Eyes on nature: Embedded vision cameras for terrestrial biodiversity monitoring

Kevin F A Darras, Marcel Balle, Wenxiu Xu, Yang Yan, Vincent G Zakka, Manuel Toledo-hernández, Dong Sheng, Wei Lin, Boyu Zhang, Zhenzhong Lan, et al.

Received: 19 January 2024 | Accepted: 4 September 2024

DOI: 10.1111/2041-210X.14436

Methods in Ecology and Evolution

RESEARCH ARTICLE

Eyes on nature: Embedded vision cameras for terrestrial biodiversity monitoring

Kevin F. A. Darras[1,2,3] | Marcel Balle[1] | Wenxiu Xu[1] | Yang Yan[1,4,5] | Vincent G. Zakka[1] | Manuel Toledo-Hernández[1,2,6] | Dong Sheng[1,7,8] | Wei Lin[1] | Boyu Zhang[9,10] | Zhenzhong Lan[5] | Li Fupeng[11] | Thomas C. Wanger[1,2,6,12,13]

[1]Sustainable Agricultural Systems & Engineering Laboratory, School of Engineering, Westlake University, Hangzhou, China; [2]Key Laboratory of Coastal Environment and Resources of Zhejiang Province, Westlake University, Hangzhou, China; [3]EFNO, INRAE, Nogent-sur-Vernisson, France; [4]College of Computer Science and Technology, Zhejiang University, Hangzhou, China; [5]Deep Learning Laboratory, School of Engineering, Westlake University, Hangzhou, China; [6]GlobalAgroforestryNetwork.com, Hangzhou, China; [7]College of Resource & Environmental Sciences, Zhejiang University, Hangzhou, China; [8]Engineering and Research Center for Industries of the Future, Westlake University, Hangzhou, China; [9]Key Laboratory of 3D Micro/Nano Fabrication and Characterization of Zhejiang Province, School of Engineering, Westlake University, Hangzhou, China; [10]School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China; [11]Spice and Beverage Research Institute, Chinese Academy of Tropical Agricultural Sciences, Wanning, Hainan, China; [12]ChinaRiceNetwork.org, Hangzhou, China and [13]Agroecology, University of Göttingen, Göttingen, Germany

Correspondence
Kevin F. A. Darras
Email: kevin.darras@inrae.fr

Li Fupeng
Email: peterlfp@163.com

Thomas C. Wanger
Email: tomcwanger@gmail.com

Funding information
Westlake University Startup Fund

Handling Editor: Xingfeng Si

Abstract

1. We need comprehensive information to manage and protect biodiversity in the face of global environmental challenges, and artificial intelligence is required to generate that information from vast amounts of biodiversity data. Currently, vision-based monitoring methods are heterogenous; they poorly cover spatial and temporal dimensions, overly depend on humans, and are not reactive enough for adaptive management.
2. To mitigate these issues, we present a portable, modular, affordable and low-power device with embedded vision for biodiversity monitoring of a wide range of terrestrial taxa. Our camera uses interchangeable lenses to resolve barely visible and remote targets, as well as customisable algorithms for blob detection, region-of-interest classification and object detection to automatically identify them. We showcase our system in six use cases from ethology, landscape ecology, agronomy, pollination ecology, conservation biology and phenology disciplines.
3. Using the same devices with different setups, we discovered bats feeding on durian tree flowers, monitored flying bats and their insect prey, identified nocturnal insect pests in paddy fields, detected bees visiting rapeseed crop flowers, triggered real-time alerts for waterfowl and tracked flower phenology over months. We measured classification accuracies (i.e. *F*1-scores) between 55% and 95% in our field surveys and used them to standardise observations over highly resolved time scales.

4. Our cameras are amenable to situations where automated vision-based monitoring is required off the grid, in natural and agricultural ecosystems, and in particular for quantifying species interactions. Embedded vision devices such as this will help addressing global biodiversity challenges and facilitate a technology-aided agricultural systems transformation.

KEYWORDS
biodiversity monitoring, camera trap, edge computing, embedded vision

1 | INTRODUCTION

This century's rapid biosphere transformations exert formidable pressure on nature and people (Steffen et al., 2015). High-resolution biodiversity data are required to inform conservation and management plans (Dove et al., 2023; Moersberger et al., 2024; UNECE, 2023; UNEP, 2020) and we are heavily reliant on artificial intelligence to analyse them and tackle Sustainable Development Goals (Beery, 2021; Klein et al., 2015; Vinuesa et al., 2020). However, comprehensive hardware and software solutions for generating and analysing high-resolution, standardised biodiversity monitoring data are still rare (Besson et al., 2022).

Much of biodiversity is sampled by vision, and we are facing critical challenges to attain comprehensive monitoring. First, heterogenous, incompatible methods are used for a variety of terrestrial taxa: point counts for birds (Ralph et al., 1995), visual encounter surveys for herpetofauna (Doan, 2003), transects for pollinators (O'Connor et al., 2019), camera traps for mammals (McCallum, 2013), etc. Second, sampling coverage and resolution along spatial and temporal scales are insufficient, resulting in considerable knowledge biases and gaps (Boakes et al., 2010; Costello et al., 2010; Meyer et al., 2015) that prevent effective conservation: notably, most of the challenging animals to monitor (amphibians, insects, mammals and reptiles) are data-deficient and likely threatened (Borgelt et al., 2022). Third, many methods still depend on human labour, and are thus time-consuming, error-prone, and hard to reproduce (Fitzpatrick et al., 2009; Gorrod & Keith, 2009; Johnston et al., 2023), a problem further compounded by the scarcity of taxonomic experts (Engel et al., 2021). Fourth and lastly, most passive vision-based methods lack real-time feedback (Whytock et al., 2023), thus ruling out immediate interventions despite increased uptake of adaptive management approaches (Jackson et al., 2010; Keith et al., 2011). Theoretically, these challenges may be solved by deploying continuously powered digital imaging devices "at the edge" for sampling multiple taxa across large scales, with embedded artificial intelligence processing data in real-time to minimise storage demands and trigger meaningful reactions when needed. While the underlying technologies exist and open hardware abounds (Oellermann et al., 2022), devices are still in early development stages and we lack integrated solutions (Høye et al., 2021; van Klink et al., 2022).

We harnessed technological advances into a field-ready, portable, low-power, modular, embedded vision camera system—dubbed "ecoEye"—thus taking advantage of recent progress in embedded computing (Dutta & Bharali, 2021) and also moving towards goals set by Besson et al. (2022). Our system can: (1) non-invasively monitor various taxa for different applications across disciplines; (2) reach high temporal resolution and coverage with solar power and be scaled up in space due to its moderate cost and size; (3) analyse images in real-time using established computer vision algorithms and performance assessment workflows that standardise observation results and thus (4) potentially link specific detections to real-time reactions for intervening in environmental processes. We finally discuss how our and other embedded vision devices fare for addressing key vision-based biodiversity monitoring challenges.

2 | MATERIALS AND METHODS

We designed a portable, weather-resistant embedded vision camera (Figure 1) consisting of (1) a low-power, expandable microcontroller printed-circuit board (i.e. PCB) with a modular image sensor (openMV H7+); (2) a custom data transfer and power management PCB (Balle et al., in press); and (3) a custom waterproof housing for the PCBs, batteries, lens mount and accessories. This advanced device development resulted in a commercial device (EcoEye–Embedded Vision Camera for Environmental Monitoring, n.d.) that can only be replicated with considerable manufacturing (e.g. injection moulding) and electronics know-how and resources (e.g. reflow soldering and PCB assembly) (Fig. 1a,b,c). However, the hardware used for the subsequently presented use cases A to E, which do not require the power management PCB developed separately, can be replicated using our instructions for building the latest open-design camera version with a 3D-printed case (Darras, n.d.; Figure 1d). We also provide an open-design CAD model (Figure 1d) to 3D-print a lens mount enabling the construction of a low-tech version using only an openMV H7+, powered with a USB power bank, inside a waterproof box. This lens mount allows placing the lens outside the box for optimal optical image quality while protecting the electronics inside the box, provided that the lens mount is sealed with silicone glue to the box and that the lens thread is sealed against it.

We evaluated six use cases representing different disciplines in temperate and tropical regions of China (Table 1, Figure 2, Figure S1). Additional licences and permits were not needed as we conducted our work either on experimental plots, within an associated

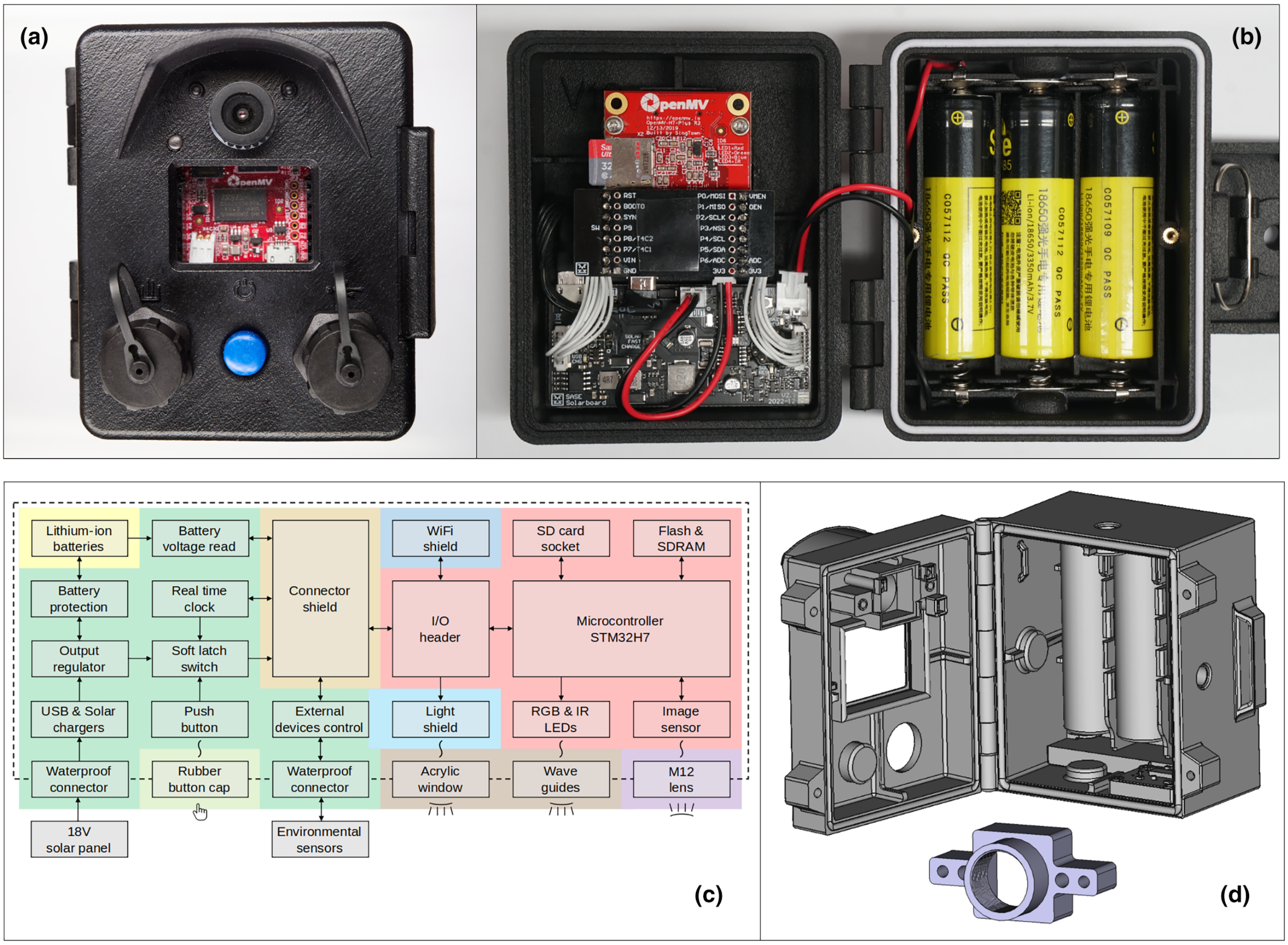

FIGURE 1 The ecoEye camera design (commercial and open-design versions). (a) Front view of the commercial version: Interchangeable lens, window for the internal light shield (not pictured), built-in LED light pipes, power button, external sensor and power connectors. (b) Inner view of the commercial version: Main OpenMV and auxiliary boards (power management system at bottom, connector shield stacked on openMV board), lithium 18,650 batteries. (c) Block diagram of the commercial version's system design, including further extensions not pictured in the previous panels, such as the WLAN and light shields. (d) The open-design CAD case model (version 49) used in use cases A to E. Open-design CAD standalone lens mount model shown at bottom, provided for DIY applications allowing the placement of the interchangeable lenses outside off-the-shelf boxes in a waterproof manner.

botanical garden, on our university campus, or on publicly accessible paths with no people in the field of view. Rather than conducting separate, fully-fledged studies, we chose to explore the broad potential of the embedded vision camera in different contexts to stimulate follow-up, in-depth studies. The same cameras were used but set up with different lenses, lighting equipment and operation settings for each use case; details are given in the Supplementary Methods. Both frame differencing-based blob detection (also called motion or region-of-interest/ROI detection) as well as deep learning convolutional neural networks (CNNs hereafter)—trained on the EdgeImpulse platform (https://www.edgeimpulse.com/) with MobileNet-v2 (Sandler et al., 2019) using images captured during test deployments—were used and combined to detect and identify targets in real-time on the cameras during survey deployments. Detection could be handled by blob detection or object detection CNNs, and classification could be handled by (blob detection-derived) ROI classification or by the latter object detection CNNs.

Although whole-image classification can also be done in camera, we did not find situations where the target would fill the entire frame. The recall and precision of the CNN-based ROI classification and object detection algorithms were evaluated with test images from the field survey deployments. For use cases involving object detection, all images were saved during the survey deployments and a subset was subsequently screened for this evaluation. For use cases involving blob detection for ROI classification, we used conservative blob detection thresholds (colour deviances were smaller than those observed in barely perceptible targets; minimum and maximum blob sizes were set beyond the observed extremes of targets' sizes) and all blobs saved during the evaluation survey interval were subsequently screened. These two approaches allowed us to obtain image series of ground-truth data reflecting the exact field of view of the cameras. The harmonic mean of the recall and precision measures obtained in the screened image datasets were reported as field accuracies (i.e. *F*1-scores; Table 1).

TABLE 1 Comparison of use case setups for the ecoEye embedded vision cameras. The main parameters, performance metrics and setup variables are listed. Optimal recall and precision correspond to the values reached at maximal field accuracy (i.e. *F*1 score).

Use case	Target	Target distance (m)	Daytime	Resolution	Colour	Computer vision	Frames per second
(A) Ethology	Bats on durian flowers	3–5	Night	2592 × 1944	Grayscale	Blob detection	0.67
(B) Landscape ecology	Flying bats and insects	3–7	Night	1944 × 1944	Grayscale	ROI classification	1.01
(C) Agronomy	Insects on light board	0.2	Night	2592 × 1944	RGB565	ROI classification	0.41
(D) Pollination ecology	Bees on rapeseed	0.4	Day	1944 × 1944	RGB565	Object detection	1.37
(E) Conservation biology	Waterfowl on water	3–5	Day	1944 × 1944	RGB565	Object detection	1.75
(F) Phenology	Flowers on ground	0.6	Day	1944 × 1944	RGB565	Object detection	0.001 (timelapse)

3 | RESULTS

3.1 | Ethology: Monitoring bats visiting durian tree flowers

Flying foxes provide pollination services for many tropical trees but little is known about their foraging behaviour, and current vision-based monitoring approaches are a compromise between automation and resolution (Darras, Yusti, Knorr, et al., 2021; Gottwald et al., 2021). In Southeast Asia, the durian tree produces the "king of fruits", a crop whose most common pollinators are fruit bats (Baqi et al., 2022). On the island of Hainan, however, flowers are usually pollinated by hand and although fruit bats inhabit the island, they have not been reported to pollinate durian there to our knowledge. We monitored the nocturnal behaviour of unknown durian flower visitors in Hainan, China (Figure 2a), using near-infrared illumination and blob detection algorithms, from the ground.

We monitored the only flowering tree of our study location for 12 nights in summer 2022 at an average of 0.68 (min: 0.56; max: 0.89) frames per second over all nights with blob detection to detect image changes within pre-set ranges of blob area and colour deviation from the background image (Figure 2a). We found the flying fox and manually screened 23,154 triggered images to determine optimal blob detection parameters at each deployment date: we computed an average maximum accuracy of 0.21 over deployment nights (range: 0.10–0.34), corresponding to the best performance that would have been obtained with optimal settings using a minimum blob area of 125,000 pixels and a maximum blob area of 631,000 pixels (Figure S2). The flying fox *Roussettus leschenaultii* (most likely species identification) was found in 122 images, flying towards or from flowers in 63 and feeding on flowers in 59 images. Flowers were visited 9 times per night on average (range: 3–32 visits, $n = 7$; flowers open only for one night). Two activity peaks were evident: early in the night (~21:00) and shortly before dawn twilight (~3:00) (Figure 2a).

3.2 | Landscape ecology: Determining bat and insect occurrence in rice fields

Landscape composition determines the occurrence of predators and their prey. Insectivorous bats, which are sensitive to forest edges and waterways (Lintott et al., 2015; Morris et al., 2010), regulate insect populations in natural and agricultural landscapes through predation, thus providing economically valuable biocontrol services, also for rice (Puig-Montserrat et al., 2015; Wanger et al., 2014). Current passive sampling methods use digital imagery without embedded artificial intelligence to sample either only flying insects or only bats (Darras, Yusti, Huang, et al., 2021; Gottwald et al., 2021; Ruczyński et al., 2020), thereby increasing the logistical complexity of sampling both taxa and limiting the spatial and temporal link between their sampling data. We used our cameras with region-of-interest classification and near-infrared illumination to simultaneously and remotely detect and distinguish both of these differently sized nocturnal targets in flight.

We monitored flying insects and bats in a mixed rice paddy and forest landscape for 6 nights in summer 2022 during astronomical dusk (between 19:20 and 20:30), in Hangzhou, China (Figure 2b). Each night, we set up one camera inside the rice field and one at its border, adjacent to a waterway with a forest on its other shore, and pointed them to the night sky, illuminated with attached near-infrared flashlights (Figure S3). The cameras were operating with a sensitive blob detection threshold at an average of 1.32 (range: 0.37–1.88) frames per second. Flying objects were detected as blobs up to an estimated 7 m above ground, and images extracted from their bounding boxes (i.e. ROIs) were classified into bats, insects, or unknown flying objects with our trained CNN. We obtained a maximum field classification accuracy (*F*-score) of 67% for bats and 93% for insects at their respective confidence thresholds of 0.6 and 0.3. After considering only detections above these thresholds, we obtained a nightly

Training images	Optimal recall	Optimal precision	Lens focal length (mm)	Lens aperture	Lighting
NA (no classification)	0.16	0.36	12	F2	2 IR flashlights
376	Bat: 0.75, insect: 0.91	Bat: 0.60, insect: 0.95	6	F2.2	4 IR flashlights
735	Chironomidae: 0.53, Coleoptera: 0.87, Curculionidae: 0.67, Delphacidae1: 0.87, Delphacidae2: 0.42, Hymenoptera1: 0.53, Hymenoptera2: 0.46	Chironomidae: 0.70, Coleoptera: 1, Curculionidae: 0.04, Delphacidae1: 0.77, Delphacidae2: 0.80, Hymenoptera1: 0.88, Hymenoptera2: 0.35	4.23	F2.8	Inner add-on white LED shield
278	0.77	0.63	12	F2	Ambient
425	Female: 0.92, male: 0.95	Female: 0.63, male: 0.97	12	F2	Ambient
118	0.99	0.89	4.23	F2.8	Ambient

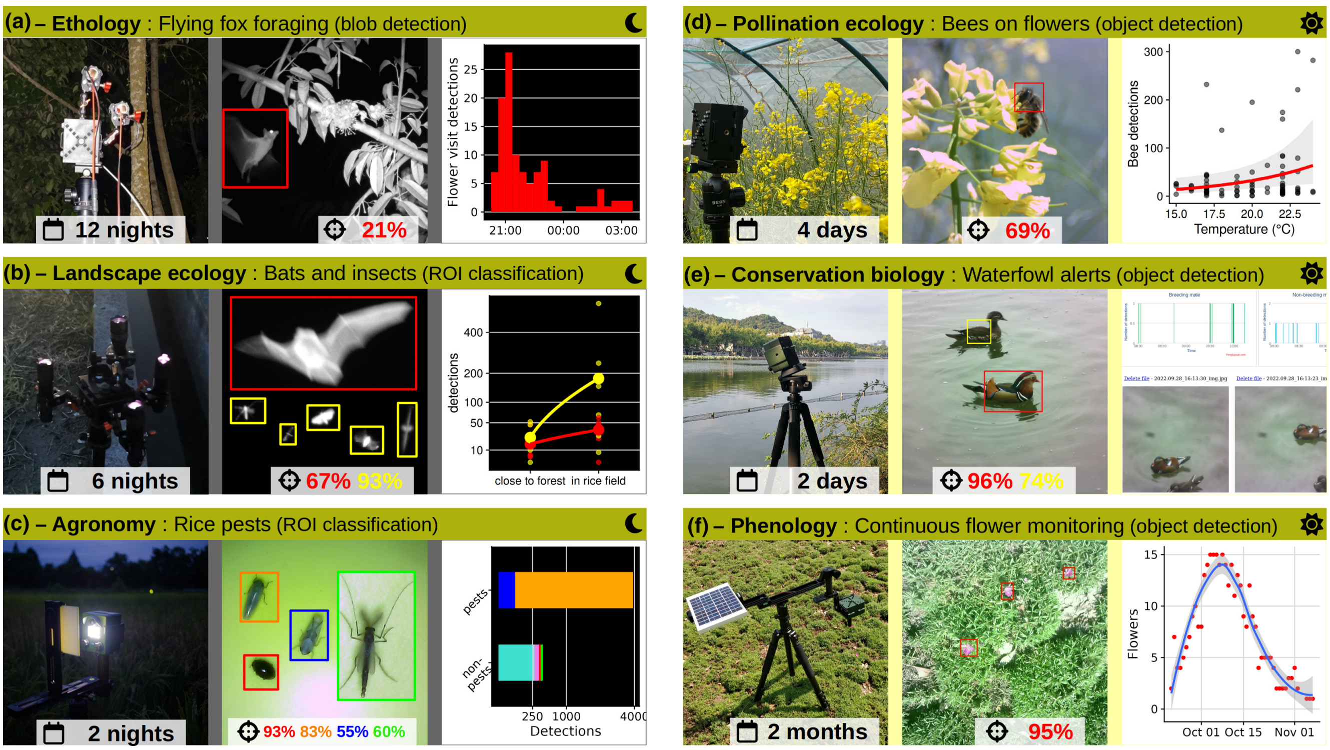

FIGURE 2 Six use cases depicting field applications of the ecoEye embedded vision camera. Middle panels for (b and c) were made from multiple pictures to depict the relative sizes of the diverse targets. Padding was added to the bounding boxes of the detections for better visibility of the targets. In left panels, calendar icons indicate deployment durations, and in middle panels, crosshairs represent field accuracies, measured with *F*1 scores. Accuracy values in (e) are based on the first deployment only. In the right panel of (d), the red line indicates the predicted temperature effect on bee detections, and the shaded grey ribbon depicts the upper and lower bounds of its 95% confidence interval. In the rightmost pane of (f), the blue line represents the smoothed relationship between time and flower counts (loess smoother, evaluated at 80 points with a span of 0.75), with its 95% confidence interval depicted as a grey ribbon.

average of 16 bat and 25 insect detections close to the forest compared to 37 bat and 180 insect detections inside the rice field, with a significantly lower bat to insect detections ratio inside the rice field than close to the forest (Estimate $= -0.34 \pm 0.17$ SD, $n = 12$, Probability$_{(\text{Estimate}=0 \mid \text{Null hypothesis})} = 0.04$).

3.3 | Agronomy: Quantifying nocturnal insect rice pests

Monitoring pests is essential for agricultural management. Light traps are a common sampling method for flying insect pests in rice paddy fields (Shimoda & Honda, 2013), a staple food throughout Asian countries that is damaged by the brown and white-backed planthoppers. Existing monitoring devices are rather large, comparatively expensive, sometimes lethal, and most have no embedded classification (Bjerge, Nielsen, et al., 2021; Yao et al., 2020). We used our cameras to demonstrate an embedded multiple classification task of nocturnal insects attracted to non-lethal, portable light traps, with real-time cloud transmission of results.

We monitored nocturnal, flying, phototactic insects attracted to white-light illuminated plastic boards. Cameras were installed in two locations inside a rice field in Hangzhou, China (Figure 2c) and worked for 2h on two nights in 2022 and 2023, starting at nautical dusk, around 17:30. Cameras operated with a sensitive blob detection threshold at an average of 0.41 (min: 0.40; max: 0.41) frames per second. Detection numbers for each detected target were sent online to a cloud server, via an internal plug-in WLAN module connected to a 4G portable router (Figure S4), with a transfer success rate between 76% and 83% (100% were saved on the SD card). Landing objects were blob-detected and their bounding boxes extracted for classification into nine classes with variable taxonomic resolution (order or family) depending on the distinctiveness of their diagnostic features. For the different morphospecies detected during survey deployments, we obtained maximum field accuracies ranging from 8% (Curculionidae—only 39 training images) to 93% (Coleoptera) for non-pest species (mean: 53%, median: 60%), and from 55% to 83% for potential planthopper pests (two Delphacidae morphospecies). After retaining only detections above the respective optimal confidence thresholds, we found 3856 detections for the Delphacidae 1 and 59 detections for the Delphacidae 2 morphospecies, putatively identified as the brown (*Nilaparvata lugens*) and white-backed (*Sogatella furcifera*) planthoppers, respectively, and 452 non-pest detections across both survey nights.

3.4 | Pollination ecology: Monitoring bees on rapeseed flowers

Global insect pollinator declines are disrupting pollination networks and crop production (Klein et al., 2006). Monitoring pollinators is essential to reap their benefits (Potts et al., 2016). Long-term diel monitoring of bees was previously conducted with motion-detection devices and off-the-shelf hardware, albeit without automated target classification (Steen, 2017). We used our cameras to monitor rapeseed, a major oil crop, with object detection models specifically trained for identifying pollination events by bees.

We monitored solitary bees (*Osmia bicornis*) on rapeseed flowers growing inside experimental enclosures with six camera deployments over 4 days in April 2022 in Fuyang district, China (Figure 2d). Cameras ran at an average of 1.37 (range: 1.13–1.91) frames per second and targets were detected in each frame by our object detection CNN at a confidence threshold of 0.5. We obtained an average field classification accuracy of 74% (range: 57%–93%) over all deployments (Figure S5). As field accuracy varied among deployments, we standardised the detections data to estimate the count of actual bee visits, which we used to derive temporal bee activity profiles that were comparable across deployments (Figure S5). We obtained the estimated visits count by multiplying the number of detections with the precision and dividing by the recall. We found a positive relationship between estimated bee visits per hour and ambient temperature (Estimate $=0.169 \pm 0.08$ SD, $n=84$, Probability$_{(\text{Estimate}=0 \mid \text{Null hypothesis})}=0.037$), in line with previous studies (Steen, 2017).

3.5 | Conservation biology: Real-time waterfowl monitoring

Adaptive species management can help to deal with unpredictable global changes that precipitate biodiversity loss (Jackson et al., 2010). Real-time passive monitoring is required for rarely encountered animals, whose population changes could trigger immediate management actions. Birds, as fast-moving, small targets are especially challenging to monitor automatically even with current vision-based technology (Latta et al., 2023; Weinstein et al., 2022). Here, we show how cameras can be networked to send real-time cloud alerts upon detection of a specific waterfowl species in an urban protected area.

We monitored swimming mandarin ducks (*Aix galericulata*) using cameras in a nature reserve, from the shore of the Xihu lake in Hangzhou city, China, over 2 days in 2022 and 2024 (Figure 2e). The cameras analysed all images with an object detection CNN at a conservative confidence threshold of 0.25. During the first deployment in 2022, a single camera ran at 1.76 frames per second. It transferred detection classes and probabilities, as well as images containing mandarin duck detections exceeding a confidence threshold of 0.5, via an internal plug-in WLAN module connected to a 4G portable router, to a cloud server where the detection time graphs and downsampled images could be checked in real-time (Figure 2e). 82% of the detections data and 80% of the images were successfully transferred from the camera to the server (100% were saved on the SD card). For the first deployment, we obtained maximum field accuracies of 74% for females, and 96% for breeding males at their optimal confidence thresholds, which allowed to filter the dataset containing 594 raw detections to yield 134 female and 405 male estimated bird passes after correcting for male and female detection accuracies (Figure S6). During the second deployment in 2024, carried out after the breeding season, males lost their colourful plumage and were not discernible from females anymore; the model classified juvenile ducks (that were not used for training the model) as male breeding ducks, thus decreasing breeding male detection accuracy (42%), while overall duck detection accuracy (age classes and sexes confounded) stayed high (90%). The male mandarin duck is an emblematic species with characteristic plumage that is easily discerned from other species—further trials should attempt to distinguish mandarin ducks of different ages and other bird species with more subtle morphological differences.

3.6 | Phenology: Flowering plants

Phenology cameras monitor vegetation phenology with high temporal resolution over long sampling durations, for instance to track climate change impacts. However, these cameras cannot analyse image contents, and even though automated analysis could happen post-capture (Mann et al., 2022), real-time monitoring methods are required for greater reactivity (Piao et al., 2019). We show here how our cameras can operate over long time periods for monitoring the phenology of plant targets with high accuracy.

We monitored ground-dwelling Carthusian pink (*Dianthus carthusianorum*) flowers with two cameras situated on our university campus in Hangzhou, China, over a period of 2 months in fall 2022 (Figure 2f). The cameras, outfitted with our custom-built power management PCBs (Balle et al., in press), captured images at intervals of 15 min from the start until the end of civil twilight (approximately 5:30 to 18:00) while they were automatically recharged with 5 W solar panels during sunny weather. The cameras saved all images and analysed them with an object detection CNN, only keeping detections above a conservative confidence threshold of 0.3. We attained a maximum field accuracy of 95% for detecting open Carthusian pink flowers (Figure S7). We automatically tracked multiple single flowers over frames and days using time series clustering in R (R Core Team, 2019) to establish a flowering peak on October 5, 2022 with 16 simultaneously open flowers and a mean lifespan of single flowers of 4 days (range: 0.4–14.5 days).

4 | DISCUSSION

Embedded vision devices such as the one proposed here have far-reaching implications for biodiversity monitoring. We demonstrated the potential of our ecoEye cameras for phenological observation, real-time species surveillance and key ecological interactions such as pollination, potential pest damage and biocontrol. Our versatile design can be utilised for multiple taxa and scaled up to large spatial and temporal extents due to its interchangeable lens options, sufficient frame rates and resolution, and low-power consumption compared to most alternative devices (Table 2), which we also discuss in the following.

4.1 | Taxonomic, spatial and temporal coverage

The interchangeable lens design with adjustable focus enabled the largest taxonomic coverage and deployment flexibility: from millimetre-scale insects to remote flying bats and flowers on the ground (Table 1). In comparison, established camera traps have a limited angle of view by design and can only detect heat displacement of medium to large homeotherms. They have been adapted with limited performance or considerable efforts beyond their core application: for far-away birds, risky arboreal sampling, nocturnal flying insect traces, with poikilothermic animal ramps, or for inanimate plants (Hobbs & Brehme, 2017; Latta et al., 2022; Mann et al., 2022; Moore et al., 2021; Wallace et al., 2022; Zhu et al., 2022). Alternative imaging devices without embedded vision can be used to detect insects (Droissart et al., 2021; Hogeweg et al., 2019; Pegoraro et al., 2020; Ruczyński et al., 2020), and recent designs with embedded vision integrate motion-detection (Bjerge, Nielsen, et al., 2021) or species classification (Bjerge, Mann, et al., 2021), but stay confined to their integrated lenses' angle of view and the corresponding targets that can be imaged. This underlines the necessity for broader image-based, embedded detection approaches which could even be extended to microscopic scales to monitor zooplankton (Lertvilai, 2020).

The presented cameras are able to sample a large temporal and spatial coverage at sufficient temporal resolution. Even though all currently available devices compared here offer high sampling resolutions over coverage times that exceed human capabilities (Table 2), ours also demonstrated successful operation in the field over several months with solar power in contrast to other embedded vision devices. The cameras ran at between 0.4 and 1.7 frames per second depending on the resolution and algorithm (thus processing between 2.3 and 12.6 MB per second). While frame rates are higher at lower resolutions and for more powerful alternative devices, we could detect ephemeral pollination events or bat passes. In contrast, camera traps may miss quickly passing and stationary targets by design (Apps & McNutt, 2018). Bulkier and more expensive embedded vision devices (Table 2) may not facilitate reaching high spatial coverage. Cheaper, more compact devices such as the PICT (when used without motion detection) or camera traps, however, can help to attain high spatial coverage (Bondi et al., 2010) through replication, but they come with challenges in the cumulative data storage and post-deployment data processing demands. Notably, due to their stationary design and fixed focal lenses, none of the devices presented here can sample areas as large as what human observers could sample. Several fields of view and taxa could be sampled concurrently by multiple cameras, but in the future dual or zoom lens designs could strike a better trade-off between coverage and resolution (Hu et al., 2013). Overall, low cost, small size and embedded vision capabilities of embedded vision devices facilitate covering large spatial and temporal scales.

4.2 | Standardised results and reactive monitoring

We used standard Artificial Intelligence (AI) evaluation procedures to measure algorithm accuracies in the field and obtain standardised biodiversity data for some use cases. Vision-based biodiversity detection processes are rarely standardised (Burton et al., 2015). Even though AI drives global camera trap syntheses (Chen et al., 2022), the data are derived from a separate sampling process that remains unquantifiable or hard to estimate with ground-truth (Boyd et al., 2023; Gilbert et al., 2021). We measured the accuracies of our algorithms by screening every saved frame over a calibration period for true and false positives and negatives to infer "field accuracies"

TABLE 2 Comparison of automated, vision-based field biodiversity monitoring devices that have computer vision-based, embedded (i.e. on-device) detection or classification processes. Camera traps, which do not have embedded vision capabilities, are shown for comparison purposes. CNN-based object detection models implicitly include classification (Insect Detect device currently uses only one class). All devices except the one based on commercial digital cameras can be enabled with wireless connectivity (WLAN or 4G), and all embedded vision devices can potentially detect any resolvable target within their field of view. The size and weight indications exclude support and continuous powering accessories. Materials cost in EUR as of 16 Jul 2024.

Author or device	Sampled taxa	Processing hardware	Detection	Classification	Maximum sensor resolution (MP)
Bjerge 2021[a]	Insects	NVIDIA Jetson Nano	Object detection	Object classification	2
Bjerge 2021, 2024[a]	Nocturnal insects	Raspberry Pi 4	Object detection	Object classification	8
Camera trap	Medium to large homeotherms	Passive infrared sensor	Heat displacement	NA	20
Corva 2022	Endotherms and ectotherms	Raspberry 3B+	Blob detection	NA	0.3
Diopsis[a] (Huijbers et al. under review)	Insects	Raspberry Pi Zero W	Object detection	Object classification	8
ecoEye	Insects, bats, birds, flowers	openMV H7+	Blob detection, object detection	Image/ROI/object classification	5
FAIR device	Insects	BeagleBone Black	Blob detection	NA	1
Insect Detect[a]	Insects	Luxonis OAK-1, raspberry Pi zero 2 W	Object detection	Object classification	12
PICT[a]	All	Raspberry Pi Zero	Blob detection	NA	5
Steen 2011	Insects	Commercial digital cameras	Blob detection	NA	0.4–12
VespAI[a]	Invasive hornets	Raspberry Pi 4	Object detection	Object classification	2

[a] Details confirmed by respective authors.

and mathematically derive standardised event numbers (e.g. bat passes, insect visits). Similar calibration approaches were employed with embedded vision devices, and additional object tracking may be crucial to distinguish individuals (Bjerge, Mann, et al., 2021). In our case, the same CNN's variable performance was corrected across multiple deployments. We did not use human observation-based methods for validating our embedded vision approach: no standard methods exist for surveying the exact same observation area as the cameras, and human observation would not have been feasible or trustworthy over such long time ranges.

We argue that since the realised performance of custom-trained CNNs is inevitably variable, standardising results with field data is necessary and potentially cost-effective (Figure S5)—possibly also for fine-tuned models based on global datasets (Shepley et al., 2021; Shimron et al., 2022). As a result, we can estimate the true number of events, circumvent complex analytical modelling approaches that attempt to statistically model them, and potentially derive animal densities over the cameras' field of views. It is thus possible to gather data for different taxa using multiple cameras and to derive comparable activity or density metrics for analyses of trophically linked taxa (such as the predatory bats and insect prey in our agronomy use case). We could thereby harmonise sampling methods across sites, ecosystems and taxa and work towards integrating existing standardisation initiatives such as camtrap DP for metadata reporting or GEO BON for Essential Biodiversity Variables (Bubnicki et al., 2023; Scholes et al., 2012). Admittedly, minimum accuracy thresholds will need to be determined before any policy- and management-relevant applications are envisaged, and every application will require different minimum recall or precision thresholds depending on their goal. For instance, invasive species detection should prioritise high recall, while invasive pest control interventions would require high precision. Ultimately, such standardised data should facilitate large-scale syntheses based on field studies with variable setups, yielding actionable density-based evidence for biodiversity management.

Embedded vision enables real-time triggered reactions. At its simplest, the embedded analysis accelerates the workflow by providing readily usable detections data with defined metrics, such as the blob characteristics or classification probabilities for pests or pollinators in our use cases. Pre-processed data such as from blob detection approaches—that are needed in every situation where training images

Operation	Power consumption (W)	Power autonomy (continuous operation)	Size and weight	Materials cost (EUR)
Day	5–10	NA (grid power)	NA	390 to 550 (grid power), 780–1100 (solar panel)
Night	5 (idle)–30 (night)	NA (grid or solar power)	NA	930 (grid power) +400 (solar panel)
Day & night	0.001	30–180 days (solar power optional)	10 × 10 × 15 cm; 700 g (typical)	150 to 500 (typical, low to mid-range)
Day & night	2.65–4.13	NA (solar power)	20 × 15 × 10 cm (estimated from photograph)	550
Day & night	2.5–6	NA (grid or solar power)	92 × 97 × 48 cm; 17 kg	NA
Day & night	1.3 (IR off)–2.1 (IR on)	23 h (10 Ah battery, IR on at night) to 2 m (solar power)	10 × 8 × 10 cm; 400 g	13 (Chinese market prices)
Day & night (with blacklight lamps)	NA	NA (grid power)	270 × 100 × 175 cm, NA	210 (excluding Malaise trap)
Day	4.4	20 h (24.8 Ah battery) to NA (solar power)	160 × 250 × 125 cm (enclosure only)	700 (solar panel, 2 batteries), 530 (one battery), 187 (core hardware)
Day & night	1.33 (day)–2.46 (night) (with MotionEyeOS)	42 h (30 Ah battery)	20 × 13 × 7 cm, 940 g (30 Ah battery)	100 to 155
Day	4.8	8–12 days	NA, 21.5 kg	463
Day	NA	48 h (27 Ah battery) or NA (grid power)	NA	100 (grid power) to 340 (with batteries, solar panel, 4G)

are not yet available—or coarse classification models may be sent to cloud servers that are less power- and computation-constrained. Then, deeper AI models may attempt accurate post-classification; this is the approach edge computing devices have been designed for. Potentially, even though detailed animal behaviour information (e.g. flying versus feeding bats) could be obtained with embedded CNNs, our ethological use case did not yield enough training images to attempt this. Beyond these reporting and networking improvements, highly confident detections could be used to trigger critical alerts for pest management or species protection. Among the alternative embedded vision devices we compared, only one was networked for real-time data transmission (Bjerge, Mann, & Høye, 2021). Although devices without embedded vision such as "cellular" camera traps can also be networked, costs for sending irrelevant images can be prohibitively high, thereby justifying "edge" solutions (Wang et al., 2020), even though hybrid solutions exist (Whytock et al., 2023). In general, real-time reactions matter most for remote locations and time-sensitive applications. One could imagine alerts when biocontrol rates fall below crop-safe levels, suction sampling of flower visitors, electrocution of specific pests on light boards, or activation of poacher deterrents. Embedded vision cameras thus offer unprecedented opportunities for meaningful actions.

4.3 | Current limitations and outlook

Our proposed devices are both enabled and limited by their underpinning technologies. Cameras are optically limited to a relatively narrow field of view compared to human observers, who will easily detect more targets within a short period of time, although this shortcoming is alleviated by longer observation times. Continuous image-by-image analysis in embedded devices, however, limits the operation time in comparison to camera traps when no lasting power source is connected. Future implementations with on-sensor AI or event cameras will push the boundaries in terms of image quality, speed and identification depth. Even though "DIY" versions of the proposed system are conceivable, robust industrial products will lower costs with economies of scale, provide access to the technology for most end-users and facilitate deployment at large scales.

Our detection and classification accuracies show mixed results in comparison to the most similar applications. The only other automated vision-based bat monitoring device to our knowledge (Gottwald et al., 2021) uses radio signals or ultrasound to trigger video capture, and the other vision-based monitoring device for flying insects (Ruczyński et al., 2020) uses timelapse photography to capture flying insects, so to our knowledge, there are no image-based detection or classification data to compare this use case's results to. While we obtained a mean classification accuracy of 58% over seven insect taxa in our agronomy use case, higher-resolution setups achieved higher taxonomic resolution for eight moth species with an average accuracy of 71% (Bjerge, Nielsen, et al., 2021). The most similar device monitoring pollinators had low precision (60% on average with PowerShot camera) and unmeasured recall. For birds, we found no other embedded vision device to compare our results to. Finally, the most similar plant phenology monitoring device achieved high accuracies of 91% (Mann et al., 2022), a similar performance to ours, albeit with offline image analysis. Overall, although our camera compares favourably overall, more robust algorithms are needed to apply them to other contexts. For instance, our flying fox detection method achieved low accuracies due to its unspecific and overly sensitive blob detection-based process, and we were not able to use object detection models instead as too few training data were available. Our waterfowl detection method also would have required more training data to handle the classification of juveniles when we repeated the deployment after the breeding season. Arguably, reusing our models elsewhere in all but our most standardised use cases, such as the nocturnal bat and insect monitoring use cases that use dark or illuminated backgrounds, would require training new models to tackle the domain shift. Finally, not all vision-based monitoring scenarios require embedded vision devices: for instance, when devices can be networked to analyse images on dedicated stations (Koger et al., 2023; Mann et al., 2022; Ratnayake et al., 2023). In any case, it is also possible to use the ecoEye camera without embedded vision algorithms, as an image capture device similar to the PICT's tested application (Droissart et al., 2021).

Embedded vision devices are poised to dethrone camera traps, currently the gold standard for vision-based monitoring, whose various constraints prevent broad scale implementation (Glover-Kapfer et al., 2019). Although camera traps currently handle much longer deployments than constantly-operating embedded vision devices and effectively sample medium and large-size homeotherms, future technological innovations such as ever-improving object detection models, on-sensor machine learning, ultra-low-power chips, higher battery energy densities and tiny machine learning will tip the scale towards higher-resolution imaging, longer operation times and more sophisticated analysis (Allan et al., 2018). Together with economies of scale and low-budget designs that should accompany the increased uptake of modern devices, prices will similarly decrease and in turn stimulate adoption. Open-set recognition will help to tackle the detection of taxonomically unresolved taxa (Geng et al., 2021), and for some species, individuals will become distinguishable (Vidal et al., 2021). Beyond this, technology-enabled monitoring for adaptive management of ecosystems can be envisaged (Evansen et al., 2021), and the data should be used for predicting trends under global change (Besson et al., 2022). However, just like new observation technologies have historically driven scientific progress, the sheer possibilities of embedded vision systems necessitate regulation to manage dual use and alleviate ethics, privacy and security issues. Likewise, harmonisation efforts should be pursued on a much higher level by devising official standards for computer vision-based monitoring of biodiversity. We hope that with tools that become ever more efficient, we will be able to address challenges faced by nature and society and focus on the implementation of solutions.

AUTHOR CONTRIBUTIONS

Kevin F. A. Darras: conceptualisation, data curation, formal analysis, investigation, methodology, project administration, software, supervision, validation, visualisation, writing—original draft preparation, writing—review & editing. Marcel Balle: data curation, investigation, methodology, software, validation, visualisation, writing—review & editing. Wenxiu Xu: data curation, investigation, validation, writing—review & editing. Yang Yan: formal analysis, investigation, methodology, software, validation, writing—review & editing. Vincent Gbouna Zakka: investigation, methodology, software, writing—review & editing. Manuel Toledo-Hernández: conceptualisation, visualisation, writing—review & editing. Dong Sheng: data curation, formal analysis, investigation, validation, writing—review & editing. Wei Lin: data curation, investigation, validation, writing – review & editing. Boyu Zhang: investigation, resources. Zhenzhong Lan: methodology, supervision. Li Fupeng: project administration, resources, supervision. Thomas Cherico Wanger: conceptualisation, funding acquisition, investigation, project administration, resources, supervision, writing—original draft preparation, writing—review & editing.

ACKNOWLEDGEMENTS

K.F.A.D. thanks his family for all the patience and help with assembling cameras and analysing data. Sincere thanks to Yang Dewei for the fieldwork in the botanical garden, Ellena F Yusti for the flying fox identification, Xueqing He and Li Juan for advice and support throughout the project. This project was funded by a Westlake University Startup Fund (T.C.W.).

CONFLICT OF INTEREST STATEMENT

The authors declare the following competing interests: T.C.W., M.B. and K.F.A.D. are involved in a company that sells the commercial camera version. T.C.W. and K.F.A.D. hold a patent for the camera design (Patent number: 202210268842.X, 2022-05-27, Vol. 38, No. 2102 Patent Gazette, No. CN114544634A.).

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.14436.

DATA AVAILABILITY STATEMENT

Data for camera assembly and operation available via the Zenodo repository https://zenodo.org/records/13739561 (Darras, 2024b). Image samples and data for reproducing results and camera operation available via the Dryad Digital Repository https://doi.org/10.5061/dryad.1ns1rn90j (Darras, 2024a). The commercial version of this camera is available from www.ecoNect.biz or for direct purchase from https://www.seeedstudio.com/EcoEye-Embedded-Vision-Camera-p-5843.html.

ORCID

Kevin F. A. Darras https://orcid.org/0000-0002-9013-3784

REFERENCES

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

Apps, P. J., & McNutt, J. W. (2018). How camera traps work and how to work them. *African Journal of Ecology*, *56*(4), 702–709. https://doi.org/10.1111/aje.12563

Balle, M., Xu, W., Darras, K. F. A., & Wanger, T. C. (in press). A power management and control system for environmental monitoring devices. *IEEE Transactions on AgriFood Electronics*. https://doi.org/10.1109/TAFE.2024.3472493

Baqi, A., Lim, V.-C., Yazid, H., Anwarali Khan, F. A., Lian, C. J., Nelson, B. R., Sathiya Seelan, J. S., Appalasamy, S., Mokhtar, S. I., & Kumaran, J. V. (2022). A review of durian plant-bat pollinator interactions. *Journal of Plant Interactions*, *17*(1), 105–126. https://doi.org/10.1080/17429145.2021.2015466

Beery, S. (2021). Scaling biodiversity monitoring for the data age. *XRDS: Crossroads, The ACM Magazine for Students*, *27*(4), 14–18. https://doi.org/10.1145/3466857

Besson, M., Alison, J., Bjerge, K., Gorochowski, T. E., Høye, T. T., Jucker, T., Mann, H. M. R., & Clements, C. F. (2022). Towards the fully automated monitoring of ecological communities. *Ecology Letters*, *25*(12), 2753–2775. https://doi.org/10.1111/ele.14123

Bjerge, K., Mann, H. M. R., & Høye, T. T. (2021). Real-time insect tracking and monitoring with computer vision and deep learning. *Remote Sensing in Ecology and Conservation*, *8*(3), 315–327. https://doi.org/10.1002/rse2.245

Bjerge, K., Nielsen, J. B., Sepstrup, M. V., Helsing-Nielsen, F., & Høye, T. T. (2021). An automated light trap to monitor moths (Lepidoptera) using computer vision-based tracking and deep learning. *Sensors*, *21*(2), Article 2. https://doi.org/10.3390/s21020343

Boakes, E. H., McGowan, P. J. K., Fuller, R. A., Chang-qing, D., Clark, N. E., O'Connor, K., & Mace, G. M. (2010). Distorted views of biodiversity: Spatial and temporal bias in species occurrence data. *PLoS Biology*, *8*(6), e1000385. https://doi.org/10.1371/journal.pbio.1000385

Bondi, N. D., White, J. G., Stevens, M., Cooke, R., Bondi, N. D., White, J. G., Stevens, M., & Cooke, R. (2010). A comparison of the effectiveness of camera trapping and live trapping for sampling terrestrial small-mammal communities. *Wildlife Research*, *37*(6), 456–465. https://doi.org/10.1071/WR10046

Borgelt, J., Dorber, M., Høiberg, M. A., & Verones, F. (2022). More than half of data deficient species predicted to be threatened by extinction. *Communications Biology*, *5*(1), Article 1. https://doi.org/10.1038/s42003-022-03638-9

Boyd, R. J., Powney, G. D., & Pescott, O. L. (2023). We need to talk about nonprobability samples. *Trends in Ecology & Evolution*, *38*(6), 521–531. https://doi.org/10.1016/j.tree.2023.01.001

Bubnicki, J. W., Norton, B., Baskauf, S. J., Bruce, T., Cagnacci, F., Casaer, J., Churski, M., Cromsigt, J. P. G. M., Farra, S. D., Fiderer, C., Forrester, T. D., Hendry, H., Heurich, M., Hofmeester, T. R., Jansen, P. A., Kays, R., Kuijper, D. P. J., Liefting, Y., Linnell, J. D. C., … Desmet, P. (2023). *Camtrap DP: An open standard for the FAIR exchange and archiving of camera trap data*. https://ecoevorxiv.org/repository/view/5593/

Burton, A. C., Neilson, E., Moreira, D., Ladle, A., Steenweg, R., Fisher, J. T., Bayne, E., & Boutin, S. (2015). REVIEW: 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

Chen, C., Brodie, J. F., Kays, R., Davies, T. J., Liu, R., Fisher, J. T., Ahumada, J., McShea, W., Sheil, D., Agwanda, B., Andrianarisoa, M. H., Appleton, R. D., Bitariho, R., Espinosa, S., Grigione, M. M., Helgen, K. M., Hubbard, A., Hurtado, C. M., Jansen, P. A., … Burton, A. C. (2022). Global camera trap synthesis highlights the importance of protected areas in maintaining mammal diversity. *Conservation Letters*, *15*(2), e12865. https://doi.org/10.1111/conl.12865

Costello, M. J., Coll, M., Danovaro, R., Halpin, P., Ojaveer, H., & Miloslavich, P. (2010). A census of marine biodiversity knowledge, resources, and future challenges. *PLoS One*, *5*(8), e12110. https://doi.org/10.1371/journal.pone.0012110

Darras, K. F. (2024b). R script and CSV data. *Dryad*. https://doi.org/10.5061/dryad.1ns1rn90j

Darras, K. F. (n.d.). *SAT-Lab-GitHub/ecoEye-open: Open CAD and EDA data for ecoEye embedded vision camera*. https://github.com/SAT-Lab-GitHub/ecoEye-open/tree/main

Darras, K. F. A. (2024a). *SAT-lab-GitHub/ecoEye-open: Original release (V49)—methods in Ecology & Evolution* (version 1.0.0). [Computer Software]. *Zenodo*. https://doi.org/10.5281/zenodo.13739561

Darras, K. F. A., Yusti, E., Huang, J. C.-C., Zemp, D.-C., Kartono, A. P., & Wanger, T. C. (2021). Bat point counts: A novel sampling method shines light on flying bat communities. *Ecology and Evolution*, *11*(23), 17179–17190. https://doi.org/10.1002/ece3.8356

Darras, K. F. A., Yusti, E., Knorr, A., Huang, J. C.-C., Kartono, A. P., & Ilham. (2021). Sampling flying bats with thermal and near-infrared imaging and ultrasound recording: Hardware and workflow for bat point counts. *F1000Research*, *10*, 189. https://doi.org/10.12688/f1000research.51195.1

Doan, T. M. (2003). Which methods are Most effective for surveying rain forest herpetofauna? *Journal of Herpetology*, *37*(1), 72–81. https://doi.org/10.1670/0022-1511(2003)037[0072:WMAMEF]2.0.CO;2

Dove, S., Bohm, M., Freeman, R., McRae, L., & Murrell, D. J. (2023). How much data do we need? Reliability and data deficiency in global vertebrate biodiversity trends (p. 2023.03.18.532273). *bioRxiv*. https://doi.org/10.1101/2023.03.18.532273

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

Dutta, L., & Bharali, S. (2021). TinyML meets IoT: A comprehensive survey. *Internet of Things*, *16*, 100461. https://doi.org/10.1016/j.iot.2021.100461

EcoEye–Embedded Vision Camera for Environmental Monitoring. (n.d.). https://www.seeedstudio.com/EcoEye-Embedded-Vision-Camera-p-5843.html

Engel, M. S., Ceríaco, L. M. P., Daniel, G. M., Dellapé, P. M., Löbl, I., Marinov, M., Reis, R. E., Young, M. T., Dubois, A., Agarwal, I., Lehmann, A. P., Alvarado, M., Alvarez, N., Andreone, F., Araujo-Vieira, K., Ascher, J. S., Baêta, D., Baldo, D., Bandeira, S. A., … Zacharie, C. K. (2021). The taxonomic impediment: A shortage of taxonomists, not the lack of technical approaches. *Zoological Journal of the Linnean Society*, *193*(2), 381–387. https://doi.org/10.1093/zoolinnean/zlab072

Evansen, M., Carter, A., & Malcom, J. (2021). A monitoring policy framework for the United States endangered species act. *Environmental Research Letters*, *16*(3), 031001. https://doi.org/10.1088/1748-9326/abe0ea

Fitzpatrick, M. C., Preisser, E. L., Ellison, A. M., & Elkinton, J. S. (2009). Observer bias and the detection of low-density populations. *Ecological Applications*, *19*(7), 1673–1679. https://doi.org/10.1890/09-0265.1

Geng, C., Huang, S.-J., & Chen, S. (2021). Recent advances in open set recognition: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *43*(10), 3614–3631. https://doi.org/10.1109/TPAMI.2020.2981604

Gilbert, N. A., Clare, J. D. J., Stenglein, J. L., & Zuckerberg, B. (2021). Abundance estimation of unmarked animals based on camera-trap data. *Conservation Biology*, *35*(1), 88–100. https://doi.org/10.1111/cobi.13517

Glover-Kapfer, P., Soto-Navarro, C. A., & Wearn, O. R. (2019). Camera-trapping version 3.0: Current constraints and future priorities for development. *Remote Sensing in Ecology and Conservation*, *5*(3), 209–223. https://doi.org/10.1002/rse2.106

Gorrod, E. J., & Keith, D. A. (2009). Observer variation in field assessments of vegetation condition: Implications for biodiversity conservation. *Ecological Management & Restoration*, *10*(1), 31–40. https://doi.org/10.1111/j.1442-8903.2009.00437.x

Gottwald, J., Lampe, P., Höchst, J., Friess, N., Maier, J., Leister, L., Neumann, B., Richter, T., Freisleben, B., & Nauss, T. (2021). BatRack: An open-source multi-sensor device for wildlife research. *Methods in Ecology and Evolution*, *12*(10), 1867–1874. https://doi.org/10.1111/2041-210X.13672

Hobbs, M. T., & Brehme, C. S. (2017). An improved camera trap for amphibians, reptiles, small mammals, and large invertebrates. *PLoS One*, *12*(10), e0185026. https://doi.org/10.1371/journal.pone.0185026

Hogeweg, L., Zeegers, T., Katramados, I., & Jongejans, E. (2019). Smart insect cameras. *Biodiversity Information Science and Standards*, *3*, e39241. https://doi.org/10.3897/biss.3.39241

Høye, T. T., Ärje, J., Bjerge, K., Hansen, O. L. P., Iosifidis, A., Leese, F., Mann, H. M. R., Meissner, K., Melvad, C., & Raitoharju, J. (2021). Deep learning and computer vision will transform entomology. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(2), e2002545117. https://doi.org/10.1073/pnas.2002545117

Hu, J., Hu, S., & Sun, Z. (2013). A real time dual-camera surveillance system based on tracking-learning-detection algorithm. In *2013 25th Chinese control and decision conference (CCDC)* (pp. 886–891). IEEE. https://doi.org/10.1109/CCDC.2013.6561048

Jackson, L., van Noordwijk, M., Bengtsson, J., Foster, W., Lipper, L., Pulleman, M., Said, M., Snaddon, J., & Vodouhe, R. (2010). Biodiversity and agricultural sustainagility: From assessment to adaptive management. *Current Opinion in Environmental Sustainability*, *2*(1), 80–87. https://doi.org/10.1016/j.cosust.2010.02.007

Johnston, A., Matechou, E., & Dennis, E. B. (2023). Outstanding challenges and future directions for biodiversity monitoring using citizen science data. *Methods in Ecology and Evolution*, *14*(1), 103–116. https://doi.org/10.1111/2041-210X.13834

Keith, D. A., Martin, T. G., McDonald-Madden, E., & Walters, C. (2011). Uncertainty and adaptive management for biodiversity conservation. *Biological Conservation*, *144*(4), 1175–1178. https://doi.org/10.1016/j.biocon.2010.11.022

Klein, A.-M., Vaissière, B. E., Cane, J. H., Steffan-Dewenter, I., Cunningham, S. A., Kremen, C., & Tscharntke, T. (2006). Importance of pollinators in changing landscapes for world crops. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1608), 303–313. https://doi.org/10.1098/rspb.2006.3721

Klein, D., Mckown, M., & Tershy, B. (2015). *Deep learning for large scale biodiversity monitoring*. https://doi.org/10.13140/RG.2.1.1051.7201

Koger, B., Hurme, E., Costelloe, B. R., O'Mara, M. T., Wikelski, M., Kays, R., & Dechmann, D. K. N. (2023). An automated approach for counting groups of flying animals applied to one of the world's largest bat colonies. *Ecosphere*, *14*(6), e4590. https://doi.org/10.1002/ecs2.4590

Latta, S. C., Michaels, M. A., Michot, T. C., Shrum, P. L., Johnson, P., Tischendorf, J., Weeks, M., Trochet, J., Scheifler, D., & Ford, B. (2023). Multiple lines of evidence suggest the persistence of the ivory-billed woodpecker (*Campephilus principalis*) in Louisiana. *Ecology and Evolution*, *13*(5), e10017. https://doi.org/10.1002/ece3.10017

Latta, S. C., Michaels, M. A., Scheifler, D., Michot, T. C., Shrum, P. L., Johnson, P., Tischendorf, J., Weeks, M., Trochet, J., & Ford, B. (2022). Multiple lines of evidence indicate survival of the ivory-billed woodpecker in Louisiana [preprint]. *BioRXiv*. https://doi.org/10.1101/2022.04.06.487399

Lertvilai, P. (2020). The in situ plankton assemblage eXplorer (IPAX): An inexpensive underwater imaging system for zooplankton study. *Methods in Ecology and Evolution*, *11*(9), 1042–1048. https://doi.org/10.1111/2041-210X.13441

Lintott, P. R., Bunnefeld, N., & Park, K. J. (2015). Opportunities for improving the foraging potential of urban waterways for bats. *Biological Conservation*, *191*, 224–233. https://doi.org/10.1016/j.biocon.2015.06.036

Mann, H. M. R., Iosifidis, A., Jepsen, J. U., Welker, J. M., Loonen, M. J. J. E., & Høye, T. T. (2022). Automatic flower detection and phenology monitoring using time-lapse cameras and deep learning. *Remote Sensing in Ecology and Conservation*, *8*(6), 765–777. https://doi.org/10.1002/rse2.275

McCallum, J. (2013). Changing use of camera traps in mammalian field research: Habitats, taxa and study types. *Mammal Review*, *43*(3), 196–206. https://doi.org/10.1111/j.1365-2907.2012.00216.x

Meyer, C., Kreft, H., Guralnick, R., & Jetz, W. (2015). Global priorities for an effective information basis of biodiversity distributions. *Nature Communications*, *6*(1), Article 1. https://doi.org/10.1038/ncomms9221

Moersberger, H., Valdez, J., Martin, J. G. C., Junker, J., Georgieva, I., Bauer, S., Beja, P., Breeze, T. D., Fernandez, M., Fernández, N., Brotons, L., Jandt, U., Bruelheide, H., Kissling, W. D., Langer, C., Liquete, C., Lumbierres, M., Solheim, A. L., Maes, J., … Bonn, A. (2024). Biodiversity monitoring in Europe: User and policy needs. *Conservation Letters*, e13038. https://doi.org/10.1111/conl.13038

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

Morris, A. D., Miller, D. A., & Kalcounis-Rueppell, M. C. (2010). Use of forest edges by bats in a managed pine Forest landscape. *The Journal of Wildlife Management*, *74*(1), 26–34. https://doi.org/10.2193/2008-471

O'Connor, R. S., Kunin, W. E., Garratt, M. P. D., Potts, S. G., Roy, H. E., Andrews, C., Jones, C. M., Peyton, J. M., Savage, J., Harvey, M. C., Morris, R. K. A., Roberts, S. P. M., Wright, I., Vanbergen, A. J., & Carvell, C. (2019). Monitoring insect pollinators and flower visitation: The effectiveness and feasibility of different survey methods. *Methods in Ecology and Evolution*, *10*(12), 2129–2140. https://doi.org/10.1111/2041-210X.13292

Oellermann, M., Jolles, J. W., Ortiz, D., Seabra, R., Wenzel, T., Wilson, H., & Tanner, R. L. (2022). Open hardware in science: The benefits of open electronics. *Integrative and Comparative Biology*, *62*(4), 1061–1075. https://doi.org/10.1093/icb/icac043

Pegoraro, L., Hidalgo, O., Leitch, I. J., Pellicer, J., & Barlow, S. E. (2020). Automated video monitoring of insect pollinators in the field. *Emerging Topics in Life Sciences*, *4*(1), 87–97. https://doi.org/10.1042/ETLS20190074

Piao, S., Liu, Q., Chen, A., Janssens, I. A., Fu, Y., Dai, J., Liu, L., Lian, X., Shen, M., & Zhu, X. (2019). Plant phenology and global climate change: Current progresses and challenges. *Global Change Biology*, *25*(6), 1922–1940. https://doi.org/10.1111/gcb.14619

Potts, S. G., Imperatriz-Fonseca, V., Ngo, H. T., Aizen, M. A., Biesmeijer, J. C., Breeze, T. D., Dicks, L. V., Garibaldi, L. A., Hill, R., Settele, J., & Vanbergen, A. J. (2016). Safeguarding pollinators and their values to human well-being. *Nature*, *540*(7632), Article 7632. https://doi.org/10.1038/nature20588

Puig-Montserrat, X., Torre, I., López-Baucells, A., Guerrieri, E., Monti, M. M., Ràfols-García, R., Ferrer, X., Gisbert, D., & Flaquer, C. (2015). Pest control service provided by bats in Mediterranean rice paddies: Linking agroecosystems structure to ecological functions. *Mammalian Biology-Zeitschrift Für Säugetierkunde*, *80*(3), 237–245. https://doi.org/10.1016/j.mambio.2015.03.008

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-project.org/

Ralph, C. J., Droege, S., & Sauer, J. R. (1995). *Managing and monitoring birds using point counts: Standards and applications* (pp. 161–168). USDA Forest Service. https://www.fs.usda.gov/psw/publications/documents/psw_gtr149/psw_gtr149_pg161_168.pdf

Ratnayake, M. N., Amarathunga, D. C., Zaman, A., Dyer, A. G., & Dorin, A. (2023). Spatial monitoring and insect Behavioural analysis using computer vision for precision pollination. *International Journal of Computer Vision*, *131*(3), 591–606. https://doi.org/10.1007/s11263-022-01715-4

Ruczyński, I., Hałat, Z., Zegarek, M., Borowik, T., & Dechmann, D. K. N. (2020). Camera transects as a method to monitor high temporal and spatial ephemerality of flying nocturnal insects. *Methods in Ecology and Evolution*, *11*(2), 294–302. https://doi.org/10.1111/2041-210X.13339

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2019). MobileNetV2: Inverted residuals and linear bottlenecks (arXiv:1801.04381). *arXiv*. https://doi.org/10.48550/arXiv.1801.04381

Scholes, R. J., Walters, M., Turak, E., Saarenmaa, H., Heip, C. H., Tuama, É. Ó., Faith, D. P., Mooney, H. A., Ferrier, S., Jongman, R. H., Harrison, I. J., Yahara, T., Pereira, H. M., Larigauderie, A., & Geller, G. (2012). Building a global observing system for biodiversity. *Current Opinion in Environmental Sustainability*, *4*(1), 139–146. https://doi.org/10.1016/j.cosust.2011.12.005

Shepley, A., Falzon, G., Meek, P., & Kwan, P. (2021). Automated location invariant animal detection in camera trap images using publicly available data sources. *Ecology and Evolution*, *11*(9), 4494–4506. https://doi.org/10.1002/ece3.7344

Shimoda, M., & Honda, K. (2013). Insect reactions to light and its applications to pest management. *Applied Entomology and Zoology*, *48*(4), 413–421. https://doi.org/10.1007/s13355-013-0219-x

Shimron, E., Tamir, J. I., Wang, K., & Lustig, M. (2022). Implicit data crimes: Machine learning bias arising from misuse of public data. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(13), e2117203119. https://doi.org/10.1073/pnas.2117203119

Steen, R. (2017). Diel activity, frequency and visit duration of pollinators in focal plants: In situ automatic camera monitoring and data processing. *Methods in Ecology and Evolution*, *8*(2), 203–213. https://doi.org/10.1111/2041-210X.12654

Steffen, W., Richardson, K., Rockström, J., Cornell, S. E., Fetzer, I., Bennett, E. M., Biggs, R., Carpenter, S. R., de Vries, W., de Wit, C. A., Folke, C., Gerten, D., Heinke, J., Mace, G. M., Persson, L. M., Ramanathan, V., Reyers, B., & Sörlin, S. (2015). Planetary boundaries: Guiding human development on a changing planet. *Science*, *347*(6223), 1259855. https://doi.org/10.1126/science.1259855

UNECE. (2023). *Guidelines for developing national biodiversity monitoring systems*. United Nations. https://unece.org/sites/default/files/2023-01/2228292_E_ECE_CEP_198_WEB_11_0.pdf

UNEP. (2020). *First draft of the post-2020 global biodiversity framework [convention on biological diversity]*. UN Environment Programme. CBD/WG2020/3/3.

van Klink, R., August, T., Bas, Y., Bodesheim, P., Bonn, A., Fossøy, F., Høye, T. T., Jongejans, E., Menz, M. H. M., Miraldo, A., Roslin, T., Roy, H. E., Ruczyński, I., Schigel, D., Schäffler, L., Sheard, J. K., Svenningsen, C., Tschan, G. F., Wäldchen, J., … Bowler, D. E. (2022). Emerging technologies revolutionise insect ecology and monitoring. *Trends in Ecology & Evolution*, *37*(10), 872–885. https://doi.org/10.1016/j.tree.2022.06.001

Vidal, M., Wolf, N., Rosenberg, B., Harris, B. P., & Mathis, A. (2021). Perspectives on individual animal identification from biology and computer vision. *Integrative and Comparative Biology*, *61*(3), 900–916. https://doi.org/10.1093/icb/icab107

Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., Felländer, A., Langhans, S. D., Tegmark, M., & Fuso Nerini, F. (2020). The role of artificial intelligence in achieving the sustainable development goals. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-019-14108-y

Wallace, J. R. A., Reber, T., Beaton, B., Dreyer, D., & Warrant, E. J. (2022). Inexpensive monitoring of flying insect activity and abundance using wildlife cameras (p. 2021.08.24.457487). *bioRxiv*. https://doi.org/10.1101/2021.08.24.457487

Wang, X., Han, Y., Leung, V. C. M., Niyato, D., Yan, X., & Chen, X. (2020). Convergence of edge computing and deep learning: A comprehensive survey. In *IEEE communications surveys & tutorials* (Vol. 22(2), pp. 869–904). IEEE Communications Surveys & Tutorials. https://doi.org/10.1109/COMST.2020.2970550

Wanger, T. C., Darras, K., Bumrungsri, S., Tscharntke, T., & Klein, A.-M. (2014). Bat pest control contributes to food security in Thailand. *Biological Conservation*, *171*, 220–223. https://doi.org/10.1016/j.biocon.2014.01.030

Weinstein, B. G., Garner, L., Saccomanno, V. R., Steinkraus, A., Ortega, A., Brush, K., Yenni, G., McKellar, A. E., Converse, R., Lippitt, C. D., Wegmann, A., Holmes, N. D., Edney, A. J., Hart, T., Jessopp, M. J., Clarke, R. H., Marchowski, D., Senyondo, H., Dotson, R., … Ernest, S. K. M. (2022). A general deep learning model for bird detection in high-resolution airborne imagery. *Ecological Applications*, *32*(8), e2694. https://doi.org/10.1002/eap.2694

Whytock, R. C., Suijten, T., van Deursen, T., Świeżewski, J., Mermiaghe, H., Madamba, N., Mouckoumou, N., Zwerts, J. A., Pambo, A. F. K., Bahaa-el-din, L., Brittain, S., Cardoso, A. W., Henschel, P., Lehmann, D., Momboua, B. R., Makaga, L., Orbell, C., White, L. J. T., Iponga, D. M., & Abernethy, K. A. (2023). Real-time alerts from AI-enabled camera traps using the iridium satellite network: A case-study in Gabon, Central Africa. *Methods in Ecology and Evolution*, *14*(3), 867–874. https://doi.org/10.1111/2041-210X.14036

Yao, Q., Feng, J., Tang, J., Xu, W., Zhu, X., Yang, B., Lü, J., Xie, Y., Yao, B., Wu, S., Kuai, N., & Wang, L. (2020). Development of an automatic monitoring system for rice light-trap pests based on machine vision. *Journal of Integrative Agriculture*, *19*(10), 2500–2513. https://doi.org/10.1016/S2095-3119(20)63168-9

Zhu, C., Li, W., Gregory, T., Wang, D., Ren, P., Zeng, D., Kang, Y., Ding, P., & Si, X. (2022). Arboreal camera trapping: A reliable tool to monitor plant-frugivore interactions in the trees on large scales. *Remote Sensing in Ecology and Conservation*, *8*(1), 92–104. https://doi.org/10.1002/rse2.232

SUPPORTING INFORMATION

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

Figure S1. Overview of the sampling locations in the Zhejiang and Hainan provinces of China (top).
Figure S2. (A) A flying fox (*Roussettus leschenaultii*) feeding on a durian flower, with four outer blob corners drawn as a polygon, and labelled blob ID. (B) Durian fruits resulting from pollinated flowers on the same tree. (C) Accuracy of the blob detection algorithm for different minimum and maximum blob pixel areas on the five different deployment nights.
Figure S3. (A) Number of detections during each survey night (at maximum accuracy confidence level), drawn for bats and insects and separated by location. (B) Camera rig, mounted on tripod, on concrete wall in rice field before dusk. (C) Example image showing blob-detected bat and insect (insect is closer to camera); blob outer corners annotated with polygon; blob IDs labelled. (D) Precision-recall curves for the classification of detected blobs based on actual survey deployments.
Figure S4. (A) Detection counts with time per detected class on each deployment night. (B) Camera rig, mounted on tripod before dusk, set up on concrete wall inside rice field. (C) Server user interface screenshot showing live counts for each detected class, including the board detections. (D) Example image showing detected blobs on the whole image with annotated outer corners as polygons; white-backed planthopper detected in blob ID 707, hymenoptera detected in other blobs; insects smaller than the minimum blob pixel area threshold were not detected (insect to the right of ID 706); stationary insects staying on the board long enough to be blended into the reference image would leave empty blobs behind after they left (ID 704). € Precision-recall curves for the classification of detected blobs based on actual survey deployments (board class omitted).
Figure S5. (A): Probability density functions for bee detection occurrence with time of the day, across multiple days that were completely sampled, separated by deployment. Raw and corrected (i.e. estimated visit) densities are shown with different line colors. (B) Robustness of F1 score measurements for each deployment depending on the analysed proportion of our 10 000 analysed images per deployment; the number of runs at each proportion increment is determined by the rounded inverse of the proportion multiplied with the run multiplicator; upper and lower ribbon boundaries represent one standard deviation from the mean, which is indicated by the coloured lines. (C) Precision-recall curve for the detection of solitary bees (*Osmia bicornis*) based on actual survey deployments, excluding one stationary detection in deployment A3—camera 1; confidence thresholds are labelled at each data point.
Figure S6. (A) Screenshot of a the server's user interface during the deployment, showing transferred images with detected mandarin ducks and bar plots depicting the number of detections for each sex with time. (B) Detections count with time, separated by identified duck sex, showing the estimated bird passes (i.e. standardised detections). (C) Precision-recall curve for the detection of each duck sex based on the first survey deployment during the breeding season, in 2022.
Figure S7. (A) Individual tracked flowers (flower IDs omitted from Y axis for clarity), separated by camera (IDs 2 and 13). Red lines represent tracked flowers over single days, and blue lines connect identical flowers through nights. (B) Probability density of the flower lifespan. (C) Precision-recall curve for the detection of Carthusian pink (*Dianthus carthusianorum*) flowers based on actual survey deployments; confidence thresholds are labelled at each data point.
Supplementary Methods. Hardware design and use cases.

How to cite this article: Darras, K. F. A., Balle, M., Xu, W., Yan, Y., Zakka, V. G., Toledo-Hernández, M., Sheng, D., Lin, W., Zhang, B., Lan, Z., Fupeng, L., & Wanger, T. C. (2024). Eyes on nature: Embedded vision cameras for terrestrial biodiversity monitoring. *Methods in Ecology and Evolution*, *00*, 1–14. https://doi.org/10.1111/2041-210X.14436