1   **Deep learning and computer vision will transform entomology**

2

3   Toke T. Høye[1,*], Johanna Ärje[1,2], Kim Bjerge[3], Oskar L. P. Hansen[1,4,5,6], Alexandros Iosifidis[7],

4   Florian Leese[8], Hjalte M. R. Mann[1], Kristian Meissner[9], Claus Melvad[10], Jenni Raitoharju[9]

5

6   *1. Department of Bioscience and Arctic Research Centre, Aarhus University, Grenåvej 14, DK-*

7   *8410 Rønde, Denmark*

8   *2. Unit of Computing Sciences, Tampere University, Finland*

9   *3. School of Engineering, Aarhus University, Finlandsgade 22, 8200 Aarhus N, Denmark*

10   *4. Natural History Museum Aarhus, Wilhelm Meyers Allé 10, DK-8000 Aarhus C*

11   *5. Department of Biology – Center for Biodiversity Dynamics in a Changing World*

12   *(BIOCHANGE), Aarhus University, Ny Munkegade 116, DK-8000 Aarhus C*

13   *6. Department of Biology - Ecoinformatics and Biodiversity, Aarhus University, Ny Munkegade*

14   *116, DK-8000 Aarhus C*

15   *7. Department of Engineering, Aarhus University, Denmark*

16   *8. Aquatic Ecosystem Research, University of Duisburg-Essen, 45141 Essen, Germany*

17   9. *Programme for Environmental Information, Finnish Environment Institute, Jyväskylä, Finland*

18   *10. School of Engineering, Aarhus University, Inge Lehmannsgade 10, 8000 Aarhus C, Denmark*

19

20   *corresponding author, tth@bios.au.dk, phone: +4587158892

21

22   ORCID: TTH 0000-0001-5387-3284, JÄ 0000-0003-0710-9044, KB 0000-0001-6742-9504, OLPH

23   0000-0002-1598-5733, AI 0000-0003-4807-1345, FL 0000-0002-5465-913X, HMRM 0000-0002-

24   4768-4767, KM 0000-0001-6316-8554, CM 0000-0002-5720-6523, JR 0000-0003-4631-9298

25   **Classification:** Biological sciences (major), Physical sciences (minor)

26

27   **Key words:** Automated monitoring, Ecology, Insects, Image-based identification, Machine learning

28

29   **Significance statement**: Insect populations are challenging to study, but computer vision and deep

30   learning provide opportunities for continuous and non-invasive monitoring of biodiversity around

31   the clock and over entire seasons. These tools can also facilitate the processing of samples in a

32   laboratory setting. Automated imaging in particular can provide an effective way of identifying and

33   counting specimens to measure abundance. We present examples of sensors and devices of

34   relevance to entomology and show how deep learning tools can convert the big data streams into

35   ecological information. We discuss the challenges that lie ahead and identify four focal areas to

36   make deep learning and computer vision game changers for entomology.

37 **ABSTRACT**

38 Most animal species on Earth are insects, and recent reports suggest that their abundance is in

39 drastic decline. Although these reports come from a wide range of insect taxa and regions, the

40 evidence to assess the extent of the phenomenon is still sparse. Insect populations are challenging to

41 study and most monitoring methods are labour intensive and inefficient. Advances in computer

42 vision and deep learning provide potential new solutions to this global challenge. Cameras and

43 other sensors that can effectively, continuously, and non-invasively perform entomological

44 observations throughout diurnal and seasonal cycles. The physical appearance of specimens can

45 also be captured by automated imaging in the lab. When trained on these data, deep learning models

46 can provide estimates of insect abundance, biomass, and diversity. Further, deep learning models

47 can quantify variation in phenotypic traits, behaviour, and interactions. Here, we connect recent

48 developments in deep learning and computer vision to the urgent demand for more cost-efficient

49 monitoring of insects and other invertebrates. We present examples of sensor-based monitoring of

50 insects. We show how deep learning tools can be applied to the big data outputs to derive ecological

51 information and discuss the challenges that lie ahead for the implementation of such solutions in

52 entomology. We identify four focal areas, which will facilitate this transformation: 1) Validation of

53 image-based taxonomic identification, 2) generation of sufficient training data, 3) development of

54 public, curated reference databases, and 4) solutions to integrate deep learning and molecular tools.

55 **INTRODUCTION**

56 We are experiencing a mass extinction of species (1), but data on changes in species diversity and

57 abundance have substantial taxonomic, spatial, and temporal biases and gaps (2, 3). The lack of data

58 holds especially true for insects despite the fact that they represent the vast majority of animal

59 species. A major reason for these shortfalls for insects and other invertebrates is that available

60 methods to study and monitor species and their population trends are antiquated and inefficient (4).

61 Nevertheless, some recent studies have demonstrated alarming rates of insect diversity and

62 abundance loss (5-7). To further explore the extent and causes of these changes, we need efficient,

63 rigorous, and reliable methods to study and monitor insects (4, 8).

64 Data to derive insect population trends are already generated as part of ongoing

65 biomonitoring programs. However, legislative terrestrial biomonitoring, e.g. in the context of the

66 EU Habitat Directive, focuses on a very small subset of individual insect species such as rare

67 butterflies and beetles because the majority of insect taxa are too difficult or too costly to monitor

68 (9). In current legislative aquatic monitoring, benthic invertebrates are commonly used in

69 assessments of ecological status (e.g. the US Clean Water Act, the EU Water Framework Directive,

70 and the EU Marine Strategy Framework Directive). Still, the spatiotemporal and taxonomic extent

71 and resolution in ongoing biomonitoring programs is coarse and does not provide information on

72 the status of the vast majority of insect populations.

73 Molecular techniques such as DNA barcoding and metabarcoding will likely become

74 valuable tools for future insect monitoring based on field collected samples (10, 11), but at the

75 moment high-throughput methods cannot provide reliable abundance estimates (12, 13) leaving a

76 critical need for other methodological approaches. The state-of-the-art in deep learning and

77 computer vision methods and image processing has matured to the point where it can aid or even

78 replace manual observation *in situ* (14) as well as in routine laboratory sample processing tasks

4

79   (15). Image-based observational methods for monitoring of vertebrates using camera traps have

80   undergone rapid development in the past decade (14, 16-18). Similar approaches using cameras and

81   other sensors for investigating diversity and abundance of insects are underway (19, 20). However,

82   despite huge attention in other domains, deep learning is only very slowly beginning to be applied

83   in invertebrate monitoring and biodiversity research (21-25).

84   Deep learning models learn features of a dataset by iteratively training on example

85   data without the need for manual feature extraction (26). In this way, deep learning is qualitatively

86   different from traditional statistical approaches to prediction (27). Deep learning models specifically

87   designed for dealing with images, so called convolutional neural networks (CNNs) can extract the

88   features of various aspects of a set of images or the objects within them, and learn to differentiate

89   among them. There is great potential in automatic detection and classification of insects in video or

90   time-lapse images with trained CNNs (20). As the methods become more refined, they will bring

91   exciting new opportunities for understanding insect ecology and for monitoring (19, 28-31).

92   Here, we argue that deep learning and computer vision can be used to develop novel,

93   high throughput systems for detection, enumeration, classification, and discovery of species as well

94   as for deriving functional traits such as biomass for biomonitoring purposes. These approaches can

95   help solve long standing challenges in ecology and biodiversity research and the pressing issues in

96   insect population monitoring (32, 33). This article has three goals. First, we present sensor-based

97   solutions for observation of invertebrates *in situ* and for specimen-based research in the laboratory,

98   which due to the volume of data generated, use or could benefit from deep learning models to

99   process data. Second, we show how deep learning models can be applied to obtained data streams to

100   derive ecologically relevant information. Last, we outline and discuss four main challenges that lie

101   ahead in the implementation of such solutions for invertebrate monitoring, ecology, and biodiversity

102   research.

103

SENSOR-BASED INSECT MONITORING

105   Sensors are widely used in ecology for gathering peripheral data such as temperature, precipitation,

106   light intensity etc., but have not yet been used much for gathering data on the insects. However,

107   solutions for sensor-based monitoring of insects and other invertebrates in their natural environment

108   are emerging (34). The innovation and development is primarily driven by agricultural research to

109   predict occurrence and abundance of beneficial and pest insect species of economic importance (35-

110   37), to provide more efficient screening of natural products for invasive insect species (38), or to

111   monitor disease vectors such as mosquitos (39, 40). The most commonly used sensors are cameras,

112   radar, and microphones. Such sensor-based monitoring is likely to generate big data, which require

113   efficient solutions for extracting relevant biological information. Deep learning could be a critical

114   tool in this respect. Below, we give examples of image-based approaches to insect monitoring,

115   which we argue has the greatest potential for integration with deep learning. We also describe

116   approaches using other types of sensors, where the integration with deep learning is less well

117   developed, but still could be relevant for detecting and classifying entomological information. We

118   further describe the ongoing efforts in the digitization of natural history collections, which could

119   generate valuable reference data for training and validating deep learning models.

120

Image-based solutions for *in situ* monitoring

122   Some case studies have already used cameras and deep learning methods for detecting single

123   species, such as the pest of the fruits of olive trees *Bactrocera oleae* (41) or for more generic pest

124   detection (42). The pest detection is based on images of insects that have been trapped with either a

125   McPhail-type trap or a trap with pheromone lure and adhesive liner. The images are collected by a

126   microcomputer and transmitted to a remote server where they are analysed. Other solutions have

6

127   embedded a digital camera and a microprocessor that can count trapped individuals in real-time

128   using object-detection based on an optimized deep learning model (37). In both these cases, deep

129   learning networks are trained to recognize and count the number of single pest species. However,

130   there are very few examples of invertebrate biodiversity-related field studies applying deep learning

131   models (23). Early attempts used feature vectors extracted from single perspective images and

132   yielded modest accuracy for 35 classes of moths (43) or used mostly coarse taxonomic resolution

133   (44). We have recently demonstrated that our custom-built time-lapse cameras can record image

134   data from which a deep learning model could accurately estimate local spatial, diurnal, and seasonal

135   dynamics of honey bees and other flower visiting insects (45; Figure 1). Time-lapse cameras are

136   less likely to create observer bias than direct observation and data collection can extend across full

137   diurnal and even seasonal time scales. Cameras can be baited just as traditional light and pheromone

138   traps or placed over ephemeral natural resources such as flowers, fruits, dung, fungi or carrion.

139   Bjerge, *et al.* (46) propose to use an automated light trap to monitor the abundance of moths and

140   other insects attracted to light. The solution is powered by a solar panel, which allows the system to

141   be installed in remote locations (Figure 2). Ultimately, true 'Internet of Things' enabled hardware

142   will make it possible to implement classification algorithms directly on the camera units to provide

143   fully autonomous systems in the field to monitor insects and report detection and classification data

144   back to the user or to online portals in real time (34).

145

146   **Radar, acoustic, and other solutions for *in situ* monitoring**

147   The use of radar technology in entomology has allowed for the study of insects at scales not

148   possible with traditional methods, specifically related to both migratory and non-migratory insects

149   flying at high altitudes (47). Utilizing data from established weather radar networks can provide

150   information at the level of continents (48), while specialized radar technology such as vertical-

151   looking radars (VLRs) can provide finer grained data albeit at a local scale (49). The VLRs can give

152   estimates of biomass and body shape of the detected object, and direction of flight, speed and body

153   orientation can be extracted from the return radar signal (50). However, VLR data provide little

154   information on community structure and conclusive species identification requires aerial trapping

155   (51, 52). Harmonic scanning radars can detect insects flying at low altitudes at a range of several

156   hundred meters, but insects need to be tagged with a radar transponder and must be within line-of-

157   sight (53, 54). Collectively, the use of radar technology in entomology can provide valuable

158   information in insect monitoring, for example on the magnitude of biomass flux stemming from

159   insect migrations (55), but requires validation with conventional monitoring methods (e.g. 56).

160   Bioacoustics is a well-established scientific discipline and acoustic signals have been

161   extensively and widely used in the field of ecology, for example for detecting presence and studying

162   behavior of marine mammals (57) and for bird species identification (58). Jeliazkov, *et al.* (59) used

163   audio recordings to study population trends of Orthoptera at large spatial and temporal scales,

164   demonstrating that bioacoustic techniques have merit in entomological monitoring. Machine

165   learning methods have proven a particularly valuable tool for deciphering noisy audio recordings

166   and detecting the signals of animals. Kiskin, *et al.* (60) demonstrated the use of a CNN to detect the

167   presence of mosquitoes by identifying the acoustic signal of their wingbeats. Other studies have

168   shown that even species classification can be done using machine learning on audio data, for

169   example for birds (58), bats (61), grasshoppers (62), and bees (63). Although, it has been argued

170   that the use of pseudo-acoustic optical sensors rather than actual acoustic sensors is a more

171   promising technology because of the much improved signal-to-noise ratio in these systems, which

172   may be a particularly important point for bioacoustics in entomology (64).

173   Other systems rely on sensor technology to automate the recording of insect activity

174   or even body mass, but without actual consideration of the subsequent processing of the data with

8

175   deep learning methods (65, 66). In (65) they use a sensor-ring of photodiodes and infrared LEDs to

176   detect large and small sized arthropods, including pollinators and pests and achieve a 95% detection

177   accuracy for live microarthropods of three different species in the size range of 0.5 – 1.1 mm. The

178   Edapholog (66) is a low-power monitoring system for real-time detection of soil microarthropods

179   where a pitfall trap is presented. Probe and sensing is based on detection of change in infrared light

180   intensity similar to (65) and it counts the organisms falling into the trap and estimates their body

181   size. The probe is connected via radio signals to a logging devices that transmits the data to a

182   central server for real-time monitoring. Similarly, others have augmented a traditional low-cost

183   trapping methods by implementing optoelectronic sensors and wireless communication to allow for

184   real-time monitoring and reporting (35). Since, such sensors do not produce images that are

185   intuitive to validate, it could be challenging to generate sufficient, validated training data for

186   implementing deep learning models, although such models could still prove useful.

187

188   **Digitizing specimens and natural history collections**

189   There are strong efforts to digitize natural history collections for multiple reasons including the

190   benefits of deep learning applications (67). The need for and benefits of digitizing natural science

191   collections have motivated the foundation of the Distributed System of Scientific Collections

192   Research Infrastructure (DISSCo RI, www.dissco.eu). DISSCo RI strives for the digital unification

193   of all European natural science assets under common curation and access policies and practices.

194   Most existing databases include single view digitisations of pinned specimens (68), while datasets

195   of insect specimens recorded using multiple sensors, 3D models, and databases on living insect

196   specimens are only just emerging (69, 70). The latter could be particularly relevant for deep

197   learning models. There is also a valuable archive of entomological data in herbarium specimens in

198   the form of signs of herbivory (71). The standard digitization of herbarium collections has proven

9

199   suitable for extracting herbivory data using machine learning techniques (72). Techniques to

200   automate digitization techniques will accelerate the development of such valuable databases and

201   enables tools for identification of non-pinned specimens and live insects *in situ* (67). The

202   BIODISCOVER machine (73) is a proposal towards the automatization of creating databases of

203   liquid preserved specimens such as most field collected insects. The process consists of four

204   automatized steps: 1) bin picking of individual insects directly from bulk samples, 2) recording the

205   specimen from multiple angles using high speed imaging, 3) saving the captured data in an

206   optimized way for deep learning algorithm training and further study, and 4) sorting specimens

207   according to size, taxonomic identity or rarity for potential further molecular processing (Figure 3).

208   Digitization efforts should carefully consider how image data of specimens can be leveraged in

209   efforts to develop deep learning models for *in situ* monitoring.

210

211   **POTENTIAL DEEP LEARNING APPLICATIONS IN ENTOMOLOGY**

212   The big data collected by sensor-based insect monitoring as described above requires efficient

213   solutions for transforming the data into biologically relevant information. Preliminary results

214   suggest that deep learning offers a valuable tool in this respect and could further inspire the

215   collection of new types of data (20, 45). Deep learning software, e.g. for ecological applications, is

216   mostly constructed using open source Python libraries and frameworks such as TensorFlow, Keras,

217   PyTorch, and Scikit-learn (24) and prototype implementations are typically publicly available e.g.

218   on www.github.com. This, in turn, makes the latest advances in other fields related to object

219   detection and fine-grained classification available also for entomological research. As such, the

220   existing deep learning toolbox is already available, but will need adaptation to entomology from the

221   domains for which the tools were developed. In the following, we provide a brief description of the

10

222   transformative potential of deep learning related to entomological data stored in images structured

223   around four main applications.

224

Detecting and tracking individuals *in situ*

226   Image-based monitoring of insect abundance and diversity could rapidly become globally

227   widespread as countries make efforts to better understand the severity of the global insect decline

228   and mitigation measures. Similarly, tracking of individual insects *in situ* even for short periods of

229   time holds exciting research potential. For example, by estimating movement speed of individual

230   insects in their natural environments and relating it to observed microclimatic variation, more

231   realistic thermal performance curves can be established and contrasted to traditional lab-derived

232   thermal performance. However, tracking insect in their natural environment is currently a highly

233   challenging task, due to e.g. the cluttered scenes and varying lighting conditions. In computer

234   vision, such tasks are termed 'detection-based online multiple object tracking', and work under a set

235   of assumptions (74). These assumptions include a precise initial detection (initialization) of the

236   objects to be tracked in a scene, a good ability to visually discriminate between the multiple tracked

237   objects, and smooth motion, velocity, and acceleration patterns of the tracked objects (75). The

238   small visual differences among individual insects and frequent hiding behaviour violate the above

239   assumptions. Moreover, current state-of-the-art deep learning models typically use millions of

240   learned parameters and can only run in near real-time with low-resolution video, which constrains

241   the visual discrimination of the targeted objects in the scene. Possible solutions to these challenges

242   include the use of non-linear motion models (76) and the development of compact (77) or

243   compressed (78) deep learning models.

244   If we manage to solve the task of individual tracking of insects it could open the doors

245   for a new individual-based ecology with profound impacts in such research fields as population,

246   behavioural, and thermal ecology as well as conservation biology. Moreover, considering the recent

247   development in low-cost powerful graphical processing units and dedicated artificial intelligence

248   processer suitable for autonomous and embedded systems (e.g. NVIDIA Jetson Nano, Google Coral

249   Edge TPU, and the Intel AI USB stick), it may soon become feasible to detect, track, and decode

250   behaviour of insects in real-time and report information back to the user.

251

252   **Detecting species interactions**

253   Species interactions are critical for the functioning of ecosystems, yet as they are ephemeral and

254   fast, the consequences of a disruption for ecological function is hard to quantify. High temporal

255   resolution image-based monitoring of consumers and resources can allow for a unique

256   quantification of species interactions (79). The use of cameras allows for continuous observations of

257   species and their interactions across entire growing seasons such as insects visiting flowers,

258   defoliation by herbivores, and predation events. There is an urgent need to develop methods to

259   observe and quantify species interactions efficiently and at ecologically relevant spatial and

260   temporal scales (80, 81). To detect such interactions, image recording should be collected at the

261   scales where individuals interact, i.e., by observing interacting individuals at intervals of seconds to

262   minutes, yet they should ideally extend over seasonal and/or multi-annual periods, which at the

263   moment is difficult to fulfil. Our preliminary results have demonstrated an exciting potential to

264   record plant-insect interactions using time-lapse cameras and deep learning (28 and Figure 1).

265

266   **Taxonomic identification**

267   Taxonomic identification can be approached as a deep learning classification problem. Deep

268   learning-based classification accuracies for image-based insect identification of specimens are

269   approaching the accuracy of human experts (82-84). Applications of gradient-weighted class

270 activation mapping can even visualize morphologically important features for CNN classification

271 (84). Classification accuracy is generally much lower when the insects are recorded live in their

272 natural environments (85, 86), but when class confidence is low at the species-level, it may still be

273 possible to confidently classify insects to a coarser taxonomic resolution (87). In recent years,

274 impressive results have been obtained by CNNs (88). They can classify huge image datasets, such

275 as the 1000-class ImageNet dataset at high accuracy and speed (89). With images of >10,000

276 species of plants, classification performance of CNNs is currently much lower than for botanical

277 experts (25), but promising results in distributed training of deep neural networks (90) and federated

278 learning (91, 92) suggest that improvements can be expected.

279   In most ecological communities, it is common for species to be rare. This often results

280 in highly imbalanced datasets, and the number of specimens representing the rarest species could be

281 insufficient for training neural networks (86, 87). As such, advancing the development of

282 algorithms and approaches for improved identification of rare classes is a key challenge for deep

283 learning-based taxonomic identification. Solutions to this challenge could be inspired by class

284 resampling and cost-sensitive training (93) or by multiset feature learning (94, 95). Class

285 resampling aims at balancing the classes by under-sampling the larger classes and/or over-sampling

286 the smaller classes, while cost-sensitive training assigns a higher loss for errors on the smaller

287 classes. In multiset feature learning, the larger classes are split into smaller subsets, which are

288 combined with the smaller classes to form separate training sets. These methods are all used to learn

289 features that can more robustly distinguish the smaller classes. Species identification performance

290 can vary widely, ranging from species which are correctly identified in most cases to species that

291 are generally difficult to identify (96). Typically, the amount of training data is a key element for

292 successful identification, although recent analyses of images of the approximately 65,000

293 specimens in the carabid beetle collection at the Natural History Museum London suggest that

13

294   imbalances in identification performance are not necessarily related to how well-represented a

295   species is in the training data (87). Further work is needed on large datasets to fully understand

296   these challenges.

297   A related challenge is formed by those species that are completely absent from the

298   reference database on which the deep learning models are trained. Detecting such species requires

299   techniques developed for multiple-class novelty/anomaly detection or open set/world recognition

300   (97, 98). A recent survey introduces various open set recognition methods with the two main

301   approaches being discriminative and generative (99). Discriminative models are based on traditional

302   machine learning techniques or deep neural networks with some additional mechanism to detect

303   outliers, while the main idea of generative models is to generate either positive or negative samples

304   for training. However, the current methods are typically applied on relatively small datasets and do

305   not scale well with the number of classes (99). Insect datasets typically have a high number of

306   classes and a very fine-grained distribution, where the phenotypic differences between species may

307   be minute while intra-species differences may be large. Such datasets are especially challenging for

308   open set recognition methods. While it will be extremely difficult to overcome this challenge for all

309   species using only phenotype based identification, combining image-based deep learning and DNA

310   barcoding techniques may help to solve the problem.

311

312   **Estimating biomass from bulk samples**

313   Deep learning models can potentially predict biomass of bulk insect samples in a lab setting.

314   Legislative aquatic monitoring efforts in the United States and Europe require information about the

315   abundance or biomass of individual taxa from bulk invertebrate samples. Using the

316   BIODISCOVER machine, Ärje, *et al.* (73) were able to estimate biomass variation of individual

317   specimens of Diptera species without destroying specimens. This was achieved from geometric

14

318   features of the specimen such as the mean area from multiple images recorded by the

319   BIODISCOVER machine and statistically relating such values to subsequently obtained dry mass

320   from the same specimens. To validate such approaches, it is necessary to have accurate information

321   about the dry mass of a large selection of taxa. In the future, deep learning models may provide

322   even more accurate estimates of biomass. Obtaining specimen-specific biomass information non-

323   destructively from bulk samples is a high priority in routine insect monitoring, since it will enable

324   more extensive insights into insect population and community dynamics and provide better

325   information for environmental management.

326

327   **FUTURE DIRECTIONS**

328   To unlock the full potential of deep learning methods for insect ecology and monitoring, four main

329   challenges need to be addressed with highest priority. We describe each of these below.

330

331   **Validating image-based taxonomic identification**

332   Validation of the detection and identification of species recorded with cameras in the field pose a

333   critical challenge for implementing deep learning tools in entomology. Often it will not be possible

334   to conclusively identify insects from images and validation of image-based species classification

335   should be done using alternative, complimentary techniques. We suggest four approaches to this

336   validation: 1) Obtaining local knowledge about the identity and relative abundance of candidate

337   species, 2) catching and manually identifying insects in the vicinity of a camera trap, 3) identifying

338   insects by environmental DNA analysis of insect DNA traces left e.g. on flowers (100), or 4) by

339   directly observing and catching insects visible to the camera. The first three approaches are indirect

340   and each come with their separate problems such as the difference in trapping efficiency of a time-

341   lapse camera trap and e.g. a pitfall trap placed to capture the same insects. However, the subsequent

342   identification of specimens from pitfall trapping can serve as validation of image-based results and

343   can further help in production of training data for optimizing deep learning models (e.g. by placing

344   specimens back under the camera). DNA techniques be able to validate image-based identification

345   since DNA can give accurate information on species identity (11, 100, 101).

346   For specific purposes, validation of insects can be done through interfaces with online

347   portals and by involving citizen science. With integrated deep learning algorithms, online portals

348   provide instant candidate species when users upload pictures of observed insect species. The most

349   prominent examples of such portals of relevance to insects are the smartphone apps connected to

350   sites such as www.iNaturalist.org and www.observation.org. Another way of using deep learning

351   models to generate data on insect occurrence in their natural environment is by involving the public

352   in the annotation and quality control of images of insects uploaded to citizen science web portals

353   such as www.zooniverse.org (102).

354

355   **Generating training data**

356   One of the main challenges with deep learning is the need for large amounts of training data, which

357   is slow, difficult, and expensive to collect and label. Deep learning models typically require

358   hundreds of training instances of a given species to learn to detect species occurrences against the

359   background (86). In a laboratory setting, the collection of data can be eased by automated imaging

360   devices, such BIODISCOVER described above, which allow imaging large amounts of insects

361   under fixed settings. The imaging of species *in situ* should be done in a wide range of conditions

362   (e.g., background, time of day, and season) to avoid that the model learns a false connection

363   between the species and the background, with resulting lower ability of the model to detect the

364   species against another background. Approaches to alleviate the challenge of moving from one

365   environment to another include multi-task learning (103), style transfer (104), image generation

16

366 (105), or domain adaptation (106). Multi-task learning aims to concurrently learn multiple different

367 tasks (e.g., segmentation, classification, detection) by sharing information leading to better data

368 representations and ultimately better results. Style transfer methods try to impose properties

369 appearing in one set of data to new data. Image generation can be used to created synthetic training

370 images with, for example, varying backgrounds. Domain adaptation aims at tuning the parameters

371 of a deep learning model trained on data following one distribution (source domain) to adapt so that

372 they can provide high performance on new data following another distribution (target domain).

373 The motion detection sensors in wildlife cameras are typically not triggered by insects

374 and species typically only occur in a small fraction of time-lapse images. A key challenge is

375 therefore to detect insects and filter out blank images from images with species of interest (102,

376 107). When it is difficult to obtain sufficient samples of rare insects, Zhong, *et al.* (108) proposed to

377 use deep learning only to detect all species of flying insects as a single class. Subsequently, the fine-

378 grained species classification can be based on manual feature extraction and support vector

379 machines, which is a machine learning technique that requires less training data and solves the

380 problem of insufficient training data.

381 The issue of scarce training data can also be alleviated with new data synthesis. Data

382 synthesis could be used specifically to augment the training set by creating artificial images of

383 segmented individual insects that are placed randomly in scenes with different backgrounds (109).

384 A promising alternative is to use deep learning models for generating artificial images belonging to

385 the class of interest. The most widely approach to date is based on generative adversarial networks

386 (110) and has shown astonishing performance results in computer vision problems in general, as

387 well as in ecological problems (111).

388

17

389 **Building reference databases**

390 Publicly available reference databases are critical for adapting deep learning tools to entomological

391 research. Initiatives like DISSCO RI and IDigBio (https://www.idigbio.org/) are important for

392 enabling the use of museum collections. However, to enable deep learning-based identification,

393 individual open datasets from entomological research and monitoring are also needed (e.g. 85, 96,

394 112). The collation of such research datasets will require dedicated projects as well as large

395 coordinated efforts that drive the open-access and reuse of research data such as the European Open

396 Science Cloud and the Research Data Alliance. Building a large insect reference dataset is laborious

397 and, therefore, it is important to maximize the benefits. To do so, non-collection datasets should

398 also use common approaches and hardware and abide to best practices in metadata and data

399 management (113-115). Further, dataset collectors and deep learning model developers should work

400 closely together and make data accessible. All the possible metadata, such as camera settings and

401 hardware, sampling location, date, and time of day, should be saved for future analysis. Similarly,

402 characteristics of the specimen, such as species identity, biomass, sex, age class, and possibly

403 derived information like dry weight should be recorded if such information exist. In particular,

404 correct labelling of species in images is critical. Using multiple experts and molecular information

405 about species identity to verify the labeling or performing subsequent validity checks through DNA

406 barcoding will improve the data quality and the performance of the deep learning models. This can

407 be done, for instance, by manually verifying the quality and labeling of images that are repeatedly

408 misclassified by the machine learning methods. Standardized imaging devices such as the

409 BIODISCOVER machine could also play a key role in building reference databases from

410 monitoring programs (73). Training classifiers with species that are currently not encountered in a

411 certain region but can possibly spread there later will naturally help to detect such changes when

412 they occur. Integration of such reference databases with field monitoring methods forms an

18

413   important future challenge. As a starting point, we provide a list of open access entomological

414   image databases (SI Appendix).

415

Integration of deep learning and DNA-based tools

417   For processing samples in the lab, molecular methods have gained increasing attention over the past

418   decade, but there are still critical challenges which remain unresolved: specimens are typically

419   destroyed, abundance cannot be accurately estimated, and key specimens cannot be identified in

420   bulk samples. Nevertheless, DNA barcoding is now an established, powerful method to reliably

421   assess biodiversity also in entomology (11). For insects, this works by sequencing a short fragment

422   of the mitochondrial cytochrome-c-oxidase I subunit gene (COI) and comparing the DNA sequence

423   to an available reference database (116). Even undescribed and morphologically cryptic species can

424   be distinguished with this approach (117), which is unlikely to be possible with deep learning. This

425   is of great importance as morphologically similar species can have distinct ecological preferences

426   (118) and thus distinguishing them unambiguously is important for monitoring, ecosystem

427   assessment and conservation biology. However, mass-sequencing based molecular methods cannot

428   provide precise abundance or biomass estimates and assign sequences to individual specimens (12).

429   Therefore, an unparalleled strength lies in combining both image-recognition and DNA

430   metabarcoding approaches: i) When building reference collections for training models for insect

431   classification, species identity can be molecularly verified and potential cryptic species can be

432   separated by the DNA barcode. ii) After image-based species identification of a whole bulk sample,

433   all specimens can be processed via DNA metabarcoding to assess taxonomic resolution at the

434   highest level. A further obvious advantage of linking computer vision and deep learning to DNA is

435   the fact that even in the absence of formal species descriptions, DNA tools can generate distinctly

436   referenced taxonomic assignments via so-called "Barcode-Index-Numbers" (BINs) (119). These

437   BINs provide referenced biodiversity units using the taxonomic backbone of the Barcode of Life

438   Data Systems (https://boldsystems.org) and represent a much greater diversity of even yet

439   undescribed species. For instance, it is typically clear that a new species belongs to the genus

440   *Astraptes* in the butterfly family Hesperiidae, but also that it represents a genetically distinct, new

441   entity (120). These units can also be directly used as part of ecosystem status assessment despite not

442   yet having Linnean names. BINs can be used for model training. Recent studies convincingly show

443   that with this more holistic approach, which includes cryptic and undescribed species, the

444   predictions of environmental status as required by several legislative monitoring programs actually

445   improve substantially (e.g. 121). For cases of cryptic species with great relevance e.g. for

446   conservation biology it is also possible to individually process specimens of a cryptic species

447   complex after automated image-based assignment to further validate identity and frequency of

448   these. Combining deep learning with DNA-based approaches could deliver detailed trait

449   information, biomass, and abundance with the best possible taxonomic resolution.

450

451   **CONCLUSION**

452   Deep learning is currently influencing a wide range of scientific disciplines (88), but has only just

453   begun to benefit entomology. While there is a vast potential for deep learning to transform insect

454   ecology and monitoring, applying deep learning to entomological research questions brings new

455   technical challenges. The complexity of deep learning models and the challenges of entomological

456   data require substantial investment in interdisciplinary efforts to unleash the potential of deep

457   learning in entomology. However, these challenges also represent ample potential for cross-

458   fertilization among the biological and computer sciences. The benefit to entomology is not only

459   more data, but also novel kinds of data. As the deep learning tools become widely available and

460   intuitive to use, they can transform field entomology by providing information that is currently

461   intractable to record by human observations (18, 33, 122). Consequently, there is a bright future for

462   entomology, with new research niches opening up and access to unforeseen scales and resolution of

463   data, vital for biodiversity assessments.

464   The shift towards automated methods may raise concerns about the future for

465   taxonomists, much like the debate concerned with developments in molecular species identification

466   (123, 124). We emphasize that the expertise of taxonomists is at the heart of and critical to these

467   developments. Initially, automated techniques will be used in the most routine-like tasks, which in

468   turn will allow the taxonomic experts to dedicate their focus on the specimens requiring more in

469   depth studies as well as the plethora of new species that need to be described and studied. To enable

470   this, we need to consider approaches that can pinpoint samples for human expert inspection in a

471   meaningful way, e.g., based on neural network classification confidences (82) or additional rare

472   species detectors (125). As deep learning becomes more closely integrated in entomological

473   research, the vision of real-time detection, tracking, and decoding of behaviour of insects could be

474   realized for a transformation of insect ecology and monitoring. In turn, efficient tracking of insect

475   biodiversity trends will aid the design of effective measures to counteract or revert biodiversity loss.

476

REFERENCES

1. G. Ceballos, P. R. Ehrlich, R. Dirzo, Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. *Proc. Natl. Acad. Sci. USA* **114**, E6089-E6096 (2017).

2. M. Dornelas *et al.*, BioTIME: A database of biodiversity time series for the Anthropocene. *Global Ecol. Biogeogr.* **27**, 760-786 (2018).

3. S. A. Blowes *et al.*, The geography of biodiversity change in marine and terrestrial assemblages. *Science* **366**, 339-345 (2019).

4. G. A. Montgomery *et al.*, Is the insect apocalypse upon us? How to find out. *Biol. Conserv.* **241**, 108327 (2020).

5. C. A. Hallmann *et al.*, More than 75 percent decline over 27 years in total flying insect biomass in protected areas. *Plos One* **12**, e0185809 (2017).

6. S. Seibold *et al.*, Arthropod decline in grasslands and forests is associated with landscape-level drivers. *Nature* **574**, 671-674 (2019).

7. R. van Klink *et al.*, Meta-analysis reveals declines in terrestrial but increases in freshwater insect abundances. *Science* **368**, 417-420 (2020).

8. D. L. Wagner, Insect declines in the Anthropocene. *Annu. Rev. Entomol.* **65**, 457-480 (2020).

9. S. Pawar, Taxonomic chauvinism and the methodologically challenged. *Bioscience* **53**, 861-864 (2003).

10. T. W. A. Braukmann *et al.*, Metabarcoding a diverse arthropod mock community. *Mol Ecol Resour* **19**, 711-727 (2019).

11. V. Elbrecht *et al.*, Validation of COI metabarcoding primers for terrestrial arthropods. *Peerj* **7** (2019).

12. V. Elbrecht, F. Leese, Can DNA-based ecosystem assessments quantify species abundance? testing primer bias and biomass-sequence relationships with an innovative metabarcoding protocol. *Plos One* **10** (2015).

13. H. Krehenwinkel *et al.*, Estimating and mitigating amplification bias in qualitative and quantitative arthropod metabarcoding. *Scientific Reports* **7** (2017).

14. H. Yousif, J. Yuan, R. Kays, Z. He, Animal Scanner: Software for classifying humans, animals, and empty frames in camera trap images. *Ecol Evol* **9**, 1578-1589 (2019).

15. J. Ärje *et al.*, Human experts vs. machines in taxa recognition. *Signal Processing: Image Communication* **87**, 115917 (2020).

16. M. S. Norouzzadeh *et al.*, Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *Proc. Natl. Acad. Sci. USA* **115**, E5716-E5725 (2018).

17. N. MacLeod, M. Benfield, P. Culverhouse, Time to automate identification. *Nature* **467**, 154-155 (2010).

18. R. Steenweg *et al.*, Scaling-up camera traps: monitoring the planet's biodiversity with networks of remote sensors. *Front. Ecol. Environ.* **15**, 26-34 (2017).

19. R. Steen, Diel activity, frequency and visit duration of pollinators in focal plants: in situ automatic camera monitoring and data processing. *Methods Ecol Evol* **8**, 203-213 (2017).

20. L. Pegoraro, O. Hidalgo, I. J. Leitch, J. Pellicer, S. E. Barlow, Automated video monitoring of insect pollinators in the field. *Emerging Topics in Life Sciences* 10.1042/etls20190074 (2020).

21. J. Wäldchen, P. Mäder, Machine learning for image based species identification. *Methods Ecol Evol* **9**, 2216-2225 (2018).

22. N. Piechaud, C. Hunt, P. F. Culverhouse, N. L. Foster, K. L. Howell, Automated identification of benthic epifauna with computer vision. *Mar. Ecol. Prog. Ser.* **615**, 15-30 (2019).

23. B. G. Weinstein, A computer vision for animal ecology. *J. Anim. Ecol.* **87**, 533-545 (2018).

24. S. Christin, É. Hervet, N. Lecomte, Applications for deep learning in ecology. *Methods Ecol Evol* **10**, 1632-1644 (2019).

25. A. Joly *et al.* (2019) Overview of LifeCLEF 2019: identification of amazonian plants, South & North American birds, and niche prediction. (Springer International Publishing, Cham), pp 387-401.

26. D. Xia, P. Chen, B. Wang, J. Zhang, C. Xie, Insect detection and classification based on an improved convolutional neural network. *Sensors* **18**, 4169 (2018).

27. D. Bzdok, N. Altman, M. Krzywinski, Statistics versus machine learning. *Nat. Methods* **15**, 233-234 (2018).

28. D. T. Tran, T. T. Høye, M. Gabbouj, A. Iosifidis, IEEE, "Automatic flower and visitor detection system" in 2018 26th European Signal Processing Conference (Eusipco). (2018), 10.23919/EUSIPCO.2018.8553494, pp. 405-409.

29. R. A. Collett, D. O. Fisher, Time-lapse camera trapping as an alternative to pitfall trapping for estimating activity of leaf litter arthropods. *Ecol Evol* **7**, 7527-7533 (2017).

30. I. Ruczyński, Z. Hałat, M. Zegarek, T. Borowik, D. K. N. Dechmann, Camera transects as a method to monitor high temporal and spatial ephemerality of flying nocturnal insects. *Methods Ecol Evol* **11**, 294-302 (2020).

31. S. E. Barlow, M. A. O'Neill, Technological advances in field studies of pollinator ecology and the future of e-ecology. *Current Opinion in Insect Science* **38**, 15-25 (2020).

32. P. Cardoso, T. L. Erwin, P. A. V. Borges, T. R. New, The seven impediments in invertebrate conservation and how to overcome them. *Biol. Conserv.* **144**, 2647-2655 (2011).

33. J. Hortal *et al.*, Seven shortfalls that beset large-scale knowledge of biodiversity. *Annual Review of Ecology, Evolution, and Systematics* **46**, 523-549 (2015).

34. I. Potamitis, P. Eliopoulos, I. Rigakis, Automated remote insect surveillance at a global scale and the internet of things. *Robotics* **6** (2017).

35. I. Potamitis, I. Rigakis, N. Vidakis, M. Petousis, M. Weber, Affordable bimodal optical sensors to spread the use of automated insect monitoring. *Journal of Sensors*, Article ID: 3949415 (2018).

36. D. J. A. Rustia, J.-J. Chao, J.-Y. Chung, T.-T. Lin (2019) An online unsupervised deep learning approach for an automated pest insect monitoring system. in *2019 ASABE Annual International Meeting* (ASABE, St. Joseph, MI), p 1.

37. Y. Sun *et al.*, Automatic in-trap pest detection using deep learning for pheromone-based Dendroctonus valens monitoring. *Biosys. Eng.* **176**, 140-150 (2018).

38. T. M. Poland, D. Rassati, Improved biosecurity surveillance of non-native forest insects: a review of current methods. *J. Pest Sci.* **92**, 37-49 (2019).

39. D. A. A. Santos, L. E. Teixeira, A. M. Alberti, V. Furtado, J. J. P. C. Rodrigues (2018) Sensitivity and noise evaluation of an optoelectronic sensor for mosquitoes monitoring. in *2018 3rd International Conference on Smart and Sustainable Technologies (SpliTech)*, pp 1-5.

40. J. Park, D. I. Kim, B. Choi, W. Kang, H. W. Kwon, Classification and morphological analysis of vector mosquitoes using deep convolutional neural networks. *Scientific Reports* **10**, 1012 (2020).

41. R. Kalamatianos, I. Karydis, D. Doukakis, M. Avlonitis, DIRT: The dacus image recognition toolkit. *Journal of Imaging* **4**, 129 (2018).

580   42.   W. Ding, G. Taylor, Automatic moth detection from trap images for pest management.
581   *Comput. Electron. Agric.* **123**, 17-28 (2016).
582   43.   M. Mayo, A. T. Watson, Automatic species identification of live moths. *Knowledge-Based*
583   *Systems* **20**, 195-202 (2007).
584   44.   J. Wang, C. Lin, L. Ji, A. Liang, A new automatic identification system of insect images at
585   the order level. *Knowledge-Based Systems* **33**, 102-110 (2012).
586   45.   T. T. Høye, H. M. R. Mann, K. Bjerge, Camera-based monitoring of insects on green roofs
587   [in Danish], *Aarhus University, DCE – National Centre for Environment and Energy*, pp.
588   18, Scientific report nr. 371 (2020).
589   46.   K. Bjerge, M. V. Sepstrup, J. B. Nielsen, F. Helsing, T. T. Høye, A light trap and computer
590   vision system to detect and classify live moths (Lepidoptera) using tracking and deep
591   learning. *bioRxiv* 10.1101/2020.03.18.996447, 2020.2003.2018.996447 (2020).
592   47.   J. W. Chapman, V. A. Drake, D. R. Reynolds, Recent insights from radar studies of insect
593   flight. *Annu. Rev. Entomol.* **56**, 337-356 (2011).
594   48.   O. Hueppop *et al.*, Perspectives and challenges for the use of radar in biological
595   conservation. *Ecography* **42**, 912-930 (2019).
596   49.   K. R. Wotton *et al.*, Mass seasonal migrations of hoverflies provide extensive pollination
597   and crop protection services. *Curr. Biol.* **29**, 2167-+ (2019).
598   50.   J. W. Chapman, A. D. Smith, I. P. Woiwod, D. R. Reynolds, J. R. Riley, Development of
599   vertical-looking radar technology for monitoring insect migration. *Comput. Electron. Agric.*
600   **35**, 95-110 (2002).
601   51.   J. W. Chapman, D. R. Reynolds, A. D. Smith, Migratory and foraging movements in
602   beneficial insects: a review of radar monitoring and tracking methods. *Int. J. Pest Manage.*
603   **50**, 225-232 (2004).
604   52.   J. W. Chapman *et al.*, High-altitude migration of the diamondback moth Plutella xylostella
605   to the UK: a study using radar, aerial netting, and ground trapping. *Ecol. Entomol.* **27**, 641-
606   650 (2002).
607   53.   W. D. Kissling, D. E. Pattemore, M. Hagen, Challenges and prospects in the telemetry of
608   insects. *Biological Reviews* **89**, 511-530 (2014).
609   54.   R. Maggiora, M. Saccani, D. Milanesio, M. Porporato, An innovative harmonic radar to
610   track flying insects: the case of *Vespa velutina*. *Scientific Reports* **9**, 11964 (2019).
611   55.   G. Hu *et al.*, Mass seasonal bioflows of high-flying insect migrants. *Science* **354**, 1584-1587
612   (2016).
613   56.   P. M. Stepanian *et al.*, Declines in an abundant aquatic insect, the burrowing mayfly, across
614   major North American waterways. *Proc. Natl. Acad. Sci. USA* **117**, 2987-2992 (2020).
615   57.   A. K. Stimpert, W. W. L. Au, S. E. Parks, T. Hurst, D. N. Wiley, Common humpback whale
616   (*Megaptera novaeangliae*) sound types for passive acoustic monitoring. *J. Acoust. Soc. Am.*
617   **129**, 476-482 (2011).
618   58.   J. Salamon, J. P. Bellol, A. Farnsworth, S. Kelling, Fusing shallow and deep learning for
619   bioacoustic bird species classification. *2017 Ieee International Conference on Acoustics,*
620   *Speech and Signal Processing (Icassp)*, 141-145 (2017).
621   59.   A. Jeliazkov *et al.*, Large-scale semi-automated acoustic monitoring allows to detect
622   temporal decline of bush-crickets. *Global Ecology and Conservation* **6**, 208-218 (2016).
623   60.   I. Kiskin *et al.*, Bioacoustic detection with wavelet-conditioned convolutional neural
624   networks. *Neural Computing & Applications* **32**, 915-927 (2020).
625   61.   O. Mac Aodha *et al.*, Bat detective - deep learning tools for bat acoustic signal detection.
626   *PLoS Comp. Biol.* **14** (2018).

627 62. E. D. Chesmore, E. Ohya, Automated identification of field-recorded songs of four British
628   grasshoppers using bioacoustic signal recognition. *Bull. Entomol. Res.* **94**, 319-330 (2004).
629 63. S. Kawakita, K. Ichikawa, Automated classification of bees and hornet using acoustic
630   analysis of their flight sounds. *Apidologie* **50**, 71-79 (2019).
631 64. Y. P. Chen, A. Why, G. Batista, A. Mafra-Neto, E. Keogh, Flying insect classification with
632   inexpensive sensors. *J. Insect Behav.* **27**, 657-677 (2014).
633 65. E. Balla *et al.*, An opto-electronic sensor-ring to detect arthropods of significantly different
634   body sizes. *Sensors* **20**, 982 (2020).
635 66. M. Dombos *et al.*, EDAPHOLOG monitoring system: automatic, real-time detection of soil
636   microarthropods. *Methods Ecol Evol* **8**, 313-321 (2017).
637 67. B. P. Hedrick *et al.*, Digitization and the future of natural history collections. *Bioscience* **70**,
638   243-251 (2020).
639 68. V. Blagoderov, I. J. Kitching, L. Livermore, T. J. Simonsen, V. S. Smith, No specimen left
640   behind: industrial scale digitization of natural history collections. *Zookeys*
641   10.3897/zookeys.209.3178, 133-146 (2012).
642 69. B. Ströbel, S. Schmelzle, N. Blüthgen, M. Heethoff, An automated device for the
643   digitization and 3D modelling of insects, combining extended-depth-of-field and all-side
644   multi-view imaging. *ZooKeys* **759** (2018).
645 70. A. E. Z. Short, T. Dikow, C. S. Moreau, Entomological collections in the age of big data.
646   *Annu. Rev. Entomol.* **63**, 513-530 (2018).
647 71. E. K. Meineke, T. J. Davies, Museum specimens provide novel insights into changing plant–
648   herbivore interactions. *Philosophical Transactions of the Royal Society B: Biological
649   Sciences* **374**, 20170393 (2019).
650 72. E. K. Meineke, C. Tomasi, S. Yuan, K. M. Pryer, Applying machine learning to investigate
651   long-term insect–plant interactions preserved on digitized herbarium specimens.
652   *Applications in Plant Sciences* **8**, e11369 (2020).
653 73. J. Ärje *et al.*, Automatic image-based identication and biomass estimation of invertebrates.
654   *Methods Ecol Evol* 10.1111/2041-210X.13428 (2020).
655 74. Q. Wang, L. Zhang, L. Bertinetto, W. Hu, P. H. S. Torr (2019) Fast online object tracking
656   and segmentation: a unifying approach. in *IEEE Conference on Computer Vision and
657   Pattern Recognition*.
658 75. W. Luo, X. Zhao, T.-K. Kim, Multiple object tracking: a review. *ArXiv* **abs/1409.7618**
659   (2014).
660 76. B. Yang, R. Nevatia, Multi-target tracking by online learning of non-linear motion patterns
661   and robust appearance models. *2012 Ieee Conference on Computer Vision and Pattern
662   Recognition (Cvpr)*, 1918-1925 (2012).
663 77. D. T. Tran, S. Kiranyaz, M. Gabbouj, A. Iosifidis, Heterogeneous multilayer generalized
664   operational perceptron. *IEEE Transactions on Neural Networks and Learning Systems*
665   10.1109/TNNLS.2019.2914082, 1-15 (2019).
666 78. D. T. Tran, A. Iosifidis, M. Gabbouj, Improving efficiency in convolutional neural networks
667   with multilinear filters. *Neural Networks* **105**, 328-339 (2018).
668 79. S. Hamel *et al.*, Towards good practice guidance in using camera-traps in ecology: influence
669   of sampling design on validity of ecological inferences. *Methods Ecol Evol* **4**, 105-113
670   (2013).
671 80. L. Estes *et al.*, The spatial and temporal domains of modern ecology. *Nature Ecology &
672   Evolution* **2**, 819-826 (2018).
673 81. A. Valiente-Banuet *et al.*, Beyond species loss: the extinction of ecological interactions in a
674   changing world. *Funct. Ecol.* **29**, 299-307 (2015).

25

675 82. J. Raitoharju, K. Meissner (2019) On confidences and their use in (semi-)automatic multi-
676   image taxa identification. in *2019 IEEE Symposium Series on Computational Intelligence*
677   *(SSCI)*, pp 1338-1343.
678 83. M. Valan, K. Makonyi, A. Maki, D. Vondráček, F. Ronquist, Automated taxonomic
679   identification of insects with expert-level accuracy using effective feature transfer from
680   convolutional networks. *Syst. Biol.* **68**, 876-895 (2019).
681 84. D. Milošević *et al.*, Application of deep learning in aquatic bioassessment: Towards
682   automated identification of non-biting midges. *Sci. Total Environ.* **711**, 135160 (2020).
683 85. X. Wu, C. Zhan, Y. Lai, M. Cheng, J. Yang (2019) IP102: a large-scale benchmark dataset
684   for insect pest recognition. in *2019 IEEE/CVF Conference on Computer Vision and Pattern*
685   *Recognition (CVPR)*, pp 8779-8788.
686 86. G. V. Horn *et al.*, The iNaturalist challenge 2017 dataset. *ArXiv* **abs/1707.06642** (2017).
687 87. O. L. P. Hansen *et al.*, Species-level image classification with convolutional neural network
688   enables insect identification from habitus images. *Ecol Evol* **10**, 737-747 (2020).
689 88. Y. LeCun, Y. Bengio, G. Hinton, Deep learning. *Nature* **521**, 436 (2015).
690 89. K. He, X. Zhang, S. Ren, J. Sun (2015) Delving deep into rectifiers: surpassing human-level
691   performance on ImageNet classification. in *2015 IEEE International Conference on*
692   *Computer Vision (ICCV)*, pp 1026-1034.
693 90. J. Dean *et al.* (2012) Large scale distributed deep networks. in *Proceedings of the 25th*
694   *International Conference on Neural Information Processing Systems - Volume 1* (Curran
695   Associates Inc., Lake Tahoe, Nevada), pp 1223–1231.
696 91. H. B. McMahan, E. Moore, D. Ramage, S. Hampson, B. A. y. Arcas (2016)
697   Communication-efficient learning of deep networks from decentralized data. in *Proceedings*
698   *of the 20th International Conference on Artificial Intelligence and Statistics*.
699 92. K. Bonawitz *et al.*, Towards federated learning at scale: system design. *ArXiv*
700   **abs/1902.01046** (2019).
701 93. C. Huang, Y. N. Li, C. C. Loy, X. O. Tang, Learning deep representation for imbalanced
702   classification. *2016 Ieee Conference on Computer Vision and Pattern Recognition (Cvpr)*
703   10.1109/Cvpr.2016.580, 5375-5384 (2016).
704 94. F. Wu, X. Y. Jing, S. G. Shan, W. M. Zuo, J. Y. Yang, Multiset feature learning for highly
705   imbalanced data classification. *Thirty-First Aaai Conference on Artificial Intelligence*,
706   1583-1589 (2017).
707 95. X. Jing *et al.*, Multiset feature learning for highly imbalanced data classification. *IEEE*
708   *Transactions on Pattern Analysis and Machine Intelligence* 10.1109/TPAMI.2019.2929166,
709   1-1 (2019).
710 96. J. Raitoharju *et al.*, Benchmark database for fine-grained image classification of benthic
711   macroinvertebrates. *Image Vision Comput.* **78**, 73-83 (2018).
712 97. M. Turkoz, S. Kim, Y. Son, M. K. Jeong, E. A. Elsayed, Generalized support vector data
713   description for anomaly detection. *Pattern Recognition* **100**, 107119 (2020).
714 98. P. Perera, V. M. Patel (2019) Deep transfer learning for multiple class novelty detection. in
715   *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp
716   11536-11544.
717 99. C. Geng, S. Huang, S. Chen, Recent advances in open set recognition: a survey. *IEEE*
718   *Transactions on Pattern Analysis and Machine Intelligence* 10.1109/TPAMI.2020.2981604,
719   1-1 (2020).
720 100. P. F. Thomsen, E. E. Sigsgaard, Environmental DNA metabarcoding of wild flowers reveals
721   diverse communities of terrestrial arthropods. *Ecol Evol* **9**, 1665-1679 (2019).

722 101. M. F. Geiger *et al.*, Testing the global malaise trap program – how well does the current
723   barcode reference library identify flying insects in Germany? *Biodiversity Data Journal* **4**
724   (2016).
725 102. M. Willi *et al.*, Identifying animal species in camera trap images using deep learning and
726   citizen science. *Methods Ecol Evol* **10**, 80-91 (2019).
727 103. A. Kendall, Y. Gal, R. Cipolla, Multi-task learning using uncertainty to weigh losses for
728   scene geometry and semantics. *2018 Ieee/Cvf Conference on Computer Vision and Pattern*
729   *Recognition (Cvpr)* 10.1109/Cvpr.2018.00781, 7482-7491 (2018).
730 104. L. A. Gatys, A. S. Ecker, M. Bethge, Image style transfer using convolutional neural
731   networks. *2016 Ieee Conference on Computer Vision and Pattern Recognition (Cvpr)*
732   10.1109/Cvpr.2016.265, 2414-2423 (2016).
733 105. J. M. Bao, D. Chen, F. Wen, H. Q. Li, G. Hua, CVAE-GAN: fine-grained image generation
734   through asymmetric training. *2017 Ieee International Conference on Computer Vision (Iccv)*
735   10.1109/Iccv.2017.299, 2764-2773 (2017).
736 106. E. Tzeng, J. Hoffman, K. Saenko, T. Darrell, Adversarial discriminative domain adaptation.
737   *30th Ieee Conference on Computer Vision and Pattern Recognition (Cvpr 2017)*
738   10.1109/Cvpr.2017.316, 2962-2971 (2017).
739 107. P. Glover-Kapfer, C. A. Soto-Navarro, O. R. Wearn, Camera-trapping version 3.0: current
740   constraints and future priorities for development. *Remote Sensing in Ecology and*
741   *Conservation* **5**, 209-223 (2019).
742 108. Y. Zhong, J. Gao, Q. Lei, Y. Zhou, A vision-based counting and recognition system for
743   flying insects in intelligent agriculture. *Sensors* **18**, 1489 (2018).
744 109. H. Inoue, Data augmentation by pairing samples for images classification. *ArXiv*
745   **abs/1801.02929** (2018).
746 110. I. J. Goodfellow *et al.*, Generative adversarial networks. *ArXiv* **abs/1406.2661** (2014).
747 111. C.-Y. Lu, D. J. Arcega Rustia, T.-T. Lin, Generative adversarial network based image
748   augmentation for insect pest classification enhancement. *IFAC-PapersOnLine* **52**, 1-5
749   (2019).
750 112. M. Martineau *et al.*, A survey on image-based insect classification. *Pattern Recognition* **65**,
751   273-284 (2017).
752 113. T. Forrester *et al.*, An open standard for camera trap data. *Biodiversity Data Journal* **4**
753   (2016).
754 114. L. Scotson *et al.*, Best practices and software for the management and sharing of camera trap
755   data for small and large scales studies. *Remote Sensing in Ecology and Conservation* **3**, 158-
756   172 (2017).
757 115. A. Nieva de la Hidalga, M. van Walsun, P. Rosin, X. Sun, A. Wijers, Quality management
758   methodologies for digitisation operations, pp. 89, 10.5281/zenodo.3469521 (2019).
759 116. S. Ratnasingham, P. D. N. Hebert, BOLD: the barcode of life data system
760   (www.barcodinglife.org). *Mol. Ecol. Notes* **7**, 355-364 (2007).
761 117. M. Hajibabaei, D. H. Janzen, J. M. Burns, W. Hallwachs, P. D. N. Hebert, DNA barcodes
762   distinguish species of tropical Lepidoptera. *Proc. Natl. Acad. Sci. USA* **103**, 968-971 (2006).
763 118. J. N. Macher *et al.*, Multiple-stressor effects on stream invertebrates: DNA barcoding
764   reveals contrasting responses of cryptic mayfly species. *Ecol. Indicators* **61**, 159-169
765   (2016).
766 119. S. Ratnasingham, P. D. N. Hebert, A DNA-based registry for all animal species: the barcode
767   index number (BIN) system. *PLOS ONE* **8**, e66213 (2013).

768 120. P. D. N. Hebert, E. H. Penton, J. M. Burns, D. H. Janzen, W. Hallwachs, Ten species in one:
769   DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptes
770   fulgerator. *Proc. Natl. Acad. Sci. USA* **101**, 14812-14817 (2004).
771 121. T. Cordier *et al.*, Supervised machine learning outperforms taxonomy-based environmental
772   DNA metabarcoding applied to biomonitoring. *Mol Ecol Resour* **18**, 1381-1391 (2018).
773 122. A. C. Burton *et al.*, REVIEW: Wildlife camera trapping: a review and recommendations for
774   linking surveys to ecological processes. *J. Appl. Ecol.* **52**, 675-685 (2015).
775 123. M. G. Kelly, S. C. Schneider, L. King, Customs, habits, and traditions: the role of
776   nonscientific factors in the development of ecological assessment methods. *Wiley*
777   *Interdisciplinary Reviews-Water* **2**, 159-165 (2015).
778 124. F. Leese *et al.*, Why we need sustainable networks bridging countries, disciplines, cultures
779   and generations for aquatic biomonitoring 2.0: a perspective derived from the DNAqua-Net
780   COST action. *Next Generation Biomonitoring, Pt 1* **58**, 63-99 (2018).
781 125. F. Sohrab, J. Raitoharju, Boosting rare benthic macroinvertebrates taxa identification with
782   one-class classification. *ArXiv* **abs/2002.10420** (2020).
783

784

785 **FIGURE LEGENDS**

786 **Figure 1**

787 We developed and tested a camera trap for monitoring flower visiting insects, which records images

788 at fixed intervals (45). (A) The setup consist of two web cameras connected to a control unit

789 containing a Raspberry Pi computer and a hard drive. In our test, ten camera traps were mounted on

790 custom built steel rod mounts c. 30cm above a green roof mix of plants in the genus *Sedum*. Images

791 were recorded every 30 sec during the entire flowering season. After training a convolutional neural

792 network (Yolo3), we detected >100,000 instances of pollinators over the course of an entire

793 growing season. (B) An example image from one of the cameras showing a scene consisting of

794 different flowering species. The locations of the insect detections varied greatly among three

795 common flower visiting species (C) the European honey bee (*Apis mellifera)*, (D) the red-tailed

796 bumblebee (*Bombus lapidarius*), and (E) the marmalade hoverfly (*Episyrphus balteatus*). Across

797 the ten camera traps, the deep learning model detected detailed variation in (F) seasonal and (G)

798 diurnal variation in the occurrence frequency among the same three species. Figure adapted with

799 permission from (45).

800

801 **Figure 2**

802 (A) To automatically monitor nocturnal moth species, we designed a light trap with an on-board

803 computer vision system (46). The light trap is equipped with three different light sources. A

804 fluorescent tube to attract moths, a light table covered by a white sheet to provide a diffuse

805 background illumination of the resting insects, and a light ring to illuminate the specimens. The

806 system is able to attract moths and automatically capture images based on motion detection. The

807 trap is designed using standard components such as a high-resolution USB web camera and a

808 Raspberry Pi computer. (B) We have proposed a computer vision algorithm that, during offline

29

809   processing of the captured images, performs tracking and counting of individual moths. A

810   customized convolutional neural network was trained to detect and classify eight different moth

811   species. The algorithm can run on the on-board computer to allow the system to automatically

812   process and submit species data via a modem to a server. The system works off grid due to a battery

813   and solar panel.

814

815   **Figure 3**

816   The BIODISCOVER machine, can automate the process of invertebrate sample sorting, species

817   identification, and biomass estimation (73). (A) The imaging system consists of an ethanol-filled

818   spectroscopic cuvette, a powerful and adjustable light source and two cameras capable of recording

819   images at 50 frames per second (B) The setup is mounted in a light proof aluminium box and fitted

820   with a pump for refilling the spectroscopic cuvette. (C) Each specimen is imaged from two angles

821   by the cameras as it is dropped into the ethanol-filled cuvette and geometric features related to size

822   and biomass are computed automatically. (D) The system has a built in flushing mechanism for

823   controlling which specimens should be kept together for subsequent storage or analysis. The results

824   for an initial dataset of images of 598 specimens across 12 species of known identity was very

825   promising with a classification accuracy of 98.0% (73). The system is generic and can easily be

826   used for other groups of invertebrates as well. As such, the BIODISCOVER machine pave the way

827   for cheap, fast, and accurate data on spatial and temporal variation in invertebrate abundance,

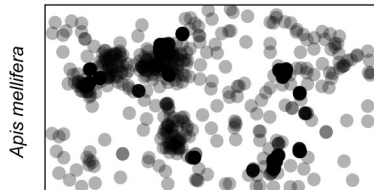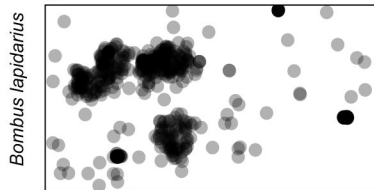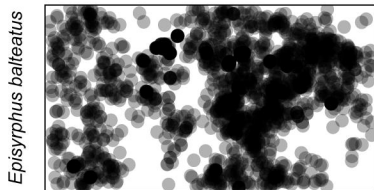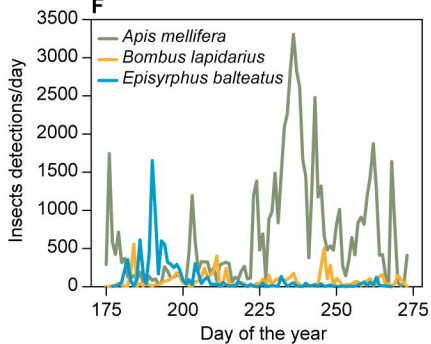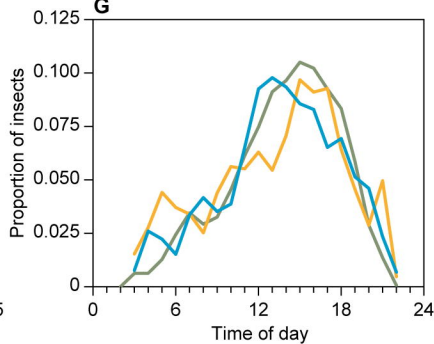828   diversity and biomass. Figure adapted with permission from (73).
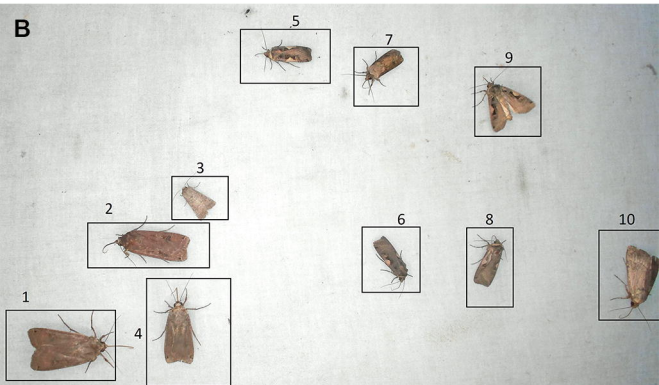
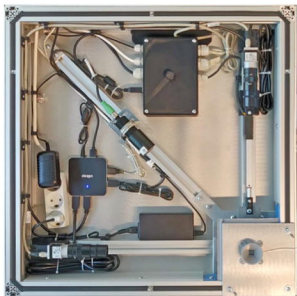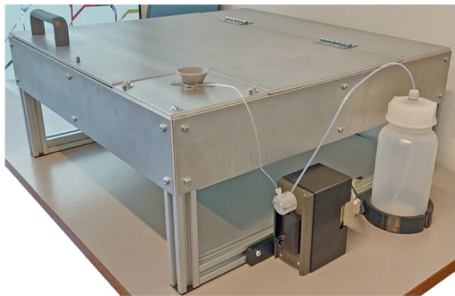829 **TABLE 1**

830 **Glossary**

831 • **Bin picking**: an industrial term for robots that pick up one of many objects randomly placed in a

832 container.

833 • **Convolutional Neural Network (CNN)**: a deep learning algorithm in the family of neural

834 networks with serval different layers commonly applied for image recognition and

835 classification. A CNN can be trained to recognize various objects and patterns in an image.

836 There are four main different operations in a CNN: convolution, activation functions, sub

837 sampling, and fully connected layer. During training the learnable parameters of each

838 convolutional and fully connected layer are adjusted so the CNN is able to recognize different

839 patterns of the training data and used for final image classification.

840 • **Data augmentation**: a technique that can be used to artificially expand the size of a training

841 dataset by creating modified images with objects of interest for classification.

842 • **Machine learning**: a subset of artificial intelligence associated with creating algorithms that can

843 change themselves without human intervention to get the desired result – by feeding themselves

844 through structured data.

845 • **Deep learning**: a subset of machine learning where algorithms are created and function

846 similarly to machine learning, but where there are many levels of these algorithms, each

847 providing a different interpretation of the data it conveys.

848 • **DNA barcoding**: Identification of a species using a short, standardised gene fragment.

849 • Initialization: description of an object to be tracked.

850 • **Training data**: classified images (e.g. images of known species identified by experts) that are

851 recorded to train a deep learning model.

852 • **Precision**: the number of true positives divided by the sum of true positives and false positives

31

853 • **Recall**: also called the true positive rate, is the number of true positives divided by the sum of

854 true positives and false negatives.

855 • **Classification accuracy**: the sum of true positives and true negatives divided by the total

856 number of specimens.

857

A

B

C *Apis mellifera*

D *Bombus lapidarius*

E *Episyrphus balteatus*

F
Insects detections/day
— *Apis mellifera*
— *Bombus lapidarius*
— *Episyrphus balteatus*
Day of the year

G
Proportion of insects
Time of day

A

B

5
7
9
3
2
6
8
10
1
4

A

B

C

D