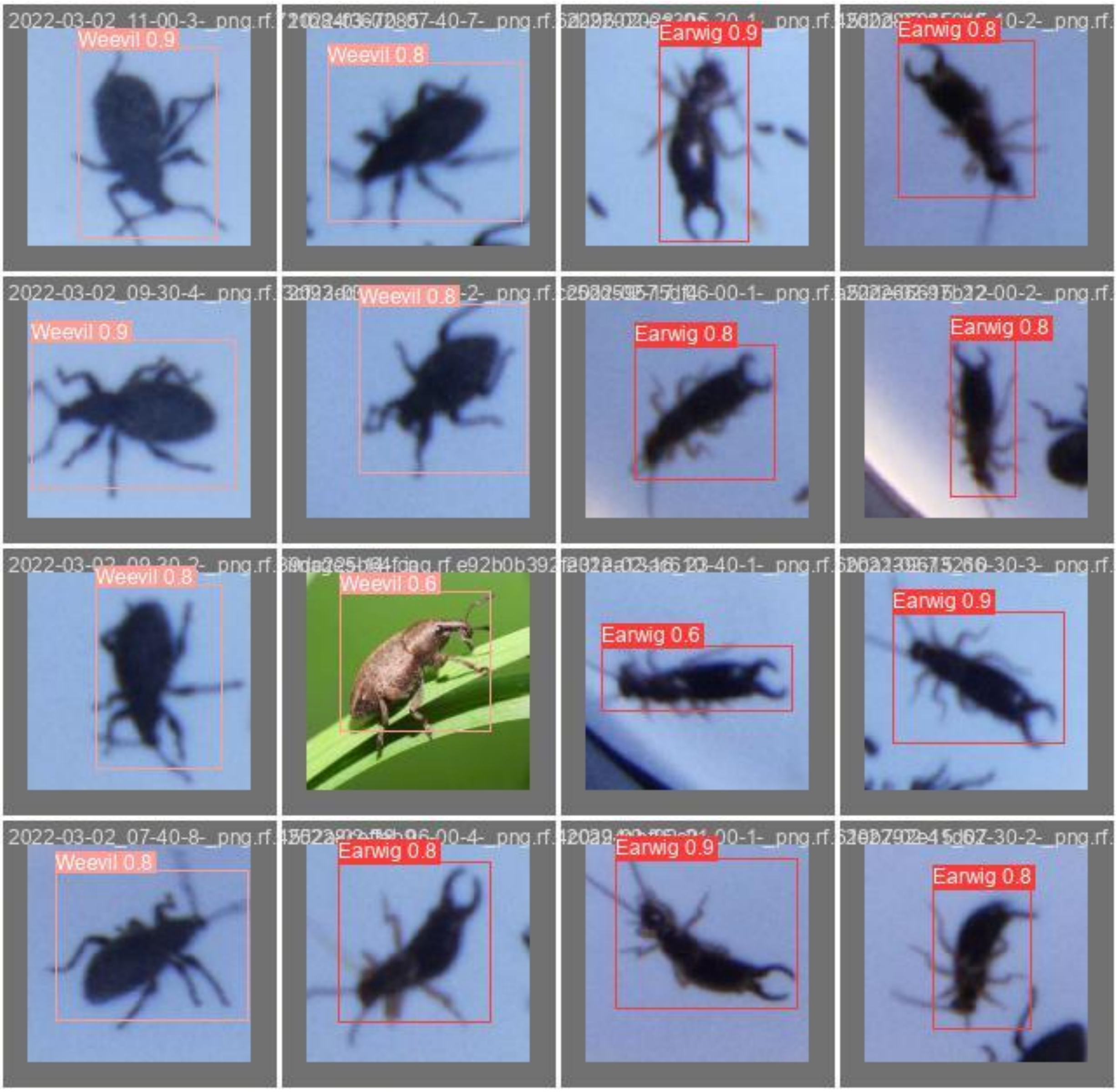



Figure 3: Images demonstrating the findings of the YOLOv8-based object detection algorithm.

Table 1: Result and evaluation metrics for batch 8 and image size 128 pixels

| Models | Precision | Recall | mAP50 | mAP50-90 |
|---|---|---|---|---|
| Yolov8n | **0.91** | **0.92** | **0.94** | **0.57** |
| Yolov8s | **0.91** | **0.92** | **0.94** | **0.57** |
| Yolov8m | 0.87 | 0.91 | 0.91 | 0.53 |
| Yolov8l | 0.81 | 0.83 | 0.82 | 0.48 |
| Yolov8x | 0.81 | 0.84 | 0.83 | 0.49 |

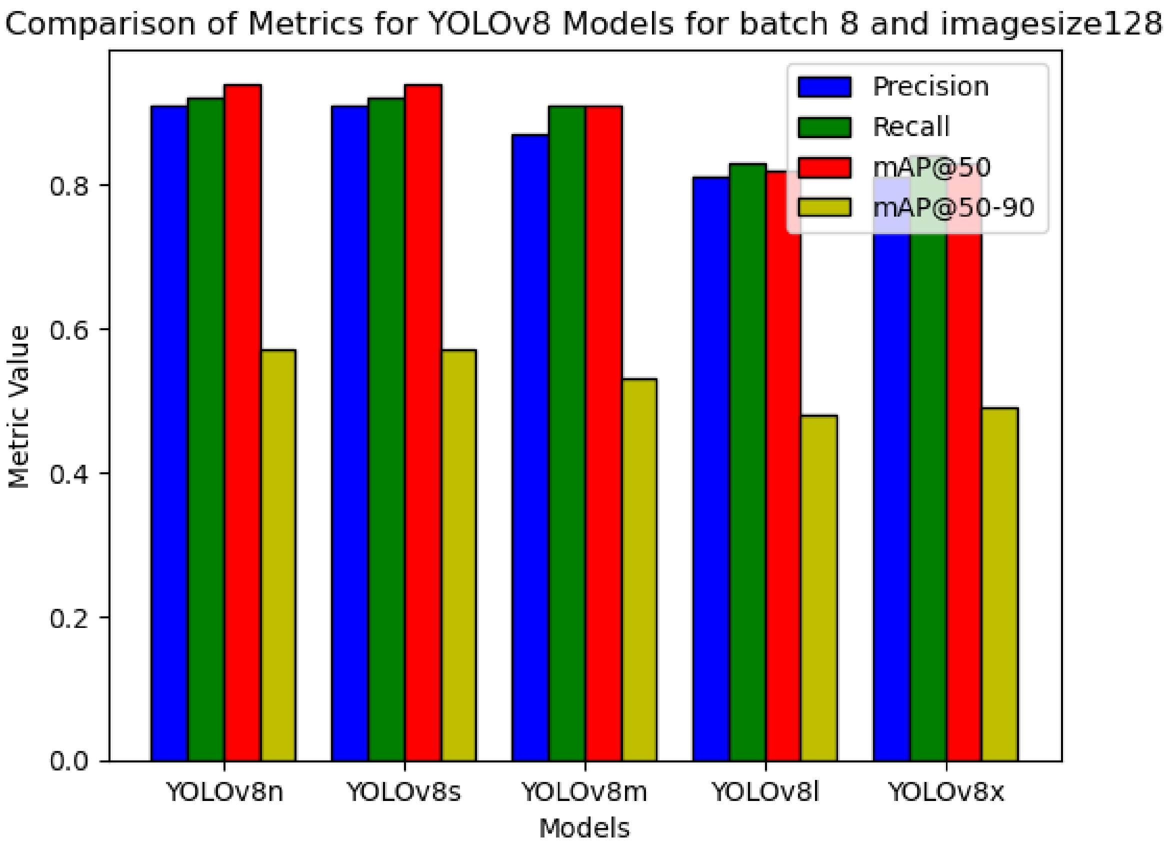



Table 2: Result and evaluation metrics for batch 8 and image size 256 pixels

| Models | Precision | Recall | mAP50 | mAP50-90 |
|---|---|---|---|---|
| Yolov8n | 0.89 | 0.92 | 0.93 | 0.56 |
| Yolov8s | 0.88 | 0.90 | 0.91 | 0.53 |
| Yolov8m | 0.86 | 0.89 | 0.89 | 0.52 |
| Yolov8l | 0.73 | 0.80 | 0.78 | 0.44 |
| Yolov8x | 0.73 | 0.79 | 0.77 | 0.444 |

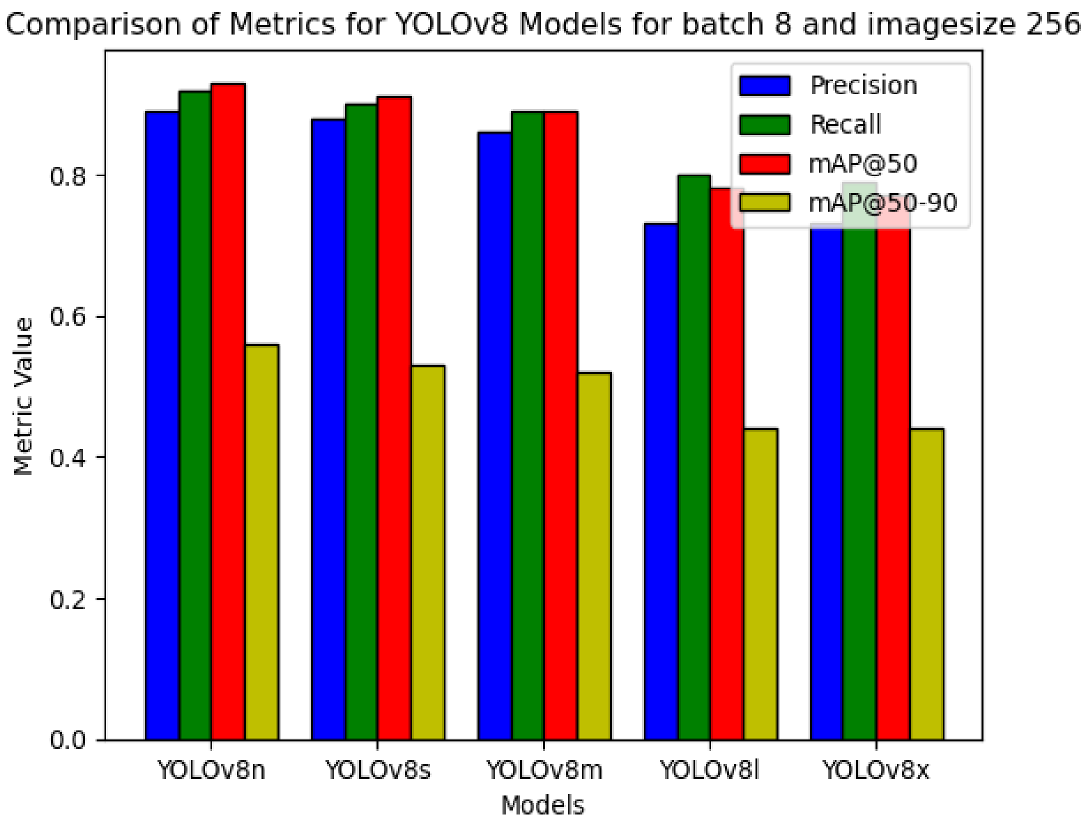



Table 3: Result and evaluation metrics for batch 16 and image size 128 pixels

| Models | Precision | Recall | mAP50 | mAP50-90 |
|---|---|---|---|---|
| Yolov8n | 0.90 | 0.92 | 0.94 | 0.56 |
| Yolov8s | 0.91 | 0.93 | 0.94 | 0.57 |
| Yolov8m | 0.89 | 0.89 | 0.90 | 0.52 |
| Yolov8l | 0.89 | 0.84 | 0.81 | 0.49 |
| Yolov8x | 0.80 | 0.82 | 0.81 | 0.48 |

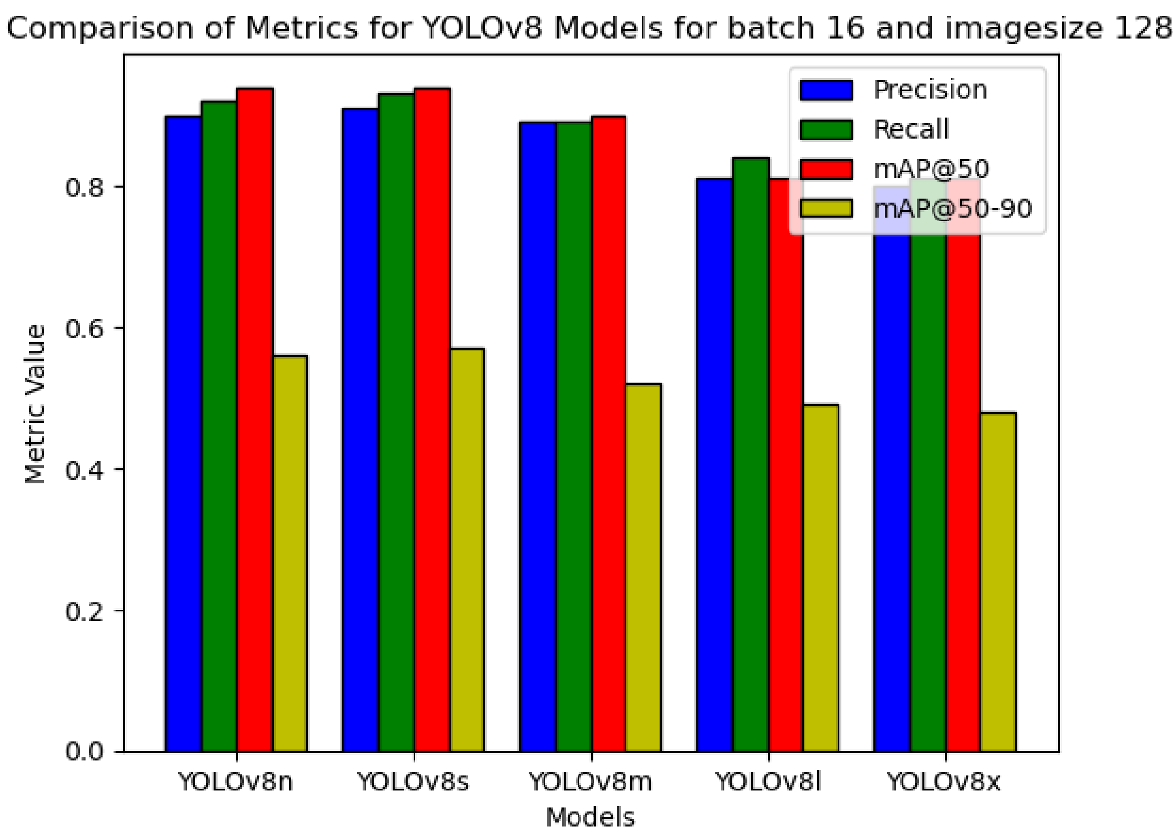
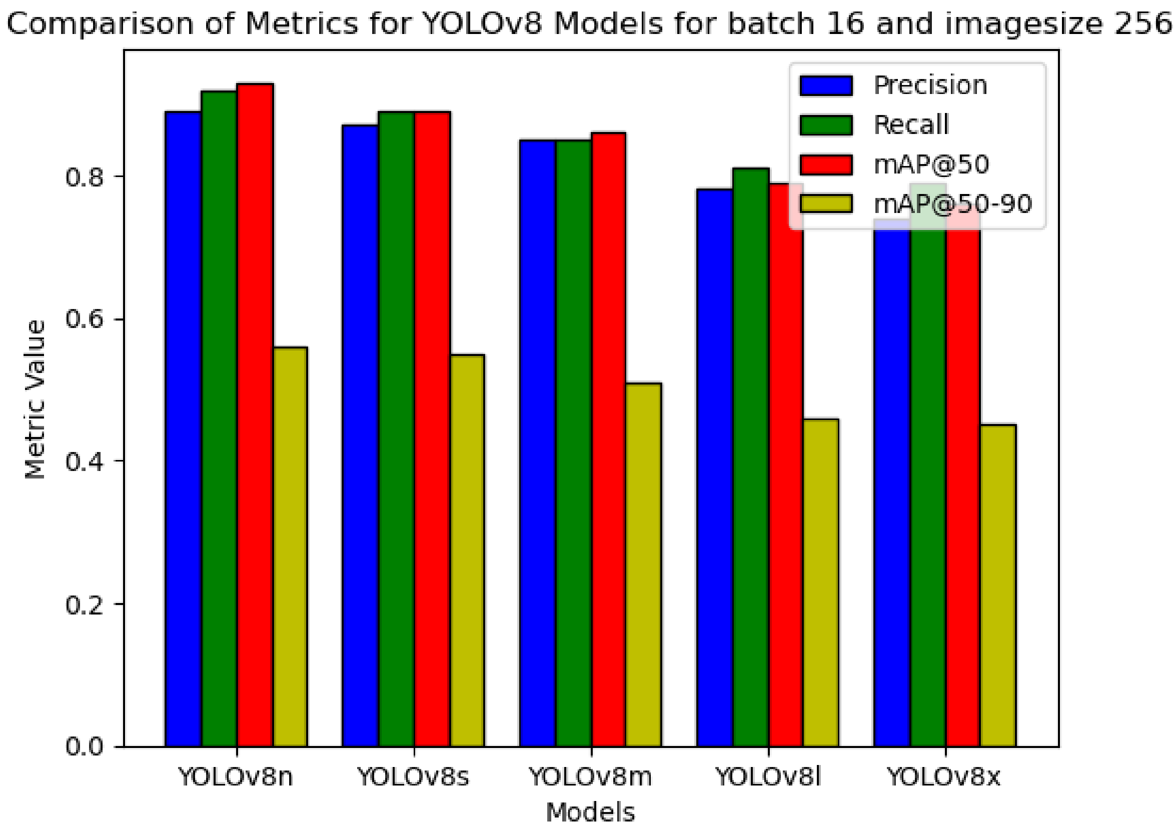



Table 4: Result and evaluation metrics for batch 16 and image size 256 pixels

| Models | Precision | Recall | mAP50 | mAP50-90 |
|---|---|---|---|---|
| Yolov8n | 0.89 | 0.92 | 0.93 | 0.56 |
| Yolov8s | 0.87 | 0.89 | 0.89 | 0.55 |
| Yolov8m | 0.85 | 0.85 | 0.86 | 0.51 |
| Yolov8l | 0.78 | 0.81 | 0.79 | 0.46 |
| Yolov8x | 0.74 | 0.79 | 0.76 | 0.45 |




The table and graph show the quantitative findings for Recall, Precision, mAP50, and mAP50-90 values for the five YOLOv8 models in Weevil and Earwig detection. When comparing the performance of several YOLOv8 models, fascinating discoveries emerge. Yolov8n and Yolov8s routinely exhibit excellent accuracy, recall, mAP50, and mAP50-90 scores, making them appropriate candidates for accurate object detection tasks in the Earwig and Weevil. The results for Yolov8m, Yolov8l, and Yolov8x are slightly lower, demonstrating an imbalance both the level of complexity and efficiency. Among the models evaluated, Yolov8n and Yolov8s appear to be the most sturdy and accurate options.

**Discussion**

This type of study is important because it addresses an important need in pest management. Weevil damages more than100 different kinds of ornamental plants in gardens, indoor or greenhouse plants, and other important potted plants. In Horticulture, accurate pest assessment is critical for maintaining aesthetic beauty of landscape plants and lawn. Pest detection has traditionally been done manually, which is time-consuming, labour-intensive, and prone to mistake. The suggested machine vision system provides a faster and more accurate alternative to manual detection. By automating the process, the system may offer farmers with quick and precise pest information, allowing them to adapt their management practises to optimise environment and biodiversity. Furthermore, the system has the potential to minimise labour costs and increases, making it a significant tool for commercial horticulture. As a result, this study is important for expanding Horticultural and Agricultural technology and increasing farmer efficiency and profitability.

Apart from using computer vision model to detect weevils, the main aim of this project is reducing the impact of pesticides on environment and biodiversity by following IPM

(Integrated pest management). IPM is not a single pest prevention rather than following a series of action to control the pest based on the setting certain threshold levels. In this context, it will eliminate the need of pesticide residuals and protect the non-targeted species in the biodiversity. This will reduce the environment risks associated with the pest management by encouraging the adoption of more ecologically control tactics.

**Conclusion**

To summarise, we created an efficient and robust object detection system YOLOv8 based on computer vision for accurate detection of Weevils and Earwigs in this study. The findings demonstrate that utilising larger batches (batch 16) and higher picture resolutions (picture size 256 pixels) leads in a minor drop in accuracy, recall, mAP50, and mAP50-90 compared to using smaller batch sizes (batch 8) and lower image resolutions (image size 128 pixels). However, the performance drop is minor, demonstrating that the models may still retain good accuracy with bigger batch sizes and better picture resolutions. The top-performing models across all configurations are the YOLOv8n and YOLOv8s. They regularly earn the greatest accuracy, recall, mAP50, and mAP50-90 scores, making them the best models for detecting Weevils and Earwigs. The excellent accuracy and recall scores of the YOLOv8n and YOLOv8s algorithms show their potential usefulness for accuracy Horticultural and agricultural applications such as pest management. Their capacity to identify and localise Weevil and Earwig occurrences can help to develop more efficient and focused pest management techniques, which will improve farming practises in the long run.

**Bibliography**

Cho, J., Choi, J., Qiao, M., Ji, C.W., Kim, H.Y., Uhm, K.B. and Chon, T.S. (2007) 'Automatic identification of whiteflies, aphids and thrips in greenhouse based on image analysis', *Red,* 346(246), pp. 244.

Ding, W. and Taylor, G. (2016) 'Automatic moth detection from trap images for pest management', *Computers and Electronics in Agriculture,* 123, pp. 17-28.

Ferri, C., Hernández-Orallo, J. and Modroiu, R., 2009. An experimental comparison of performance measures for classification. *Pattern recognition letters*, *30*(1), pp.27-38.

Gassoumi, H., Prasad, N.R. and Ellington, J.J. (2000) *Neural network-based approach for insect classification in cotton ecosystems.* InTech 2000 Bangkok, Thailand.

Hazarika, L.K., Bhuyan, M. and Hazarika, B.N. (2009a) 'Insect Pests of Tea and Their Management', *Annual Review of Entomology,* 54(1), pp. 267-284. doi: 10.1146/annurev.ento.53.103106.093359.

Hazarika, L.K., Bhuyan, M. and Hazarika, B.N. (2009b) 'Insect pests of tea and their management', *Annual Review of Entomology,* 54, pp. 267-284.

Hoye, T. T., et al. (2021). Potential of Deep Learning Tools for Entomology: A Perspective. Insects, 12(3), 270.

Jiang, P., Ergu, D., Liu, F., Cai, Y. and Ma, B., 2022. A Review of Yolo algorithm developments. *Procedia Computer Science*, *199*, pp.1066-1073.

Kang, S., Cho, J. and Lee, S. (2014) 'Identification of butterfly based on their shapes when viewed from different angles using an artificial neural network', *Journal of Asia-Pacific Entomology,* 17(2), pp. 143-149.

Larios, N., Soran, B., Shapiro, L.G., Martínez-Muñoz, G., Lin, J. and Dietterich, T.G. (2010) *Haar random forest features and SVM spatial matching kernel for stonefly species identification.* IEEE, pp. 2624.

Li, W., Zheng, T., Yang, Z., Li, M., Sun, C. and Yang, X. (2021) 'Classification and detection of insects from field images using deep learning for smart pest management: A systematic review', *Ecological Informatics,* 66, pp. 101460. doi: 10.1016/j.ecoinf.2021.101460.

Maharlooei, M., Sivarajan, S., Bajwa, S.G., Harmon, J.P. and Nowatzki, J. (2017) 'Detection of soybean aphids in a greenhouse using an image processing technique', *Computers and Electronics in Agriculture,* 132, pp. 63-70.

Miller, G.T. and Spoolman, S. (2014) *Sustaining the earth.* Cengage Learning.

Moorhouse, E.R., Charnley, A.K. and Gillespie, A.T., 1992. A review of the biology and control of the vine weevil, Otiorhynchus sulcatus (Coleoptera: Curculionidae). *Annals of Applied Biology*, *121*(2), pp.431-454.

Muralidharan, K. and Pasalu, I.C. (2006) 'Assessments of crop losses in rice ecosystems due to stem borer damage (Lepidoptera: Pyralidae)', *Crop Protection,* 25(5), pp. 409-417. doi: 10.1016/j.cropro.2005.06.007.

Nanni, L., Maguolo, G. and Pancino, F. (2020) 'Insect pest image detection and recognition based on bio-inspired methods', *Ecological Informatics,* 57, pp. 101089. doi: 10.1016/j.ecoinf.2020.101089.

Redmon, J., Divvala, S., Girshick, R. and Farhadi, A., 2016. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788).

Stern, V.M.R.F., Smith, R., van den Bosch, R. and Hagen, K., 1959. The integration of chemical and biological control of the spotted alfalfa aphid: the integrated control concept. *Hilgardia*, *29*(2), pp.81-101.

Wang, J., Lin, C., Ji, L. and Liang, A. (2012) 'A new automatic identification system of insect images at the order level', *Knowledge-Based Systems,* 33, pp. 102-110.

Zhang, W., and Swinton, S.M. (2009) 'Incorporating natural enemies in an economic threshold for dynamically optimal pest management', *Ecological Modelling,* 220(9-10), pp. 1315-1324.