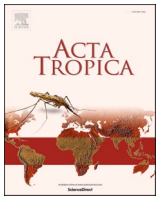
Acta Tropica 279 (2026) 108120

Contents lists available at ScienceDirect

Acta Tropica

journal homepage: www.elsevier.com/locate/actatropica



Mapping mosquito flight dynamics and directional responses: A scalable deep learning model for behavioural research

Manuela Carnaghi [a,1,*], Khaled Mostafa [b,1], Mohamed Hany [b,1], Ayman Atia [b]

[a] Agriculture, Health, and Environment Department, Natural Resources Institute, University of Greenwich, United Kingdom
[b] Faculty of Computer Science, MSA University Giza, Egypt

ARTICLE INFO

Keywords:
Behavioural analysis
Machine learning
AI
Flight track
Mosquitoes
Directional analysis

ABSTRACT

Mosquitoes are important vectors of pathogens that affect millions of people worldwide. Understanding their flight patterns and behaviours is crucial for developing novel control and surveillance tools. Mosquito flight track analysis using traditional manual methods can be time-consuming and laborious. In this study, we employed advanced image processing and deep learning techniques, specifically using a GRU model, to analyse 2D video recordings of three medically important mosquito vector species. Videos were recorded in controlled laboratory settings using simple, low-tech equipment. Our model integrates background subtraction, YOLOv5 detection, and the DeepSORT matching strategy to detect and track mosquitoes with high accuracy, achieving detection rates between 99.7% and 99.9% depending on the mosquito species, effectively mitigating challenges posed by background noise, occlusions, and tracking labels inconsistencies. In addition, a Gated Recurrent Unit (GRU) model was employed to classify mosquito movement directions with ~97.6% accuracy. The system also generates visual outputs, including heatmaps and videos that illustrate mosquito flight trajectories, facilitating the interpretation of mosquito behavioural responses under experimental conditions. These findings demonstrate that integrating computer vision with deep learning techniques provides an effective method for tracking mosquito flight paths, classifying movement patterns, and assessing behavioural responses. This approach offers a promising avenue for automating video analysis in mosquito research and may be adapted for studying other small flying insect species.

1. Introduction

Mosquitoes transmit pathogens that affect nearly 700 million people annually, resulting in over 700,000 death every year (World Health Organization, 2024). These staggering statistics underscore the global health threat posed by mosquito vectors. Understanding the biology, ecology, and behaviour of mosquitoes is crucial for developing effective control strategies against mosquito-borne illnesses (Carnaghi et al., 2021). In particular, studying mosquito movements and behavioural patterns can drive the creation of new strategies or the improvement on current strategies used for disease prevention and control. Furthermore, studying their flight parameters could also shed light on the factors that influence their dispersal, which is a key aspect that interests ecologists and medical entomologists. However, tracking mosquito movements and analysing its behaviour can be challenging, as behavioural studies generally require continuous observation, making the monitoring process difficult, time-consuming, and labour intensive (Morris et al., 2022). Direct observation of the mosquitoes is impractical due to their small size and rapid flight speed. Furthermore, as mosquitoes are hematophagous and naturally attracted to humans, the presence of an observer in the experimental room can influence their behavioural responses. These constraints significantly limit experimental designs. A viable solution to this problem is to record mosquitoes on video and analyse the footage. Multiple tracking systems have been developed and applied to record mosquito flight tracks (Bredt et al., 2025; Javed et al., 2024b; Spitzen and Takken, 2018). However, these systems often require complex high-tech setups and expensive equipment, thus limiting their use in neglected tropical disease vector research where resources are generally scarce. There is a need to develop simple recording systems that can be paired with automated methods of video analysis. By utilizing automated approaches capable of rapidly processing large volumes of data, it may be possible to uncover minor

* Corresponding author.
E-mail address: manuela.carnaghi@greenwich.ac.uk (M. Carnaghi).
[1] These authors contributed equally to the work.

https://doi.org/10.1016/j.actatropica.2026.108120
Received 18 December 2025; Received in revised form 18 March 2026; Accepted 4 May 2026
Available online 7 May 2026

patterns or connections that might have otherwise gone unnoticed with human analysis alone (Hashimoto et al., 2018). Automated video analysis methods also enable continuous observations, saving time and resources (Kong and Peng, 2023).

Recent developments in artificial intelligence (AI) offer promising solutions for automated video analysis. Artificial Intelligence has emerged as a transformative force across industries, revolutionizing task execution and decision-making processes (Xu et al., 2021). Systems using AI are capable of automating tasks, provide new insights, and facilitate complex problem-solving through the analysis of vast datasets, which enables them to identify hidden patterns (Joksimovic et al., 2023). Two AI subdomains, machine learning (ML) and computer vision, are particularly relevant. Machine learning enables machines to learn from data and improve their performance without explicit programming, whilst computer vision is used to train machines to analyse visual information, e.g. photos and videos (Kitaguchi et al., 2021) Over the past decade, significant advancements have been made in object detection and tracking by improving algorithms and computer vision technologies. Particular advancements have been made in recognizing objects in various states of action, i.e. objects moving or remaining still (Murthy et al., 2020). One of the most prominent algorithmic structures in this field is You Only Look Once (YOLO), which delivers high accuracy and rapid processing speeds through an efficient neural network design (Diwan et al., 2023). The adoption of this neural network design facilitates real-time object detection and tracking across various applications (Rane et al., 2023). Moreover, it enables efficient resource allocation and scalability for large-scale object detection tasks.

The remarkable diversity of behaviours exhibited by insects, ranging from complex social systems to intricate mating rituals, has long attracted scientists' attention. The use of AI could significantly enhance researchers' ability to conduct research related to insect classification, biology, and behaviour. For example, the use of AI could enable detailed investigation of swarm dynamics, feeding methods, and communication patterns, all complex factors that are crucial in the study of insect communities (Sumpter and Pratt, 2003). In recent years, several studies have demonstrated the value of AI, including ML, in behavioural and classification studies on animals (Hu et al., 2023; Jiang et al., 2020; Maekawa et al., 2020), including mosquitoes (Javed et al., 2024b). Recent advances in AI have enabled more detailed investigation of mosquito flight behaviour. For instance, a recent study applied ML models to classify flight trajectories of *Anopheles gambiae* around insecticide-treated nets, revealing an immediate onset of erratic behaviour upon exposure to permethrin, including in resistant strains (Qureshi et al., 2025). Using SHAP analysis, their model identified trajectory features predictive of insecticidal disruption with high classification accuracy. Similarly, FlightTrackAI (Javed et al., 2024a) integrated convolutional neural networks with multi-object tracking and spline interpolation to achieve 99.9% tracking accuracy and over 91% success in identity reassignment following occlusions in videos of *Aedes aegypti*. This represented a substantial improvement over the group's earlier work (Javed et al., 2023), in which a Mask RCNN model was used under similar conditions but was less resilient to occlusions and background interference. Collectively, these studies addressed key challenges posed by mosquito morphology and flight variability. Other applications of ML in mosquito research have focused primarily on identification and taxonomy, where these approaches have demonstrated useful for classification tasks. For example, a YOLOv3 deep learning model accurately distinguished mosquitoes from other insects, and when mosquitoes were detected, the model identified the species and the sex of the specimen with an average precision rate of 99% and a sensitivity rate of ~92% (Kittichai et al., 2021). Similarly, a model based on deep learning algorithms classified different *Aedes* species using wing interferential patterns, with identification accuracy exceeding 70% for eight out of the ten species studied (Cannet et al., 2023).

Here, we present a model that uses computer vision and deep learning algorithms to analyse 2D videos of mosquitoes in laboratory settings using low-tech equipment. In addition to mosquito detection and tracking, the model incorporates a directional movement classification feature, enabling it to identify whether a mosquito is moving to the right, to the left, upwards, downwards, in a zigzag pattern, or not moving at all; the model can then display such information as superimposed tags on videoframes. The directional movement classification can be informative when analysing the behavioural approach that mosquitoes display close to baited traps or control tools. We tested the model on three mosquito species, all major vectors of human pathogens: *Ae. aegypti, Anopheles coluzzii*, and *Culex quinquefasciatus*. The model proposed here provides further tools to track and analyse mosquito flight paths and behaviours, thus offering insights that could inform innovative disease prevention strategies.

2. Materials and methods

2.1. Mosquito rearing and experimental set up

All three mosquito species were maintained under the controlled laboratory conditions with a 12: 12 h light: dark photoperiods, temperature of 26 ± 2 °C, and a relative humidity of 60 ± 5%. Rearing was carried out following established protocols (Carnaghi et al., 2024a). In summary, mature female mosquitoes were fed defibrinated horse blood using artificial feeders (Hemotek, UK). Gravid females were allowed to oviposit eggs in oviposition dishes, which consisted of a Petri dish lined with a layer of damp cotton wool covered by moist filter paper. Eggs were transferred into 1 L of deionized water solution with 0.1% of aquarium salts (Tropic Marin, Germany). Larvae were fed ad libitum until pupation with organic baby rice (4–6+ months, Aptamil, Netherlands) and a source of protein, which was either fish flakes (TetraMin Tetra Werke, Germany) for *Anopheles* mosquitoes, or dog food (Purina Bakers Adult 100% Complete, UK) for *Culex* and *Aedes* mosquitoes. Pupae were transferred daily into 30 × 30 × 30 cm cages (BugDorms, Taiwan), and emerged adults were provided a 10% sucrose solution, until ready to be used in experiments. Mosquitoes used in the assay were between five to seven days post-emergence.

To enhance responsiveness during video recording, *An. coluzzii* and *Cx. quinquefasciatus* mosquitoes were starved for five hours prior to the experiment commencement, whilst *Ae. aegypti* mosquitoes were starved for 24 h. One hour prior to the experiment, mosquitoes were selected and transferred into 25 mL tubes (i.e. habituation vials) and were kept in light habituation conditions as follows: *Ae. aegypti* mosquitoes were subjected to dusk light for 30 min, whilst *An. coluzzii* and *Cx. quinquefasciatus* were habituated to night darkness conditions by keeping them in darkness for one hour. Lighting conditions during the habituation and experimental time were designed to simulate natural light levels at which mosquitoes are most active. As *Ae. aegypti* is known to be a crepuscular or day-time biting mosquito, a suffused warm-white light that mimicked dusk lighting was placed at a one metre distance from the experimental cage. On the other hand, as *An. coluzzii* and *Cx. quinquefasciatus* mosquitoes are known to be nocturnal or night-time biting mosquitoes, an array of warm-white LEDs was placed around the cage, mimic light levels equivalent to starlight (Carnaghi et al., 2024b). For each experimental run a new group of mosquitoes were transferred from the habituation vial into the experimental cage (20 × 20 × 45 cm). The experimental cage consisted of a metal frame covered with 100% nylon sheer tights (15 deniers, nude; George ASDA, UK), secured on the right side with a small knot, producing a slightly darker region in the lower-right corner. The fine seams of the nylon introduced minor linear patterns detectable by the camera, which did not affect recording quality. Mosquitoes were left in the experimental cage for five to ten minutes to habituate before recording. Each video lasted for two to five minutes, after which, the recording was stopped, mosquitoes were left to rest undisturbed for ten minutes before starting a new recording. Each batch of mosquitoes was filmed three times, each video was considered a technical replicate. At the end of the third recording, each batch of

mosquitoes was removed from the cage, the room was left to air for 15 min before commencing a new set of recordings which were carried out using a freshly assembled cage, i.e. the nylon case was changed for each mosquito batch. For each species the following experimental groups were tested: 1) group of six male mosquitoes; 2) group of six female mosquitoes; and 3) group of six female mosquitoes to which host cues were presented. To present host cues, a black terrarium heating pad (ProRep, UK), of 10 × 15 cm surface area, set at 35°C was presented 3 cm away from the cage, along with a nylon sock which was worn for 24 h by the experimenter (Carnaghi et al., 2021). This set up presented a combination of thermal, visual, and olfactory cue, which are known to be attractants for host-seeking female mosquitoes. For each experimental group, four batches of mosquitoes were used, each group representing a biological replicate. Thus, a total of 72 mosquitoes were used per species.

The recording system consisted of a simple camera (acA2440 – 75 μm, Basler, Germany), with 2448 × 2048 pixel resolution, fitted with HF6XA-5 M lenses (Fujinon, Japan), positioned 15 cm from the cage. The camera was connected to a laptop using Basler Video Recording Software (Basler, Germany) to record the videos at a rate of 25 frames per second. The camera's field of view captured the entire cage. To ensure visibility in low-light conditions, three infrared (IR) lamps (IRINB20UK, JC, China), each containing an array of twelve LEDs emitting light > 800 nm, were arranged around the cage to create a diffused and evenly distributed IR light illumination visible to the camera but undetectable to mosquitoes (Gibson, 1995). Note that the addition of IR LEDs is not necessary when recording in daylight conditions.

2.2. Overall model training, testing, and output

The detection, tracking, and behavioural classification pipeline employed in the model presented here used a combination of the following algorithms: YOLOv5, background subtraction, DeepSORT, and deep learning, as indicated in the flow chart outlined in Fig. 1.

Training and testing of the YOLOv5 and the deep-learning classifiers were conducted independently. For the YOLO model, training and testing were performed at the frame level, where a frame is intended as a single static image, using a total of 782 video samples (i.e. 15,640 frames). Each video sample consisted of 20 sequential frames extracted from the six original recordings. The data was divided into a training set, comprising of 80% of the data (625 video samples, for a total of 12,500

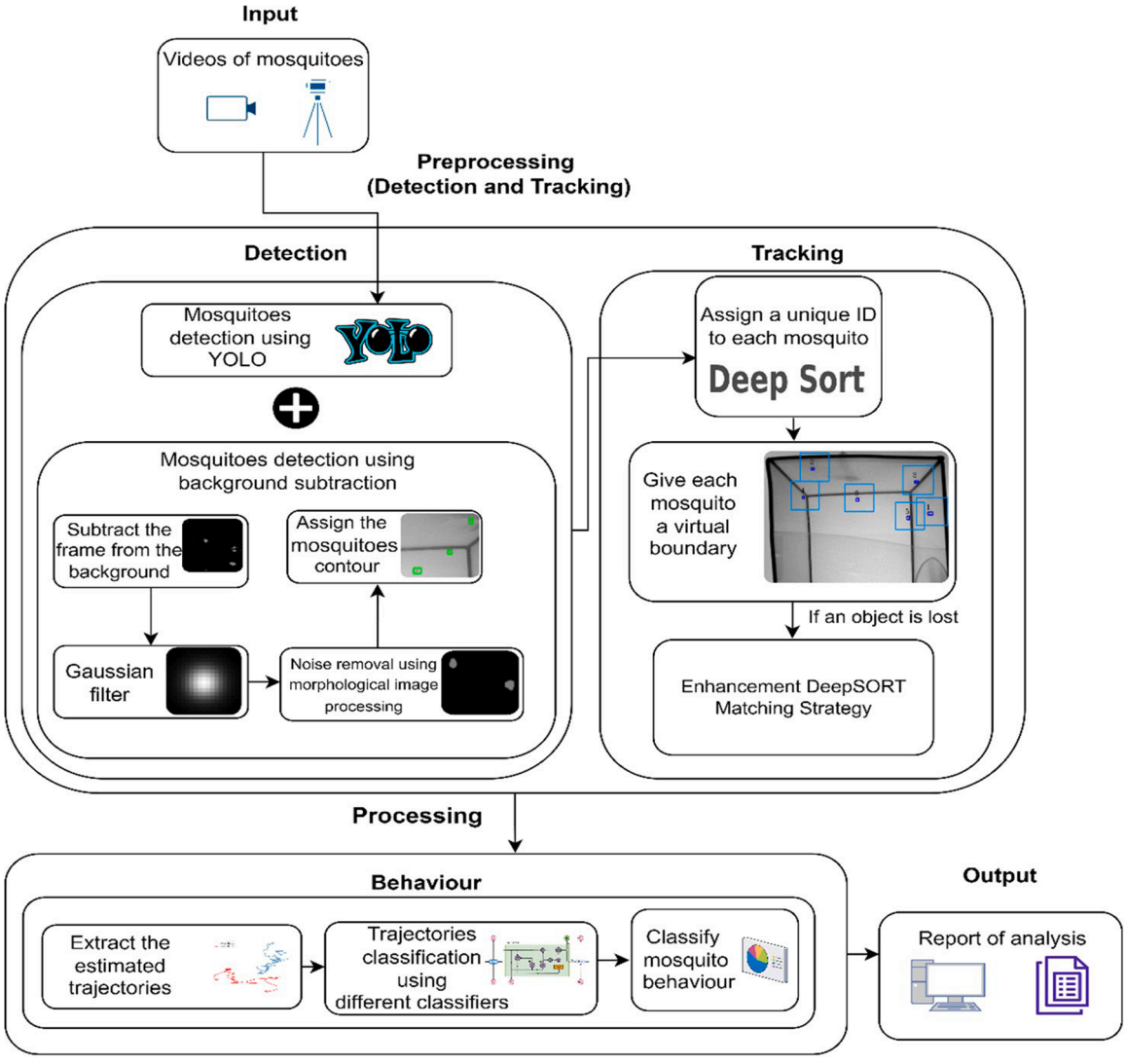


Fig. 1. Representation of the detection, tracking and behaviour classification pipeline. The data is captured as video format, which is then input in the model. The detection and tracking sequence uses YOLOv5, background subtraction, and Deep Sort algorithms. The data is then fed into the behavioural classification analysis where deep-learning classifiers produce the directional movement classification outcome. The final output of the model consists of a comprehensive report including spreadsheets and videos.

frames), and a test set, which contained the remaining 20% of the data (157 video samples, 3,140 frames). The model was trained on 100 epochs, with a detection threshold set to 70%, and all detections below this threshold were excluded from the analysis, following the protocol outlined in previous studies (Javed et al., 2023). The dataset was annotated using 'LabelImg' tool (Tzutalin, 2015). Similarly, training and testing for the deep-learning classifiers, used for the mosquito movement classification, were done using 1,248 video samples extracted from ten recorded videos. The dataset was annotated manually. Of those samples, one third (416) were used for training of the model, i.e. supervised training, one third (416) were used for testing the model, and one third (416) were used for the validation process. Each frame included the x and y coordinates of the detected mosquito's centroid (the centre of the detected mosquito) along with the rate of positional changes in both axes. The deep-learning models were trained on 150 epochs using an early stopping technique to prevent overfitting. The model outputs a spreadsheet with information regarding the frame number, the mosquito identification code (ID number), and the coordinates of the centroid, i.e. x and y coordinates.

2.2.1. Ground truth and metric based evaluation

To validate the tracking algorithm's performance and evaluate the accuracy and reliability of the model, the ground truth centroids were compared to predicted centroids. Ground truth centroids were extracted manually from the videos. These were analysed frame-by-frame, and the experimenter used a mouse cursor to mark the centroid values to the nearest pixel. The evaluation employed two metrics previously described (Javed et al., 2023): the centre error, and the average accuracy. The centre error calculates the average (Eq. (1)) between the Euclidean distance (Eq. (2)) of the predicted mosquito centroids and the ground truth centroid across n frames. As values of the centre error are averaged across multiple frames, the parameter could result in numbers that are fractions of pixels. Here, d_f denotes the Euclidean distance between estimated and ground truth pixel in frame f, where x_p and y_p represent the estimated x-axis and y-axis coordinates, and x_g and y_g represent the ground truth coordinates.

$$\text{Average} = 1 \Big/ n \sum_{f=1}^{n} df \tag{1}$$

$$d_f = \sqrt{\left((x_p - x_g)^2 + \left(y_p - y_g\right)^2\right)} \tag{2}$$

The second metric used was the average accuracy, which quantifies the absolute positional error in both axes. This metric reports an estimation of the average accuracy of the model predictions. Accurate centroid predictions have been previously defined as those in which the error between estimated and ground truth fell within 8 pixels from each other (Javed et al., 2023), therefore this value was taken as threshold for the model presented here. This corresponded to approx. 0.7 mm at the centre plane of the flight cage, based on the spatial resolution of the imaging system.

2.3. Detection and tracking of mosquitoes

Mosquito detection was achieved through the combined use of YOLOv5 and a background subtraction algorithm. While YOLOv5 is capable of operating as a standalone system on raw video data, its integration with background subtraction enhanced sensitivity to small, moving objects such as mosquitoes. This combined approach mitigated the limitations of each individual method and aimed to improve accuracy and consistency in detection outcomes. The components are described in detail below. Following detection, individual mosquito identities were assigned using the DeepSORT strategy. Trajectories were then computed by quantifying mosquito displacement over time, with outputs exported as annotated videos or spreadsheets for downstream analysis.

2.3.1. Mosquito detection based on background subtraction

The algorithm was fed raw video data, and Gaussian blur was applied to reduce high-frequency sensor noise and smooth local intensity variations prior to object detection. Subsequently, the absolute difference between the current frame (f_t) and the previous frame (f_{t-1}) was computed to highlight areas of object movement. To separate the moving object, i.e. the mosquito, from the background, a binary threshold was applied. The threshold value (T) was set to 30, after establishing it in exploratory analyses. Pixels exceeding T were classified as foreground, and were classed as white, while pixels below T were considered part of the background and were assigned the colour black. This resulted in a binary image where white regions represented moving objects while black regions indicated the background. To overcome erroneous attribution of white spots, the morphological opening operation sequence, comprising erosion and dilation, was used to eliminate noise and enhance detection (Said and Jambek, 2021). Erosion was applied to remove small white regions while dilation was set to expand white areas, thus closing gaps between regions. White regions with an area between 350 and 3500 pixels were recorded as mosquito detections; this was based on measurements of average mosquito size from sixty mosquitoes across ten videos. The contours of these valid regions were extracted and classified as mosquitoes.

2.3.2. Mosquito detection using YOLO

Given the small size of mosquito silhouettes, YOLOv5 was selected due to its superior performance in small object detection (Liu et al., 2023). The model was fed raw video data, which generated predictions for the boundary boxes, i.e. the spatial regions where mosquitoes were detected. The model provided information on the position and size of the mosquito in each image, along with the classification probability of each detection, i.e. the likelihood that the detection corresponded to a mosquito, which was set to 70% following protocols indicated in previous studies (Javed et al., 2023) (see Overall model training, testing, and output section above). The system obtained detections with high certainty and accuracy, as overlapping or duplicate detections were removed using non-maximum suppression techniques (Saidani, 2023). This process ensured that the final output, i.e. the boundary box coordinates and the classification scores, were precise and reliable.

2.3.3. Integration of the two detection models

To achieve optimal results, the background subtraction algorithm and YOLOv5 outputs were integrated. Both models processed the raw video data independently, and their outputs were merged in a complementary manner. Specifically, once the background subtraction algorithm detected a mosquito, a confirmation-check was performed by YOLOv5. If both algorithms detected a mosquito in the same area, i.e. if the areas of interest in both algorithms overlapped at any point, then the detection was confirmed (Fig. 2). In instances where the two algorithms detected the same mosquito but the outlined areas of interest did not intersect, or in cases when the YOLOv5 model failed to detect the mosquito, localisation was done using the nearest virtual tracking region created by the background subtraction algorithm (Fig. 2). A virtual tracking region was the area corresponding to the last known positions of the mosquito. Once the mosquito was identified, either by both systems or by using the nearest tracking region, the exact position of the mosquito within the boundary box was extracted using the background subtraction algorithm, as this algorithm presented a higher localisation accuracy. When the mosquito remained stationary, the background subtraction algorithm could not detect it, so only the YOLOv5 model was used (Fig. 2). This dual-model integration ensured accurate and consistent detection across scenarios, including moving and stationary mosquitoes. After detection, each mosquito was assigned a unique virtual tracking region three frames after initial detection, and it retained it for multiple frames to provide spatial references for re-identification if

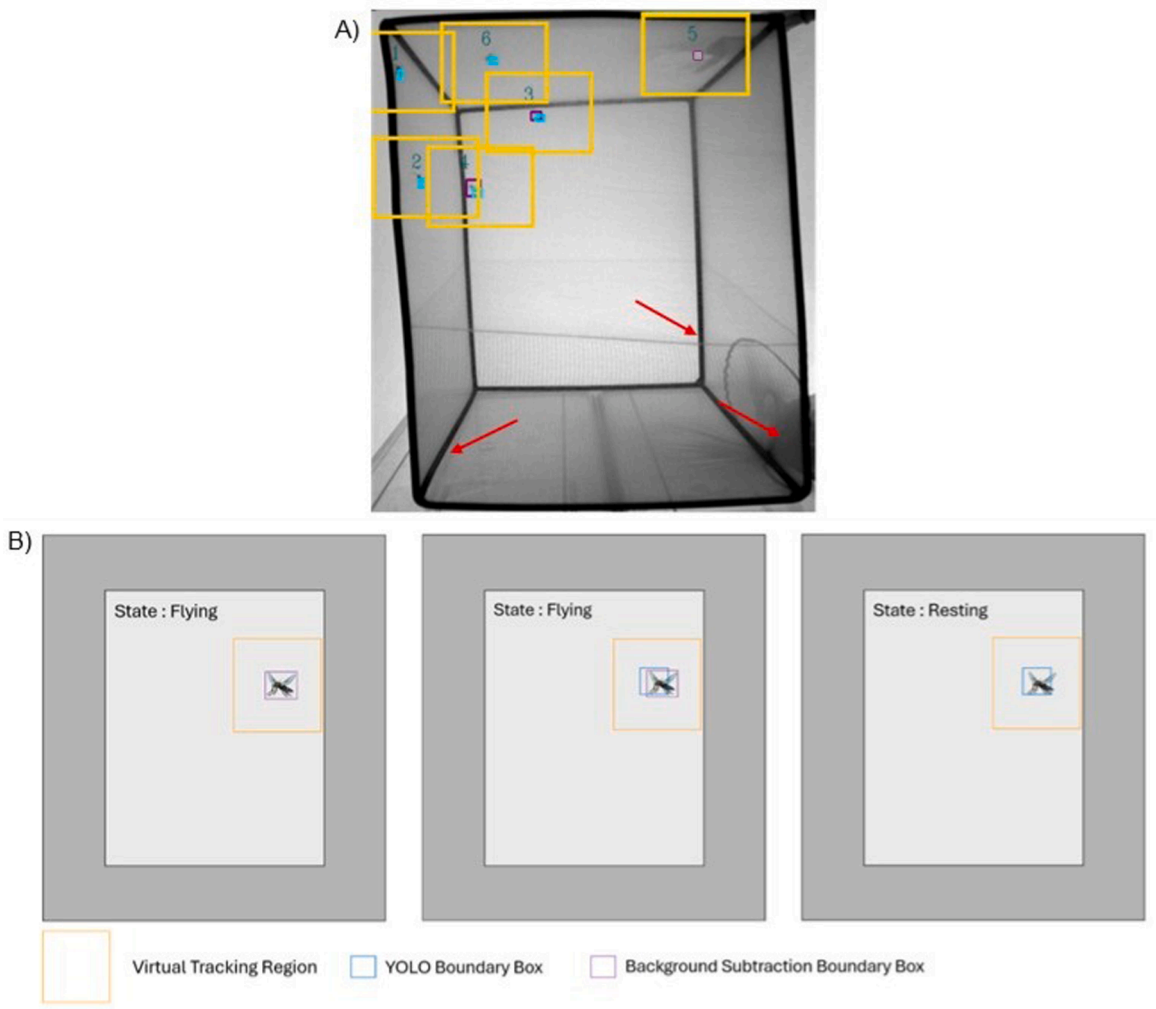


Fig. 2. Mosquito detection using the combination of YOLOv5 and background subtraction algorithm. (A) Example view of a still frame extracted from a video of Ae. aegypti. The number above the mosquito represents the mosquito ID. The yellow squares represent the virtual tracking region of each mosquito. The light blue squares indicate the boundary box from YOLOv5, while the purple squares indicate areas where mosquitoes were detected using background subtraction algorithm. Mosquitoes with ID number 1, 2, and 6 are localised using solely the YOLO model, mosquito 5 is localised using solely the background subtraction, mosquitoes 3 and 4 are localised by both methods, as the output areas of both models overlap. The red arrows highlight the black metal frames or shadowed areas that obstructed mosquito detection, as when mosquitoes flew in front of these areas their dark silhouettes merged with the dark areas. (B) Mosquito localisation based on movement state. The left diagram represents a moving mosquito localised solely using the background subtraction system (purple box), the middle diagram shows combined detection of a moving mosquito using YOLOv5 (blue box) and the subtraction model (purple box), note the overlap of the boundary boxes. The right diagram shows a resting mosquito, which is detected only by YOLOv5, as the background subtraction model cannot detect stationary objects.

detection by either model failed (Wojke et al., 2017). The distance between the virtual tracking region and the mosquito was calculated using the Euclidean distance equation (Eq. (2)).

2.3.4. Enhancement of tracking using the DeepSORT matching strategy

To enhance the tracking, a DeepSORT strategy was used. This is an advanced version of the SORT algorithm (Bewley et al., 2016). Mosquito position and coordinates were extracted from the background subtraction and YOLOv5 models and used as input for the DeepSORT strategy. This strategy used Kalman filters and Hungarian algorithms to track the objects, predict their future position in the subsequent frames, and assign a unique ID to each mosquito, and ensured that the IDs remained consistent throughout the tracking sequence. The DeepSORT model recorded the individual positions of each mosquito along with the timestamp, thus allowing for the reconstruction of their spatial displacement in x and y axes. During occlusion events the algorithm presented difficulties keeping track of the mosquitoes ID. Occlusion events occurred when two or more mosquitoes crossed paths, causing their silhouettes to merge temporarily, or when mosquito silhouettes overlapped with dark objects in the background, such as the cage frame (Fig. 2). When an occlusion event occurred, a mosquito was considered lost if it was not detected for more than three consecutive frames. In such cases, the model scanned the area around its last known position to reacquire detection (Fig. 3). Upon re-detection, a temporary new ID was assigned to the mosquito, as if the new ID were made permanent, the new track would not be matched with the previous track, thus resulting in a sequence of short fragmented contiguous tracks. To avoid this, the model checked in every frame if any mosquito's identity had changed by comparing the IDs present in the frame with those attributed at the start of the video. If a new ID was found, the model cross-checked it against original IDs, and the temporary ID was replaced with missing original ID. In instances where two or more mosquitoes' flight paths intersected, the model applied the same re-identification logic described above, with an additional step to reduce identity-swapping errors. Specifically, to preserve original ID continuity, the model analyzed the recent trajectory segments of each mosquito prior to the intersection. Based on the

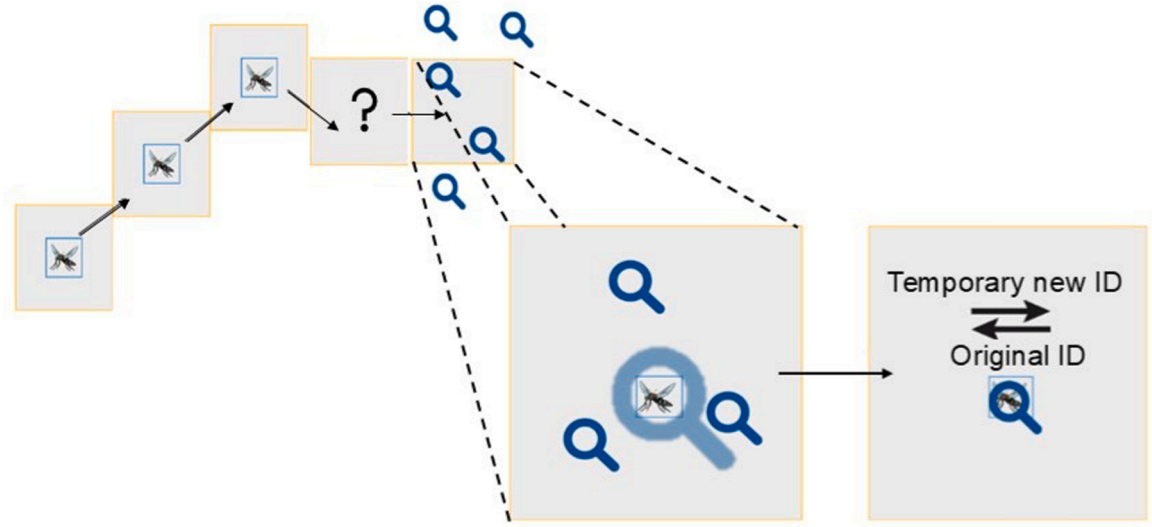


Fig. 3. Diagram showing the DeepSORT matching strategy flow with ID enhancement. The sequence shows a mosquito detected for three consecutive frames (blue boxes indicating model detection area), then becomes undetected. The model searches for the mosquito in the virtual tracking region (yellow boxes) or adjacent areas until the mosquito is found again, and the original ID is restored.

directional consistency and continuity of each path, the model inferred the most probable trajectory for each individual and reassigned the IDs accordingly. The ID re-assignation when two mosquitoes intercepted paths was manually verified in 47 cases, and it was noted that the correct ID was assigned in over 75% of the cases. While this approach did not entirely eliminate ID swaps during path crossings, it substantially reduced their occurrence, thereby improving the overall accuracy of identity maintenance across frames. Overall, this enhanced strategy mitigated the impact of occlusion events on the tracking results.

2.4. Mosquito movement classification

The section below presents the three different models used to classify mosquito directional movement, intended as the direction of displacement. Movement was divided into nine categories: movement upwards, downwards, to the left, to the right, up-left, up-right, down-left, down-right, or no movement. All three models were fed the time points, x and y coordinates of the mosquito's centroids, and the rate of change along both axes. The models assessed the direction of the movements at each timepoint. A timepoint was defined as the conjunction of 20 consecutive frames extracted from the video, with the first twelve frames deriving from the preceding timepoint, and the remaining eight frames consisting of new frames, i.e. new information. This approach allowed the model to incorporate temporal context of the previous movement, while continuously updating the input data. The output (directional movement) is saved in individual spreadsheet for each mosquito, and it can also be displayed as a stamp next to each mosquito in the video format. Given the method in which the direction of movement is assessed, the direction stamp in the video format appears with a lag time of approx. 0.8 s. A detailed explanation of each model is provided below.

2.4.1. Long short-term memory

Long Short-Term Memory (LSTM) is an advanced variant of Recurrent Neural Networks (RNN) designed to bypass the obstacle of capturing long-term information inherent in basic RNN. The LSTM cell consists of three gates: input, forget, and output, each controlling the flow of information. The input gate determines how to update the internal state using the current input and the prior state information; the forget gate decides how much of the previous state should be retained, and the output gate controls how memory data is used to generate the output (Liu et al., 2020). In this study, a Multi-Layer Sequential LSTM (MLS LSTM) was employed to analyse video sequences and classify mosquito movements. The MLS LSTM consisted of four layers, as this was found to be the optimal number of layers for accurate results (Md et al., 2023). The first layer had 64 cells, the second and third layer had 128 cells each, and the fourth contained 64 cells. The output of the four LSTM layers was then processed by three dense layers, each with 64 units. Following this, information passed a dropout layer to reduce overfitting and then was fed to a dense layer with a softmax activation function (Eq. (3)) to classify the direction of the movements. Here, where z_i represents the logit for class *i*, *e* denotes the Euler's number, and *j* indexes the summation over all classes, where *k* is the total number of classes.

$$\mathrm{softmax}(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{k} e^{z_j}} \tag{3}$$

2.4.2. Gated recurrent unit

Gated Recurrent Unit (GRU) is a variant of RNN architecture that addresses short-term memory limitations, while offering a simpler structure compared to the LSTM model (Abbaspour et al., 2020). The GRU model merges the input and the forget gates into a single update gate. A GRU unit consists of three main parts: an update gate, a reset gate, and current memory content. The Multi-Layer Sequential GRU model used here contained four GRU layers responsible for capturing temporal dependencies from input data. The first two layers had 64 cells each, while the last two had 128 cells each. This was followed by three dense layers, activated by the Rectified Linear Unit (ReLU) function (Eq. (4)), which outputs the input value if it is positive and zero if negative. In Eq. (4), *x* represent the scalar input of the activation function, *0* denotes the constant threshold used to suppress negative values, while max is the maximum function that returns the larger value between 0 and x. To reduce overfitting, a dropout layer with a rate of 0.3 was applied after each dense layer, randomly setting 30% of the input units to zero, thereby preventing the model from overly rely on any one unit. Finally, the information was moved to a dense layer with five neurons activated by the softmax function, and the final movement classification was chosen based on the highest probability.

$$ReLU(x) = \max(0, x) \tag{4}$$

2.4.3. Convolutional neural networks and long short-term memory network

The Convolutional Neural Networks and Long Short-Term Memory (CNN-LSTM) is a hybrid architecture that combines CNN and LSTM

layers for classification tasks (de Costa, 2020). The CNN component consisted of two convolutional layers, each using the ReLU activation function to introduce non-linearity by setting negative values to zero. These layers were followed by max-pooling layers, which reduced spatial dimensions by selecting the maximum value within each pooling window, thereby down-sampling the input and enhancing computational efficiency. The LSTM component captured temporal dependencies within sequential data. It comprised of four LSTM layers with 50, 128, 128, and 64 cells, respectively, enabling the model to remember temporal patterns over different time spans. A dropout layer followed the LSTM layers to prevent overfitting, after which a final dense layer activated by softmax function produced the classification of the direction of movement.

2.5. Heatmap extraction

This method was employed to generate a comprehensive heatmap that visually represented mosquito activity levels by tracking their physical displacement over time. The model returned an image or video that shows mosquitoes trail, where the trails were colour-coded based on the movement rate per unit area, with blue indicating low activity, green and yellow indicating moderate activity, and red denoting high activity. Resting mosquitoes did not generate any colour trail and were instead visualised as raw static objects in the video. As the system employed a single camera, the image generated was bidimensional; consequently, in the rare event that a mosquito moved exclusively along the third plane (i.e. depth), it would have been recorded as stationary. Based on behavioural observations, such movement patters were considered unlikely, and this limitation was therefore deemed acceptable for the purpose of the system. To create the heatmap, firstly, each mosquito had to be detected and assigned a unique identifier. The pixels of the boundary box containing the mosquitoes were rendered black, i.e. were set to an intensity value of 1. Subsequently, the intensity value of the pixels within the boundary box was set to dynamically increase based on movement rate per frame. This step was crucial for standardising movement per unit area and ensuring that the final output accurately reflected the frequency of mosquito movements within specific regions. Independent heatmaps were initially generated for each mosquito and subsequently combined for all mosquitoes in the frame into a single composite heatmap. In cases where mosquito paths intersected, the pixel colour at the intersection was set to the highest intensity value from among the overlapping images. This approach facilitated a clearer representation of movement patterns and areas of preference, emphasizing regions with high activity. The resulting composite image, initially black, was normalised by dividing each pixel intensity by the maximum intensity observed in the image, thus ensuring consistent pixel intensities across the heatmap, allowing for meaningful comparisons. Following normalisation, pixel intensity was scaled by multiplying each value by 255, thus adjusting the values to range from 0 (black) to 255 (white) according to the universal pixel format. This step enhanced the visual clarity of the heatmap, as it provided information at a higher resolution. Finally, the grayscale image was converted to the BGR colour scale and then transformed into a jet colour map. This visual enhancement was particularly effective for identifying hotspots of mosquito activity, offering clear and intuitive representation of movement patterns.

3. Results

3.1. Mosquito detection and tracking

The model developed and tested as described above was then used to analyse an independent additional set of videos, two per species, for a total of six videos analysed. For *Ae. aegypti* and *An. coluzzii*, a total of 1,000 frames were examined for each video, with four mosquitoes flying and two mosquitoes resting in all frames. For *Cx. quinquefasciatus*, a total of 1,050 frames per video were analysed, and in all of them three mosquitoes were flying and three were resting. By combining YOLOv5 and the background subtraction algorithms, the proposed model successfully detected and tracked mosquitoes in the x and y axes, regardless of whether individuals were in motion or stationary. When comparing the model flight paths to ground truth trajectories, results indicated a high reliability and precision which remained consistent across all species (Fig. 4), demonstrating that the model can be used across species with different morphological and behavioural characteristics, e.g. body size, and flight speed.

Furthermore, the integration of the DeepSORT matching strategy was instrumental in maintaining mosquito identity over time, including after occlusion events. Without this enhancement, the model would assign a new ID to mosquitoes re-detected after occlusions, resulting in fragmented tracks and incorrect mosquito counts (Video S1). It is important to note that although DeepSORT re-identification was consistently accurate in cases of background-induced occlusion, its reliability was reduced during mosquito path interceptions, where occasional ID swaps were observed (correct ID reassignment confirmed in over 75% of manually verified cases). Nonetheless, the DeepSORT enhancement allowed for long, continuous tracks and preserved the original identification codes (Video S1).

The performance of the tracking system varied depending on the mosquito species and the detection methodology employed, specifically, on whether only the standalone YOLOv5 model was used or whether the new combination of YOLOv5 and background subtraction method was used. Each video was evaluated using different three threshold tolerance values: eight, six, and four pixels. For each threshold, the models were run in triplicates, and the average accuracy was calculated as a mean of the three runs. Detection accuracy for resting mosquitoes was 100% across all conditions. While both methodologies accurately detected and tracked mosquitoes, the proposed model offered the advantage of better managing occlusion and variable light conditions, thus yielding a better performance. Even when using the most restrictive threshold value (four pixels), the proposed methodology maintained an average accuracy above 97%, and in five cases out of six the average accuracy was over 99%. In contrast, YOLOv5 alone achieved average accuracies below 99% in five cases out of six (Table 1). The superiority in the performance of the proposed methodology was also reflected in the mean centre error, which was consistently lower in the combined methodology compared to the use of YOLOv5 alone (Table 1). Particularly, in video number two of *Ae. aegypti*, the mean centre error fell over two-fold, from 1.7 to 0.72, when the methodology incorporated background subtraction. Similarly, in video number one of *An. coluzzii*, the metric decreased over two-fold, from 0.294 to 0.089, by incorporating the use of the background subtraction. Overall, the proposed methodology outperformed the standalone YOLOv5 model, yielding overall higher accuracy and lower centre error values across all tested conditions.

3.2. Directional movement classification

For *Ae. aegypti, An. coluzzii*, and *Cx. quinquefasciatus*, the model was evaluated on 4,260, 4,520, and 4,860 frames respectively. Ground truth movement directions were manually annotated for validation. The results indicate that the models were able to detect mosquito movement direction, with the GRU model achieving the highest performance across all evaluation metrics (Table 2). Specifically, the GRU model obtained values over 97% accuracy for the three mosquito species, whereas the other models produced lower accuracies, ranging from 65% to 72%. These findings highlight the GRU superior ability to capture and classify directional movement patterns, establishing it as the best model for this application. In addition to quantitative outputs, the model generated annotated videos displaying the directional movement tags adjacent to each mosquito (Video S2), rendering possible to visually validating the movement predictions, and facilitating intuitive interpretation of directional movement over time.

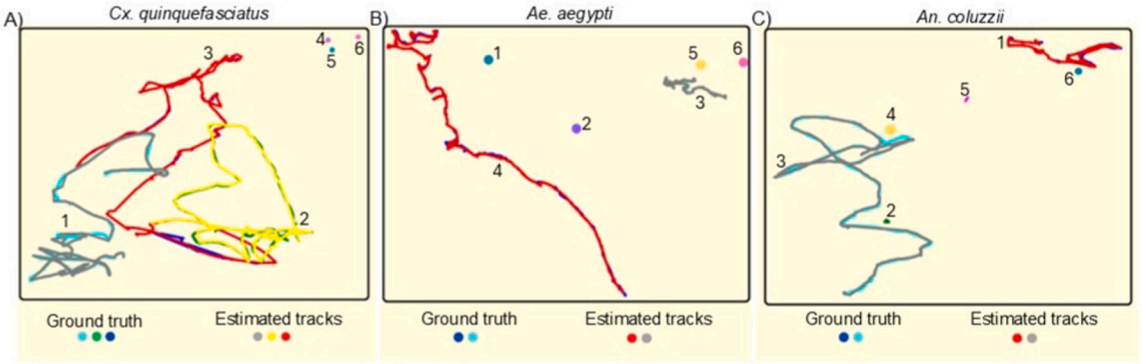


Fig. 4. Comparison between the model-estimated and ground truth trajectories for flying and resting mosquitoes across the three tested species. (A) Cx. quinquefasciatus, (B) Ae. aegypti, and (C) An. coluzzii. In each panel, the estimated trajectories closely overlap with the ground truth trajectories, underscoring the high accuracy and reliability of the proposed model.

Table 1
Mosquito tracking evaluation results comparing average accuracies and mean centre errors for the two methodologies, the standalone YOLOv5 and the proposed methodology combining YOLOv5 and background subtraction. Accuracy was calculated for three different pixel tolerances (eight, six, and four). The mean centre error indicates the deviation between the model-predicted centre point and the ground truth centroid position.

Methodology	Mosquito species	Video	Average Accuracy 8-Pixels	Average Accuracy 6-Pixels	Average Accuracy 4-Pixels	Mean centre error
YOLOv5	*Ae. aegypti*	One	99.43%	99.23%	98.3%	0.62
Proposed methodology	*Ae. aegypti*	One	99.8%	99.56 %	99.03%	0.5
YOLOv5	*Ae. aegypti*	Two	97.76%	97.06%	94%	1.7
Proposed methodology	*Ae. aegypti*	Two	99.4%	99.16%	97.1%	0.72
YOLOv5	*Cx. quinquefasciatus*	One	99.7%	99.3%	98.3%	0.34
Proposed methodology	*Cx. quinquefasciatus*	One	99.87%	99.5%	99%	0.174
YOLOv5	*Cx. quinquefasciatus*	Two	99.0%	98.7%	98%	0.44
Proposed methodology	*Cx. quinquefasciatus*	Two	99.8%	99.6%	99.2%	0.3
YOLOv5	*An. coluzzii*	One	99.46%	99.2%	98.9%	0.294
Proposed methodology	*An. coluzzii*	One	99.94%	99.91%	99.87%	0.089
YOLOv5	*An. coluzzii*	Two	99.8%	99.7%	99.6%	0.353
Proposed methodology	*An. coluzzii*	Two	99.83%	99.83%	99.7%	0.241

Table 2
Results of the directional movement classification models. Models were evaluated using the following metrics: accuracy, precision, recall, and F1-Score values.

Mosquito species	Model	Accuracy	Precision	Recall	F1-Score
Ae. aegypti	GRU	97.56%	97.825%	97.56%	97.65%
	LSTM	66.5%	69.05%	62%	65.33%
	CNN-LSTM	69%	68%	63.7%	65.78%
An. coluzzii	GRU	97.78%	97.9%	97.8%	97.74%
	LSTM	72.44%	65.4%	63%	62.4%
	CNN-LSTM	71.56%	73%	72%	71%
Cx. quinquefasciatus	GRU	97.4%	97.3%	97.1%	97.1%
	LSTM	65.9%	60.4%	65%	60%
	CNN-LSTM	66.5%	68.5%	66%	66%

3.3. Heatmaps

Activity patterns were visualised using heatmaps, which highlighted mosquito movement rate per space unit. The system was tested with videos of cages containing either male or female mosquitoes. Heatmaps were generated consistently in all cases where mosquitoes were detected and tracked successfully, and this was irrespective of the species or sex of the mosquitoes. Heatmaps provided a clear visual interpretation of activity levels, capturing both static positions and dynamic flight behaviours, and revealed spatial patterns of high and low activity within the cage (Fig. 5). To test whether heatmaps could provide biologically meaningful information, such as spatial preference, heatmaps were generated from videos of female mosquitoes exposed to host cues, as host-seeking behaviour is well-characterised and therefore could serve as positive control. The resulting heatmaps indicated that host-seeking females consistently concentrated their activity in areas near the host cue source (Fig. 5, Video S3), with repeated visits and hovering behaviour in the relevant area. This outcome aligned with expectations and supported the value of heatmaps for rapid visual interpretation of mosquito activity levels, and by proxy, behavioural patterns.

4. Discussion

The model developed in this study integrates three algorithms, namely background subtraction, YOLOv5, and DeepSORT, for the detection and trajectory tracking of mosquitoes. The integration leverages the strengths and the different functions of each component (Huang et al., 2020). YOLOv5 offers robust object detection, enhanced by the background subtraction component, and DeepSORT resolves the identity inconsistencies created after occlusion events, particularly those caused by background clutter, thus providing continuous long tracks. However, the integration of multiple algorithms requires careful handling to ensure compatibility of information flow, and it may also impose significant computational demands, with high-resolution video data. Therefore, balancing model complexity and system performance is essential. The model presented here achieves this balance, operating without noticeable delays while delivering precise tracking, as highlighted by the validation using ground truth data manually extracted from the videos.

Behavioural classification has been a key focus in ethological studies, with automated extraction of information from large datasets being a

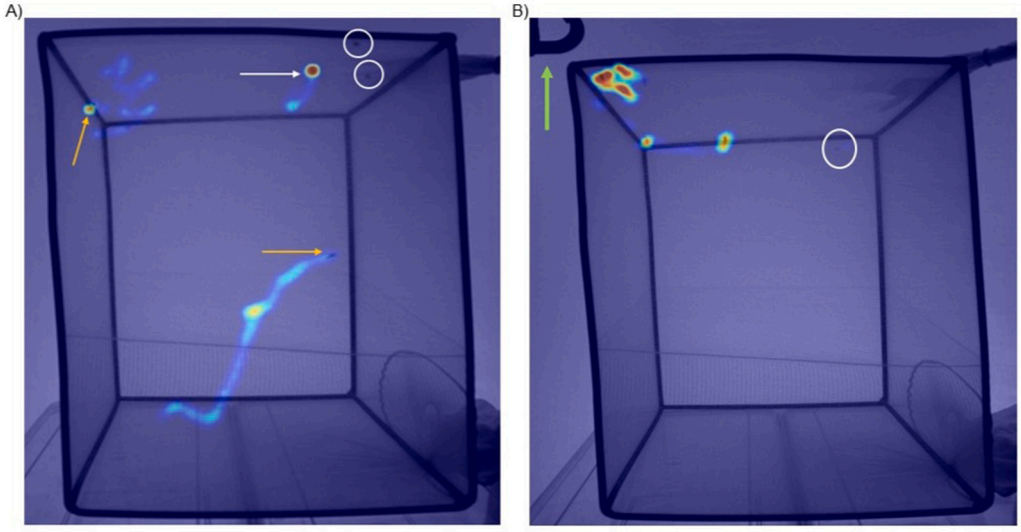


Fig. 5. Example of a heatmaps produced from two-second videos of six Ae. aegypti females. (A) Non-host-seeking females, (B) host-seeking females. Heatmap use colour gradient to display movement intensity, with blue indicating low activity, green and yellow moderate activity, and red high activity rate per unit area. In (A) the white circles indicate the position of two static mosquitoes, the white arrow indicates a mosquito hovering near the top of the cage (red area), and the orange arrows indicate two flying mosquitoes. In (B) three mosquitoes can be seen hovering on the top left corner of the cage, near the host cue source, indicated by the green arrow. Two mosquitoes cam be seen flying near the top side of the cage, and the sixth mosquito remained static at the rear of the cage, position indicated by the white circle.

desirable goal for researchers studying different organisms (Kain et al., 2013). Advancements in tracking technologies and equipment have permitted simultaneous observation of multiple individuals and have enhanced the efficiency and the complexity of experimental designs (Wang et al., 2024). Recent applications of deep learning models to animal tracking and behaviour demonstrated the versatility and accuracy of AI-based classification methods across different organisms and experimental settings. For example, YOLOv4 and Region-Based Convolutional Neural Networks (RCNN) were employed to analyse the following behaviours in pen-kept goats: moving, eating, drinking, and stationary (Jiang et al., 2020). They achieved high accuracy rates of 96.86%, 97.87%, 98.27%, and 96.92%, respectively. Similarly, in another study applied YOLOv5 was applied to recognize three sheep behaviours, namely feeding, standing, and lying down (Hu et al., 2023). Using a cross-validation approach, in which the model was trained on images from handheld cameras and evaluated on images from fixed cameras (and vice versa), they obtained mean average precisions ranging from 72.4% to 83.5%. In the present study, directional movement classification of mosquitoes was achieved using three different models, with classification accuracies spanning from 65.9% to 97.7%, and precisions ranging from 60.4% to 97.9%, depending on the model used and the mosquito species analysed. Cross-validation in this case was not used to avoid data leakage and bias resulting from cross-validation on correlated samples; instead, the dataset was split into separate samples to be used independently for either training, validation, and testing. The GRU architecture consistently outperformed the two other models, across all tested species and evaluation metrics. Due to its simpler architecture and fewer parameters, GRU is particularly advantageous when working with smaller datasets or under limited computational resources (Yang et al., 2020). Here we show that GRU delivered stable and robust performance with minimal need for adjustments, further supporting its suitability in systems aimed at streamlining the workflow. Directional classification can be a useful resource in the interpretation of mosquito movement to or from a specific treatment. For example, this classification could quantify the number or proportion of mosquitoes approaching an attractant or trap, or avoiding specific stimulus, thus offering a practical tool for evaluating intervention efficacy. By embedding temporal information alongside spatial coordinates, the model also enables assessments of mosquitoes responsiveness over time, allowing to differentiate between transient and sustained behavioural responses, which is critical when evaluating dynamics of vector-stimulus interaction. Moreover, heatmaps generated from mosquito flight data allow for intuitive and rapid interpretation of activity levels and spatial preferences. For example, when applied to assays with host-seeking females, the heatmaps accurately indicated areas of aggregation near the host-cues. These dual visualisation tools, i.e. heatmaps and directional classification videos, enable rapid identification of activity patterns and behavioural hotspots, and could be incorporated into behavioural assays aimed at evaluating mosquito response to attractants or repellents.

Scalability is an important consideration in automated behavioural analysis pipelines, particularly with respect to the number of tracked individuals and recording duration. Object detection using YOLO-based architectures enables high frame rates inferences, making them suitable for long recordings. The DeepSORT tracking stage introduces relatively limited overhead because it relies on efficient Kalman filtering for motion prediction and the Hungarian algorithm for frame-to-frame association, both of which scale efficiently with the number of objects. As a result, the computational complexity of the tracking component increases approximately linearly with the number of individuals and video length, allowing the pipeline to remain practical for behavioural experiments involving extended recordings or moderate group sizes. In our setup, videos were recorded at 25 frames per second and analysed frame-by-frame without noticeable delays, indicating that the approach is suitable for routine behavioural assays involving small mosquito groups. Importantly, the computational load was primarily driven by the detection stage rather than by trajectory linking, suggesting that further optimisation, such as the use of other YOLO variants, could allow the pipeline to scale to longer recordings or larger insect groups. Similar optimisation strategies have been proposed in recent multi-object tracking studies to maintain efficiency in computationally constrained environment (Magdin and Balogh, 2025).

The design of the current study has some limitations. Models that use YOLO require a well-annotated training dataset to achieve optimal performance (Terven et al., 2023), which could represent a considerable burden for the experimenter, especially at early stages of the study when labelled data are not yet available. However, the initial effort investment is offset by the speed and scalability at which, after the model is trained,

the data is analysed. The proposed model is better suited for projects with large datasets or that involve repeated measurements, whereas it would be less appropriate for those where only a few videos require analysis. Additionally, upscaling the system to monitor large areas would necessitate greater computational power, which comes associated with increased operational costs (Jain et al., 2019). A second limitation is that summary statistics of contiguous track length were not recorded, as this metric was not defined within the original analysis framework. Such measures could provide additional insight into trajectory characteristics and should be considered in future studies. Another limitation is the current restriction to two-dimensional tracking. While the current methodology proved effective in tracking mosquitoes in the x and y axes, future studies should aim to incorporate 3D tracking and 3D directional classification outputs. Lastly, maintaining accurate mosquito identity across frames is contingent on several factors, including video quality and the number of insects tracked. In videos featuring higher densities of flying mosquitoes, particularly when multiple individuals move simultaneously or intersect flight paths, the likelihood of ID swaps increases, resulting in occasional interruptions to trajectory continuity. This challenge is exacerbated under visually complex backgrounds or suboptimal lighting, which can hinder detection and feature matching. While our experiments focused on relatively short laboratory recordings, previous studies have successfully tracked mosquitoes in more complex or extended scenarios, including mating swarms in the field (Butail et al., 2012), flights near insecticide-treated bed nets with heterogeneous lighting (Angarita-Jaimes et al., 2016; Voloshin et al., 2020), and flights around house entry points in field settings (Spitzen et al., 2025). While the current model demonstrated reliable identity preservation in low-to-moderate density settings, further studies using larger and more complex datasets are recommended to assess scalability and improve tracking robustness under more demanding experimental conditions.

5. Conclusion

The work presented here builds upon recent advance in deep learning technologies and focusses on mosquito activity and behaviour, offering a streamlined, scalable solution for automated behavioural analysis. Crucially, our methodology does not require specialized high-tech equipment or complex set-ups. Accurate data were obtained from simple 2D videos recorded using a single camera, making this approach accessible for laboratory operating with limited resources. This is particularly relevant for entomological research in low-resource or field settings, where high-end imaging equipment may not be available. Additionally, the model provides automated analysis with minimal intervention, as once trained, the model returns features such as tracking, directional classification, and heatmaps in a fully automated manner. Although the system was tailored to mosquitoes or small flying insects, it can be used to study other organisms. For instance, the same model was successfully employed to track rat movements (Mostafa et al., 2024), demonstrating its broader potential. Overall, our streamlined robust approach constitutes a valuable tool for mosquito behavioural research, and more broadly, for behavioural entomology, enabling detailed, quantitative analysis of mosquito flight dynamics and activity levels, which could be used in the study of novel surveillance or control tools.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Manuela Carnaghi: Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Investigation, Data curation. **Khaled Mostafa:** Writing – original draft, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Mohamed Hany:** Writing – original draft, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Ayman Atia:** Supervision, Software, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank Prof. Richard Hopkins for early discussions that helped shape the conceptual framework of this project. We are also grateful to Natalie Morley for her assistance in insect rearing. Finally, we extend our sincere thanks to Dr Felix Hol for his constructive feedback during manuscript preparation.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.actatropica.2026.108120.

Data availability

Data will be made available on request.

References

Abbaspour, S., Fotouhi, F., Sedaghatbaf, A., Fotouhi, H., Vahabi, M., Linden, M., 2020. A comparative analysis of hybrid deep learning models for human activity recognition. Sensors 20, 5707.

Angarita-Jaimes, N.C., Parker, J.E.A., Abe, M., Mashauri, F., Martine, J., Towers, C.E., McCall, P.J., Towers, D.P., 2016. A novel video-tracking system to quantify the behaviour of nocturnal mosquitoes attacking human hosts in the field. J. R. Soc. Interface 13.

Bewley, A., Ge, Z., Ott, L., Ramos, F., Upcroft, B., 2016. Simple online and realtime tracking. In: IEEE International Conference on Image Processing (ICIP). IEEE, pp. 3464–3468.

Bredt, B.H., Tripet, F., Müller, P., 2025. Revealing complex mosquito behaviour: a review of current automated video tracking systems suitable for tracking mosquitoes in the field. Parasit. Vect. 18, 66.

Butail, S., Manoukis, N., Diallo, M., Ribeiro, J.M., Lehmann, T., Paley, D.A., 2012. Reconstructing the flight kinematics of swarming and mating in wild mosquitoes. J. R. Soc. Interf. 9, 2624–2638.

Cannet, A., Simon-Chane, C., Akhoundi, M., Histace, A., Romain, O., Souchaud, M., Jacob, P., Sereno, D., Mouline, K., Barnabe, C., Lardeux, F., Boussès, P., Sereno, D., 2023. Deep learning and wing interferential patterns identify Anopheles species and discriminate amongst Gambiae complex species. Sci. Rep. 13, 13895.

Carnaghi, M., Belmain, S.R., Hopkins, R.J., Hawkes, F.M., 2021. Multimodal synergisms in host stimuli drive landing response in malaria mosquitoes. Sci. Rep. 11. https://doi.org/10.1038/s41598-021-86772-4.

Carnaghi, M., Mandelli, F., Feugère, L., Joiner, J., Belmain, S.R., Hopkins, R.J., Hawkes, F.M., 2024a. Protocol for rearing and using mosquitoes for flight path tracking and behavioral characterization in wind tunnel bioassays. STAR Protoc. 5, 103180.

Carnaghi, M., Mandelli, F., Feugère, L., Joiner, J., Young, S., Belmain, S.R., Hopkins, R.J., Hawkes, F.M., 2024b. Visual and thermal stimuli modulate mosquito-host contact with implications for improving malaria vector control tools. iScience 27. https://doi.org/10.1016/j.isci.2023.108578.

de Costa, R.L., 2020. Convolutional-LSTM networks and generalization in forecasting of household photovoltaic generation. Eng. Appl. Artif. Intell. 116, 105458.

Diwan, T., Anirudh, G., Tembhurne, J.V., 2023. Object detection using YOLO: challenges, architectural successors, datasets and applications. Multimed. Tools Appl. 82, 9243–9275.

Gibson, G., 1995. A behavioural test of the sensitivity of a nocturnal mosquito, Anopheles gambiae, to dim white, red and infra-red light. Physiol. Entomol. 20, 224–228.

Hashimoto, D.A., Rosman, G., Rus, D., Meireles, O.R., 2018. Artificial intelligence in surgery: promises and perils. Ann. Surg. 268, 70–76.

Hu, T., Yan, R., Jiang, C., Chand, N.V., Bai, T., Guo, L., Qi, J., 2023. Grazing sheep behaviour recognition based on improved yolov5. Sensors 23, 4752.

Huang, J.C., Ko, K.M., Shu, M.H., Hsu, B.M., 2020. Application and comparison of several machine learning algorithms and their integration models in regression problems. Neural Comput. Appl. 32, 5461–5469.
Jain, S., Ananthanarayanan, G., Jiang, J., Shu, Y., Gonzalez, J., 2019. Scaling video analytics systems to large camera deployments. In: Proceedings of the 20th International Workshop on Mobile Computing Systems and Applications. Santa Cruz, CA, USA. Association for Computing Machinery, pp. 9–14.
Javed, N., López-Denman, A.J., Paradkar, P.N., Bhatti, A., 2024a. FlightTrackAI: a robust convolutional neural network-based tool for tracking the flight behaviour of Aedes aegypti mosquitoes. R. Soc. Open Sci. 11, 240923.
Javed, N., Paradkar, P.N., Bhatti, A., 2024b. An overview of technologies available to monitor behaviours of mosquitoes. Acta Trop. 258, 107347.
Javed, N., Paradkar, P.N., Bhatti, A., 2023. Flight behaviour monitoring and quantification of Aedes aegypti using convolution neural network. PLoS One 18, e0284819.
Jiang, M., Rao, Y., Zhang, J., Shen, Y., 2020. Automatic behavior recognition of group-housed goats using deep learning. Comput. Electron. Agric. 177, 105706.
Joksimovic, S., Ifenthaler, D., Marrone, R., De Laat, M., Siemens, G., 2023. Opportunities of artificial intelligence for supporting complex problem-solving: findings from a scoping review. Comput. Educ. Artif. Intell. 4, 100138.
Kain, J., Stokes, C., Gaudry, Q., Song, X., Foley, J., Wilson, R., de Bivort, B., 2013. Leg-tracking and automated behavioural classification in Drosophila. Nat. Commun. 4, 1910.
Kitaguchi, D., Takeshita, N., Hasegawa, H., Ito, M., 2021. Artificial intelligence-based computer vision in surgery: recent advances and future perspectives. Ann. Gastroenterol. Surg. 6, 29–36.
Kittichai, V., Pengsakul, T., Chumchuen, K., Samung, Y., Sriwichai, P., Phatthamolrat, N., Tongloy, T., Jaksukam, K., Chuwongin, S., Boonsang, S., 2021. Deep learning approaches for challenging species and gender identification of mosquito vectors. Sci. Rep. 11, 4838.
Kong, M., Peng, R., 2023. Automated Animal Tracking for Behavioral Experiments. In: Proceedings of the 2022 9th International Conference on Biomedical and Bioinformatics Engineering, ICBBE. Association for Computing Machinery, New York, pp. 288–291.
Liu, P., Xie, Z., Li, T., 2023. Traffic light detection based on improved YOLOV5. In: 8th International Conference on Intelligent Computing and Signal Processing (ICSP). IEEE, pp. 1891–1895.
Liu, W., Pan, J., Ren, Y., Wu, Z., Wang, J., 2020. Coupling prediction model for long-term displacements of arch dams based on long short-term memory network. Struct. Control Heal. Monit. 27, e2548.
Maekawa, T., Ohara, K., Zhang, Y., Fukutomi, M., Matsumoto, S., Matsumura, K., Shidara, H., Yamazaki, S.J., Fujisawa, R., Ide, K., Nagaya, N., Yamazaki, K., Koike, S., Miyatake, T., Kimura, K.D., Ogawa, H., Takahashi, S., Yoda, K., 2020. Deep learning-assisted comparative analysis of animal trajectories with DeepHL. Nat. Commun. 11, 5316.
Magdin, M., Balogh, Z., 2025. Comparison classification algorithms and the YOLO method for video analysis and object detection. Sci. Rep. 15, 25432.
Md, A.Q., Kapoor, S., Junni, A.V.C., Sivaraman, A.K., Tee, K.F., 2023. Novel optimization approach for stock price forecasting using multi-layered sequential LSTM. Appl. Soft Comput. 134, 109830.
Morris, B.I., Kittredge, M.J., Casey, B., Meng, O., Chagas, A.M., Lamparter, M., Thul, T., Pask, G.M., 2022. PiSpy: an affordable, accessible, and flexible imaging platform for the automated observation of organismal biology and behavior. PLoS One 17, e0276652.
Mostafa, K., Hany, M., Ashraf, A., Zaafan, M., Atia, A., 2024. Mapping Rats Movements: tracking and Visualizing Heatmaps and Trajectories Using Computer Vision. Intelligent Methods, Systems, and Applications (IMSA). IEEE, pp. 160–165.
Murthy, C.B., Hashmi, M.F., Bokde, N.D., Geem, Z.W., 2020. Investigations of object detection in images/videos using various deep learning techniques and embedded platforms—a comprehensive review. Appl. Sci. 10, 3280.
Qureshi, Y.M., Voloshin, V., Guy, A., Ranson, H., McCall, P.J., Covington, J.A., Towers, C.E., Towers, D.P., 2025. Machine learning reveals immediate disruption in mosquito flight when exposed to Olyset nets. Curr. Res. Parasitol. Vector-Borne Dis. 7, 100273.
Rane, N., Choudhary, S., Rane, J., 2023. YOLO and Faster R-CNN object detection in architecture, engineering and construction (AEC): applications, challenges, and future prospects. SSRN [Preprint].
Said, K.A.M., Jambek, A.B., 2021. Analysis of image processing using morphological erosion and dilation. J. Phys. Conf. Ser. 2071, 012033.
Saidani, T., 2023. Deep learning approach: yOLOv5-based custom object detection. Eng. Technol. Appl. Sci. Res. 13, 12158–12163.
Spitzen, J., Lankheet, M.J., Pieters, R.P., Gadamika, M., Phiri, I., Cribellier, A., Logan, J. G., Constantianus, J.M., Phiri, K.S., Muijres, F.T., McCann, R.S., 2025. The effect of eave and window modifications on house entry behavior of Anopheles gambiae. Parasit. Vect. 18, 251.
Spitzen, J., Takken, W., 2018. Keeping track of mosquitoes: a review of tools to track, record and analyse mosquito flight. Parasit. Vect. 11. https://doi.org/10.1186/s13071-018-2735-6.
Sumpter, D., Pratt, S., 2003. A modelling framework for understanding social insect foraging. Behav. Ecol. Sociobiol. 53, 131–144.
Terven, J., Córdova-Esparza, D.M., Romero-González, J.A., 2023. A comprehensive review of yolo architectures in computer vision: from yolov1 to yolov8 and yolo-nas. Mach. Learn. Knowl. Extr. 5, 1680–1716.
Tzutalin D., 2015. Human signal label image. URL: https://github.com/tzutalin/labelImg.
Voloshin, V., Kröner, C., Seniya, C., Murray, G.P., Guy, A., Towers, C.E., McCall, P.J., Towers, D.P., 2020. Diffuse retro-reflective imaging for improved video tracking of mosquitoes at human baited bednets. R. Soc. Open Sci. 7.
Wang, Y.H., Hsieh, J.W., Chen, P.Y., Chang, M.C., So, H.H., Li, X., 2024. SMILEtrack: similarity learning for occlusion-aware multiple object tracking. In: Proceedings of the AAAI Conference on Artificial Intelligence, 38, pp. 5740–5748 p.
Wojke, N., Bewley, A., Paulus, D., 2017. Simple online and realtime tracking with a deep association metric. In: IEEE International Conference on Image Processing (ICIP). IEEE, pp. 3645–3649.
World Health Organization, 2024. Vector-borne diseases [WWW Document]. https://www.who.int/news-room/fact-sheets/detail/vector-borne-diseases.
Xu, Y., Liu, X., Cao, X., Huang, C., Liu, E., Qian, S., Liu, X., Wu, Y., Dong, F., Qiu, C.W., Qiu, J., Hua, K., Su, W., Wu, J., Xu, H., Han, Y., Fu, C., Yin, Z., Liu, M., Roepman, R., Dietmann, S., Virta, M., Kengara, F., Zhang, Z., Zhang, L., Zhao, T., Dai, J., Yang, J., Lan, L., Luo, M., Liu, Z., An, T., Zhang, B., He, X., Cong, S., Liu, X., Zhang, W., Lewis, J.P., Tiedje, J.M., Wang, Q., An, Z., Wang, F., Zhang, L., Huang, T., Lu, C., Cai, Z., Wang, F., Zhang, J., Research, A., Intelligence: A. powerful paradigm for scientific, 2021. Artificial intelligence: a powerful paradigm for scientific research. Innovation 2, 100179.
Yang, S., Yu, X., Zhou, Y., 2020. Lstm and gru neural network performance comparison study: taking yelp review dataset as an example. International Workshop on Electronic Communication and Artificial Intelligence (IWECAI). IEEE, pp. 98–101.