

Review Article

Identifying disease vector images in the Americas in the age of artificial intelligence

Vinícius Lima de Miranda^[1] , José Fabrício de Carvalho Leal^[1] ,
Isadora Ribeiro de Carvalho Gomes^{[1],[2]} , Taís Oliveira de Araújo^{[1],[2]} ,
and Rodrigo Gurgel-Gonçalves^[1] 

[1]. Universidade de Brasília, Faculdade de Medicina, Laboratório de Parasitologia Médica e Biologia de Vetores, Brasília, DF, Brasil.

[2]. Fundação Oswaldo Cruz, Núcleo de Epidemiologia e Vigilância em Saúde, Programa de Residência Multiprofissional em Vigilância em Saúde, Brasília, DF, Brasil.

ABSTRACT

Digital technologies and artificial intelligence (AI) have become integral in many fields, including medicine. Neglected tropical diseases transmitted by vectors, such as arboviral diseases, spotted fever, Chagas disease, and leishmaniasis, pose a significant impact on public health, particularly in the Americas. Strengthening surveillance and control requires the use of digital technology to identify vectors. In this study, we explored how AI can aid in identifying vectors in the Americas and strengthen disease surveillance and control efforts. We reviewed the literature on the automated identification of triatomines, mosquitoes, sand flies, and ticks, focusing on advances in the Americas over the last 10 years, and provided a critical analysis of the automated identification systems for each group. Moreover, we analyzed the development stages of each study: image acquisition, image processing, algorithm training, algorithm testing, app development, app availability, and AI-based devices for vector identification and surveillance. Most studies have applied AI to identify mosquito species. The vector species databases were not diverse, and the most representative group was Triatominae, comprising 65 species (41% of all described species). Currently, approximately 30 algorithms are used for automated vector identification, with the most common being AlexNet, MobileNet, and ResNet. Most studies are in the algorithm training stage, and in the Americas, only one study has progressed to the development of applications or devices. These results highlight the potential of AI for identifying vectors in the Americas, supporting the use of automated visual identification systems as a promising approach to improve vector surveillance, while also promoting citizen science.

Keywords: Neglected tropical diseases Vectors. Public health. Automated identification.

INTRODUCTION

Digital technologies and artificial intelligence (AI) have become integral to many fields, including medicine¹. The ability of machine learning algorithms to identify images plays a critical role in the surveillance of vector-borne diseases². Neglected tropical diseases transmitted by vectors, such as arboviral diseases, spotted fever, Chagas disease, and leishmaniasis, pose a significant impact on public health, particularly in the Americas^{3,4}. Strengthening surveillance and control efforts requires the use of digital technologies to identify vectors, such as mosquitoes, ticks, kissing

bugs, and sand flies, as vector control remains a cornerstone of vector-borne disease prevention⁵.

Currently, vector identification services in the Americas rely on skilled professionals and specialized equipment, which are not always available. Strengthening vector surveillance could benefit from community participation, such as household notifications that support vector detection. In this context, automated visual identification systems offer a promising approach to improving surveillance while also fostering citizen science⁶.

Over the past decade, scientists worldwide have made significant progress in developing digital technologies for vector identification². However, there remains a notable gap in such advancements in the Americas, a region that harbors the highest diversity of vector species globally, including triatomines⁷. In this review, we explored how AI can aid in identifying vectors in the Americas and strengthen disease surveillance and control efforts. We reviewed the literature on the automated identification of triatomines, mosquitoes, sand flies, and ticks, focusing on advances in the Americas over the last 10 years, and provided a critical analysis of the automated identification systems for each group. Furthermore, we presented an overview of the current trends and future directions in this field.

 Dr. Rodrigo Gurgel Gonçalves. **E-mail:** gurgelrg@hotmail.com

Data Availability Statement: Data-in-article: Research data is available in the body of the document (Table 1).

Financial Support: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior and Conselho Nacional de Desenvolvimento Científico e Tecnológico.

Conflict of Interest: The authors declare no conflicts of interest.

Editor-in-Chief: Prof. Dalmo Correia Filho 

Received 22 July 2025

Accepted 06 September 2025

We searched for literature containing the terms 'automated identification,' 'mosquitoes,' 'ticks,' 'triatomines,' or 'sand flies.' All relevant article types, such as original articles, reviews, commentaries, opinion pieces (from PubMed), grey literature, reports, and digital media—in English, Portuguese, and Spanish, were included. Additional materials identified through manual searches were also incorporated. To organize the findings, we compiled a table summarizing key information extracted from each study, including the taxonomic group, number of species analyzed, algorithms used, training hyperparameters, and accuracy. We also assessed the development stage of each study: image acquisition, image processing, algorithm training, algorithm testing, app development, app availability, and AI-based devices for vector identification and surveillance. The review is structured by taxonomic group, according to the amount of available scientific evidence.

CULICIDAE (MOSQUITOES)

Mosquito-borne diseases represent a major public health concern, with more than half of the world's population exposed to arboviruses such as dengue⁸. Recently, the application of AI in mosquito identification has advanced considerably, particularly through convolutional neural networks (CNNs). A PubMed search using the keywords 'automated identification' and 'mosquito' yielded at least 38 articles, nine of which originated from the Americas (**Table 1**). Most studies on automated mosquito identification have been conducted in Asia.

In Korea, Kim et al.⁹ presented an automated system that detects mosquitoes using either a single image classifier (52% accuracy) or multiple deep-learning networks (84% accuracy). Okayasu et al.¹⁰ compared machine learning and CNN-based approaches using images of *Aedes*, *Culex*, and *Anopheles* in Japan. Models such as AlexNet, VGGNet, and ResNet achieved accuracy rates up to 95.5%, outperforming traditional data mining methods (82.4%). Similarly, Park et al.¹¹ in South Korea tested VGG-16, ResNet-50, and SqueezeNet CNNs, achieving identification accuracies of 97%, 96%, and 90%, respectively, for eight species. Their findings demonstrated that fine-tuning and data augmentation substantially enhanced network performance. Heat-activation maps further showed that the networks focused on morphological features comparable to those employed by taxonomists, including leg structure and coloration patterns.

In the Americas, Sanchez-Ortiz et al.¹² conducted a pioneering study on automated mosquito identification. They developed a larval mosquito image database and applied CNNs for classification. Arista-Jalife et al.¹³ effectively recognize larval samples with an accuracy of 94% and automatically cropped the larval region of interest to classify them as *Aedes*-positive or -negative without human intervention. Motta et al.¹⁴ used the LeNet, AlexNet, and GoogleNet networks to identify adult *Aedes aegypti*, *Aedes albopictus*, and *Culex quinquefasciatus* in the field, achieving 76.2% accuracy, which later improved to 93.5% after hyperparameter adjustments¹⁵.

TABLE 1: Highest accuracies obtained by studies on automated identification of disease vector images considering taxonomic groups, number of taxa, algorithms, and training hyperparameters.

Group	Taxonomic level	Number of taxa	Algorithms	Highest accuracy (%)	Training hyperparameters	Author
Culicidae	Species	2	AlexNet	97	Learning rate = 0.001; epochs = 200; training: scratch with small dataset	Sanchez-Ortiz et al. ¹²
Culicidae	Species	3	LeNet AlexNet GoogLeNet	76	Variable learning rate and epochs; optimizer: NAG, Adam, AdaGrad, SGD.	Motta et al. ¹⁴
Culicidae	Species Genera	2	Perceptron	95	Learning rate = 0.5; training method: delta rule	Carrillo et al. ¹⁶
Culicidae	Species	3	ResNet-50 VGG-16 XceptionV3 DenseNet-201	97	Variable learning rates and epochs; batch size = 32–256; optimizer: RMSprop, Adam, AdaGrad, SGD.	Motta et al. ¹⁵
Culicidae	Species	67	Xception	97	Learning rate = 0.01; epochs = 60; optimizer = Ranger (RAdam + LookAhead); Batch size = 16.	Goodwin et al. ¹⁹
Culicidae	Species	2	Dual-Source DNN Shallow-NN Slim-DNN Robust-DNN	94	Learning rate = 0.00001; epochs = 1000; optimizer = RMSProp.	Arista-Jalife et al. ¹³
Culicidae	Species	53	Xception	94	Variable learning rate. Monte Carlo simulation used for unknown species detection	Gupta et al. ²⁰
Culicidae	Species Genera	5	VectorBrain (EfficientNet-26)	94	Learning rate = 0.005; epochs = 100; optimizer: AdamW.	Li et al. ¹⁸
Culicidae	Species	4	AlexNet	94	Learning rate = 0.001; epochs = 50	Araújo et al. ¹⁷

Continue...

TABLE 1: Continuation.

Group	Taxonomic level	Number of taxa	Algorithms	Highest accuracy (%)	Training hyperparameters	Author
Ixodidae	Species	3	CNN "TickIDNet"	88	Learning rate = 0.004; mini-batch size = 32; optimizer = SGD.	Justen et al. ³⁰
Ixodidae	Species	4	ResNet-50 Custom CNN	80	Learning rate = 0.01; epochs = 50; batch size = 64; optimizer = Adam;	Omodior et al. ³¹
Ixodidae	Species	31	LeNet (TickPhone)	96	Epochs = 60; optimizer = support vector machine (SVM);	Xu et al. ³²
Ixodidae	Species	7	CNNs (Inception-ResNet)	92	Learning rate = 0.001; batch size = 64; optimizer = Adam.	Akbarian et al. ³³
Ixodidae	Species	3	VGG-16 ResNet-50 InceptionV3 DenseNet-121 MobileNetV2	99	Learning rate = 0.00001; batch size = 32; Adam10-fold cross-validation.	Luo et al. ³⁴
Ixodidae	Species	6	AlexNet ResNet-50 MobileNetV2	94	Learning rate = 0.001; number of epochs = 50; optimizer:Adam.	Gomes et al. ³⁵
Triatominae	Species	51	Statistical Classifiers	88	Classification strategy = pairwise binary classifiers;	Gurgel-Gonçalves et al. ⁴⁰
Triatominae	Species	51	InceptionV3	83	Training steps = 2000-12000	Khalighifar et al. ⁴¹
Triatominae	Species lineages	3	Multilayer Perceptron Neural Networks	91	Learning rate = 0.1; Number of harmonics = 25; number of PCA components = 30;	Cruz et al. ⁴²
Triatominae	Subfamily	51	VGG-16	99	Learning rate = 0.0001; epochs = 20;	Abdelghani et al. ⁴³
Triatominae	Species	51	VGG-16 7-layer CNN	97	Learning rate = 0.0003; weight decay = 0.002; epochs = 350; batch size = 64-128	Parsons & Banitaan ⁴⁴
Triatominae	Subfamily	2	ResNet-34	96	Learning rate = lr_find(); training steps = 50; training strategy = discriminative layer training;	Cochero et al. ⁴⁵
Triatominae	Genera	3	AlexNet AdaBoost, Gradient Boosting Histogram-based Gradient Boosting Linear Discriminant Model	100	Learning rate = 0.15-0.001; epochs = 15	Miranda et al. ⁷
Triatominae	Species	65	AlexNet	93	Learning rate = 0.001; epochs = 50	Miranda et al. ⁴⁶
Triatominae	Order Subfamily	3 2	AlexNet MobileNetV2 ResNet-50	96	Learning rate = 0.001; epochs = 100	Miranda & Gurgel-Gonçalves ⁴⁷
Phlebotominae	Species Genera Subgenera Family	12 3 6 1	MobileNet ResNet-20 YOLOv2 DarkNet	96	Batch normalization = enabled; SURF descriptors + Bag of Features (BoF) + SVM; dictionary size (BoF) = 4000 codewords; SVM kernel = polynomial	Cannet et al. ⁵¹

Most studies aimed at automating the identification of mosquitoes have focused on distinguishing species belonging to the genera *Aedes*, *Culex*, and *Anopheles*; however, some studies have explored other mosquito vectors. Carrillo et al.¹⁶ achieved 95% accuracy in distinguishing *Limatus durhamii* from *Wyeomyia* spp. based on wing characteristics using a perceptron network. Araujo et al.¹⁷ applied AlexNet network to identify four wild mosquito species belonging to the genera *Sabethes*, *Haemagogus*, and *Aedes*, which are vectors of yellow fever, achieving >90% accuracy across different body regions. Li et al.¹⁸ developed VectorBrain, a lightweight deep-learning model capable of identifying *Anopheles* species, sex, and abdominal status from field sites in Africa and North America with high accuracy (Table 1). Goodwin et al.¹⁹ developed a CNN that achieved accuracies between 86% and 97% across 2,696 images of 67 species. Based on these models, Gupta et al.²⁰ developed an IDX device, integrating high-resolution optics with computer vision algorithms, achieving 80–93% accuracy for 53 adult mosquito species. When tested on data from the United States and Papua New Guinea, the accuracies ranged from 55.3% to 93.6%.

Future studies in the Americas could leverage technologies already developed in the Old World, including algorithm optimization and robotics for vector identification. For instance, Jeyakodi et al.²¹ developed the mAedesID, an Android-based application using a customized CNN to classify *Ae. aegypti* and *Ae. albopictus*. The model was supported by techniques that reduced redundancies in the images and achieved an accuracy of 84.87%. Other studies have explored lighter models to avoid the need for large amounts of computational power. For example, Ong et al.²² used MobileNet and achieved an accuracy of 98.06% by using a model applicable to smartphones. Montalbo et al.²³ used EfficientNet B0, which has low memory consumption (437 KB) and achieved 99% accuracy. Maruf²⁴ proposed TransembleNet, an ensemble approach based on transfer learning, to classify mosquito vector species. This model combined InceptionV3, VGG-16, and ResNet-50, and outperformed previous methods with high accuracy and robust metrics.

CNNs and digital sensors have transformed mosquito identification and surveillance systems. Smart traps integrated with computer vision and neural networks can accurately identify species, such as *Ae. aegypti* and *Cx. quinquefasciatus*²⁵. Other automated computer vision-based devices can capture and identify images of mosquitoes. A high-resolution optical system integrated with computer vision algorithms provides an accurate identification of adult mosquito species. The system acquires two distinctly oriented images of each specimen and immediately aggregates the data into a template report format²⁶. They demonstrated the importance of refining algorithms and expanding datasets to cover more species and regions, which could enhance identification systems. Finally, AI-enabled mosquito surveillance using CNNs and robots designed to capture mosquitoes using multistage mosquito attractants is proposed. Semwal et al.²⁷ developed the robot 'Dragonfly,' which captures mosquitoes and uses the YOLOV4 algorithm trained to identify *Ae. aegypti*, *Ae. albopictus*, and *Culex*. The 'Dragonfly' detected and identified mosquitoes with 82% confidence level in a real-time field trial. Moreover, the robot generated a map showing the distribution of mosquitoes. Overall, these advances highlight the potential of AI to automate mosquito identification, facilitating rapid interventions and real-time surveillance.

IXODIDAE (TICKS)

Ticks are blood-sucking ectoparasites found in mammals, birds, reptiles and amphibians²⁸. Identifying ticks is crucial for controlling diseases, such as spotted fever, Lyme disease, granulocytic anaplasmosis and babesiosis²⁹. However, the traditional tick identification methods are time-consuming. With the advancement of CNNs, researchers have explored new approaches for automating tick identification with greater speed and efficiency. Our search yielded six articles from the Americas with accuracies ranging from 80–99% (Table 1).

Most studies on automated identification of ticks have been conducted in the United States. Justen et al.³⁰ used the CNN TickIDNet model to identify three tick species, *Amblyomma americanum*, *Dermacentor variabilis* and *Ixodes scapularis*, by considering their stages of development (adult, nymph, or larva), sex, and feeding status. More than 12,000 images were analyzed and enlarged to 90,000 using data augmentation techniques. Images of ticks were captured from different backgrounds and angles. The overall accuracy was 87.8%, surpassing the identification accuracy of members of the public and health professionals, but falling short of that of entomology experts. This study highlights that the image quality, particularly the size of the tick in the photograph, significantly influences the accuracy of the model in identifying arthropods.

Omodior et al.³¹ compared the performance of the ResNet-50 model with a custom CNN to identify four tick species: *I. scapularis*, *D. variabilis*, *A. americanum*, and *Haemaphysalis* sp. This study used approximately 2,000 images, considering the sex and life stages of the ticks. Images were captured against different backgrounds, showing the dorsal and ventral positions. The customized CNN achieved 80% accuracy, surpassing that of the pretrained ResNet-50 model. Xu et al.³² developed the TickPhone app, a deep learning-based system for identifying ticks via smartphones using the LeNet network. The model was trained using more than 2,000 images, achieved 95.7% accuracy in identifying 31 species, including *I. scapularis*, *A. americanum*, and *D. variabilis*. This study emphasizes the importance of rapid identification for predicting the risk of tick-borne diseases.

Akbarian et al.³³ used a computer vision approach to differentiate *I. scapularis*, *A. americanum*, and *D. variabilis*, considering characteristics such as developmental stage, sex, and feeding status. This study involved capturing 12,588 images using smartphones. Additionally, 1,000 high-resolution tick images were captured using a camera mounted on a stereomicroscope to enhance the quality of the dataset. CNNs such as Inception-ResNet and knowledge-transfer techniques have been used to achieve 92% accuracy. Luo et al.³⁴ tested five CNN models (VGG16, ResNet50, InceptionV3, DenseNet121, and MobileNetV2) to identify *A. americanum*, *D. variabilis* and *I. scapularis*. They used a set of 12,000 images from the TickReport passive tick surveillance program, which received samples from individuals, veterinary clinics, and partner agencies. The images were captured from multiple angles, including the dorsal and ventral angles, to assess the impact of perspective on model accuracy. Considering characteristics, such as developmental stage, sex, and feeding status, InceptionV3 model obtained the highest accuracy (99.5%).

In South America, Gomes et al.³⁵ analyzed the ability of AlexNet, ResNet-50, and MobileNetV2 networks to identify vectors of spotted fever, such as *Amblyomma aureolatum*, *A. ovale*, and

A. sculptum. This study organized a bank of 826 images, categorized according to sex, dorsal and ventral positions, and image resolution. The CNN models achieved accuracies of 94%, 92%, and 94% with sensitivities ranging from 59% to 100%, depending on the species, sex, position, and image resolution. In Asia, Akgül and Kaya³⁶ used the YOLO-V3 network to detect ticks in a set of 1,500 images sourced from the Internet. These images were captured at various locations, such as skin and vegetation, at different resolutions. As the images were obtained from various online sources, they had no fixed positions. The model achieved an accuracy of 98.5%, demonstrating its efficacy for automatically identifying ticks in various environments. In Africa, Mudau et al.³⁷ applied CNNs to quantify the parasitic load of ticks in cattle using infrared thermal imaging. The ConvNet and MobileNet models were trained using 1,124 thermographic images of animals captured at different angles, achieving training and validation accuracies of approximately 90% and 60% and 95% and 75%, respectively. Images of adult ticks (fed and unfed females and males) belonging to *Amblyomma hebraeum*, *Hyalomma marginatum*, *Rhipicephalus appendiculatus*, *Rhipicephalus (Boophilus) spp.*, and *Rhipicephalus evertsi* were captured. This study highlights the potential of AI in monitoring tick infestations in livestock. Overall, these results indicated the high reliability of classifying important tick species using different algorithms (Table 1). In addition, some studies performed photo position tests, background variation tests, segmentation or Grad-CAM tests, and other applications. To date, no AI-driven robots have been developed to capture and identify ticks (Figure 1).

TRIATOMINAE (KISSING BUGS)

The subfamily Triatominae comprises blood-sucking insects that transmit *Trypanosoma cruzi*, the etiological agent of Chagas' disease³⁸, which remains a major public health challenge. As triatomines continue to invade and infest human dwellings in the Americas³⁹, developing technologies that can distinguish them from similar bugs is essential for improving Chagas disease vector surveillance⁶.

Our search revealed nine studies on the automated identification of triatomines in the Americas (Table 1). Gurgel-Gonçalves et al.⁴⁰ first explored this approach, consistently achieving accuracies above 80%. They processed high-quality dorsal photographs taken with a specialized apparatus into landmarks and used ratios derived from these measurements to distinguish between 12 Mexican and 39 Brazilian triatomine species. Building on this, Khalighifar et al.⁴¹ employed the same image database alongside TensorFlow, an open-source software platform for deep learning algorithms, to improve identification rates. Using InceptionV3, they achieved identification rates of 83.0% and 86.7% for the Mexican and Brazilian species, respectively. When distributional data were incorporated, the identification rates increased markedly—to 95.8% for Mexican species and 98.9% for Brazilian species. Subsequent studies further improved accuracy by applying other algorithms to the same image bank of Gurgel-Gonçalves et al.⁴⁰ and additional triatomine datasets, also yielding strong results^{7,42–46}.

In later stages, the scope of automated triatomine identification broadened in both methodology and taxonomy. More complex deep learning algorithms, diverse datasets, and images of varying quality were incorporated. CNN-based models such as AlexNet and VGG16 outperformed classical statistical approaches, achieving accuracies between 93% and 100%, even for mobile

phone images captured at different angles and resolutions^{7,43,46}. Miranda et al.⁷ demonstrated that the AlexNet model can identify three genera (*Triatoma*, *Panstrongylus* or *Rhodnius*) with 100% accuracy, regardless of image orientation, while Miranda et al.⁴⁶ reported 94.6% accuracy in identifying 48 vector species of medical importance. Similarly, Cocherio et al.⁴⁵ achieved 96% accuracy in classifying triatomines, using a CNN trained on mobile phone images, underscoring the feasibility of citizen-science-based systems. However, variations in performance across studies reflected differences in algorithms, image quality, number of taxonomic classes, and database curation.

Recent advances in the identification of triatomines have prompted Miranda and Gurgel-Gonçalves⁴⁷ to propose a new challenge for automated identification systems: the need to distinguish triatomines from morphologically similar insects, such as phytophagous and predatory bugs, which are frequently misidentified as triatomines. To address this, they evaluated three CNNs (AlexNet, MobileNetV2, and ResNet-50) trained via transfer learning on a database of 707 dorsal bug images categorized by feeding habit. All models achieved over 94% accuracy, with ResNet-50 reaching 96.5% accuracy and perfect sensitivity and specificity for certain categories. Grad-CAM visualizations indicated that the models relied on different morphological features—such as the abdomen (connexivum) and hemelytra for blood-sucking bugs, and the head and pronotum for phytophagous and predatory bugs. These findings demonstrate that CNNs can accurately discriminate triatomines from other Heteroptera species even without lateral visualization of the mouthparts traditionally used by experts, marking a significant advance in automated entomological surveillance for both citizen science and fieldwork applications⁴⁷.

To further advance automated triatomine identification, expanding the representativeness of image databases remains a crucial step. Miranda et al.⁴⁶ showed that training AlexNet algorithm on a larger and more diverse dataset of 6,397 images from 65 species across seven genera and 27 countries markedly improved performance. The inclusion of images with varying resolutions, backgrounds, and lighting conditions enhanced model robustness against both morphological and photographic variability. The model achieved an overall accuracy of 92.8% across all species, rising to 94.6% when limited to the 48 most epidemiologically relevant species. The authors emphasized that incorporating species from diverse geographical regions and vectorial capacities is key to developing identification systems suitable for broad application. They also suggested that continuously expanding the database is fundamental to creating open-source tools that can be accessed via mobile devices and provide accurate real-time identification, thereby strengthening entomological surveillance programs throughout Latin America.

PSYCHODIDAE (SAND FLIES)

Identifying phlebotomine sand flies is essential because they are vectors of several pathogens, such as *Leishmania* spp.⁴⁸. Female sand flies transmit *Leishmania* species in a species-specific manner. Manual identification of internal structures is difficult; however, few studies have attempted to automate this process. We found four studies from the Old World. Fraiwan⁴⁹ developed a deep learning AI system to identify three species (*Phlebotomus alexandri*, *Phlebotomus papatasi*, and *Phlebotomus sergenti*) in northern and central Jordan. Fraiwan et al.⁵⁰ used images of the genitalia and pharynxes of sand flies captured in the field over a period of

two years in Jordan and CNNs to identify them, achieving an accuracy of approximately 95%. Cannet et al.⁵¹ identified sand flies with an accuracy of over 77%, reaching 96% for American sand flies (**Table 1**). A new dataset was introduced to train and evaluate the recognition system for dipterans of medical and veterinary importance (Calliphoridae, Muscidae, Tabanidae, Ceratopogonidae, and Psychodidae). This system can identify dipterans based on wing morphology, thereby eliminating the need for internal organ inspection. This system can be implemented under both field and near-field conditions⁵².

Deep learning has shown promising results for the automated identification of cutaneous leishmaniasis lesion images⁵³. Therefore, it is important to expand the leishmaniasis vector dataset to include neotropical phlebotomines and advance the development of application for sandfly identification based on image analysis. This application could help researchers and health professionals conduct surveillance of leishmaniasis vectors.

CONCLUSIONS AND OUTLOOK

AI has revolutionized the identification and surveillance of disease vectors. Most studies have focused on applying AI to identify mosquito species, although the majority included fewer than ten species, except for two that analyzed 53 and 67 species. The vector species databases were generally limited in diversity, with the most representative group being the Triatominae, which included 65 species (41% of all described species). Approximately 30 algorithms have been used for automated vector identification, the most common being AlexNet, MobileNet, and ResNet. Moreover, these studies employed a wide range of hyperparameters.

Using CNNs for insect identification requires adaptations that extend beyond the original network architecture and typically

involves three main stages. First, a suitable pretrained model must be selected (e.g., ResNet, EfficientNet, or MobileNet). Next, the convolutional layers that extract general features are frozen, and new classification layers are added. Finally, fine-tuning is conducted through hyperparameter optimization. Although most reviewed studies did not provide a comprehensive description of these steps, we summarized the algorithms and hyperparameters used in each case. Image resolutions varied both between and within studies. Several investigations have explored the impact of resolution, indicating that markedly low resolutions significantly reduce accuracy. However, in real-world scenarios, AI systems for vector identification cannot rely solely on high-resolution images. Identifying kissing bugs using low-resolution mobile phone images yielded similar accuracy to that obtained with high-resolution images. Comparable findings were reported for ticks, suggesting that resolution may not be a critical factor.

Most studies reported accuracy as the primary performance metric, with limited discussion of the implications of false positives and false negatives. A false positive (where a non-vector is misclassified as a vector) could trigger unnecessary public health alerts and resource expenditures, while a false negative (where a true vector is missed) could delay control interventions and facilitate pathogen transmission. Future studies should therefore incorporate a broader range of metrics, such as precision, recall, the F1 score, and the area under the curve (AUC).

Another important limitation across studies is the lack of systematic evaluation of how dataset characteristics influence performance. Factors such as database size, image resolution, background variability, and species diversity often have a greater impact on outcomes than the algorithm itself. These findings underscore the importance of dataset composition in determining model outcomes, and overlooking this factor can

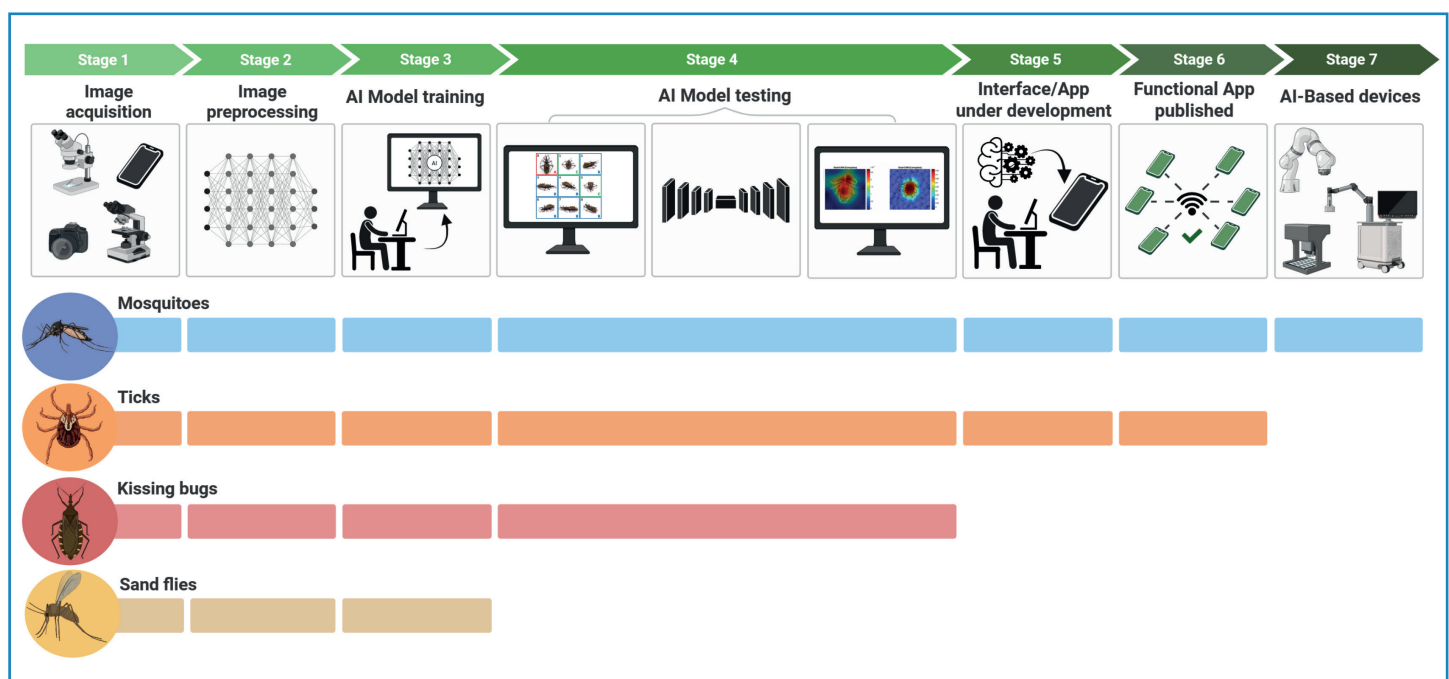


FIGURE 1: Advances in the development of automated identification systems for tropical disease vectors. The figure illustrates the seven stages of developing automated identification systems for four groups of vectors: mosquitoes, ticks, triatomines, and sand flies. The studies on mosquitoes and ticks are more advanced, with robust systems and AI-based devices already developed. Triatomine studies are close behind, with the prospect of developing applications within the next few years. Sand fly studies remain in the beginning stages.

lead to overestimation of algorithmic performance. The results also depend heavily on hyperparameter settings, including architectural parameters (e.g., number of convolutional layers, filters, dropout rates, and activation functions) and training parameters (e.g., learning rate, optimizer, batch size, number of epochs, and regularization strategies). However, most reviewed studies did not report these configurations, limiting reproducibility. Some exceptions exist, such as Motta et al.¹⁵, who optimized hyperparameters for mosquito classification, and Miranda et al.⁴⁶, who evaluated multiple CNN architectures. Future research should systematically report hyperparameter choices in standardized tables to facilitate comparison across studies.

Most investigations remain in the algorithm training phase, with only one study in the Americas advancing to the development of applications or devices. Robots equipped with computer vision and CNNs, such as YOLOV4 have been used to capture mosquito images and identify species with high accuracy (**Figure 1**). Advancing this field requires expanding datasets to cover more species and regions, thereby enhancing identification systems. Significant progress has been made over the past seven years, highlighting the potential of AI in automated mosquito identification and contributing to rapid interventions and real-time surveillance.

Although encouraging advances have been made in automating tick identification, current studies primarily focus on tick species relevant to public health in North America. However, recent research has explored the application of AI to identify ticks in other regions where these ectoparasites play a key role in transmitting infectious agents. This review reinforces the need for further studies on spotted fever vectors in South America, which would strengthen epidemiological surveillance and support the development of deep learning tools for tick identification, benefiting both researchers and public health professionals in the region.

Following the inclusion of more triatomine species, automated identification systems could be made available through open-access applications (**Figure 1**). With an Internet-connected smartphone, users could photograph a triatomine, upload the image, and receive immediate identification and vector-related information. A key advancement in this direction is GeoVin (Geovin.com.ar), a platform that gathers geographic information on Argentinean triatomines via citizen reports of bug findings. Occurrence data (photos and geographic coordinates) online, stored, and automatically integrated into the GeoVin dataset. Applications based on automated visual identification systems may substantially strengthen Chagas disease vector surveillance while promoting citizen science.

Research on the automated identification of sand flies remains in its early stages (**Figure 1**). We found only four studies applying deep learning to sandfly identification. Several challenges remain, including expanding databases, testing optimal algorithms for this group, and determining the most informative morphological regions for males and females. Preliminary findings suggest that automated sandfly identification is promising. We also propose further studies on the automated identification of other vector groups, including snails, ceratopogonids, and simuliids in the Americas. Our team is currently developing databases and planning future studies on these groups. We hope that technologies already applied to triatomines, culicids, and ticks can soon be extended to these additional vectors.

ACKNOWLEDGMENTS

We would like to thank Rebecca Kerr, Tessa Cook, Dallin Karl Mead, and José Eduardo Gonçalves Guimarães for reviewing the manuscript.

AUTHORS' CONTRIBUTION

VLM: Conceptualization, Methodology, Formal analysis, Validation, Writing - review & editing; JFCL: Methodology, Writing - original draft, Writing - review & editing; IRCG: Writing - original draft, Writing - review & editing. TOA: Writing - original draft, Writing - review & editing. RGG: Conceptualization, Data curation, Investigation, Methodology, Supervision, Writing - original draft, Writing - review & editing;

REFERENCES

- Chen X, Wang X, Zhang K, Fung KM, Thai TC, Moore K, et al. Recent advances and clinical applications of deep learning in medical image analysis. *Med Image Anal.* 2022;79:102444.
- Bai S, Shi L, Yang K. Deep learning in disease vector image identification. *Pest Manag Sci.* 2025;81(2):527–9.
- Wilson AL, Courtenay O, Kelly-Hope LA, Scott TW, Takken W, Torr SJ, et al. The importance of vector control for the control and elimination of vector-borne diseases. *PLoS Negl Trop Dis.* 2020;14(1):e0007831.
- Hotez PJ, Aksoy S, Brindley PJ, Kamhawi S. What constitutes a neglected tropical disease? *PLoS Negl Trop Dis.* 2020;14(1):e0008001.
- World Health Organization, UNICEF. Global vector control response 2017-2030. Geneva: WHO; 2017.
- Gurgel-Gonçalves R. Stronger control-surveillance systems for vector-borne Chagas disease. *Mem Inst Oswaldo Cruz.* 2022;117:e210130chgsb.
- Miranda VL, de Souza EP, Bambil D, Khalighifar A, Peterson AT, de Oliveira Nascimento FA, et al. Cellphone picture-based, genus-level automated identification of Chagas disease vectors: Effects of picture orientation on the performance of five machine-learning algorithms. *Ecol Inform.* 2024;79:102430.
- World Health Organization. Dengue and severe dengue [Internet]. Geneva: WHO; [cited 2025 Jun 25]. Available from: <https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue>.
- Kim K, Hyun J, Kim H, Lim H, Myung H. A deep learning-based automatic mosquito sensing and control system for urban mosquito habitats. *Sensors.* 2019;19:2785.
- Okayasu K, Yoshida K, Fuchida M, Nakamura A. Vision-based classification of mosquito species: Comparison of conventional and deep learning methods. *Appl Sci.* 2019;9(18):3935.
- Park J, Kim DI, Choi B, Kang W, Kwon HW. Classification and morphological analysis of vector mosquitoes using deep convolutional neural networks. *Sci Rep.* 2020;10:1012.
- Sanchez-Ortiz A, Fierro-Radilla A, Arista-Jalife A, Cedillo-Hernandez M, Nakano-Miyatake M, Robles-Camarillo D, et al. Mosquito larva classification method based on convolutional neural networks. In: 2017 International Conference on Electronics, Communications and Computers (CONIELECOMP). IEEE; 2017. p.1-6.
- Arista-Jalife A, Nakano M, Garcia-Nonoal Z, Robles-Camarillo D, Perez-Meana H, Arista-Viveros HA. *Aedes* mosquito detection in its larval stage using deep neural networks. *Knowl Based Syst.* 2020;189:104841.
- Motta D, Santos AÁB, Winkler I, Machado BAS, Pereira DADI, Cavalcanti AM, et al. Application of convolutional neural networks for classification of adult mosquitoes in the field. *PLoS One.* 2019;14(1):e0210829.

15. Motta D, Santos AÁB, Machado BAS, Ribeiro-Filho OGV, Camargo LOA, Valdenegro-Toro MA, et al. Optimization of convolutional neural network hyperparameters for automatic classification of adult mosquitoes. *PLoS One*. 2020;15(7):e0234959.
16. Carrillo D, Benítez DS, Ramón G, Pérez N. Towards automatic classification of mosquito species based on wing geometrical features. In: 2020 IEEE ANDESCON. IEEE; 2020. p.1-6.
17. Araújo TO, de Miranda VL, Gurgel-Gonçalves R. AI-driven convolutional neural networks for accurate identification of yellow fever vectors. *Parasit Vectors*. 2024;17(1):329.
18. Li D, Hegde S, Kumar AS, Zacharias A, Mehta P, Mukthinini V, et al. Towards transforming malaria vector surveillance using VectorBrain: a novel convolutional neural network for mosquito species, sex, and abdomen status identifications. *Sci Rep*. 2024;14(1):23647.
19. Goodwin A, Padmanabhan S, Hira S, Glancey M, Slinowsky M, Immidiseti R, et al. Mosquito species identification using convolutional neural networks with a multitiered ensemble model for novel species detection. *Sci Rep*. 2021;11(1):13656.
20. Gupta KA, Ikonomidou VN, Glancey M, Faiman R, Talafha S, Ford T, et al. Mosquito species identification accuracy of early deployed algorithms in IDX, a vector identification tool. *Acta Trop*. 2024;260:107392.
21. Jeyakodi G, Agarwal T, Bala PS. mAedesID: Android application for *Aedes* mosquito species identification using convolutional neural network. *arXiv [Preprint]*. 2023. Available from: <https://arxiv.org/abs/2305.07664>.
22. Ong SQ, Ahmad H, Nair G, Isawasan P, Majid AHA. Implementation of a deep learning model for automated classification of *Aedes aegypti* (Linnaeus) and *Aedes albopictus* (Skuse) in real time. *Sci Rep*. 2021;11(1):9908.
23. Montalbo FJP. Automating mosquito taxonomy by compressing and enhancing a feature fused EfficientNet with knowledge distillation and a novel residual skip block. *MethodsX*. 2023;10:102072.
24. Maruf A, Mahmudul Haque M, Ara Romy R, Jahan Puspo J, Aung Z. TransembleNet: Enhancing vector mosquito species classification through transfer learning-based ensemble model. *PLoS One*. 2025;20(5):e0322171.
25. Rajak P, Ganguly A, Adhikary S, Bhattacharya S. Smart technology for mosquito control: recent developments, challenges, and future prospects. *Acta Trop*. 2024;258:107348.
26. Brey J, Sai Sudhakar BMM, Gersch K, Ford T, Glancey M, West J, et al. Modified mosquito programs' surveillance needs and an image-based identification tool to address them. *Front Trop Dis*. 2022;2:810062.
27. Semwal A, Melvin LMJ, Mohan RE, Ramalingam B, Pathmakumar T. AI-enabled mosquito surveillance and population mapping using Dragonfly robot. *Sensors (Basel)*. 2022;22(13):4921.
28. Guglielmone AA, Nava S, Robbins RG. Geographic distribution of the hard ticks (Acari: Ixodida: Ixodidae) of the world by countries and territories. *Zootaxa*. 2023;5251(1):1–274.
29. Dantas-Torres F, Chomel BB, Otranto D. Ticks and tick-borne diseases: a One Health perspective. *Trends Parasitol*. 2012;28(10):437–46.
30. JJusten L, Carls Smith D, Paskewitz SM, Bartholomay LC, Bron GM. Identification of public submitted tick images: a neural network approach. *PLoS One*. 2021;16(12):e0260622.
31. Omodior O, Saeedpour-Parizi MR, Rahman MK, Azad A, Clay K. Using convolutional neural networks for tick image recognition—a preliminary exploration. *Exp Appl Acarol*. 2021;84:607–22.
32. Xu Z, Ding X, Yin K, Li Z, Smyth JA, Sims MB, et al. Tickphone App: A smartphone application for rapid tick identification using deep learning. *Appl Sci*. 2021;11(16):7355.
33. Akbarian S, Nelder MP, Russell CB, Cawston T, Moreno L, Patel SN, et al. A computer vision approach to identifying ticks related to Lyme disease. *IEEE J Transl Eng Health Med*. 2021;10:1–8.
34. Luo CY, Pearson P, Xu G, Rich SM. A computer vision-based approach for tick identification using deep learning models. *Insects*. 2022;13(2):116.
35. Gomes IRC, Miranda VL, Leal JFC, Oliveira IP, Silva PJ, Bitencourth K, et al. Automated identification of spotted-fever tick vectors using convolutional neural networks. *Med Vet Entomol*. 2025.
36. Akgül İ, Kaya V. Realization of a YOLO-V3-based application for the detection of tick-borne cases. *Eng Sci*. 2022;17(3):35–41.
37. Mudau F, Van Zyl TL, Molotsi AH, Waldmann P, Dzama K, Marufu MC. Application of convolutional neural networks to the quantification of tick burdens on cattle using infrared thermographic imaging. In: 2022 IST-Africa Conference (IST-Africa). IEEE; 2022. p. 1–9.
38. Lent H, Wygodzinsky P. Revision of the Triatominae (Hemiptera, Reduviidae), and their significance as vectors of Chagas disease. *Bull Am Mus Nat Hist*. 1979;163(3):123–520.
39. Cucunubá UM, Gutiérrez-Romero SA, Ramírez JD, Velásquez-Ortiz N, Ceccarelli S, Parra-Henao G, et al. The epidemiology of Chagas disease in the Americas. *Lancet Reg Health Am*. 2024;37:100881.
40. Gurgel-Gonçalves R, Komp E, Campbell LP, Khalighifar A, Mellenbruch J, Mendonça VJ, et al. Automated identification of insect vectors of Chagas disease in Brazil and Mexico: the Virtual Vector Lab. *PeerJ*. 2017;5:e3040.
41. Khalighifar A, Komp E, Ramsey JM, Gurgel-Gonçalves R, Peterson AT. Deep learning algorithms improve automated identification of Chagas disease vectors. *J Med Entomol*. 2019;56(5):1404–10.
42. Cruz DD, Arellano E, Ávila DD, Ibarra-Cerdeña CN. Identifying Chagas disease vectors using elliptic Fourier descriptors of body contour: a case for the cryptic *dimidiata* complex. *Parasites Vectors*. 2020;13:332.
43. Abdelghani BA, Banitaan S, Maleki M, Mazen A. Kissing bugs identification using convolutional neural network. *IEEE Access*. 2021;9:140539–48.
44. Parsons Z, Banitaan S. Automatic identification of Chagas disease vectors using data mining and deep learning techniques. *Ecol Inform*. 2021;62:101270.
45. Cocherio J, Pattori L, Balsalobre A, Ceccarelli S, Marti G. A convolutional neural network to recognize Chagas disease vectors using mobile phone images. *Ecol Inform*. 2022;68:101587.
46. Miranda VL, Oliveira-Correia JP, Galvao C, Obara MT, Peterson AT, Gurgel-Gonçalves R. Automated identification of Chagas disease vectors using AlexNet pre-trained convolutional neural networks. *Med Vet Entomol*. 2024;39(2):291–300.
47. Miranda VL, Gurgel-Gonçalves R. Phytophagous, blood-suckers or predators? Automated identification of Chagas disease vectors and similar bugs using convolutional neural network algorithms. *Acta Trop*. 2025;265:107621.
48. Galati EAB, de Andrade AJ, Perveen F, Loyer M, Vongphayloth K, Randrianambinintsoa FJ, et al. Phlebotomine sand flies (Diptera, Psychodidae) of the world. *Parasites Vectors*. 2025;18(1):220.
49. Fraiwan M. Efficient convolutional neural network model for the taxonomy and sex identification of three phlebotomine sandfly species (Diptera, Psychodidae, and Phlebotominae). *Animals (Basel)*. 2024;14(24):3712.
50. Fraiwan M, Mukbel R, Kanaan D. Using deep learning artificial intelligence for sex identification and taxonomy of sand fly species. *PLoS One*. 2025;20(4): e0320224.

51. Cannet A, Simon-Chane C, Histace A, Akhoundi M, Romain O, Souchaud M, et al. Species identification of phlebotomine sandflies using deep learning and wing interferential pattern (WIP). *Sci Rep.* 2023;13(1):21389.
52. Cannet A, Simon-Chane C, Histace A, Akhoundi M, Romain O, Souchaud M, et al. An annotated wing interferential pattern dataset of dipteran insects of medical interest for deep learning. *Sci Data.* 2024;11(1):4.
53. Leal JFC, Barroso DH, Trindade NS, Miranda VL, Gurgel-Gonçalves R. Automated identification of cutaneous leishmaniasis lesions using deep-learning-based artificial intelligence. *Biomedicine.* 2023;12(1):12.