EARLY DETECTION OF CATERPILLARS USING ARTIFICIAL INTELLIGENCE

Sergejs Kodors[1], Ilmars Apeinans[1], Imants Vancans[1], Toms Bartulsons[2], Imants Zarembo[1]
[1]Rezekne Academy of Riga Technical University, Latvia; [2]Institute of Horticulture (LatHort), Latvia
sergejs.kodors@rta.lv, ilmars.apeinans@rtu.lv, toms.bartulsons@gmail.com, imants.zarembo@rta.lv

Abstract. Entomofauna are essential for agricultural productivity and ecological balance. Understanding interactions among pollinators and pests is key to sustainable agriculture. Caterpillars, as pests, damage crops by feeding on tissues and spreading diseases, and reducing the yield and its quality. Therefore, early detection of caterpillars is essential for agriculture, it can be achieved using artificial intelligence embedded into agrobots or unmanned aerial vehicles following the concept of precision farming. This study introduces the dataset Caterpillar640, which contains 1,300 images of caterpillars, which were captured in natural conditions in Latvia. The images are annotated and prepared for artificial intelligence training by using YOLO architecture. The collected dataset is publicly available and distributed under an open license, providing valuable resources for research in precision farming and biodiversity monitoring. The artificial intelligence was developed using three convolution neural network architectures: YOLOv8, YOLOv9 and YOLOv10. The medium-sized models (YOLOv8m, YOLOv9m, and YOLOv10m) were selected for the comparison experiment. The comparison showed that YOLOv8m achieved the best results 0.887 mAP@0.50, YOLOv9m and YOLOv10m achieved 0.873 mAP@0.50 and 0.859 mAP@0.50, respectively. The experiment was conducted by repeating the training of each neural network model five times. The trained models can be embedded into IoT sensors, agrobots and drones for autonomous monitoring of fields, orchards and gardens offering an efficient solution for pest and biodiversity monitoring. Additionally, the dataset Caterpillar640 can be applied by researchers to train general convolutional neural networks for agricultural tasks combining it with other agricultural datasets that can impact on the accuracy of pre-trained models.

Keywords: artificial intelligence, caterpillar, object detection, pest insects, precision farming.

Introduction

Entomofauna are essential for agricultural productivity and ecological balance. Diverse insect populations increase agroecosystem resilience and support conservation. Understanding interactions among pollinators and pests is key to sustainable agriculture. Pollinators – such as bees, butterflies, and beetles – enable fruit and seed production through pollen transfer. Approximately 75% of global food crops depend on animal pollination, underscoring their importance for food security and biodiversity. Insufficient pollination could significantly reduce yields of key crops posing risks to both farmers and consumers [1; 2].

In contrast, pest insects, including aphids, caterpillars, and beetles, feed on crop tissues and act as vectors for disease transmission, causing significant damage that reduces yield and quality while impacting global food security, biodiversity, and ecosystem stability. Insect injury to leaves induces metabolic and physical changes in host, including altered CO_2 assimilation, increased water loss, and modified nutrient concentrations that influence plant growth, yield, and fitness. The environmental and economic costs of plant diseases and pests are substantial. This dichotomy highlights the critical need for managing insect populations to promote beneficial species in agriculture while mitigating the impact of harmful pests [1; 3].

The interactions between plants, pathogens, and pests present challenges in agriculture, requiring sustainable management strategies. Overreliance on chemical pesticides has led to environmental pollution, pesticide resistance, and harm to beneficial organisms. Indiscriminate use contaminates soil and water, affecting biodiversity. Sustainable pest control methods are essential to maintaining agricultural productivity while preserving ecosystem health [2].

The agricultural industry faces growing challenges from rising population demands and climate change, which accelerate the spread of pests and pathogens. Smart horticulture offers a solution by integrating publicly available data, sensor inputs, and decision-making systems. This study enhances sustainable pest management by combining technology and ecology, focusing on early detection of caterpillars by using artificial intelligence. The early detection is achieved through the application of drones and agrobots, which continuously monitor fields and orchards. Automation provides advantages over human monitoring, including faster and more consistent coverage of large areas. Considering precision farming principle, this approach allows human focuses on only areas where pests are detected, optimizing human resources rather than conducting broad and time-consuming inspections. Also, agrobots will be capable of processing pests autonomously without human involvement in the future.

Lima et al. (2020) presented the comprehensive review of automatic insect pest detection techniques [4]. The design of an autonomous early warning system was described by Liao et al. (2012) [5] and On and Abubacker (2024) [6]. The system consists of a monitoring sensor or smart insect trap, insect recognition using artificial intelligence, local (edge-computing) or remote (cloud solution), wireless communication, warning and decision support software.

The modern development of artificial intelligence is based on the usage of annotated datasets and application of robust architectures. The most popular object detection architecture is YOLO (You Only Look Once). YOLO is a real-time object detection solution introduced by Redmon et.al. (2016) [7]. It was introduced in 2015 and since then it has been continuously improved by large groups of people. Each model comes with some benefit that outperforms previous models, for example, YOLOv8 has improved architecture for work with small objects [8], that is often important for agriculture domain. For example, On and Abubacker (2024) trained insect detection neural networks using the dataset with 29 species of insect pests and three architectures YOLOv8, YOLOv9 and YOLOv10 [6]. Meanwhile, Tetila et al. (2024) applied the dataset with 12 species and trained YOLOv3 [9].

In this study a new dataset called Caterpillar640 is presented. Caterpillar640 includes the annotated natural images of caterpillars. In this article, three YOLO architectures will be experimentally compared: YOLOv8 [10], YOLOv9 [11] and YOLOv10 [12]; with a goal to find effective architecture for caterpillar detection. The experiment showed the next results with the dataset Caterpillar640: YOLOv8m achieved the best accuracy 0.887 mAP:50, but YOLOv9m and YOLOv10m achieved 0.873 mAP:50 and 0.859 mAP:50 respectively.

Materials and methods

This study was a pilot experiment to understand the specifics of domain and select experimentally the most appropriate YOLO architecture. The team collected all species of caterpillars, as the artificial intelligence was trained to detect caterpillars in general. However, the future dataset must be fine-grained to improve detection accuracy and divide caterpillars on pests and pollinators (e.g. butterflies).

The dataset was called “Caterpillar640”. The images were collected manually using mobile phones. Image annotation was done using a web tool “makesense.ai”. A single category “caterpillar” was labelled independently on the species. The images were cropped from original images (3000x4000px) to 640x640px resolution images, which are optimal for the pretrained YOLO models. The images, which did not contain caterpillars, were removed. In the result, the dataset contains 1300 annotated images. The examples of the images are presented in Fig. 1. The dataset Caterpillar640 is available under CC-BY4.0 license in Kaggle [13].

Fig. 1. **Examples of the images from Caterpillar640**

When the dataset was prepared, the actual YOLO architectures – YOLOv8, YOLOv9, and YOLOv10 – were selected at the time of the experiment. The middle size models pretrained on COCO dataset: YOLOv8m, YOLOv9m and YOLOv10m, were applied in the experiment as the trade-off between an accuracy and a latency. The training of neural networks was completed using GigaByte GeForce RTX™ 2060 WINDFORCE OC 6G with 6GB VRAM. The default augmentation of frameworks was used for model training.

Before training of YOLO models, 10% of the original dataset (130 images) was removed and saved as a testing dataset. The testing dataset is used to understand the accuracy in real-life. The remaining images were randomly divided into training and validation datasets for each of the training attempt using the random shuffle method (using Python script). This process was repeated five times for each YOLO model selected for the experiment. The division was 80% (936 images) for the training dataset and 20% (234 images) for the validation dataset. YOLOv8m, YOLOv9m and YOLOv10m models were trained separately, leading to the creation of five distinct trained models for each architecture. It was done to construct a box-plot diagram and understand the robustness of accuracy dependent on random training factors.

Results and discussion

The experiment showed that YOLOv8m achieved the best results, 0.887 mAP:0.50 and 0.634 mAP:0.95. Meanwhile, YOLOv9m and YOLOv10m achieved 0.873 mAP:0.50 and 0.859 mAP:0.50 respectively (see Fig. 2 and Table 1). The model YOLOv8m provided the most stable results, but the lowest results are presented by YOLOv9m.

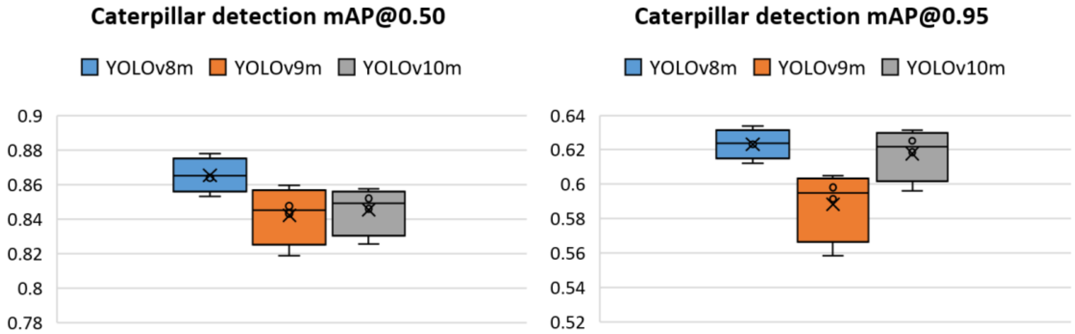

Fig. 2. **Accuracy of YOLO models trained on Caterpillar640 (testing dataset)**

Table 1

Accuracy of YOLO models trained on Caterpillar640 (testing dataset)

Model	Accuracy	Min	Mean	Median	Max
YOLOv8m	mAP@0.50	**0.853**	**0.866**	**0.864**	**0.878**
YOLOv9m	mAP@0.50	0.819	0.843	0.847	0.860
YOLOv10m	mAP@0.50	0.825	0.846	0.852	0.857
YOLOv8m	mAP@0.50:0.95	**0.612**	**0.624**	0.623	**0.634**
YOLOv9m	mAP@0.50:0.95	0.558	0.592	0.598	0.605
YOLOv10m	mAP@0.50:0.95	0.596	0.619	**0.625**	0.631

On and Abubacker (2024) obtained comparable results with insect detection: YOLOv8-10 showed the range 0.881 – 0.884 mAP@0.50 and 0.629 – 0.643 mAP@0.50:0.95; with the baseline models. However, if the models are specially tuned, the accuracy can achieve 0.951 – 0.967 mAP@0.50 and 0.733 – 0.771 mAP@0.50:0.95 [6]. The simple increase of dataset is not an effective solution to improve the accuracy of models. Apeinans (2024) studied the impact of the dataset size on the accuracy of YOLOv8n [14]. His experiment showed that the maximal accuracy increase is obtained within the first 500 images. Meanwhile, the addition of new category provides more sufficient accuracy increase, if the size of the dataset is larger than 500 images. Therefore, the fine-grained dataset can be an effective solution to improve accuracy. Another approach, YOLO models can be pretrained on the large domain dataset like IP102 [15].

Caterpillar640 dataset contains only one category "caterpillar". The future research should be focused on the granularity of the dataset collecting and annotating pest species of caterpillars. The autonomous monitoring of caterpillars requires overcoming challenges like their concealment in foliage, placement in the ground, or small size. Fig. 3 illustrates the conceptual framework for categorizing caterpillar species based on the detection methods using computer vision.

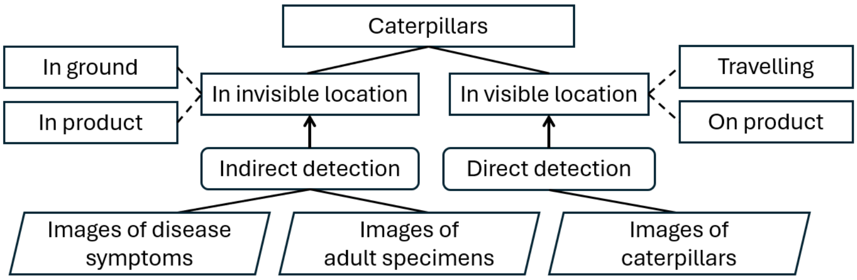

Fig. 3. **Detection methods of caterpillar using computer vision**

Regarding Latvia, the caterpillars, which can be directly detected in visible locations, are leaf-feeding caterpillars like Cone pitch moth (*Argyroploce variegana*), Turnip sawfly (*Athalia rosae*) or Currant sawfly (*Nematus ribesii*). The detection of caterpillars, which are in the ground or in product, is more complex due to their invisible state. However, they can be detected through the external symptoms or adult specimens (distributors). For example, the larvae of the green pug moth (*Pasiphila rectangulata*) damage pear buds, leaves, and flowers. They precisely weave both edges of the leaves together from the petiole to the tip and feed within the woven structure. The larvae of *Phytomyza horticola* mine legume leaves beneath the leaf epidermis. As the larva grows, the mines expand and spread in various directions. The larvae of the pear midge (*Contarinia pyrivora*) damage the young fruit pulp as they feed, creating a cavity inside. The affected fruits initially grow faster and appear rounder than healthy ones. The larvae of the apple sawfly (*Hoplocampa testudinea*) tunnel small, curved galleries just beneath the fruit epidermis, causing the affected tissue to cork over. Later, the larvae burrow deeper into the fruit, consuming the core. The walls of the tunnels turn brown or black.

There are different monitoring methods and techniques. Lima et al. (2020) described the detection techniques, which are based on usage of smart traps [4]. The weakness of smart traps is their static location. Another approach is mobile application. Christakakis et al. (2024) described the technologies and design of cross-platform mobile application, which can recognize plant diseases and insect pests. AI was developed using YOLOv8 achieving accuracy 0.70 mAP@0.50 [16]. The disadvantage of mobile application is characterized by two factors: (1) it is useful for non-professional farmers and infrequent usage; (2) it is not useful for large commercial fields or orchards, because the periodic manual monitoring is time-consuming. Crupi et al. (2025) proposed the idea to apply the swarm of autonomous nano-UAVs for pest control [17]. Considering the effectiveness and usability, the application of UAVs and UGVs is the most appropriate, because vehicles can move around monitoring location. However, the technology is not sufficiently developed at this moment, and it is impacted by different risks: crash possibility, precision of coordinates, sensitivity to weather conditions, etc.

Conclusions

1. The new dataset Caterpillar640 is presented in the article. The dataset contains the annotated natural images of caterpillars. The caterpillars are not classified on species (single-class dataset). The dataset is available in Kaggle under CC-BY4.0 license.
2. Three middle-size YOLO models were compared using the dataset Caterpillar640. The best accuracy was shown by YOLOv8m model: 0.887 mAP:0.50 and 0.634 mAP:0.95.
3. The caterpillars must be grouped into two categories: (1) caterpillars, which can be directly detected by video/photo camera; (2) caterpillars, which can be detected indirectly through disease symptoms or adult specimens.
4. The future study can be continued by developing the dataset of insect pests actual to Latvia.

Author contributions

Conceptualization, S.K.; methodology, S.K.; software, I.V. and I.A.; validation, S.K., I.Z. and I.A; formal analysis, I.A.; investigation, S.K. and I.A.; data curation, S.K., I.V. and I.A.; writing – original draft preparation, S.K., T.B., I.A.; writing – review and editing, S.K., I.Z., I.A. and T.B.; visualization, S.K. and I.A. All authors have read and agreed to the published version of the manuscript.

References

[1] Gazi S. Entomofauna of Agricultural Crops: Roles, Impacts, and Ecological Significance. Nature & Science International Scientific Journal, vol. 6, issue 11, 2024, pp. 15-19.

[2] Cohen M., Wohlmuth H., Williams C., Clarke P. The Impact of Plant Pathology: Examining Diseases and Pests in The Plant Kingdom and Strategies for Effective Control and Management. Australian Herbal, 2023, pp 1-6.

[3] Oleksyn J., Karolewski P., Giertych M.J., Zytkowiak R., Reich P.B., Tjoelker M.G. Primary and secondary host plants differ in leaf-level photosynthetic response to herbivory: evidence from Alnus and Betula grazed by the alder beetle, Agelastica alni. New Phytologist, vol. 140, 1998, pp 239-249.

[4] Lima M. C. F., De Almeida Leandro M. E. D., Valero C., Coronel L. C. P., Bazzo C. O. G. Automatic Detection and Monitoring of Insect Pests – A Review. Agriculture, vol. 10, issue 5, 2020, pp. 161-184.

[5] Liao M., Chuang C., Lin T., Chen C., Zheng X., Chen P., Liao K., Jiang J. Development of an autonomous early warning system for Bactrocera dorsalis (Hendel) outbreaks in remote fruit orchards, Computers and Electronics in Agriculture, vol. 88, 2012, pp. 1-12, ISSN 0168-1699.

[6] On W. M., Abubacker N. F. YOLO-Driven Lightweight Mobile Real-Time Pest Detection and Web-Based Monitoring for Sustainable Agriculture. International Journal of Advanced Computer Science and Applications(ijacsa), vol. 15, Issue 12, 2024.

[7] Redmon J., Divvala S., Girshick R., Farhadi A. You Only Look Once: Unified, Real-Time Object Detection. Proceedings of the “2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)”, June 27-30, 2016, Las Vegas, NV, USA, pp. 779 – 788.

[8] Bai R., Shen F., Wang M., Lu J., Zhang Z. Improving Detection Capabilities of YOLOv8-n for Small Objects in Remote Sensing Imagery: Towards Better Precision with Simplified Model Complexity. June 22, 2023.

[9] Tetila E. C., Da Silveira F. a. G., Da Costa A. B., Amorim W. P., Astolfi G., Pistori H., Barbedo J. G. A. YOLO performance analysis for real-time detection of soybean pests. Smart Agricultural Technology, vol. 7, 2024.

[10] Ultralytics, YOLOv8 GitHub repository. [online] [01.03.2025]. Available at: https://github.com/ultralytics/ultralytics

[11] WongKinYiu, YOLOv9 GitHub repository. [online] [01.03.2025]. Available at: https://github.com/WongKinYiu/yolov9

[12] THU-MIG, YOLOv10 GitHub repository. [online] [01.03.2025]. Available at: https://github.com/THU-MIG/yolov10

[13] Caterpillar640, Caterpillar640 dataset Kaggle repository. [online] [24.02.2025]. Available at: https://www.kaggle.com/datasets/projectlzp201910094/caterpillar640

[14] Apeinans I. Optimal Size of Agricultural Dataset for YOLOv8 Training. Proceedings of the 15th International Scientific and Practical Conference “Environment. Technology. Resources.”, June 27-28, 2024, Veliko Tarnovo, Bulgaria, pp. 38-42.

[15] Wu X., Zhan C., Lai Y., Cheng M.-M., Yang J. IP102: A Large-Scale Benchmark Dataset for Insect Pest Recognition. Proceedings of the “2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)”, June 15-20, 2019, Long Beach, CA, USA, pp. 8779-8788.

[16] Christakakis P., Papadopoulou G., Mikos G., Kalogiannidis N., Ioannidis D., Tzovaras D., Pechlivani E. M. Smartphone-Based Citizen Science Tool for Plant Disease and Insect Pest Detection Using Artificial Intelligence. Technologies, vol. 12, issue 7, 2020.

[17] Crupi L., Butera L., Ferrante A., Giusti A. An Efficient Ground-aerial Transportation System for Pest Control Enabled by AI-based Autonomous Nano-UAVs. 2025 ACM Journal on Autonomous Transportation Systems, vol. 1, 2025.