

Monitoring fast-moving animals—Building a customized camera system and evaluation toolset

Katharina Wittmann¹  | Mohamed Gamal Ibrahim² | Andrew David Straw³  |
Alexandra-Maria Klein¹  | Michael Staab^{1,4} 

¹Nature Conservation and Landscape Ecology, Faculty of Environment and Natural Resources, University of Freiburg, Freiburg, Germany

²Computer Science, Faculty of Engineering, University of Freiburg, Freiburg, Germany

³Institute of Biology I and Bernstein Center Freiburg, Faculty of Biology, University of Freiburg, Freiburg, Germany

⁴Ecological Networks, Technical University Darmstadt, Darmstadt, Germany

Correspondence

Katharina Wittmann
Email: katharina_wittmann@gmx.de

Funding information

Baden-Württemberg Stiftung, Grant/Award Number: 1.16101.19; Volkswagen Foundation, Grant/Award Number: 98692; Wissenschaftliche Gesellschaft Freiburg

Handling Editor: Miguel Acevedo

Abstract

1. Automated cameras (including camera traps) are an established observation tool, allowing, for example the identification of behaviours and monitoring without harming organisms. However, limitations including imperfect detection, insufficient data storage and power supply restrict the use of camera traps, making inexpensive and customizable solutions desirable. We describe a camera system and evaluation toolset based on Raspberry Pi computers and YOLOv5 that can overcome those shortcomings with its modular properties. We facilitate the set-up and modification for researchers via detailed step-by-step guides.
2. A customized camera system prototype was constructed to monitor fast-moving organisms on a continuous schedule. For testing and benchmarking, we recorded mason bees (*Osmia cornuta*) approaching nesting aids on 20 sites. To efficiently process the extensive video material, we developed an evaluation toolset utilizing the convolutional neural network YOLOv5 to detect bees in the videos.
3. In the field test, the camera system performed reliably for more than a week (2 h per day) under varying weather conditions. YOLOv5 detected and classified bees with only 775 original training images. Overall detection reliability varied depending on camera perspective, site and weather conditions, but a high average detection precision (78%) was achieved, which was confirmed by a human observer (80% of algorithm-based detections confirmed).
4. The customized camera system mitigates several disadvantages of commercial camera traps by using interchangeable components and incorporates all major requirements a researcher has for working in the field including moderate costs, easy assembly and an external energy source. We provide detailed user guides to bridge the gap between ecology, computer science and engineering.

KEYWORDS

artificial intelligence, convolutional neural network, Hymenoptera, low-budget, object classification, Raspberry Pi, videos, YOLOv5

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Authors. *Methods in Ecology and Evolution* published by John Wiley & Sons Ltd on behalf of British Ecological Society.

1 | INTRODUCTION

Camera systems including camera traps are valuable tools to collect standardized and replicated data, because they allow remote observations without interfering with study objects (Caravaggi et al., 2017; Meek et al., 2014). Also, data can be collected simultaneously over large areas, particularly when direct human observation is not feasible (Moore et al., 2021). Camera trap studies have, however, mainly focused on vertebrates (Delisle et al., 2021) and several technical difficulties continue to limit their applicability on other taxa. Most wildlife cameras trigger via a passive infrared motion sensor, which can cause imperfect detection depending on camera placement, background, animal body temperature and movement speed (Burton et al., 2015; Krauss et al., 2018). Being small ectothermic organisms, monitoring insects and other arthropods is particularly challenging (Tobler et al., 2008; Welbourne et al., 2016). Thus, only a few studies focused on arthropods, predominantly for pest management (Cardim Ferreira Lima et al., 2020).

Scheduled filming, that is, recording continuously within a fixed schedule, is an alternative to overcome detection issues as no individual is missed during the recording (Bonelli et al., 2020). However, almost all commercial wildlife cameras have limited programmability, cannot film for long durations (usually 60s) and have limited power (usually eight AA batteries or power grid access required) or data storage (SD card compatibility usually <128GB). This can be overcome by customized camera systems with interchangeable hardware tailored to specific needs, for example by extended data storage and long filming schedules. The broad availability of inexpensive low-energy single board computers like Raspberry Pi (Raspberry Pi Foundation, Cambridge, UK) offers a solution. Such systems are modular and flexible, allowing to assemble a customized camera system aided by the availability of detailed guides and documentation (Halfacre, 2018; Monk, 2020). By applying efficient power management, customized systems can monitor for extended periods of time without maintenance, thus generating continuous data that surpass commercial camera traps (e.g. PICT camera system, Droissart et al., 2021).

Shortcomings of scheduled recordings are the long video sequences without target species and the data quantities produced, which require substantial storage capacity. Machine learning techniques can be applied to recognize and filter relevant images automatically and have a potentially higher detection rate than human observers (Kulyukin & Mukherjee, 2019; Naqvi et al., 2022). Due to their relatively easy set-up, flexibility and modest hardware requirements, convolutional neural networks (CNNs) like YOLO (Redmon et al., 2015) are increasingly used to support the selection of images with target species (Knauer et al., 2022; Wäldchen & Mäder, 2018).

By combining a customized camera trap design with a CNN, we present a modular, automatic and cost-efficient camera system, whose construction does not require advanced technical skills and enables filming and evaluation of small and fast-moving insects. We are not the first to conceive a customized camera system for monitoring (see Appendix S1 for other examples) and built on previous

attempts to adapt Raspberry Pi for recording honeybees (Kulyukin & Reka, 2016). The key advantages of the new system are a broad applicability that is not restricted to single species, minimal interference with the natural behaviour of the recorded individual, high-quality videos facilitating the identification of individuals and affordability. For easy replication and modification of the system, we provide detailed step-by-step guides, assembly video and extensive material. With this, we aim to make this technology available to researchers with limited informatics or engineering background. We exemplarily built and tested a camera system research prototype under field conditions by filming the solitary bee *Osmia cornuta*, an important pollinator in fruit orchards (Marqu ez et al., 1994) flying in and out of a nesting aid. The customized CNN processes the video material in a CNN compatible format and filters target species out of scheduled videos.

2 | MATERIALS AND METHODS

2.1 | Assembling the camera system

We describe three components to construct a fully functioning camera system: power, hardware and software. A detailed documentation of all components including assembling guide is available (Appendix S2).

As maintenance should at most be weekly, a sufficient and exchangeable *power source* was necessary. We utilized 12V 60AH lead-acid accumulators (ABH-Nord GmbH, Flintbek, Germany) that provide energy for at least 2 weeks with the planned filming schedule (see below). A 12VDC to 5VDC converter with four USB ports (QSKY, Shenzhen, China) was used to provide the 5V voltage required by a Raspberry Pi computer (Raspberry Pi Foundation, Cambridge, UK). The converter was connected to the accumulator by soldering standard electric clips (JZK, Uhlirsk e Janovice, Czech Republic) to a piece of car jumper cable.

General requirements for *hardware* were easy customization without advanced technological knowledge and a low price. We chose the Raspberry Pi single board computer version 3B+ that provides four USB ports, WIFI and HDMI. The USB ports were used for computer mouse, keyboard and power access for a portable screen (7-inch portable monitor 1024×600p; UPERFECT, Shenzhen, China) during set-up and maintenance (checking accumulator status, camera system functionality and securing data). A Raspberry Pi camera module (v.2 with Sony IMX219 8-megapixel sensor, Raspberry Pi Foundation, Cambridge, UK) was connected with the computer via included ribbon cable. An energy-saving module, which included a real-time clock (WittyPi rev2, UUGear, Prague, Czech Republic), was mounted on top of the Raspberry Pi to finalize a *camera unit* comprising of Raspberry Pi, camera and energy-saving module. The camera unit was powered via a standard USB-A to USB-C cable (1.8m length; Anker, Shenzhen, China) that connected it to the converter. The recorded videos were saved on a 256GB USB stick (Ultra Line, Intenso, Vechta, Germany).

The software of the energy-saving module was programmed to regulate the power of the Raspberry Pi to save energy when not recording. A 'crontab', a UNIX-based command-line utility for Raspberry Pi (in operating system Raspbian GNU/Linux 10 (buster)), was created to execute a python script at specific times, in the test case at 10:00 and 15:00 CEST, the main flight times of the study subject *O. cornuta* (Vicens & Bosch, 2000). The executed script recorded a video of a nesting aid entrance for 1 h each. A timestamp was directly imprinted in every frame.

Cost of a camera system (status: February 2021) with two camera units was 325 € and 197 € with one camera unit (updated costs February 2024: 339 € for two units, 220 € for one unit). The sum includes all costs, including rain protection boxes, cables and electric clips (Table S2-1).

2.2 | Testing of a camera system research prototype

To test the functionality of the camera system under field conditions (Figure 1), 15 male and 15 female *O. cornuta* cocoons were placed each in a nesting aid (Staab et al., 2018). A customized mounting device was constructed and preassembled to record the bees flying in and out of the nesting aid. We used two camera units (Figure 1) that were stored in waterproof boxes (1.75 L; Vani, Hamburg, Germany) to record the nesting aid from the top perspective (view of bee tag) and from the side (view of carried pollen or clay). Field permits were issued by Regierungspräsidium Freiburg, Referat 55 (AZ 55-8852.15/00). The camera lens was placed in a hole drilled in the plastic box and secured by masking tape (protection of circuit board) and duct tape. A notice reading 'GPS tracked' was included in the box to discourage theft.

A pre-test showed that complex and moving background (e.g. vegetation during wind) in addition to varying light conditions led to larger video files and a higher difficulty to detect bees by human observers (object classification not tested on the corresponding videos). Thus, a wooden screen treated with transparent wood varnish was placed opposite of each camera to reduce background

movement and to increase contrast. The screen is an optional optimization not comprising the applicability and functioning of the research prototype. With sufficient training data and high data storage capacity, object classification can be likewise conducted without a background screen.

The accumulators were placed in a plastic box (Eurobox NextGen Portable, Müller & Son GmbH & Co. KG, Twistetal, Germany) and secured with a lock. A hole was drilled for the cable providing power, which was then sealed with plasticine. The accumulator box was further secured by a tent peg (10×280 mm; EisenRon.de, Delitzsch, Germany).

The assembled components were first tested indoors for 7 days, filming two 1-h videos daily. Next, the camera lens was adjusted to focus on the centre of the nesting aid (adjustment tool included in Raspberry Pi camera module) and the camera system was run outdoors in field-like conditions for seven additional days. Camera settings were compared by moving a dead bee in and out of the camera field of view. Best results were achieved each with the 'sports' mode on 1024×768 pixels and 60 fps. The camera settings were optimized to depict the colours of tags (top camera) and to depict carried pollen or building material (side camera) (Table S2-3).

All components were replicated and 21 additional units were assembled (20 field sites in south-west Germany + two spare units). As final field test, all camera systems recorded from 10 April until 14 May 2021 for 2 h daily (10:00–11:00, 15:00–16:00 CEST). The energy-saving tool powered the Raspberry Pi 5 min before the start and shut down 5 min after the recording. We recorded videos with H.264 compression. Maintenance and data collection were conducted weekly. During each maintenance interval, the nesting aid was also observed for 15 min and female *O. cornuta* were caught and tagged with a coloured and numbered plastic tag as a potential monitoring application for individual animals.

2.3 | Video post-processing and training the CNN

Recording videos within a schedule quickly produces material to an extent that is beyond evaluation capabilities of human observers.

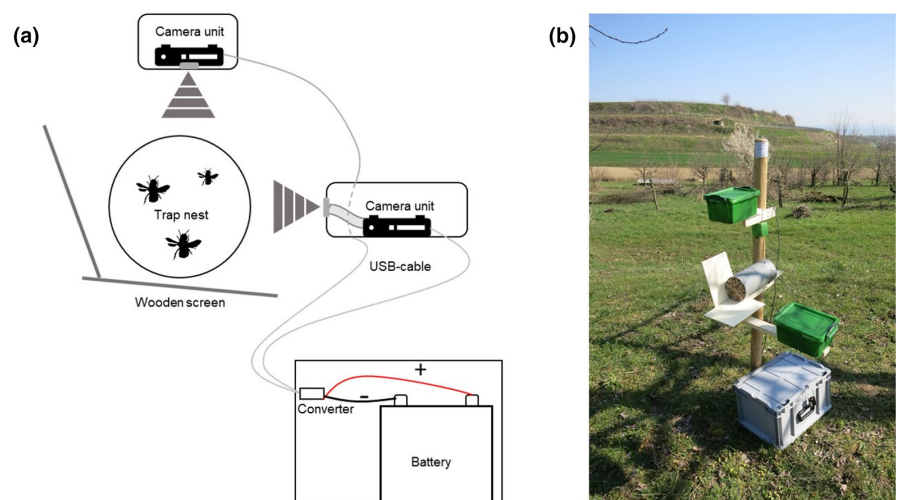


FIGURE 1 Illustration of (a) a customized camera system including a camera unit with Raspberry Pi computer, Raspberry Pi camera module and energy-saving module. As potential application, the filming of bees in (b) front of a nesting-aid entrance is shown.

Here, a CNN can be used to pre-filter sequences, in our test bees. A detailed installation guide and all necessary files can be found in Appendix S2 and the repository (Wittmann et al., 2024).

To identify representative sequences to train YOLOv5 (v.7.0, Jocher et al., 2022), videos were manually examined on 8× speed. A 'snapshot' tool (VLC player, v.3.0.12 Vetinari) was used to generate mp4 videos with relevant bee occurrences. Files were converted to jpg images by selecting eight frames per second with the command-line tool ffmpeg (v.5.0.1). We chose sequences from varying weather conditions, time windows, cameras and sites to maximize the learning efficiency of the object detection algorithm (following Ultralytics, 2022; Table S2-4).

All images were annotated for YOLOv5 with command-line software labellmg (v.1.8.5, Tzutalin, 2015), that is, a bounding box was drawn around the bee and classified as 'O. cornuta' (one bee=one instance). Images of partly hidden bees were included in the training dataset when the bee was recognizable. Only few instances of bees clearly carrying pollen or clay and/or wore coloured tags could be classified manually, which led to an unreliable detection. Thus, those classes were excluded. Additional annotated images from different cameras (O. cornuta filmed in front of a wooden nesting block) were added to diversify the training dataset (Table S2-4). Images showing similar bee instances were removed to avoid model overfitting (Yamashita et al., 2018). Image augmentation (automatic and random rotation, flipping and change of brightness; adapted from Paperspace, 2020) was conducted to further improve the training dataset (Appendix S2: [Optional] Image augmentation). We generated 2128 training images (775 original +1416 augmented) in total and added 63 background only pictures (~3% of total images; Table S2-4).

The compiled images were afterwards separated in training (70%), validation (20%) and test (10%) images for the automatic training, self-validation and testing feature of YOLOv5. The YOLOv5x6 network structure (following Ultralytics, 2022) was then trained with four NVIDIA RTX 2080 Ti GPUs (NVIDIA, Santa Clara, USA) and default hyperparameters (Table S2-5), which took around 4 h for 300 epochs and until detection did not significantly improve (optimum at 295 epochs).

2.4 | Modified CNN YOLOv5

YOLOv5 was selected due to its easy use, high accuracy and the ongoing support by the developer (Ultralytics, 2022). We adapted the YOLOv5 algorithm to take a video in mp4 or H.264 format and segment it into frames according to a predefined fps number. The resulting images are passed into the CNN that detects bees (or any other organism, depending on the training data). Results are compiled in a csv file containing image file name, X and Y coordinates of the detected target species in the image, confidence level of the CNN about the detected class, class number (in our case Class 0) and class name ('O. cornuta'). Also, a video containing the detected target species and a folder containing all images with bees including

annotation files is generated, which can be used to re-train YOLOv5. The modified CNN requires up to 4 h for a 1-h video with 60 frames (216,000 images) when processed on a NVIDIA RTX3070 GPU, depending on input format and settings.

2.5 | Manual evaluation of CNN

The final manual evaluation of YOLOv5 was conducted by one human observer for one randomly selected video per site and camera with one fps on novel videos ($N=40$; video examples in the repository, Wittmann et al., 2024) to verify the detection quality.

3 | RESULTS

Both camera units of the research prototype filmed automatically twice a day on 20 sites and did not require maintenance for at least a week. The camera system withstood snowfall, heavy rain, storm and fluctuating temperatures (ranging from -6 to 45°C) and the 40 deployed cameras recorded 2527 of the scheduled 2760 h (=92%). Average video size was 3.1 GB (SD=0.5, range=0.3–4.0). Camera failures include low accumulator capacity after cold nights (~94 h lost) or occasional malfunctioning of the energy-saving tool (~62 h). Other failures were due to damage by livestock, malfunctioning of USB-sticks and unknown reasons.

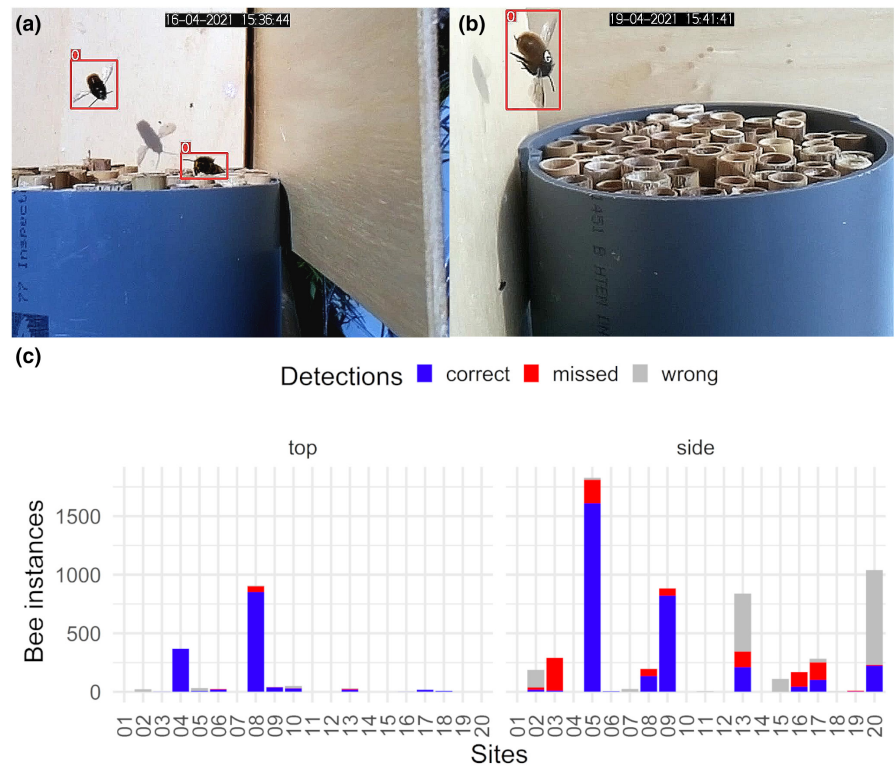
In the videos of the top perspective, the final CNN detected nearly all bees ($n=20$ sites; mean=94%; median=100%; SD=9%; range 69%–100%) with ≤1% misclassifications (Figure 2). The video recordings of the side perspective were less accurate ($n=20$ sites; mean=59%; median=61%; SD=32%; range=4%–100%), mainly due to varying light conditions and different angles a bee can be filmed from the side. Additionally, ~22% classifications were wrong in side videos, mostly due to shadows and other insects (Figure 2).

The average self-evaluation of YOLOv5 (mAP@0.5:0.95) on images of the training dataset (78% of bee instances correctly detected, with precision 0.98 and recall 0.97) was similar to human evaluation of the correctly detected bees on novel images (80%). The true positive ratio of YOLOv5 was 0.97 (to 0.03 false negative ratio) and the false positive ratio 1. Objectiveness loss was 0.003 and box loss 0.013.

4 | DISCUSSION

The camera system based on a single board computer is easy to build and is adaptable for many ecological questions including behavioural studies. It was tested quantitatively for a month by recording flying bees under field conditions. The camera system offers appropriate video quality to monitor fast-moving individuals with minimal disturbance of behaviour at an affordable price. To enable researchers with limited informatics and engineering background access to this technology, we provide detailed material to facilitate the replication,

FIGURE 2 Bee instances and their detection with YOLOv5 per site and per camera in one randomly selected video were manually validated by a human observer. The videos were filmed from a (a) top perspective ($n=20$ sites) and from a (b) side perspective ($n=20$ sites). (c) The randomly selected videos were validated by classifying correct detections of present bees (blue), present bees not being detected by YOLOv5 (red) and misdetections (grey; predominantly due to bee shadows, and other insects).



use and adaptation of our camera system and evaluation toolset. As such, we bridge the gap between technology and ecology at a level of detail that is not available yet. The manual, assembly video, source code and repository are freely available under a GPL-3.0 licence (Wittmann et al., 2024).

4.1 | Newly developed camera system

The research prototype performed well in the field with over 90% of all scheduled videos recorded, indicating that the principle of the introduced camera system can be successfully operated independent of a specific context. Using long-term storage accumulators, the system is reliably powered for more than 2 weeks, effectively filming 28 h with temperatures fluctuating below 0°C and above 40°C. In remote sites, power supply could be extended by using solar panels (depending on local conditions) and wireless networking to reduce physical maintenance. Wireless networking could, for instance, provide information about battery and storage status and newly created data remotely (Abas et al., 2018; Miller et al., 2015). The energy-saving tool minimized energy consumption and offered quality-of-life features for field conditions (e.g. power button, real-time clock). Other peripheral devices like sensors to monitor the environment could be additionally added to, for example, record environmental conditions including temperature and humidity but also monitor devices like wingbeat sound recorders for insects (Kim et al., 2021).

Maintaining high video quality is a challenge for all camera traps. Fluctuating temperatures can shift the field of view as the material used to place the camera expands or contracts. Occasionally, fogging

of the lens occurred, which could be mitigated by placing a protective window in front of the camera and adding silica gel in the container.

4.2 | Perspectives for data analysis

Apart from customized hardware components, the software for scheduled filming (repository, Wittmann et al., 2024) provides a more reliable solution than videos recorded with a passive infrared motion sensor. Consequently, ecological data would be less biased by, for example, camera types and site selection, which is important due to limited observation possibilities in field studies (Hofmeester et al., 2019).

Nevertheless, extracting information from collected videos is challenging, and manual evaluation is inefficient. Use of machine learning techniques, sometimes called artificial intelligence ('AI') is on the rise in ecology (Borowiec et al., 2022) and can minimize evaluation effort by automatically filtering videos. We integrated the CNN YOLOv5 in an evaluation toolset. By providing source code and a comprehensive guide (repository, Wittmann et al., 2024), the CNN can be trained for different object detection tasks across ecology. For instance, there are large image collections of many taxonomic groups available at the Global Biodiversity Information Facility (GBIF) or iNaturalist that can be annotated (Gilles & Hick, 2021) and used to improve the pool of training images. The performance of object detection, however, depends on various aspects including study subject, weather and camera placement. If a project requires information on all animals filmed, such as in general monitoring or in resource-choice experiments, the scheduled videos could be filtered by removing background-only pictures, which is especially feasible

when only few images are available to train the algorithm (Cunha et al., 2021). In this case, a background screen, as tested with our research prototype, can limit undesired background distractions and enhance detection and classification results. Also, behavioural studies can profit by the evaluation of the recorded scheduled videos in the field with minimal maintenance (e.g. free software anTraX; Gal et al., 2020).

We emphasize the benefits and synergies of interdisciplinary collaborations from scientists with engineering or computer-science background (Allan et al., 2018). We also encourage to integrate features of alternative camera systems with or without evaluation toolset (Appendix S1) into our system, as the field of video monitoring and object classification is rapidly changing and improved tools can be available already tomorrow (Pichler & Hartig, 2023).

AUTHOR CONTRIBUTIONS

Katharina Wittmann and Michael Staab conceived the idea. Mohamed Gamal Ibrahim modified YOLOv5 and programmed the evaluation toolset. Katharina Wittmann designed the camera system and annotated images, and wrote the manuscript with input from all authors that gave approval for publication.

ACKNOWLEDGEMENTS

We thank the Regierungspräsidium Freiburg, Referat 55, for field permits, and Felix Baab, Clemens Koch, all fieldwork helpers, students and landowners for their support. Florian Hartig and Maximilian Pichler provided GPU computing resources and support. Thomas Brox supervised Mohamed Gamal Ibrahim and provided GPU resources for training. Andrew David Straw was supported by the Volkswagen Foundation project 98692. Michael Staab is indebted to the Baden-Württemberg Foundation for support by the Eliteprogramm für Postdoktorandinnen und Postdoktoranden (grant number 1.16101.19). We acknowledge the support by the Open Access Publication Fund of the University of Freiburg and the Wissenschaftliche Gesellschaft Freiburg. We also thank three anonymous reviewers and the associate editor for valuable and critical feedback. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflict of interest.

PEER REVIEW

The peer review history for this article is available at <https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/2041-210X.14322>.

DATA AVAILABILITY STATEMENT

Manuals, code and dataset are available in figshare: <https://doi.org/10.6084/m9.figshare.22262986> (Wittmann et al., 2024). An assembly video can be found on YouTube: <https://youtu.be/YDqYuXQCYE8>.

ORCID

Katharina Wittmann  <https://orcid.org/0000-0003-3334-8196>
 Andrew David Straw  <https://orcid.org/0000-0001-8381-0858>
 Alexandra-Maria Klein  <https://orcid.org/0000-0003-2139-8575>
 Michael Staab  <https://orcid.org/0000-0003-0894-7576>

REFERENCES

- Abas, K., Obraczka, K., & Miller, L. (2018). Solar-powered, wireless smart camera network: An IoT solution for outdoor video monitoring. *Computer Communications*, 118, 217–233. <https://doi.org/10.1016/j.comcom.2018.01.007>
- Allan, B. M., Nimmo, D. G., Ierodiaconou, D., VanDerWal, J., Koh, L. P., & Ritchie, E. G. (2018). Futurecasting ecological research: The rise of technoecology. *Ecosphere*, 9(5), e02163. <https://doi.org/10.1002/ecs2.2163>
- Bonelli, M., Melotto, A., Minici, A., Eustacchio, E., Gianfranceschi, L., Gobbi, M., Casartelli, M., & Caccianiga, M. (2020). Manual sampling and video observations: An integrated approach to studying flower-visiting arthropods in high-mountain environments. *Insects*, 11(12), 881. <https://doi.org/10.3390/insects11120881>
- Borowiec, M. L., Dikow, R. B., Frandsen, P. B., McKeeken, A., Valentini, G., & White, A. E. (2022). Deep learning as a tool for ecology and evolution. *Methods in Ecology and Evolution*, 13, 1640–1660. <https://doi.org/10.1111/2041-210X.13901>
- Burton, A. C., Neilson, E., Moreira, D., Ladle, A., Steenweg, R., Fisher, J. T., Bayne, E., & Boutin, S. (2015). Wildlife camera trapping: A review and recommendations for linking surveys to ecological processes. *Journal of Applied Ecology*, 52(3), 675–685. <https://doi.org/10.1111/1365-2664.12432>
- Caravaggi, A., Banks, P. B., Burton, C., Finlay, C. M., Haswell, P. M., Hayward, M. W., Rowcliffe, M. J., & Wood, M. D. (2017). A review of camera trapping for conservation behaviour research. *Remote Sensing in Ecology and Conservation*, 3(3), 109–122. <https://doi.org/10.1002/rse2.48>
- Cardim Ferreira Lima, M., de Almeida Leandro, M. E., Valero, C., Pereira Coronel, L. C., & Gonçalves Bazzo, C. O. (2020). Automatic detection and monitoring of insect pests—A review. *Agriculture*, 10(5), 161. <https://doi.org/10.3390/agriculture10050161>
- Cunha, F., dos Santos, E. M., Barreto, R., & Colonna, J. G. (2021). Filtering empty camera trap images in embedded systems. 2015 *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Nashville, TN, USA. <https://doi.org/10.1109/CVPRW.53098.2021.00276>
- Delisle, Z. J., Flaherty, E. A., Nobbe, M. R., Wzientek, C. M., & Swihart, R. K. (2021). Next-generation camera trapping: Systematic review of historic trends suggests keys to expanded research applications in ecology and conservation. *Frontiers in Ecology and Evolution*, 9, 617996. <https://doi.org/10.3389/fevo.2021.617996>
- Droissart, V., Azandi, L., Onguene, E. R., Savignac, M., Smith, T. B., & Deblauwe, V. (2021). PICT: A low-cost, modular, open-source camera trap system to study plant–insect interactions. *Methods in Ecology and Evolution*, 12(8), 1389–1396. <https://doi.org/10.1111/2041-210X.13618>
- Gal, A., Saragosti, J., & Kronauer, D. J. (2020). anTraX, a software package for high-throughput video tracking of color-tagged insects. *eLife*, 9, e58145. <https://doi.org/10.7554/eLife.58145>
- Gilles, C., & Hick, S. (2021). *Automated nocturnal insect monitoring in real-time using YOLOv4*. <http://hdl.handle.net/2078.1/thesis:30560>
- Halfacre, G. (2018). *The official Raspberry Pi Beginner's Guide: How to use your new computer*. Raspberry Pi Press.
- Hofmeester, T. R., Crowsigt, J. P., Odden, J., Andrén, H., Kindberg, J., & Linnell, J. D. (2019). Framing pictures: A conceptual framework to identify and correct for biases in detection probability of camera

- traps enabling multi-species comparison. *Ecology and Evolution*, 9(4), 2320–2336. <https://doi.org/10.1002/ece3.4878>
- Jocher, G., Chaurasia, A., Stoken, A., Borovec, J., NanoCode012, Kwon, Y., Michael, K., TaoXie, Fang, J., imyhxy, Lorna, Yifu, Z., Wong, C., Abhiram, V., Montes, M., Wang, Z., Fati, C., Nadar, J., Laughing, ... Jain, M. (2022). ultralytics/yolov5: v7.0—YOLOv5 SOTA realtime instance segmentation. *Zenodo*. <https://doi.org/10.5281/zenodo.7347926>
- Kim, D., DeBriere, T. J., Cherukumalli, S., White, G. S., & Burkett-Cadena, N. D. (2021). Infrared light sensors permit rapid recording of wing-beat frequency and bioacoustic species identification of mosquitoes. *Scientific Reports*, 11, 10042. <https://doi.org/10.1038/s41598-021-89644-z>
- Knauer, A. C., Gallmann, J., & Albrecht, M. (2022). Bee Tracker—An open-source machine learning-based video analysis software for the assessment of nesting and foraging performance of cavity-nesting solitary bees. *Ecology and Evolution*, 12(3), e8575. <https://doi.org/10.1002/ece3.8575>
- Krauss, S. L., Roberts, D. G., Phillips, R. D., & Edwards, C. (2018). Effectiveness of camera traps for quantifying daytime and nighttime visitation by vertebrate pollinators. *Ecology and Evolution*, 8(18), 9304–9314. <https://doi.org/10.1002/ece3.4438>
- Kulyukin, V., & Mukherjee, S. (2019). On video analysis of omnidirectional bee traffic: Counting bee motions with motion detection and image classification. *Applied Sciences*, 9(18), 3743. <https://doi.org/10.3390/app9183743>
- Kulyukin, V., & Reka, S. K. (2016). Toward sustainable electronic beehive monitoring: Algorithms for omnidirectional bee counting from images and harmonic analysis of buzzing signals. *Engineering Letters*, 24(3), 317–327.
- Marquéz, J., Bosch, J., & Vicens, N. (1994). Pollens collected by wild and managed populations of the potential orchard pollinator *Osmia cornuta* (Latr.) (Hym., Megachilidae). *Journal of Applied Ecology*, 117(1–5), 353–359. <https://doi.org/10.1111/j.1439-0418.1994.tb00746.x>
- Meek, P. D., Ballard, G.-A., Fleming, P. J., Schaefer, M., Williams, W., & Falzon, G. (2014). Camera traps can be heard and seen by animals. *PLoS One*, 9(10), e110832. <https://doi.org/10.1371/journal.pone.0110832>
- Miller, L., Abas, K., & Obraczka, K. (2015). SCmesh: Solar-powered wireless smart camera mesh network. In L. Miller, K. Abas, & K. Obraczka (Chairs), *2015 24th International Conference on Computer Communication and Networks (ICCCN)*, Las Vegas, USA.
- Monk, S. (2020). *Raspberry Pi Cookbook: Software and hardware problems & solutions* (3rd ed.). O'Reilly.
- Moore, J. F., Soanes, K., Balbuena, D., Beirne, C., Bowler, M., Carrasco-Rueda, F., Cheyne, S. M., Coutant, O., Forget, P.-M., Haysom, J. K., Houlihan, P. R., Olson, E. R., Lindshield, S., Martin, J., Tobler, M., Whitworth, A., & Gregory, T. (2021). The potential and practice of arboreal camera trapping. *Methods in Ecology and Evolution*, 12(10), 1768–1779. <https://doi.org/10.1111/2041-210X.13666>
- Naqvi, Q., Wolff, P. J., Molano-Flores, B., & Sperry, J. H. (2022). Camera traps are an effective tool for monitoring insect–plant interactions. *Ecology and Evolution*, 12(6), e8962. <https://doi.org/10.1002/ece3.8962>
- Paperspace. (2020). *Data augmentation for object detection*. <https://github.com/Paperspace/DataAugmentationForObjectDetection?ref=blog.paperspace.com>
- Pichler, M., & Hartig, F. (2023). Machine learning and deep learning—A review for ecologists. *Methods in Ecology and Evolution*, 14(4), 994–1016. <https://doi.org/10.1111/2041-210X.14061>
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2015). *You only look once: Unified, real-time object detection*. <http://arxiv.org/pdf/1506.02640v5>
- Staab, M., Pufal, G., Tschardtke, T., & Klein, A.-M. (2018). Trap nests for bees and wasps to analyse trophic interactions in changing environments—A systematic overview and user guide. *Methods in Ecology and Evolution*, 9(11), 2226–2239. <https://doi.org/10.1111/2041-210X.13070>
- Tobler, M. W., Carrillo-Percestequi, S. E., Leite Pitman, R., Mares, R., & Powell, G. (2008). An evaluation of camera traps for inventorying large- and medium-sized terrestrial rainforest mammals. *Animal Conservation*, 11(3), 169–178. <https://doi.org/10.1111/j.1469-1795.2008.00169.x>
- Tzutalin, D. (2015). *LabelImg*. <https://github.com/tzutalin/labelImg>
- Ultralytics. (2022). *YOLOv5 documentation*. <https://docs.ultralytics.com/tutorials/training-tips-best-results/#dataset>
- Vicens, N., & Bosch, J. (2000). Weather-dependent pollinator activity in an apple orchard, with special reference to *Osmia cornuta* and *Apis mellifera* (Hymenoptera: Megachilidae and Apidae). *Environmental Entomology*, 29(3), 413–420. <https://doi.org/10.1603/0046-225X-29.3.413>
- Wäldchen, J., & Mäder, P. (2018). Machine learning for image based species identification. *Methods in Ecology and Evolution*, 9(11), 2216–2225. <https://doi.org/10.1111/2041-210X.13075>
- Welbourne, D. J., Claridge, A. W., Paull, D. J., & Lambert, A. (2016). How do passive infrared triggered camera traps operate and why does it matter? Breaking down common misconceptions. *Remote Sensing in Ecology and Conservation*, 2(2), 77–83. <https://doi.org/10.1002/rse2.20>
- Wittmann, K., Ibrahim, M. G., Straw, A. D., Klein, A.-M., & Staab, M. (2024). Repository to monitoring fast-moving animals—Building a customized camera system and evaluation toolset. *Figshare*. <https://doi.org/10.6084/m9.figshare.22262986>
- Yamashita, R., Nishio, M., Do, R. K., & Togashi, K. (2018). Convolutional neural networks: An overview and application in radiology. *Insights into Imaging*, 9(4), 611–629. <https://doi.org/10.1007/s13244-018-0639-9>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Appendix S1: Examples of recently developed camera systems and evaluation tools.

Appendix S2: Step-by-step guide and additional details.

How to cite this article: Wittmann, K., Gamal Ibrahim, M., Straw, A. D., Klein, A.-M., & Staab, M. (2024). Monitoring fast-moving animals—Building a customized camera system and evaluation toolset. *Methods in Ecology and Evolution*, 00, 1–7. <https://doi.org/10.1111/2041-210X.14322>