Exploring YOLO Algorithm application in Smart Traps for Fruit Pest Detection

Eduardo M. Behling[1], **Marilton S. de Aguiar**[1]

[1]Universidade Federal de Pelotas (UFPEL)
Pelotas - RS - Brazil

{embehling,marilton}@inf.ufpel.edu.br

Abstract. *Agricultural practices are essential for modern civilization and have significantly contributed to population growth in recent years. Food production depends heavily on the natural environment; however, pests can severely damage crops and are often difficult to control. This emphasizes the urgent need for effective pest control measures to protect agriculture. One common way to manage insect pests is by using traps. These traps typically employ bait to attract insects, often consisting of sexual pheromones or food, and they have a mechanism that captures the insect when it attempts to enter. This work explores the training and application of various size variants and versions of the modern object classification algorithm, You Only Look Once (YOLO), during the classification phase of real-time object detection. The focus is on two predominant pests associated with fruit cultivation: Ceratitis capitata and Grapholita molesta. The study utilizes a limited dataset to maximize performance and computational metrics in an intelligent trap-controlled environment.*

1. Introduction

Agricultural cultivation has been the backbone of modern society, contributing to significant demographic growth in recent years. However, this population increase has resulted in a high demand for agricultural products, particularly food. The natural environment heavily influences food cultivation, and pests present a significant challenge, often causing considerable damage to crops and proving difficult to manage. This underscores the urgent need for effective systems to control pest populations and mitigate their impact on agriculture [Tudi et al. 2021].

One of the most common solutions for pest management is the use of pesticides. While pesticides can effectively control pests, research has linked them to adverse effects on human health and the environment, as noted by [Kim et al. 2017, Ghafarifarsani et al. 2024]. The Integrated Pest Management (IPM) framework has been developed to reduce reliance on chemical solutions. This approach emphasizes preventive measures, monitoring, and decision-making based on established thresholds rather than solely relying on reactive pesticide applications.

A widely used technique for controlling insect pest populations is using traps. These traps typically consist of an attractive bait, which often includes sexual pheromones or food, as well as a system that captures the insect upon entry. This paper explores the results of employing the YOLO family of classification algorithms in the classification phase of the pipeline introduced by the smart variant of this technique, designed by [Freitas et al. 2022]. This variant automates the monitoring task that would need to be performed manually. It utilizes a V2 8 MP camera driven by a Raspberry Pi III, aiming to detect the pests *Ceratitis capitata* and *Grapholita molesta*. These pests are significant nuisances to fruit, adapting well to modern urban and rural environments, and they are prevalent in various regions across Brazil, particularly in the states of Rio Grande do Sul and São Paulo [Dias et al. 2023].

2. Material and Method

2.1. Dataset

The dataset comprises 22475 sections of images, which have been segmented and obtained using an instance of the smart trap equipped with a V2 8 MP camera. Each section represents a 100×100 pixel area of interest, and the images are labeled as either *Ceratitis* (C. capitata), *Grapholita* (G. molesta), or classified as other (including debris or different insects). Furthermore the dataset was arbitrarily divided into three subsets, maintaining the original class proportion (*Ceratitis*: 53.02%, *Grapholita*: 27.46% and *Others*: 19.53%), 15731 (70%) entries were allocated to the training set, 2248 (10%) to the validation set, and 4496 (20%) to the test set. The training set was used to train the model, the validation set was used to validate and adjust the model at end of each epoch and the test set was used to assess the model's performance on an isolated dataset to better estimate its real-world effectiveness. In Figure 1, is presented a sample of images in the dataset.

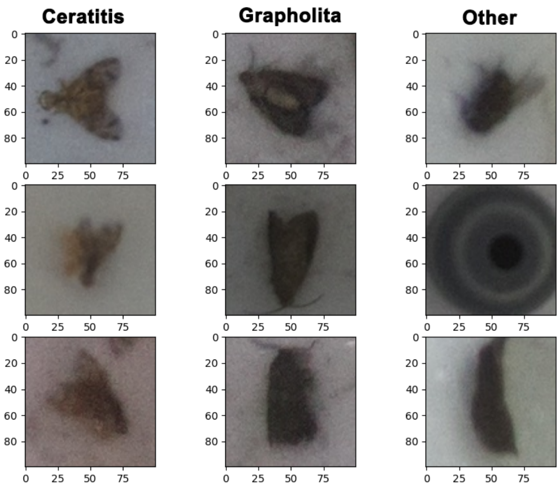

Figure 1. Sample of images provided by the dataset.

2.2. Object Detection

Object detection involves two key tasks: object classification, which assigns classes to objects within an image, and object localization, which consists of placing an enclosing bounding box around each object [Vijayakumar and Subramaniyaswamy 2024]. Real-time object detection has become a trending task, significantly enhanced by the improved performance of deep learning algorithms in this area [Ragab et al. 2024].

An essential aspect of these tasks is classifying areas of interest identified during the segmentation phase. In this application, the accuracy and speed of the classifier are crucial. Therefore, this article focuses on utilizing variants from the YOLO algorithm family, which are well-known for their fast and lightweight solutions for various computer vision tasks. In particular, this paper analyzes the classification task variants of YOLOv8 [Reis et al. 2024] and YOLOv11 [Khanam and Hussain 2024]. It employs the Nano and Small versions to determine the most suitable option, using precision and inference time as the primary metrics.

2.3. Development Environment

The development environment used the YOLO framework in Python, which Ultralytics provides. Since the target application is intended for embedded devices, we separated the environments into two hardware setups: training and testing. This distinction optimizes training performance, reduces training time, and offers a more accurate estimate of performance on devices with lower specifications. The training hardware consisted of a desktop featuring an Intel Core i7-2600 processor operating at eight cores, each at 3.8 GHz, and an NVIDIA RTX 3060 GPU with 12GB of RAM. This setup significantly accelerated training time. In contrast, the testing hardware was a desktop equipped with an AMD Ryzen 3400G CPU and integrated AMD Radeon RX Vega 11 graphics.

3. Results and Discussion

Analyzing the findings expressed in Table 1 and the matrices in Figure 2, it is evident that the newer YOLOv11 demonstrates superior accuracy when comparing the different versions. Additionally, there is a tendency for accuracy to increase with the size of the model. However, this increase in accuracy comes at the cost of significantly lower frames per second (FPS). As a result, this application yields diminishing returns, especially in embedded devices, making the Small and Nano variants ideal choices for these systems.

Table 1. Benchmark results per model

Model	Accuracy (Top 1)	Macro F1-Score	FPS	Training Time (hh:mm:ss)
YOLOv8-cls nano	0.9911	0.9888	**185**	01:05:07
YOLOv8-cls small	0.9931	0.9915	71	**01:05:01**
YOLOv11-cls nano	0.9893	0.9868	177	01:19:59
YOLOv11-cls small	**0.9935**	**0.9919**	69	01:26:50

4. Conclusion and future works

In conclusion, the Small and Nano variants in real-time applications on devices with limited capabilities offer excellent frames per second and high accuracy. This enables synchronous monitoring, significantly reducing the need for pesticides and promoting a more effective and quicker method for pest control. In future work, adapting the dataset would be beneficial to facilitate the training of proper object detection algorithms, such as the mainline object detection YOLO models, while also expanding the classification techniques by exploring other deep learning approaches like MambaVision.

References

Dias, N. P., Silva, F. F. D., Abreu, J. A. D., Pazini, J. D. B., and Botta, R. A. (2023). Infestation level of fruit flies in the boundary zone in rio grande do sul, brazil. *Rev. Ceres*.

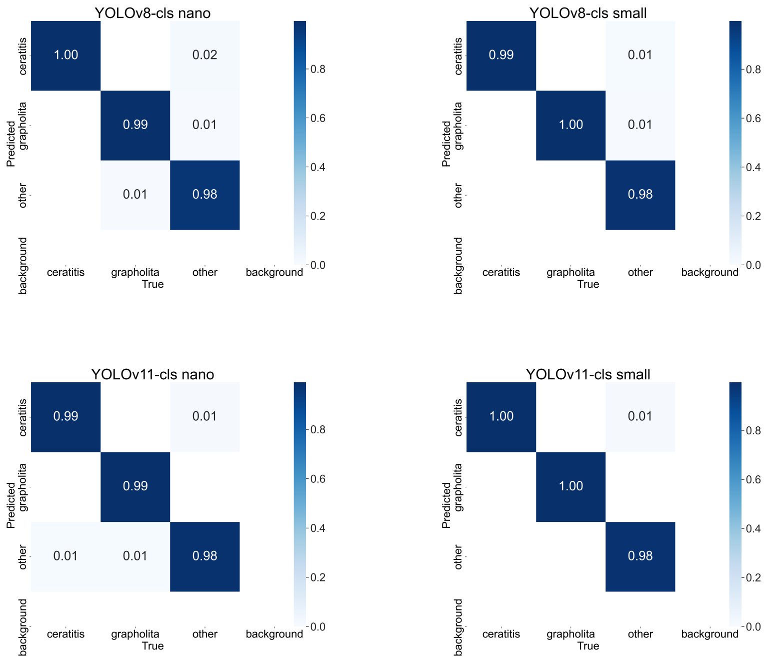

Figure 2. Confusion matrices.

Freitas, L., Martins, V., de Aguiar, M., de Brisolara, L., and Ferreira, P. (2022). Deep learning embedded into smart traps for fruit insect pests detection. *ACM Trans. Intell. Syst. Technol.*, 14(1).

Ghafarifarsani, H., Rohani, M. F., Raeeszadeh, M., Ahani, S., Yousefi, M., Talebi, M., and Md, S. H. (2024). Pesticides and heavy metal toxicity in fish and possible remediation – a review. *Annals of Animal Science*, 24(4):1007–1024.

Khanam, R. and Hussain, M. (2024). Yolov11: An overview of the key architectural enhancements.

Kim, K.-H., Kabir, E., and Jahan, S. A. (2017). Exposure to pesticides and the associated human health effects. *Science of The Total Environment*, 575:525–535.

Ragab, M. G., Abdulkadir, S. J., Muneer, A., Alqushaibi, A., Sumiea, E. H., Qureshi, R., Al-Selwi, S. M., and Alhussian, H. (2024). A comprehensive systematic review of yolo for medical object detection (2018 to 2023). *IEEE Access*, 12:57815–57836.

Reis, D., Kupec, J., Hong, J., and Daoudi, A. (2024). Real-time flying object detection with yolov8.

Tudi, M., Ruan, H. D., Wang, L., Lyu, J., Sadler, R., Connell, D., Chu, C., and Phung, D. T. (2021). Agriculture development, pesticide application and its impact on the environment. *International Journal of Environmental Research and Public Health*, 18(3):1112. Research Support, Non-U.S. Gov't, Review.

Vijayakumar, Ajantha, V. and Subramaniyaswamy (2024). Yolo-based object detection models: A review and its applications. *Multimedia Tools and Applications*, 83(35):83535–83574.