

Ethoflow: computer vision and artificial intelligence-based software for automatic behavior analysis

Rodrigo Cupertino Bernardes¹, Maria Augusta Pereira Lima², Raul Narciso C. Guedes¹, Gustavo Ferreira Martins³

¹Departamento de Entomologia, Universidade Federal de Viçosa, Viçosa, MG 36570-900, Brazil

²Departamento de Biologia Animal, Universidade Federal de Viçosa, Viçosa, MG 36570-900, Brazil

³Departamento de Biologia Geral, Universidade Federal de Viçosa, Viçosa, MG 36570-900, Brazil

* Corresponding author: R.C. Bernardes;

E-mail addresses: rodrigo.bernardes@ufv.br

ORCID: <https://orcid.org/0000-0001-9481-036X>

Abstract

1. Manual monitoring of animal behavior is time-consuming and prone to bias. An alternative to such limitations is the use of computational resources in behavioral assessments, such as a tracking system, to facilitate accurate and long-term evaluations. There is a demand for robust software that addresses analysis in heterogeneous environments (such as in field conditions) and evaluates multiple individuals in groups while maintaining their identities.

2. The Ethoflow software was developed using computer vision and artificial intelligence (AI) tools to automatically monitor various behavioral parameters. A state-of-the-art object detection algorithm based on instance segmentation was implemented, allowing behavior monitoring in the field under heterogeneous environments. Moreover, a convolutional neural network was implemented to assess complex behaviors, thus expanding the possibilities of animal behavior analyses.

3. The heuristics used to automatically generate training data for the AI models are described, and the models trained with these datasets exhibited high accuracy in detecting individuals in heterogeneous

environments and assessing complex behavior. Ethoflow was employed for kinematic assessments and to detect trophallaxis in social bees. The software runs on the Linux, Microsoft Windows, and IOS operating systems with an intuitive graphical interface.

4. In the Ethoflow algorithm, the processing with AI is separate from the other modules, which facilitates kinematic measurements on an ordinary computer and the assessment of complex behavior on machines with graphics processing units (GPUs). Thus, Ethoflow is a useful support tool for applications in biology and related fields.

Keywords: animal monitoring, convolutional neural networks, deep learning, machine learning, object detection, tracking

1 Introduction

Behavioral studies are critical to understanding the fundamental aspects of animal ecology (Anderson & Perona, 2014; Dell et al., 2014). The assessment of animal behavior by visual inspection is limited and subjective, and does not allow observations over long periods (Noldus, Spink, & Tegelenbosch, 2002). The use of computational tools in behavioral assessments, such as automatic tracking systems, allows accurate and long-term evaluations of animals (Dell et al., 2014; Valletta, Torney, Kings, Thornton, & Madden, 2017). Calculation of important kinematic measurements, including the tracked distance, is feasible when tracking animals, and the evaluation of complex behaviors can provide relevant insights about animal biology. For example, the evaluation of complex behaviors among social insects, such as changes in grooming and trophallaxis, is important for understanding their response to stress agents such as pesticides (Gandra, Amaral, Couceiro, Della Lucia, & Guedes, 2016; Boff, Friedel, Mussury, Lenis, & Raizer, 2018).

Robust systems are needed for animal monitoring in heterogeneous environments, such as in the field (Dell et al., 2014). The greatest challenge in heterogeneous environments involves the extraction of target objects from the background (segmentation) (Zou, Shi, Guo, & Ye, 2019). Animal tracking software operates by background subtraction or thresholding (Yamanaka & Takeuchi, 2018; Sridhar, Roche, & Gingsins, 2019). As these approaches require video recordings with sufficient contrast between the object and homogeneous background, they are not applicable in the field. Software using artificial intelligence (AI) can be sufficiently robust for assessments in heterogeneous environments, as AI models can be trained to learn the detection of target objects in different environments (He, Gkioxari, Dollár, & Girshick, 2018).

Convolutional neural networks (CNNs) are the most conventional AI models employed in computer vision tasks. While these models have exhibited outstanding performance in computer vision applications and tracking software, including idtracker.ai and DeepLabCut, which use AI models in their algorithms, they still exhibit some limitations. The idtracker.ai software applies a CNN to maintain the identity of individuals in a group, but its application is limited to homogeneous environments (Romero-Ferrero, Bergomi, Hinz, Heras, & de Polavieja, 2019). While DeepLabCut uses a CNN for animal pose estimation and tracks parts of objects in heterogeneous environments (Nath et al., 2019), it does not maintain the identity of individuals in a group.

Given the potential applications of AI and the demand for a robust system that fulfills the requirement for studying animal behavior (Dell et al., 2014), the Ethoflow software was developed. AI was incorporated into Ethoflow for object detection, enabling evaluations of the complex behaviors of individuals in groups living in heterogeneous environments. To validate Ethoflow, bioassays with two species of eusocial bees were performed. In addition, parameters associated with the performance of the Ethoflow software were evaluated during its execution.

74 2 Software features

75 The Ethoflow software runs on Linux, Microsoft Windows, and IOS operating systems. This
76 software is registered with the Brazilian National Institute of Intellectual Property (Instituto Nacional de
77 Propriedade Industrial, INPI, Ministério da Economia, Brazil, reg. no. BR 51 2020 000737-6). Ethoflow
78 is user-friendly and does not require the use of line commands because of the intuitive graphical user
79 interface (GUI) (Fig. 1A). There are three tabs (Settings, Analyses, and Deep analyses) on the Ethoflow
80 GUI. Under the ‘Settings’ tab, the parameters can be set. Once the parameters have been set, they can be
81 saved as a protocol (as a .txt file) and loaded for the next step. When setting the parameters, the interface
82 allows real-time monitoring of the effects of parameter changes; for example, detected objects (i.e.,
83 animals) will be marked with a red mask. A region of interest (Fig. 1A; blue box on the bottom right) can
84 also be defined to assess how long the individuals stayed in this region. The ‘Analyses’ tab corresponds
85 to the video processing step and renders images to train the instance segmentation model **Erro! Fonte de**
86 **referência não encontrada..** During video processing, the user can visually monitor the processing in
87 real time. Following completion of processing, the program prints the processing speed and detection
88 rate on the GUI. Under the ‘Deep analyses’ tab, models for complex behavior analysis are loaded, which
89 allows the recognition of several specific behaviors on the condition that an AI model is set up. The
90 median and standard deviation of some measurements of the body of individuals (for e.g., area and
91 length) can also be calculated. These parameters can serve as a basis for generating labeled images to
92 train the specific behavior model and assist in the protocol definition.

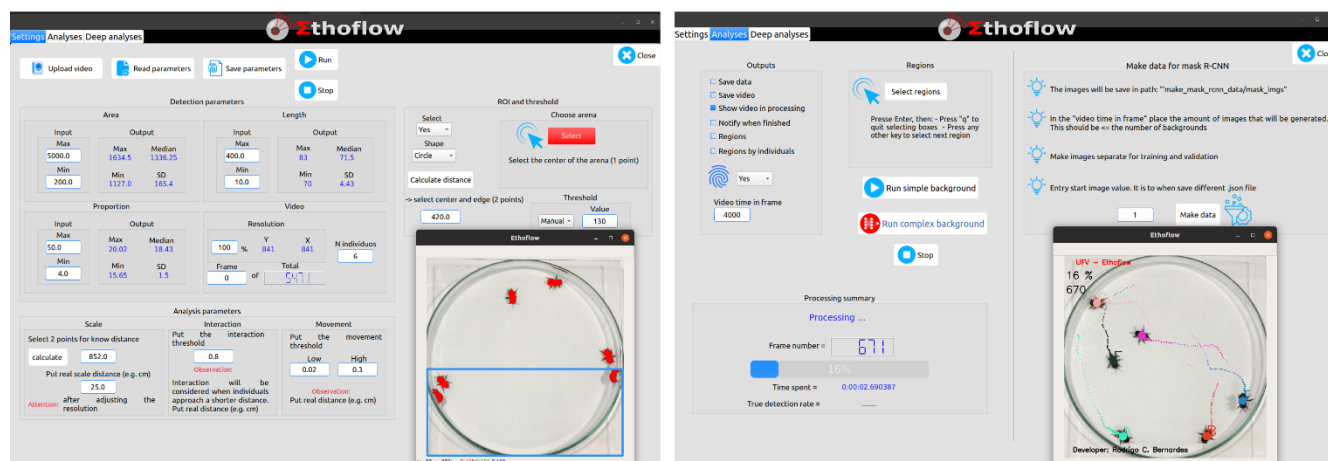
93 In the Ethoflow algorithm, multi-threaded reading was implemented. This avoids the delay
94 between frame reading and other processing steps of the algorithm, whereby frames are always available,
95 thus making it possible to obtain high rates in frames per second (fps) (Supporting Information A.1). In
96 addition, the AI processing is separate from the other modules in the Ethoflow algorithm. Therefore,

97 Ethoflow enables kinematic measurements on an ordinary computer and the assessment of more complex
98 behavior using a GPU.

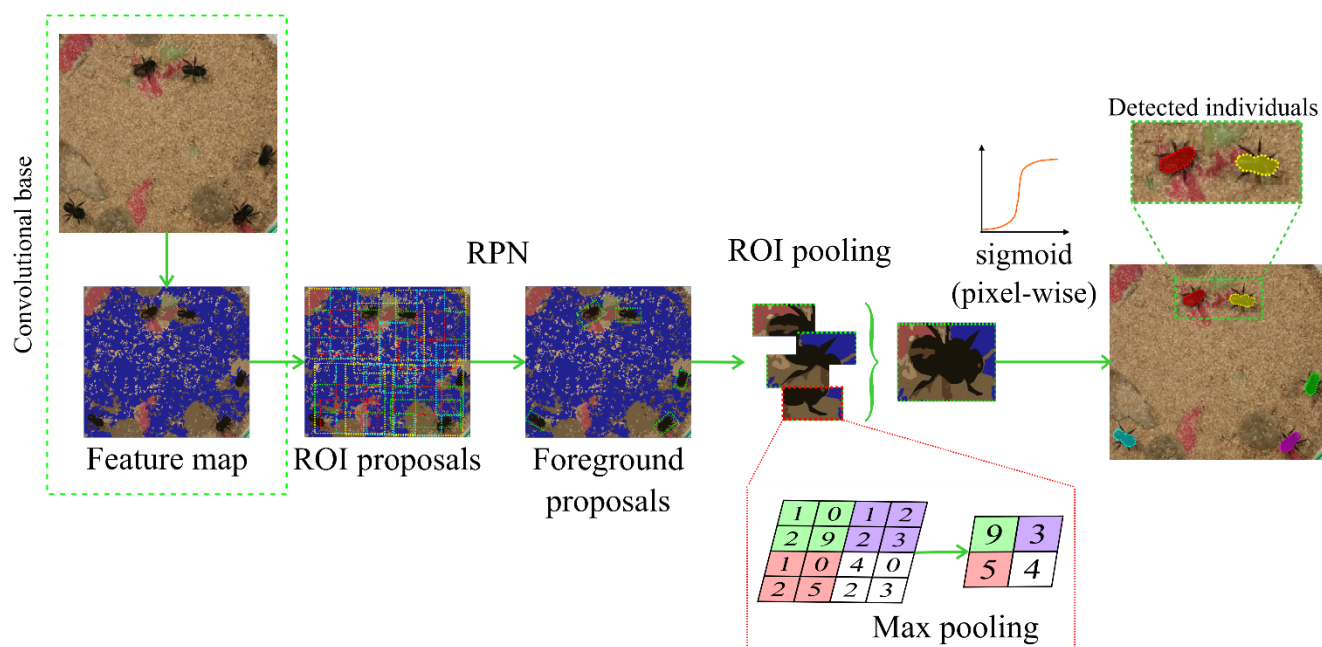
99 Manual and automatic thresholds were applied in Ethoflow to detect individuals. In manual
100 thresholding, the threshold value is defined by the user. One of the automatic thresholding options is
101 based on Otsu's method (Otsu, 1979), wherein the optimal threshold minimizes the within-class variance
102 (Supporting Information A.3). The other automatic thresholding option involves instance segmentation
103 (IS), which is based on the Mask R-CNN AI model (He et al., 2018). This state-of-the-art model for
104 object detection allows Ethoflow to work in field conditions, detecting individuals in a heterogeneous
105 background (Fig 1. B). The hyperparameters and architecture of the IS model are based on the
106 implementation reported by Abdulla (2017) (Supporting Information A.3). The IS model should be
107 trained to learn to detect the animal of interest. Thus, a heuristic was used to automatically generate the
108 training data. This heuristic functions by extracting the contours of individuals in a homogeneous
109 background and subsequently pasting them in a heterogeneous background. (Supporting Information
110 A.3).

111 Ethoflow monitors animals in groups, maintaining the identity of individuals. In this step of the
112 algorithm, the nonhierarchical clustering k-means is applied to separate merged individuals.
113 Subsequently, a combinatorial optimization algorithm is applied, which provides the optimal identity
114 assignment between individuals (Supporting Information A.5). Ethoflow records the coordinates of
115 movement in time and calculates various kinematic parameters associated with the behavior of
116 individuals and groups (Supporting Information A.6). Moreover, Ethoflow measures specific behaviors
117 using a CNN model (Fig. 1C); different hyperparameter configurations were tested to define the CNN
118 model, which can be used to recognize binary behaviors (Supporting Information A.7). The data used to
119 train the CNN model was also generated with a heuristic based on the animal body size and length
120 (Supporting Information A.7).

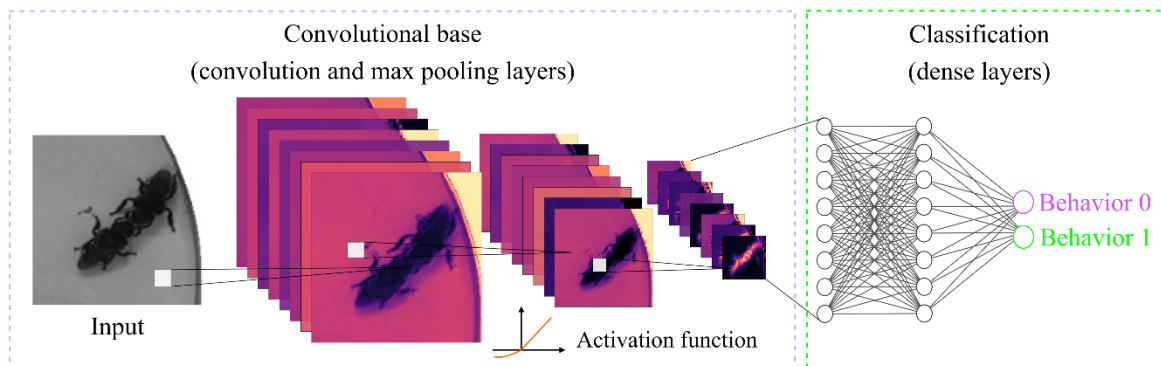
A



B



C



121

122

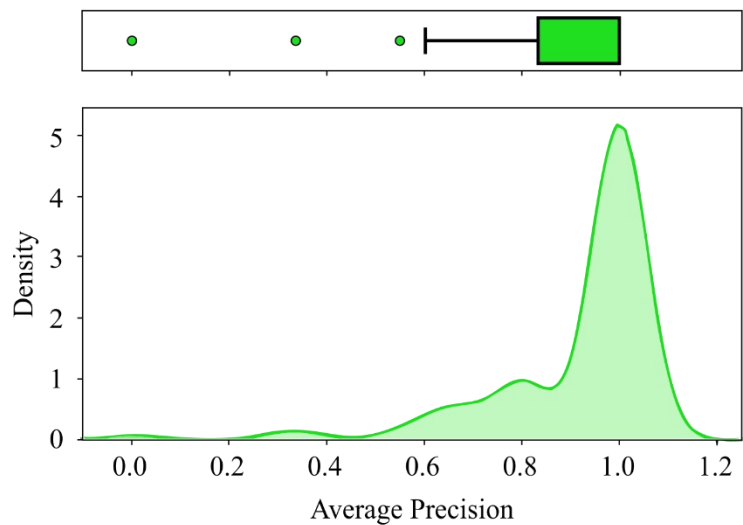
Figure 1. Features of the Ethoflow software. (A) Graphical user interface (GUI). In the left panel (Settings tab), the detected individuals are marked with a red mask after adjusting the parameters (protocol definition). Blue box on the bottom right shows the selected region of interest, which is used to assess how long the individuals remain inside the defined region. In the right panel (Analyses tab), monitoring of the processing step is demonstrated. Each individual is randomly assigned an identity and color. In this example, the analyzed objects correspond to workers of the stingless bee species *Melipona quadrifasciata* present in a Petri dish (15 cm diameter, 2 cm height). (B) Diagram depicting the operations involved in instance segmentation. The input passes through a convolutional base for feature extraction, leading to the generation of a feature map. The region proposal network (RPN) is applied, which provides several candidate boxes (ROI proposals). As several ROIs are generated, the model classifies these boxes into foreground proposals (objects) and backgrounds. ROI pooling is applied to standardize the size of the foreground proposals, slicing each foreground into a fixed amount of parts, and max pooling is applied to standardize the size. Then, the boxes labeled as real objects (the individuals) are instantiated using a pixel-wise sigmoid function. (C) Workflow of the convolutional neural networks used in Ethoflow to assess complex behaviors. In the convolutional base, the input passes through the convolutional and max pooling layers for feature extraction. Then, behavior classification is performed in the dense layers. The activation function is applied to the output of each layer to introduce non-linearity.

3 Performance and applications

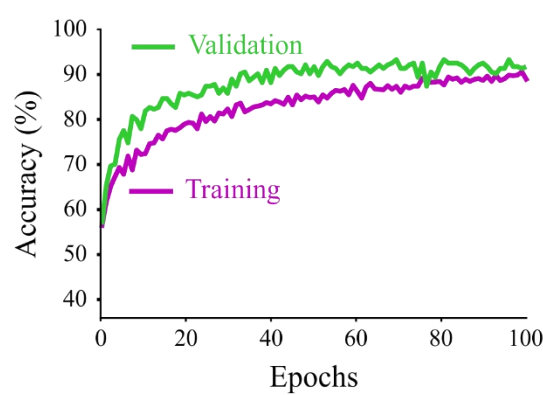
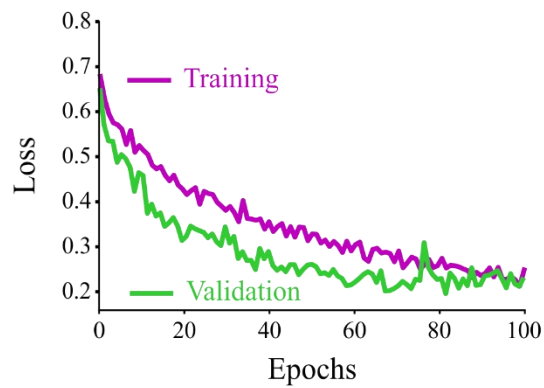
Ethoflow was trained to detect the bee *Melipona quadrifasciata* in several field conditions (Supporting Information B.1). The proposed model was efficient in detecting all animals in heterogeneous backgrounds with high precision (average precision \pm standard error = 0.916 ± 0.02 ; Fig. 2A). Ethoflow was also trained to learn the detection of trophallaxis, the complex social behavior of food

147 exchange among nestmates, in *M. quadrifasciata* (Supporting Information B.1). This model exhibited
 148 high accuracy in the validation process (global accuracy = 92.13%, Kappa index = 0.84, Z = 24.74, Fig.
 149 2B).

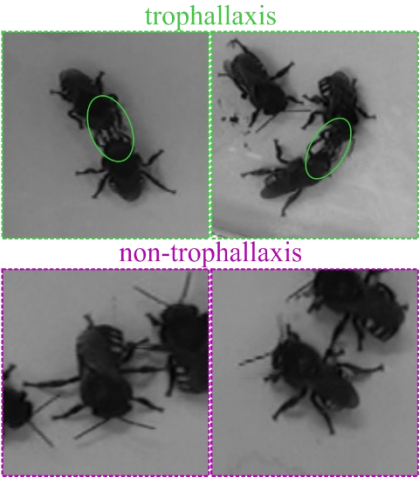
A



B



		0 25 50 75 100	
		predicted class	
target class	trophallaxis	95	5
	non-trophallaxis	10	90
		trophallaxis	non-trophallaxis



150
151
152

Figure 2. Performance of the AI models used in Ethoflow. (A) Object detection in heterogeneous backgrounds based on instance segmentation (IS). The high average precision (left panel; $n = 100$) implies that the model precisely detects real objects in the scenes with no false positives, as demonstrated by (right panel) the detected objects (*Melipona quadrifasciata* bees) with masks in random colors. (B) The training process of the CNN model (top panel) and cross-validation (confusion matrix; bottom left panel) ($n = 127$) for the monitoring of trophallaxis (green circles) in bees.

To validate Ethoflow, a behavioral assay was performed with the stingless bee species *M. quadrifasciata* and *Partamona helleri* (Supporting Information B.2) and different kinematic variables were measured (definitions of the variables are detailed in Supporting Information A.6). In both species, the centrality increased with the polarization of the group ($F_{1,35} = 25.1, p < 0.0001$) and decreased with milling ($F_{1,35} = 46.2, p < 0.0001$) (Fig. 3A). Meandering was influenced by the statistical interaction between the variables resting and bee species ($F_{1,33} = 4.71, p = 0.037$; Fig. 3B). Moreover, a difference between species was observed in the tracked distance ($F_{1,35} = 13.6, p = 0.0008$; Fig. 3C).

A toxicological bioassay was performed with *M. quadrifasciata* to demonstrate the recognition of trophallaxis under pesticide stress conditions. The pesticide imidacloprid, which is usually associated with the decline of bees, was used (Lima, Martins, Oliveira, & Guedes, 2016). The exposure protocol was based on that reported by Botina et al. (2020) (Supporting Information B.2). Bees exposed to the pesticide exhibited significantly reduced trophallaxis ($\chi^2 = 94.9, df = 58, p < 0.0001$; Fig. 3D).

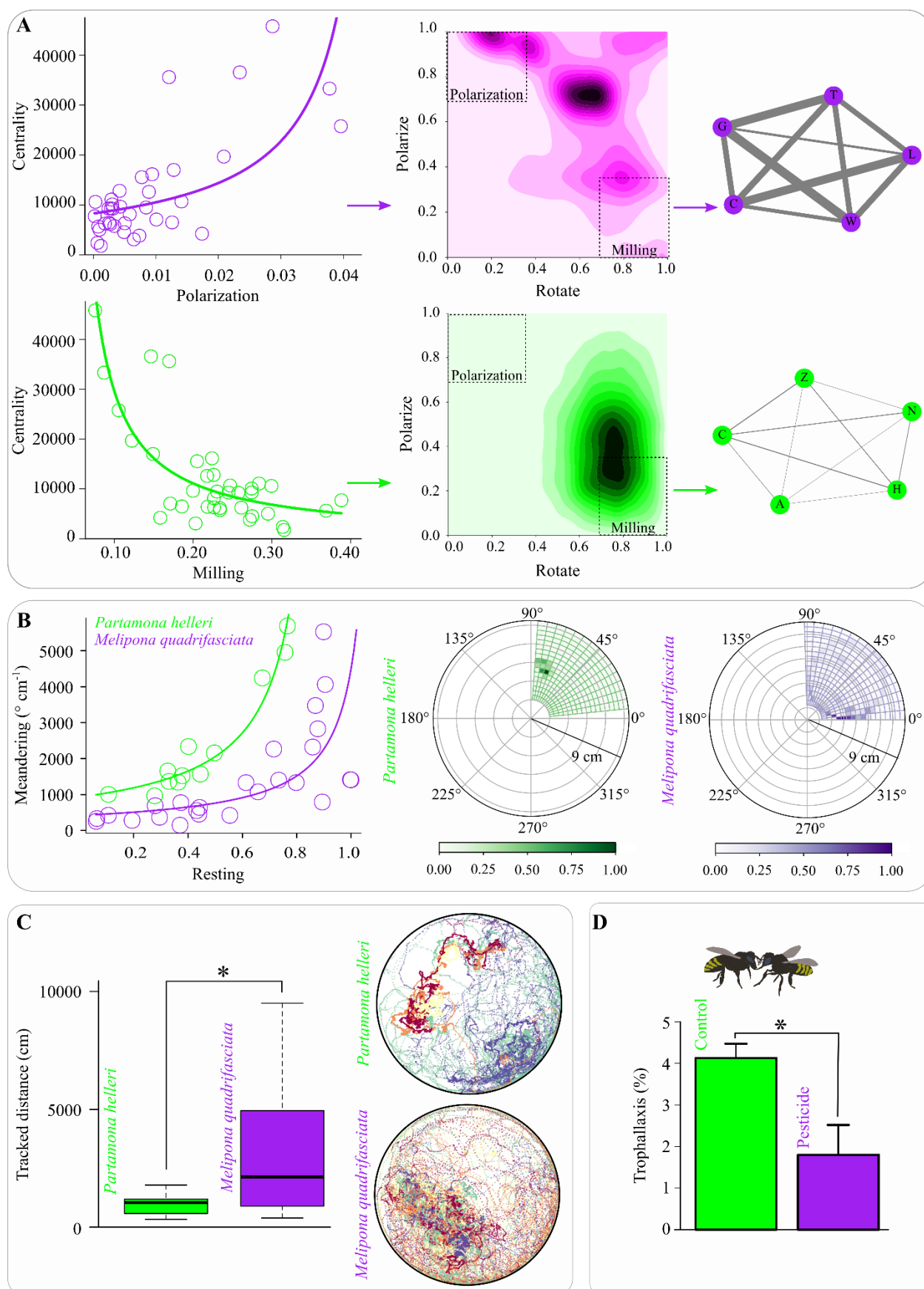
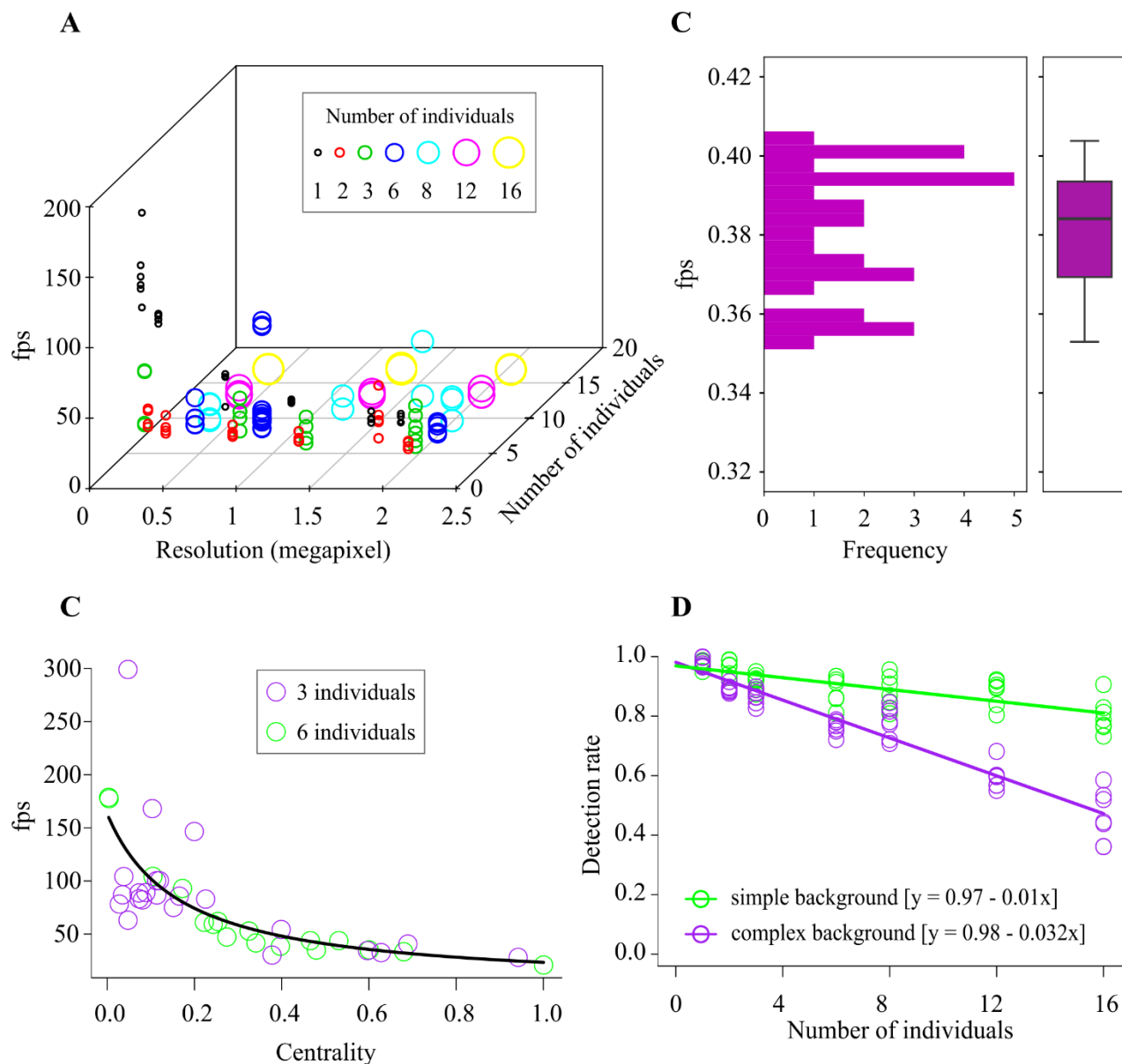


Figure 3. Behavioral assessments conducted using the Ethoflow software. (A) Association between centrality and group dynamics polarization (top panel) and milling (bottom panel) ($n = 37$). The 2D density plots and network diagrams showed that in more polarized bee groups, a higher interaction exists among individuals, while this interaction is reduced in the milling groups. In the networks, the circles represent individuals and connections correspond to the edges, where their widths are proportional to the frequency of interactions. (B) Meandering behavior is associated with resting proportions (left panel) ($n = 37$) and histograms of polar coordinates (rays and azimuth angles) for the two bee species (right panel). (C) The tracked distance of the assessed bee species ($n = 37$). In group representative tracks, the track color reflects the individual identity (right panel). (D) Trophallaxis alteration in *M. quadrifasciata* after pesticide exposure ($n = 60$). * $p < 0.05$ in the generalized linear model. When an explanatory variable had no significant effect, the model was simplified, and the results were plotted as a function of the significant variable.

Considering the parameters associated with the software performance (Supporting Information B.3), Ethoflow achieves a high processing speed, reaching 300 fps (Fig. 4). In the homogeneous backgrounds, statistical interaction was observed between the variables video resolution and group size in the range of fps ($F_{1,130} = 12.81, p = 0.0005$, Fig. 4A). The heterogeneous environment quantification was not influenced by the resolution or number of individuals ($F_{1,28} = 0.81, p = 0.37$, Fig. 4B). The fps decreased with an increase in the centrality of individuals ($F_{1,38} = 81.24, p < 0.0001$, Fig. 4C). There was no significant effect on the number of individuals ($F_{1,37} = 0.009, p = 0.93$), and no interaction was observed between the centrality and individuals ($F_{1,36} = 1.62, p = 0.21$). In addition, the software exhibited high detection rates (Fig. 4D). A significant interaction was observed between the number of individuals and type of background ($F_{1,94} = 137.85, p < 0.0001$), where an increase in the number of individuals had a greater influence on the heterogeneous environments. A few videos processed with Ethoflow are available in B.4 of the Supporting Information.

199



200

201 Figure 4. Quantification of the performance of Ethoflow. (A) Fps in response to the video resolution (in
 202 pixels) and number of individuals in homogeneous backgrounds; the dots ($n = 134$) represent the raw
 203 data. (B) Histogram of the fps in heterogeneous environments ($n = 30$). The box plot indicates the median
 204 and range of dispersion (lower and upper quartiles and outliers). (C) Fps in response to centrality. The
 205 proportion of group interaction per frame was used to quantify the centrality ($n = 40$). (D) Accuracy of
 206 the software as a function of the interaction between the number of individuals and type of environment
 207 (homogeneous and heterogeneous); the symbols represent the raw data (circles; $n = 98$). When an

explanatory variable had no significant effect, the model was simplified, and the results were plotted as a function of the significant variable.

4 Conclusion

The Ethoflow software was developed using modern computer vision techniques and AI. This software exhibited consistent speed rates and processing accuracy. The developed software is suitable for behavioral assessments in heterogeneous environments, to track groups with individuals maintaining their identities, and can be trained to learn specific behaviors. Ethoflow was applied to biological assessments, and some possibilities of data analysis and representation were demonstrated with Ethoflow's output. Accurate AI models have been implemented to expand the possibilities of animal behavior analyses to other fields, including the behavioral monitoring of domestic animals in precision livestock farming. Therefore, Ethoflow is a useful support tool for technical and scientific applications in biology and related fields.

Acknowledgements

We thank the National Council of Scientific and Technological Development (CNPq 142206/2017-2 and 301725/2019-5) and CAPES-PROEX for financial support (Financial Code 001).

Authors' contributions

RCB developed the software, performed the experiments, analyzed the data; RCB and GFM wrote the manuscript. All authors designed the work and contributed to the drafts and gave final approval for publication.

Data archiving statement

The data have been archived at <https://github.com/bernardesrodrigoc/Ethoflow>; DOI: <https://doi.org/10.5281/zenodo.3956831>

5 References

235 Abdulla, W. (2017). Mask R-CNN for object detection and instance segmentation on Keras and
236 TensorFlow. *GitHub Repository*. Github.

237 Anderson, D. J., & Perona, P. (2014). Toward a Science of Computational Ethology. *Neuron*, 84(1), 18–
238 31. doi:10.1016/J.NEURON.2014.09.005

239 Boff, S., Friedel, A., Mussury, R. M., Lenis, P. R., & Raizer, J. (2018). Changes in social behavior are
240 induced by pesticide ingestion in a Neotropical stingless bee. *Ecotoxicology and Environmental*
241 *Safety*, 164, 548–553. doi:https://doi.org/10.1016/j.ecoenv.2018.08.061

242 Botina, L. L., Bernardes, R. C., Barbosa, W. F., Lima, M. A. P., Guedes, R. N. C., & Martins, G. F. (2020).
243 Toxicological assessments of agrochemical effects on stingless bees (Apidae, Meliponini).
244 *MethodsX*, 100906. doi:https://doi.org/10.1016/j.mex.2020.100906

245 Dell, A. I., Bender, J. A., Branson, K., Couzin, I. D., de Polavieja, G. G., Noldus, L. P. J. J., ... Brose, U.
246 (2014). Automated image-based tracking and its application in ecology. *Trends in Ecology &*
247 *Evolution*, 29(7), 417–428. doi:10.1016/J.TREE.2014.05.004

248 Gandra, L. C., Amaral, K. D., Couceiro, J. C., Della Lucia, T. M., & Guedes, R. N. (2016). Mechanism
249 of leaf-cutting ant colony suppression by fipronil used in attractive toxic baits. *Pest Management*
250 *Science*, 72(8), 1475–1481. doi:10.1002/ps.4239

251 He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2018). Mask R-CNN. *IEEE Transactions on Pattern*
252 *Analysis and Machine Intelligence*, 42(2), 386–397. doi:10.1109/TPAMI.2018.2844175

253 Lima, M. A. P., Martins, G. F., Oliveira, E. E., & Guedes, R. N. C. (2016). Agrochemical-induced stress
254 in stingless bees: peculiarities, underlying basis, and challenges. *Journal of Comparative Physiology*
255 *A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 202(9–10), 733–747.
256 doi:10.1007/s00359-016-1110-3

257 Nath, T., Mathis, A., Chen, A. C., Patel, A., Bethge, M., & Mathis, M. W. (2019). Using DeepLabCut for
258 3D markerless pose estimation across species and behaviors. *Nature Protocols*, 14(7), 2152–2176.
259 doi:10.1038/s41596-019-0176-0

260 Noldus, L. P. J. ., Spink, A. J., & Tegelenbosch, R. A. . (2002). Computerised video tracking, movement
261 analysis and behaviour recognition in insects. *Computers and Electronics in Agriculture*, 35(2–3),
262 201–227. doi:10.1016/S0168-1699(02)00019-4

263 Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on*
264 *Systems, Man, and Cybernetics*, 9(1), 62–66. doi:10.1109/TSMC.1979.4310076

265 Romero-Ferrero, F., Bergomi, M. G., Hinz, R. C., Heras, F. J. H., & de Polavieja, G. G. (2019).
266 idtracker.ai: tracking all individuals in small or large collectives of unmarked animals. *Nature*
267 *Methods*, 16(2), 179–182. doi:10.1038/s41592-018-0295-5

268 Sridhar, V. H., Roche, D. G., & Gings, S. (2019). Tracktor: Image-based automated tracking of animal
269 movement and behaviour. *Methods in Ecology and Evolution*, 10(6), 815–820. doi:10.1111/2041-
270 210X.13166

271 Valletta, J. J., Torney, C., Kings, M., Thornton, A., & Madden, J. (2017). Applications of machine
272 learning in animal behaviour studies. *Animal Behaviour*, 124, 203–220.
273 doi:<https://doi.org/10.1016/j.anbehav.2016.12.005>

274 Yamanaka, O., & Takeuchi, R. (2018). UMATracker: an intuitive image-based tracking platform. *The*
275 *Journal of Experimental Biology*, 221(16).

276 Zou, Z., Shi, Z., Guo, Y., & Ye, J. (2019). Object Detection in 20 Years: A Survey. Retrieved from
277 <http://arxiv.org/abs/1905.05055>
278