

INTELIGÊNCIA ARTIFICIAL EXPLICÁVEL APLICÁVEL À IDENTIFICAÇÃO DE SEXOS DO MOSQUITO Aedes Aegypti: UMA ABORDAGEM COM REDES NEURAIS CONVOLUCIONAIS

EXPLAINABLE ARTIFICIAL INTELLIGENCE APPLICABLE TO SEX
IDENTIFICATION OF Aedes Aegypti MOSQUITOES: A CONVOLUTIONAL
NEURAL NETWORKS APPROACH

INTELENCIA ARTIFICIAL EXPLICABLE APLICABLE A LA IDENTIFICACIÓN
DE SEXO DE LOS MOSQUITOS Aedes Aegypti: UN ENFOQUE CON REDES
NEURONALES CONVOLUCIONALES

João Francisco Rangel Pinheiro Cavalcanti¹, Erika Carlos Medeiros², Patricia Cristina Moser³, Jorge
Cavalcanti Barbosa Fonsêca⁴, Rômulo César Dias de Andrade⁵, Fernando Ferreira de Carvalho⁶,
Fernando Pontual de Souza Leão Junior⁷

DOI: 10.54899/dcs.v22i85.3886

Recibido: 26/11/2025 | Aceptado: 27/11/2025 | Publicación en línea: 12/12/2025.

RESUMO

Os arbovírus transmitidos por *Aedes aegypti* representam um desafio na saúde pública, tornando crucial a identificação sexual dos mosquitos, pois somente as fêmeas são hematófagas e vetoras. Embora a plumosidade antenal e o comprimento dos palpos sejam os marcadores diagnósticos clássicos, observa-se que estas estruturas são pequenas, limitando sua exploração por modelos de visão por computador em imagens de corpo inteiro. O presente estudo teve como objetivo construir e avaliar uma Rede Neural Convolucional para a predição de sexo, treinada com 2.000 imagens balanceadas de *A. aegypti*. A Rede Neural Convolucional, estruturada em três blocos convolucionais, foi avaliada com métricas padrão (precisão, acurácia, precisão, *recall*, pontuação F1 e matriz de confusão); sua interpretabilidade foi analisada com *Grad-CAM* e *Grad-CAM++*. Os resultados indicaram uma precisão de 0,81 após o ajuste fino, revelando assimetria: as fêmeas apresentaram maior precisão (0,90) mas menor *recall* (0,69), enquanto os machos obtiveram alta *recall* (0,92). Os mapas *Grad-CAM* demonstraram que a Rede Neural Convolucional se baseou em regiões anatômicas maiores (tórax, asas e abdômen), não explorando os caracteres

¹ Bacharel em Sistemas de Informação, Universidade de Pernambuco (UPE), Caruaru, Pernambuco, Brasil.
E-mail: joaofrancisco1015@gmail.com

² Doutora em Ciência da Computação, Universidade Federal de Pernambuco (UFPE), Recife, Pernambuco, Brasil.
E-mail: erika.medeiros@upe.br Orcid: <https://orcid.org/0000-0003-2506-7116>

³ Doutora em Ciência da Computação, Universidade Federal de Pernambuco (UFPE), Recife, Pernambuco, Brasil.
E-mail: patricia.moser@upe.br

⁴ Doutor em Ciência da Computação, Universidade Federal de Pernambuco (UFPE), Recife, Pernambuco, Brasil.
E-mail: jorge.fonseca@upe.br Orcid: <https://orcid.org/0000-0001-7954-2766>

⁵ Doutor em Ciência da Computação, Universidade Federal de Pernambuco (UFPE), Recife, Pernambuco, Brasil.
E-mail: romulo.andrade@upe.br Orcid: <https://orcid.org/0000-0002-0561-7507>

⁶ Doutor em Ciência da Computação, Universidade Federal de Pernambuco (UFPE), Recife, Pernambuco, Brasil.
E-mail: fernando.carvalho@upe.br Orcid: <https://orcid.org/0009-0005-4677-6917>

⁷ Doutor em Administração, Universidade Federal de Pernambuco (UFPE), Recife, Pernambuco, Brasil.
E-mail: fernando.pontual@upe.br

morfológicos cefálicos clássicos. Concluiu-se que as Redes Neurais Convolucionais treinadas com imagens de corpo inteiro alcançam um desempenho razoável, mas não reproduzem os critérios entomológicos clássicos de identificação sexual. Para uma aplicação prática em vigilância, recomenda-se o pré-processamento orientado à cabeça, a normalização da escala e experimentos de ablação, para direcionar o foco do modelo para os marcadores morfológicos relevantes.

Palavras-chave: *Aedes aegypti*. Sexagem. Redes Neurais Convolucionais. *Grad-CAM*. *Grad-CAM++*. XAI.

ABSTRACT

Arboviruses transmitted by *Aedes aegypti* represent a challenge in public health, making the sexual identification of mosquitoes crucial, since only females are hematophagous and vectors. Although antennal plumosity and palp length are the classic diagnostic markers, it is observed that these structures are small, limiting their exploitation by computer vision models in whole-body images. The present study had as objective to construct and evaluate a Convolutional Neural Network for sex prediction, trained with 2,000 balanced images of *A. aegypti*. The Convolutional Neural Network, structured in three convolutional blocks, was evaluated with standard metrics (precision accuracy, precision, recall, *F1-score*, and confusion matrix); its interpretability was analyzed with *Grad-CAM* and *Grad-CAM++*. The results indicated an accuracy of 0.81 after fine-tuning, revealing asymmetry: females presented higher precision (0.90) but lower recall (recall) (0.69), while males obtained high recall (0.92). The *Grad-CAM* maps demonstrated that the Convolutional Neural Network relied on larger anatomical regions (thorax, wings, and abdomen), not exploiting the classic cephalic morphological characters. It was concluded that Convolutional Neural Networks trained with whole-body images achieve reasonable performance, but do not reproduce the classic entomological criteria for sexual identification. For a practical application in surveillance, head-oriented preprocessing, scale normalization, and ablation experiments are recommended, to direct the model's focus toward the relevant morphological markers.

Keywords: *Aedes aegypti*. Sexing/Sex Identification. Convolutional Neural Networks. *Grad-CAM*. *Grad-CAM++*. XAI.

RESUMEN

Los arbovirus transmitidos por *Aedes aegypti* representan un desafío en salud pública, haciendo crucial la identificación sexual de los mosquitos, pues solo las hembras son hematófagas y vectores. Si bien la plumosidad antenal y la longitud de los palpos son los marcadores diagnósticos clásicos, se observa que estas estructuras son pequeñas, limitando su explotación por modelos de visión por computadora en imágenes de cuerpo entero. El presente estudio tuvo como objetivo construir y evaluar una Red Neuronal Convolucional para la predicción de sexo, entrenada con 2,000 imágenes balanceadas de *A. aegypti*. La Red Neuronal Convolucional, estructurada en tres bloques convolucionales, fue evaluada con métricas estándar (precisión *accuracy*, precisión, *recall*, puntuación F1 y matriz de confusión); su interpretabilidad fue analizada con *Grad-CAM* y *Grad-CAM++*. Los resultados indicaron una precisión de 0.81 tras el ajuste fino, revelando asimetría: las hembras presentaron mayor precisión (0.90) pero menor *recall* (0.69), mientras que los machos obtuvieron alto *recall* (0.92). Los mapas *Grad-CAM* demostraron que la Red Neuronal Convolucional se basó en regiones anatómicas más grandes

(tórax, alas y abdomen), no explotando los caracteres morfológicos cefálicos clásicos. Se concluyó que las Redes Neuronales Convolucionales entrenadas con imágenes de cuerpo entero logran un rendimiento razonable, pero no reproducen los criterios entomológicos clásicos de identificación sexual. Para una aplicación práctica en vigilancia, se recomienda el preprocesamiento orientado a la cabeza, la normalización de la escala y experimentos de ablación, para dirigir el enfoque del modelo hacia los marcadores morfológicos relevantes.

Palabras clave: *Aedes aegypti*. Sexado. Redes Neuronales Convolucionales. Grad-CAM. Grad-CAM++. XAI.



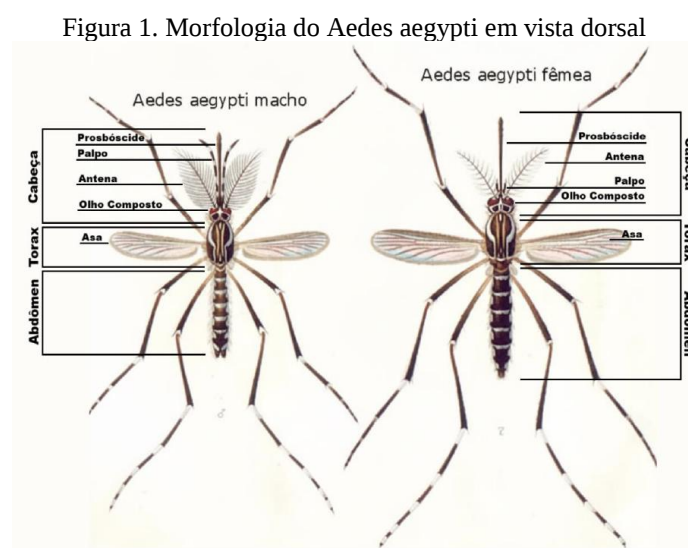
Esta obra está bajo una [Licencia Creative Commons Atribución- NoComercial 4.0 Internacional](https://creativecommons.org/licenses/by-nc/4.0/)

INTRODUÇÃO

As arboviroses transmitidas por mosquitos do gênero *Aedes* permanecem como desafio prioritário da saúde pública. Em especial, destaca-se a dengue, transmitida pela fêmea do *Aedes aegypti*, cuja hematofagia permite a aquisição e transmissão do vírus ao ser humano (World Health Organization, 2024). Em diversos contextos urbanos e periurbanos, o *Aedes aegypti* também é vetor de outros patógenos (zika, chikungunya e febre amarela), o que reforça a centralidade da vigilância entomológica rápida e confiável (European Centre for Disease Prevention and Control, 2023). A identificação correta do sexo do mosquito é operacionalmente relevante: somente as fêmeas picam e, portanto, compõem o contingente vetor. Diferenciar machos e fêmeas apoia decisões de campo, como a priorização de ações de controle, a calibração de armadilhas e programas baseados na liberação de machos. Tradicionalmente, a sexagem em Culicidae baseia-se em caracteres da cabeça — sobretudo a plumosidade das antenas e o comprimento ou ornamentação dos palpos maxilares —, mais evidentes nos machos (Burkett-Cadena, 2013; Li *et al.*, 2022). Para a sexagem do *Aedes aegypti*, o especialista concentra a inspeção na região cefálica (antenas e palpos maxilares junto à probóscide). Os machos apresentam antenas plumosas (flagelômeros com tufo denso de cerdas, aspecto "penachado") e palpos longos, que se aproximam do comprimento da probóscide; já as fêmeas têm antenas não plumosas/pilosas (cerdas curtas e esparsas, aspecto liso) e palpos curtos, nitidamente menores que a probóscide (Burkett-Cadena, 2013; Li *et al.*, 2022). Especialistas do Instituto Aggeu Magalhães – FIOCRUZ, consultados pelos pesquisadores, confirmaram que as estruturas morfológicas utilizadas para a sexagem dos mosquitos são as antenas, os palpos e a probóscide

(Instituto Aggeu Magalhães – FIOCRUZ, s.d.). Na Figura 1, observa-se a morfologia do *Aedes aegypti* em vista dorsal.

Nos últimos anos, a visão computacional baseada em aprendizado profundo tem sido aplicada à entomologia médica para a classificação de mosquitos por espécie e, em alguns estudos, também por sexo. Modelos do tipo Redes Neurais Convolucionais (CNN) e variantes modernas (por exemplo, *You Only Look Once - YOLO* e *Transformers* visuais) alcançaram altas acurácias em imagens de corpo inteiro, de asas, e até em sinais acústicos de batimento de asas, indicando potencial para automatizar rotinas de vigilância (Kittichai *et al.*, 2021; Zhao *et al.*, 2022; Sauer *et al.*, 2024; Supratak *et al.*, 2024). Em especial, Kittichai *et al.* (2021) demonstraram detecção simultânea de espécie e sexo em amostras de campo com *YOLO*; Zhao *et al.* (2023) reportaram que “as chaves morfológicas usadas pelo modelo são semelhantes, mas não idênticas, às usadas por humanos”; e abordagens por áudio têm discriminado espécie/sexo com robustez e ruído de campo.



Fonte: Pinheiro, 2025

A adoção de ferramentas de Inteligência Artificial Explicável (*Explainable AI – XAI*) é crucial para avaliar a confiabilidade desses sistemas em cenários de saúde pública. O Mapeamento de Ativação de Classe Ponderado por Gradiente (*Gradient-weighted Class Activation Mapping – Grad-CAM*) projeta mapas de ativação sobre a imagem de entrada, indicando regiões que mais contribuíram para a decisão do modelo (Selvaraju *et al.*, 2017). Em aplicações recentes com mosquitos, o *Grad-CAM* tem sido utilizado para elucidar decisões em

diferentes tarefas (espécie, estágios fisiológicos e sexagem) (Azam *et al.*, 2023; Zhao *et al.*, 2022).

Segundo Andrew e Bar (2013), observa-se que no macho são maiores o palpo maxilar ($\approx 0,77 \pm 0,06$ mm *versus* $0,08 \pm 0,01$ mm na fêmea), a probóscide ($\approx 0,76 \pm 0,04$ mm *versus* $0,66 \pm 0,03$ mm), a largura da cabeça ($\approx 0,81 \pm 0,13$ mm *versus* $0,73 \pm 0,11$ mm) e o abdômen (comprimento $\approx 3,03 \pm 0,18$ mm; largura $\approx 0,51 \pm 0,07$ mm *versus* $2,94 \pm 0,20$ mm e $0,41 \pm 0,06$ mm nas fêmeas). Na fêmea, predominam tórax maior ($\approx 0,50 \pm 0,08$ $\times$ $0,35 \pm 0,07$ mm *versus* $0,41 \pm 0,06$ $\times$ $0,29 \pm 0,02$ mm no macho) e asas mais longas, além de maior comprimento ântero-posterior da cabeça ($\approx 0,55 \pm 0,09$ mm). Quanto às antenas, o comprimento é muito próximo entre os sexos ($\approx 0,57 \pm 0,03$ mm no macho *versus* $0,52 \pm 0,07$ mm na fêmea) e não é útil para sexagem; o que realmente distingue é a plumosidade antenal (densa no macho, discreta na fêmea).

O uso de visão computacional para a sexagem do mosquito *Aedes aegypti* suscita a seguinte pergunta de pesquisa: a identificação do sexo por Redes Neurais Convolucionais usa as mesmas características morfológicas que os especialistas?

Para responder à pergunta de pesquisa segue o objetivo geral: Construir uma CNN treinada com imagens de corpo inteiro de *Aedes aegypti*, para predição de sexo e checar se a decisão é guiada predominantemente pelas características de diagnósticos clássicas (antenas/palpos) usadas pelos especialistas.

Para atingir o objetivo geral, seguem quatro objetivos específicos:

1. Escolha do conjunto de dados de corpo inteiro de *Aedes aegypti*;
2. Construir uma CNN;
3. Avaliar o desempenho da CNN usando métricas de acurácia, precisão, recall, *F1-score* e matriz de confusão;
4. Analisar a confiabilidade através de técnicas de XAI.

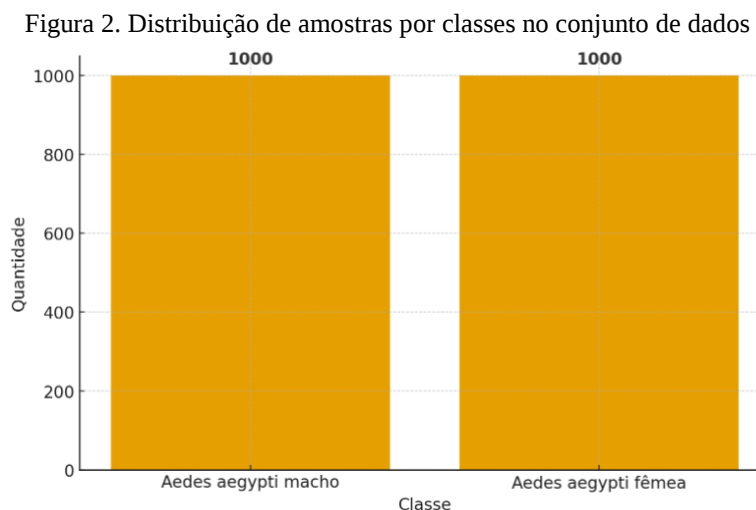
METODOLOGIA

Adotou-se como diretriz metodológica o modelo Cross-Industry Standard Process for Data Mining (CRISP-DM), amplamente reconhecido em projetos de mineração de dados e aprendizado de máquina. Neste trabalho, são contempladas as cinco etapas iniciais - compreensão

do negócio, compreensão dos dados, preparação dos dados, modelagem e avaliação, ficando a etapa de implantação fora do escopo.

A compreensão do negócio foi construída a partir da revisão de literatura e conversa com especialistas, com foco em mapear avanços, desafios e oportunidades no uso de modelos de aprendizado de máquina para antecipar a identificação de fêmeas de *Aedes aegypti*, vetor primário da dengue.

Na etapa de compreensão dos dados, caracteriza-se o conjunto utilizado: um acervo de 2.000 imagens de *Aedes aegypti*. As imagens estão em formato png e são coloridas. As imagens apresentam dimensões variadas contendo o espécime centrado, apresentando apenas o corpo do mosquito e sem o plano de fundo e quaisquer bordas residuais. O conjunto de dados foi criado por Júnior (2022) e tratado por Gomes (2025).



Fonte: Elaborado pelos autores.

O conjunto de dados é organizado de forma balanceada entre 1.000 machos e 1.000 fêmeas. Esse equilíbrio entre classes é essencial para reduzir vieses no treinamento e favorecer a precisão e a imparcialidade dos modelos de classificação a serem desenvolvidos. A Figura 2 ilustra a distribuição das classes, evidenciando a paridade entre as amostras. As imagens foram coletadas no Instituto Aggeu Magalhães, por meio de um microscópio digital HT-60S (5 MP). Para padronizar o registro, os espécimes foram manipulados com pinças de precisão e posicionados sobre uma base quadriculada de papel ofício, fixada na plataforma do microscópio. O protocolo de captura contemplou múltiplos ângulos dos mosquitos, com o objetivo de ampliar

a variabilidade do banco de dados e, conseqüentemente, a robustez do processo de modelagem (Júnior, 2022).

Na fase de preparação dos dados, as imagens foram redimensionadas para 224x224 pixels. Os pixels das imagens foram normalizados para o intervalo [0,1] por meio da divisão por 255, a fim de padronizar a escala de entrada.

Usou-se uma proporção de 70/15/15 para dados de treino, validação e teste. Para o conjunto de treino, aplicou-se aumento de dados *on-the-fly* com o objetivo de ampliar a variabilidade e mitigar o *overfitting*, preservando a semântica das classes. As transformações contemplaram: rotações aleatórias de até 40°; Translações horizontais e verticais de até 20% da dimensão da imagem; Cisalhamento de 0,2; Zoom aleatório entre 0,8 e 1,2; Inversões horizontal e vertical; E variação de brilho no intervalo [0.8, 1.2] para simular diferentes condições de iluminação. As regiões expostas por transformações geométricas foram preenchidas pelo método *nearest neighbor*, evitando artefatos de interpolação que possam introduzir padrões artificiais. Esse conjunto de operações promove invariância a variações de orientação, posição, escala e iluminação, características plausíveis no registro das imagens dos mosquitos.

Nos conjuntos de validação e teste, empregou-se apenas a normalização (sem aumentos), garantindo uma avaliação fidedigna do desempenho do modelo em dados não vistos e evitando vazamento de informação entre as partes.

Para a etapa de modelagem, empregou-se uma CNN voltada à classificação binária (machos vs. fêmeas). As entradas consistem em imagens com dimensões 224x224x3, coerentes com o pré-processamento descrito anteriormente. A arquitetura segue um desenho em três blocos convolucionais com complexidade crescente, cada um composto por uma convolução 3x3 com ativação ReLU, seguida de max pooling 2x2 para redução espacial e ganho de invariância translacional: o primeiro bloco utiliza 32 filtros, o segundo 64 filtros e o terceiro 128 filtros. Essa progressão deve permitir a captura desde padrões de baixa frequência (bordas, texturas simples) até características mais discriminativas associadas à morfologia do mosquito.

Após a extração de atributos, o mapa de ativação é achatado (*Flatten*) e encaminhado a uma camada densa de 512 neurônios com ReLU, que atua como cabeça classificadora ao combinar as características convolucionais em representações de maior nível. A camada de saída possui 2 neurônios com *softmax*, produzindo as probabilidades normalizadas por classe, condizentes com a codificação *one-hot* do rótulo.

O treinamento é formulado como minimização da perda *categorical cross-entropy*, apropriada para duas saídas com *softmax*, e otimizado por *Adam*, devido à adaptação de taxas de aprendizado por parâmetro e ao bom desempenho empírico em visão computacional. A acurácia é utilizada como métrica de monitoramento. Em alinhamento ao pipeline de dados, o treinamento ocorre em mini-lotes de 32 imagens.

A arquitetura da CNN é apresentada na tabela da Figura 3.

Na fase de avaliação são aplicadas as métricas acurácia, precisão, *recall*, *F1-score* e matriz de confusão, além dos algoritmos de XAI: *GRAD-CAM* e *GRAD-CAM++*. Os resultados são discutidos na próximo seção.

Figura 3. Arquitetura da CNN

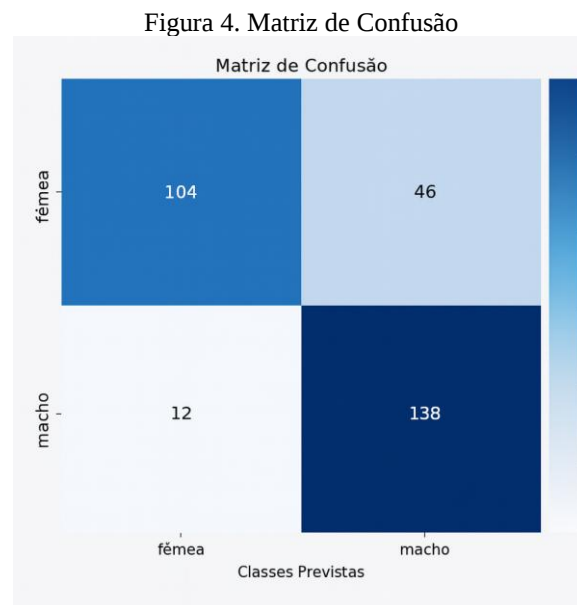
Layer (type)	Output shape	Param #
input_layer_1 (InputLayer)	(None, 224, 224, 3)	0
conv2d_1 (Conv2D)	(None, 222, 222, 32)	896
max_pooling2d_1 (MaxPooling2D)	(None, 111, 111, 32)	0
conv2d_2 (Conv2D)	(None, 109, 109, 64)	18.496
max_pooling2d_2 (MaxPooling2D)	(None, 54, 54, 64)	0
conv2d_3 (Conv2D)	(None, 52, 52, 128)	73,856
max_pooling2d_3 (MaxPooling2D)	(None, 26, 26, 128)	0
flatten_1 (Flatten)	(None, 86528)	0
dense_2 (Dense)	(None, 512)	44.392.848
dense_3 (Dense)	(None, 2)	1.026

Fonte: Elaborada pelos autores.

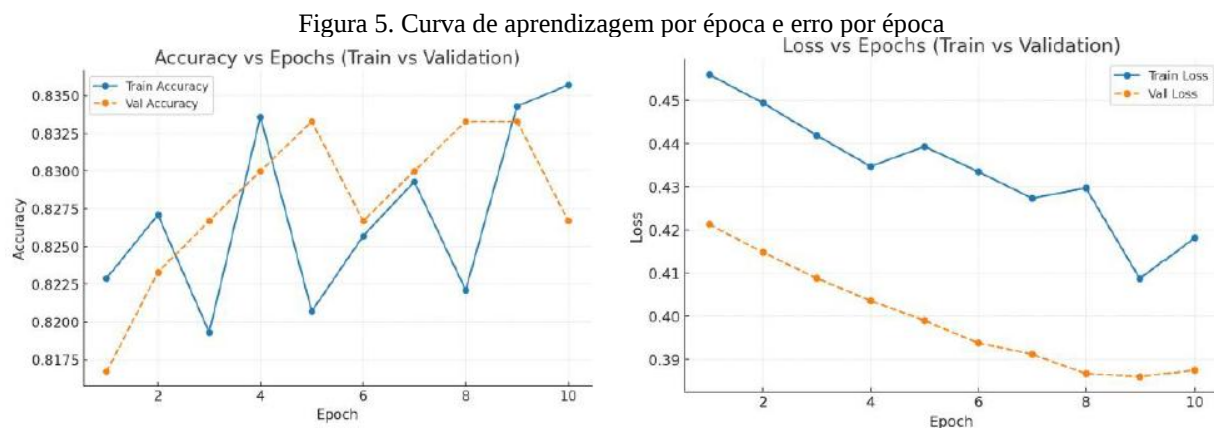
RESULTADOS E DISCUSSÃO

Nessa seção os resultados são discutidos. Para treinamento do modelo de CNN, como dito anteriormente, o conjunto de dados foi particionado em treino/validação/teste na proporção 70/15/15, assegurando balanceamento por classe em todas as divisões. Após a etapa de ajuste de hiperparâmetros, obteve-se 0,81 de acurácia, com médias macro e ponderada também em 0,81/0,80. Em termos de classe, o cenário pós-ajuste apresenta, para fêmeas, precisão =0,90, recall =0,69 e *F1-score* =0,78; Para machos, precisão =0,75, *recall* =0,92 e *F1-score* =0,83. A matriz de confusão, mostrada na Figura 4, corresponde (n=300; 150 amostras por classe) registra 104 fêmeas corretamente reconhecidas e 46 confundidas como machos, além de 138 machos corretos e 12 confundidos como fêmeas.

A distribuição dos erros observada na Figura 4 (46 fêmeas preditas como machos e 12 machos preditos como fêmeas) é explicada por uma combinação de dois fatores: visibilidade reduzida dos marcadores sexuais (antenas e palpos) em parte das imagens por ângulo, oclusão, desfoque ou enquadramento — uma vez que o recorte prioriza o corpo e, em alguns casos, a cabeça fica parcial ou totalmente fora da área útil e predominância de pistas texturais no tórax/abdômen, semelhantes entre os sexos, que tendem a guiar a rede em detrimento de estruturas cefálicas mais discriminativas. Em conjunto, esses elementos explicam a concentração de confusões, sobretudo no sentido fêmea→macho, apesar do balanceamento do conjunto de teste. Isso foi observado no teste de XAI feito através do *Grad-CAM* e *Grad-CAM++*, explicado adiante.



Fonte: Elaborada pelos autores.

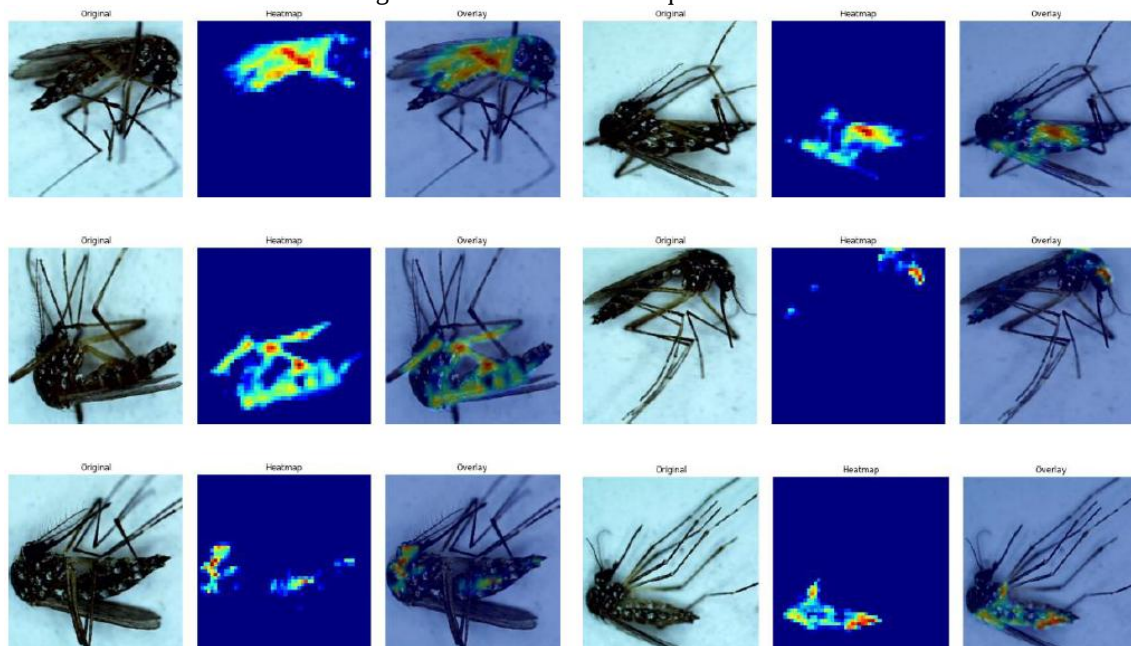


Fonte: Elaborado pelos autores.

Na figura 5 pode-se ver curvas de aprendizado ao longo de 10 épocas. A acurácia de validação acompanha de perto a de treino oscilando em torno de 0,83 e encerrando ligeiramente acima do ponto de partida, sem divergências sustentadas que caracterizariam sobreajuste. Já as curvas de perda (*loss*) apresentam tendência decrescente em treino e em validação, com a perda de validação discretamente inferior à de treino em vários momentos, um comportamento compatível com regularizações e/ou aumento de dados, e indicativo de que o modelo não memoriza o conjunto de treino.

É crucial compreender quais regiões do inseto estão efetivamente sustentando as decisões do modelo; por isso, no tópico seguinte discute-se a análise XAI via *Grad-CAM* e *Grad-CAM++* examinando se o padrão de atenção converge para os marcadores morfológicos empregados na sexagem, que são antenas e palpos.

Figura 6. *Grad-CAM* de mosquitos fêmea



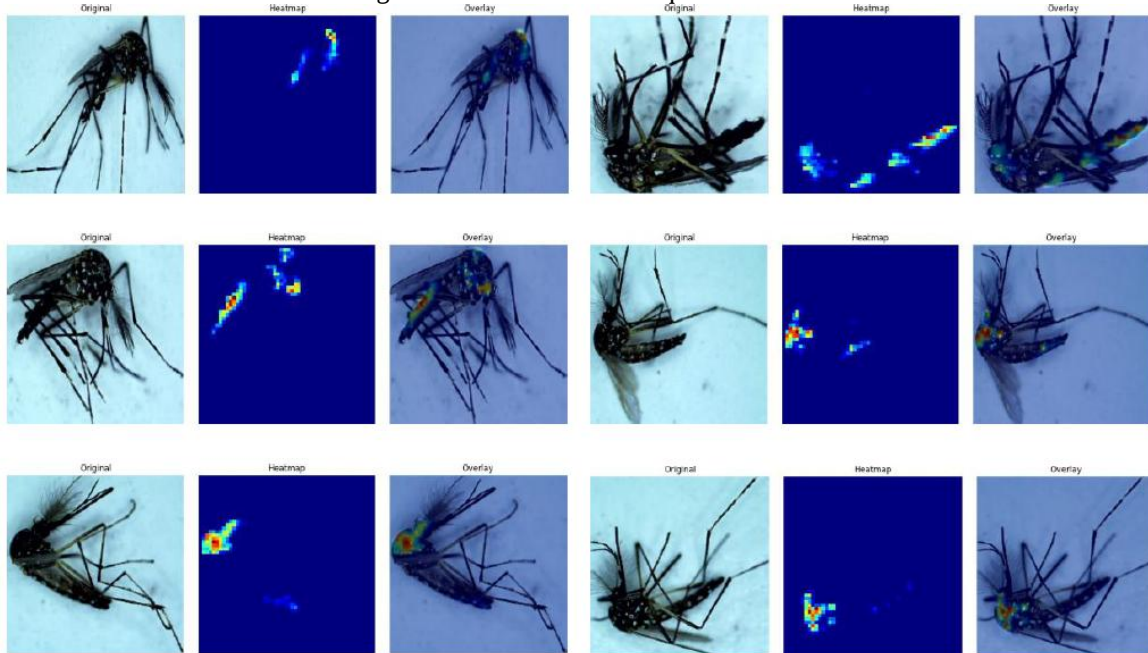
Fonte: Elaborado pelos autores.

A Figura 6 apresenta três imagens: Original, a foto bruta do mosquito, sem marcações; *Heatmap* (*Grad-CAM*), um mapa de saliência em que cores quentes (amarelo/vermelho) indicam maior atenção da CNN e cores frias (azul) menor atenção; e *Overlay*, o *heatmap* sobreposto à foto original para visualizar exatamente onde o modelo “olhou”.

Nas visualizações *Grad-CAM* de fêmeas (Figura 6), a saliência é predominantemente torácica, concentrada no mesonoto (dorso) e nas inserções das asas, com prolongamentos

ocasionais para o abdome. Em algumas amostras há foco secundário na cabeça, tipicamente sobre o olho composto e a base da probóscide. Antenas e palpos não aparecem como áreas quentes nos painéis *heatmap* e *overlay*. Como os marcadores sexuais empregados por especialistas para fêmeas dependem justamente da ausência de plumosidade antenal e da morfologia dos palpos, a ausência de saliência nessas estruturas, somada ao predomínio no tórax, indica desalinhamento entre as pistas usadas pela CNN e os atributos diagnósticos relevantes, o que é compatível com a assimetria nas métricas (maior confusão fêmea → macho).

Figura 7. *Grad-CAM* de mosquitos macho



Fonte: Elaborado pelos autores.

Nos machos, os mapas *Grad-CAM*, ilustrados na Figura 7, exibem um padrão multirregional de saliência, com focos distribuídos pela cabeça, tórax/mesonoto (incluindo a inserção das asas) e abdome. Na cabeça, os picos concentram-se sobretudo no olho composto e na probóscide, não recaindo sobre as antenas plumosas e palpos, que seriam os marcadores morfológicos mais discriminativos para sexagem. No tórax e no abdome, a atenção parece ser atraída por texturas escamadas e contrastes amplos, possivelmente reforçados por reflexos e variação de pose/iluminação. Esse comportamento sugere que a CNN decide com base em pistas genéricas dessas três regiões — o que pode sustentar o recall elevado (0,92) para machos, mas revela desalinhamento em relação aos atributos diagnósticos clássicos, indicando a necessidade de estratégias que direcionam o modelo às estruturas cefálicas relevantes.

Em síntese, o *Grad-CAM* sugere que o classificador não está explorando diretamente os marcadores sexuais diagnósticos — plumosidade antenal e morfologia dos palpos — nem em machos nem em fêmeas; em vez disso, privilegia padrões de tórax, asas e regiões cefálicas não antenais.

Os indícios reunidos pelas métricas e pelos mapas *Grad-CAM* não nos permitem afirmar que uma CNN treinada com imagens de corpo inteiro seja inadequada para a sexagem de *Aedes aegypti*; todavia, sugerem fortemente que, nesse arranjo, o modelo tende a priorizar pistas genéricas (tórax, asas e porções do abdome) e desconsiderar os marcadores morfológicos mais diagnósticos (antenas e palpos). Esse desalinhamento pode explicar a persistência de confusões fêmea→macho, pois as evidências realmente específicas do sexo ficam, muitas vezes, pouco visíveis ou sub-representadas no campo de visão efetivo da rede. Importa frisar que o *Grad-CAM* tem resolução limitada e indica importância relativa — portanto, as conclusões devem ser lidas como hipóteses apoiadas por evidência, não como prova definitiva.

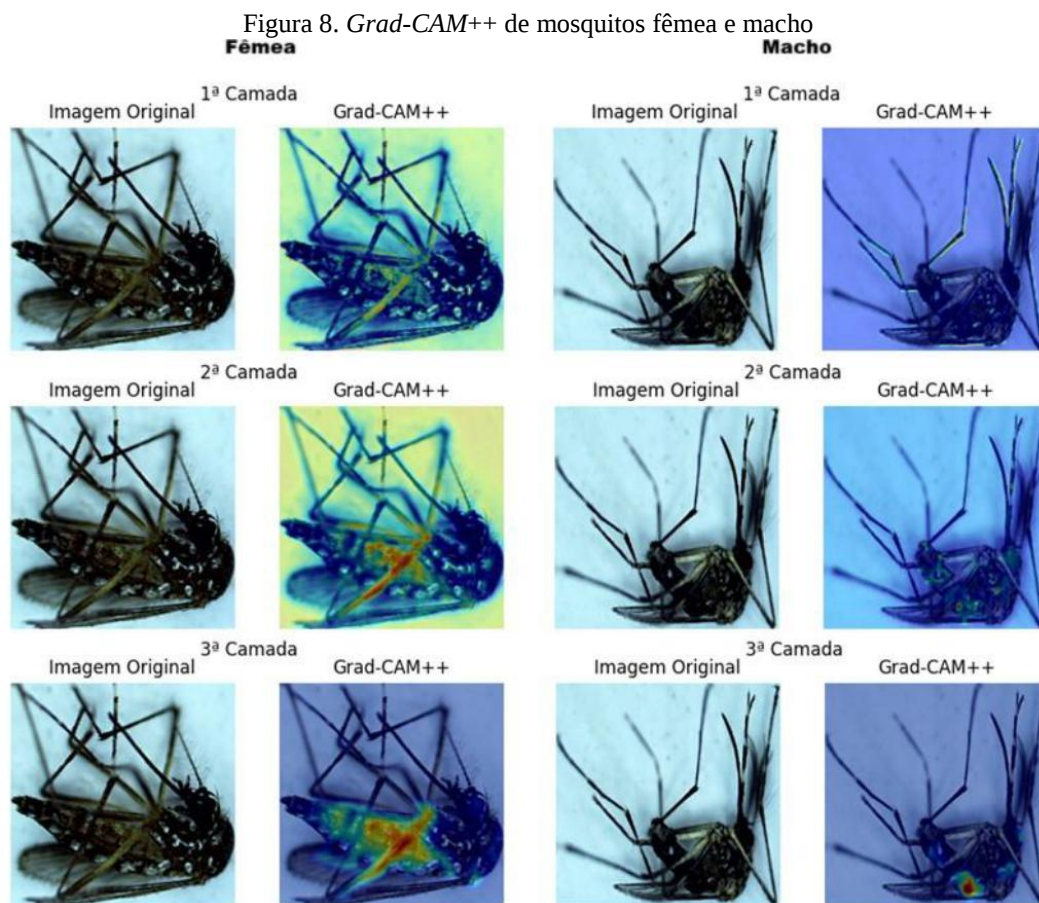
Considerando que o *Grad-CAM* não revelou alinhamento entre as pistas de sexagem aprendidas pela CNN e as usadas por especialistas, a técnica de XAI *Grad-CAM++* foi explorada. Diferentemente do protocolo com *Grad-CAM*, ele viabilizou a visualização consistente de um conjunto mais amplo de camadas (iniciais, intermediárias e finais), permitindo contrastar pistas de baixo e alto nível. Para assegurar comparação direta, manteve-se o mesmo pré-processamento. Os mapas correspondentes estão apresentados na Figura 8.

Nos resultados para fêmeas, observa-se que, na primeira camada da convolução, a ativação é difusa sobre o corpo do mosquito, com foco fraco e distribuído nas asas e pernas, coerente com a extração inicial de contornos e texturas gerais; na segunda camada da convolução, a saliência concentra-se mais no abdômen e na porção central do corpo, sobretudo em regiões pigmentadas, indicando o início da distinção de padrões morfológicos característicos; por fim, na terceira camada da convolução, a ativação torna-se fortemente concentrada em abdômen e tórax.

Para machos, na primeira camada convolucional, há foco inicial em pernas e palpos, sugerindo atenção a estruturas externas finas e captação de bordas e formas alongadas; na segunda camada da convolução, as ativações tornam-se mais distribuídas, com maior intensidade na cabeça e nas junções das pernas, surgindo indícios de reconhecimento de simetria e textura; na terceira camada convolucional, a saliência converge intensamente para a base do tórax.

Em síntese, os mapas de *Grad-CAM++* preservam o desalinhamento já apontado pelo *Grad-CAM* em relação aos marcadores diagnósticos de sexagem — plumosidade antenal e

morfologia dos palpos. Em fêmeas, a progressão das camadas leva de padrões difusos à ênfase em abdômen/tórax, deixando antenas e palpos em segundo plano; em machos, apesar de haver um sinal promissor de atenção aos palpos na camada inicial, a saliência migra nas camadas subsequentes para o tórax, reduzindo a evidência direta nas estruturas cefálicas mais discriminativas. Assim, embora o *Grad-CAM++* ofereça maior pontualidade e revele a trajetória de atenção ao longo da hierarquia, ele reforça a conclusão de que a rede privilegia padrões genéricos em tórax/abdômen.



Fonte: Elaborado pelos autores.

Apesar da acurácia global de 0,81, as visualizações (*Grad-CAM/Grad-CAM++*) e o padrão de erros indicam que o modelo atinge esse desempenho sobretudo explorando pistas genéricas no tórax/abdômen e a proporção corporal — as fêmeas tendem a ser maiores, e a rede aprendeu esse atalho de tamanho/textura em vez dos marcadores diagnósticos (plumosidade antenal e morfologia dos palpos).

Diante disso, a linha de ação mais promissora passa por pré-processamento orientado à região cefálica. Em termos práticos: (i) aplicar recorte/*region of interest* (ROI) automático da cabeça (via detector/segmentador leve ou heurísticas de contorno) antes da classificação; (ii) padronizar escala, orientação e iluminação para maximizar a nitidez de antenas e palpos; (iii) empregar *data augmentation* que preserve a cabeça (p. ex., proibir *crops* que removam a região cefálica e limitar rotações que a expulsem do quadro); e (iv) considerar ablação sistemática (*masking*) de tórax/abdome durante o treino para reduzir a dependência de pistas inespecíficas. Em conjunto, essas intervenções não negam a utilidade de CNNs de corpo inteiro, mas reorientam o sinal de treinamento para que o modelo aprenda, de fato, os marcadores utilizados pelos especialistas.

CONSIDERAÇÕES FINAIS

Este trabalho partiu da necessidade de apoiar a vigilância entomológica com métodos automáticos para distinguir, com rapidez e confiabilidade, o sexo de *Aedes aegypti*. Revisitando a pergunta de pesquisa à luz dos resultados: ao treinar CNNs com imagens de corpo inteiro, as decisões convergem para os marcadores cefálicos utilizados por especialistas (plumosidade antenal e morfologia dos palpos)?

No plano quantitativo, a CNN alcançou acurácia de 0,81 em teste, com assimetria de erros concentrada em fêmea → macho. A análise de interpretabilidade com *Grad-CAM* e, de forma mais abrangente, com *Grad-CAM++* — que permite inspecionar múltiplas camadas (iniciais, intermediárias e finais) — indicou um padrão consistente de saliência em tórax/mesonoto e porções do abdome, com atenção cefálica secundária dirigida principalmente a olho composto e probóscide. Antenas e palpos raramente emergem como regiões quentes. Em fêmeas, a progressão de camadas no *Grad-CAM++* vai de mapas difusos para ênfase marcada em abdômen/tórax; em machos, sinais iniciais em estruturas finas (palpos/pernas) migram nas camadas subsequentes para o tórax. Assim, embora o *Grad-CAM++* forneça mapas mais pontuais e revele a trajetória de atenção ao longo da hierarquia, ele corrobora o desalinhamento entre as pistas usadas pela rede e os critérios diagnósticos de especialistas.

Os resultados sugerem que a rede sustenta parte do desempenho explorando pistas genéricas de textura/contorno em tórax/abdômen e proporção corporal (com fêmeas frequentemente maiores nas imagens), isto é, atalhos que não traduzem os marcadores

morfológicos relevantes. Tal estratégia é operacionalmente inadequada, pois é sensível a variações de escala, pose e iluminação, e não garante que o sistema esteja de fato “olhando” para as estruturas diagnósticas.

Diante desse quadro, os direcionamentos para aproximar a decisão automática dos critérios periciais são:

- a) Pré-processamento orientado à cabeça (ROI): detectar/recortar automaticamente a região cefálica (antenas, palpos, probóscide) antes da classificação;
- b) Padronização de escala e pose: normalizar tamanho e orientação (aproveitando a referência quadriculada) para neutralizar o uso de proporção/tamanho como proxy de sexo;
- c) *Data augmentation* conservadora da cabeça: impedir crops que excluam a região cefálica e limitar rotações que a expulsem do quadro;
- d) Ablação dirigida: mascarar sistematicamente tórax/abdômen em parte do treino para desincentivar a dependência de pistas inespecíficas;
- e) Supervisão de atenção (quando viável): usar máscaras/caixas fracas da região antenas-palpos para regularizar a perda (e.g., penalizando saliência dominante fora da cabeça);
- f) XAI complementar e métricas de alinhamento: combinar *Grad-CAM/Grad-CAM++* com occlusion/RISE, LIME/SHAP por superpixels e quantificar o alinhamento.

Por fim, reconhece-se limitações: (i) dependência do protocolo único de aquisição e enquadramento de corpo inteiro, que pode subexpor antenas/palpos; (ii) resolução espacial limitada dos mapas quando a decisão depende de estruturas finas. Essas limitações reforçam a utilidade do *Grad-CAM++* para inspecionar camadas múltiplas, mas também a necessidade de coletar anotações leves da cabeça e de validar em domínios distintos (outros dispositivos/iluminações).

A CNN de corpo inteiro demonstrou desempenho razoável, porém não fundamentado nos marcadores cefálicos de referência. A adoção das intervenções propostas — especialmente o pipeline em duas etapas com ROI cefálica, a normalização de escala/pose e a regularização de atenção — constitui o próximo passo para reorientar o aprendizado às estruturas diagnósticas (antenas e palpos), reduzir a dependência de pistas genéricas em tórax/abdômen e aumentar a robustez e a confiabilidade do sistema em cenários operacionais de vigilância entomológica.

REFERÊNCIAS

AL MARUF, A., *et al.* (2025). TransembleNet: Enhancing vector mosquito species classification through transfer learning-based ensemble model. **PLOS One**, 20(5), e0322171. Disponível em: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0322171>. Acesso em: 13 jan. 2025 [i].

ANDREW, J., & BAR, A. (2013). Morphology and Morphometry of *Aedes aegypti* Adult Mosquito. **Annual Research & Review in Biology**, 3(1), 52–69. Disponível em: <https://journalarrb.com/index.php/ARRB/article/view/1903>. Acesso em: 17 dez. 2024 [i].

AZAM, F. B., *et al.* (2023). Classifying stages in the gonotrophic cycle of mosquitoes using deep learning. **Parasites & Vectors**, 16, art. 336. Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC10719391/>. Acesso em: 18 dez. 2024 [i].

BURKETT-CADENA, N. (2013). **Morphology of Adult and Larval Mosquitoes**. University of Florida/IFAS. Disponível em: https://fmel.ifas.ufl.edu/media/fmelifasufledu/workshop/Mosquito_Morphology.pdf. Acesso em: 5 out. 2024 [i].

CANNET, A., *et al.* (2023). Deep learning and wing interferential patterns identify *Anopheles* species and discriminate amongst *Gambiae* complex species. **Scientific Reports**, 13(1), 13895. Disponível em: <https://www.nature.com/articles/s41598-023-41114-4>. Acesso em: 27 jan. 2025 [i].

ECDC – EUROPEAN CENTRE FOR DISEASE PREVENTION AND CONTROL. (2023). **Aedes aegypti—Factsheet for experts**. Disponível em: <https://www.ecdc.europa.eu/en/disease-vectors/facts/mosquito-factsheets/aedes-aegypti>. Acesso em: 22 set. 2024 [i].

GOMES, A. N. B., *et al.* (2025). Identificação automatizada da fêmea do *Aedes aegypti* com redes neurais convolucionais: uma abordagem para o controle epidemiológico da dengue. **Revista Caderno Pedagógico**, Curitiba, 22(7), 1–20. DOI: 10.54033/cadpedv22n7-100. Disponível em: <https://ojs.studiespublicacoes.com.br/ojs/index.php/cadped/article/view/16208/9034>. Acesso em: 6 mai. 2025 [i].

GONZÁLEZ-PÉREZ, A., OLIVER-VILLANUEVA, J. V., ALCARAZ, R., GARCÍA-MARTÍNEZ, J. M., FERNÁNDEZ-DÍAZ, M., HERNÁNDEZ-CAMACHO, R., LÓPEZ-LÓPEZ, J. A., & MARTÍNEZ-ÁLVAREZ, F. (2022). Classification of mosquitoes of the *Aedes* and *Culex* genera and determination of their sex by optical sensors and machine learning. **Parasites & Vectors**, 15(1), 1-14. DOI: 10.1186/s13071-022-05324-5. Disponível em: <https://parasitesandvectors.biomedcentral.com/articles/10.1186/s13071-022-05324-5>. Acesso em: 14 dez. 2024 [i].

INSTITUTO AGGEU MAGALHÃES. Fundação Oswaldo Cruz. (s.d.). **Instituto Aggeu Magalhães**. Disponível em: <https://www.cpqam.fiocruz.br>. Acesso em: 23 set. 2025 [i].

JÚNIOR, G. P. C. (2022). **Reconhecimento de espécies de mosquitos por meio de métodos computacionais inteligentes baseados em aprendizagem profunda** [Dissertação de Mestrado não publicada]. Universidade Federal de Pernambuco, Recife. Disponível em: <https://repositorio.ufpe.br/handle/123456789/46727>. Acesso em: 15 mai. 2025 [i].

KITTICHA, V., *et al.* (2021). Deep learning approaches for challenging species and gender identification of mosquito vectors. **Scientific Reports**, 11, art. 4838. DOI: 10.1038/s41598-021-84219-4. Disponível em: <https://www.nature.com/articles/s41598-021-84219-4>. Acesso em: 2 nov. 2024 [i].

LI, D., *et al.* (2024). Towards transforming malaria vector surveillance using VectorBrain: a novel convolutional neural network for mosquito species, sex, and abdomen status identifications. **Scientific Reports**, 14, art. 23647. DOI: 10.1038/s41598-024-71856-8. Disponível em: <https://www.nature.com/articles/s41598-024-71856-8>. Acesso em: 8 fev. 2025 [i].

LI, J., *et al.* (2022). Neuroanatomical basis of sexual dimorphism in the brain of two disease vector mosquitoes. **iScience**, 25(10), 105113. Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC9583127/>. Acesso em: 12 out. 2024 [i].

PINHEIRO, P. (2025, 15 jun.). Aedes aegypti (fotos): como é o mosquito da dengue? **MD.Saúde**. Disponível em: <https://www.mdsaude.com/doencas-infecciosas/fotos-mosquito-dengue/>. Acesso em: 17 out. 2024 [i].

SAUER, F. G., *et al.* (2024). A convolutional neural network to identify mosquito species from wing images. **Scientific Reports**, 14, art. 1454. Disponível em: <https://www.nature.com/articles/s41598-024-53631-x>. Acesso em: 25 nov. 2024 [i].

SELVARAJU, R. R., *et al.* (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. In: **PROCEEDINGS OF THE IEEE INTERNATIONAL CONFERENCE ON COMPUTER VISION (ICCV)** (pp. 618–626). IEEE. Disponível em: https://openaccess.thecvf.com/content_ICCV_2017/papers/Selvaraju_Grad-CAM_Visual_Explanations_ICCV_2017_paper.pdf. Acesso em: 20 out. 2024 [i].

SUPRATAK, A., *et al.* (2024). MosquitoSong+: A noise-robust deep learning model for mosquito species and sex classification from raw wingbeat sounds. **PLOS ONE**, 19(6), e0310121. Disponível em: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0310121>. Acesso em: 5 dez. 2024 [i].

WHO – WORLD HEALTH ORGANIZATION. (2024). **Dengue and severe dengue**. Disponível em: <https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue>. Acesso em: 15 set. 2024 [i].

ZHAO, D., *et al.* (2022). A Swin Transformer-based model for mosquito species identification. **Scientific Reports**, 12, art. 18664. DOI: 10.1038/s41598-022-21017-6. Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC9636261/>. Acesso em: 15 nov. 2024 [i].