# High Resolution Environmental Monitoring of Pollinator Insects through Macro Camera Trapping and AI

Mohammad Sa'doun[1], Daniel Dalton[2], Vanessa Berger[2], Gernot Paulus[1], Anders Karl-Heinrich[1], Ilja Svetnik[2], Johanna Schultz[2] and Peter Unglaub[2]

[1]*Carinthia University of Applied Science, SIENA – Spatial Informatics for Environmental Applications, Villach, Austria*

[2]*Carinthia University of Applied Sciences, UNESCO Chair on Sustainable Management of Conservation Areas, Villach, Austria*

## Abstract

Insect identification using camera traps has been neglected due to traditional sensors' inability to detect cold-blooded organisms. However, insects are essential for ecosystem services like pollination, making accurate detection crucial. This study uses advancements in computer vision, specifically convolutional neural networks (CNNs), to develop an automated insect identification algorithm. High-resolution time-lapse images from camera traps in the Austrian Alps were annotated and processed with LabelBox and YOLOv8, achieving order-level insect classification. Results showed a high diversity of insects with a peak mean Average Precision (mAP) of 54.23% at an Intersection over Union (IoU) threshold of 0.5, indicating the method's potential for precise species-level data collection. Despite challenges like background misclassification, the approach offers a robust framework for insect detection and valuable insights for refinement. YOLO-generated data enable comprehensive time series analysis, aiding in effective monitoring and management strategies. This study supports CNN-based methods in ecological monitoring, relevant for pest management and identifying key species.

## Keywords

Pollination ecosystem service, CNN, YOLOv8, Machine learning, Insect identification, Camera traps

## 1. Introduction

Automatic identification of insects using camera traps has been largely neglected because most camera trap sensors work based on heat differential between the target animal and the surrounding environment. Insects, being cold-blooded, largely assume the temperature of their environment; therefore, this heat differential is negligible. However, insects play a critical role in the environment, and their detection and classification are crucial for understanding their impact. Recent advancements in computer vision technology, particularly convolutional neural networks (CNNs), have enabled accurate insect detection and classification. As a pilot action towards developing an automated insect identification algorithm supported by the FFG-funded project BioMONITec, we proceeded with a camera trapping design focused on pollinator insects. We gathered our own images using a time-lapse approach to document insect visitors on blooming flowers. Images were annotated using the data management platform LabelBox (LabelBox Inc., San Francisco, CA, USA) and processed for automated recognition through the open-source program YOLOv8 (Ultralytics Inc., Los Angeles, CA, USA), achieving order-level classification of insects. This approach aligns with other studies demonstrating the potential of CNNs in agricultural pest management. For instance, [1] showed that YOLO architectures, including YOLOv5, could effectively detect and identify insect pests with high precision, which is essential for reducing pesticide use and promoting spot spraying. Furthermore, [2] employed Faster R-CNN with inception ResNet v2 to detect and count pests in greenhouse tomato crops, achieving strong correlations between network detections and human observations. These studies highlight the importance of monitoring pests and the potential of deep learning to automate this process, making it more efficient and accurate. The integration of AI and automation in agriculture, including computer vision and other remote monitoring technologies, offers promising solutions for integrated pest management. Despite challenges such as detecting small

✉ m.sadoun@fh-kaernten.at (M. Sa'doun)

insects and improving real-time detection accuracy, advancements in CNN-based methods are paving the way for broader applications, including the identification of economically and medically important species like pollinators and mosquitoes.

## 2. Materials and Methods

Our four-phase methodology included data acquisition, annotation, YOLOv8 training/validation and post processing. During the data acquisition phase, high-resolution camera traps are utilized to capture insects. Subsequently, the annotation phase involved annotating the acquired images using the LabelBox platform. Then, the output from the previous phases, comprising of the images and their corresponding labels, was fed into YOLO for training and validation purposes. The objective of this phase was to assess the efficacy of AI in detecting and accurately classifying insects. The last phase was the post-processing of AI detections where we performed a time series analysis to plot the occurrences of different insect types for a particular dataset.

### 2.1. Data acquisition and study sites

Waterproof outdoor digital cameras (WG-70, Ricoh Co., Ltd., Tokyo, Japan) were installed in Weinitzen and Sittersdorf (Carinthia), and St. Margarethen im Lungau (Salzburg) using tripods (Joby GorillaPod 1K, Vivendum Plc, Richmond, UK). Cameras targeted 10 blooming annual plant species over 13 dates ( Table 1). Plant species had a range of flower colors and phenology. These factors were assumed to impact the different types of insects over the season. Cameras utilized a macro setting at the highest photo quality (16MP) and were programmed to trigger at 30- to 45-second intervals. Devices exposed to intense sunlight occasionally shut down due to overheating, thus impacting the planned schedule of time lapse photo capture.

**Table 1**
Summary statistics on the number of images taken at specific host plants at the three study sites.

| Camera | Date | Location | Host | Start time | End time | Photos taken |
|---|---|---|---|---|---|---|
| 1 | 09-May | St. Margarethen | Borago officinalis | 09:50 | 19:19 | 1116 |
| 2 | 09-May | St. Margarethen | Borago officinalis | 09:52 | 19:20 | 1116 |
| 3 | 26-Jun | Sittersdorf | Achillea millefolium | 05:06 | 20:03 | 1001 |
| 4 | 26-Jun | Sittersdorf | Achillea millefolium | 05:12 | 20:06 | 1001 |
| 5 | 26-Jun | Sittersdorf | Borago officinalis | 05:06 | 20:03 | 1001 |
| 6 | 02-Jul | St. Margarethen | Inula helenium | 07:51 | 16:11 | 666 |
| 1 | 03-Jul | St. Margarethen | Filipendula ulmaria | 07:54 | 20:23 | 1000 |
| 5 | 03-Jul | Sittersdorf | Achillea millefolium | 09:39 | 17:59 | 1000 |
| 6 | 09-Jul | St. Margarethen | Inula helenium | 09:02 | 17:21 | 1000 |
| 7 | 10-Jul | St. Margarethen | Inula helenium | 07:06 | 15:35 | 640 |
| 8 | 13-Jul | Weinitzen | Dianthus carthusianorum | 09:06 | 21:35 | 1000 |
| 9 | 13-Jul | Weinitzen | Knautia arvensis | 09:29 | 21:58 | 1000 |
| 1 | 17-Jul | St. Margarethen | Inula helenium | 10:54 | 19:14 | 1000 |
| 6 | 17-Jul | St. Margarethen | Filipendula ulmaria | 10:52 | 19:11 | 1000 |
| 7 | 17-Jul | St. Margarethen | Inula helenium | 10:57 | 19:16 | 1000 |
| 10 | 17-Jul | St. Margarethen | Hypericum perforatum | 11:01 | 19:20 | 1000 |
| 8 | 19-Jul | Weinitzen | Tragopogon pratensis | 09:34 | 19:17 | 1000 |
| 9 | 20-Jul | Weinitzen | Oreoselinum majus | 08:05 | 20:34 | 1000 |
| 9 | 27-Jul | Weinitzen | Prunella grandiflora | 07:04 | 19:33 | 1000 |
| 8 | 02-Aug | Weinitzen | Oreoselinum majus | 09:24 | 19:07 | 1000 |
| 3 | 09-Aug | Weinitzen | Oreoselinum majus | 09:50 | 19:32 | 1000 |
| 5 | 09-Aug | Weinitzen | Oreoselinum majus | 09:50 | 19:32 | 1000 |
| Total | | | | | | 24,280* |

## 2.2. Image Annotation with LabelBox

In the current study, all visible parts of individual insects were fitted as tightly as possible within corresponding annotated bounding box. LabelBox was used as the annotation program that fed into YOLOv8 analysis. The process of drawing a bounding box is simply to click-and-drag over the object of interest. There are several important features that contribute to a good annotation in LabelBox, specifically in the context of bounding boxes. A good bounding box annotation should accurately capture the entire object of interest, without cutting off any parts or leaving out any areas. This is important for downstream tasks like object recognition and classification, as well as for visualizing the annotated objects. Additionally, a good annotation should have a clear and appropriate ratio of object to background. The bounding box should tightly surround the object of interest, while minimizing the inclusion of unnecessary background elements that can interfere with object recognition or classification. Furthermore, a good annotation should be consistent across all instances of a given object, as well as accurate in terms of location and size. Consistency is important for training machine learning models, as it helps to ensure that the model is able to recognize the same object across different images. Accuracy is crucial for ensuring that the annotated data is of high quality and can be used for downstream tasks with confidence. The data were converted in a pre-processing step into x-y coordinates with a calculated width and height of each object and saved as a small .txt file. Annotations were then coded, where each unique classification was assigned a number. Numbers corresponded to a mapping function where the classification was designated to an order of insect. Next, the annotations were randomly divided into 'training' and 'validation' data sets. To address the class imbalance and potential overfitting, we implemented a stratified random sampling technique for the train-validation split. This approach ensured that each insect class (Chelicerata, Coleoptera, Diptera, Hemiptera, Hymenoptera, and Lepidoptera) was represented proportionally in both sets. While we aimed to mitigate the imbalance, particularly between Diptera and Hemiptera, we acknowledge that the initial disparity might have persisted to some extent due to the nature of the data. For instance, multiple insect classes, especially Diptera, can often appear within the same image, making perfect class balance challenging. While the stratified sampling strategy helped to alleviate the impact of class imbalance to some extent, it's important to note that the model's performance might still be influenced by the inherent limitations of the dataset, particularly regarding underrepresented classes. The actual counts used in both training and validation are shown (Table 2).

**Table 2**
Training and validation sets sizes used in the study.

| Order | Training count | Validation count | Total count |
|---|---|---|---|
| Helicerata | 216 | 32 | 248 |
| Coleoptera | 261 | 43 | 304 |
| Diptera | 2874 | 523 | 3397 |
| Hemiptera | 200 | 37 | 237 |
| Hymenoptera | 343 | 64 | 307 |
| Lepidoptera | 215 | 31 | 246 |

# 3. Results

## 3.1. YOLOv8

YOLO v8 was used in the pilot action. YOLO is an open-source object detection algorithm created by [3] that can be modified and customized by developers to suit their specific needs. It takes a single pass over an image and can detect multiple objects at once. It uses a neural network that is trained on a large data set of labelled images, and it breaks down the image into a grid of cells. Each cell is responsible for predicting bounding boxes which are then classified into the categories of objects found in the dataset.

YOLO can accurately detect objects in an image even when the object is partially occluded. The model was trained for 100 epochs, peaking at epoch 88 with 54.23% mAP50 (Figure 1)
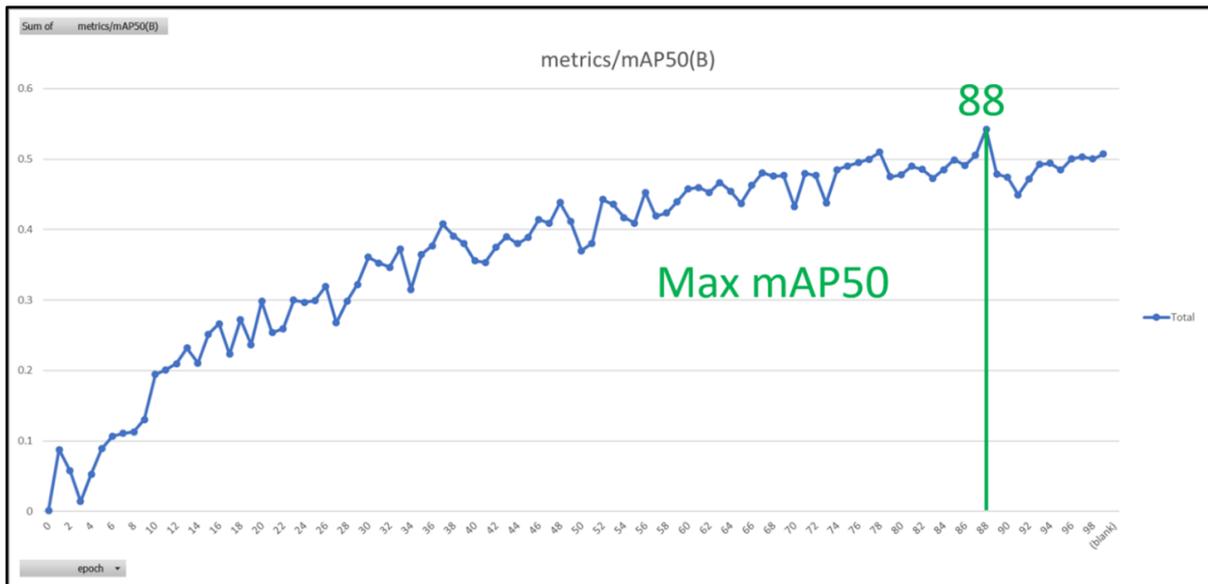


**Figure 1:** Training process progress with mAP50 saturates at epoch 88.

The validation process involves evaluating the model on the validation set and computing how close was the model predictions to the ground truth. The validation process is represented in the form of a confusion matrix (Figure 2). Each cell in the confusion matrix is bounded by a color-coded border to represent different aspects of the model's performance. The matrix is organized based on normalized numbers related to the Ground Truth count of classes in the validation set. Green cells signify true positives, yellow cells indicate true negatives or misclassification, red cells represent undetected instances, and black cells denote background misclassified as insects. This confusion matrix offers a detailed breakdown of the model's detection and classification outcomes, providing valuable insights into its strengths and areas requiring improvement. Some examples of correctly classified insects are provided (Figure 3).

### 3.2. Time Series Analysis

The analysis processes detection text files generated by YOLO, organizing the information into a table grouped by time detected and insect taxonomy. Subsequently, charts are generated to illustrate the frequency of insect sightings for each species at different times of the day (Figure 4). This chart serves as a tool for understanding when various types of insects are most active, contributing valuable insights for insect monitoring and management strategies.

## 4. Discussion and Conclusion

Camera trapping for automated identification of insects is in its infancy, although some promising case studies exist (e.g., [4]). Our study made significant strides towards advancing this goal, but much work remains. Our results show an impressive diversity of insects in a narrow window of time, considering only a single season of effort. We found a method to gather highly precise species-level data, allowing researchers and museum curators to showcase the recovered species to a wider audience in multiple ways. The well-defined data annotation process in LabelBox is a key contributor to the pipeline, ensuring the model is trained on accurately labelled data and promoting precise object detection. At the current status, while YOLO demonstrated a baseline level of performance in detecting and classifying insects, the achieved mean Average Precision (mAP) of 50% indicates room for improvement. The confusion
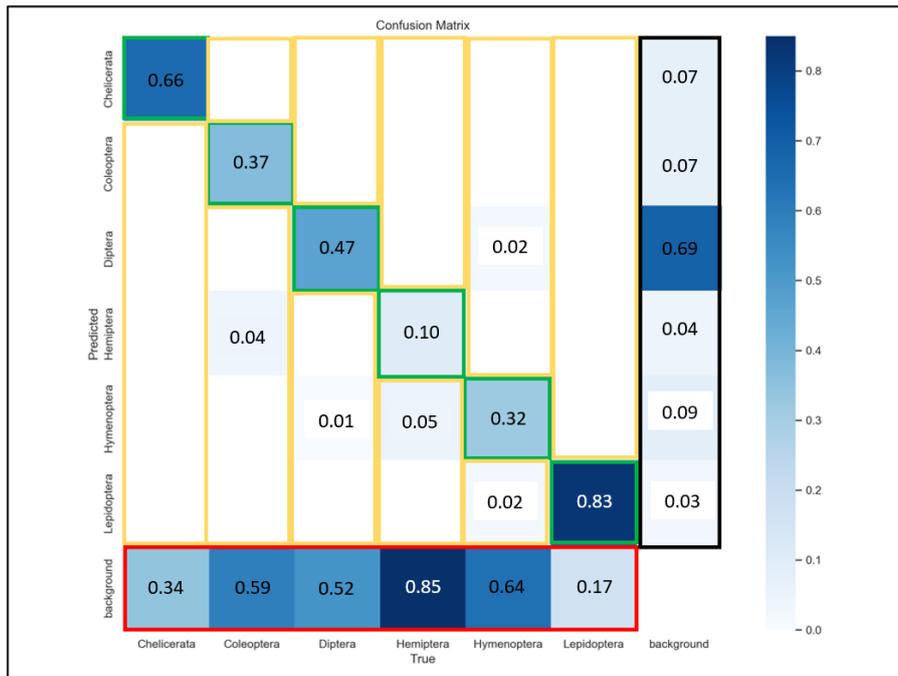
**Figure 2:** Confusion matrix. Green (Correct classification), Yellow (Misclassified), Red (Undetected), Black (Background detected as insects)
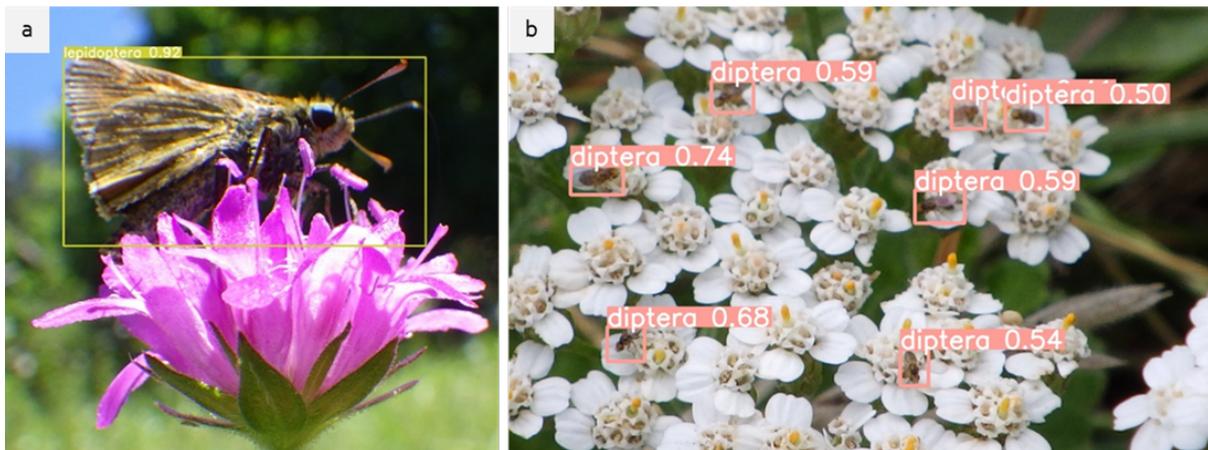


**Figure 3:** Visual results sample. a) Lepidoptera; b) Diptera.

matrix further highlights specific areas where the model struggled to differentiate between certain insect classes. It involves pre-processing to split the data into training and validation sets, training to minimize the loss function, and evaluation to assess performance. YOLOv8's advantages include quick detection and classification, flexibility, adaptability, and lightweight architecture. However, the model's tendency to detect background elements as insects, especially in complex backgrounds, highlights a challenge that needs fine-tuning in future iterations. Addressing this issue could enhance the model's precision and reliability. The structured data generated by YOLO allow for comprehensive time series analysis of temporal patterns in insect detection, providing insights into the temporal behavior of different species and aiding in developing informed monitoring and management strategies. In conclusion, the successful YOLO upgrade, effective data processing pipeline, and well-defined annotation process contribute to a robust framework for insect detection. The results allow for meaningful time series analysis and provide valuable insights for ongoing improvements to address challenges such as background detection. This ongoing refinement ensures that the model remains a powerful tool for accurate and insightful insect
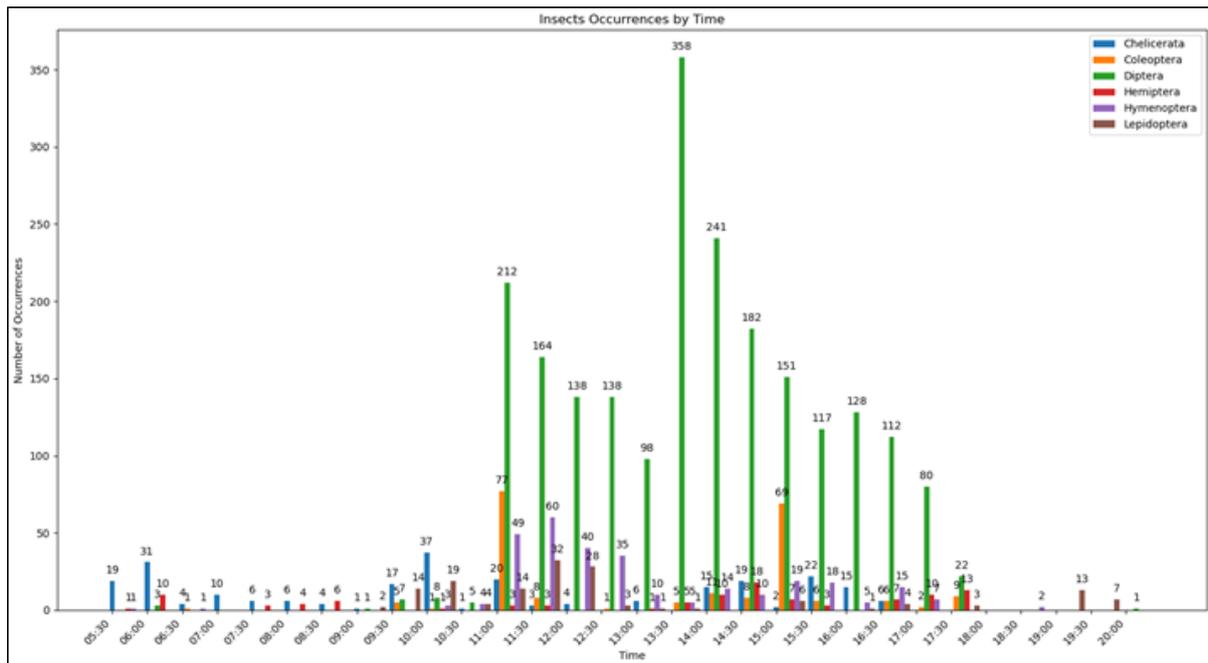
**Figure 4:** Bar chart time series analysis of the insect dataset.

monitoring and management.

## 5. Outlook

For the next steps, we plan to complete the annotation phase in LabelBox to fully incorporate the data into the model. Following this, we will initiate the training and evaluation process. Throughout these stages, we will regularly monitor the model's performance, making any necessary updates and adjustments. Once these steps are successfully completed, we aim to deploy the model in a practical application for real-world use.

## Acknowledgements

## References

[1] 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, Applied Sciences 12 (2022) 10167.

[2] A. Nieuwenhuizen, J. Hemming, H. K. Suh, Detection and classification of insects on stick-traps in a tomato crop using faster r-cnn, The Netherlands Conference on Computer Vision (2018).

[3] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You only look once: Unified, real-time object

detection, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 779–788.

[4] J. Alison, J. M. Alexander, N. Diaz Zeugin, Y. L. Dupont, E. Iseli, H. M. Mann, T. T. Høye, Moths complement bumblebee pollination of red clover: a case for day-and-night insect surveillance, Biology Letters 18 (2022) 20220187.