Smart Agricultural Technology 15 (2026) 102698

Contents lists available at ScienceDirect

Smart Agricultural Technology

journal homepage: www.journals.elsevier.com/smart-agricultural-technology

Benchmarking one-stage object detectors for field-based detection of Flavescence dorée vectors

Matteo Dho [a], Alessia De Luigi [a], Elena Gonella [a,*], Gianluca Manduca [b,c], Donato Romano [b,c], Alberto Alma [a]

[a] Department of Agricultural, Forest and Food Sciences, University of Torino, Italy
[b] The BioRobotics Institute, Sant'Anna School of Advanced Studies, Pontedera, PI, Italy
[c] Department of Excellence in Robotics and AI, Sant'Anna School of Advanced Studies, 56127, Pisa, Italy

ARTICLE INFO

Keywords:
Convolutional neural network
Smart trap
Grapevine
Integrated pest management
Phytoplasma vector

ABSTRACT

Deep learning-based object detection has emerged as a promising paradigm for automating high-throughput insect monitoring in agricultural environments. In this context, we propose a computer vision framework for automated detection of Flavescence dorée (FD) vector insects from yellow sticky trap (YST) imagery, addressing the scalability limitations of manual entomological inspection.

Flavescence dorée is a severe phytosanitary threat to viticulture in Europe, primarily transmitted by the leafhopper *Scaphoideus titanus*, with *Orientus ishidae* increasingly recognized as a secondary vector due to its role in phytoplasma transmission dynamics. Current monitoring strategies rely on field-deployed YSTs followed by manual expert analysis, which is labor-intensive, time-consuming, and difficult to scale to large vineyard networks.

To address these limitations, a custom dataset of images acquired from field-deployed YSTs using a fixed-distance laboratory imaging setup was developed and used to fine-tune and benchmark multiple modern one-stage object detection architectures. Repeated training runs showed generally stable performance, with average mAP50 values above 0.93 across all configurations. For the best-performing configuration, the highest performing run achieved mAP50 = 0.965, mAP50-95 = 0.690, precision = 0.949, and recall = 0.914. Edge benchmarking further quantified deployment feasibility, with YOLO26n reaching 11.41 FPS on Raspberry Pi 5.

Class-wise evaluation confirmed accurate detection of both *S. titanus* and *O. ishidae*, despite heterogeneous trap backgrounds and illumination conditions. These image-level results support further evaluation of automated FD-vector monitoring, while trap-level independence, whole-trap counting, and fully operational field acquisition remain to be validated. The annotated dataset and benchmark provide a basis for further research on data-efficient agricultural computer vision systems.

1. Introduction

Flavescence dorée (FD) is among the most economically significant grapevine diseases in Europe, posing a critical threat to viticultural systems and wine production stability [1]. The causal agents are phytoplasmas belonging to the 16SrV-C and 16SrV-D subgroups, classified as quarantine organisms by the European Food Safety Authority (EFSA) [2]. Since antibiotic treatments are prohibited in European agriculture [3,4], disease containment relies primarily on uprooting infected plants and mandatory insecticide interventions targeting the principal vector *Scaphoideus titanus* Ball (Hemiptera: Cicadellidae) [5,6].

Scaphoideus titanus is a univoltine, ampelophagous leafhopper native to North America [7] and first reported in Europe in 1958 [8]. Adult populations occur in vineyards from mid-June to late October, with females laying approximately 60–65 eggs [9,10]. In addition to this primary vector, multiple leafhopper species have been implicated in the epidemiology [11]. Among them, the mosaic leafhopper *Orientus ishidae* (Matsumura) (Hemiptera: Cicadellidae) has gained increasing attention due to its feeding behavior on potential phytoplasma reservoir plants and its mobility across vineyard edges and adjacent habitats, which may facilitate pathogen spread at the landscape scale [12,13]. Other vectors are *Dictyophara europaea* (L.) (Hemiptera: Dictyopharidae), *Allygus* spp.

* Corresponding author.
E-mail address: elena.gonella@unito.it (E. Gonella).

https://doi.org/10.1016/j.atech.2026.102698
Received 14 July 2026; Received in revised form 29 October 2026; Accepted 1 November 2026
Available online 2 November 2026
2772-3755/© 2026 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

or recently *Lamprotettix nitidulus* (Fabricius) (Hemiptera: Cicadellidae) [14,15]. In northern Italy, only *D. europaea* was previously found infected by FD-related phytoplasmas in vineyards; however, the main phytoplasma strain found in symptomatic vines, namely M54, has not been identified so far in field-collected individuals of this species [16, 17].

Surveillance of vectors is a mandatory component of Integrated Pest Management (IPM) programs and is regulated across European vineyards under Commission Implementing Regulation (EU) 2022/1630. Monitoring is primarily conducted using yellow sticky traps (YSTs) [5, 13]. However, due to the coexistence of primary and secondary infection pathways, coupled with heterogeneous vineyard landscapes and vector dispersal dynamics, dense spatial deployment of traps is required to reliably characterize population dynamics and identify invasion corridors from surrounding environments [10,18].

Despite its importance, current monitoring practice mostly relies on manual inspection of collected traps by trained entomologists. This process does not scale efficiently to large vineyard areas and introduces operational bottlenecks in surveillance pipelines. Furthermore, morphological similarity among co-occurring leafhopper species [19, 20] increases the difficulty of accurate field-level identification, reinforcing the need for expert taxonomic validation.

To address these limitations, recent research has explored the integration of computer vision into entomological monitoring systems. In particular, convolutional neural network (CNN)-based object detection frameworks have emerged as a robust solution for automated insect recognition across agricultural and ecological domains [21–23]. Early applications demonstrated feasibility in detecting *Cydia pomonella* in pheromone traps [21], while subsequent studies extended CNN-based detection to broader pest groups, often without fine-grained species classification [24,25].

More recently, data-driven pest monitoring systems have shifted toward end-to-end detection pipelines capable of operating directly on trap imagery. Li et al. [26] demonstrated automated detection of thrips and whiteflies on YSTs with strong agreement between model outputs and manual counts. In parallel, one-stage detectors have become increasingly prevalent due to their favorable trade-off between inference speed and accuracy, making them particularly suitable for real-time field applications [27–29].

Benchmark datasets and comparative studies have further advanced the field. Jain et al [30] developed a dataset curated by entomologists for moth species detection in wild conditions. Using a previously published dataset, Deserno and Briassouli [31,32] achieved high performance on multiple greenhouse pest species, albeit requiring background preprocessing. YOLO-based architectures have also demonstrated superior performance compared to alternative detection frameworks such as EfficientDet, SSD, and Faster R-CNN in detecting *Scirtothrips dorsalis* [33].

Additional studies have confirmed the feasibility of species-level identification in complex trap environments, including real-time detection of thrips and whiteflies [34]. Cost-efficient smart-trap systems have been developed for field monitoring of Dipteran pests such as *Hippobosca equina* and *Stomoxys calcitrans* [35], while recent findings indicate that high detection accuracy can also be achieved under non-ideal imaging conditions, including low-quality field-acquired data [36].

Despite this progress, only limited work has addressed FD vector detection specifically on YST imagery. To our knowledge, a single study conducted in northeastern Italy has investigated *S. titanus* and *O. ishidae* detection using a prototype smart-trap for background images and mainly scanned traps for object-containing images, comparing single-stage and two-stage detectors and reporting superior performance for the former [27]. Consequently, the availability of correctly annotated datasets derived from field-deployed YSTs and acquired under a standardized imaging configuration remains limited. Such datasets are needed to develop and evaluate automated systems for both in-field monitoring and laboratory-based YST inspection. Furthermore, new generations of the popular YOLO architectures have recently been released [37,38], and the benefits of architectural advances remain either poorly or yet not benchmarked in entomological applications [39, 40].

Accordingly, this study makes two main contributions: i) the construction of an annotated field-derived FD vectors dataset originated from YSTs deployed in one of Italy's major viticultural areas and photographed at a fixed distance; and ii) a comparative benchmark of more recent and already established one-stage detectors of the YOLO family under a common experimental protocol.

2. Materials and methods

2.1. Data collection

Images were mostly taken from field-placed YSTs positioned in intensive vineyard agroecosystems. YSTs were placed during Summer 2025 in Dogliani (Cuneo Province, CN, Piedmont, Italy, 44°33′25.6"N 7°57′45.7"E), Serralunga d'Alba (CN, Italy, 44°38′36.0"N 7°58′53.1"E), Chieri (Torino Province, TO, Piedmont, Italy, 45°00′58.5"N 7°47′20.0"E) and Grugliasco (TO, Italy, 45°03′51.5"N 7°35′35.7"E) inside and on the perimeter of different vineyards with cultivars Dolcetto, Chardonnay, Barbera, Pinot noir, Sauvignon, Moscato and Nebbiolo. Most of the traps used came from vineyard borders, due to a general higher presence of both targets on YSTs [12,41]. Traps were collected two weeks after field hanging. Vineyards were selected as previous monitoring activities revealed moderate to high abundance of both vectors. Vineyards in Cuneo Province were selected for the presence of hazelnut groves (a highly preferred host plant for *O. ishidae*) close to the borders, thus increasing the chance of finding a composite leafhopper community, and more specifically to retrieve *O. ishidae* adults [12]. To increase the diversity of sources for image collection, two types of YSTs from different manufacturers were used, namely Glutor (Biogard®, 40 × 25 cm) and Goldentrap (GEA®, 16 × 25.5 cm). Details about collection dates, number and type of YSTs are reported in Table 1. Moreover, one laboratory-prepared trap was assembled by artificially positioning insects to reproduce a range of realistic conditions. Adult insects were collected in the same vineyards using sweep nets (30 cm diameter), stored at −20 °C, and subsequently placed on a clean Glutor trap in dorsal, ventral, and lateral positions.

Images were taken with a 16 MP GXVISION USB camera module based on a Sony IMX298 sensor equipped with a manually adjustable

Table 1
Yellow sticky traps (YSTs) collection sites and dates. Collection date: dates are expressed as month/day/year; Perimetral and Internal traps: trap numbers are expressed as "X" or as "X + Y", with X indicating the number of Glutor YSTs and with Y the number of Goldentrap YSTs.

Vineyard location	Collection date	Perimetral traps	Internal traps	Total traps
Dogliani	07/14/2025	5 + 5	2 + 2	14
	07/28/2025	5 + 5	2 + 2	14
	08/11/2025	4 + 5	1 + 1	11
	08/25/2025	3 + 4	1 + 1	9
Serralunga d'Alba	07/15/2025	5	3	8
	07/29/2025	5	3	7
	08/12/2025	3	1	4
	08/26/2025	3	1	3
Chieri	07/21/2025	7	2	10
	08/04/2025	8	2	10
	08/18/2025	7	2	9
	09/01/2025	4	2	6
Grugliasco	07/21/2025	7	2	10
	08/04/2025	8	2	10
	08/18/2025	7	2	9
	09/01/2025	4	2	6
Total		85 + 19	30 + 6	140

fixed-focus M12 lens (6 mm focal length, f/2.0, IR filter), allowing for a minimum focusing distance of 2 cm according to manufacturer's specification. Distance between camera lens and trap surface was maintained at 7 cm for all the acquisitions. Pictures were captured with an aspect ratio of 4:3, with resolutions of 640 × 480 and 1280 × 960 pixels to reduce final dataset weight, covering an area of 7.0 × 5.0 cm. Prior to image acquisition, focus was manually adjusted observing a real-time video stream using OpenCV (v. 4.11.0.86) and confirmed by inspection of representative captured frames. To keep a fixed distance between camera lens and YSTs, a wooden structure was realized (Fig. 1).

Images acquisition was focused on increasing dataset variability. Captured areas were selected to cover different target insects densities, presence of non-target organisms, variable background conditions such as presence of debris, text printed on traps, presence of holes used for trap hanging and different traps regions, from borders up to central parts. No systematic coverage of YST surface was implied. A total of 140 field-placed YSTs (115 Glutor, 25 Goldentrap) were selected for image acquisition following manual confirmation of *S. titanus* and *O. ishidae* presence by expert personnel. Images were captured inside laboratory at different times of the day, and light conditions were not standardized, meaning that either only natural light was used or light in the room was turned on. Seldomly, a lamp was used to mimic increased light conditions. Avoiding the standardization of YSTs illumination was preferred to increase the variability of the dataset in terms of saturation and brightness. Images were taken from portions of YSTs containing variable density of *S. titanus* and/or *O. ishidae*, from one target insect up to 18 *S. titanus* and 17 *O. ishidae*. An average of 11 images were taken for each trap. The dataset was randomly split into training, validation, and test subsets at the image level; images originating from the same YST were not explicitly grouped during splitting. Details about total images acquired, number of images within the three datasets and number of *S. titanus* and *O. ishidae* individuals photographed are reported in Table 2. To quantify potential bias introduced by image-level random splitting, a subset of the full dataset was re-partitioned using both random and grouped splitting strategies. In the random split, images were randomly assigned to the training, validation, and test sets regardless of their trap of origin. In the grouped split, all images originating from YSTs included in the test partition were excluded from both the training and validation sets, thereby ensuring trap-level independence between the test set and the data used for model development.

Pictures were taken in a Python 3.13.12 environment using OpenCV library (v. 4.11.0.86) with a custom script. Labelling was done with LabelImg annotation tool (v. 1.8.6), and no image enhancement was applied as preprocessing.

Table 2
Number of images and target insects individuals used in the Training, Validation and Test sets for models training and performance evaluation. Field YSTs: Number of pictures taken from field-placed YSTs having at least one target insect; Laboratory YST: number of pictures taken from laboratory prepared YST having at least one target insect; Background: number of pictures taken from field-placed YSTs having no target insects; *S. titanus*: total number of *S. titanus* individuals present in the pictures for each set; *O. ishidae*: total number of *O. ishidae* individuals present in the pictures for each set.

Dataset	Field YSTs	Laboratory YST	Background	*S. titanus*	*O. ishidae*
Train	943	23	188	1135	1167
Validation	143	1	27	156	130
Test	211	0	33	272	275
Total	1569			1563	1572

2.2. CNN models selection and training

Labelled images were divided into train, validation and test datasets. Datasets composition in terms of number of images and number of *S. titanus* and *O. ishidae* individuals is reported in Table 2. Models training was performed on a HP Pavillion Gaming TG01 computer with Intel® Core™ i7-11700F processor (2.50 GHz, 16 MB Cache, 8 Cores, 16 Threads), 16 GB RAM, GPU NVIDIA GeForce RTX 3060 Ti (8GB VRAM). CUDA v11.8 and CUDNN v8.9.7 were used for GPU acceleration.

YOLO family architectures were selected for fine-tuning on the custom dataset due to the proved high performances in insect recognition tasks [27,36,42]. We focused on the widely used YOLOv8 [36,43,44] and on the more recent YOLO11 and YOLO26 [37] architectures.

Differently from older variants, the backbone of YOLOv8 architecture uses a modified version of CSPDarknet53 with C2f module to enhance detection accuracy and the head performs an anchor-free detection, simplifying the model and improving accuracy on small objects [45]. YOLO11 architecture has improved backbone and neck structure, with C2f module replaced by C3k2 for better computational efficiency and introducing C2PSA module to improve feature extraction [46]. YOLO26, the most recent architecture, reduced models latency by limiting head complexity. Distribution Focal Loss (DFL) for bounding boxes regression was replaced by a lighter parametrization and predictions are generated directly, without the need for NMS post-processing step. Improvements in models training involve the adoption of ProgLoss and STAL, enhancing performances on small or partially occluded objects. Finally, the implementation of MuSGD optimizer allows for a faster and stable convergence also resulting in a reduction of training time [37,47].

Medium-size variants from three YOLO generations were fine-tuned

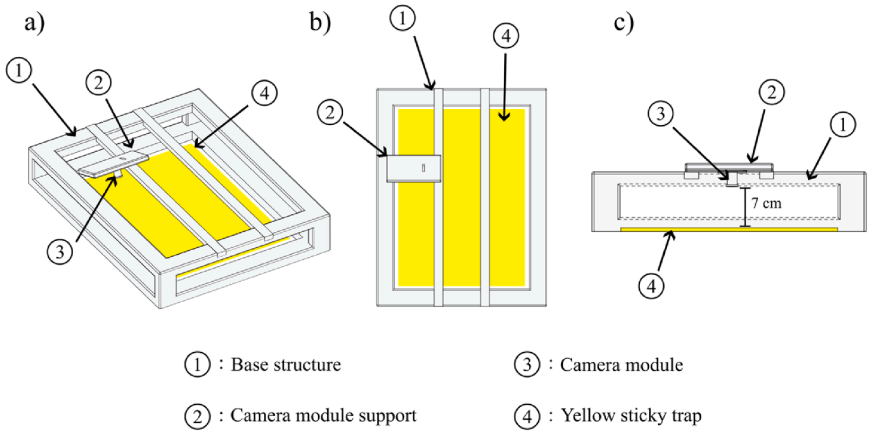


Fig. 1. Schematization of the structure used for capturing YSTs pictures from a fixed distance of 7 cm. a) Rotated view; b) Top view; c) front view with camera-trap distance. 1) Wooden box frame with spacers perpendicular to the short side; 2) base for camera module adjustment to be laid between box borders and/or spacers; 3) camera module; 4) YST. Image drawn with FreeCAD (v 1.0).

and benchmarked on the custom dataset. Detectors size was chosen considering possible applications on edge devices, aiming to compromise between detection accuracy and computational requirements. To reduce the influence of stochastic training effects a repeated runs approach was followed, with each architecture trained using five random seeds. The same seed values were applied to all evaluated models. Two variants of YOLO26m architecture were trained, one with default AdamW (YOLO26m_AdamW) optimizer and one with SGD (YOLO26_SGD). For the best performing configuration, small and nano model variants have been trained and compared with the medium-size variant.

To define a common augmentation strategy to be used for all the benchmarked models, preliminary augmentation-tuning phase was conducted using YOLOv8m. Reduced subsets of the final training and validation datasets, containing 600 and 90 images respectively, were considered. The objective of this step was to identify augmentation settings to improve models generalization while preserving morphologically important characteristics of target species. Augmentation parameters were selected via trial-and-error process, and no formal hyperparameter optimization procedure was used. For this preliminary step, models were trained for a maximum of 200 epochs and patience value was set to 30 to prevent overfitting. Starting from default Ultralytics augmentation parameters, resulting fine-tuned models detections on a subset of training and validation sets were visually inspected to identify detection errors and guide augmentation choices. As an example, default scaling factor (scale = 0.5) generated false positive (FP) detections of smaller or larger insects, resulting detrimental for a fixed-distance camera setup. Saturation parameter was reduced from 0.7 to 0.15 to preserve colour patterns important for *S. titanus* recognition. Images from the test set were not used in this step.

Once defined, the augmentation configuration was applied unchanged to all benchmarked detectors, ensuring that performance differences reflected architectural and optimization differences. Training was conducted for all the models setting a maximum of 300 epochs with a patience value of 40 to prevent overfitting. Images were automatically resized by the Ultralytics framework to match the specified input size (640 pixels), while preserving their original aspect ratio. Details of training parameters are reported Table S1.

2.3. Performance assessment

Models performance evaluation was conducted with ultralytics (v. 8.4.19) built-in validation function using the resulting best model weights for each repeated run, and the test dataset to provide ground-truth annotations. Computed metrics for each trained model were mAP50, mAP50-95, mean Precision, mean Recall and inference speed computed considering both *S. titanus* and *O. ishidae* classes, while AP50, AP50-95, Precision and Recall for single class.

Precision measures the accuracy of positive predictions by considering both true positive (TP) and FP identifications and is computed as follows:

$$Precision = \frac{TP}{TP + FP}$$

High precision of a model indicates high reliability of positive predictions. Considering the presence of disturbing elements on field-collected YSTs such as non-target insects or simultaneous presence of *S. titanus* and *O. ishidae*, high precision is important for correct monitoring activities.

Recall measures the ability of a model to correctly identify all the target objects in an image by considering both TP and false negative (FN) predictions and is calculated as:

$$Recall = \frac{TP}{TP + FN}$$

High recall is important in the case of monitoring activity to limit the risk of losing target insects, thus underestimating vectors population in field.

Since predicted bounding boxes are unlikely to perfectly match ground-truth ones provided by test dataset, the definition of Intersection over Union (IoU) is important to correctly identify positive and negative predictions. IoU is described as:

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union}$$

Area of Overlap indicates the intersection between ground-truth and predicted bounding boxes. A prediction is considered as TP at a given IoU threshold if its IoU with a ground-truth annotation of the same class exceeds the threshold value.

Average Precision (AP) is defined as the area under the Precision-Recall curve, thus considering model performance according to both previous metrics. AP is normally expressed for different IoU values. AP50 indicates AP at IoU = 0.50, while AP50-95 calculates the AP averaged on different IoU thresholds ranging from 0.50 to 0.95 at intervals of 0.05.

The mean Average Precision (mAP) is calculated averaging the AP of each single class for the different IoU considered and is computed as follows:

$$mAP = \frac{\sum_{i=1}^{N} APi}{N}$$

With N the number of classes and APi the AP value of each class.

Inference speed, expressed as frame per second (FPS), was computed as:

$$FPS = \frac{1000}{t_{preprocess} + t_{inference} + t_{postprocess}}$$

Performance metrics were summarized as mean and 95% confidence intervals across the five repeated runs. Models detection on test dataset were visualized and manually checked to identify most common sources of errors in detections.

2.4. Embedded deployment and performance evaluation

To assess the suitability of the proposed classifiers for edge deployment, the YOLO26 family was selected as the most recent architecture considered in this study and based on its identification performance observed in the previous evaluation. Accordingly, the best-performing YOLO26n, YOLO26s, and YOLO26m models were exported in NCNN format and benchmarked on two embedded platforms: a Raspberry Pi 4 Model B (Rev. 1.5), equipped with 8 GB of RAM and a quad-core ARM Cortex-A72 processor, and a Raspberry Pi 5 Model B (Rev. 1.0), equipped with 8 GB of RAM and a quad-core ARM Cortex-A76 processor. Both platforms used a 64-bit aarch64 architecture and ran Debian GNU/Linux 12 (Bookworm) and Debian GNU/Linux 13 (Trixie), respectively. The same inference environment, based on Ultralytics v8.4.79 and NCNN v1.0.20260526, was used on both platforms. Benchmarking was performed on the complete test dataset comprising 244 images, 33 of which backgrounds, and containing 272 *S. titanus* and 275 *O. ishidae* individuals (Table 2), with images automatically resized to an input size of 640 pixels by the Ultralytics framework. Prior to performance measurement, 10 inference runs were performed as warm-up and excluded from the benchmark statistics. For each model and hardware platform, inference performance was characterized in terms of average latency per image, latency variability expressed as standard deviation and 95th percentile, throughput in FPS, and memory consumption. Inference latency was measured individually for each image, while memory consumption was monitored as the resident set size (RSS) of the inference process. All experiments were conducted using the same input images and inference settings to ensure a consistent comparison across models and hardware platforms.

3. Results

3.1. Training results and performance analysis

Training dynamics across models indicated a general convergence for all the tested models, with decreases in both localization loss (box_loss) and classification loss (cls_loss) (Figs. 2, 3). In addition, Distribution Focal Loss (DFL) exhibited a decreasing trend for YOLOv8m and YOLO11m (Fig. 2). Correspondingly, evaluation metrics (Precision, Recall, mAP50, and mAP50–95) increased steadily during training, confirming effective optimization of detection performance (Figs. 2, 3). YOLO_v8m showed more variability among runs in the initial epochs, without impacting metrics convergence in later steps (Fig. 2a). YOLO_11m models showed some divergence among runs for the localization performances, reflected in mAP50–95 values, starting around epoch 80. This difference did not impact different runs performances on Precision, Recall and mAP50 metrics (Fig. 2b). Considering YOLO26 models, YOLO26m_AdamW runs showed a divergent behavior between epochs 50 and 100 for the localization loss, yet models converged (Fig. 3a). YOLO26m_SGD resulted in the most robust training metrics among different runs (Fig. 3b).

Model evaluation on the test dataset (ground-truth annotated) confirmed strong overall detection performance across all architectures, with variability among repeated runs generally limited, demonstrating

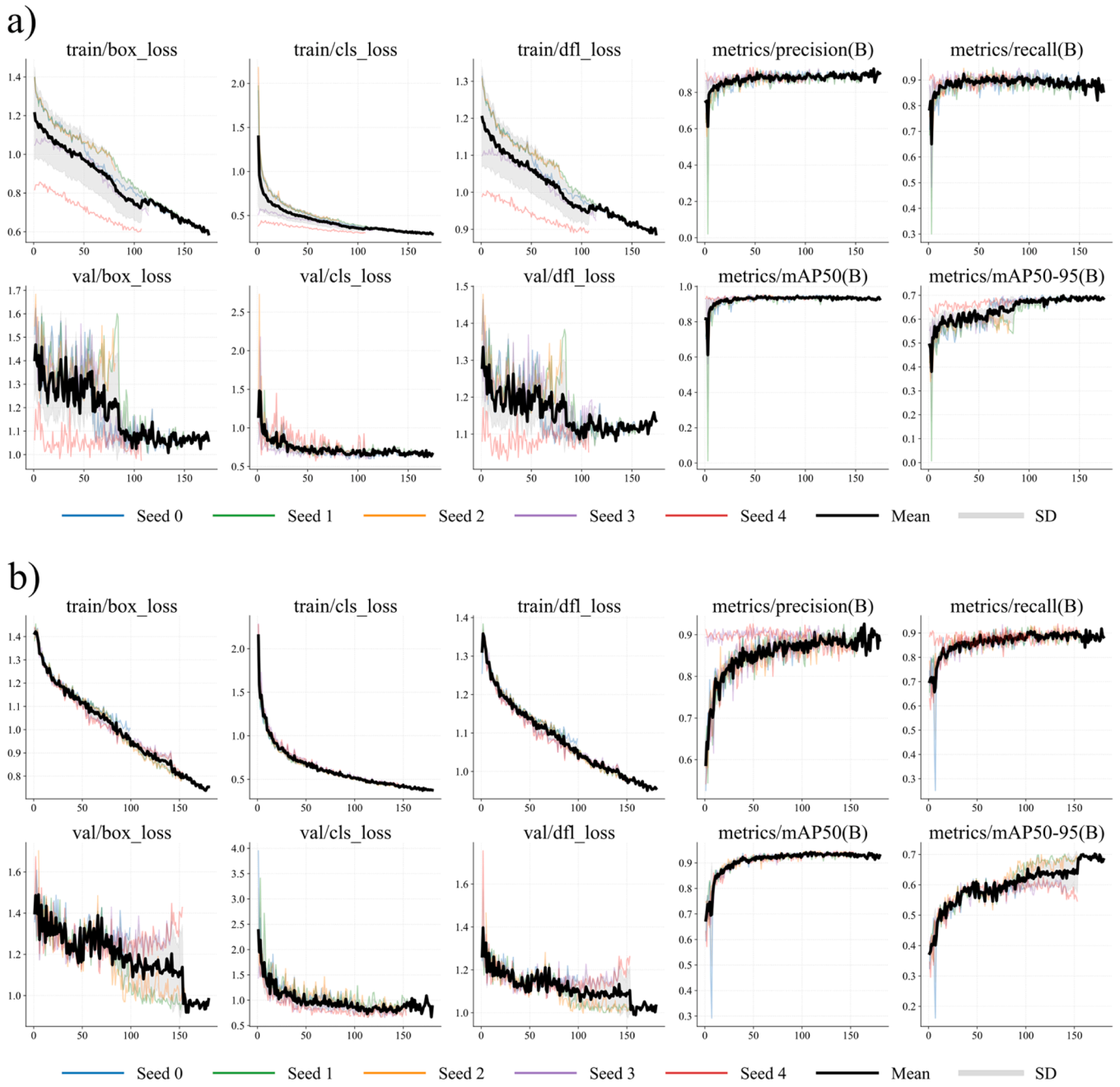


Fig. 2. Training dynamics of detectors variants for repeated runs. Top left two plots represent train box loss and classification loss for all models developed. Two plots on the bottom left represent validation box loss and classification loss for all models. Central plots represent distribution focal loss. Top right graphs represent, in order, model precision and recall. Bottom right graphs represent mean average precision (mAP) at intersection over union (IoU) of 0.50 and mAP averaged at IoU thresholds ranging from 0.50 and 0.95. Coloured lines refer to single runs with different seeds, black line to the mean value per epoch, grey area to standard deviation. a) YOLO_v8m runs; b) YOLO_11m runs.

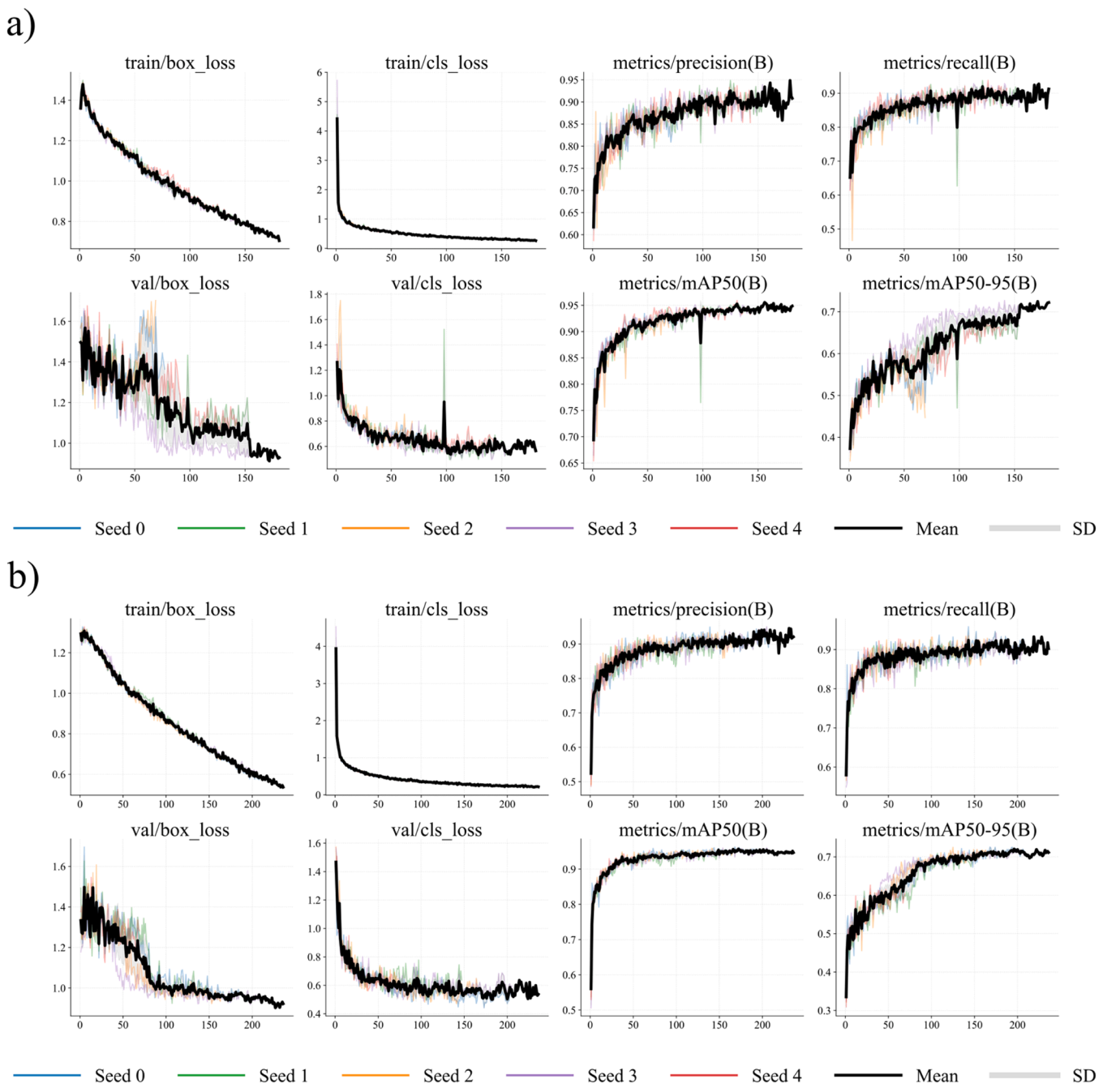


Fig. 3. Training dynamics of YOLO26m detectors. No distribution focal loss plot is reported due to DFL removal in YOLO26. Coloured lines refer to single runs with different seeds, black line to the mean value per epoch, grey area to standard deviation. a) YOLO_26m models trained with default optimizer; b) YOLO_26m models trained with SGD optimizer.

good reproducibility of the training process (Fig. 4). All models achieved average mAP50 values above 0.930, demonstrating robust detection capability under field-derived imaging conditions. YOLOv8m showed the highest reproducibility for different random seeds in terms of mAP50-95 (M = 0.652, 95% CI = 0.635 – 0.670) and the highest average precision (M = 0.920, 95% CI = 0.888 – 0.952). Considering inference speed, YOLOv8m and YOLO11m resulted in average lower FPS (respectively: 84.30 ± 3.74 and 81.00 ± 2.71) compared to both most recent training configurations (YOLO26m_SGD: 93.20 ± 5.91; YOLO26m_AdamW: 93.40 ± 3.70). The best average performances in terms of mAP50 and mAP50-95 were achieved by YOLO26m_SGD (0.953, 95% CI = 0.932 – 0.974; 0.668, 95% CI = 0.618 – 0.718 respectively). YOLO-26m_AdamW displayed the largest variability, particularly for Recall (M = 0.863, 95% CI = 0.796 – 0.930) and mAP50-95 (M = 0.630, 95% CI = 0.563 – 0.698), indicating greater sensitivity to random initialization. Detailed values represented in Fig. 4 are reported in Supplementary Table 2. Medium-size models size are reported in Supplementary Table 3. Details about performances at run-level are reported in Supplementary Table 4.

To quantify potential bias introduced by image-level random splitting for the overall best-performing architecture, YOLO26m_SGD configuration was fine-tuned with random and grouped splits and the performance evaluated on the respective test sets. Performance metrics were lower for the grouped split compared to random one. mAP50 declined from average 0.972 (95% CI = 0.944 – 1.000) to 0.923 (95% CI = 0.909 – 0.937) and mAP50-95 from average 0.738 (95% CI = 0.686

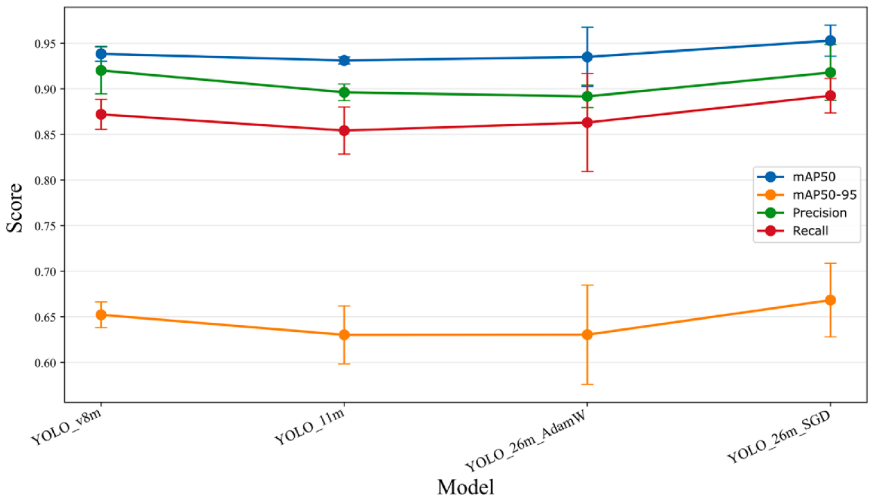


Fig. 4. Performance metrics comparison of the four fine-tuned YOLO configurations: mAP50, mAP50-95, precision, and recall. Different line colours represent different metric. Results are reported as mean values among five repeated runs, error bars indicate 95% confidence intervals.

– 0.789) to 0.657 (95% CI = 0.632 – 0.682) (Table 5). Similar reductions were observed for average precision (0.952 to 0.893) and recall (0.926 to 0.847) (Table 5).

3.2. Class-wise performance analysis

When considering per-class performance, all the models generally showed higher average metrics for *S. titanus* compared to *O. ishidae* (Table 3). Per class performance was consistent with global metrics (Fig. 4), with YOLO26m_SGD achieving the average best performance across both classes, yielding the highest AP50 and AP50–95 values overall (Table 3). YOLOv8m resulted in slightly higher average Precision on *S. titanus* (0.940, 95% CI = 0.904 – 0.976) compared to the best performing architecture (0.908, 95% CI = 0.862 – 0.955). YOLO11m and YOLO26m_AdamW resulted in high average metrics values, yet lower compared to the other detectors. YOLO26m_AdamW generally showed wider confidence intervals, confirming a higher sensibility to random initialization when trained with the custom dataset proposed in this study (Table 3).

Table 3
Fine-tuned model metrics for each insect species used as class. Values are reported as mean among five repeated runs with different random seeds. 95% confidence intervals are reported in parentheses.

Model	Class	AP50	AP50-95	Precision	Recall
YOLOv8m	*S. titanus*	0.943 (0.933 - 0.954)	0.668 (0.650 - 0.687)	0.940 (0.904 - 0.976)	0.883 (0.847 - 0.919)
	O. ishidae	0.933 (0.914 - 0.952)	0.636 (0.612 - 0.660)	0.901 (0.864 - 0.937)	0.861 (0.85 - 0.871)
YOLO11m	*S. titanus*	0.939 (0.925 - 0.953)	0.646 (0.609- 0.683)	0.892 (0.88 - 0.905)	0.889 (0.874 - 0.899)
	O. ishidae	0.925 (0.917 - 0.933)	0.598 (0.554 - 0.642)	0.902 (0.876 - 0.929)	0.811 (0.775 - 0.847)
YOLO26m_AdamW	*S. titanus*	0.947 (0.911 - 0.982)	0.655 (0.595 - 0.716)	0.888 (0.847 - 0.930)	0.901 (0.847 - 0.955)
	O. ishidae	0.923 (0.877 - 0.969)	0.605 (0.530 - 0.680)	0.895 (0.854 - 0.936)	0.825 (0.739 - 0.912)
YOLO26m_SGD	*S. titanus*	0.951 (0.932 - 0.970)	0.683 (0.637 - 0.730)	0.908 (0.862 - 0.955)	0.895 (0.878 - 0.913)
	O. ishidae	0.955 (0.932 - 0.978)	0.653 (0.597 - 0.709)	0.927 (0.894 - 0.961)	0.889 (0.846 - 0.933)

3.3. Model size performance comparison

After identifying the best fine-tuning parameters, small and nano variants of YOLO26 architecture with SGD optimizer were trained. Smaller variants generally showed higher training stability compared to medium-size one (Table 4). All the configurations resulted in high performances metrics, with AP and mAP average values slightly lower for smaller variants compared to medium-size one (Table 3; Table 4). Details about smaller-size models performances at run-level are reported in Supplementary Table 4. As expected by model sizes, smaller variants showed higher inference speed (YOLO26s_SGD: 121.00 FPS, 95% CI = 114.81 – 127.912; YOLO26n_SGD: 162.00 FPS, 95% CI = 139.45 –

Table 4
General and class-specific performance metrics for YOLO26 based models trained with SGD optimizer. Mean performances among five repeated runs with different random seeds are reported, with values in parentheses representing the 95% confidence interval.

General performance metrics

Model	mAP50	mAP50-95	Precision	Recall	FPS
YOLO26m	0.953 (0.932 – 0.974)	0.668 (0.618 – 0.718)	0.918 (0.880 – 0.956)	0.892 (0.869 – 0.916)	93.20 (85.89 – 100.57)
YOLO26s	0.948 (0.943 – 0.952)	0.664 (0.653 – 0.674)	0.925 (0.908 – 0.943)	0.871 (0.856 – 0.885)	121.00 (114.18 – 127.91)
YOLO26n	0.944 (0.937 – 0.951)	0.651 (0.638 – 0.665)	0.925 (0.911 – 0.939)	0.882 (0.863 – 0.901)	162.00 (139.45 – 184.50)

Per-class performance

Model	Class	AP50	AP50-95	Precision	Recall
YOLO26s	*S. titanus*	0.945 (0.931 – 0.958)	0.675 (0.657 – 0.693)	0.926 (0.891 – 0.960)	0.868 (0.842 – 0.893)
	O. ishidae	0.950 (0.943 – 0.958)	0.652 (0.642 – 0.661)	0.925 (0.902 – 0.948)	0.874 (0.837 – 0.911)
YOLO26n	*S. titanus*	0.950 (0.941 – 0.959)	0.663 (0.639 – 0.688)	0.933 (0.906 – 0.960)	0.891 (0.879 – 0.903)
	O. ishidae	0.939 (0.92 – 0.957)	0.639 (0.625 – 0.654)	0.918 (0.897 – 0.939)	0.874 (0.837 – 0.911)

184.50) compared to medium size one. Small and nano-size model size are reported in Supplementary Table 3.

3.4. Error analysis and qualitative evaluation

To summarize classification performances across repeated runs, confusion matrices reporting the mean and standard deviation of the cell values on the test set were computed. YOLO26m_SGD confirmed superior robustness compared to alternative configurations (Fig. 5; Figs. S1–S3).

As reported by overall metrics, all the benchmarked models confirmed high accuracy in target insects detection among repeated runs. The majority of instances were correctly classified, with very limited mutual misclassification between target classes. Average detections on target classes were similar between YOLO26m_SGD and YOLOv8m, with the former showing a lower number of false positive (FP) *O. ishidae* on average (Fig. 5; Fig. S1). YOLO11m and YOLO26-m_AdamW showed similar average values for true positive (TP) and true negative (TN) for *S. titanus* compared to the other models, but slightly less accurate detections on *O. ishidae* (Figs. S2–S3).

In addition to quantitative comparisons for the benchmarking study and considering possible future deployments of one model on edge devices, a qualitative evaluation of models errors was performed by visualizing predictions on the test dataset. For each configuration, the best performing model in terms of mAP50, mAP50-95, Precision and Recall was selected. YOLO26m_SGD (seed 0) resulted in the most accurate predictions compared to other models. The majority of instances were correctly classified (Fig. 6a-b), with very limited mutual misclassification between target classes. Notably, only two cases of reciprocal confusion between *S. titanus* and *O. ishidae* were observed, both associated with partial visibility of the insect at image boundaries (one representative example in Fig. 6c). A limited number of duplicate detections (multiple bounding boxes for a single insect) was observed in three test images, as illustrated in Fig. 6d. This behavior was primarily associated with non-ideal pose configurations, particularly cases in which one wing was extended and spatially separated from the body, likely increasing local feature ambiguity.

FP analysis indicates that most errors for *S. titanus* were associated with partially cropped individuals located at image boundaries (Fig. 6e). For *O. ishidae*, FP cases were similarly linked to truncated specimens (Fig. 6f), although additional misclassifications were observed for dark non-target objects exhibiting low-level visual similarity (Fig. 6g). False negatives (FN) were predominantly associated with degraded or partially obscured individuals, where critical morphological features (e.g., wings) were missing or occluded (Fig. 6h). Confusion matrix with absolute and normalized predictions are reported in supplementary materials (Fig. S4).

3.5. Model-specific error patterns

Across alternative models, distinct error distributions were observed. For YOLOv8m (seed 0), FP detections for *O. ishidae* were mainly associated with dark background artifacts (Fig. 7a), while this model also exhibited occasional multiple detections in high-density insect scenarios (Fig. 7b) and sporadic class attribution errors (Fig. 7c; Fig. S1).

YOLO11m (seed 2) exhibited the highest overall FP and FN rates (Fig. S2). Most of detection errors were related to *O. ishidae*, with generally more FN compared to best performing models of other YOLO variants. A notable failure mode consisted of dual-class ambiguity, where *O. ishidae* instances were simultaneously assigned as both categories. Similarly to YOLOv8m errors in *S. titanus* detections, YOLO11m best model showed multiple *O. ishidae* counting in pictures with high density of this target class. An example of those common errors is reported in Fig. 7d.

The best performing YOLO26_AdamW model (seed 3) showed similar error patterns to YOLO26m_SGD, with overall more FP *S. titanus* and FP and FN *O. ishidae*. FP *S. titanus* detections were mainly related to insect body parts cropped by image borders (Fig. 7e). Dark insects were incorrectly classified as *O. ishidae*, suggesting limitations in fine-grained discrimination under domain overlap conditions (Fig. 7f). For each medium-size model, both absolute value and normalized confusion matrix are reported in supplementary materials (Figs. S5 – S7).

Apart from detections errors related to insects at picture edges, smaller YOLO26_SGD variants generally showed more *O. ishidae* FP detections compared to the medium-size model (Figs. S8 - S9), with either other dark insects or background elements such as leaf debris incorrectly classified (Fig. 7g). Both best YOLO26s_SGD (seed 4) and YOLO26n_SGD (seed 3) showed a higher number of *S. titanus* FN, due to failed detections in image areas with strong light reflections from the transparent plastic film covering the YSTs (Fig. 7h). Confusion matrices of small-size models can be found in supplementary materials (Figs. S8 - S9).

3.6. Edge deployment performance

The benchmarking results showed substantial differences in inference performance across model sizes and embedded platforms. On Raspberry Pi 4, mean inference latency increased from 382.46 ms/frame

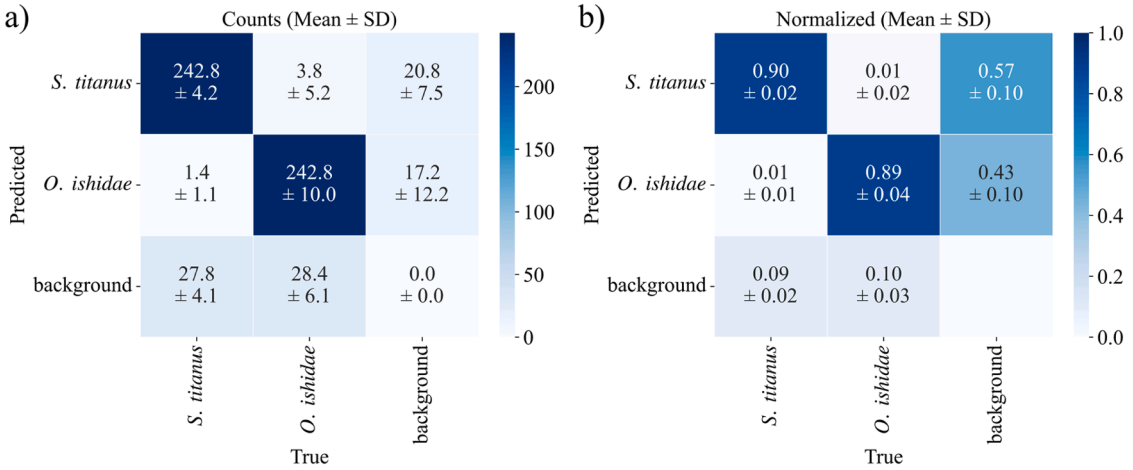


Fig. 5. Confusion matrices summarizing YOLO26m based models with SGD optimizer predictions on the test dataset. Values are reported as mean ± standard deviation for each cell across five repeated runs with different random seeds. a) Confusion matrix with counts of ground-truth assignations and predicted detections; b) confusion matrix with normalized values over the ground-truth assignations.

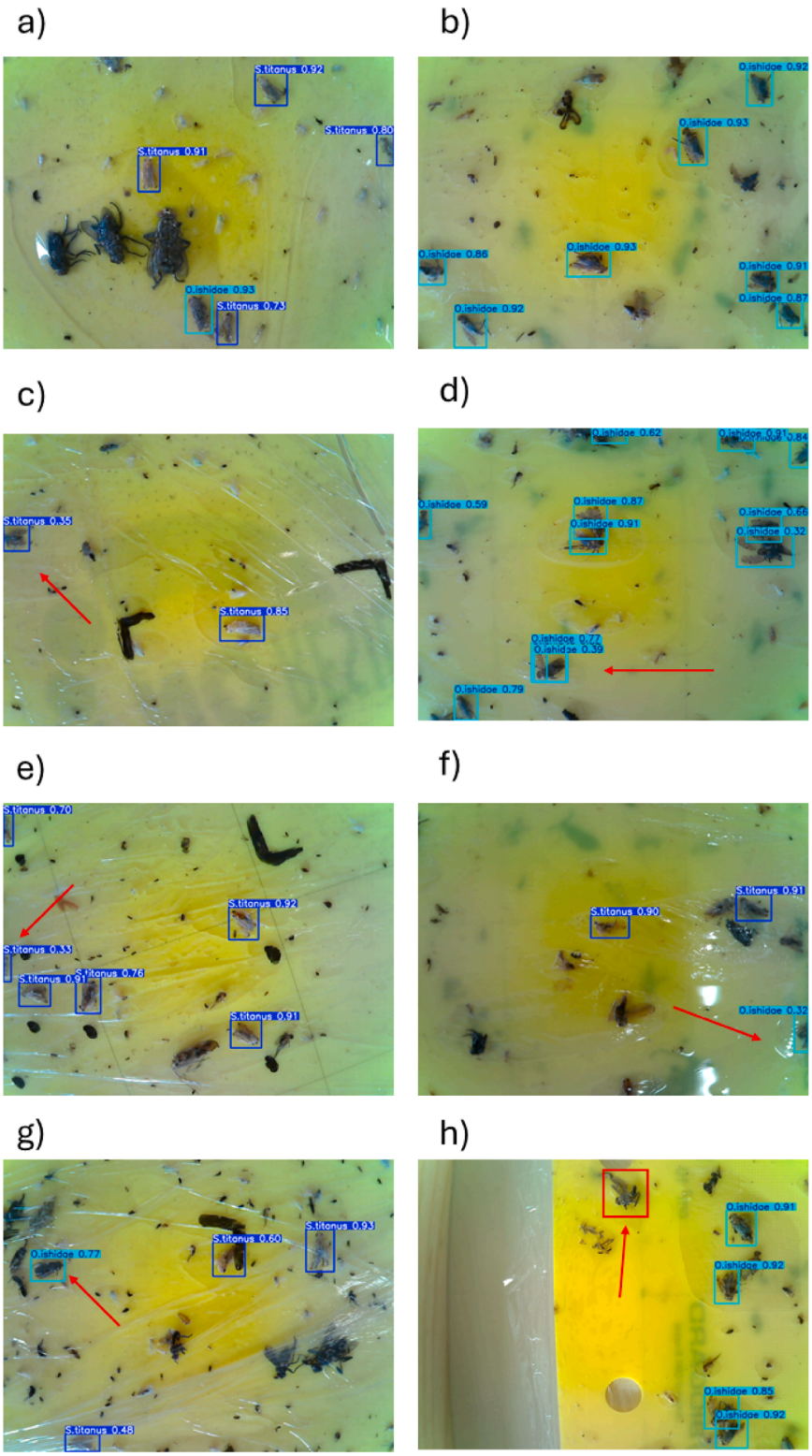


Fig. 6. Visual representation of YOLO26m_SGD detections on test dataset images. a) example of correct identifications in an image with higher density of *S. titanus*; b) correct *O. ishidae* identifications; c) misclassification of target insect partially cropped by image border leading to FN *O. ishidae* and FP *S. titanus*; d) multiple identifications of a single *O. ishidae* with one wing placed roughly perpendicular to insect body; e) identification of insect body parts cropped by image edge leading to FP *S. titanus* identification; f) identification of insect body parts cropped by image edge leading to FP *O. ishidae* identification; g) dark off-target insect incorrectly identified as *O. ishidae*; h) example of partially degraded *O. ishidae* not recognized by the model leading to FN.

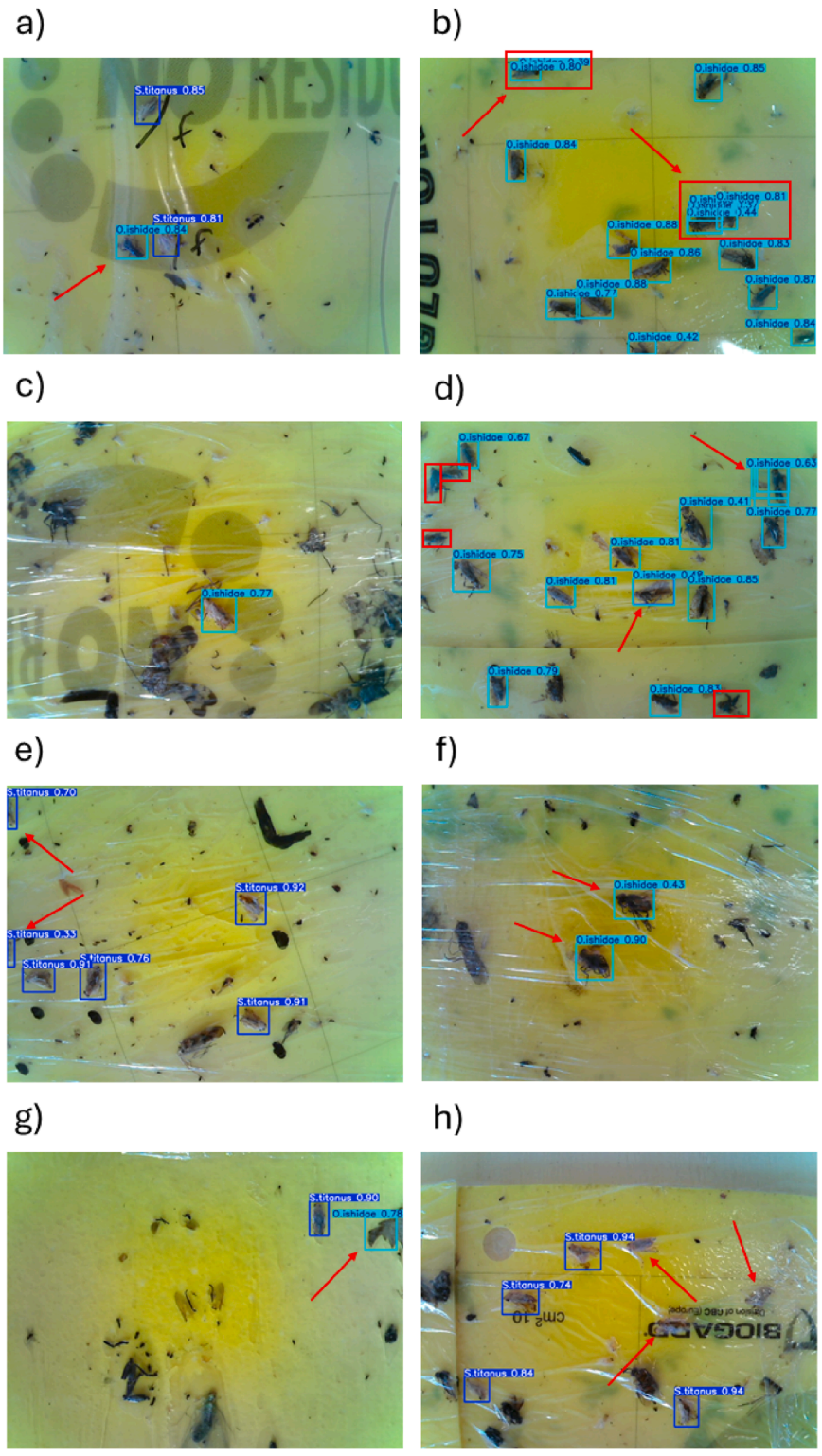


Fig. 7. Visualization of common detection errors of trained models other than the best performing YOLO26m_SGD. a) dark insect incorrectly identified as *O. ishidae* by YOLOv8m model; b) multiple *O. ishidae* identifications in image with high density of target insect by YOLOv8m model; c) *S. titanus* incorrectly identified as *O. ishidae* by YOLOv8m model; d) example of YOLO11m best model errors, with multiple detections of the same insect (upper arrow), *O. ishidae* false negative (squares) and double detection of the same insect as both classes (lower arrow); e) detection of insect body parts cropped by image border erroneously predicted as *S. titanus* by the best YOLO26m_AdamW model; f) dark non-target insects leading to *O. ishidae* FP detections by the best performing YOLO26m_AdamW model; g) example of disturbing background element leading to FP *O. ishidae* detection by the best performing YOLO26s_SGD model; h) FN YOLO26n_SGD detections of *S. titanus* in one frame with light reflections.

for YOLO26n to 1019.71 ms/frame for YOLO26s and 4060.29 ms/frame for YOLO26m, corresponding to throughputs of 2.61, 0.98, and 0.25 FPS, respectively. The Raspberry Pi 5 considerably reduced inference latency across all model configurations, reaching 87.64 ms/frame for YOLO26n, 198.05 ms/frame for YOLO26s, and 648.13 ms/frame for YOLO26m, with corresponding throughputs of 11.41, 5.05, and 1.54 FPS. Compared with Raspberry Pi 4, this corresponds to approximately 4.4x, 5.1x, and 6.3x lower mean latency for YOLO26n, YOLO26s, and YOLO26m, respectively (Table 6).

Latency variability remained relatively limited across configurations, with P95 values ranging from 101.63 ms for YOLO26n on the Raspberry Pi 5 to 4263.27 ms for YOLO26m on the Raspberry Pi 4. Memory consumption increased with model size on both platforms. Mean RAM usage ranged from 837.5 to 1143.1 MB on Raspberry Pi 4 and from 770.0 to 943.9 MB on Raspberry Pi 5. These results highlight the combined influence of model complexity and hardware platform on embedded inference performance, with YOLO26n on the Raspberry Pi 5 achieving the highest throughput among the evaluated configurations.

4. Discussion

Given the high economic and phytosanitary impact of Flavescence dorée (FD) in European viticulture and the absence of curative field treatments [7,48], management of *Scaphoideus titanus* populations through mandatory insecticide applications remains a central control strategy in several countries [49]. In Italy, compulsory monitoring of adult *S. titanus* within demarcated zones (Ministerial Order no. 4 of 22nd June 2023) is predominantly implemented through YSTs [5,50]. However, the subsequent manual inspection of traps constitutes a significant operational bottleneck due to its time-intensive nature and the requirement for expert entomological expertise [5]. In this context, automated object detection pipelines represent a practical mechanism to accelerate surveillance workflows while maintaining diagnostic reliability. Given that FD phytoplasma haplotype found in symptomatic grapevines is seldomly present in hazelnut plants [16], the competence of *O. ishidae* in transmission [12] and the morphological similarity with other Cicadellidae commonly found in YSTs [19], this insect was selected as target along with the main vector. Other vectors were not considered as target classes since their epidemiological role in viticulture is considered marginal. *D. europaea* was not included since it is easily distinguishable from other vectors [51] and it harbors phytoplasma haplotypes not found in symptomatic grapevines in Northen-Italy so far [16,17].

From a computer vision perspective, deep learning-based detection frameworks offer improved robustness under unconstrained field conditions and enhanced capability for small object localization compared to classical image processing approaches [52,53]. Prior studies in agricultural entomology consistently indicate that convolutional detection architectures outperform alternative approaches for pest identification tasks, particularly in visually complex backgrounds [27,33,36]. Accordingly, this study adopted a comparative evaluation strategy across four representative modern one-stage detection architectures, selected for their balance between computational efficiency and detection accuracy.

A dedicated dataset was constructed from field-deployed YSTs to retain variability generated during field exposure, including differences in insect density and the presence of confounding elements such as non-target arthropods, leaf fragments, and particulate debris. Image acquisition itself was performed in the laboratory, with fixed camera-to-trap geometry but intentionally non-standardized illumination, and therefore reproduces only part of the variability expected in an operational field-imaging system. In addition, given the known sensitivity of single-stage detectors to small object representation [54], image acquisition was standardized at a fixed camera-to-trap distance of 7 cm to increase the effective pixel footprint of target organisms. Since the only publicly available dataset for these target species was collected under markedly different experimental conditions [27], external validation was beyond the scope of the present benchmark.

The custom dataset proposed in this study differs from other previously developed publicly available ones for entomological applications in terms of image acquisition strategy, background composition and scopes. IP102 dataset [55] comprises 102 different classes of insect pests hampering various crops. Images for this dataset were retrieved from different online sources, therefore acquired under highly heterogeneous conditions. In contrast, the proposed dataset was generated with a partially standardized acquisition protocol and focuses only on two major viticultural pests. The MassID45 dataset [56] was developed from Malaise trap collections, comprises images captured with standardized background conditions, deals with instance segmentation and is meant for the classification of bulk arthropod samples. In comparison, the proposed dataset is meant for object detection of target insects directly from YSTs, providing realistic background conditions and not requiring pre-processing steps for the insects. The AgriPest dataset [57] focuses on 14 insect species harming wheat, rice, corn, and rape production. Differently from the here proposed dataset, image acquisition was conducted in-field, directly on target crops at varying distances from the camera. The recently published Haly.ID dataset [58] tackles field-detection of *Halyomorpha halys* (Stål) (Hemiptera: Pentatomidae) using different acquisition strategies (UAV, smartphone, stationary cameras) with the majority of the pictures captured directly in a pear orchard. By contrast, the frames of the proposed dataset were captured only from YSTs, from fixed distance and the image scale was lower compared to the average size of Haly.ID ones. Other studies have focused on sticky traps images, Ong and Høye [59] developed a dataset comprising four traps colours and with insects artificially placed on them, making it unrealistic compared to the one used for this study. Pest Sticky Traps database [60] deals with whiteflies counting from YSTs and does not involve multi-class object detection. The dataset from Checola et al. [27] is the only one available targeting the same vector species, but image acquisition strategy markedly differs from the one of this study.

Compared to full-trap single-image acquisition strategies commonly adopted in smart-trap systems [27,61], the multi-view acquisition approach improves the spatial resolution of target insects, albeit at the cost of increased image sampling requirements. Similar trade-offs have been reported in related applications involving morphologically small or densely distributed insects such as tephritid flies and thrips [33,62]. This study, however, was conducted at the image level rather than at the trap level. Further work is required to involve systematic spatial coverage of whole traps and compare expert personnel insect counts with automated counts.

No image enhancement or preprocessing operations (e.g., sharpening or contrast amplification) were applied prior to annotation or training. This decision was motivated by the need to avoid artificially induced high-frequency artifacts that may bias feature learning and degrade generalization performance [63,64]. Notably, prior work has shown that such preprocessing can negatively affect detection robustness in YOLO-based models for *S. titanus* and *O. ishidae* when applied inconsistently [27].

During dataset creation, multiple images were captured from each trap and randomly assigned to the training, validation, and test subsets at the image level. Consequently, images originating from the same YST could occur in more than one subset, and shared trap-level background characteristics may reduce the independence of the evaluation sets. The reported metrics should therefore be interpreted as an image-level benchmark under the present acquisition protocol rather than as an independent trap-level or external generalization estimate. Future studies should increase dataset size and adopt grouped splits by trap, site, and/or sampling date, together with external field validation on previously unseen traps.

An interesting outcome of this benchmarking study is that more recently proposed architectural variants did not consistently outperform earlier-generation but widely adopted detection backbones. Even

though all the models resulted in high-performance metrics values, neither YOLO11m nor YOLO26m_AdamW outperformed YOLOv8m considering average mAP50 and mAP50-95 metrics. The latter variant, largely used for entomological detection tasks [36,65,66], resulted in the highest training stability overall, as indicated by the narrow confidence interval for most metrics (Fig. 4). In contrast, YOLO26m-based models generally showed greater variability across repeated runs, although this effect was mitigated by optimizer selection. This observation is in line with previous findings [67], where the training strategy proved to mitigate the effect of backbone and neck architectural changes allowing older YOLO variants to compete with more recent ones. The higher between-run variability suggests a stronger influence of random initialization and other stochastic training factors on the final model performance. This observation suggests that improvements introduced in large-scale benchmark optimization contexts (e.g., MS COCO) do not necessarily transfer directly to small-scale, domain-specific datasets with limited sample diversity. Indeed, while recent architectural updates introduce enhanced feature extraction blocks and modified optimization strategies aimed at improving convergence efficiency and inference speed on large datasets [37], their effectiveness appears sensitive to dataset scale and distributional complexity. Consistent with the reduced head complexity [68], the most recent YOLO variants achieved higher inference speed than models from earlier generations.

A further relevant observation concerns optimization strategy sensitivity. Even considering overlaps of confidence intervals among architectures, YOLO26m_SGD consistently achieved among the highest mean values for the considered performance metrics. In the present case, characterized by a relatively small and domain-specific dataset, the use of a more conservative stochastic optimization strategy resulted in higher test-set performance in the present experiments, suggesting that slower but more stable convergence may be advantageous for fine-grained insect detection under this dataset and protocol. Optimizer selection could have facilitated a more effective exploitation of the architectural improvements introduced in recent YOLO variants, allowing them to better adapt to the specific characteristics of the dataset developed in this study. This finding is consistent with prior observations that large-scale optimization heuristics may not be optimal in low-data regimes [47].

Additionally, loss formulation and attention mechanisms integrated into recent architectures likely contributed to improved localization performance for partially occluded or small-scale targets [47]. This is particularly relevant for YST imagery, where insects are frequently observed at trap edges or partially overlapping among each other or with debris. Such effects are also reflected in qualitative results, where false positive detections often correspond to partially visible insects incorrectly classified as full instances (Fig. 6e–f, Fig. 7e). From an operational perspective, this limitation could be mitigated through acquisition strategies incorporating overlapping spatial coverage of traps to reduce boundary truncation effects that remain untested in this study [33].

Since complete dataset was randomly split at image-level for the benchmarking study, images originating from the same trap could be used for both detector fine-tuning and test phases. To assess potential bias associated with image-level random splitting, grouped and random subsets of complete dataset were used for fine-tuning and evaluating performance metrics of YOLO26m_SGD models. Results confirm that random partitioning produces more optimistic estimates. Nevertheless, the detector maintained high performance when evaluated on traps not seen during training, supporting its potential applicability under more realistic deployment conditions that will be assessed in future studies (Table 5).

Error analysis indicates high detection accuracy for the benchmarked models, with generally higher *O. ishidae* FP compared to *S. titanus* (Fig. 5; Figs. S1 – S3). These errors are primarily attributable to visually confounding background structures and low-contrast objects. In particular, the characteristic wing mosaic pattern of *O. ishidae* may be confounded with particulate noise in darker regions of the trap surface, leading to feature ambiguity. Furthermore, the inherently darker coloration of this species introduces additional bias, increasing the likelihood of misclassification with non-target dark objects. Among the best-performing models of each architectural variant, YOLO26m_SGD achieved the highest detection accuracy for *O. ishidae* (Fig. S4), while YOLOv8m and YOLO26m_AdamW showed multiple detections in images with high target density (Fig. 7b-d). This behavior is particularly critical in monitoring contexts, where accurate abundance estimation is required rather than purely detection-based evaluation. From an operational deployment perspective, such behavior limits suitability for automated counting pipelines, which are essential for population monitoring and threshold-based pest management decisions. Given that high densities of *O. ishidae* are typically expected at vineyard edges [12] and that field conditions do not allow strict control over background variability, the most robust configuration is the one that maintains stable detection behavior under density variation and background noise.

Table 5
General performance metrics for YOLO26m models fine-tuned with random and grouped dataset splits for optimistic bias quantification. Mean performances among five repeated runs with different random seeds are reported, with values in parentheses representing the 95% confidence interval.

General performance metrics				
Dataset split	mAP50	mAP50-95	Precision	Recall
Grouped	0.923 (0.909 – 0.937)	0.657 (0.632 – 0.682)	0.893 (0.864 – 0.923)	0.847 (0.819 – 0.784)
Random	0.972 (0.944 – 1.000)	0.737 (0.686 – 0.789)	0.952 (0.930 – 0.973)	0.926 (0.879 – 0.973)

For *S. titanus*, misclassification patterns vary depending on model configuration. On average, benchmarked architectures showed similar *S. titanus* FN, while YOLO26m_SGD and YOLOv8m resulted in less FP. By comparing best performing models for each architecture, two different subsets of errors were identified. Older variants (YOLOv8m, YOLO11m) misclassifications were associated with confusion among morphologically similar Deltocephalinae taxa, indicating limitations in fine-grained feature discrimination. For YOLO26m based models, FP errors were mostly related to insect body parts cropped by picture edges, potentially caused by architectural improvements to enhance recognition of occluded objects [47].

Standardization of image acquisition has previously been identified as a key factor influencing detection performance in YST-based monitoring systems [27]. While object detection frameworks are inherently scale-robust to some degree [38], consistent imaging geometry remains important for reducing intra-class variability. In this study, only camera-to-trap distance was standardized, while illumination conditions were intentionally left unconstrained to increase photometric variability represented in the dataset. This choice was also intended to increase the potential transferability of the models to future smart-trap implementations, although no deployment on prototype or operational smart-trap systems was performed within the scope of the present study. To further investigate future deployment possibilities, smaller variants of the best-performing configuration were included. Small and nano variants of YOLO26 family are particularly suited for resource-constrained edge devices [47,68], and all the fine-tuned models resulted in high detection accuracy (Table 4). Embedded benchmarking confirmed that smaller models variants exhibited substantially lower inference latency (Table 6), making them more suited for applications where fast image-processing is required. Nevertheless, by visualizing small and nano models predictions on test dataset, FN detections of *S. titanus* were occasionally observed in images with strong light reflections. This may represent a limitation when YSTs are analysed after field recollection, as the transparent plastic film commonly used to cover and store traps can generate reflections that could interfere with

Table 6
Edge-deployment performance of YOLO26 based models for different size variants (medium, small and nano). All the models were fine-tuned using the SGD optimizer. Mean inference latency (± SD), 95th-percentile latency (P95), frames per second (FPS), and mean RAM utilization are reported.

Model	Platform	Mean latency (ms/frame)	P95 (ms)	FPS	Mean RAM (MB)
YOLO26m	Raspberry Pi 4	4060.29 ± 246.26	4263.27	0.25	1143.1
	Raspberry Pi 5	648.13 ± 51.47	708.87	1.54	943.9
YOLO26s	Raspberry Pi 4	1019.71 ± 45.04	1084.68	0.98	981.8
	Raspberry Pi 5	198.05 ± 16.02	224.81	5.05	854.2
YOLO26n	Raspberry Pi 4	382.46 ± 16.02	400.83	2.61	837.5
	Raspberry Pi 5	87.64 ± 10.56	101.63	11.41	770.0

visual features of the insects (Fig. 7h). Conversely, YOLO26m_SGD model achieved the highest detection metrics in the present benchmark and may therefore provide greater tolerance to challenging imaging conditions, albeit with substantially higher inference latency. Alternative illumination strategies such as transmitted light, as explored in previous studies [33], were not adopted due to their limited applicability to single-sided trap configurations and reduced generalizability in field conditions.

Benchmarking studies have evaluated on-edge performance differences of YOLO-based models on Raspberry Pi and other CPU-only devices. As previously reported [69], different model variants in terms of size and architectural generation display substantial differences in latency and memory usage on-edge devices. Comparing performance results with other studies [70,71], average latency of YOLO26s resulted even lower than nano variants of YOLOv8, v9, v10 and v11, while RAM usage did not differ substantially among most recent and previous architectures when same model sizes are considered. Even though direct numerical comparison should be interpreted cautiously given different software environments, datasets and different deployment backends, those observations are in line with YOLO26 improvements [47], further supporting its applicability to resource-constrained edge platforms.

Finally, the image-level performance metrics obtained in this study are numerically higher than those previously reported for the same target species [27]. However, direct comparison must be interpreted cautiously, as prior work relied only on light-weight architecture and employed different imaging conditions that reduce the relative object size within the frame. These factors collectively influence detection difficulty and must be considered when benchmarking performance across studies.

5. Conclusion

This study showed that optimization strategy and training configuration influenced object-detection performance within the tested vineyard-pest monitoring setup. Benchmarking multiple YOLO architectures revealed differences in training stability and computational behavior across model generations. Overall performance was high, with the best-performing run achieving mAP50 = 0.965 and mAP50-95 = 0.69.

Evaluation on images acquired from field-deployed yellow sticky traps supports further investigation of the proposed approach under operational monitoring scenarios characterized by heterogeneous illumination, variable insect density, and complex visual backgrounds. The present results indicate the potential of deep learning-based detection to reduce the burden of manual inspection, but this potential should be confirmed through trap-level validation, systematic comparison with expert counts, and end-to-end field deployment.

From a systems perspective, the proposed framework should be regarded as an initial step toward fully automated monitoring pipelines for Flavescence dorée vector surveillance. Future work should focus on integrating detection outputs with standardized ground-truth manual inspection protocols to quantitatively assess agreement in operational conditions, as well as on automating the image acquisition process to enable end-to-end smart trap deployment.

Finally, to support reproducibility and facilitate further research on automated entomological monitoring, the annotated dataset used in this study is released to the scientific community. This resource is expected to contribute to the development of more generalized and data-efficient models for insect detection in agricultural environments.

Data availability

The dataset used in this study along with models predictions on the Test dataset will be made publicly available in Zenodo after publishing (DOI: 10.5281/zenodo.20553173).

CRediT authorship contribution statement

Matteo Dho: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Alessia De Luigi:** Writing – review & editing, Visualization, Validation, Methodology, Investigation, Data curation, Conceptualization. **Elena Gonella:** Writing – review & editing, Resources, Project administration, Funding acquisition, Conceptualization. **Gianluca Manduca:** Writing – review & editing, Visualization, Software, Methodology, Formal analysis, Conceptualization. **Donato Romano:** Writing – review & editing, Supervision, Project administration, Methodology, Conceptualization. **Alberto Alma:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare no competing interests

Acknowledgment

The authors acknowledge Doctors Badr-Eddine Jabri and Marco Manta for the support in YSTs placing and reading for the development of the datasets.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.atech.2026.102698.

References

[1] A. Bene, M. Vergine, A. Pedrelli, L. De Bellis, A. Luvisi, Flavescence dorée and grapevine susceptibility: from host–pathogen interaction to cultivar categorization, Pathogens 14 (2025) 939, https://doi.org/10.3390/pathogens14090939.
[2] European Food Safety Authority (EFSA), S. Tramontini, A. Delbianco, S. Vos, Pest survey card on flavescence dorée phytoplasma and its vector *Scaphoideus titanus*, EFSA Support. Publ. 17 (2020) 1909E, https://doi.org/10.2903/sp.efsa.2020.EN-1909.
[3] P. Taylor, R. Reeder, Antibiotic use on crops in low and middle-income countries based on recommendations made by agricultural advisors, CABI Agric. Biosci. 1 (2020) 1, https://doi.org/10.1186/s43170-020-00001-y.
[4] M. Verhaegen, J. Mahillon, S. Caulier, M. Mingeot-Leclercq, C. Bragard, Data collection on antibiotics for control of plant pathogenic bacteria, EFSA Support. Publ. 21 (2024) 8522E, https://doi.org/10.2903/sp.efsa.2024.EN-8522.
[5] E. Gonella, G. Benelli, N. Arricau-Bouvery, D. Bosco, C. Duso, C.H. Dietrich, L. Galetto, A. Rizzoli, J. Jovic, V. Mazzoni, *Scaphoideus titanus* forecasting and management: quo vadis? Entomol. Gen. 44 (2024) 497–510, https://doi.org/10.1127/entomologia/2023/2598.
[6] European Food Safety Authority (EFSA), A. Nougadère, D. Makowski, D. Rzepecka, M. Scala, B. Sánchez, F. Baldassarre, S. Tramontini, S. Vos, Council for Agricultural

Research and Economics (CREA), Grapevine flavescence dorée phytoplasma Pest Report to support the ranking of EU candidate priority pests, Wiley Online Library, 2025, https://doi.org/10.2903/sp.efsa.2025.EN-9567.
[7] J. Chuche, D. Thiéry, Biology and ecology of the Flavescence dorée vector *Scaphoideus titanus*: a review, Agron. Sustain. Dev. 34 (2014) 381–403, https://doi.org/10.1007/s13593-014-0208-7.
[8] J. Bonfils, D. Schvester, The leafhoppers (Homoptera: Auchenorrhyncha) and their relationship with vineyards in south-western France, Ann. Epiphyt. 11 (1960) 325–336.
[9] F. Bocca, L. Picciau, A. Alma, New insights on *Scaphoideus titanus* biology and their implications for integrated pest management, Entomol. Gen. 40 (2020), https://doi.org/10.1127/entomologia/2020/0977.
[10] E. Gonella, G. Benelli, N. Arricau-Bouvery, D. Bosco, C. Duso, C.H. Dietrich, L. Galetto, A. Rizzoli, J. Jović, V. Mazzoni, N. Mori, R. Nieri, P.F. Roversi, G. Strauss, D. Thiéry, V. Trivellone, M. Virant-Doberlet, A. Lucchi, A. Alma, *Scaphoideus titanus* up-to-the-minute: biology, ecology, and role as a vector, Entomol. Gen. 44 (2024) 481–496, https://doi.org/10.1127/entomologia/2023/2597.
[11] S. Malembic-Maher, D. Desqué, D. Khalil, P. Salar, B. Bergey, J.-L. Danet, S. Duret, M.-P. Dubrana-Ourabah, L. Beven, I. Ember, Z. Acs, M. Della Bartola, A. Materazzi, L. Filippin, S. Krnjajic, O. Krstić, I. Toševski, F. Lang, B. Jarausch, M. Kölber, J. Jović, E. Angelini, N. Arricau-Bouvery, M. Maixner, X. Foissac, When a Palearctic bacterium meets a Nearctic insect vector: genetic and ecological insights into the emergence of the grapevine Flavescence dorée epidemics in Europe, PLOS Pathog. 16 (2020) e1007967, https://doi.org/10.1371/journal.ppat.1007967.
[12] F. Lessio, L. Picciau, E. Gonella, F. Tota, M. Mandrioli, A. Alma, The mosaic leafhopper *Orientus ishidae*: host plants, spatial distribution, infectivity, and transmission of 16SrV phytoplasmas to vines, Bull. Insectol. 69 (2016) 277–289.
[13] A. Oggier, M. Conedera, A. Rizzoli, Habitat management as an integrative strategy for Flavescence dorée: a case study of wild-growing common hazels hosting the alternative vector *Orientus ishidae* (Hemiptera: Cicadellidae), J. Econ. Entomol. 118 (2025) 3094–3102, https://doi.org/10.1093/jee/toaf249.
[14] L. Filippin, J. Jović, T. Cvrković, V. Forte, D. Clair, I. Toševski, E. Boudon-Padieu, M. Borgo, E. Angelini, Molecular characteristics of phytoplasmas associated with Flavescence dorée in clematis and grapevine and preliminary results on the role of *Dictyophara europaea* as a vector, Plant Pathol. 58 (2009) 826–837, https://doi.org/10.1111/j.1365-3059.2009.02092.x.
[15] B. Jarausch, C. Abidon, S. Biancu, P. Salar, D. Desqué, M. Roques, M. Runne, W. Jarausch, A. Markheiser, M. Maixner, E. Meistermann, X. Foissac, S. Malembic-Maher, The role of reservoir plants and alternative vectors for the spread of Flavescence dorée (FD) phytoplasmas in FD-free regions: *Lamprotettix nitidulus* identified as a new vector, J. Pest Sci. 99 (2026) 35, https://doi.org/10.1007/s10340-025-02009-7.
[16] I.E. Rigamonti, M. Salvetti, P. Girgenti, P.A. Bianco, F. Quaglino, Investigation on flavescence dorée in north-western Italy identifies Map-M54 (16SrV-D/Map-FD2) as the only phytoplasma genotype in *Vitis vinifera* L. and reveals the presence of new putative reservoir plants, Biology 12 (2023) 1216, https://doi.org/10.3390/biology12091216.
[17] A. Pedrelli, M. Carli, A. Panattoni, E. Pellegrini, D. Rizzo, C. Nali, L. Cotrozzi, Investigating a new alarming outbreak of flavescence dorée in Tuscany (Central Italy): molecular characterization and map gene typing elucidate the complex phytoplasma ecology in the vineyard agroecosystem, Front. Plant Sci. 15 (2024), https://doi.org/10.3389/fpls.2024.1489790.
[18] M. Ripamonti, M. Pegoraro, M. Rossi, N. Bodino, D. Beal, L. Panero, C. Marzachì, D. Bosco, Prevalence of flavescence dorée phytoplasma-infected *Scaphoideus titanus* in different vineyard agroecosystems of northwestern Italy, Insects 11 (2020) 301, https://doi.org/10.3390/insects11050301.
[19] D. Bosco, A. Alma, A. Arzone, Studies on population dynamics and spatial distribution of leafhoppers in vineyards (Homoptera: Cicadellidae), Ann. Appl. Biol. 130 (1997) 1–11, https://doi.org/10.1111/j.1744-7348.1997.tb05778.x.
[20] E. Ruzzier, F. Lessio, F. Cinquatti, F. Poggi, A. Alma, A. Galli, L. Bani, F. Sanna, First record of the non-native *Osbornellus auronitens* (Provancher, 1889) (Hemiptera, Cicadellidae, Deltocephalinae) in Italy, Biodivers. Data J. 11 (2023) e106166, https://doi.org/10.3897/BDJ.11.e106166.
[21] W. Ding, G. Taylor, Automatic moth detection from trap images for pest management, Comput. Electron. Agric. 123 (2016) 17–28, https://doi.org/10.1016/j.compag.2016.02.003.
[22] J. An, Y. Du, P. Hong, L. Zhang, X. Weng, Insect recognition based on complementary features from multiple views, Sci. Rep. 13 (2023) 2966, https://doi.org/10.1038/s41598-023-29600-1.
[23] A. Raza, F. Hanif, H.A. Mohammed, Analyzing the enhancement of CNN-YOLO and transformer based architectures for real-time animal detection in complex ecological environments, Sci. Rep. 15 (2025) 39142, https://doi.org/10.1038/s41598-025-26645-2.
[24] B. Ramalingam, R.E. Mohan, S. Pookkuttath, B.F. Gómez, C.S.C. Sairam Borusu, T. Wee Teng, Y.K. Tamilselvam, Remote insects trap monitoring system using deep learning framework and IoT, Sensors 20 (2020) 5280, https://doi.org/10.3390/s20185280.
[25] I. Ahmad, Y. Yang, Y. Yue, C. Ye, M. Hassan, X. Cheng, Y. Wu, Y. Zhang, Deep learning based detector YOLOv5 for identifying insect pests, Appl. Sci. 12 (2022) 10167, https://doi.org/10.3390/app121910167.
[26] W. Li, D. Wang, M. Li, Y. Gao, J. Wu, X. Yang, Field detection of tiny pests from sticky trap images using deep learning in agricultural greenhouse, Comput. Electron. Agric. 183 (2021) 106048, https://doi.org/10.1016/j.compag.2021.106048.
[27] G. Checola, P. Sonego, R. Zorer, V. Mazzoni, F. Ghidoni, A. Gelmetti, P. Franceschi, A novel dataset and deep learning object detection benchmark for grapevine pest surveillance, Front. Plant Sci. 15 (2024), https://doi.org/10.3389/fpls.2024.1485216.
[28] Y. Han, Comparative analysis of two-stage and one-stage object detection models, Proc. 2nd Int. Conf. Data Anal. Mach. Learn 1 (2025) 289–294, https://doi.org/10.5220/0013515900004619.
[29] S.Y. Mohammed, Architecture review: two-stage and one-stage object detection, Frankl. Open 12 (2025) 100322, https://doi.org/10.1016/j.fraope.2025.100322.
[30] A. Jain, F. Cunha, M.J. Bunsen, J.S. Cañas, L. Pasi, N. Pinoy, F. Helsing, J. Russo, M. Botham, M. Sabourin, Insect identification in the wild: the AMI dataset, Springer, 2024, pp. 55–73.
[31] A.T. Nieuwenhuizen, J. Hemming, D. Janssen, H.K. Suh, L. Bosmans, V. Sluydts, N. Brenard, E. Rodríguez, M.D.M. Tellez, Raw data from yellow sticky traps with insects for training of deep learning convolutional neural network for object detection, (2019).
[32] M. Deserno, A. Briassouli, Faster R-CNN and EfficientNet for accurate insect identification in a relabeled yellow sticky traps dataset, in: 2021 IEEE Int. Workshop Metrol. Agric. For. MetroAgriFor, 2021, pp. 209–214, https://doi.org/10.1109/MetroAgriFor52389.2021.9628708.
[33] G. Niyigena, S. Lee, S. Kwon, D. Song, B.-K. Cho, Real-time detection and classification of *Scirtothrips dorsalis* on fruit crops with smartphone-based deep learning system: preliminary results, Insects 14 (2023) 523, https://doi.org/10.3390/insects14060523.
[34] B. Laekeman, J. Bonte, W. Dermauw, A. Christiaens, B. Gobin, J. Van Huylenbroeck, E. Dhooghe, P. Lootens, Species-level detection of thrips and whiteflies on yellow sticky traps using YOLO-based deep learning detection models, Front. Plant Sci. 16 (2025), https://doi.org/10.3389/fpls.2025.1668795.
[35] G. Santaera, V. Zeni, G. Manduca, A. Canale, M. Mele, G. Benelli, C. Stefanini, D. Romano, Development of an autonomous smart trap for precision monitoring of hematophagous flies on cattle, Smart Agric. Technol. 10 (2025) 100842, https://doi.org/10.1016/j.atech.2025.100842.
[36] N. Wang, S. Fu, Q. Rao, G. Zhang, M. Ding, Insect-YOLO: A new method of crop insect detection, Comput. Electron. Agric. 232 (2025) 110085, https://doi.org/10.1016/j.compag.2025.110085.
[37] R. Sapkota, M. Karkee, Ultralytics YOLO Evolution: an overview of YOLO26, YOLO11, YOLOv8 and YOLOv5 object detectors for computer vision and pattern recognition, (2026). https://doi.org/10.48550/arXiv.2510.09653.
[38] P. Hidayatullah, R. Tubagus, YOLO26: a comprehensive architecture overview and key improvements, ArXiv Prepr. ArXiv260214582 (2026).
[39] J. Song, K. Cheng, F. Chen, X. Hua, RDW-YOLO: a deep learning framework for scalable agricultural pest monitoring and control, Insects 16 (2025) 545, https://doi.org/10.3390/insects16050545.
[40] Z. Li, Z. Lin, X. Liu, Y. Li, A YOLO11-MSC-based method for small pest detection in insect pest monitoring lamp images, Sensors 26 (2026) 4576, https://doi.org/10.3390/s26144576.
[41] F. Pavan, N. Mori, G. Bigot, P. Zandigiacomo, Border effect in spatial distribution of Flavescence dorée affected grapevines and outside source of *Scaphoideus titanus* vectors, Bull. Insectol. 65 (2012) 281–290.
[42] G. Manduca, L.T. Wilson, C. Stefanini, D. Romano, Automated detection of larval stages of the black soldier fly (*Hermetia illucens* Linnaeus) through deep learning augmented with optical flow, Inf. Process. Agric. 12 (2025) 501–510, https://doi.org/10.1016/j.inpa.2025.05.001.
[43] P.V. Jamali, E. Nambi, L. M, S. Saravanan, C. V, Rice-YOLO: an automated insect monitoring in rice storage warehouses with the deep learning model, ACS Agric. Sci. Technol. 5 (2025) 206–221, https://doi.org/10.1021/acsagscitech.4c00633.
[44] X. Zhuang, B. Wang, Z. Xie, Z. Qiao, J. Feng, J.E. Behm, C. Schneider, B. Shi, J. Zheng, Y. Gu, Deep learning method for collembola identification using single species and community combinations images, Soil Ecol. Lett. 7 (2025) 250352, https://doi.org/10.1007/s42832-025-0352-9.
[45] J. Terven, D.-M. Córdova-Esparza, J.-A. Romero-González, A comprehensive review of YOLO architectures in computer vision: from YOLOv1 to YOLOv8 and YOLO-NAS, Mach. Learn. Knowl. Extr. 5 (2023) 1680–1716, https://doi.org/10.3390/make5040083.
[46] R. Sapkota, M. Flores-Calero, R. Qureshi, C. Badgujar, U. Nepal, A. Poulose, P. Zeno, U.B.P. Vaddevolu, S. Khan, M. Shoman, H. Yan, M. Karkee, YOLO advances to its genesis: a decadal and comprehensive review of the You Only Look Once (YOLO) series, Artif. Intell. Rev. 58 (2025) 274, https://doi.org/10.1007/s10462-025-11253-3.
[47] R. Sapkota, R.H. Cheppally, A. Sharda, M. Karkee, YOLO26: key architectural enhancements and performance benchmarking for real-time object detection, (2026). https://doi.org/10.48550/arXiv.2509.25164.
[48] E. Gonella, A. Arpellino, L. Picciau, F. Lessio, M. Prati, L. Bucci, M. Dho, A. Alma, Protective and curative effect of an experimental biocomplex against Flavescence dorée phytoplasma infection and transmission by *Scaphoideus titanus* under laboratory conditions, Crop Prot. 176 (2024) 106472, https://doi.org/10.1016/j.cropro.2023.106472.
[49] S.C. Prazaru, L. D'Ambrogio, M. Dal Cero, M. Rasera, G. Cenedese, E. Guerrieri, M. Pavasini, N. Mori, F. Pavan, C. Duso, efficacy of conventional and organic insecticides against *Scaphoideus titanus*: field and semi-field trials, Insects 14 (2023) 101, https://doi.org/10.3390/insects14020101.
[50] B. Bagnoli, R. Ricciardi, L. De Fazi, G. D'Anna, P. Braccini, G. Benelli, A. Lucchi, Monitoring and management of the Nearctic leafhopper *Scaphoideus titanus* (Hemiptera: Cicadellidae) in Italian vineyards, J. Econ. Entomol. 117 (2024) 2281–2291, https://doi.org/10.1093/jee/toae250.

[51] Z.-S. Song, A.-P. Liang, The Palaearctic planthopper genus *Dictyophara* Germar, 1833 (Hemiptera: Fulgoroidea: Dictyopharidae) in China. Museum and Institute of Zoology, Polish Academy of Sciences, 2008, pp. 537–549.
[52] H. Patel, A comprehensive study on object detection techniques in unconstrained environments, ArXiv Prepr. ArXiv230405295 (2023). https://doi.org/10.48550/arXiv.2304.05295.
[53] S. Wang, D. Xu, H. Liang, Y. Bai, X. Li, J. Zhou, C. Su, W. Wei, Advances in deep learning applications for plant disease and pest detection: a review, Remote Sens. 17 (2025) 698, https://doi.org/10.3390/rs17040698.
[54] X. Peng, H. Jiang, A review of small object detection based on deep learning, in: Proc. 2025 2nd Int. Conf. Big Data Anal. Artif. Intell. Appl., Association for Computing Machinery, New York, NY, USA, 2026, pp. 89–96, https://doi.org/10.1145/3788108.3788510.
[55] X. Wu, C. Zhan, Y.-K. Lai, M.-M. Cheng, J. Yang, IP102: a large-scale benchmark dataset for insect pest recognition, in: 2019 IEEECVF Conf. Comput. Vis. Pattern Recognit. CVPR, 2019, pp. 8779–8788, https://doi.org/10.1109/CVPR.2019.00899.
[56] J. Orsholm, J. Quinto, H. Autto, G. Banelyte, N. Chazot, J. deWaard, S. deWaard, A. Farrell, B. Furneaux, B. Hardwick, N. Ito, A. Kar, O. Kalttopää, D. Kerdraon, E. Kristensen, J. McKeown, T. Mononen, E. Nein, H. Rogers, T. Roslin, P. Schmitz, J. Sones, M. Sujala, A. Thompson, E.V. Zakharov, I. Zarubiieva, A. Gupta, S. C. Lowe, G.W. Taylor, A multi-modal dataset for insect biodiversity with imagery and DNA at the trap and individual level, Sci. Data 13 (2026) 630, https://doi.org/10.1038/s41597-026-07251-x.
[57] R. Wang, L. Liu, C. Xie, P. Yang, R. Li, M. Zhou, AgriPest: a large-scale domain-specific benchmark dataset for practical agricultural pest detection in the wild, Sensors 21 (2021) 1601, https://doi.org/10.3390/s21051601.
[58] L. Palazzetti, F. Betti Sorbelli, P. Das, L. Almstedt, D. Niederprüm, L. Wolf, C. M. Pinotti, A comprehensive dataset of *Halyomorpha halys* for automated detection and monitoring in orchards, Sci. Data 13 (2026) 1059, https://doi.org/10.1038/s41597-026-07774-3.
[59] S.-Q. Ong, T.T. Høye, An annotated image dataset of pests on different coloured sticky traps acquired with different imaging devices, Data Brief. 55 (2024) 110741, https://doi.org/10.1016/j.dib.2024.110741.
[60] L. Ciampi, V. Zeni, L. Incrocci, A. Canale, G. Benelli, F. Falchi, G. Amato, S. Chessa, A deep learning-based pipeline for whitefly pest abundance estimation on chromotropic sticky traps, Ecol. Inform. 78 (2023) 102384, https://doi.org/10.1016/j.ecoinf.2023.102384.
[61] G. Ortis, C. Duso, N. Mori, New tools to monitor *Scaphoideus titanus*, the main vector of phytoplasmas associated with Flavescence dorée in vineyards, Entomol. Exp. Appl. 173 (2025) 330–335, https://doi.org/10.1111/eea.13543.
[62] M. Tannous, C. Stefanini, D. Romano, A deep-learning-based detection approach for the identification of insect species of economic importance, Insects 14 (2023) 148, https://doi.org/10.3390/insects14020148.
[63] N. Kumar, F.Flammini Nagarathna, YOLO-based light-weight deep learning models for insect detection system with field adaption, Agriculture 13 (2023) 741, https://doi.org/10.3390/agriculture13030741.
[64] R. Yang, Y. Guo, J. Yan, Q. Zhou, H. Yang, D. Gong, SHRS-YOLO: small target detection in hazy remote sensing images, IEEE Trans. Geosci. Remote Sens. 64 (2026) 1–15, https://doi.org/10.1109/TGRS.2026.3651676.
[65] M. Masuko, S. Kikuta, Image recognition-based deep learning model for identifying the developmental stages of *Acyrthosiphon pisum* (Hemiptera: Aphididae), Appl. Entomol. Zool. 59 (2024) 251–259, https://doi.org/10.1007/s13355-024-00873-w.
[66] Y.S. Şahin, N.S. Gençer, H. Şahin, Real-time detection and segmentation of tomato pests with YOLOv8, J. Agric. Sci. 32 (2026) 119–129.
[67] I. Lazarevich, M. Grimaldi, R. Kumar, S. Mitra, S. Khan, S. Sah, YOLOBench: benchmarking efficient object detectors on embedded systems, in: 2023 IEEECVF Int. Conf. Comput. Vis. Workshop ICCVW, 2023, pp. 1161–1170, https://doi.org/10.1109/ICCVW60793.2023.00126.
[68] G. Jocher, J. Qiu, M. Liu, S. Lyu, F.C. Akyon, M.E. Kalfaoglu, Ultralytics YOLO26: unified real-time end-to-end vision models, ArXiv Prepr. ArXiv260603748 (2026).
[69] D.K. Alqahtani, M.A. Cheema, M.A. Rodriguez, A.N. Toosi, A comprehensive evaluation of deep learning object detection models on heterogeneous edge devices, ArXiv Prepr. ArXiv240916808 (2024).
[70] M.T. Okano, W.A.C. Lopes, S.M. Ruggero, Comparative benchmarking of YOLO-based object detection on Raspberry Pi 3B, 4B, and 5B under CPU-only edge AI deployment, in: 2026 12th Int. Conf. Control Decis. Inf. Technol. CoDIT, 2026, pp. 80–85, https://doi.org/10.1109/CoDIT70676.2026.11631305.
[71] L. Palazzetti, D. Giannetti, A. Verolino, D.A. Grasso, C.M. Pinotti, F. Betti Sorbelli, AntPi: a Raspberry Pi based edge–cloud system for real-time ant species detection using YOLO, Ecol. Inform. 91 (2025) 103383, https://doi.org/10.1016/j.ecoinf.2025.103383.