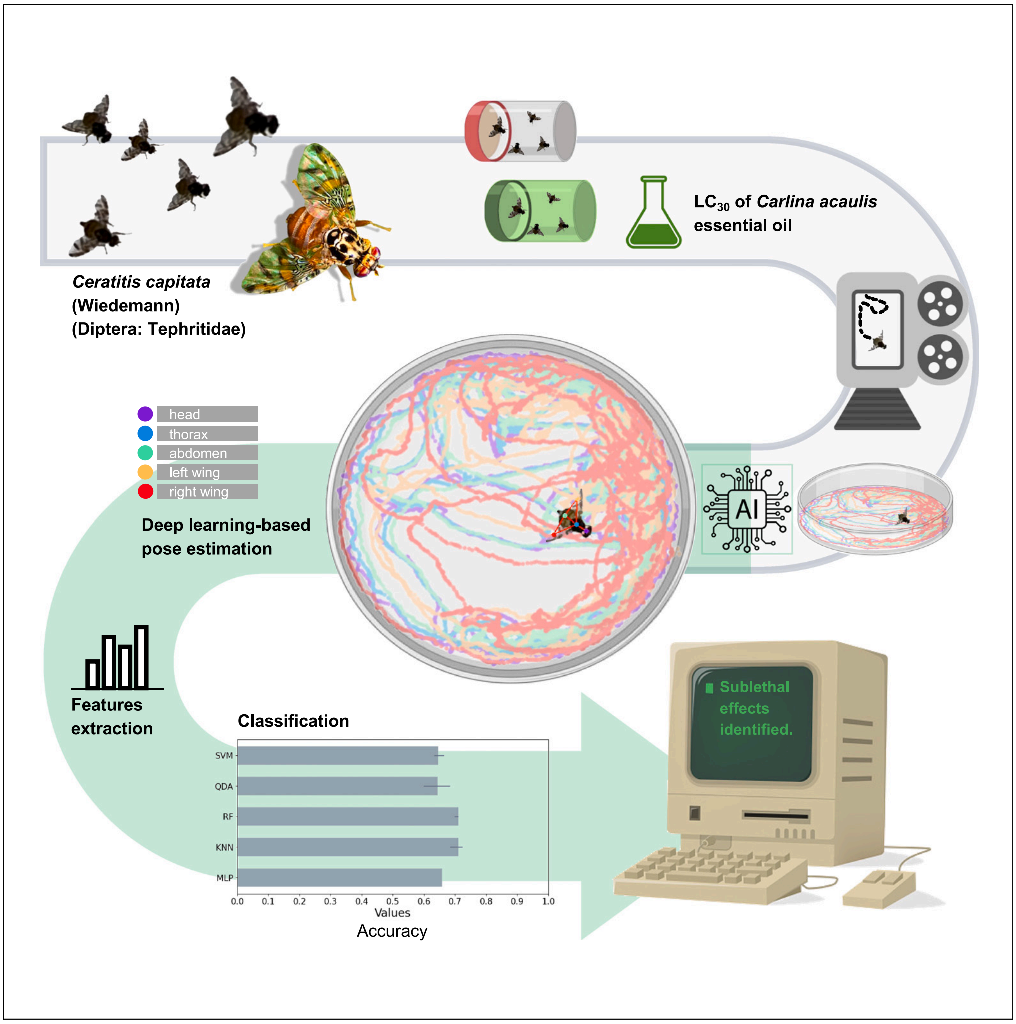
iScience

CellPress
OPEN ACCESS

Article

Learning algorithms estimate pose and detect motor anomalies in flies exposed to minimal doses of a toxicant



Gianluca Manduca, Valeria Zeni, Sara Moccia, ..., Giovanni Benelli, Cesare Stefanini, Donato Romano

gianluca.manduca@santannapisa.it (G.M.)
donato.romano@santannapisa.it (D.R.)

Highlights

Even minimal pesticide exposure can have detrimental effect on organisms

AI empowers traditional methodologies for studying animal behavior

Learning algorithms reveals toxicant effects on motor activity in a fly

The proposed automated method holds potential for future ecotoxicology research

Manduca et al., iScience *26*, 108349
December 15, 2023
https://doi.org/10.1016/j.isci.2023.108349

iScience

CellPress OPEN ACCESS

Article

Learning algorithms estimate pose and detect motor anomalies in flies exposed to minimal doses of a toxicant

Gianluca Manduca,[1,2,*] Valeria Zeni,[3] Sara Moccia,[1,2] Beatrice A. Milano,[4,5] Angelo Canale,[3] Giovanni Benelli,[3] Cesare Stefanini,[1,2] and Donato Romano[1,2,6,*]

SUMMARY

Pesticide exposure, even at low doses, can have detrimental effects on ecosystems. This study aimed at validating the use of machine learning for recognizing motor anomalies, produced by minimal insecticide exposure on a model insect species. The Mediterranean fruit fly, *Ceratitis capitata* (Diptera: Tephritidae), was exposed to food contaminated with low concentrations of *Carlina acaulis* essential oil (EO). A deep learning approach enabled fly pose estimation on video recordings in a custom-built arena. Five machine learning algorithms were trained on handcrafted features, extracted from the predicted pose, to distinguish treated individuals. Random Forest and K-Nearest Neighbor algorithms best performed, with an area under the receiver operating characteristic (ROC) curve of 0.75 and 0.73, respectively. Both algorithms achieved an accuracy of 0.71. Results show the machine learning potential for detecting sublethal effects arising from insecticide exposure on fly motor behavior, which could also affect other organisms and environmental health.

INTRODUCTION

The use of insecticides in modern agriculture is essential for crop protection,[1] but often leads to abuse, resulting in negative effects on the overall environment and non-target organisms, including humans.[2] Insecticides, indeed, are designed to especially control harmful and pathogen-carrier organisms. However, insecticides are mainly considered for their lethal effects, while their sublethal impact caused through exposure to low doses is still understudied.[3] Of note, insecticides can represent a risk to non-target species, e.g., pollinators, birds, fish, and others.[4] The overuse or misuse of insecticides can also contaminate water, air, and soil, therefore, accumulating in the tissues of animals and humans.[5,6] It is, therefore, crucial to acknowledge that such consequences may be subtle, neither evident nor easily quantifiable in the short term.[7]

Over the years, however, long-term effects of insecticide exposure have been observed, though their impact varies based on several factors. These include the specific insecticide used, the duration and frequency of exposure, and the general health condition of the individual.[8] Such effects include neurological problems, namely, headaches, dizziness, memory loss, tremors, and increased risk of developing Parkinson's disease.[6] Additionally, exposure to certain insecticides has been linked to increased risk of various forms of cancer,[8] including non-Hodgkin's lymphoma, lung cancer, leukemia, and brain cancer.[9] Reproductive problems, such as infertility, birth defects, and miscarriage, have also been associated with pesticide exposure.[10] In terms of respiratory problems, insecticides can irritate the respiratory system, causing symptoms such as coughing, wheezing, and shortness of breath.[6,11] Skin problems, such as rashes and other skin irritations, can also result from exposure to insecticides.[12] Furthermore, exposure to insecticides has the potential to disrupt the delicate balance of hormones in the body, leading to thyroid disorders and other endocrine-related issues.[6] It is precisely due to the numerous potential health consequences of exposure to insecticides that further research in this area is critically important.[13,14]

Studying the toxicological impact of biocides is crucial to achieving sustainable management of agricultural, ecological, and urban environments. Besides assessing the contribution of subtle toxic effects arising from being exposed to low doses of insecticides on pests, several insect species serve as an accessible and simple model system to assess the effects of these chemicals on invertebrates.[15] From a biomedical point of view, Diptera Brachycera as *Drosophila* spp., have been employed as models for pre-clinical studies,[16–18] and neuroscience research advancements,[19–21] just to cite some examples. By studying the subtle effects of insecticides, we can gain insights into their potential

[1]The BioRobotics Institute, Sant'Anna School of Advanced Studies, Viale Rinaldo Piaggio 34, 56025, Pontedera, Pisa, Italy
[2]Department of Excellence in Robotics and AI, Sant'Anna School of Advanced Studies, Piazza Martiri della Libertà 33, 56127, Pisa, Italy
[3]Department of Agriculture, Food and Environment, University of Pisa, Via del Borghetto 80, 56124, Pisa, Italy
[4]Institute of Life Sciences, Sant'Anna School of Advanced Studies, Piazza Martiri della Libertà 33, 56127, Pisa, Italy
[5]Faculty of Medicine and Surgery, University of Pisa, Via Roma 55/Building 57, 56126, Pisa, Italy
[6]Lead contact
*Correspondence: gianluca.manduca@santannapisa.it (G.M.), donato.romano@santannapisa.it (D.R.)
https://doi.org/10.1016/j.isci.2023.108349

environmental impact. The utilization of green insecticides, which are derived from natural sources, presents a feasible alternative to synthetic insecticides based on their inherent properties. These insecticides, such as essential oils (EOs), have the potential to mitigate adverse effects on both non-target organisms and the environment.[22,23] Although EOs activity has been studied on many insect species, research on their effect when administered at low doses is still limited.[24] Low dosages of green pesticides have been shown to influence insect life span, development, sex, ratio, fertility, fecundity and behavioral features.[25]

Insect motor function analysis represents a valuable and complementary approach to traditional biochemical and neurophysiological analyses, providing a comprehensive understanding of the insecticide toxicity on living organisms.[26–28] The interplay among biochemical, physiological, and behavioral responses to insecticide exposure could be helpful to fully evaluate the safety of a given product. Various approaches have been proposed in literature for data collection and analysis, to examine motor anomalies caused by insecticide exposure. Radio Frequency Identification (RFID) tags are a valuable method due to the capability of analyzing many test subjects simultaneously being able to put markers on each individual test subject.[29] However, physical marker systems may potentially alter subject behavior.[30] Markerless video capture systems are a cost-effective and versatile alternative to physical markers.[31–33] The approaches suggested, however, only focus on tracking individuals. Pose estimation, which involves analyzing the geometric configuration of different body parts, provides us with quantitative information on motor analysis for studying the effects of insecticides. With the help of deep learning algorithms, pose estimation has become more accurate and automated.[30,34–36] After data collection, statistical analysis is the most frequently used method to identify significant data from motor analysis that can evidence the effects of insecticides.[26,29,31,32] On the other hand, mathematical and fractal methods can be useful tools.[33] Machine learning (ML) has shown in other fields good potential for handling large amounts of data and variability, however, in this context, it has been used for genetic analysis[37] or for classifying crop pests,[38,39] but has not yet been used to investigate motor anomalies.

In our research, we used ML algorithms to investigate the impact of low-dose insecticide exposure on *Ceratitis capitata* (Wiedemann) (Diptera: Tephritidae), also known as the Mediterranean fruit fly (medfly).[40] Despite being widely used as a model organism for biology and behavioral research,[41–43] there has been limited investigation into the effect of biocides exposure on the locomotion behavior of this polyphagous fruit pest. We believe that this insect species holds significant as a model organism for biology and behavioral research.[44,45] Herein, medflies from a single population were divided into two groups. The control group remained untreated, while the second group was exposed to food contaminated with an LC_{30} dose of *Carlina acaulis* EO.[46] The latter was selected as one of the most bioactive insecticides of botanical origin currently known. Several studies demonstrated a variety of biological activities of *C. acaulis* EO, including a highly promising insecticidal potential, as well as antimicrobial and antiprotozoal activities.[47] This EO is unique as a natural substance, and it deserves additional research to enable its practical implementation in pesticide formulations since it is extremely active against multiple insect species and is composed almost solely of a component, i.e., carlina oxide (>97%). Although a green insecticide was chosen for this study, the proposed methodology can be applied to all kind of xenobiotics. Medflies were isolated into a custom-built arena. 152 videos, balanced between treated and control specimens and by sex, were recorded. By analyzing videos of treated and untreated specimens using deep learning-based software for pose estimation, we were able to obtain a series of relevant data to further analyze the effect of the green insecticide on the medfly motor behavior.[48] A deep learning approach enabled fly pose estimation on video recordings, recognizing and labeling five key points on the specimen body for each frame. Temporal and vector data were converted into features. A statistical analysis revealed more significant features in distinguishing the two groups of medflies. Five different classifiers were trained on the selected features to compare and evaluate the effects of the pesticide on the model organism. Figure 1 shows the workflow of the proposed approach. The use of ML techniques in this study allowed us to gain new insights into the acute effects of low-dose insecticide exposure on medfly, providing valuable information for understanding the risks of pesticide abuse on non-target organisms and the environment. Our work, therefore, substantially contributes to biomedical research as it integrates the principles of One Health and EcoHealth,[49] which highlight the interconnectedness of human health and the environment. Indeed, by investigating the sublethal effects of biocides on non-human organisms, we can better understand and prevent their potential impact on human health and ecosystems.

RESULTS

Deep learning-based pose estimation allowed us to obtain data over 152 videos with a duration of 5 min, balanced between treated and control specimens and by sex. The trained convolutional neural network recognized five key points on the insect body for each frame. Collected data are relevant for studying locomotory and grooming activities.

Temporal and vector data were then converted into features by evaluating median, variance, absolute maximum value, and number of peaks. We used these features to analyze the effects of insecticide exposure on the motor function of medflies. Table 1 collects all the examined features and presents the results of the statistical analysis. Among the 44 features, 19 were found to be significant (*t*-test for normally distributed features and the Wilcoxon test for non-normally distributed ones, $p = 0.05$) in distinguishing between the treated and control specimens, indicating that these features were affected by insecticide exposure. The features that showed the greatest significance in our analysis of the effects of insecticide exposure on medfly motor function were related to wing activity, specifically the median values of aperture for both the right ($\chi^2 = 8.8213$; $p = 0.0030$) and left wings ($\chi^2 = 10.0196$; $p = 0.0015$). In particular, the right wing showed greater disruption in its movement, with the variance of aperture being the feature that exhibited the greatest significance between treated and control specimens ($\chi^2 = 12.2549$; $p = 0.0005$). Additionally, the median value of velocity ($\chi^2 = 6.9224$; $p = 0.0085$) and the number of positive acceleration peaks ($\chi^2 = 7.1453$; $p = 0.0075$) were also significantly affected by insecticide exposure. Another feature of notable significance was the maximum duration of time that specimens remained immobile during the experiment ($\chi^2 = 8.1665$; $p = 0.0043$). Features related to the orientation of the insect's body and the angles it traveled, while still relevant, showed less significance than wing activity.

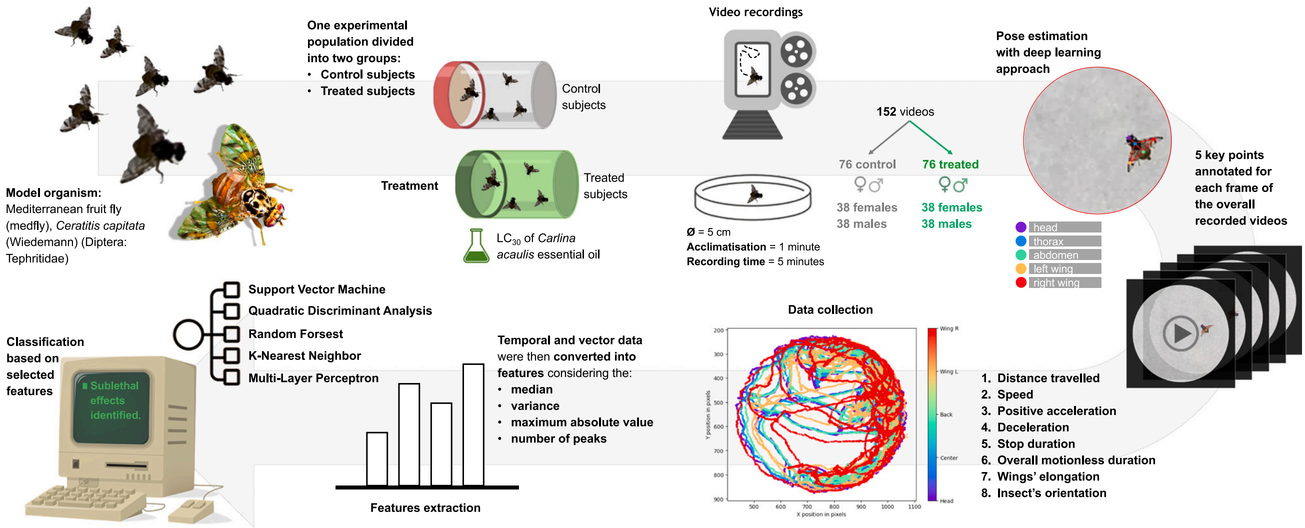


Figure 1. Process workflow

Medflies from a single population were divided into two groups. The control group was untreated, while the second group was exposed to food contaminated with an LC_{30} of *Carlina acaulis* EO. Medfly individuals were then isolated into a custom-built arena, and 152 videos, balanced between treated and control specimens and by sex, were recorded. A deep learning approach enabled fly pose estimation on video recordings, recognizing and labeling five key points on the specimen body for each frame. Relevant motor activities data were collected, and features extracted. Based on these features, five different machine learning algorithms (Support Vector Machine, Quadratic Discriminant Analysis, Random Forest, K-Nearest Neighbor, and Multi-Layer Perceptron) were trained to distinguish treated subjects.

Figure 2 presents a detailed comparison between the treated and control groups using boxplots. This comparison is based on six distinct features that emerged as particularly significant from the statistical analysis. These features include maximum time duration of immobility within the arena (A), median speed (B), number of positive acceleration peaks (C), median aperture of the left (D) and right wings (E), and variance of the latter (F). From the data observation, it is evident that treated specimens tend to remain motionless for shorter durations compared to the control group, while simultaneously exhibiting a higher median speed and less pronounced re-starts. Intriguingly, the most pronounced effect of insecticide exposure appears to revolve around the features associated with wing activity. Treated specimens display a more substantial aperture of both wings during the test. Focusing specifically on the right wing, the treated medflies exhibit reduced variance in wing aperture. This observation concerns the feature with the highest significance according to the statistical analysis.

The significant features were then used for training Support Vector Machine (SVM), Quadratic Discriminant Analysis (QDA), Random Forest (RF), K-Nearest Neighbor (KNN), and Multi-Layer Perceptron (MLP). Each of these five machine learning algorithms used in the study has a different approach to classification and prediction. By using a diverse set of algorithms, the study aimed to cover a broad range of classification methods and provide a more comprehensive, accurate and reliable analysis of the effects of green insecticide exposure on fly motor behavior. To cope with the limited size of our dataset, we used 4-fold cross-validation for both model testing and retrieval of optimal hyper-parameters. Please refer to the methods for more information about the optimization grid used. Figure 3 shows the performance of the classifiers across precision, recall, accuracy, and f1-score. Results are expressed in terms of median and inter-quartile range among the 4-folds. Overall, the classifiers achieved an average precision of 0.66, recall of 0.75, accuracy of 0.67, and f1-score of 0.7. Among the five classifiers, RF and KNN best performed in distinguishing the insecticide-exposed medfly group from the control group, indeed both achieved an accuracy of 0.71. RF achieved high performance in all the metrics considered, surpassing a score of 0.7 with a precision of 0.7, a recall of 0.74, and an f1-score of 0.72. This indicates that RF was able to accurately identify the treated medfly group from the control one, with a high level of precision, recall and f1-score. The precision score of 0.7 indicates that out of all the specimens predicted to be in the treated medfly group, 70% of them were in that group. The recall score of 0.74 indicates that out of all the specimens in the treated medfly group, 74% of them were correctly identified by the classifier. The f1-score of 0.72, which is the harmonic mean of precision and recall, combines both metrics to provide an overall measure of the classifier's performance. In comparison, KNN achieved a precision score of 0.68, a recall score of 0.79, and an f1-score of 0.74. Although KNN had a slightly lower precision score than RF, it achieved a higher recall score, indicating that it was better at correctly identifying the specimens in the treated medfly group. The f1-score of 0.74 for KNN is also relatively high, indicating that it achieved a good balance between precision and recall. We also calculated the area under the curve (AUC) of the receiver operating characteristic (ROC). The mean ROC curves of the five classifiers employed are shown in Figure 4A. The overall average ROC curve had an AUC of 0.69. However, RF and KNN outperformed the other classifiers, with AUC values of 0.75 and 0.72, respectively. Additionally, SVM achieved a good result in terms of AUC (0.71). The relative ROC curves obtained during the iterative cross-validation process for RF and KNN are shown in Figure 4B and 4C. Notably, while RF slightly surpassed KNN in performance, both algorithms consistently achieved high metric scores. Overall, these results demonstrate the effectiveness of machine learning in distinguishing between treated and control medfly specimens.

Table 1. Statistical analysis of the examined features

Features	Feature selected number	χ^2	*p*-value
Stops			
Number of stops []		0.5354	0.4643
Sum of stop times (overall motionless duration) [s]	1	5.0362	0.0248*
Duration			
Median [s]		0.6203	0.4310
Variance [s^2]	2	5.8257	0.0158*
Maximum [s]	3	8.1665	0.0043**
Distance			
Traveled distance [mm]		3.5596	0.0592
Speed			
Median [mm/s]	4	6.9224	0.0085**
Variance [mm^2/s^2]		0.1385	0.7098
Number of peaks []	5	4.7890	0.0286*
Maximum [mm/s]		0.0275	0.8683
Acceleration			
Median [mm/s^2]		3.6717	0.0553
Variance [mm^2/s^4]		0.6453	0.4218
Number of peaks []	6	7.1453	0.0075**
Maximum [mm/s^2]		0.4161	0.5189
Deceleration			
Median [mm/s^2]	7	4.4583	0.0347*
Variance [mm^2/s^4]		0.2054	0.6504
Number of peaks []	8	4.8124	0.0283*
Minimum [mm/s^2]		0.1414	0.7069
Left wing normalized aperture ($L_l(t)$)			
Median [0–1]	9	10.0196	0.0015**
Variance [0-0.25]		3.3406	0.0676
Number of peaks []		1.1317	0.2874
Right wing normalized aperture ($L_r(t)$)			
Median [0–1]	10	8.8213	0.0030**
Variance [0–0.25]	11	12.2549	0.0005***
Number of peaks []		1.7282	0.1886
Module of the left and right wing normalized aperture difference ($\lvert L_l(t) - L_r(t)\rvert$)			
Median [0–1]	12	6.3901	0.0115*
Variance [0–0.25]	13	3.9596	0.0466*
Number of peaks []	14	5.7693	0.0163*
Maximum [0–1]		1.6069	0.2049
Angle of rotation on itself when stationary ($\theta_s(t)$)			
Clockwise			
Median [rad]	15	6.1320	0.0133*
Variance [rad^2]		0.0130	0.9091
Number of peaks []	16	6.3647	0.0116*
Maximum [rad]		1.1037	0.2935
Counterclockwise			

(*Continued on next page*)

Table 1. *Continued*

Features	Feature selected number	χ^2	*p*-value
Median [rad]	17	4.5365	0.0332*
Variance [rad^2]		0.2475	0.6189
Number of peaks []	18	6.2993	0.0121*
Minimum [rad]		0.0217	0.8828
Angle of rotation when traveling ($\theta_t(t)$)			
Clockwise			
Median [rad]		1.6162	0.2036
Variance [rad^2]		0.7692	0.3805
Number of peaks []	19	5.3414	0.0208*
Maximum [rad]		0.8493	0.3568
Counterclockwise			
Median [rad]		2.9361	0.0866
Variance [rad^2]		0.6753	0.4112
Number of peaks []		2.6421	0.1041
Minimum [rad]		0.0827	0.7737

The table presents the features evaluated and statistical analysis results to filter out relevant data. Shapiro-Wilk test was used to test for normal distribution (p = 0.01) for each feature. For trials with two conditions (e.g., LC_{30}–exposed flies vs. control), statistical significance was established using the t-test for normally distributed features and the Wilcoxon test for non-normally distributed features. The threshold for statistical significance was set at $p = 0.05$. The last column of the table shows the *p*-value for each feature (different numbers of asterisks indicate the level of statistical significance, specifically * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$). Features with statistical significance have been highlighted in bold.

Pose estimation using a deep learning approach made it possible to extrapolate important data about the motor activity of medflies. Our analysis indicated that exposure to pesticides had a notable impact on various features, including fly wing aperture. Classifiers trained on these features clearly identified motor anomalies, distinguishing between treated and control specimens. These findings are important for identifying the potential impacts of insecticide exposure.

DISCUSSION

In this study, we employed ML techniques to analyze the motor behavior of medflies exposed to an LC_{30} of an insecticidal formulation based on *C. acaulis* EO. This approach offers a comprehensive understanding of the impact of this biocide on medfly locomotor activity. The objective of the present research was to investigate the sublethal effects of an insecticide, typically evaluated at concentration ranging from LC_5 to LC_{30}. Sublethal effects are defined as effects on individuals that survive exposure to an insecticide, and it induces no apparent mortality in the experimental population.[40,44]

Earlier research evaluated the effects of insecticides through various methods, including biochemical and neurophysiological analyses.[50–52] Our approach provides a novel and innovative way to examine the effects of a green insecticide on insect locomotion and is less intrusive than traditional methods that use physical markers.[30] Behavioral alterations due to insecticide treatment have been previously investigated in bees, among others. Further studies have used different methods to monitor the activity of treated specimens, such as RFID technology, which detects movements of specimens that are equipped with microchips,[29] or automated video-tracking systems such as EthoVisionXT.[31] In these studies, the presence of food and the interaction between test subjects were considered during the tests. The temporal data obtained were analyzed using statistics, and variables such as distance traveled, time spent in the food zone, and interaction time were considered. Similar methods were applied in another study where motor anomalies were studied in larval amphibians exposed to an insecticide.[32] Tenorio et al.[33] used a video-tracking approach with ImageJ 1.49v software to investigate the locomotion of shrimps, employing mathematical and fractal methods. Our research proposes a different approach using deep learning for the pose estimation of medflies from videos. To ensure the accuracy of our study, we isolated the specimens during the experiments and controlled for any external factors that could have influenced our results. This allowed us to minimize the impact of confounding variables and obtain more reliable and precise data. Furthermore, our approach allowed us to investigate the effects of low-dose insecticide exposure on the locomotion of medflies, which would have been difficult to detect using traditional toxicological studies. Our approach offers a more nuanced understanding of the effects of insecticides on insect behavior, providing valuable insights into the potential risks of insecticide exposure and emphasizing the need for more comprehensive risk assessments, particularly for plant extracts. Overall, our study demonstrates the importance of carefully designed experiments and the potential of advanced technologies such as deep learning for studying the effects of insecticides on a pest of high agricultural importance. Our image-based approach utilizes advanced deep learning algorithms to provide a more detailed and complex analysis of medfly motor behavior. By analyzing various features such as wing aperture, we were able to detect subtle changes in insect locomotion

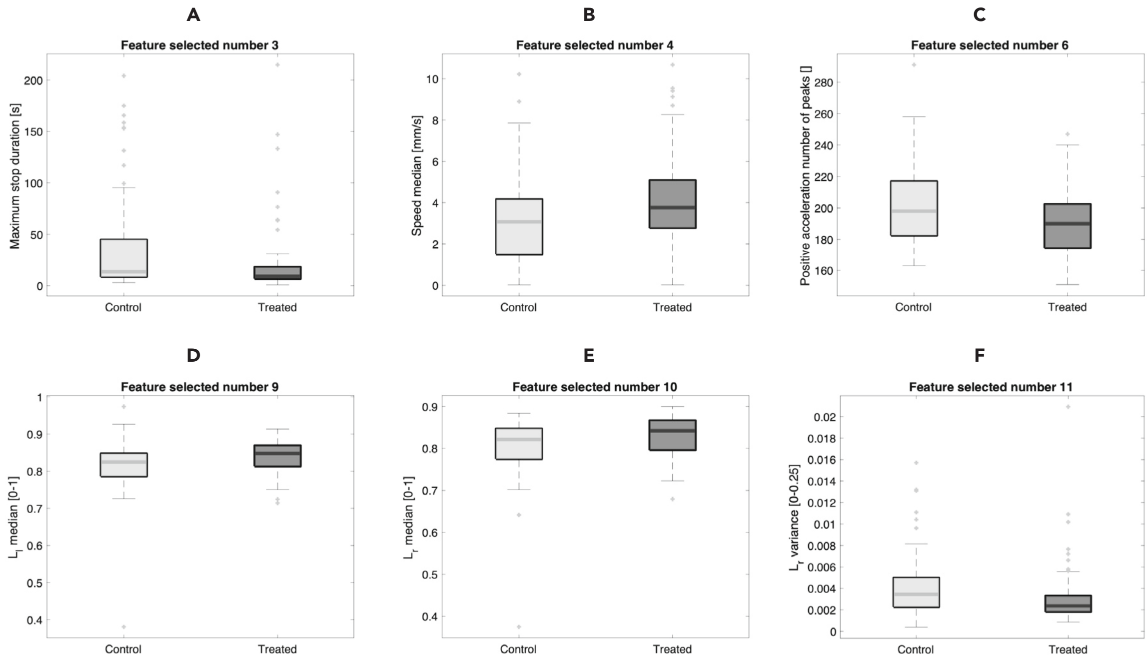


Figure 2. Most significant features comparison between control and treated flies

Comparison through boxplots between control and treated subjects considering the six most relevant features according to the statistical analysis: maximum time duration of immobility within the arena (χ^2 = 8.1665; p = 0.0043) (A), median speed (χ^2 = 6.9224; p = 0.0085) (B), number of positive acceleration peaks (χ^2 = 7.1453; p = 0.0075) (C), median aperture of the left (χ^2 = 10.0196; p = 0.0015) (D) and right wings (χ^2 = 8.8213; p = 0.0030) (E), and variance of the latter (χ^2 = 12.2549; p = 0.0005) (F).

that would have been difficult to observe using traditional methods. This allowed us to investigate the effects of low doses of *C. acaulis* EO exposure on medfly behavior, providing valuable insights into its potential as green insecticides in IPM programs. Our approach also allowed us to quantify the effects of green insecticide exposure on medfly locomotion, providing a more objective and reliable assessment of the overall biological activity of the EO. Furthermore, our methodology offers several advantages over traditional toxicological studies. Our

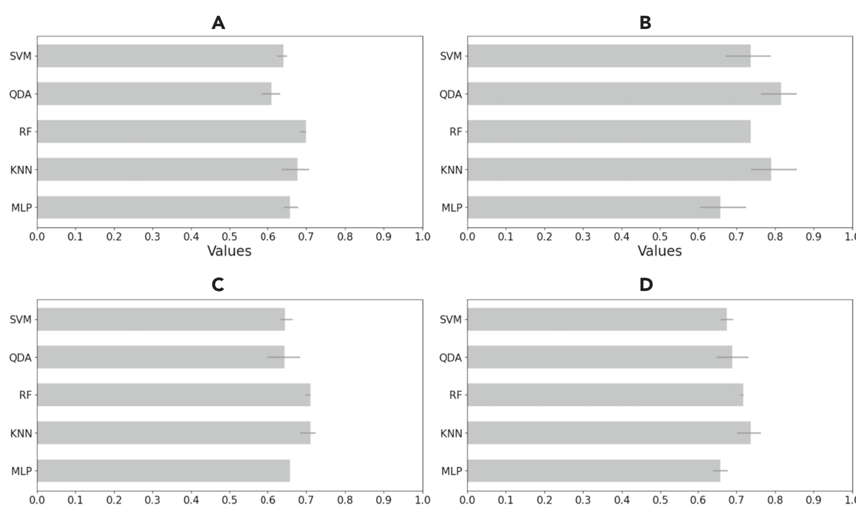


Figure 3. Comparative analysis of performance metrics across classifiers

The performances of the five classifiers employed were compared using different evaluation metrics: precision (A), recall (B), accuracy (C), and f1-score (D). The results, obtained through cross-validation, are presented in terms of median and inter-quartile range among the 4-folds. Precision measures the proportion of true positive predictions among all positive predictions, while recall measures the proportion of true positive predictions among all actual positive cases. Accuracy measures the proportion of correct predictions among all cases, while f1-score is the harmonic mean of precision and recall.

CellPress
OPEN ACCESS

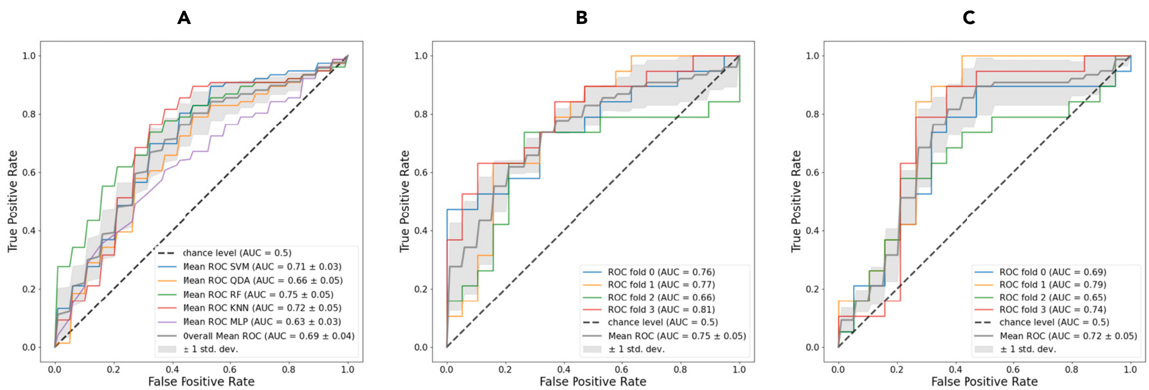


Figure 4. ROC curves

Mean ROC curves of the five classifiers employed and overall mean ROC curve are shown (A). The ROC curves plot the true positive rate (TPR) against the false positive rate (FPR) at different classification thresholds, allowing for the assessment of the trade-off between sensitivity and specificity. The curves are obtained by averaging the four cross-validation testing results for each classifier. Panels (B) and (C) highlight the performance of Random Forest and K-Nearest Neighbor respectively, both of which achieved superior AUC scores compared to the other tested algorithms. Each panel showcases both the individual ROC curves and their mean values derived from 4-fold cross-validation testing.

approach is less invasive, as it does not require physical markers or other invasive procedures that could affect insect behavior. Additionally, our approach is more sensitive and accurate, as it can detect subtle changes in insect behavior that may be missed by human observers. Our protocol also offers the potential for high-throughput screening of green insecticides, allowing for more rapid and efficient assessment of their toxicity. Though this study focused on the effects of a green insecticide, the introduced methodology undeniably offers potential for evaluating the sublethal effects of various xenobiotics on insect motor behavior.

The pose estimation algorithm used in our study allowed for a detailed and complex motor analysis of medflies. We conducted statistical analysis on the collected data and identified 19 significant features that could potentially indicate motor anomalies due to insecticide administration. Several studies have reported significant patterns in motor behavior through statistical analysis, highlighting the effects of insecticides on various organisms.[26,29,31,32] To detect and highlight the effects of insecticide exposure on medfly behavior, we used machine learning algorithms, including SVM, QDA, RF, KNN, and MLP. Each algorithm was trained to classify the treated and control fly groups. The performances of the classifiers were compared, and results provided valuable insights into the effects of insecticide exposure on medfly behavior. Overall, the classifiers achieved an average accuracy of 0.66, a recall of 0.75, an accuracy of 0.67, and an f1-score of 0.7. RF and KNN algorithms performed the best in distinguishing the treated medfly group from the control group, achieving an accuracy of 0.71. ROC curves analysis further showed that both algorithms had clear effects of the insecticide on the fly motor behavior, with an AUC value of 0.75 (RF) and 0.72 (KNN). Our methodology, which achieved an accuracy of 71% in classifying treated flies, is comparable to initial results obtained in similar systems.[53–55] It's essential to note that while higher accuracy levels have been reported in some machine learning applications, the unique challenges posed by medfly motor activities classification, particularly when discerning subtle changes due to toxicant minimal doses exposure, make our results promising for this specific context. The achieved results indicate a substantial potential for practical applications, even in field experiments. While not flawless, this level of accuracy can support researchers in rapid preliminary field screenings, where even human expertise might miss the effects of pesticides dispersed in the environment. There are several avenues to potentially improve the classification accuracy: Increasing the dataset size might enhance the model's generalizability. Diving deeper into the behavioral attributes of the medflies might reveal additional features that can improve classification performances.

This study findings suggest that our image-based approach, combined with machine learning algorithms, can provide a powerful tool for studying the sublethal effects of insecticides on fly motor behavior. Furthermore, the proposed automated method can offer a streamlined and efficient approach to ecotoxicology research. To ensure the reliability of the method, a focus was placed on fundamental features, with locomotion being the central aspect of this study. Integrating evaluations of intricate behaviors, such as courtship, territoriality, and oviposition, at this initial stage could have jeopardized the integrity of the data obtained. Following the successful validation for basic behaviors, subsequent studies will explore the effects on these more complex behavioral patterns.

Limitations of the study

A limitation of our study may be seen in the fact that we only assessed the acute effects of low-dose green insecticide exposure on medfly motor function. Acute effects can be more difficult to detect than chronic effects, as they may be transient or reversible and may require sensitive and specific assays to measure. Additionally, acute effects may not accurately reflect the long-term impacts of insecticide exposure on medfly health and behavior. Therefore, while our study provides important insights into the immediate effects of insecticide exposure on medfly motor function, it is important to recognize that these effects may not fully capture medfly behavior. Future studies that incorporate longer-term assessments and more complex behavioral assays may be needed to fully understand the broader implications of insecticide use.

Another limitation is that our study focused on analyzing the free behavior of the flies, without any external stimuli. As such, the results obtained might not fully reflect the complex and dynamic nature of medfly behavior in their natural environment. Future studies could incorporate external stimuli such as attracting and repelling factors, or interactions with other individuals (e.g., agonistic behavior and courtship), to better understand the impact of insecticide exposure on medfly behavior.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Insect treatment
  - Experimental set-up and video recordings
  - Pose estimation
  - Features extraction
  - Machine learning classifiers
  - Performance assessment
- QUANTIFICATION AND STATISTICAL ANALYSIS

ACKNOWLEDGMENTS

This research was supported by the EU H2020 FETOPEN Project 'Robocoenosis - ROBOts in cooperation with a bioCOENOSIS' [899520], and the Department of Excellence in Robotics and AI of Sant'Anna School of Advanced Studies, Pisa, Italy. Funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the article. The authors thank Ms. Federica Durini and Mr. Raffaele Picichè for their kind assistance in visualization.

AUTHOR CONTRIBUTIONS

Conceptualization: G.M., and D.R.; methodology: G.M., S.M., G.B., C.S., and D.R.; experiment: G.M., V.Z., S.M., B.A.M., and A.C.; formal analysis: G.M., V.Z., G.B., C.S., and D.R.; resources: G.B., C.S., and D.R.; writing – original draft: G.M.; writing – review & editing: G.M., V.Z., S.M., B.A.M., A.C., G.B., C.S., and D.R.; supervision: D.R.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: May 10, 2023
Revised: October 4, 2023
Accepted: October 24, 2023
Published: October 27, 2023

REFERENCES

1. Tudi, M., Daniel Ruan, H., Wang, L., Lyu, J., Sadler, R., Connell, D., Chu, C., and Phung, D.T. (2021). Agriculture development, pesticide application and its impact on the environment. Int. J. Environ. Res. Publ. Health *18*, 1112. https://doi.org/10.3390/ijerph18031112.
2. Damalas, C.A., and Eleftherohorinos, I.G. (2011). Pesticide exposure, safety issues, and risk assessment indicators. Int. J. Environ. Res. Publ. Health *8*, 1402–1419. https://doi.org/10.3390/ijerph8051402.
3. Guedes, R.N.C., Benelli, G., and Agathokleous, E. (2022). Arthropod outbreaks, stressors, and sublethal stress. Current Opinion in Environmental Science & Health *28*, 100371. https://doi.org/10.1016/j.coesh.2022.100371.
4. Hernández, M., and Margalida, A. (2008). Pesticide abuse in Europe: effects on the Cinereous vulture (*Aegypius monachus*) population in Spain. Ecotoxicology *17*, 264–272. https://doi.org/10.1007/s10646-008-0193-1.
5. Yin, R., Yao, S., and Huo, X. (2013). China's forest tenure reform and institutional change in the new century: What has been implemented and what remains to be pursued? Land Use Pol. *30*, 825–833. https://doi.org/10.1016/j.landusepol.2012.06.010.
6. Curl, C.L., Spivak, M., Phinney, R., and Montrose, L. (2020). Synthetic pesticides and health in vulnerable populations: agricultural workers. Curr. Environ. Health Rep. *7*, 13–29. https://doi.org/10.1007/s40572-020-00266-5.

7. Shen, Y., Zhao, E., Zhang, W., Baccarelli, A.A., and Gao, F. (2022). Predicting pesticide dissipation half-life intervals in plants with machine learning models. J. Hazard Mater. *436*, 129177. https://doi.org/10.1016/j.jhazmat.2022.129177.
8. Blair, A., and Zahm, S.H. (1995). Agricultural exposures and cancer. Environ. Health Perspect. *103*, 205–208. https://doi.org/10.1289/ehp.95103s8205.
9. Morrison, H.I., Wilkins, K., Semenciw, R., Mao, Y., and Wigle, D. (1992). Herbicides and cancer. J. Natl. Cancer Inst. *84*, 1866–1874. https://doi.org/10.1093/jnci/84.24.1866.
10. Bag, D. (2000). Pesticides and Health Risks. Econ. Polit. Wkly. *35*, 3381–3383. 38. http://www.jstor.org/stable/4409739.
11. Varghese, J.V., Sebastian, E.M., Iqbal, T., and Tom, A.A. (2021). Pesticide applicators and cancer: a systematic review. Rev. Environ. Health *36*, 467–476. https://doi.org/10.1515/reveh-2020-0121.
12. Kim, K.H., Kabir, E., and Jahan, S.A. (2017). Exposure to pesticides and the associated human health effects. Sci. Total Environ. *575*, 525–535. https://doi.org/10.1016/j.scitotenv.2016.09.009.
13. Jepson, P.C., Murray, K., Bach, O., Bonilla, M.A., and Neumeister, L. (2020). Selection of pesticides to reduce human and environmental health risks: a global guideline and minimum pesticides list. Lancet Planet. Health *4*, e56–e63. https://doi.org/10.1016/S2542-5196(19)30266-9.
14. Baghel, A.S., Bhardwaj, A., and Ibrahim, W. (2022). Optimization of Pesticides Spray on Crops in Agriculture Using Machine Learning. Comput. Intell. Neurosci. *2022*, 9408535. https://doi.org/10.1155/2022/9408535.
15. Spong, K.E., Andrew, R.D., and Robertson, R.M. (2016). Mechanisms of spreading depolarization in vertebrate and insect central nervous systems. J. Neurophysiol. *116*, 1117–1127. https://doi.org/10.1152/jn.00352.2016.
16. Rudrapatna, V.A., Cagan, R.L., and Das, T.K. (2012). *Drosophila* cancer models. Dev. Dynam. *241*, 107–118. https://doi.org/10.1002/dvdy.22771.
17. Enomoto, M., Siow, C., and Igaki, T. (2018). *Drosophila* as a cancer model. Drosophila Models for Human Diseases, pp. 173–194. https://doi.org/10.1007/978-981-13-0529-0_10.
18. Gladstone, M., and Su, T.T. (2011). Chemical genetics and drug screening in *Drosophila* cancer models. Journal of Genetics and Genomics *38*, 497–504. https://doi.org/10.1016/j.jgg.2011.09.003.
19. Bellen, H.J., Tong, C., and Tsuda, H. (2010). 100 years of *Drosophila* research and its impact on vertebrate neuroscience: a history lesson for the future. Nat. Rev. Neurosci. *11*, 514–522. https://doi.org/10.1038/nrn2839.
20. Kazama, H. (2015). Systems neuroscience in *Drosophila*: Conceptual and technical advantages. Neuroscience *296*, 3–14. https://doi.org/10.1016/j.neuroscience.2014.06.035.
21. Davis, R.L. (2005). Olfactory memory formation in *Drosophila*: from molecular to systems neuroscience. Annu. Rev. Neurosci. *28*, 275–302. https://doi.org/10.1146/annurev.neuro.28.061604.135651.
22. Pavela, R., and Benelli, G. (2016). Essential oils as ecofriendly biopesticides? Challenges and constraints. Trends Plant Sci. *21*, 1000–1007. https://doi.org/10.1016/j.tplants.2016.10.005.
23. Isman, M.B. (2020). Bioinsecticides based on plant essential oils: A short overview. Z. Naturforsch., C: J. Biosci. *75*, 179–182. https://doi.org/10.1515/znc-2020-0038.
24. Giunti, G., Benelli, G., Palmeri, V., Laudani, F., Ricupero, M., Ricciardi, R., Maggi, F., Lucchi, A., Guedes, R.N.C., Desneux, N., and Campolo, O. (2022). Non-target effects of essential oil-based biopesticides for crop protection: Impact on natural enemies, pollinators, and soil invertebrates. Biol. Control *176*, 105071. https://doi.org/10.1016/j.biocontrol.2022.105071.
25. Isman, M.B. (2020). Commercial development of plant essential oils and their constituents as active ingredients in bioinsecticides. Phytochem. Rev. *19*, 235–241. https://doi.org/10.1007/s11101-019-09653-9.
26. Tosi, S., and Nieh, J.C. (2017). A common neonicotinoid pesticide, thiamethoxam, alters honey bee activity, motor functions, and movement to light. Sci. Rep. *7*, 15132. https://doi.org/10.1038/s41598-017-15308-6.
27. Zaluski, R., Kadri, S.M., Alonso, D.P., Martins Ribolla, P.E., and de Oliveira Orsi, R. (2015). Fipronil promotes motor and behavioral changes in honey bees (*Apis mellifera*) and affects the development of colonies exposed to sublethal doses. Environ. Toxicol. Chem. *34*, 1062–1069. https://doi.org/10.1002/etc.2889.
28. Gómez-Giménez, B., Felipo, V., Cabrera-Pastor, A., Agustí, A., Hernández-Rabaza, V., and Llansola, M. (2018). Developmental exposure to pesticides alters motor activity and coordination in rats: sex differences and underlying mechanisms. Neurotox. Res. *33*, 247–258. https://doi.org/10.1007/s12640-017-9823-9.
29. Decourtye, A., Devillers, J., Aupinel, P., Brun, F., Bagnis, C., Fourrier, J., and Gauthier, M. (2011). Honeybee tracking with microchips: a new methodology to measure the effects of pesticides. Ecotoxicology *20*, 429–437. https://doi.org/10.1007/s10646-011-0594-4.
30. Mathis, A., Mamidanna, P., Cury, K.M., Abe, T., Murthy, V.N., Mathis, M.W., and Bethge, M. (2018). DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. Nat. Neurosci. *21*, 1281–1289. https://doi.org/10.1038/s41593-018-0209-y.
31. Teeters, B.S., Johnson, R.M., Ellis, M.D., and Siegfried, B.D. (2012). Using video-tracking to assess sublethal effects of pesticides on honey bees (*Apis mellifera* L.). Environ. Toxicol. Chem. *31*, 1349–1354. https://doi.org/10.1002/etc.1830.
32. Denoël, M., Libon, S., Kestemont, P., Brasseur, C., Focant, J.F., and De Pauw, E. (2013). Effects of a sublethal pesticide exposure on locomotor behavior: a video-tracking analysis in larval amphibians. Chemosphere *90*, 945–951. https://doi.org/10.1016/j.chemosphere.2012.06.037.
33. Tenorio, B.M., da Silva Filho, E.A., Neiva, G.S.M., da Silva, V.A., Tenorio, F.D.C.A.M., da Silva, T.D.J., Silva, E.C.S.E., and Nogueira, R.d.A. (2017). Can fractal methods applied to video tracking detect the effects of deltamethrin pesticide or mercury on the locomotion behavior of shrimps? Ecotoxicol. Environ. Saf. *142*, 243–249. https://doi.org/10.1016/j.ecoenv.2017.03.051.
34. Høye, T.T., Ärje, J., Bjerge, K., Hansen, O.L.P., Iosifidis, A., Leese, F., Mann, H.M.R., Meissner, K., Melvad, C., and Raitoharju, J. (2021). Deep learning and computer vision will transform entomology. Proc. Natl. Acad. Sci. USA *118*, e2002545117. https://doi.org/10.1073/pnas.2002545117.
35. Graving, J.M., Chae, D., Naik, H., Li, L., Koger, B., Costelloe, B.R., and Couzin, I.D. (2019). DeepPoseKit, a software toolkit for fast and robust animal pose estimation using deep learning. Elife *8*, e47994. https://doi.org/10.7554/eLife.47994.
36. Walter, T., and Couzin, I.D. (2021). TRex, a fast multi-animal tracking system with markerless identification, and 2D estimation of posture and visual fields. Elife *10*, e64000. https://doi.org/10.7554/eLife.64000.
37. Tomiazzi, J.S., Judai, M.A., Nai, G.A., Pereira, D.R., Antunes, P.A., and Favareto, A.P.A. (2018). Evaluation of genotoxic effects in Brazilian agricultural workers exposed to pesticides and cigarette smoke using machine-learning algorithms. Environ. Sci. Pollut. Res. Int. *25*, 1259–1269. https://doi.org/10.1007/s11356-017-0496-y.
38. Thenmozhi, K., and Srinivasulu Reddy, U. (2019). Crop pest classification based on deep convolutional neural network and transfer learning. Comput. Electron. Agric. *164*, 104906. https://doi.org/10.1016/j.compag.2019.104906.
39. Tannous, M., Stefanini, C., and Romano, D. (2023). A Deep-Learning-Based Detection Approach for the Identification of Insect Species of Economic Importance. Insects *14*, 148. https://doi.org/10.3390/insects14020148.
40. Desneux, N., Decourtye, A., and Delpuech, J.M. (2007). The sublethal effects of pesticides on beneficial arthropods. Annu. Rev. Entomol. *52*, 81–106. https://doi.org/10.1146/annurev.ento.52.110405.091440.
41. Louâpre, P., van Alphen, J.J.M., and Pierre, J.S. (2010). Humans and insects decide in similar ways. PLoS One *5*, e14251. https://doi.org/10.1371/journal.pone.0014251.
42. Huang, Y., Liu, Z., and Rong, Y.S. (2016). Genome editing: from *Drosophila* to non-model insects and beyond. Journal of Genetics and Genomics *43*, 263–272. https://doi.org/10.1016/j.jgg.2016.04.007.
43. Gomulski, L.M., Dimopoulos, G., Xi, Z., Soares, M.B., Bonaldo, M.F., Malacrida, A.R., and Gasperi, G. (2008). Gene discovery in an invasive tephritid model pest species, the Mediterranean fruit fly, *Ceratitis capitata*. BMC Genom. *9*, 243. https://doi.org/10.1186/1471-2164-9-243.
44. Benelli, G., Ceccarelli, C., Zeni, V., Rizzo, R., Lo Verde, G., Sinacori, M., Boukouvala, M.C., Kavallieratos, N.G., Ubaldi, M., Tomassoni, D., et al. (2022). Lethal and behavioural effects of a green insecticide against an invasive polyphagous fruit fly pest and its safety to mammals. Chemosphere *287*, 132089. https://doi.org/10.1016/j.chemosphere.2021.132089.
45. Barud, F.J., Gómez, M.P., Ruiz, M.J., Bachmann, G.E., Goane, L., Segura, D.F., Lara, N., Murúa, F., Asfennato, A., Gómez, E., et al. (2023). Sexual competitiveness of sterile *Ceratitis capitata* males exposed to essential oils from non-host plant species native to Argentina. Entomol. Exp. Appl. *171*, 146–155. https://doi.org/10.1111/eea.13255.
46. Benelli, G., Rizzo, R., Zeni, V., Govigli, A., Samková, A., Sinacori, M., Lo Verde, G., Pavela, R., Cappellacci, L., Petrelli, R., et al. (2021). *Carlina acaulis* and *Trachyspermum ammi* essential oils formulated in protein baits are highly toxic and reduce aggressiveness in the medfly, *Ceratitis capitata*. Ind. Crop. Prod. *161*, 113191.

https://doi.org/10.1016/j.indcrop.2020.113191.

47. Spinozzi, E., Ferrati, M., Cappellacci, L., Caselli, A., Perinelli, D.R., Bonacucina, G., Maggi, F., Strzemski, M., Petrelli, R., Pavela, R., et al. (2023). *Carlina acaulis* L. (Asteraceae): biology, phytochemistry, and application as a promising source of effective green insecticides and acaricides. Ind. Crop. Prod. *192*, 116076. https://doi.org/10.1016/j.indcrop.2022.116076.
48. Martins, V.A., Freitas, L.C., de Aguiar, M.S., de Brisolara, L.B., and Ferreira, P.R. (2019). Deep learning applied to the identification of fruit fly in intelligent traps. In 2019 IX Brazilian Symposium on Computing Systems Engineering (SBESC) (IEEE), pp. 1–8. https://doi.org/10.1109/SBESC49506.2019.9046088.
49. Pallathadka, H., Mustafa, M., Sanchez, D.T., Sekhar Sajja, G., Gour, S., and Naved, M. (2023). Impact of machine learning on management, healthcare and agriculture. Mater. Today: Proc. *80*, 2803–2806. https://doi.org/10.1016/j.matpr.2021.07.042.
50. Papaefthimiou, C., and Theophilidis, G. (2001). The cardiotoxic action of the pyrethroid insecticide deltamethrin, the azole fungicide prochloraz, and their synergy on the semi-isolated heart of the bee *Apis mellifera macedonica*. Pestic. Biochem. Physiol. *69*, 77–91. https://doi.org/10.1006/pest.2000.2519.
51. Pilling, E.D., Bromleychallenor, K.A.C., Walker, C.H., and Jepson, P.C. (1995). Mechanism of synergism between the pyrethroid insecticide λ-cyhalothrin and the imidazole fungicide prochloraz, in the honeybee (*Apis mellifera* L.). Pestic. Biochem. Physiol. *51*, 1–11. https://doi.org/10.1006/pest.1995.1001.
52. Rumpf, S., Hetzel, F., and Frampton, C. (1997). Lacewings (Neuroptera: Hemerobiidae and Chrysopidae) and integrated pest management: enzyme activity as biomarker of sublethal insecticide exposure. J. Econ. Entomol. *90*, 102–108. https://doi.org/10.1093/jee/90.1.102.
53. Symons, M., Feeney, G.F.X., Gallagher, M.R., Young, R.M., and Connor, J.P. (2020). Predicting alcohol dependence treatment outcomes: a prospective comparative study of clinical psychologists versus 'trained' machine learning models. Addiction *115*, 2164–2175. https://doi.org/10.1111/add.15038.
54. Yassin, W., Nakatani, H., Zhu, Y., Kojima, M., Owada, K., Kuwabara, H., Gonoi, W., Aoki, Y., Takao, H., Natsubori, T., et al. (2020). Machine-learning classification using neuroimaging data in schizophrenia, autism, ultra-high risk and first-episode psychosis. Transl. Psychiatry *10*, 278. https://doi.org/10.1038/s41398-020-00965-5.
55. Munsell, B.C., Wee, C.Y., Keller, S.S., Weber, B., Elger, C., da Silva, L.A.T., Nesland, T., Styner, M., Shen, D., and Bonilha, L. (2015). Evaluation of machine learning algorithms for treatment outcome prediction in patients with epilepsy based on structural connectome data. Neuroimage *118*, 219–230. https://doi.org/10.1016/j.neuroimage.2015.06.008.
56. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn: Machine learning in Python. the Journal of machine Learning research *12*, 2825–2830. https://doi.org/10.48550/arXiv.1201.0490.
57. Canale, A., and Benelli, G. (2012). Impact of mass-rearing on the host seeking behaviour and parasitism by the fruit fly parasitoid *Psyttalia concolor* (Szépligeti)(Hymenoptera: Braconidae). J. Pest. Sci. *85*, 65–74. https://doi.org/10.1007/s10340-011-0382-1.
58. Nath, T., Mathis, A., Chen, A.C., Patel, A., Bethge, M., and Mathis, M.W. (2019). Using DeepLabCut for 3D markerless pose estimation across species and behaviors. Nat. Protoc. *14*, 2152–2176. https://doi.org/10.1038/s41596-019-0176-0.
59. Mathis, M.W., and Mathis, A. (2020). Deep learning tools for the measurement of animal behavior in neuroscience. Curr. Opin. Neurobiol. *60*, 1–11. https://doi.org/10.1016/j.conb.2019.10.008.
60. Ambrosini, E., Caielli, M., Milis, M., Loizou, C., Azzolino, D., Damanti, S., Bertagnoli, L., Cesari, M., Moccia, S., Cid, M., et al. (2019). Automatic speech analysis to early detect functional cognitive decline in elderly population. In 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (IEEE), pp. 212–216. https://doi.org/10.1109/EMBC.2019.8856768.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
Carlina acaulis EO	Pharmaceutical Biology, University of Camerino	Prof. Filippo Maggi
Experimental models: Organisms/strains		
Ceratitis capitata (Wiedemann) (Diptera: Tephritidae)	Department of Agriculture, Food and Environment (DAFE), University of Pisa	Prof. Angelo Canale
Software and algorithms		
DeepLabCut	Mathis et al.[30]	RRID:SCR_021391
MATLAB	MathWorks	RRID:SCR_001622
Scikit-learn	Pedregosa et al.[56]	RRID:SCR_002577
JMP Pro	JMP Statistical Discovery	RRID:SCR_014242

RESOURCE AVAILABILITY

Lead contact

Further information should be directed to and will be fulfilled by the Lead Contact, Donato Romano (donato.romano@santannapisa.it).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Data reported in this paper will be shared by the lead contact upon request.
- Original code will be shared by the lead contact upon request.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

For all the experiments, *Ceratitis capitata* (Wiedemann) (Diptera: Tephritidae) were reared at University of Pisa, as described by Canale and Benelli.[57] Mass-rearing conditions were 25 $\pm$ 1°C, 45 % R.H., and 16:8 (L:D). Adult medflies (both sexes) used in bioassays were 10–12 days old. The present research adheres to the guidelines for the treatment of animals in behavioral research and teaching (ASAB/ABS 2014), the Italian laws (D.M. 116,192), and the regulations of the European Union (European Commission 2007). For tests involving *C. capitata*, no particular permits were needed by the Italian government.

METHOD DETAILS

Insect treatment

We aimed to determine whether a low dose of *C. acaulis* EO elicited a sublethal effect on the locomotory and grooming behavior of medfly adults. Following Benelli et al.,[46] medflies were fed for 96 h with 2 mL of mucilage containing the LC_{30} dose (716 ppm) of *C. acaulis* EO. The mucilage was prepared by emulsifying the EO with DMSO (1:1), then 2% of carboxymethylcellulose sodium salt, 12.5 % of sucrose and 1% of protein bait (NuBait, Biogard, Italy) were added. The *C. acaulis* EO was kindly provided by Prof. Filippo Maggi (University of Camerino, Italy); the EO was chiefly composed of carlina oxide (97.7%), its GC-MS analysis was detailed in our recent study Benelli et al.[46] The mucilage was administered inside a bakelite cup (∅ =30 mm), covered with a cotton disk (∅ =30 mm). A negative control was carried out testing the viscous carrier without EO.

Experimental set-up and video recordings

Test subjects were observed in a circular arena (∅ = 5 cm). To better investigate the non-aerial behaviors, including walking and grooming, medflies were not allowed to fly. To do this, the edges of the arena were 3D printed in Anycubic 3D Printing UV Sensitive Resin with three different heights (h = 2.7, 3.5 and 4.5 mm) to adapt the arena to the size of the tested subject. Before testing, medflies were acclimated for 1 min in a Petri dish (Ø = 5 cm). An RGB camera (12MP camera, telephoto: f /2.8 aperture) was placed over the arena to record medfly locomotory and grooming activities. A total of 152 videos with a duration of 5 min were recorded. Videos were recorded at 30 fps with a resolution of 1920x1080 pixels. The overall videos were balanced between treated and control specimens and by sex. Treated medflies were kept separate from control ones before the experiment. All recordings occurred between 11:00 and 17:00 to reduce the effect of circadian rhythms, at 25 $\pm$ 1°C.

Pose estimation

We annotated the poses of insects semi-automatically at the frame level with DeepLabCut (DLC), a deep learning-based pose estimation software available online.[30] First, we selected a total of 1157 frames from five videos. We manually labelled five key points (KPs) on the insect's body for each frame: head, thorax, abdomen, and two wings (Figure 1). The five KPs were selected for an analysis not only of the centroid, but also of the angles travelled, and the motor activity of the wings. We trained a convolutional neural network (CNN) to recognize KPs on medflies and annotate the remaining 1,366,843 frames from the 152 videos. To train the network, 95% of the manually labelled images were employed, while the remaining 5% were used for testing. We started with the ResNet50 backbone pre-trained on ImageNet and fine-tuned the network with 600,000 training iterations using a batch size of one. The number of training set images was balanced based on the insect size and sex. The number of selected frames and the training iterations were chosen based on values recommended in the literature.[30,58,59] The software's default settings were used to determine the split percentage between labeled images for training and those for testing, as well as the selection of the batch size.

Features extraction

We focused our attention on medfly movement data with possible behavioral relevance. Temporal KPs coordinates were low-pass filtered with a passband frequency of 0.05 Hz and a steepness of 0.95 avoiding possible pose estimation inaccuracies. We investigated the insect's distance travelled, speed, positive acceleration, and deceleration through the medfly thorax tracking. We also considered the durations of the times the medfly remained stationary during the experiment (stop duration) and their sum (overall motionless duration). The wings' aperture was analyzed during tests. The distance between the key points of the wing and abdomen was taken as the wing aperture. As the insect size varies among individuals, the aperture was normalized to the maximum value measured during the experiment. Through the thorax-abdomen skeleton we investigated the insect's rotation angles. The angles travelled clockwise and counterclockwise were considered separately. The case in which the insect rotated on itself when stationary was distinguished from when it was in motion. Temporal and vector data were then converted into features considering the median, variance, maximum absolute value, and number of peaks. The data obtained from DLC were processed, and features were then extrapolated using MATLAB (MathWorks Inc., MA) software.

Machine learning classifiers

Following the statistical analysis, we selected important features to train various learning-based models in order to investigate the effects of pesticide exposure, as proposed by.[60] We considered five models, including Support Vector Machine, Quadratic Discriminant Analysis, Random Forest, K-Nearest Neighbor, and Multi-Layer Perceptron. For robust testing, we performed a nested cross-validation (CV) with both 4-fold external and internal loops. In each outer iteration, 114 subjects were used for training, and 38 for testing. Hyperparameter tuning in the internal loop was performed using grid search CV. All analyses were conducted using scikit-learn in Python.[56] For SVM, we used a Radial Basis Function (RBF) kernel and considered a range of C values (0.01 to 100,000) with scaling. The gamma parameter was treated as the inverse of the radius of influence of support vectors. Larger C values prioritize correct classification of training examples, while smaller values encourage a larger decision function margin. For QDA, we used default hyperparameter values. For RF, we tuned the number of tree estimators and maximum tree depth with grid-search spaces of [20, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500] and [3–7], respectively. For KNN, we used a k-range between 10 and 15. For MLP, we considered a maximum of 5 nodes and two layers.

Performance assessment

To thoroughly evaluate the performance of the classifiers, we adopted multiple metrics that capture different aspects of classification efficacy. The following metrics were considered:

1. Accuracy measures the ratio of correct predictions to the total predictions made.
2. Precision evaluates the ratio of correctly predicted positive observations, true positives (TP), to the total predicted positives, true and false positives (FP).

$$Precision = \frac{TP}{TP+FP}$$

3. Recall (Sensitivity) represents the ratio of correctly predicted positive observations, over all the positive cases in the data, TP and false negatives (FN).

$$Recall = \frac{TP}{TP+FN}$$

4. F1-Score provides the harmonic mean of precision and recall, offering a balance between them.

$$F1 - Score = 2 * \frac{Precision * Recall}{Precision+Recall}$$

5. Receiver operating characteristic (ROC) curves plot the TP rate (sensitivity) against the FP rate (1-specificity) across different threshold values. The area under the curve (AUC) of the ROC provides a comprehensive measure of the classifier's performance across all thresholds.

Each of these metrics offers a unique perspective on classifier performance, ensuring a well-rounded evaluation. Through this multifaceted assessment, we aimed to provide a robust analysis of the classifier's capabilities in the context of our study.

QUANTIFICATION AND STATISTICAL ANALYSIS

We performed statistical analysis to filter out relevant data. A Shapiro-Wilk test was used for each feature to verify for normal distribution (p = 0.01). For trials with two conditions (e.g., LC_{30}–exposed flies vs. control), statistical significance was established using the t-test for normally distributed features and the Wilcoxon test for non-normally distributed ones. Statistical analyses were performed using JMP Pro 16 (SAS) software. The threshold was set at p = 0.05.