Morphological classification of hematophagous Diptera with Convolutional Neural Networks: A mapping of literature

ARTICLE HISTORY
Received 14 October 2025
Accepted 23 February 2026
Published 7 July 2026

Benjamín Paulino Mendoza Contreras
Veracruzana University
Faculty of Statistics and Informatics
Xalapa, Veracruz
benjaminpaulinom6@gmail.com
ORCID: 0009-0000-3491-6234

Emmanuel Morales García
Veracruzana University
Faculty of Statistics and Informatics
Xalapa, Veracruz
emmorales@uv.mx
ORCID: 0000-0002-6837-9227

Cecilia Cruz López
Veracruzana University
Faculty of Statistics and Informatics
Xalapa, Veracruz
ceccruz@uv.mx
ORCID: 0000-0002-9156-5669

Luis Enrique Gomez Medina
Veracruzana University
Institute for Research and Higher Studies in Administrative Sciences
Xalapa, Veracruz
luisgomez04@uv.mx
ORCID: 0009-0009-1324-389X

B. Mendoza, E. Morales, C. Cruz, and L. Gomez, "Morphological classification of hematophagous Diptera with Convolutional Neural Networks: A mapping of literature," Latin-American Journal of Computing (LAJC), vol. 13, no. 2, 2026.

ISSN:1390-9266 e-ISSN:1390-9134 LAJC 2026

Morphological classification of hematophagous Diptera with Convolutional Neural Networks: A mapping of literature

Benjamín Paulino Mendoza Contreras
Veracruzana University
Faculty of Statistics and Informatics
Xalapa, Veracruz
benjaminpaulinom6@gmail.com

Emmanuel Morales García
Veracruzana University
Faculty of Statistics and Informatics
Xalapa, Veracruz
emmorales@uv.mx

Cecilia Cruz López
Veracruzana University
Faculty of Statistics and Informatics
Xalapa, Veracruz
ceccruz@uv.mx

Luis Enrique Gomez Medina
Veracruzana University
Institute for Research and Higher Studies in Administrative Sciences
Xalapa, Veracruz
luisgomez04@uv.mx

Abstract**—** This review analyzes studies that primarily address the morphological classification of hematophagous Diptera, with limited mention of other insects. These networks have become increasingly important in morphological analysis through the accurate and efficient automatic identification of species, surpassing even traditional methods based on human observation. The main architectures used, such as VGG-16, YOLOv5, Faster R-CNN, Mask R-CNN, ResNet, and Swin Transformer-L are reviewed, highlighting their applications in the detection and identification of different anatomical parts. Common limitations are also mentioned, such as the need for large volumes of classified data and variability in image quality. Finally, current trends have been identified that point to the development of more robust hybrid models capable of recognizing new species and improving accuracy under real-world conditions. This literature mapping provides greater certainty and evidence regarding the most important identification methods in the field of entomology. These findings highlight the gap in literature related to the availability of public data, parameters used, data volume, image quality, and model evaluation, providing a solid foundation to guide future research in the field of entomology.

Keywords**—***Entomology, Organism Classification, Deep Learning, Species, Identification, Morphology*

I. Introduction

Image-based classification of hematophagous Diptera is fundamental and has evolved thanks to the various Convolutional Neural Networks (CNNs) used for image recognition. These networks allow for the analysis of visual characteristics (morphology) of species at a higher level and more quickly, even under varying image conditions [1].

Classifying hematophagous Diptera from digital images allows for faster, more accurate, and scalable identification than traditional methods based on human observation. These networks are capable of automatically extracting complex visual features (shapes, textures, or patterns) that determine the differences between species, without requiring the researcher to manually define the important characteristics [1].

One of the main advantages of CNNs is their ability to process large volumes of data from different image capture devices. This represents a significant change in field data collection, as it automates repetitive tasks and reduces the human workload, allowing experts and researchers to achieve greater accuracy.

Some authors have developed hierarchical architectures that incorporate taxonomic relationships between genera and species within the model itself, improving performance and reducing errors when classifying at more specific taxonomic levels [2],[3]. This type of innovation is important for classifying morphologically similar organisms, where visual differences may be minimal.

Convolutional neural networks (CNNs) are one of the most influential innovations in the field of deep learning, due to their ability to automatically and efficiently process, analyze, and classify visual data. Their main advantage lies in their ability to extract hierarchical features directly from images, reducing the need for human intervention in selecting relevant features [4].

In the scientific and technological fields, CNNs have demonstrated exceptional performance in tasks such as facial recognition, medical imaging, organism classification, object detection, and autonomous driving. Their structure, based on convolutional, clustering, and fully connected layers, allows for the identification of complex patterns [5],[6].

The objective of this literature mapping is to show the trends, methodological approaches, data types, and CNN architecture used in the morphological classification of Diptera reported in recent research. Although morphological classification includes various biological groups, recent studies show a significant concentration on hematophagous Diptera, due to their relevance to public health issues. Therefore, this research focuses on this group.

LATIN-AMERICAN JOURNAL OF COMPUTING (LAJC), Vol XIII, Issue 2, July 2026

II. METHODOLOGY

This study employs an exploratory and descriptive literature review approach, aiming to identify trends, methodologies, convolutional neural network (CNN) architectures, data types used, and gaps in the literature regarding morphological classification of dipterans using images. This type of review seeks to provide a structured overview of the current state of the art. Fig. 1 shows the general phases of the mapping review process followed in this study.

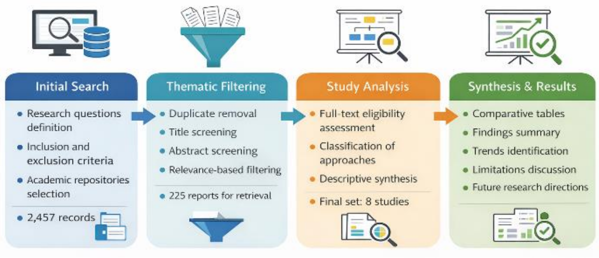

Fig. 1. Article selection process.

A. Research Questions

RQ1. What pre-trained convolutional neural network architectures have been used to detect morphology in Diptera?

RQ2. What CNN-based approaches analyze Diptera morphology?

RQ3. What types of results are reported in studies on wing morphology?

RQ4. What computational approaches have been proposed for detecting morphological patterns in Diptera?

B. Inclusion and exclusion criteria

TABLE I. SELECTION CRITERIA FOR PRIMARY STUDIES

Category	Inclusion	Exclusion
Type of research	Practical research on image classification methods using convolutional neural networks	Non-primary studies: literature review.
Publication year interval	Articles published from 2020 to 2025 to ensure current relevance	Studies published before 2020.
Language	Articles in English	Articles in Spanish or another language.
Search engines	Primary Search Engines (Publishers)	Search engines of dubious scientific quality
Applications	Emphasis on medical imaging and other areas of study	
Thematic relevance	Research on the classification of diptera	Investigations that divert attention from the main topic

C. Search Strategy

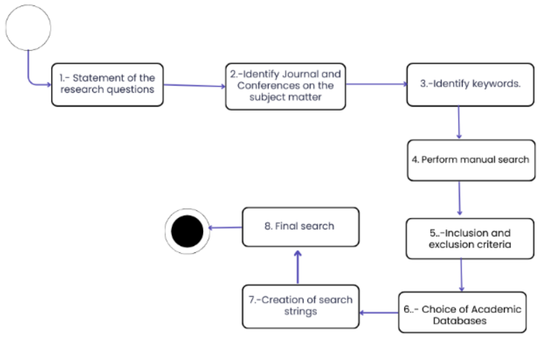

Fig. 2. Distribution of selected articles by digital library.

For this study, a search strategy was employed that was designed to achieve broad and specific coverage of the relevant literature, following a descriptive process (Fig. 2). The following elements were considered:

1. Repositories used: ACM Digital Library, IEEE Xplore, SpringerLink, and ScienceDirect.
2. Year range: Publications between 2020 and 2025, with the aim of including recent work.
3. Language: Only publications in English.
4. Keywords: Terms related to CNNs, image classification, Diptera, and public health.

To address the variability of terms present in the literature, a set of keywords and synonyms were proposed, which were combined using Boolean operators. The general search string used was:

"convolutional neural network" OR CNN OR "deep learning" OR "computer vision") AND ("image classification" OR "image recognition" OR "object detection") AND (morphology OR "morphological identification" OR "morphometric analysis") AND (Diptera OR mosquito OR mosquitoes OR Culicidae OR "hematophagous insects") AND (wing OR winVuelgs OR "wing morphology" OR "body morphology"

This string is adapted to the specific syntax of each repository used.

The annual trend in publications in the selected subsample was analyzed, as shown in Fig. 3. The frequency analysis by year for the period 2020-2025 shows a growing trend in the publication of articles that use Convolutional Neural Networks (CNN) in the analyzed repositories.

LATIN-AMERICAN JOURNAL OF COMPUTING (LAJC), Vol XIII, Issue 2, July 2026

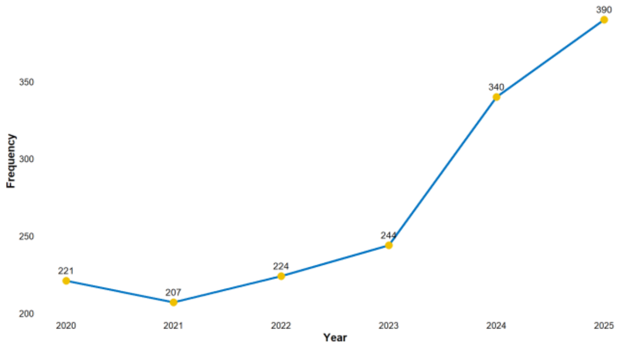

Fig. 3. Annual trend in publication frequency (2020-2025) of the subsample of articles selected from the four repositories

In addition to the trend analysis, the authors with the highest number of scientific publications were identified using CNNs to classify organisms in the four academic repositories shown in Fig. 4. The word cloud in Fig. 4 reveals a clear concentration of publications by Asian authors, particularly those with the surnames Zhang, Liu, Li, and Wang.

Fig. 4. Word cloud of the most frequent authors in the publication of articles in the four repositories (2020-2025)

D. *Study Selection Process*

The search conducted across the selected academic repositories yielded a total of 2,457 records (Table II). To ensure the relevance of the studies included in this mapping review, a multi-stage filtering process was applied following the PRISMA 2020 guidelines (Fig. 5). First, duplicate records were removed, resulting in the exclusion of 412 articles, leaving 2,045 unique records for the screening phase.

Second, a thematic filtering stage was performed based on the analysis of titles, abstracts, and keywords, which led to the exclusion of 1,820 records that were not aligned with the objectives of this study. As a result, 225 articles were considered potentially relevant and were retrieved for full-text evaluation.

Third, during the eligibility assessment, the full texts of the 225 remaining studies were examined in detail. At this stage, 217 studies were excluded because they did not meet the inclusion criteria. The main reasons for exclusion included studies not focused on morphological classification using images, studies that did not employ deep learning methods, articles lacking experimental evaluation or performance metrics, studies based on non-image data, and research outside the scope of insect identification. After applying these criteria, a final set of eight studies was selected for qualitative analysis. These studies were analyzed to identify the architecture, datasets, and methodological trends used in automated morphological classification of insects using artificial intelligence.

TABLE II. NUMBER OF ARTICLES RECOVERED FROM SELECTED ACADEMIC REPOSITORIES (2020–2025)

Repository	Total, number of items
ACM	165
IEEE Xplore	116
SpringerLink	397
ScienceDirect	1779

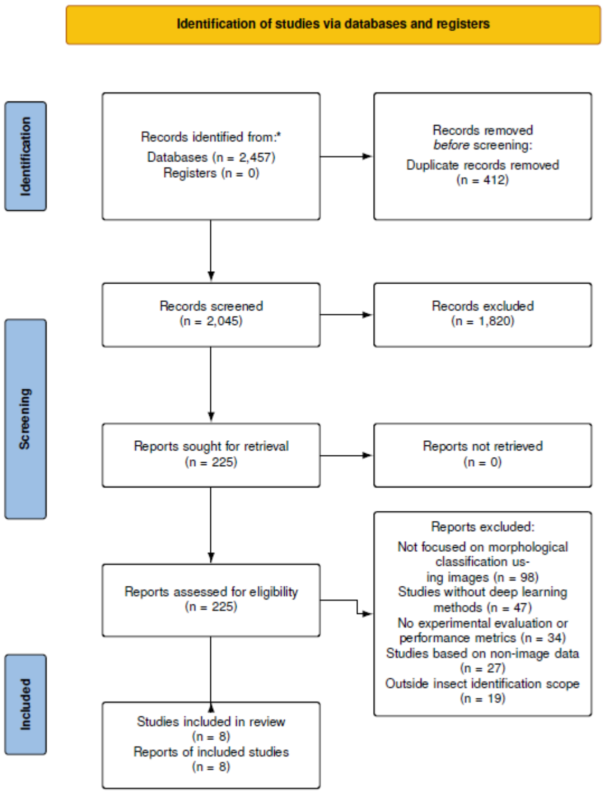

Fig. 5. Study filtering process

E. *Descriptive analysis of the studies*

Table III below shows information about the selected studies, allowing you to view key information about each one. It should be noted that no quantitative comparison is made.

LATIN-AMERICAN JOURNAL OF COMPUTING (LAJC), Vol XIII, Issue 2, July 2026

TABLE III. ARTICLES DATA MATRIX

Author(s)	Year	Organism	Analyzed part	Dataset	Task type	Metrics
Adhane, G., Dehshibi, M. M., & Masip, D. [1]	2021	Mosquitoes (A. albopictus)	Body (legs, abdomen)	Public	Classification	94.61% Accuracy
Minakshi, M. et al. [21]	2020	Mosquitoes (9 species)	Thorax, wings, abdomen	Private	Detection/Feature extraction	95% Accuracy, 60% mAP
Sauer, F. G. et al. [22]	2024	Mosquitoes (Aedes)	Wings	Private	Classification	99% F1-score
Nolte, K. et al. [23]	2025	Mosquitoes (4 species)	Wings vs body	Private	Classification	87.6% (Wings) / 78.9% (Body)
Cannet, A. et al. [25]	2023	Mosquitoes (Aedes genus)	Interference patterns (wings)	Private	Classification	95% Accuracy
Zhao, D. et al. [28]	2022	Mosquitoes (17 species)	Whole body	Private	Classification	99.04% Accuracy
Lee, S., Kim, H., & Cho, B.-K. [29]	2023	Mosquitoes (11 species)	Whole body	Mixed	Detection	97.1% F1-score
Goodwin, A. et al. [30]	2021	Wildlife species	Body	Private	Multilevel	97.04% Accuracy (Known classes)

III. CONCEPTUAL FOUNDATIONS

A Convolutional Neural Network (CNN) is a type of deep learning model designed primarily to process data with a grid-like structure, such as images or spatial and temporal signals. CNNs are inspired by the functioning of the visual cortex of the human brain, which responds selectively to visual patterns such as edges, textures, and shapes [4].

It is a machine learning system that mimics human visual perception, capable of learning hierarchical patterns and complex representations from large volumes of visual data. Thanks to their generalization capacity and efficiency, CNNs have become an essential tool in modern artificial intelligence and in various areas of application [4].

The fundamental principle of a CNN is the convolution operation, whereby the model applies filters (also called kernels) to images to automatically extract relevant features at different levels of abstraction. In the first layers, simple features such as edges or colors are detected, while in the deeper layers, more complex shapes such as complete objects are recognized [4].

These networks have demonstrated robust performance in computer vision tasks such as facial recognition, medical image diagnosis, organism classification, object detection, and autonomous driving. Unlike traditional methods, CNNs do not require an expert to manually define visual features, as they learn directly from the data, which increases accuracy and reduces human bias. In addition, CNNs achieve very high levels of accuracy in visual recognition tasks, outperforming image classification or object detection. Thanks to the use of graphics processing units (GPUs), these networks can handle large volumes of visual information in a short time, making them efficient and scalable [7].

However, CNNs also have limitations. They require large amounts of labeled data to achieve adequate performance, as well as high computational costs, which can limit their application in contexts with limited data and resources. Furthermore, when the dataset is small or lacks variety, the model may overfit, i.e., learn specific patterns from the training that do not generalize correctly to new data [3].

Another limitation relates to the lack of interpretability of the results. CNNs are considered "black box" models because it is not always possible to know exactly how they make decisions, which creates uncertainty in areas where explaining the process is as important as the prediction, such as in medicine or biology. Similarly, their performance can be affected by variability in external conditions, such as camera angle, the image capture device used, or the background of the images. Finally, if the training data is biased or imbalanced between classes, the model may reproduce those same biases in its results, affecting the accuracy of the predictions [8].

CNNs are a fundamental tool for automated image processing and analysis, with applications in numerous fields of knowledge. They have transformed the analysis of biological images by offering an automated, efficient, and accurate way to process large volumes of visual information. In recent years, these architectures have become an important tool in computational biology, ecology, and digital taxonomy, enabling species identification, organism counting, and morphological characterization from photographs or video sequences. Likewise, they have shown outstanding results in the identification of animal species captured by camera traps or drones, achieving greater accuracy than experts [9].

Another limitation in the analysis of biological images is the scarcity of labeled data, since obtaining high-quality images for each species or taxonomic group is costly and requires expert knowledge. This problem becomes more relevant when there are unbalanced classes, causing the model to favor the classes with more images and reduce the accuracy of the minority classes [10].

IV. LITERATURE REVIEW AND DESCRIPTIVE SYNTHESIS OF THE RESEARCH QUESTIONS

A. RQ1. What pretrained convolutional neural network architectures have been used to detect morphology in diptera?

It is designed for large-scale image classification and recognition. It consists of 16 weighted layers (13 convolutional and 3 dense). Its key architectural principle is the exclusive use of 3×3 convolutional filters stacked in blocks, followed by Max Pooling layers (2×2). The process begins with an input image (224×224 RGB), from which features are extracted to be finally classified by the dense layers with Softmax activation [11].

Faster R-CNN

It is a two-stage object detection model based on Convolutional Neural Networks (CNN):

1. Region Proposal Network (RPN): Uses anchors with various scales and aspect ratios to propose regions, classifying them and adjusting their coordinates (regression).
2. Detection (Fast R-CNN): It uses ROI Pooling/Align to standardize the proposals, then performs the final classification of the object and the final refinement of the bounding box.

The parameters include the choice of backbone, the configuration of the anchors, and the use of Soft L1 Loss for box regression [12].

YOLOv5

It is a single-stage real-time object detector with an architecture divided into three parts:

1. Backbone: Uses CSPDarknet53 for efficient feature extraction.
2. Neck: Employs SPPF and PANet to fuse and enhance features at different scales, which is key to its high performance.
3. Head: This is the final layer that performs simultaneous prediction of class, objectivity, and bounding box coordinates.

The parameters are defined by model variants that control the depth and width of the network, using the CIoU loss function (for localization) and BCE (for classification) [13].

Darknet

It is the convolutional architecture that serves as the backbone for the YOLO algorithm, optimized for real-time object detection. The Darknet process (especially in versions such as Darknet-53) is based on a fully convolutional architecture that incorporates residual connections (like ResNet) to increase depth. Its key parameters include the use of 1×1 layers for dimensionality reduction and the Leaky ReLU (or Mish) activation function to improve gradient flow, allowing YOLO to predict bounding boxes and classes in a single pass through the network [14].

ResNet101 DC-5

It is a 101-layer Deep Residual Neural Network that uses residual connections to prevent accuracy degradation in deep learning model training.

It uses ResNet base architecture with bottleneck blocks. The DC-5 feature involves replacing the pooling layers in Stage 5 with dilated convolutions. Its main application is image classification, although it can be optimized for detection and semantic segmentation tasks [15].

ResNeXt101

It is an architecture that extends ResNet to improve accuracy by introducing cardinality as a new dimension of scalability. It is based on Transformation Aggregation. Each residual block divides it into multiple parallel and identical branches that process different aspects of the input before merging. Its main parameter is Cardinality (C), which is the number of independent groups or branches, usually using a value of C=32 [15].

ResNet18

It is the smallest and most efficient version of the Residual Neural Network family, designed for image classification and to serve as a fast backbone. Its most important process is the use of residual connections so that the gradient flows directly through its 18 layers with weights, solving the problem of accuracy degradation in deep networks. Unlike larger versions, it uses basic residual blocks (two 3×3 layers) and has approximately 11.7 million parameters, which gives it speed and efficiency in training [16].

RetinaNet

It is a single-stage image detector designed to achieve high accuracy in dense object detection. Its architecture combines a ResNet backbone with a Feature Pyramid Network (FPN) to process information at multiple scales. Its key process and characteristic parameter are Focal Loss, a modified loss function that resolves the extreme imbalance between positive and negative examples by heavily weighing difficult (misclassified) samples and reducing the weight of easy (background) samples, enabling effective training [17].

Mask R-CNN

It is an advanced model for instance detection and segmentation that extends Faster R-CNN, which detects objects and generates a precise pixel mask for each one. The architecture uses a ResNet-101 backbone combined with a Feature Pyramid Network (FPN) to create a multiscale feature pyramid (neck).

Its process is ROI Align, a parameter that replaces ROI Pooling to ensure precise alignment of the characteristics of the Regions of Interest (ROI) in order to achieve accuracy in the mask. It operates with three parallel heads (classification, box regression, and an FCN for the mask) for application in Instance Segmentation [18].

Swin Transformer-L

It is a hierarchical Vision Transformer architecture with high capacity, designed to be a backbone for achieving high performance in tasks such as object detection and segmentation. Its hierarchical architecture mimics CNNs and its process is Shifted Window Attention (SW-MSA). This mechanism limits the self-attention calculation to local windows, solving the quadratic complexity of traditional Transformers and allowing communication between neighboring windows, resulting in a network with approximately 197 million parameters and superior efficiency and accuracy in vision applications [19].

ISSN:1390-9266 e-ISSN:1390-9134 LAJC 2026
DOI: https://doi.org/10.33333/lajc.vol13n2.03

LATIN-AMERICAN JOURNAL OF COMPUTING (LAJC), Vol XIII, Issue 2, July 2026

MobileNet

It is a lightweight CNN architecture for detection, classification, and segmentation on mobile and edge devices with limited resources. Its process and characteristic feature are the use of Depth Separable Convolutions (DSC), which divide the convolution operation into two steps (depth and point), significantly reducing the number of parameters and the computational load compared to standard convolutions. Its scaling parameters (Width Factor and Resolution Factor) allow the model to be adjusted for different latency and power constraints [20].

B. *RQ2. What CNN based approaches analyze Diptera morphology?*

Automatic identification of mosquito species has become a central field of research, focusing on the use of convolutional neural networks (CNN) and other deep learning techniques to overcome the challenges of manual classification. Recent studies demonstrate the feasibility and high accuracy of these methods by evaluating the effectiveness of different model architectures and specific anatomical parts used for identification [1].

A key approach focuses on the use of citizen science images, which introduce variability and field conditions. [1] proposed a method for the automated classification of Aedes albopictus mosquitoes using the VGG16 architecture. Unlike other studies based on laboratory images, this model was trained with a large dataset of field images collected by volunteers through the citizen science initiative "Mosquito Alert." Despite the way these images were obtained, the model's VGG16 architecture achieved a test accuracy of 94.61%, demonstrating the feasibility of using neural networks to classify mosquitoes. They applied the Grad-CAM algorithm, and this analysis revealed that the model focuses on the white stripes located on the mosquito's legs, abdomen, and thorax, the same characteristics that entomologists use for identification. It was found that classification errors were directly related to poor image quality, such as lack of clarity, occlusion, or damage to key parts of the mosquito's body. As a result, images of non-tiger mosquitoes with morphological similarities could be misclassified, highlighting the importance of image quality for model accuracy.

In the field of anatomical feature extraction, [21] developed a Mask R-CNN-based framework to automatically detect and extract anatomical components of mosquitoes (thorax, wings, abdomen, and legs) from images obtained with smartphones. They used 1,600 images of nine species for training and validation, and evaluated performance with metrics such as Precision, Recall, IoU, and mAP. The system used ResNet-101 combined with Feature Pyramid Network (FPN) and achieved 95% accuracy for thorax, abdomen, and wings, with a mAP of 60% in validation and 52% in testing.

C. *RQ3. What types of results are reported in studies on wing morphology?*

Wing morphology has proven to be a valuable feature for automatic species classification. [22] developed a convolutional neural network (CNN) to identify seven species of Aedes from wing images. They used 1,155 images of Aedes and 554 non-Aedes mosquitoes, and trained CNNs in grayscale and RGB. This model achieved an F1-score of 99% for differentiating Aedes from other mosquitoes and around 90–91% for classifying the seven species, with 100% accuracy for Aedes albopictus. Classification errors occurred mainly among similar native species.

Comparing effectiveness, [23] evaluated full-body and wing images, finding that models trained with wing images achieved higher accuracy (87.6%) than with body images (78.9%) for the identification of four morphologically similar Aedes species. Wing-based models required fewer images for reliable performance. Likewise, model performance decreased significantly when evaluated with images from devices not included in training, although the study highlights the viability of body- and wing-based classification methods.

[24] evaluated the geometric morphometry of wings to identify six species of the genus Aedes in northeastern France. They used 18 reference points on the wings, applied Procrustes overlap analysis and Canonical Variant Analysis, achieving 98% accuracy in reclassification.

[25] developed an automatic system to identify Aedes species using wing interference patterns (WIPs) and deep learning. With a set of 494 images, they trained several CNN architectures, including MobileNet, ResNet18, and reduced versions of DarkNet. The models achieved 95% accuracy at the genus level, with perfect classification in half of the species.

[26] proposed a two-stage method for the automatic classification of midge species of the genus Culicoides based on morphological analysis of their wings. They applied image preprocessing techniques (filters and morphological operations) and machine learning, achieving 95.31% accuracy for wing segmentation and 94.79% for particle segmentation.

Additionally, [27] focused on the analysis of wingbeat patterns, presenting a hybrid method for the classification of mosquito species that combines different machine learning and deep learning architectures (SVM, MLP, Random Forest, Gradient Boosting, and kNN). They showed that hybrid architecture outperforms individual algorithms, as they achieved high accuracy and balanced performance in multi-class classification.

D. *RQ4. What computational approaches have been proposed for morphological patterns in Diptera?*

The integration of sophisticated architectures, such as Transformers, has set new standards for accuracy. [27] developed a deep learning model for automatic mosquito species identification based on Swin Transformer. They created a balanced dataset of 9,900 high-resolution images of 17 species and 3 subspecies. When comparing various convolutional networks and transformer-based models, the Swin Transformer-L variant was selected for its higher accuracy (called Swin MSI), which achieved 99.04% accuracy. In addition, this model demonstrated robustness by achieving 96.26% accuracy when classifying species not included in the training.

Along the same lines of advanced models, [29] developed a deep learning image analysis method to identify eleven mosquito species in Korea. They trained and compared five object detection models: Faster R-CNN with Swin Transformer, YOLOv5, ResNet101 DC-5, ResNeXt 101, and

RetinaNet. The results showed that the combination of Swin Transformer + Faster R-CNN achieved the highest accuracy with an F1-score of 97.1%, and YOLOv5 with 96.4%. They found that the combined use of RGB and fluorescent images, together with the non-maximum suppression (NMS) technique, improved the performance of all models. They identified the small sample size in some species as a limitation.

Finally, [29] developed a system to identify species using convolutional neural networks with a multilevel model that detects unknown species. They used a database of 12,977 images of 2,696 wild species, many with morphological damage. The system combines CNNs for feature extraction with classifiers (SVM, Random Forest) and a Gaussian mixture model for low-confidence cases. It achieves 97.04% accuracy in classifying 16 known species and 89.50% accuracy in detecting novel species.

One of the trends for future research is to improve the robustness and generalization of deep learning models. The study by [1] already highlighted that, despite using a large field dataset from the "Mosquito Alert" initiative, classification errors were directly related to poor image quality. This points to the need to develop models that are more robust to variability. Future research should focus on: preprocessing and data augmentation techniques; image quality detection models that can filter or warn about problematic images before classification and data collection, as suggested by [29] when identifying the small sample size in some species as a limitation, and the suggestion to improve the capture process to increase the volume of training data.

There is a trend to further investigate and validate the effectiveness of specific anatomical parts as descriptors, particularly wings. Findings from [22], [23] demonstrated that models trained with wing images achieved higher accuracy and required less data than full-body images. This will aid research into: the systematic comparison between the geometric morphometrics of [25], wing interference patterns (WIPs, Cannet et al., 2023), and wing-based CNNs to determine the most efficient technique; the application of anatomical extraction frameworks [20] to isolate and improve the image quality of wings and legs before classification; and the exploration of new motion-based descriptors, such as the flapping pattern analysis proposed by [26].

The implementation of architectures such as Transformer-based models to improve the identification of unclassified species. The Swin MSI model [27] has already demonstrated superior performance (99.04% accuracy) and generalization ability (96.26% on unseen species). Future research will focus on: exploring hybrid models that combine the high accuracy of Transformers with the ability to classify unknown or low-confidence species, as did the multilevel model of [30], which achieved 89.50% accuracy in detecting novel species; optimizing the combination of different image modalities, following the example of [28] with the use of RGB and fluorescent images to improve performance.

V. Conclusions

This research presents a mapping of the literature on the application of convolutional neural networks (CNNs) for classifying dipteran morphology. A descriptive analysis of the included studies was used, based on which the predominant architectures, the anatomical parts used, and the most frequently employed computational approaches were mapped, providing a current overview of the field.

According to the results found, the literature relies on pre-trained CNN architectures, such as VGG-16 and ResNet, as well as object detection models such as YoloV5, Faster R-CNN, and Mask R-CNN, which have demonstrated good performance in classification and morphological detection tasks. These architectures are mainly used through transfer learning, which allows for robust results even in scenarios with limited datasets. In this regard, the analysis shows evidence that wing morphology-based approaches consistently report better performance metrics, with accuracy and F1 scores exceeding 90% in most of the consulted articles. This confirms that the wings constitute a highly discriminating anatomical region, allowing for more precise and efficient models compared to using the whole body. In contrast, using whole-body images tends to employ more complex pipelines, based on higher detection or other strategies, to handle visual and structural variability.

Methodologically, the mapping shows a predominance of direct classification approaches, complemented by detection and segmentation, especially when it is necessary to locate specific structures or work with damaged specimens. Articles were identified that study hybrid models based on CNNs and Transforms, which represent a promising line of research for capturing complex spatial relationships, although their adoption is limited. The findings of this mapping are useful for researchers in computer vision and machine learning, as well as for entomologists, biologists, and digital taxonomy specialists, as it provides an overview of current trends in CNNs applied to image classification of dipteran morphology. This mapping contributes to a better understanding of the current state of applications in dipteran morphological classification, providing a solid foundation for future research in medical entomology.

References

[1] Adhane, G., Dehshibi, M. M., & Masip, D. (2021). A deep convolutional neural network for classification of Aedes albopictus mosquitoes. *IEEE Access*, 9, 72681–72690.

[2] Elhamod, M., et al. (2022). Hierarchy-guided neural network for species classification. *Methods in Ecology and Evolution*, 13(3), 642–652.

[3] Zhou, Z., et al. (2023). EchoAI: A deep-learning based model for classification of echinoderms in global oceans. *Frontiers in Marine Science*, 10, 1147690.

[4] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444.

[5] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84–90.

[6] Oliveira, M. B., et al. (2025). Classification of animal species via deep neural networks and species distribution modeling: A systematic review. *Artificial Intelligence Review*, 58, 230.

LATIN-AMERICAN JOURNAL OF COMPUTING (LAJC), Vol XIII, Issue 2, July 2026

[7] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.

[8] Litjens, G., et al. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, 60–88.

[9] Miao, Z., et al. (2019). Insights and approaches using deep learning to classify wildlife. *Scientific Reports*, 9, 8137.

[10] Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. In *Proceedings of the International Conference on Learning Representations (ICLR)*.

[11] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. In *Advances in Neural Information Processing Systems* (Vol. 28, pp. 91–99).

[12] Khanam, R., & Hussain, M. (2024). What is YOLOv5: A deep look into the internal features of the popular object detector. arXiv:2407.20892.

[13] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 779–788).

[14] Xie, S., et al. (2017). Aggregated residual transformations for deep neural networks. In Proceedings of the *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 5987–5995).

[15] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 770–778).

[16] Lin, T. Y., et al. (2017). Focal loss for dense object detection. In Proceedings of the *IEEE International Conference on Computer Vision (ICCV)* (pp. 2980–2988).

[17] He, K., et al. (2017). Mask R-CNN. In Proceedings of the *IEEE International Conference on Computer Vision (ICCV)* (pp. 2961–2969).

[18] Liu, Z., et al. (2021). Swin transformer: Hierarchical vision transformer using shifted windows. In Proceedings of the *IEEE/CVF International Conference on Computer Vision (ICCV)* (pp. 10012–10022).

[19] Howard, A. G., et al. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. arXiv:1704.04861.

[20] Minakshi, M., et al. (2020). A framework based on deep neural networks to extract anatomy of mosquitoes from images. arXiv:2007.14725.

[21] Sauer, F. G., et al. (2024). A convolutional neural network to identify mosquito species (Diptera: Culicidae) of the genus Aedes by wing images. *Scientific Reports,* 14(1), 3094.

[22] Nolte, K., et al. (2025). Potentials and limitations in the application of convolutional neural networks for mosquito species identification using wing images. *PLoS Computational Biology*, 21(9), e1013435.

[23] Martinet, J.-P., et al. (2021). Wing morphometrics of Aedes mosquitoes from north-eastern France. *Insects*, 12(4), 341.

[24] Cannet, A., et al. (2023). Wing interferential patterns (WIPs) and machine learning for the classification of some Aedes species of medical interest. *Scientific Reports*, 13(1), 11956.

[25] Venegas, P., et al. (2020). An approach to automatic classification of Culicoides species by learning the wing morphology. *PLoS ONE*, 15(11), e0241798.

[26] Gireesh, A., & Noortaj, S. (2025). Hybrid machine learning approach for mosquito species classification using wing beat analysis. International Journal of Scientific Research in Science, Engineering and Technology, 12(3), 1062–1070. https://ijsrset.com/index.php/home/article/view/IJSRSET2512124

[27] Zhao, D., et al. (2022). A Swin transformer-based model for mosquito species identification. *Scientific Reports*, 12(1), 18664.

[28] Lee, S., Kim, H., & Cho, B.-K. (2023). Deep learning-based image classification for major mosquito species inhabiting Korea. *Insects*, 14(6), 526.

[29] Goodwin, A., et al. (2021). Mosquito species identification using convolutional neural networks with a multitiered ensemble model for novel species detection. *Scientific Reports*, 11(1), 24040.

AUTHORS

Benjamín Mendoza

Benjamín Paulino Mendoza Contreras es estudiante de octavo semestre de la Licenciatura en Estadística en la Universidad Veracruzana, con sede en Xalapa, Veracruz, México. Sus intereses de investigación se centran en el aprendizaje automático (machine learning) y el aprendizaje profundo (deep learning), particularmente en el desarrollo y aplicación de modelos estadísticos y computacionales para el análisis de datos. Actualmente participa en actividades académicas relacionadas con la ciencia de datos y la modelación estadística.

Emmanuel Morales

Licenciado en Ciencias y Técnicas Estadísticas y Especialista en Métodos Estadísticos por la Universidad Veracruzana, con Maestría en Ciencias de la Información Geoespacial por el Centro Geo, CDMX (Centro CONAHCYT), actualmente cursando el Doctorado en Ciencias de la Computación en la Universidad Veracruzana. Profesor en la Licenciatura en Estadística, en la Especialidad en Métodos Estadísticos y la Maestría en Economía y Sociedad de China y América Latina en la misma universidad, donde he dirigido 15 tesis de licenciatura y 4 de especialidad. También tengo experiencia como Analista Estadístico en la Oficina del Programa de Gobierno del Estado de Veracruz. Mis líneas de investigación incluyen metodologías de cómputo, programación estadística, estadística multivariada, análisis espacial, ciencia de datos y modelos estadísticos, con aplicaciones en biología, medicina, ciencias administrativas y sociales. He participado en diversos congresos nacionales e internacionales.

ISSN:1390-9266 e-ISSN:1390-9134 LAJC 2026

AUTHORS

Cecilia Cruz

Profesora de tiempo completo en la Facultad de Estadística e Informática de la Universidad Veracruzana (UV). Es Doctora en Investigación Educativa por la UV, Maestra en Ciencias con especialidad en Estadística Aplicada por el ITESM Campus Monterrey, así como Especialista en Métodos Estadísticos y Licenciada en Estadística por la UV. Actualmente coordina la Especialización en Métodos Estadísticos y pertenece al Sistema Nacional de Investigadores (SNII) como Candidata.

Sus líneas de investigación abarcan la Educación Estadística, la Metodología Estadística Aplicada y la integración de la estadística con tecnologías emergentes como Machine Learning, Ciencia de Datos y Análisis Espacial. Ha colaborado en proyectos sobre sustentabilidad, alfabetización digital y actitudes hacia la estadística en estudiantes latinoamericanos.

Entre sus publicaciones recientes destacan trabajos sobre aprendizaje supervisado, formación en consultoría estadística y aplicaciones multivariantes, además de capítulos sobre alfabetización digital y redes sociales. Ha dirigido más de 70 tesis, combinando docencia, investigación e impulso al uso estratégico de la estadística para el desarrollo sostenible.

Luis Gomez

Es doctor en Administración y doctor en Finanzas Públicas. Estudió la Maestría en Impuestos, la Maestría en Contabilidad, así como la Especialidad en Administración. Es licenciado en Contaduría Pública Certificado por el IMCP y Licenciado en Psicología.

En su experiencia docente se desempeña como coordinador del Posgrado en Administración modalidad Virtual, asimismo es integrante del Cuerpo Académico Consolidado “Las organizaciones y su entorno”. Investigador de Tiempo Completo del Instituto de Investigaciones y Estudios Superiores de las Ciencias Administrativas. El Dr. Luis Enrique es miembro del sistema nacional de investigadores e investigadoras, miembro certificado por Conocer, perfil deseable PRODEP y nivel 6 de productividad UV.

Además de ser autor y coautor de varios libros y revistas.

B. Mendoza, E. Morales, C. Cruz, and L. Gomez,
“Morphological classification of hematophagous Diptera with Convolutional Neural Networks: A mapping of literature,”
Latin-American Journal of Computing (LAJC), vol. 13, no. 2, 2026.