Computing&AI Connect
ISSN: 3104-4719
2025, Article ID. 2025.0026, Cite as: https://doi.org/10.69709/CAIC.2025.150921

SCIFINITI PUBLISHING

Research Article

OPEN ACCESS

Enhancing Pest Management in Precision Agriculture: Integration of Improved YOLOv5 and IoT Technology for Efficient Codling Moth Detection

Mohamed Zarboubi[1,*] **Abdelaaziz Bellout**[2] **Samira Chabaa**[1,3] **Azzedine Dliou**[2,4]

[1] Systems Engineering and Decision Support Laboratory (LISAD), National School of Applied Sciences, Ibn Zohr University, Agadir 80000, Morocco
[2] Systems Engineering and Information Technology Laboratory (LISTI), National School of Applied Sciences, Ibn Zohr University, Agadir 80000, Morocco
[3] Instrumentation, Signals and Physical Systems (I2SP) Team, Faculty of Sciences Semlalia, Cadi Ayyad University, Marrakech 40000, Morocco
[4] Innovation in Mathematics and Intelligent Systems (IMIS) Laboratory, Faculty of Applied Sciences, Ibn Zohr University, Agadir 80000, Morocco

Article History
Received: May 02, 2025 Accepted: December 09, 2025 Published: December 19, 2025

Abstract
Effective pest management remains a persistent challenge in precision agriculture, particularly due to the difficulty of accurately detecting small insects in cluttered trap environments under varying lighting and background conditions. Traditional methods often require labor-intensive inspections, while existing deep learning-based object detectors, such as standard YOLO models, face trade-offs between detection accuracy and computational feasibility on edge devices. This study introduces an improved version of the YOLOv5m model tailored for deployment on a Raspberry Pi-based smart insect trap, targeting codling moth (Cydia pomonella) detection. The proposed architecture incorporates a Convolutional Block Attention Module (CBAM) to enhance feature representation and reduce background interference, along with strategic filter reduction to lower computational complexity. As a result, the model achieves a maximum confidence level of 95% and an average of 91.36%, with reduced parameter count and a FLOPs value of 26.88 billion. Integration with the Firebase IoT platform enables real-time monitoring and remote data management. Comparative analysis with YOLOv5–YOLOv12 variants demonstrates that the improved YOLOv5m offers the best balance between accuracy and efficiency for low-power deployment. These findings highlight the potential of combining lightweight deep learning and IoT infrastructure to create scalable, energy-efficient, and sustainable pest detection systems for real-world agricultural applications.

Keywords:
precision agriculture; improved YOLOv5; pest detection; IoT (Internet of Things); sustainable farming; codling moth

1. Introduction

The global population has experienced a significant increase and is projected to maintain its upward trajectory [1–3]. This upsurge in population has fueled an escalating demand for agricultural products, leading to a notable expansion in cultivated areas, all in an effort to augment annual crop yields [4,5]. A variety of biotic and abiotic factors affect agricultural productivity worldwide. Approximately 40% of agricultural output is negatively affected by insects, pests, diseases, and weed infestations [6]. It is often difficult to achieve the desired level of pest and disease control due to a lack of timely and accurate diagnosis. The improper and excessive use of agrochemicals can result in financial and environmental problems [7]. According to the Food and Agriculture Organization (FAO),

*** Corresponding Author:**
Mohamed Zarboubi, Systems Engineering and Decision Support Laboratory (LISAD), National School of Applied Sciences, Ibn Zohr University, Agadir 80000, Morocco, mohamed.zarboubi@edu.uiz.ac.ma

pest damage accounts for approximately 20% to 40% of global production losses each year [8–10]. In addition, the annual economic impact of pest infestations worldwide is estimated at approximately $220 billion, while invasive insects alone contribute to financial losses of around $70 billion [8,11]. As a result, farmers often resort to various pesticides to improve the productivity and longevity of their crops. However, prolonged pesticide use can lead to environmental contamination and may increase the risk of severe health conditions such as cancer, respiratory diseases, genetic disorders, and adverse pregnancy outcomes [12].

One of these pests is Cydia pomonella, commonly known as the codling moth, which is a significant pest that severely affects crops, particularly apple and pear orchards. The larvae of this moth feed on the fruits, leading to yield loss and quality deterioration. Infestations of codling moths can cause significant economic damage to farmers, with losses estimated to be in the billions of dollars worldwide [13–15].

Sophisticated technological methods are essential in agriculture to detect pests at early stages and reduce reliance on harmful pesticides [16]. Farmers traditionally rely on their expertise and insight to diagnose pest infestations, resulting in excessive pesticide spraying [17]. However, growing environmental and health concerns have underscored the need to reduce pesticide use. One crucial approach to achieving this is by spraying pesticides only where necessary. Manual pest detection methods are laborious and error-prone, often require substantial human effort, and are susceptible to inaccuracies [18–20]. Fortunately, recent advances in the Internet of Things (IoT) and computer vision for precision agriculture have made insect pest and disease detection an essential component in monitoring crop growth and health [9,21–24]. This detection is significant for estimating future yields, activating intelligent spraying systems, and overseeing autonomous pesticide-spraying robots on expansive farms and orchards. However, accurately detecting target objects, such as pests, is difficult due to factors including shape similarity, complex backgrounds, object overlap, variable lighting, and vast orchard topography [25–27]. Yet, thanks to technological progress, particularly in image processing, identifying insect pests has become achievable [28,29]. Precision agriculture has become increasingly popular as a solution to overcome these challenges, aiming to improve the precision and accuracy of pest detection [30–38]. The acquisition and processing of visual information via computer vision have become indispensable for effective pest detection.

Recent advances in machine vision and learning technologies have opened new possibilities for addressing object detection challenges. Significantly, the development of deep object detector algorithms has demonstrated remarkable effectiveness in diverse fields, such as agriculture [39–45]. Specifically, counting insects can be viewed as a particular application of object detection, making CNN-based object detectors an optimal solution [46,47]. Consequently, a considerable number of researchers are currently engaged in the rigorous exploration and development of image detection methodologies utilizing convolutional neural networks (CNNs) [48–53]. This surge in activity underscores the significance and potential of CNN-based approaches to advance image detection capabilities. The CNN-based YOLO (You Only Look Once) model is the ideal solution for this task, providing real-time responsiveness and high accuracy. One of the key advantages of using the YOLO model is its ability to detect objects in real time, making it particularly suitable for applications that require fast, reliable performance and immediate processing. Additionally, the YOLO model achieves high accuracy by performing object detection and classification in a single pass, resulting in efficient dependable performance. This makes it highly beneficial for tasks such as insect counting in agriculture, where timely and accurate detection is crucial for effective pest management. Furthermore, the YOLO model is known for its robustness against variations in object size, orientation, and occlusion, making it a versatile and reliable choice for object detection tasks in complex environments. Notably, the YOLO algorithms have been developed in twelve different versions: YOLOv1 [54], YOLOv2 [55], YOLOv3 [56], YOLOv4 [57], YOLOv5 [58], YOLOv6 [59], YOLOv7 [60], YOLOv8 [61], YOLOv9 [62], YOLOv10 [63], YOLOv11 [64], and YOLOv12 [65].

This study presents a novel pest monitoring system that integrates an innovative pheromone trap design with a CBAM-enhanced YOLOv5m model, deployed on a low-cost Raspberry Pi platform. Although standard YOLOv5m and other YOLO variants deliver strong detection performance, they face key limitations when deployed on edge devices like the Raspberry Pi. These include large parameter sizes, high FLOP requirements, and reduced ability to focus on small insect targets in cluttered, variable trap environments. To address these challenges, we modified the YOLOv5m architecture by incorporating the Convolutional Block Attention Module (CBAM) to refine feature representation and suppress irrelevant background information. Additionally, we strategically reduced the number of filters in selected layers to minimize computational complexity without compromising accuracy. The resulting model achieves real-time, energy-efficient detection on lightweight hardware, making it highly suitable for embedded pest detection systems.

Coupled with the Firebase IoT platform, our system supports seamless remote monitoring, automated data management, and precision-driven pest control interventions. This enables more informed insecticide deployment, reduces unnecessary pesticide usage, and ultimately promotes sustainable, environmentally responsible farming practices.

The primary objective of this research is to optimize pest management strategies by accurately determining the most effective timing for insecticide application. By harnessing the computational efficiency of the Raspberry Pi, the detection capabilities of the enhanced YOLOv5m model, and the real-time data synchronization features of the Firebase IoT platform, the study proposes a fully automated and cost-effective framework for monitoring the population dynamics of Cydia pomonella. This integrated system empowers farmers with timely, actionable insights, enabling more precise, evidence-based pesticide interventions. In doing so, it supports reducing unnecessary chemical use, lowers environmental impact, and advances the principles of sustainable and intelligent agriculture.

To better contextualize the contributions of this study, it is essential to examine the limitations of current pest-monitoring techniques. Conventional methods, such as sticky and pheromone traps, require frequent manual inspections, making them time-consuming, labor-intensive, and impractical for large-scale orchards or remote agricultural sites [66,67]. Even though automated trapping systems equipped with image sensors and wireless communication technologies have demonstrated the potential to reduce field visits and enable real-time monitoring, they typically rely on centralized cloud-based infrastructure for data processing and analysis [68–70]. This dependence introduces challenges, including high energy consumption, increased latency, and elevated operational costs. In parallel, alternative approaches, such as acoustic traps that detect wingbeat frequencies [71,72], tend to exhibit reduced reliability due to their sensitivity to environmental noise and to overlapping frequency patterns among insect species. In contrast, the proposed system provides a lightweight, edge-computing solution that performs real-time image processing locally on a Raspberry Pi, transmitting only essential detection results via the Firebase IoT platform. This architecture significantly reduces data transmission demands and energy consumption while maintaining robust detection accuracy, making it highly adaptable for scalable, sustainable, and cost-effective deployment in precision agriculture.

The remaining sections of this paper are structured as follows: Section 2 outlines the methodology employed in this study, detailing the dataset collection and preparation process, the object detection model utilized, the integration of the Firebase IoT platform, and the innovative trap design. Section 3 discusses the experimental results and provides insights. Section 4 presents the perspectives and future research directions. Finally, Section 5 provides the overall conclusions drawn from this research.

2. Methodology and Performance Enhancements of Improved YOLOv5m Architecture

Our proposed approach unfolds through a sequence of seven distinct phases, illustrated in Figure 1:

- **Stage N°1**: Gather a collection of insect pest images to train and evaluate the deep learning model.
- **Stage N°2**: Prepare the dataset by resizing images to 640x640 and applying data augmentation techniques to expand the number of training samples.
- **Stage N°3**: Label the images to create a dataset suitable for object detection tasks.
- **Stage N°4**: Train the improved YOLOv5 model using the processed dataset.
- **Stage N°5**: Validate the model's performance with a subset of the dataset and analyze the results.
- **Stage N°6**: Select the most accurate model for integration with the Raspberry Pi-based trap.
- **Stage N°7**: Transmit the detection results to the Firebase platform for real-time monitoring.

2.1. Dataset

To develop and validate our insect pest detection system, we compiled a diverse dataset from three main sources. The first source was the team of Akroute et al. [73] at the National Institute of Agricultural Research of Morocco, whose study investigates correlations between apple variety characteristics and codling moth susceptibility to support sustainable pest management. Their research, which

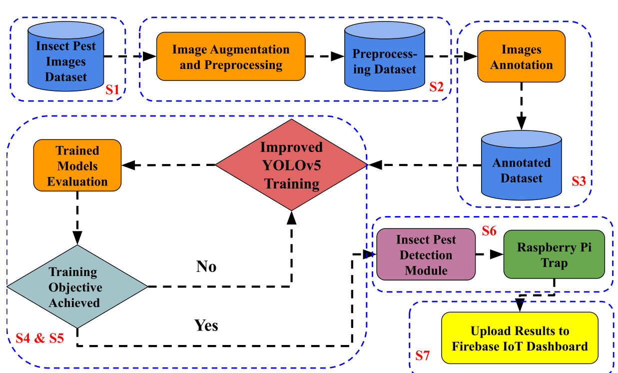


Figure 1: Flowchart illustration of the research methodology [20].

involved laboratory rearing of insects and field monitoring using pheromone traps, yielded a rich collection of codling moth images for use as training data. The second source was the database employed by Sütő [74], further enhancing the dataset's variety. Lastly, we gathered additional images from online platforms, including Google, Bing, Flickr, IPM Images, iStock, and Lepiforum. To ensure uniformity, all photos were standardized to 640x640 pixels, and representative samples from the dataset are shown in Figure 2.

Deep learning models perform better with large datasets, but collecting sufficient data can be challenging. Small datasets can lead to problems like overfitting and reduced accuracy. To address this, techniques such as flipping and shifting are often used to generate more training data, though they can sometimes cause pixel loss at image edges. In this study, insect pest images were augmented by rotating them at 90°, 180°, and 270°, creating three new images for each original, as shown in Figure 3. This method maintained image quality while varying the insect's position, increasing the dataset to 1011 Images. These additional images help the model generalize better and improve its pest identification performance.

During the image preprocessing stage for deep learning model training, image annotation plays a pivotal role. This essential process entails extracting significant features from an image and subsequently assigning appropriate labels to these features based on selected inputs. The significance of image annotation lies in its ability to provide labeled data that serves as the foundation for supervised learning, facilitating the model's understanding of relevant patterns and characteristics in images.

The dataset used in this study consists of 1011 images containing 2551 annotated insect instances. These instances represent different poses, sizes, and lighting conditions of codling moths captured in realistic trap environments, making the dataset both diverse and representative of real-world field conditions.

Figure 2: Images of the Cydia pomonella dataset we employed for training and validation the object detection model (Adopted from online platforms, including Google, Bing, Flickr, IPM Images, iStock, and Lepiforum).

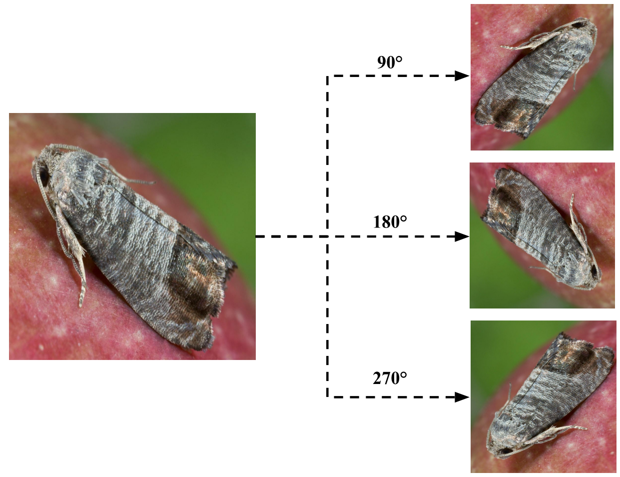


Figure 3: Data augmentation (Adopted from online platforms, including Google, Bing, Flickr, IPM Images, iStock, and Lepiforum).

In the context of this study, the process of image annotation is illustrated in Figure 4, where we utilize the Makesense online platform to accomplish this task. This platform streamlines the annotation process, enabling efficient, accurate labeling of extracted features and thereby enhancing the overall effectiveness and quality of the training dataset for the deep learning model.

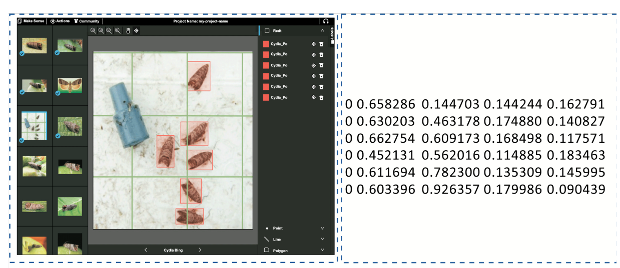


Figure 4: Image annotation in the makesense platform.

2.2. Improvement of YOLOv5m Network Architecture Design

This study presents an enhanced YOLOv5m network architecture that integrates the Convolutional Block Attention Module (CBAM) to improve object detection performance. The original YOLOv5m architecture, shown in Figure 5, consists of a backbone for feature extraction, a neck for feature aggregation, and a head for final detection. The backbone uses CSPNet-based feature-extraction blocks (C3) and a Spatial Pyramid Pooling-Fast (SPPF) module to enhance receptive fields.

2025, Vol. 2, Article ID. 2025.0026
https://doi.org/10.69709/CAIC.2025.150921

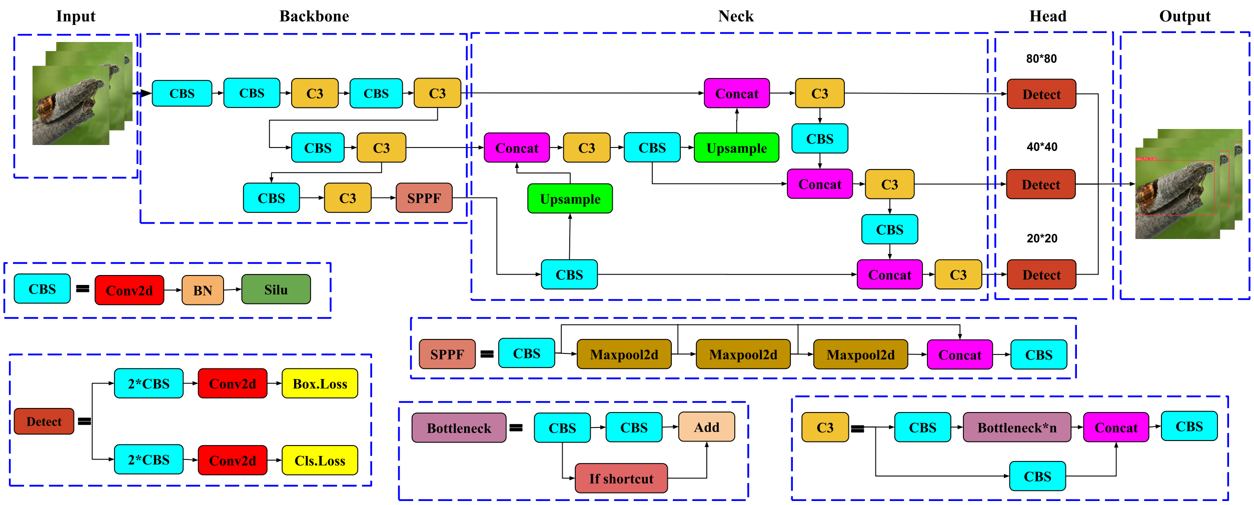


Figure 5: YOLOv5 network structure.

In the improved design, depicted in Figure 6, CBAM is incorporated into the YOLOv5m architecture. CBAM, illustrated in Figure 7, consists of a Channel Attention Module and a Spatial Attention Module. These modules sequentially refine the input features by focusing on 'what' and 'where', respectively. The Channel Attention Module adaptively recalibrates channel-wise feature responses, while the Spatial Attention Module emphasizes informative regions in the feature map.

CBAM is integrated at multiple strategic points within the backbone and neck of the network. Specifically, CBAM modules are inserted after key convolutional layers and bottleneck structures, enabling the network to focus on more relevant features during detection. This integration enhances the model's ability to detect and classify objects, particularly in challenging scenarios involving occlusions or varying object scales.

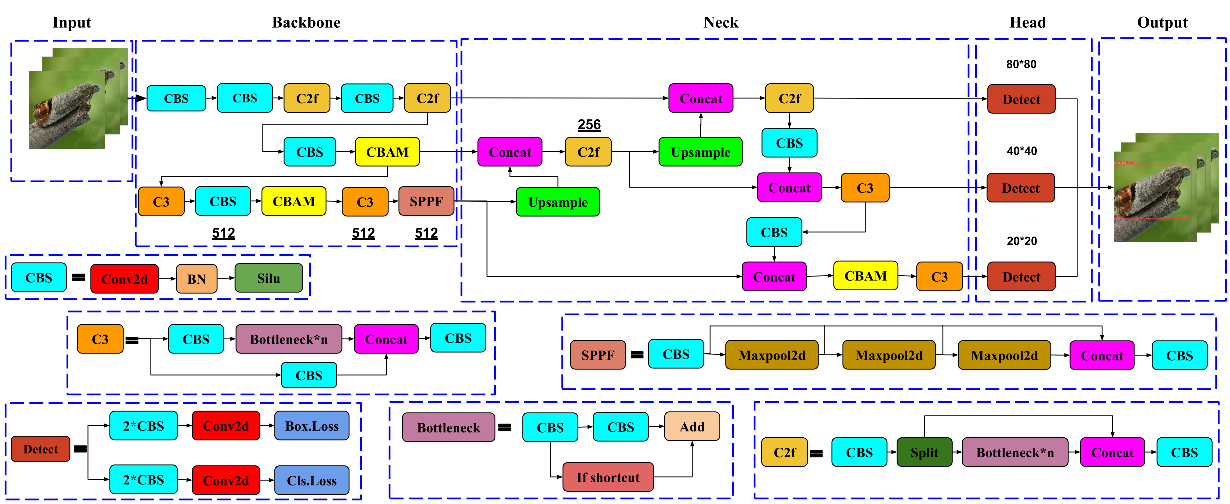


Figure 6: Improved YOLOv5 network structure.

Additionally, the number of filters in specific layers was reduced to optimize computational efficiency. For instance, the number of filters in the last backbone layer was decreased from 1024 to 512 for each of the SPPF, C3, and CBS modules. Similarly, at the beginning of the neck, the filters in the C2f modules were reduced from 512 to 256. These adjustments help balance the model's complexity and performance, ensuring efficient processing without compromising accuracy.

Moreover, we replace the C3 modules with the C2f modules in the improved design to further boost feature learning capacity. The C2f modules, combined with CBAM, provide a more robust feature extraction process, thereby improving detection accuracy.

Overall, adding CBAM to the YOLOv5m architecture, along with strategic reductions in filter counts, results in a more powerful and efficient object detection model, that delivers higher accuracy and better generalization across diverse datasets. This improved architecture has significant potential for applications that require precise and reliable object detection, such as automated pest-monitoring systems.

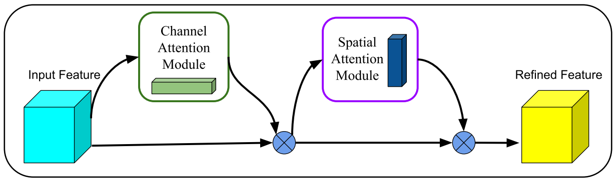


Figure 7: Convolutional block attention module.

2.3. Experimental Impact of CBAM on Feature Extraction

Following the architectural enhancements discussed in section 2.2, we conducted experiments to evaluate the effect of integrating the Convolutional Block Attention Module (CBAM) on feature extraction. This section provides a visual analysis of CBAM's effectiveness in guiding the network to focus on the most relevant regions of an image.

To illustrate the role of CBAM, we applied it at various stages of the network and compared the results with those of standard convolutional layers. The visual outputs in Figure 8 demonstrate how CBAM progressively enhances feature focus as the model processes the image.

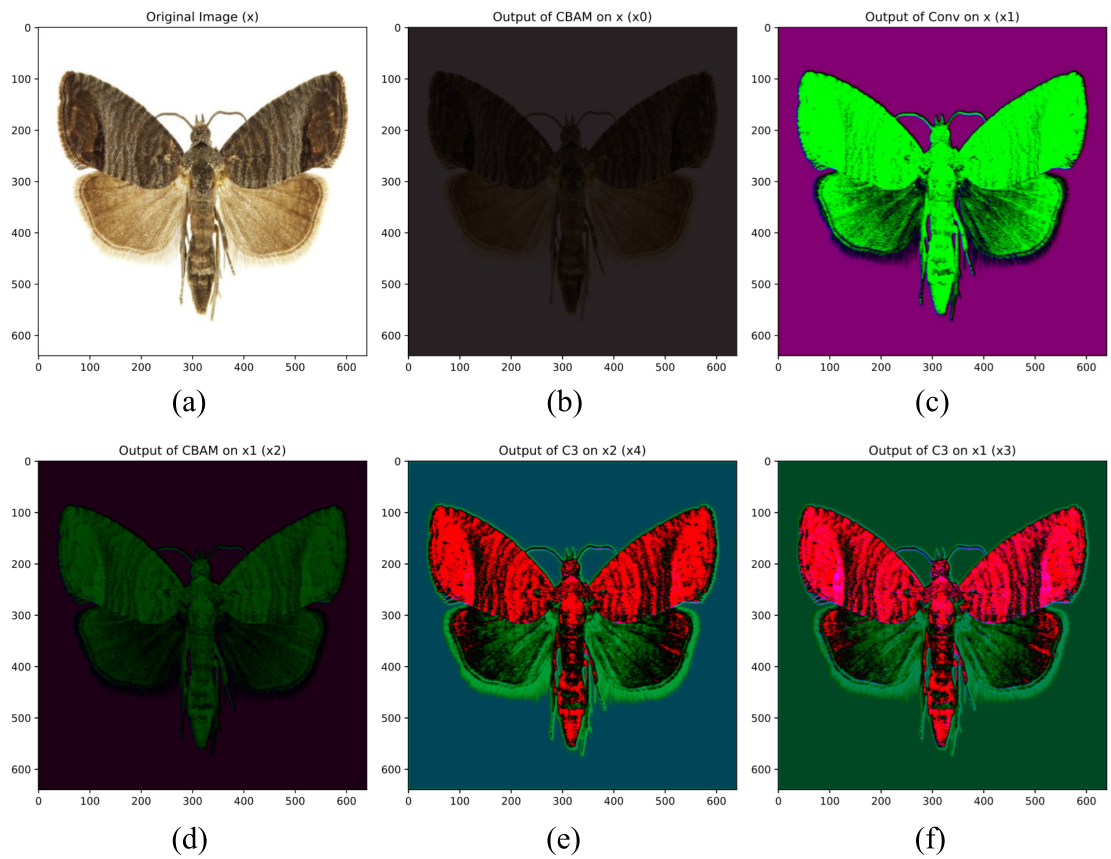


Figure 8: Comparison of outputs at different stages of the network with and without CBAM.

- **Original Image** (Figure 8(a)): This is the input image of a moth, which is passed through the model for object detection.
- **CBAM Output on Input** (Figure 8(b)): Applying CBAM directly to the input results in a darker image with minimal focus. This demonstrates that CBAM alone, without prior feature extraction, lacks sufficient information to focus effectively.
- **Conv Output** (Figure 8(c)): After applying a convolution layer, basic image features such as edges and textures become visible. These features serve as the foundation for subsequent layers to build more complex patterns.

- **CBAM Output on Conv** (Figure 8(d)): When CBAM is applied after convolution, the attention mechanism improves. CBAM directs the model's focus to key regions of the moth (e.g., wings and body) while suppressing background noise. This indicates that CBAM enhances feature extraction when paired with preliminary convolutional operations.
- **C3 Output on CBAM** (Figure 8(e)): Applying the C3 block after CBAM leads to the most precise and most focused output. The moth's essential features, particularly the wings and body, are highlighted, confirming that the attention-guided feature maps from CBAM significantly enhance the model's ability to extract meaningful information.
- **C3 Output on Conv** (Figure 8(f)): The C3 block further processes the features from the convolution, refining details like texture and structure. The moth's wings and body become more pronounced, indicating that C3 captures more complex features than a single convolutional layer.

The experimental results clearly show that adding CBAM after convolution improves the network's focus on relevant image regions. When combined with the C3 block, CBAM helps extract more complex and accurate features. These findings demonstrate that CBAM plays a crucial role in improving the model's ability to focus on key regions, thereby enhancing feature extraction for object detection tasks.

2.4. Firebase IoT Platform

The Firebase IoT Platform provides a flexible and comprehensive infrastructure for connecting and managing IoT devices and their associated data. Its intuitive interface, coupled with real-time database capabilities, simplifies the creation and scaling of IoT applications. With secure data transmission via authentication services and cloud functions, Firebase ensures reliable, optimized data handling, enhancing the performance and reliability of IoT solutions. Whether applied in home automation or industrial settings, Firebase equips developers with essential tools to streamline project development and improve the overall user experience. Firebase's cloud services are used to consolidate detection results from the improved YOLOv5 model deployed across various traps, enabling seamless remote monitoring and data management.

2.5. Raspberry Pi Trap

The Raspberry Pi-based insect trap is designed specifically for detecting codling moth (Cydia pomonella), offering an advanced, efficient solution for farmers to monitor pest populations in their fields. The trap uses pheromones to attract the moths, ensuring that targeted pests are effectively drawn for accurate monitoring. Once the moths are attracted, the integrated camera and the improved YOLOv5m model, enhanced with the Convolutional Block Attention Module (CBAM) to improve feature extraction and focus on relevant image areas, accurately detect and count the number of captured codling moths.

The trap is programmed to capture images at two specific times each day: 08:00 after sunrise and 18:00 just before sunset. This schedule is carefully chosen to coincide with the onset of codling moth activity, as moths typically begin flying and mating shortly before sunset and continue into the night under suitable conditions, avoiding direct sunlight and high temperatures. The Raspberry Pi's activation schedule adjusts throughout the year to align with seasonal variations in moth activity. This setup maximizes energy efficiency, enabling the system to operate autonomously with minimal power consumption.

Locally, the Raspberry Pi processes each image using the improved YOLOv5m model for real-time object detection and classification, identifying and counting codling moths with high precision. The detection pipeline achieves an inference speed of approximately 0.8 frames per second (FPS), with an average energy consumption of around 7 watts during operation. The results are transmitted over the internet via built-in WiFi, using the HTTP protocol to upload data to the Firebase IoT Platform for remote monitoring.

This IoT-enabled setup allows farmers to track pest populations in real time, providing critical insights into field conditions and enabling timely, targeted interventions. By adopting this technology, farmers can swiftly detect fluctuations in pest populations, reducing crop damage and pesticide use. The complete Raspberry Pi parameters used in the trap are listed in Table 1, while an overview of the system is illustrated in Figure 9.

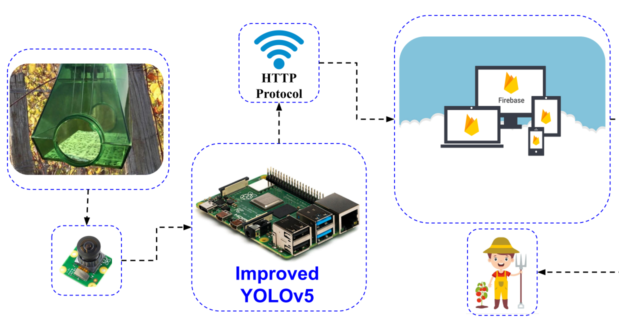


Figure 9: Raspberry Pi-Based insect trap with improved YOLOv5 and firebase IoT for remote monitoring [3].

Table 1: Parameters of raspberry Pi.

Configuration	Parameter
Raspberry Pi	Raspberry Pi 4B
RAM	2 GB
CPU	Broadcom BCM2711
Camera resolution	2592*1944
Wi-Fi	2.4 GHz and 5.0 GHz
Language	Python 3.9
Framework	Torch 1.13, Torchvision 0.14.0
Operating system	Raspberry Pi OS 64 bit

3. Results and Discussion

The improved YOLOv5m model has been selected for pest detection on a Raspberry Pi 4B because it provides an ideal balance between processing speed and detection accuracy. This model offers the necessary accuracy to effectively detect small pests while maintaining computational efficiency, making it well-suited for use with the Raspberry Pi. Its ability to perform well within the hardware constraints of the Raspberry Pi 4B ensures reliable, efficient pest detection across various agricultural settings.

3.1. Training Environment and Evaluation Metrics

Completing of the data labeling for both the training and validation sets marks a crucial step in the research. Afterward, we trained the improved YOLOv5m model in Google Colab, using an NVIDIA Tesla T4 GPU with 16 GB of memory, running CUDA 12.2 and driver 535.104.05. The training images were resized to 640 pixels. Key information, such as class labels and names, was organized in the data-yaml file and applied to both the training and validation datasets. The dataset was partitioned into 85% for training and 15% for validation, and the model was trained for 100 epochs.

The performance of the improved YOLOv5m was measured using several evaluation metrics, including precision, recall, mean average precision (mAP) at an IoU threshold of 0.5, mAP averaged across IoU thresholds from 0.5 to 0.95, and parameter count.

Precision measures the accuracy of positive predictions by determining the ratio of correctly predicted positive samples to the total number of predicted positive samples. The precision calculation is expressed as:

$$\text{Precision (P)} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \tag{1}$$

Recall measures the proportion of correct predictions out of all actual positive targets, and it is defined as follows:

$$\text{Recall (R)} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \tag{2}$$

The calculation formulas for mAPval 0.5 and mAPval 0.5:0.95 are provided below:

$$AP = \int_0^1 P(R)dR \tag{3}$$

$$mAP = \frac{\sum_{i=1}^{N} AP_i}{\text{N}} \tag{4}$$

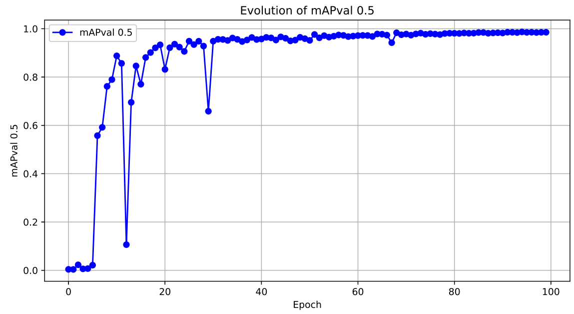
The mAPval 0.5 metric represents the average precision for detections with confidence scores above 0.5, which helps calculate precision and recall. Conversely, mAPval 0.5:0.95 evaluates the average precision over a range of confidence thresholds, from 0.5 to 0.95, in increments of 0.05. This provides a more comprehensive view of the model's performance across varying confidence levels. Model size refers to the storage required after the final phase of training. FLOPs (Floating Point Operations Per Second) measure the computational efficiency of a model, indicating the total number of arithmetic operations needed to process a single image. This metric is essential for assessing the feasibility of deploying models in resource-constrained environments.

To better assess the model's performance and training stability, Figure 10 presents the validation curves of mean Average Precision (mAP) over 100 epochs. Figure 10(a) shows the mAPval 0.5 trend, while Figure 10(b) displays the more stringent mAPval 0.5:0.95. Both metrics demonstrate a consistent upward trend during the initial training phase, followed by a convergence around epoch 75. This indicates that the model steadily learns to detect insect instances more accurately over time, with minimal oscillation or degradation, reflecting strong generalization capability. The smoothness of the curves suggests that the training process is stable and free from overfitting.

To further investigate model robustness and learning progression, Figure 11 illustrates the evolution of both training and validation loss curves for box loss (a) and objectness loss (b). The box loss, which quantifies the accuracy of predicted bounding box coordinates, exhibits a sharp decline in the early epochs for both training and validation sets before plateauing, indicating rapid spatial learning and consistent localization. Similarly, the objectness loss, which evaluates the model's confidence in identifying objects, shows a smooth and synchronized decrease across both datasets with no divergence or instability. These trends confirm that the proposed improved YOLOv5m architecture achieves efficient convergence and maintains training stability throughout the learning process.

2025, Vol. 2, Article ID. 2025.0026
https://doi.org/10.69709/CAIC.2025.150921

(a) Validation mAPval 0.5



(b) Validation mAPval 0.5:0.95

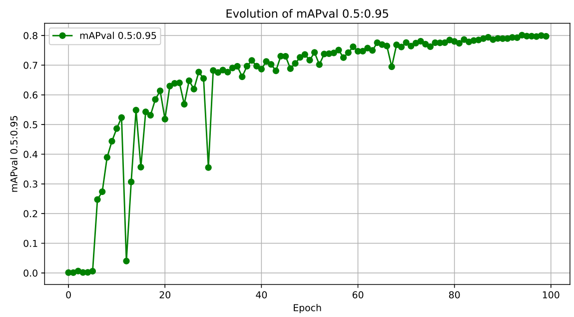
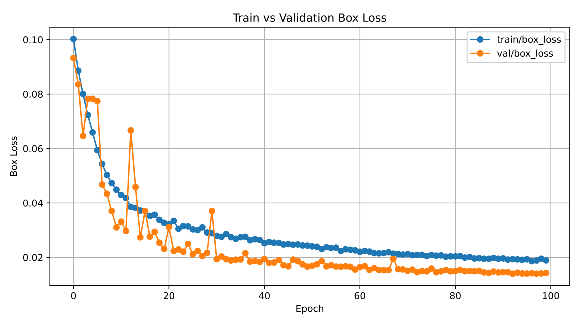


Figure 10: Validation mean Average Precision (mAP) curves over 100 epochs. (a) mAP evaluated at IoU threshold 0.5; (b) mAP evaluated at multiple thresholds from 0.5 to 0.95.

(a) Box Loss



(b) Objectness Loss

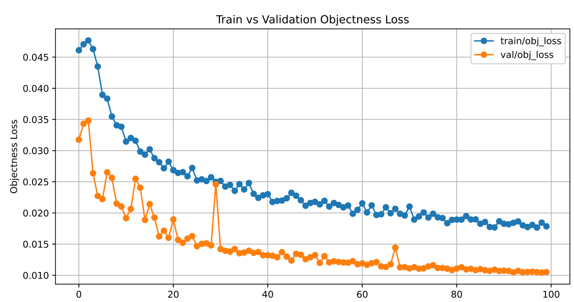


Figure 11: Training and validation loss curves over 100 epochs. (a) Box loss measures the accuracy of bounding box predictions; (b) Objectness loss reflects the model's confidence in detecting objects.

3.2. Enhanced Detection Performance with Improved YOLOv5m

In the development of pest monitoring systems, the deployment of the improved YOLOv5m model has significantly advanced pest detection accuracy, computational efficiency, and sustainable computing practices. Demonstrating substantial improvements over previous models, Figure 12 shows the original YOLOv5m model, achieving an average confidence level of 83.36% and a maximum of 88% while using 25.06 million parameters. In contrast:

- **YOLOv6m** (Figure 13) despite having the highest parameter count at 51.99 million, displays slightly lower average and minimum confidence levels of 81% and 67%, respectively, but reaches a higher maximum of 91%.
- **YOLOv8m** (Figure 14) maintains consistent performance with an average confidence level of 84% and a maximum of 88%, similar to the original YOLOv5m, but with slightly more efficient parameter usage at 25.85 million.
- **YOLOv9m** (Figure 15) offers a balance with 20.15 million parameters, an average confidence level of 83%, and a maximum of 90%.
- **YOLOv10m** (Figure 16) stands out for achieving the highest maximum confidence level of 94% with the lowest parameter count of 16.4 million, though its minimum confidence level drops to 52%.
- **YOLOv11m** (Figure 17) provides reliable performance with 20.05 million parameters, achieving an average confidence level of 81.2%, a maximum of 88%, and a minimum of 70%.
- **YOLOv12m** (Figure 18) — With 20.10 million parameters, this latest version of the YOLO model achieves a maximum confidence of 89%, a minimum of 57%, and an average confidence of 80.45%.

However, the improved YOLOv5m model, illustrated in Figure 19, elevated the maximum confidence level to 95% and the average to 91.36%, all while significantly reducing the parameter count to 16.98 million. This not only showcases enhanced model efficiency and accuracy but also highlights the model's contribution to sustainable computing by reducing computational demands.

Table 2 details the comparative performance of the YOLO series, underscoring the superior accuracy and efficiency of the improved YOLOv5m. This model stands out for its high performance and low energy consumption, achieving a significant reduction in FLOPs (26.88 billion) compared to other models and aligning with sustainable computing objectives that aim to minimize the environmental impacts of technology.

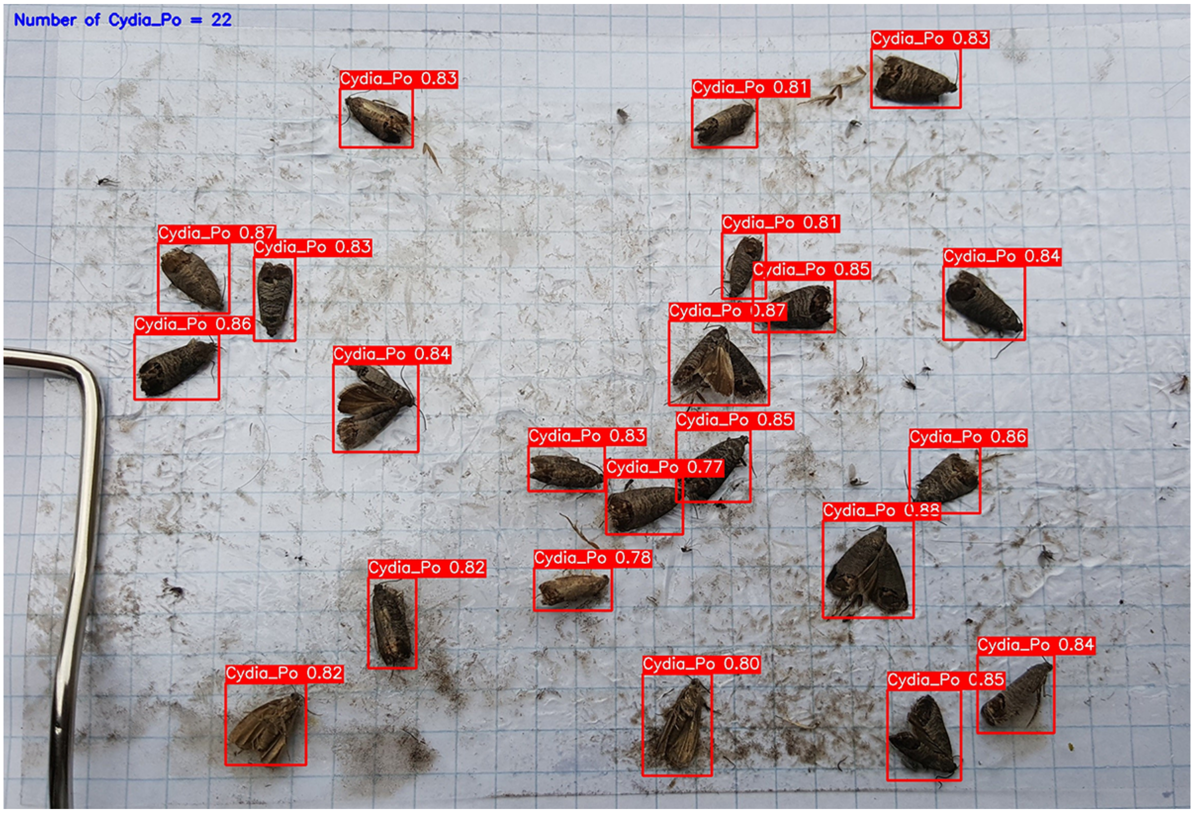


Figure 12: Detection performance of original YOLOv5m model (adapted from [74]).

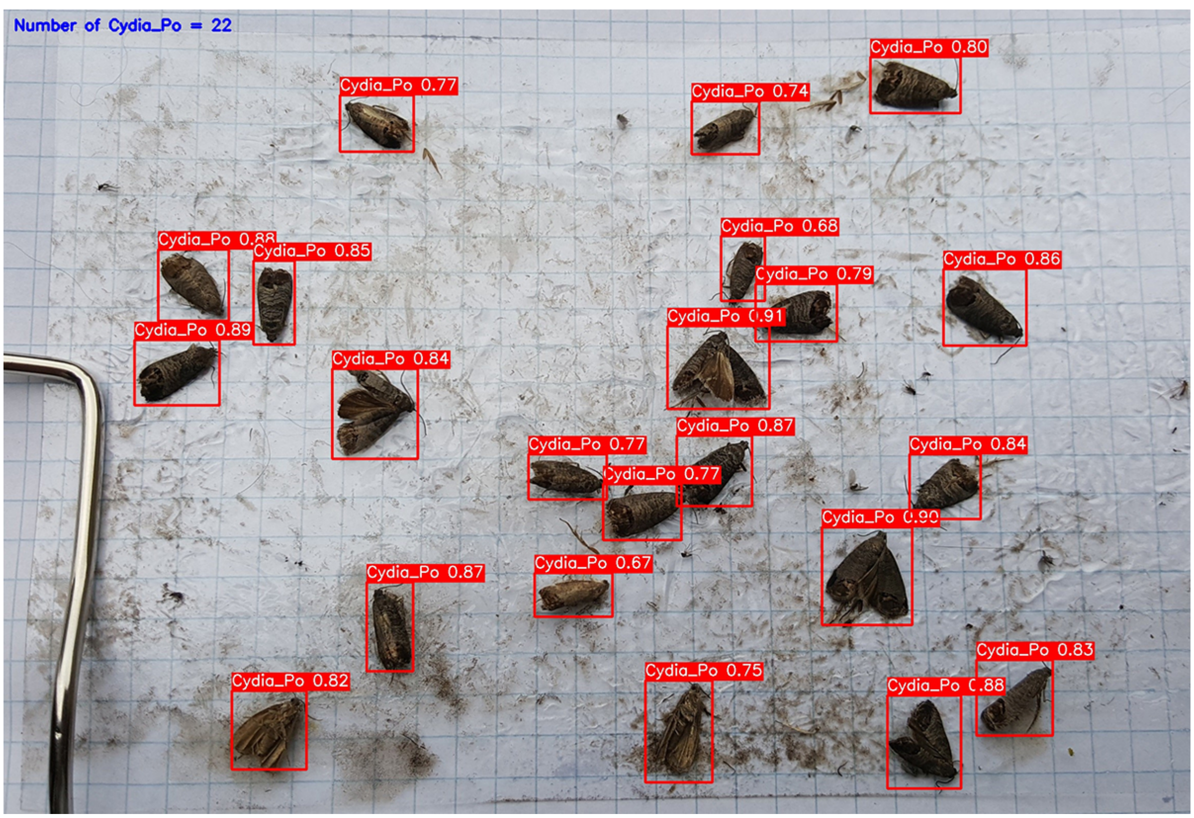


Figure 13: Performance evaluation of the YOLOv6m model in detecting cydia pomonella pests (adapted from [74]).

2025, Vol. 2, Article ID. 2025.0026
https://doi.org/10.69709/CAIC.2025.150921

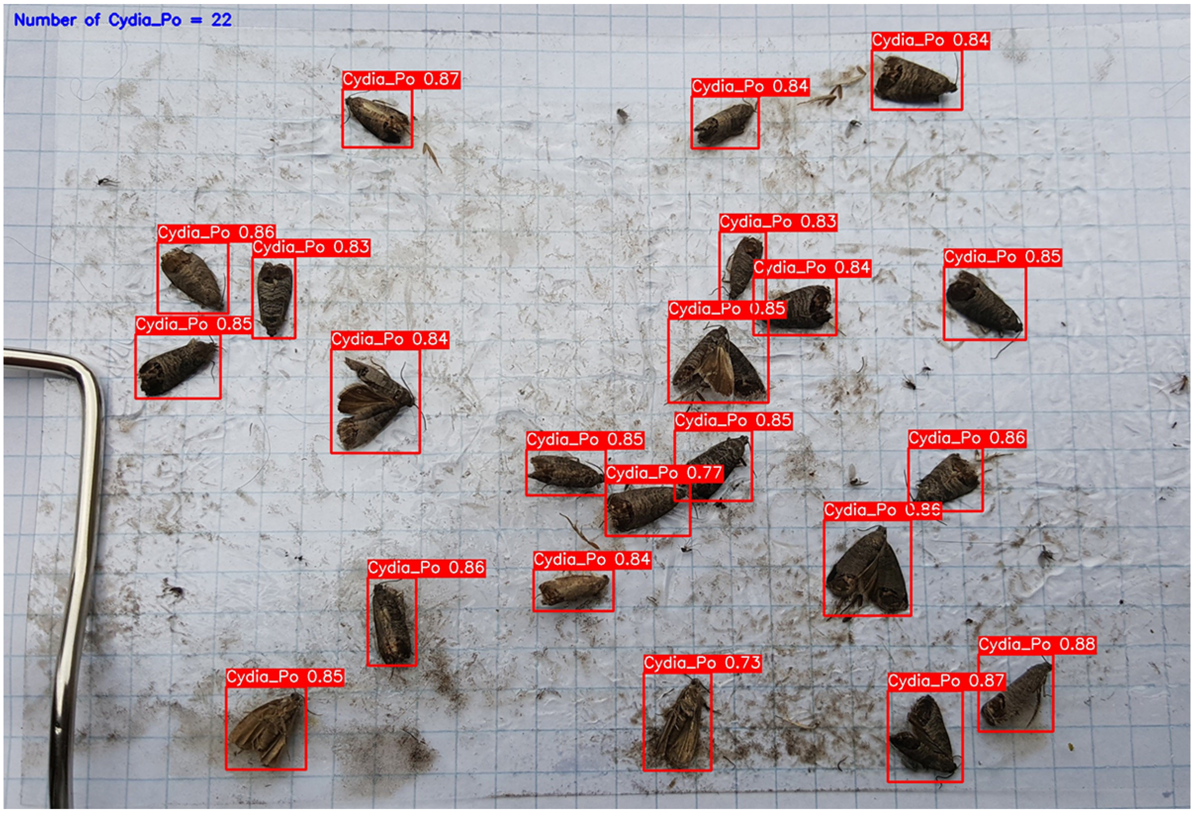


Figure 14: Performance evaluation of the YOLOv8m model in detecting cydia pomonella pests (adapted from [74]).

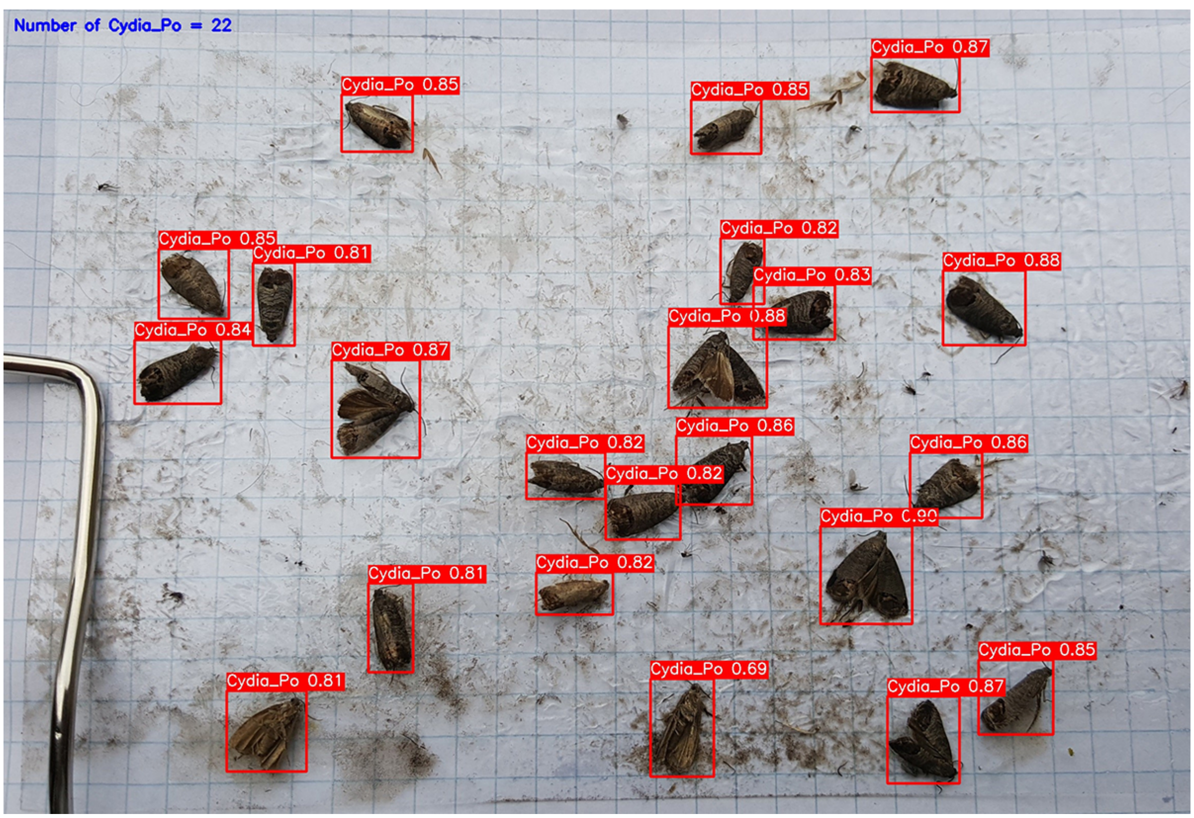


Figure 15: Performance evaluation of the YOLOv9m model in detecting cydia pomonella pests (adapted from [74]).

SCIFINITI PUBLISHING

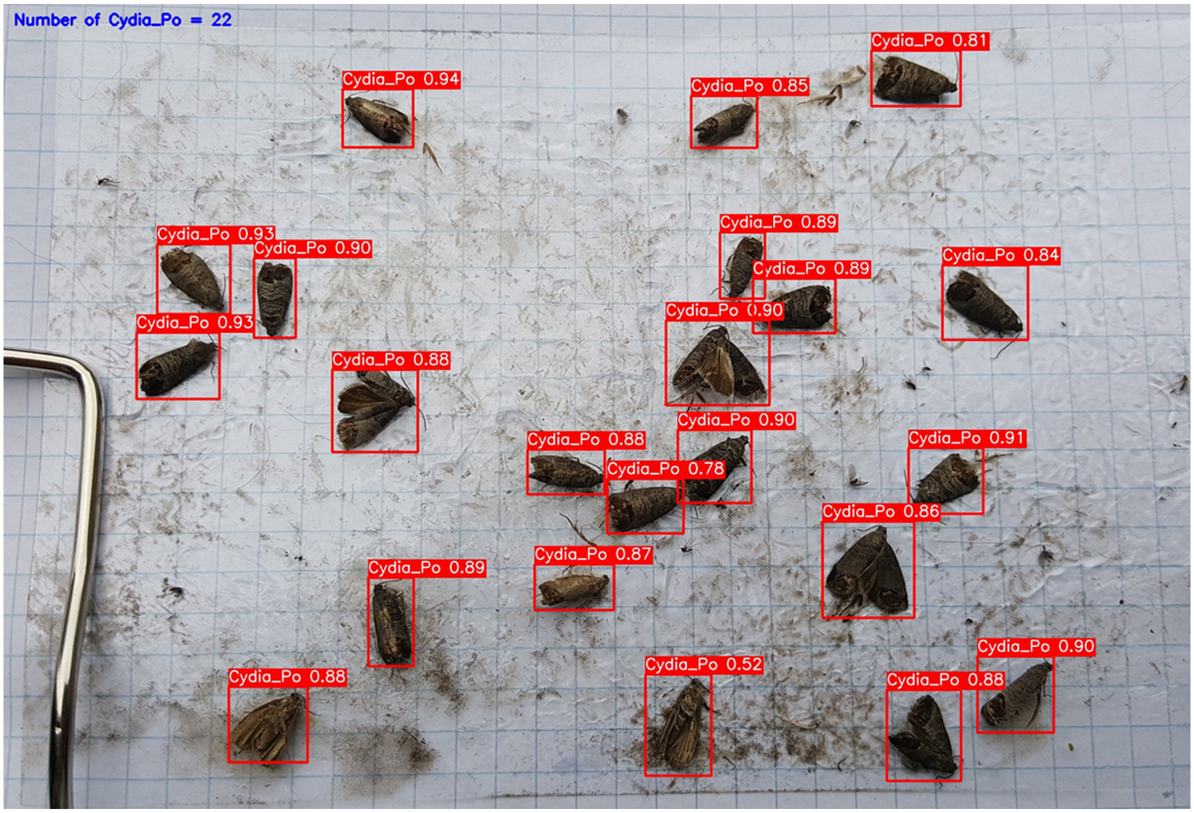


Figure 16: Performance evaluation of the YOLOv10m model in detecting cydia pomonella pests (adapted from [74]).

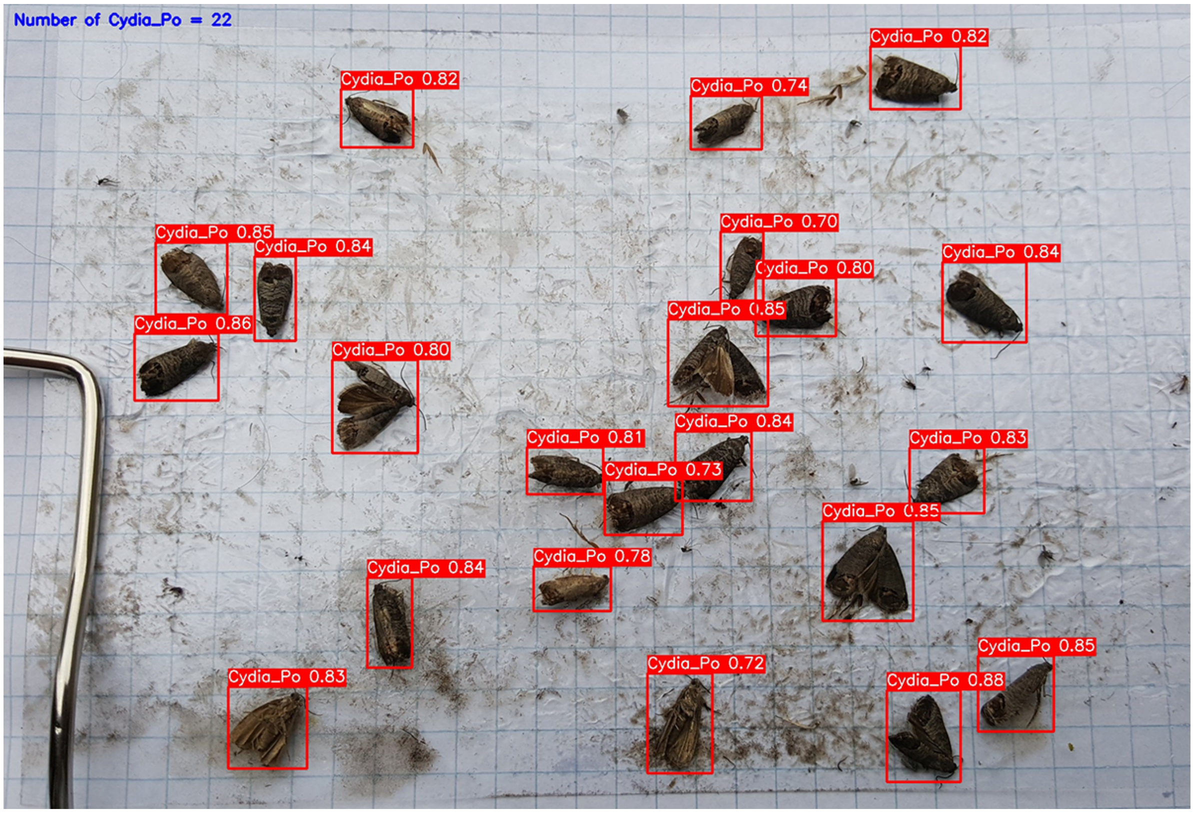


Figure 17: Performance evaluation of the YOLOv11m model in detecting cydia pomonella pests (adapted from [74]).

2025, Vol. 2, Article ID. 2025.0026
https://doi.org/10.69709/CAIC.2025.150921

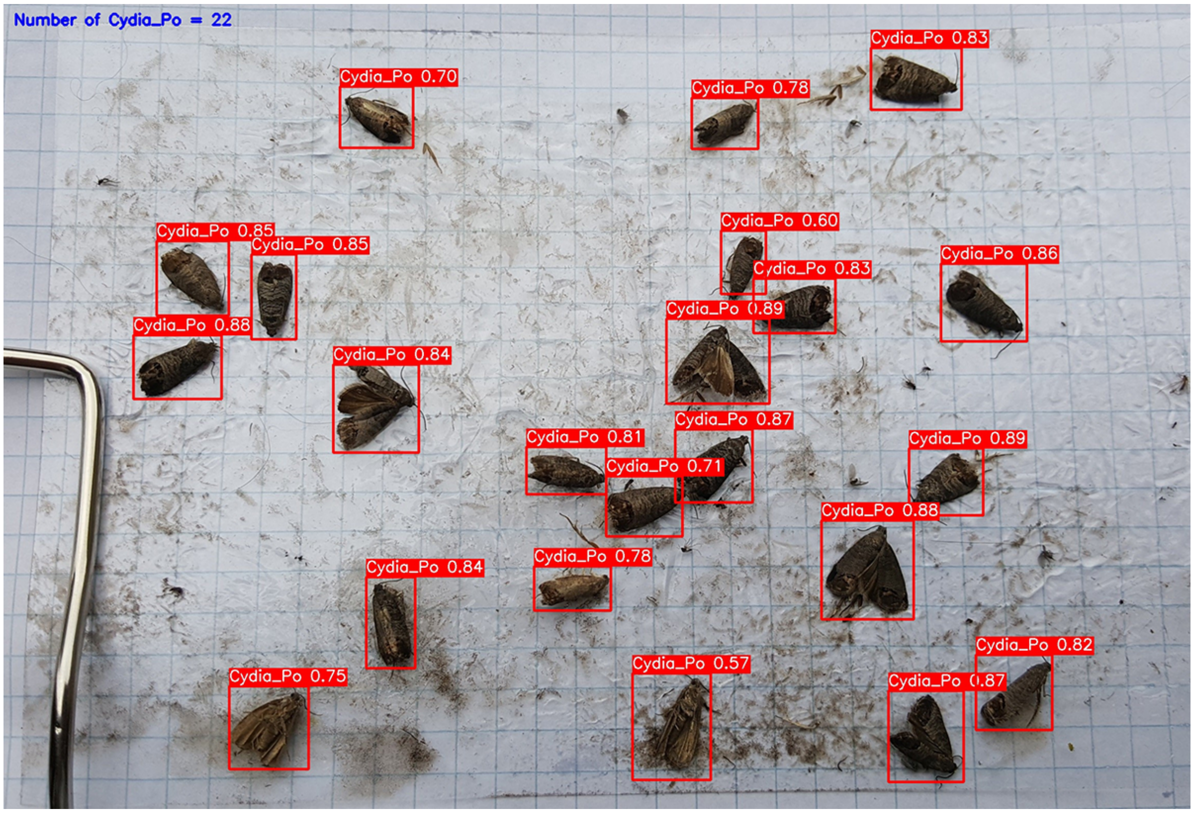


Figure 18: Performance evaluation of the YOLOv12m model in detecting cydia pomonella pests (adapted from [74]).

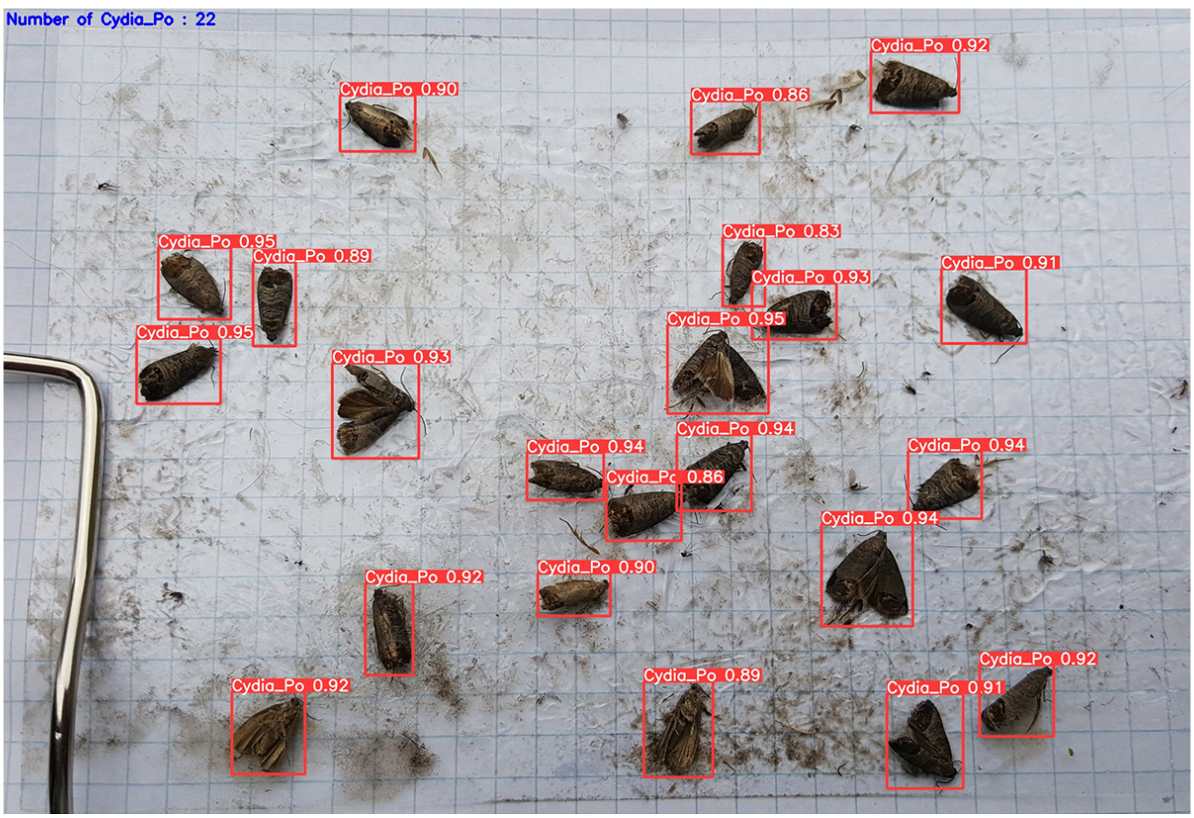


Figure 19: Enhanced detection performance of improved YOLOv5m model (adapted from [74]).

2025, Vol. 2, Article ID. 2025.0026
https://doi.org/10.69709/CAIC.2025.150921

Table 2: Comparative performance metrics of YOLO models.

Models	FLOPs (B)	Number of Parameters (M)	Average Confidence Level (%)	Min Confidence Level (%)	Max Confidence Level (%)
YOLOv5m	32.08	25.06	83.36	77	88
YOLOv6m	80.32	51.99	81	67	91
YOLOv8m	39.44	25.85	84	73	88
YOLOv9m	38.64	20.15	83	69	90
YOLOv10m	31.91	16.4	86.5	52	94
YOLOv11m	34.02	20.05	81.2	70	88
YOLOv12m	35.35	20.10	80.45	57	89
Improved YOLOv5m	**26.88**	**16.98**	**91.36**	**83**	**95**

The integration of this model into a Raspberry Pi-based trap system has revolutionized pest management strategies. Utilizing the Firebase IoT platform for real-time data transmission, the system enables farmers to remotely monitor pest activity and respond promptly, applying pesticides only where necessary. This precise and targeted approach drastically reduces pesticide use, lowering environmental footprints and promoting sustainable agricultural practices. Additionally, the robust data analytics provided by the system help farmers achieve higher yields by maintaining optimal pest control, thereby supporting a balance between enhanced productivity and environmental sustainability.

3.3. Firebase IoT Platform Integration for Pest Management

The integration of the Firebase IoT platform is crucial for efficiently monitoring pest-detection data collected from the trap. The Firebase storage is organized into two main folders:

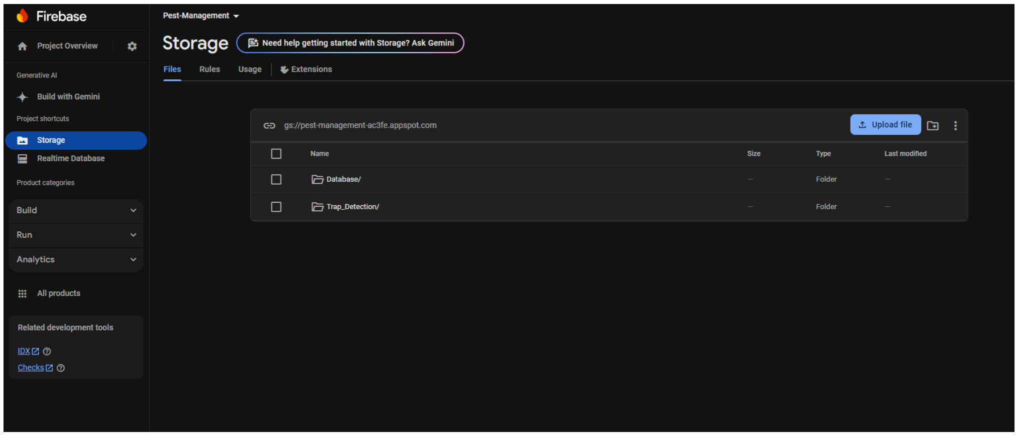
- **Database Folder:** This folder systematically stores all images captured by the trap at two specific intervals each day—08:00 after sunrise and 18:00 just before sunset. This ensures targeted data collection during peak codling moth activity periods, which is essential for periodic retraining of the detection model. As new images are consistently added to the dataset, the model can be retrained to maintain its accuracy and effectiveness in identifying pests, adapting to changing environmental conditions, and seasonal variations.
- **Trap_Detection Folder:** This folder contains only the most recent image captured and processed by the improved YOLOv5m model. It provides real-time updates on the pest activity within the trap. This feature enables farmers or agricultural professionals to access up-to-date information remotely, supporting timely decision-making for pest management interventions.

The Figure 20 below illustrates the Firebase platform's structure and data organization:

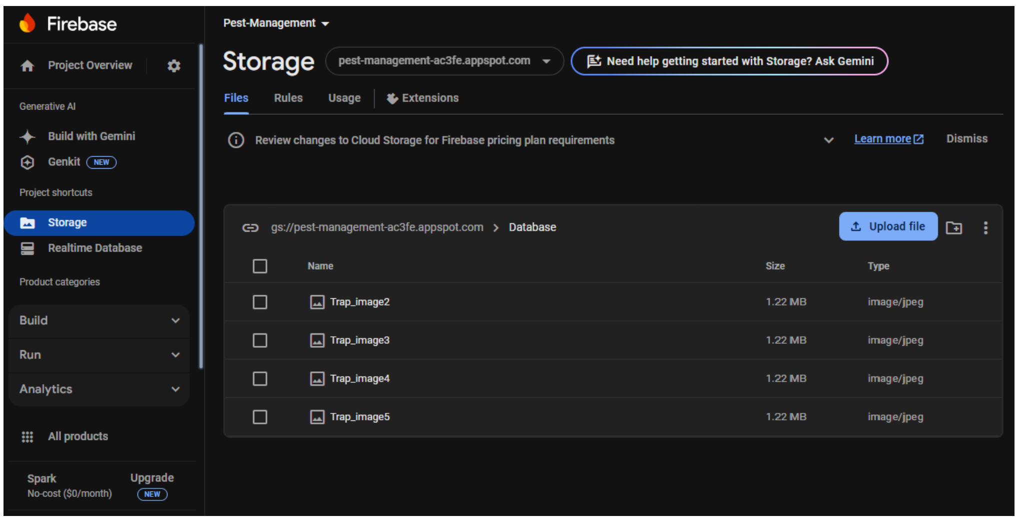
- Figure 20(a): Displays the overall folder structure, showing both the Database and Trap_Detection folders within Firebase storage.
- Figure 20(b): Depicts the contents of the Database folder, which includes a historical archive of all captured images, forming the dataset for retraining the model with updated data.
- Figure 20(c): Shows the most recent image processed by the model, with bounding boxes highlighting detected pests, enabling real-time monitoring.

The proposed Firebase IoT-based system significantly enhances pest detection by integrating intelligent pest recognition models directly within the traps. This localized image processing reduces data transmission loads, minimizes power consumption, and ensures rapid pest identification. Unlike conventional methods, this system enables real-time tracking of insect populations in the field, providing more precise and timely data for decision-making. The combination of embedded processing and cloud-based analytics ensures scalability and accessibility, allowing farmers to optimize pest management strategies efficiently while reducing environmental impact.

(a)



(b)



(c)

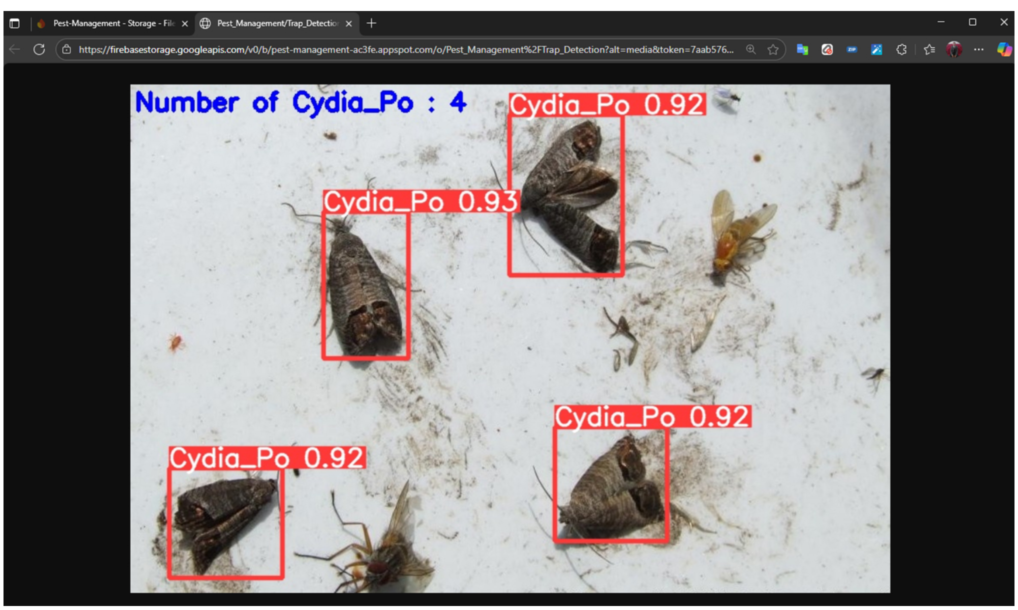


Figure 20: Firebase IoT platform: (a) folder structure with database and trap detection; (b) database folder showing image archive; (c) recent detection image with bounding boxes (adapted from [74]).

4. Future Work

This research establishes a foundation for the development of more adaptive, efficient, and intelligent pest monitoring systems within the context of precision agriculture. Future work will focus on enhancing the system's robustness, generalization capabilities, and ecological relevance by addressing practical challenges observed in real-world agricultural deployments.

1. **Integration of Environmental and Weather Data for Context-Aware Pest Management:** To enhance the system's adaptability and decision-making capabilities, future work will focus on integrating both on-site environmental sensing and regional weather forecasts. IoT-enabled sensors will be deployed to collect real-time data on key variables such as temperature, humidity, and ambient light, in the field. In parallel, regional meteorological data—including wind speed, precipitation, and forecasted conditions—will be incorporated into the system's notification and control module. This dual-source integration will enable the system to correlate pest population behavior with environmental triggers, support predictive modeling, and recommend optimal timing for pesticide application. Such alignment with favorable conditions will reduce chemical waste, improve treatment efficacy, and promote environmentally sustainable pest control practices.
2. **Incorporation of Test-Time Adaptation (TTA) for Robustness to Domain Shifts:** Given the inherent variability in field conditions—such as changes in lighting, background clutter, trap designs, and crop types—we propose the integration of lightweight Test-Time Adaptation (TTA) techniques. These methods enable the model to adjust its internal representations dynamically at inference time, without requiring access to labeled data or retraining. By adapting to previously unseen distributions during deployment, TTA can significantly improve detection reliability in heterogeneous and evolving agricultural environments [75].
3. **Exploration of Advanced Data Augmentation Techniques to Enhance Generalization:** In addition to biologically realistic rotations, future experiments will incorporate more diverse augmentation strategies to improve model generalization, especially under few-shot or class-imbalanced conditions. Notably, the Random Interpolation Resize (RIR) technique will be evaluated [75], which introduces interpolation variability during resizing to improve robustness to distributional shifts. Furthermore, style-based transformations inspired by content-style disentanglement and contrastive learning [76] will be explored to simulate environmental variability—such as lighting, color tone, and background texture—thereby improving the model's ability to generalize across different agricultural contexts.

5. Conclusions

This study has demonstrated significant advances in pest detection technology by optimizing the YOLOv5m model. The improved YOLOv5m, enhanced with the Convolutional Block Attention Module (CBAM) and filter reduction, achieved a maximum confidence level of 95% and an average confidence level of 91.36%, surpassing the original YOLOv5m. These enhancements were achieved by reducing the parameter count and lowering FLOPs to 26.88 billion, further demonstrating the model's suitability for deployment on low-power edge devices and alignment with sustainable computing principles.

The integration of this optimized model into a Raspberry Pi-based trapping system, coupled with real-time data management, enhances pest monitoring efficiency. This system enables real-time tracking of pest populations and supports precision interventions, ultimately reducing pesticide use, mitigating environmental impacts, and improving crop yields.

Additionally, the comparative evaluation of YOLO variants provided valuable insights into the trade-offs between accuracy, computational complexity, and generalization capacity. The improved YOLOv5m demonstrated an ideal balance between detection performance and resource efficiency, making it a strong candidate for real-world applications in smart agriculture.

Overall, the findings underscore the transformative potential of integrating advanced deep learning techniques with IoT infrastructure for pest management. This approach offers a scalable path toward more sustainable, automated, and intelligent farming systems.

List of Abbreviations

C2f	Modified feature extraction block
C3	CSPNet-based feature extraction block
CBAM	Convolutional Block Attention Module
CNN	Convolutional Neural Network
FAO	Food and Agriculture Organization
FPS	Frames Per Second
GPU	Graphics Processing Unit
IoT	Internet of Things
IoU	Intersection over Union

mAP	Mean Average Precision
SPPF	Spatial Pyramid Pooling-Fast
YOLO	You Only Look Once

Author Contributions

Writing-review & editing, writing-original draft, visualization, validation, methodology, investigation, formal analysis, data curation, funding acquisition, conceptualization: M.Z.; Writing-review & editing, resources, validation: A.B.; Supervision, writing-review & editing, validation: S.C.; Supervision, writing-review & editing, validation: A.D. All authors have read and agreed to the published version of the manuscript.

Availability of Data and Materials

The dataset and source code used in this work will be available on request.

Conflicts of Interest

The authors declare no conflicts of interest.

Funding

The authors are grateful to the National Center for Scientific and Technical Research of Morocco (CNRST) for their financial assistance and guidance provided through their supervision (grant number: 26 UIZ2022).

Acknowledgments

The authors would like to acknowledge the National School of Applied Sciences, Ibn Zohr University, Agadir, Morocco, for providing research facilities and support throughout this study.

AI Declaration

The authors confirm that no AI tools were used to generate any content of this manuscript.

References

[1] M. Roser and L. Rodés-Guirao, “Future population growth,” 2013. Accessed: Apr. 2, 2025. [Online]. Available: https://ourworldindata.org/future-population-growth

[2] Y. Yang and N. Zhao, “Future projections of compound temperature and precipitation extremes and corresponding population exposure over global land,” *Glob. Planet. Change*, vol. 236, 2024, Art. no. 104427. [CrossRef]

[3] M. Zarboubi, A. Bellout, S. Chabaa, and A. Dliou, “Custombottleneck-vggnet: Advanced tomato leaf disease identification for sustainable agriculture,” *Comput. Electron. Agric.*, vol. 232, 2025, Art. no. 110066. [CrossRef]

[4] D. Fróna, J. Szenderák, and M. Harangi-Rákos, “The challenge of feeding the world,” *Sustainability*, vol. 11, 2019, Art. no. 5816. [CrossRef]

[5] I. Zunzunegui, J. Martín-García, Ó. Santamaría, and J. Poveda, “Insect frass as an agricultural resource against abiotic and biotic crop stresses: Mechanisms of action and possible negative effects,” *Appl. Sci.*, vol. 15, 2025, Art. no. 3606. [CrossRef]

[6] M. Carvajal-Yepes et al., “A global surveillance system for crop diseases,” *Science*, vol. 364, pp. 1237–1239, 2019. [CrossRef]

[7] B. S. Faiçal et al., “An adaptive approach for uav-based pesticide spraying in dynamic environments,” *Comput. Electron. Agric.*, vol. 138, pp. 210–223, 2017. [CrossRef]

[8] FAO, ”New standards to curb the global spread of plant pests and diseases,” 2019. Accessed: Apr. 2, 2025. [Online]. Available: https://www.fao.org/newsroom/detail/New-standards-to-curb-the-globalspread-of-plant-pests-and-diseases/en

[9] M. E. Karar, F. Alsunaydi, S. Albusaymi, and S. Alotaibi, “A new mobile application of agricultural pests recognition using deep learning in cloud computing system,” *Alex. Eng. J.*, vol. 60, pp. 4423–4432, 2021. [CrossRef]

[10] S. Arora, T. Husain, and S. M. Prasad, “Allelochemicals as biocontrol agents: Promising aspects, challenges and opportunities,” *South Afr. J. Bot.*, vol. 166, pp. 503–511, 2024. [CrossRef]

[11] H. Gautham and M. Bhardwaj, “International year of plant health 2020 to focus on threat of pests on food security,” *Curr. Sci.*, vol. 118, 2020, Art. no. 857. [View Online]

[12] S. Narenderan, S. Meyyanathan, and B. Babu, “Review of pesticide residue analysis in fruits and vegetables. pre-treatment, extraction and detection techniques,” *Food Res. Int.*, vol. 133, 2020, Art. no. 109141. [CrossRef]

[13] I. Pajač, I. Pejić, and B. Barić, “Codling moth, cydia pomonella (lepidoptera: Tortricidae)–major pest in apple production: An overview of its biology, resistance, genetic structure and control strategies,” *Agric. Conspec. Sci.*, vol. 76, pp. 87–92, 2011. [View Online]

[14] A. Y. Khaled et al., “Non-destructive detection of codling moth infestation in apples using acoustic impulse response signals,” *Biosyst. Eng.*, vol. 224, pp. 68–79, 2022. [CrossRef]

[15] M. Zarboubi, A. Bellout, S. Chabaa, A. Dliou, and A. Zeroual, “Iot-based pest detection in agriculture using raspberry pi and yolov10m for precision farming,” *ITM Web Conf.*, vol. 69, 2024, Art. no. 04017. [CrossRef]

[16] R. P. Sharma, R. Dharavath, and D. R. Edla, “Ioftfis: Internet of farm things based prediction for crop pest infestation using optimized fuzzy inference system,” *Internet Things*, vol. 21, 2023, Art. no. 100658. [CrossRef]

[17] J. She et al., "A method for automatic real-time detection and counting of fruit fly pests in orchards by trap bottles via convolutional neural network with attention mechanism added," *Ecol. Inform.*, vol. 70, 2022, Art. no. 101690. [CrossRef]

[18] M. Lippi, N. Bonucci, R. F. Carpio, M. Contarini, S. Speranza, and A. Gasparri, "A YOLO-based pest detection system for precision agriculture," in *Proc. 2021 29th Mediterr. Conf. Control Autom. (MED)*, Puglia, Italy, 2021, pp. 342–347. [CrossRef]

[19] M.-L. Huang, T.-C. Chuang, and Y.-C. Liao, "Application of transfer learning and image augmentation technology for tomato pest identification," *Sustain. Comput. Inform. Syst.*, vol. 33, 2022, Art. no. 100646. [CrossRef]

[20] M. Zarboubi, S. Chabaa, and A. Dliou, "Advancing precision agriculture with deep learning and IoT integration for effective tomato pest management," in *Proc. 2023 IEEE Int. Conf. Adv.-Data-Driven Anal. Intell. Syst. (ADACIS)*, Marrakesh, Morocco, 2023, pp. 1–6. [CrossRef]

[21] A. Bellout, M. Zarboubi, M. Elhoseny, A. Dliou, R. Latif, and A. Saddik, "Lt-YOLOv10n: A lightweight IoT-integrated deep learning model for real-time tomato leaf disease detection and management," *Internet Things*, vol. 33, 2025, Art. no. 101663. [CrossRef]

[22] A. Bellout, M. Zarboubi, A. Dliou, R. Latif, and A. Saddik, "Yolov10-enabled IoT robot car for accurate disease detection in strawberry cultivation," *ITM Web Conf.*, vol. 69, 2024, Art. no. 01011. [CrossRef]

[23] G. G. Devarajan, S. M. Nagarajan, T. Ramana, T. Vignesh, U. Ghosh, and W. Alnumay, "Ddnsas: Deep reinforcement learning based deep q-learning network for smart agriculture system," *Sustain. Comput. Inform. Syst.*, vol. 39, 2023, Art. no. 100890. [CrossRef]

[24] L. M. Dang et al., "An efficient zero-labeling segmentation approach for pest monitoring on smartphone-based images," *Eur. J. Agron.*, vol. 160, 2024, Art. no. 127331. [CrossRef]

[25] T. Wang, L. Zhao, B. Li, X. Liu, W. Xu, and J. Li, "Recognition and counting of typical apple pests based on deep learning," *Ecol. Inform.*, vol. 68, 2022, Art. no. 101556. [CrossRef]

[26] L. Nanni, G. Maguolo, and F. Pancino, "Insect pest image detection and recognition based on bio-inspired methods," *Ecol. Inform.*, vol. 57, 2020, Art. no. 101089. [CrossRef]

[27] M. D. Chodey and C. Noorullah Shariff, "Hybrid deep learning model for in-field pest detection on real-time field monitoring," *J. Plant Dis. Prot.*, vol. 129, pp. 635–650, 2022. [CrossRef]

[28] W. Li, T. Zheng, Z. Yang, M. Li, C. Sun, and X. Yang, "Classification and detection of insects from field images using deep learning for smart pest management: A systematic review," *Ecol. Inform.*, vol. 66, 2021, Art. no. 101460. [CrossRef]

[29] T. Ayoub Shaikh, T. Rasool, and F. Rasheed Lone, "Towards leveraging the role of machine learning and artificial intelligence in precision agriculture and smart farming," *Comput. Electron. Agric.*, vol. 198, 2022, Art. no. 107119. [CrossRef]

[30] K. Li, J. Wang, H. Jalil, and H. Wang, "A fast and lightweight detection algorithm for passion fruit pests based on improved yolov5," *Comput. Electron. Agric.*, vol. 204, 2023, Art. no. 107534. [CrossRef]

[31] W. Xu et al., "A lightweight ssv2-yolo based model for detection of sugarcane aphids in unstructured natural environments," *Comput. Electron. Agric.*, vol. 211, 2023, Art. no. 107961. [CrossRef]

[32] M. Zarboubi, A. Bellout, S. Chabaa, and A. Dliou, "Revolutionizing tomato pest management: Synergy of deep learning, IoT, and precision agriculture," *Math. Model. Comput.*, vol. 11, pp. 1167–1178, 2024. [CrossRef]

[33] H. M. Abdullah, M. N. Islam, M. H. Saikat, and M. A. Bhuiyan, "Precision agriculture practices from planting to postharvest: Scopes, opportunities, and challenges of innovation in developing countries," in *Remote Sensing in Precision Agriculture*, Amsterdam, The Netherlands: Elsevier, 2024, pp. 3–26. [CrossRef]

[34] A. Taseer and X. Han, "Advancements in variable rate spraying for precise spray requirements in precision agriculture using unmanned aerial spraying systems: A review," *Comput. Electron. Agric.*, vol. 219, 2024, Art. no. 108841. [CrossRef]

[35] W. Ding, M. Abdel-Basset, I. Alrashdi, and H. Hawash, "Next generation of computer vision for plant disease monitoring in precision agriculture: A contemporary survey, taxonomy, experiments, and future direction," *Inf. Sci.*, vol. 665, 2024, Art. no. 120338. [CrossRef]

[36] A. Amrani, D. Diepeveen, D. Murray, M. G. Jones, and F. Sohel, "Multi-task learning model for agricultural pest detection from crop-plant imagery: A bayesian approach," *Comput. Electron. Agric.*, vol. 218, 2024, Art. no. 108719. [CrossRef]

[37] G. A. Senthil, S. U. Suganthi, L. Prinslin, R. Selvi, and R. Prabha, "Generative AI in agri: Sustainability in smart precision farming yield prediction mapping system based on gis using deep learning and gps," *Procedia Comput. Sci.*, vol. 252, pp. 365–380, 2025. [CrossRef]

[38] K. Joshi et al., "Precision diagnosis of tomato diseases for sustainable agriculture through deep learning approach with hybrid data augmentation," *Curr. Plant Biol.*, vol. 41, 2025, Art. no. 100437. [CrossRef]

[39] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Comput. Electron. Agric.*, vol. 147, pp. 70–90, 2018. [CrossRef]

[40] H. Li, E. Yan, J. Jiang, and D. Mo, "Monitoring of key camellia oleifera phenology features using field cameras and deep learning," *Comput. Electron. Agric.*, vol. 219, 2024, Art. no. 108748. [CrossRef]

[41] H. Liu et al., "Upgrading swin-b transformer-based model for accurately identifying ripe strawberries by coupling task-aligned one-stage object detection

mechanism," *Comput. Electron. Agric.*, vol. 218, 2024, Art. no. 108674. [CrossRef]
[42] W. Zhou, Y. Cui, H. Huang, H. Huang, and C. Wang, "A fast and data-efficient deep learning framework for multi-class fruit blossom detection," *Comput. Electron. Agric.*, vol. 217, 2024, Art. no. 108592. [CrossRef]
[43] M. Ariza-Sentís, S. Vélez, R. Martínez-Peña, H. Baja, and J. Valente, "Object detection and tracking in precision farming: A systematic review," *Comput. Electron. Agric.*, vol. 219, 2024, Art. no. 108757. [CrossRef]
[44] K. Hu et al., "Deep learning techniques for in-crop weed recognition in large-scale grain production systems: A review," *Precis. Agric.*, vol. 25, pp. 1–29, 2024. [CrossRef]
[45] H. Li and F. Shi, "A detr-like detector-based semi-supervised object detection method for brassica chinensis growth monitoring," *Comput. Electron. Agric.*, vol. 219, 2024, Art. no. 108788. [CrossRef]
[46] I. Kalfas, B. De Ketelaere, K. Bunkens, and W. Saeys, "Towards automatic insect monitoring on witloof chicory fields using sticky plate image analysis," *Ecol. Inform.*, vol. 75, 2023, Art. no. 102037. [CrossRef]
[47] C. Chen, Y. Liang, L. Zhou, X. Tang, and M. Dai, "An automatic inspection system for pest detection in granaries using yolov4," *Comput. Electron. Agric.*, vol. 201, 2022, Art. no. 107302. [CrossRef]
[48] F. Zhang, R. Bao, B. Yan, M. Wang, Y. Zhang, and S. Fu, "Lsannet: A lightweight convolutional neural network for maize leaf disease identification," *Biosyst. Eng.*, vol. 248, pp. 97–107, 2024. [CrossRef]
[49] M. Zarboubi, S. Chabaa, A. Dliou, and A. Zeroual, "Smart pest control in grain warehouses: Yolov8-powered iot robot car for precision agriculture," in *Proc. 2024 Int. Conf. Glob. Aeronaut. Eng. Satell. Technol. (GAST)*, Marrakesh, Morocco, 2024, pp. 1–6. [CrossRef]
[50] Y. Peng and Y. Wang, "Cnn and transformer framework for insect pest classification," *Ecol. Inform.*, vol. 72, 2022, Art. no. 101846. [CrossRef]
[51] W. Fang et al., "Identification of wormholes in soybean leaves based on multi-feature structure and attention mechanism," *J. Plant Dis. Prot.*, vol. 130, pp. 401–412, 2023. [CrossRef]
[52] S. Dong, Y. Teng, L. Jiao, J. Du, K. Liu, and R. Wang, "Esa-net: An efficient scale-aware network for small crop pest detection," *Expert Syst. Appl.*, vol. 236, 2024, Art. no. 121308. [CrossRef]
[53] A. Bellout, M. Zarboubi, A. Dliou, R. Latif, and A. Saddik, "Advanced YOLO models for real-time detection of tomato leaf diseases," *Math. Model. Comput.*, vol. 11, pp. 1198–1210, 2024. [CrossRef]
[54] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Las Vegas, NV, USA, 2016, pp. 779–788. [CrossRef]
[55] J. Redmon and A. Farhadi, "Yolo9000: Better, faster, stronger," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Honolulu, HI, USA, 2017, pp. 7263–7271. [CrossRef]
[56] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," 2018. [CrossRef]
[57] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "Yolov4: Optimal speed and accuracy of object detection," 2020. [CrossRef]
[58] G. Jocher et al., "ultralytics/yolov5: v3. 0," *Zenodo*, 2020. [CrossRef]
[59] C. Li et al., "Yolov6: A single-stage object detection framework for industrial applications," 2022. [CrossRef]
[60] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Vancouver, BC, Canada, 2023, pp. 7464–7475. [CrossRef]
[61] G. Jocher, A. Chaurasia, and J. Qiu. "Yolo by ultralytics," *Github* 2023. [View Online]
[62] C.-Y. Wang, I.-H. Yeh, and H.-Y. M. Liao, "Yolov9: Learning what you want to learn using programmable gradient information," 2024. [CrossRef]
[63] A. Wang et al., "Yolov10: Real-time end-to-end object detection," 2024. [CrossRef]
[64] Ultralytics, "Yolov11—Key features," 2024. Accessed: Apr. 2, 2025. [Online]. Available: https://docs.ultralytics.com/models/yolo11/#key-features
[65] Y. Tian, Q. Ye, and D. Doermann, "Yolov12: Attention-centric real-time object detectors," 2025. [CrossRef]
[66] M. M. Rahman et al., "Combined use of sticky traps and biopesticides as a sustainable tool to manage aleurocanthus rugosa (hemiptera: Aleyrodidae) infesting betel vine," *Crop Prot.*, vol. 172, 2023, Art. no. 106299. [CrossRef]
[67] Y. K. Shin, S. B. Kim, and D.-S. Kim, "Attraction characteristics of insect pests and natural enemies according to the vertical position of yellow sticky traps in a strawberry farm with high-raised bed cultivation," *J.-Asia-Pac. Entomol.*, vol. 23, pp. 1062–1066, 2020. [CrossRef]
[68] J. Suto, "Codling moth monitoring with camera-equipped automated traps: A review," *Agriculture*, vol. 12, 2022, Art. no. 1721. [CrossRef]
[69] F. Lello et al., "Fruit fly automatic detection and monitoring techniques: A review," *Smart Agric. Technol.*, vol. 5, 2023, Art. no. 100294. [CrossRef]
[70] W. Li, D. Wang, M. Li, Y. Gao, J. Wu, and X. Yang, "Field detection of tiny pests from sticky trap images using deep learning in agricultural greenhouse," *Comput. Electron. Agric.*, vol. 183, 2021, Art. no. 106048. [CrossRef]
[71] I. Potamitis, I. Rigakis, N. Vidakis, M. Petousis, and M. Weber, "Affordable bimodal optical sensors to

spread the use of automated insect monitoring," *J. Sens.*, vol. 2018, 2018, Art. no. 3949415. [CrossRef]

[72] M. Cardim Ferreira Lima, M. E. Damascena de Almeida Leandro, C. Valero, L. C. Pereira Coronel, and C. O. Gonçalves Bazzo, "Automatic detection and monitoring of insect pests—A review," *Agriculture*, vol. 10, 2020, Art. no.161. [CrossRef]

[73] D. Akroute et al., "Influence of some fruit traits on codling moth (*Cydia pomonella* L.) preference among apple varieties in two contrasted climatic conditions," *Horticulturae*, vol. 9, 2023, Art. no. 788. [CrossRef]

[74] J. Sütő, "Embedded system-based sticky paper trap with deep learning-based insect-counting algorithm," *Electronics*, vol. 10, 2021, Art. no. 1754. [CrossRef]

[75] P. Li, H. Tao, H. Zhou, P. Zhou, and Y. Deng, "Enhanced Multiview attention network with random interpolation resize for few-shot surface defect detection," *Multimed. Syst.*, vol. 31, 2024, Art. no. 36. [CrossRef]

[76] Z. Wang, H. Tao, H. Zhou, Y. Deng, and P. Zhou, "A content-style control network with style contrastive learning for underwater image enhancement," *Multimed. Syst.*, vol. 31, 2025, Art. no. 60. [CrossRef]