



Universidade de Brasília

Núcleo de Medicina Tropical - NMT

Programa de Pós-graduação em Medicina Tropical - PPGMT

TAÍS OLIVEIRA DE ARAÚJO

**APLICAÇÃO DA INTELIGÊNCIA ARTIFICIAL PARA IDENTIFICAÇÃO DE
MOSQUITOS VETORES DA FEBRE AMARELA**

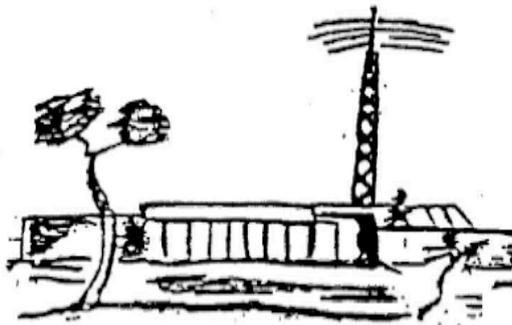
Brasília/DF

2025

TAÍS OLIVEIRA DE ARAÚJO

**APLICAÇÃO DA INTELIGÊNCIA ARTIFICIAL PARA IDENTIFICAÇÃO DE
MOSQUITOS VETORES DA FEBRE AMARELA**

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Medicina Tropical da Universidade de Brasília, para a obtenção do título de Mestre em Medicina Tropical, na área de concentração: Biologia das doenças infecciosas e parasitárias.



Brasília/DF

2025

OA663a Oliveira de Araújo, Taís
APLICAÇÃO DA INTELIGÊNCIA ARTIFICIAL PARA IDENTIFICAÇÃO
DE MOSQUITOS VETORES DA FEBRE AMARELA / Taís Oliveira de
Araújo; orientador Rodrigo Gurgel-Gonçalves . --
Brasília, 2025.
79 p.

Dissertação (Mestrado em Medicina Tropical) --
Universidade de Brasília, 2025.

1. Aprendizado profundo. 2. Aprendizado de máquina. 3.
Identificação automatizada. 4. Culicidae. 5. Febre amarela.
I. Gurgel-Gonçalves , Rodrigo , orient. II. Título.

TAÍS OLIVEIRA DE ARAÚJO

**APLICAÇÃO DA INTELIGÊNCIA ARTIFICIAL PARA IDENTIFICAÇÃO DE
MOSQUITOS VETORES DA FEBRE AMARELA**

Banca examinadora

Prof. Dr. Rodrigo Gurgel Gonçalves (Orientador)

Universidade de Brasília

Prof. Dr. Rafael de Oliveira Christe (Membro externo)

Faculdade de Saúde Pública da Universidade de São Paulo

Prof. Dr. Filipe Vieira Santos de Abreu (Membro externo)

Instituto Federal do Norte de Minas Gerais

Prof. Dr. Marcos Takashi Obara (Membro Interno - Suplente)

Universidade de Brasília

Dedico esta dissertação aos
meus pais, Alcina e Manoel.

AGRADECIMENTOS

Agradeço primeiramente aos meus pais pelo amor incondicional, pela paciência e pelo incentivo constante, especialmente nos momentos mais desafiadores. Sem vocês, nada disso seria possível.

Aos meus amigos Paulo, Nathali, Giulieta, Anna e João, sou grata por estarem sempre ao meu lado, celebrando cada conquista e compartilhando esta jornada comigo.

Ao meu orientador, Prof. Dr. Rodrigo Gurgel Gonçalves, registro minha profunda gratidão pela paciência, orientação e dedicação ao longo de toda esta pesquisa. Sua presença foi fundamental desde a graduação até a pós-graduação, guiando meus passos e contribuindo enormemente para o meu crescimento acadêmico e pessoal.

Ao Prof. Dr. Marcos Takashi Obara, pelos ensinamentos e conselhos que foram indispensáveis nesta trajetória.

Ao Prof. Dr. Vinícius Lima, sou imensamente grata pelo conhecimento compartilhado e pelas boas risadas que tornaram esta jornada mais leve.

Aos colegas de laboratório Fabricio, Isadora, Karine e Elias, obrigada pelas trocas de ideias, pelo apoio mútuo e pelos momentos de descontração que tornaram este percurso mais significativo e alegre.

Agradeço também àquelas pessoas que passaram brevemente pela minha vida, mas que, em momentos cruciais, me ofereceram apoio e palavras de incentivo.

À Universidade de Brasília (UnB) e ao Programa de Pós-Graduação em Medicina Tropical, pelo suporte acadêmico e pela oportunidade de desenvolver este projeto.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela bolsa, e ao auxílio PROAP/CAPES, por todo o apoio financeiro que contribuiu para o desenvolvimento deste trabalho.

A todos vocês, minha mais sincera gratidão.

LISTA DE TABELAS

Tabela 1 - Estudos que aplicaram CNNs para a identificação de espécies de mosquitos, destacando as redes utilizadas, as espécies identificadas e médias gerais de acurácia.....	37
---	----

LISTA DE FIGURAS

Figura 1 - Ciclos de transmissão do vírus da febre amarela na América do Sul. Fonte: CDC	15
Figura 2 - Disseminação do vírus da febre amarela nos biomas brasileiros. Linha tracejada preta: propagação viral do final de 1980 até 2010; linha branca: disseminação viral da metade de 2010 até 2018. Fonte: POSSAS et al., 2018.	16
Figura 3 - Filogenia de gêneros de mosquitos com base na morfologia segundo Harbach & Kitching, 1998 (Harbach,2007)	19
Figura 4 - Aspectos gerais de <i>Haemagogous leucocelaenus</i>	23
Figura 5- Aspectos gerais de <i>Sabethes albiprivus</i>	23
Figura 6 - Aspectos gerais de <i>Aedes scapularis</i>	24
Figura 7- Aspectos gerais <i>Aedes serratus</i>	24
Figura 8 - Primeira chave taxonômica para identificação de mosquitos, A Monograph of the Culicidae, Or Mosquitoes de Frederick Vincent Theobald... 25	
Figura 9 - Bibliografias utilizadas para identificação de mosquitos a partir de chaves dicotômicas. A) Neotropical Culicidae (LANE., 1953); B) Culicidologia Médica, Volume 2 (FORATTINI., 2002); C) Principais Mosquitos de Importância Sanitária no Brasil (CONSOLI & LOURENÇO., 1994).	26
Figura 10 - Chave eletrônica online na plataforma Xper para identificação de fêmeas do gênero <i>Sabethes</i> . Fonte: NEVES et al., 2024.	28
Figura 11 - Esquema ilustrativo comparativo entre redes neurais biológicas e redes neurais artificiais. Fonte: Wäldchen & Mäde., 2018.	30
Figura 12 - Arquitetura de uma rede de detecção de objetos.1) Imagem de entrada; 2) reconhecimento de objetos e isolamento do mesmo; 3) Extração das características de cada objeto, 4) classificação de cada objeto. Fonte: AZIZ et al., 2020	32
Figura 13 - Bee Machine: Aplicativo móvel para identificação de abelhas da América do Norte. Fonte: SPIESMAN et al., 2021. (https://beemachine.ai/)	33
Figura 14 . IDX um microscópio digital de captura imagens e identificação automatizada de mosquitos e carrapatos. Fonte: Gupta et al 2024	40

Figura 15 - Exemplos de imagens usadas para treinar a rede AlexNet. Mosquitos fotografados com estereomicroscópio e celular. (A) <i>Aedes scapularis</i> , (B) <i>Aedes serratus</i> (C) <i>Haemagogus leucocelaenus</i> , (D) <i>Sabethes albiprivus</i>	44
Figura 16 - Matrizes de confusão ilustrando os acertos e erros de classificação para os 10% das imagens utilizadas no conjunto de teste em cada um dos quatro experimentos. Para cada experimento foi realizado 10 pseudo-replicatas.	48
Figura 17 - Acurácia média geral (pontos) e intervalos de confiança de 95% (barras) do desempenho da rede AlexNet para identificação de mosquitos vetores da febre amarela.....	49
Figura 18 - Acurácia média do algoritmo AlexNet na identificação interespecífica de espécies. Os experimentos levaram em consideração a identificação a partir de imagens de regiões isoladas de partes do corpo, com intervalos de confiança de 95%	50
Figura 19- Exemplos de imagens identificadas incorretamente, devido a espécimes danificados. (AC e H) <i>Ae. serratus</i> foi confundido com <i>Ae. scapularis</i> . (D e E) <i>Ae. scapularis</i> foi confundido com <i>Ae. serratus</i> . (F) <i>Ae. serratus</i> foi confundido com <i>Sa. albiprivus</i> . (G) <i>Sa. albiprivus</i> foi confundido com <i>Hg. leucocelaenus</i> . (I) <i>Ae. scapularis</i> foi identificado erroneamente como <i>Hg. leucocelaenus</i>	51

LISTA DE ABREVIATURAS E SIGLAS UTILIZADAS

CNN's – Convolutional Neural Networks

IA – Inteligência artificial

IC – Intervalo de confiança

DL – Deep learning

FA – Febre amarela

ML – Machine Learning

PNH – Primatas não humanos

RGB – Vermelho-verde-azul

ReLU – Rectified Linear Unit (em português, unidade linear retificada)

YFV - Yellow fever virus (em português, vírus da febre amarela)

FINANCIAMENTO

Apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) através da bolsa de estudo para auxílio nas atividades envolvidas no mestrado. Além do PROAP/CAPES, auxílio financeiro a discentes de pós-graduação para o desenvolvimento de dissertações e de teses.

SUMÁRIO

1 INTRODUÇÃO	15
1.1 Febre Amarela no Brasil	15
1.2 Mosquitos.....	18
1.2.1 Características gerais	18
1.2.2 Relações Filogenéticas.....	19
1.2.3 Bionomia e importância médica	19
1.2.4 Mosquitos vetores da febre amarela.....	21
1.3 Identificação de culicídeos.....	25
1.4 Aprendizado de máquina	29
1.5 Identificação automatizada de insetos e outros invertebrados	32
1.6 Identificação automatizada de Culicídeos.....	36
2 JUSTIFICATIVA	41
3 OBJETIVO	42
3.1 Objetivos específicos	42
4 METODOLOGIA	42
4.1 Banco de imagens	42
4.2 Algoritmo	45
4.3 Treinamento e teste do algoritmo.....	46
4.4 Análise de dados	47
5 RESULTADOS	48
5.1 Matriz de confusão	48
5.2 Acurácia geral	49
5.3 Acurácia específica	50
6 DISCUSSÃO	52
7 CONCLUSÃO	56
REFERÊNCIAS	57
APÊNDICES – acurácias observadas para cada pseudo-replicata	68
ANEXO A – Artigo Publicado	71

RESUMO

A identificação de mosquitos vetores é fundamental para o controle de doenças. Estudos de identificação automatizada usando Redes Neurais Convolucionais (CNNs) já foram realizados para alguns mosquitos vetores urbanos, mas ainda não para mosquitos silvestres que transmitem a febre amarela e outras arbovíroses. Avaliamos a capacidade da rede CNN AlexNet de identificar quatro espécies de mosquitos: *Aedes serratus*, *Aedes scapularis*, *Haemagogus leucocelaenus* e *Sabethes albiprivus* e se há variação na capacidade da AlexNet de classificar mosquitos com base em fotos de quatro regiões diferentes do corpo. Os espécimes foram fotografados usando um telefone celular conectado a um estereoscópio. Foram tiradas fotografias de corpo inteiro, pronoto e vista lateral do tórax, que foram pré-processadas para treinar o algoritmo AlexNet. A avaliação foi baseada na matriz de confusão, na média de acurácia (10 pseudo-replicatas) e no intervalo de confiança de cada experimento. Nosso estudo constatou que o AlexNet pode identificar com precisão imagens de mosquitos dos gêneros *Aedes*, *Sabethes* e *Haemagogus* com mais de 90% de precisão. Além disso, o desempenho do algoritmo não mudou de acordo com as regiões do corpo apresentadas. É importante observar que o estado de preservação dos mosquitos, que muitas vezes estavam danificados, pode ter afetado a capacidade da rede de diferenciar essas espécies e, portanto, as taxas de precisão poderiam ter sido ainda maiores. Nossos resultados apoiam a ideia de aplicar CNNs orientadas por IA para identificação automatizada de mosquitos vetores de doenças tropicais. Essa abordagem pode ser usada na vigilância de vetores da febre amarela pelos serviços de saúde e na vigilância de base comunitária.

Palavras-chaves: Aprendizado profundo, Aprendizado de máquina, Culicidae

ABSTRACT

Identifying mosquito vectors is crucial for controlling diseases. Automated identification studies using Convolutional Neural Network (CNN) have been conducted for some urban mosquito vectors but not yet for sylvatic mosquito vectors that transmit the yellow fever and other arboviruses. We evaluated the ability of the AlexNet CNN to identify four mosquito species: *Aedes serratus*, *Aedes scapularis*, *Haemagogus leucocelaenus* and *Sabethes albiprivus* and whether there is variation in AlexNet's ability to classify mosquitoes based on pictures of four different body regions. The specimens were photographed using a cell phone connected to a stereoscope. Photographs were taken of the full-body, pronotum and lateral view of the thorax, which were pre-processed to train the AlexNet algorithm. The evaluation was based on the confusion matrix, the accuracy (10 pseudo-replicates) and the confidence interval for each experiment. Our study found that the AlexNet can accurately identify mosquito pictures of the genus *Aedes*, *Sabethes* and *Haemagogus* with over 90% accuracy. Furthermore, the algorithm performance did not change according to the body regions submitted. It is worth noting that the state of preservation of the mosquitoes, which were often damaged, may have affected the network's ability to differentiate between these species and thus accuracy rates could have been even higher. Our results support the idea of applying CNNs for AI-driven identification of mosquito vectors of tropical diseases. This approach can potentially be used in the surveillance of yellow fever vectors by health services and the population as well.

Keywords: Deep Learning, Machine learning, Culicidae

1 INTRODUÇÃO

1.1 Febre Amarela no Brasil

A Febre Amarela (FA) é uma doença infecciosa, não contagiosa, originária da África, cujo agente infeccioso é o arbovírus YFV (*Yellow Fever Virus*), pertencente ao gênero *Orthoflavivirus*, da família *Flaviviridae*, sendo transmitida ao homem por meio da picada de mosquitos hematófagos da família *Culicidae*. Nas américas o YFV mantém-se em dois ciclos: o ciclo urbano, transmitido pelo mosquito *Aedes aegypti*, e o ciclo silvestre, transmitido principalmente por mosquitos dos gêneros *Haemagogus* e *Sabethes* (VASCONCELOS et al., 2001, 2002) (Figura 1). Em humanos, a FA é uma doença aguda grave, caracterizada por febre, náuseas, vômitos, dor epigástrica, hepatite com icterícia, insuficiência renal e hemorragia, levando à morte em 20 a 60% dos casos (GIOVANETTI et al., 2019; TUELLS et al., 2022).

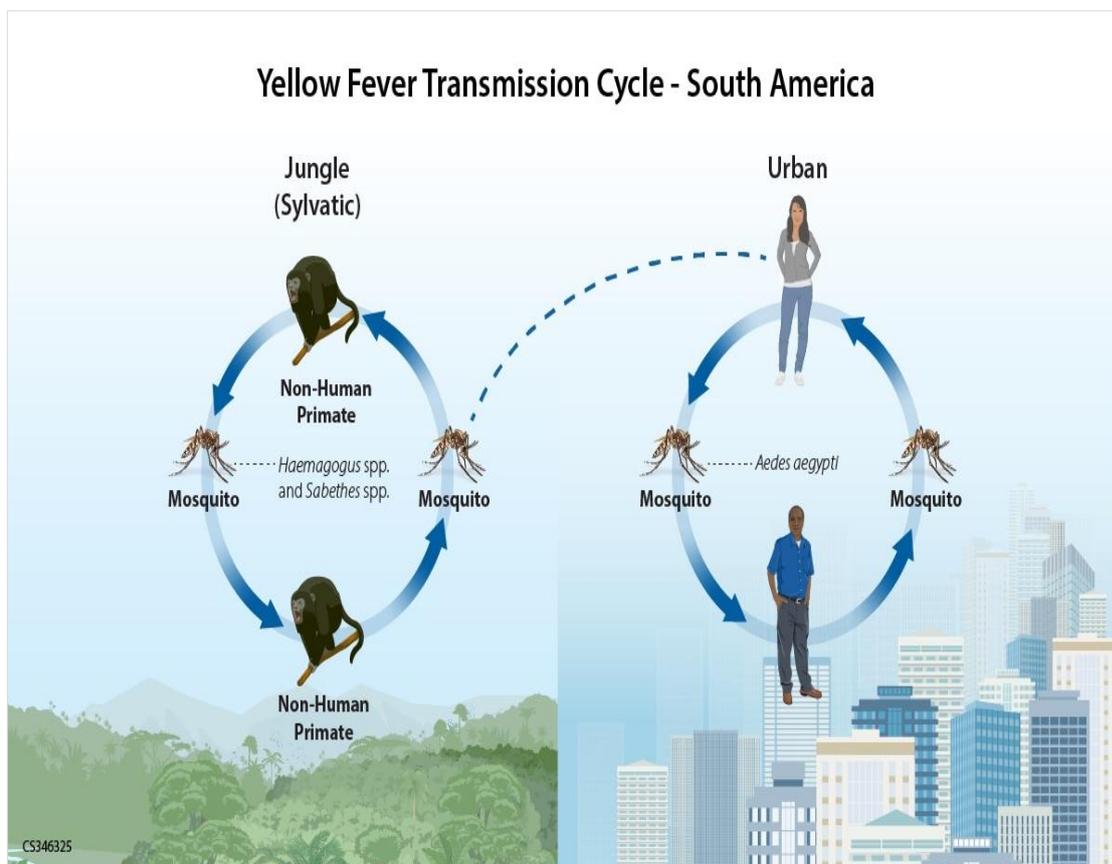


Figura 1- Ciclos de transmissão do vírus da febre amarela na América do Sul. Fonte: CDC

A FA é endêmica em regiões tropicais e foi responsável por epidemias na África, Europa e nas Américas do Norte e Sul, durante o século XVII até o início do século XX. A primeira epidemia de FA descrita no Brasil ocorreu em 1685, em Recife, onde o vírus teria sido levado por barcos procedentes da África, permanecendo na região por pelo menos dez anos. No ano seguinte, atingiu Salvador, período em que cerca de 25 mil pessoas adoeceram e 900 evoluíram para óbito, com relatos de casos até 1692. Os últimos casos urbanos foram reportados no município de Sena Madureira, no estado do Acre, em 1942. Desde então, não há registro do ciclo urbano de FA no Brasil (POSSAS et al., 2018)

Nas últimas décadas, tem-se acompanhado ressurgimentos periódicos da FA em áreas da África Subsaariana e América do Sul tropical, incluindo o Brasil, sendo mais frequente na Região Amazônica. Após o último caso de FA urbana em 1942, o vírus YFV passou a circular predominantemente entre primatas não humanos (PNH) na Amazônia. No entanto, a partir de 2000, a doença expandiu para regiões do Sul e Sudeste do Brasil, onde não havia registros anteriores (GIOVANETTI et al., 2023; GARCIA-OLIVEIRA et al., 2023) (Figura 2).

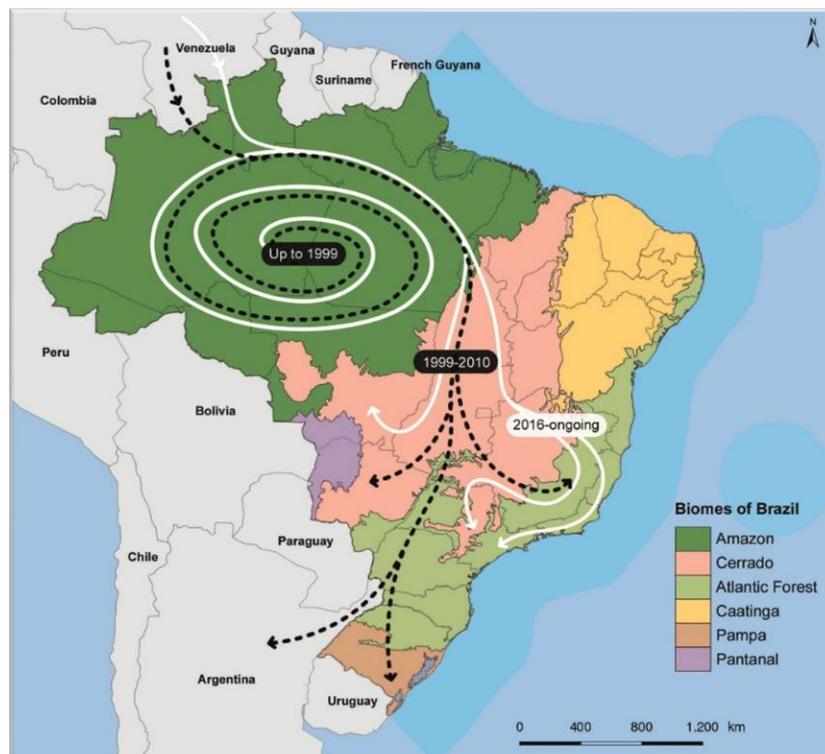


Figura 2 - Disseminação do vírus da febre amarela nos biomas brasileiros. Linha tracejada preta: propagação viral do final de 1980 até 2010; linha branca: disseminação viral da metade de 2010 até 2018. Fonte: POSSAS et al., 2018.

O período de 2016-2018 marcou um dos maiores surtos de FA no Brasil, com os primeiros casos registrados na Mata Atlântica. Em apenas um ano, a doença se espalhou por diversos estados do Sudeste, atingindo áreas que estavam livres da FA há cerca de 80 anos (ABREU et al., 2019, 2022). Entre 2016 e 2018, milhares de casos foram registrados em PNH e humanos, com um número significativo de óbitos. Minas Gerais foi o estado mais afetado. Apesar da redução nos casos em PNH nesse período, o número de casos humanos aumentou. No entanto, a partir de 2018, houve uma diminuição nos casos, tanto em PNH quanto em humanos. Todos os casos humanos foram associados à transmissão silvestre (SILVA et al., 2020; SACCHETTO et al., 2020; GARCIA-OLIVEIRA et al., 2023).

Após esses eventos, em 2019, a FA continuou a se espalhar para o sul, chegando aos estados do Paraná e Santa Catarina, causando casos humanos e centenas de epizootias em PNH. No final de 2020 e início de 2021, o vírus foi detectado causando mortes de PNH no Rio Grande do Sul (ANDRADE et al., 2021). No mesmo período, uma nova reemergência da FA foi observada nos estados de Goiás, Distrito Federal e sudeste e norte de Minas Gerais (ANDRADE et al., 2022; OLIVEIRA et al., 2023; GIOVANETTI et al., 2023). Em 2024, o Brasil registrou quatro casos confirmados de febre amarela, sendo três fatais, nos estados de Roraima, Amazonas e São Paulo. Em janeiro de 2025, foram confirmados 11 casos humanos de febre amarela silvestre em São Paulo com local provável de infecção regiões rurais de Campinas e Bauru (SÃO PAULO.,2025). A vacinação é fundamental para evitar as infecções e óbitos por FA, principalmente para quem vive ou visita regiões com risco de transmissão (BRASIL., 2024). A vacina, oferecida gratuitamente pelo SUS, aliada à vigilância epidemiológica e ao controle de vetores, permanece como a principal estratégia para prevenir novos casos e surtos.

1.2 Mosquitos

Os mosquitos, pertencem à família Culicidae da ordem Diptera, são insetos de grande diversidade e abundância, encontrados em regiões tropicais e temperadas de todo o mundo, incluindo áreas além do Círculo Polar Ártico. Atualmente, são reconhecidas 3.727 espécies viventes (HARBACH; 2024) classificadas em duas subfamílias: Anophelinae, com três gêneros, e Culicinae, com 110 gêneros distribuídos em 11 tribos. Além disso, 31 espécies extintas já foram descritas, incluindo seis fósseis da subfamília extinta Burmaculicinae, do Cretáceo Superior, e 25 fósseis de gêneros existentes como *Anopheles*, *Culex*, *Mansonia* e *Toxorhynchites*, oriundos da Era Cenozóica. A subfamília Anophelinae inclui 525 espécies existentes formalmente reconhecidas. A subfamília é dividida em três gêneros: *Anopheles*, *Bironella* e *Chagasia*. Anophelinae inclui duas espécies fósseis, uma questionavelmente colocada em *Anopheles* e outra reconhecida como uma espécie do subgênero *Nyssorhynchus*. Os mosquitos pertencentes a esses gêneros são chamados de 'anofelinos'. Culicinae é a maior subfamília de mosquitos, contendo 3.202 espécies existentes em 110 gêneros e dois grupos de *incertae sedis*, '*Aedes*' *sensu auctorum*, '*Ochlerotatus*' *sensu auctorum*. Os gêneros são agrupados em 11 tribos. Culicinae também inclui 23 espécies fósseis, 18 em cinco gêneros existentes, cinco em três gêneros extintos. (HAWKES & HOPKINS., 2022; HARBACH, 2007, 2024)

1.2.1 Características gerais

Os mosquitos são insetos delgados e de pernas compridas que são facilmente reconhecidos por sua longa probóscide e pela presença de escamas em várias partes do corpo. As larvas se distinguem de outros insetos aquáticos pela presença de cabeça distinta com escovas bucais e antenas, tórax bulboso papilas anais posteriores e um par de aberturas respiratórias (subfamília Anophelinae) ou um sifão respiratório alongado (subfamília Culicinae) presente perto do final do abdômen. Os mosquitos são identificados, de forma confiável pelas suas formas adultas e por larvas de quarto instar (HARBACH, 2007, 2024).

1.2.2 Relações Filogenéticas

A família Culicidae é monofilética, mas as relações mais profundas não foram resolvidas. A subfamília Anophelinae é uma linhagem monofilética basal a todos os outros culicídeos, e o gênero *Chagasia* é uma linhagem monofilética basal a outros Anophelinae. A subfamília Culicinae não é comprovadamente monofilética em relação ao gênero *Toxorhynchites*. As tribos Aedini, Culicini e Sabethini são monofiléticas, mas a monofilia das outras tribos não foi testada e suas relações filéticas são incertas (Harbach, 2007) (Figura 3).

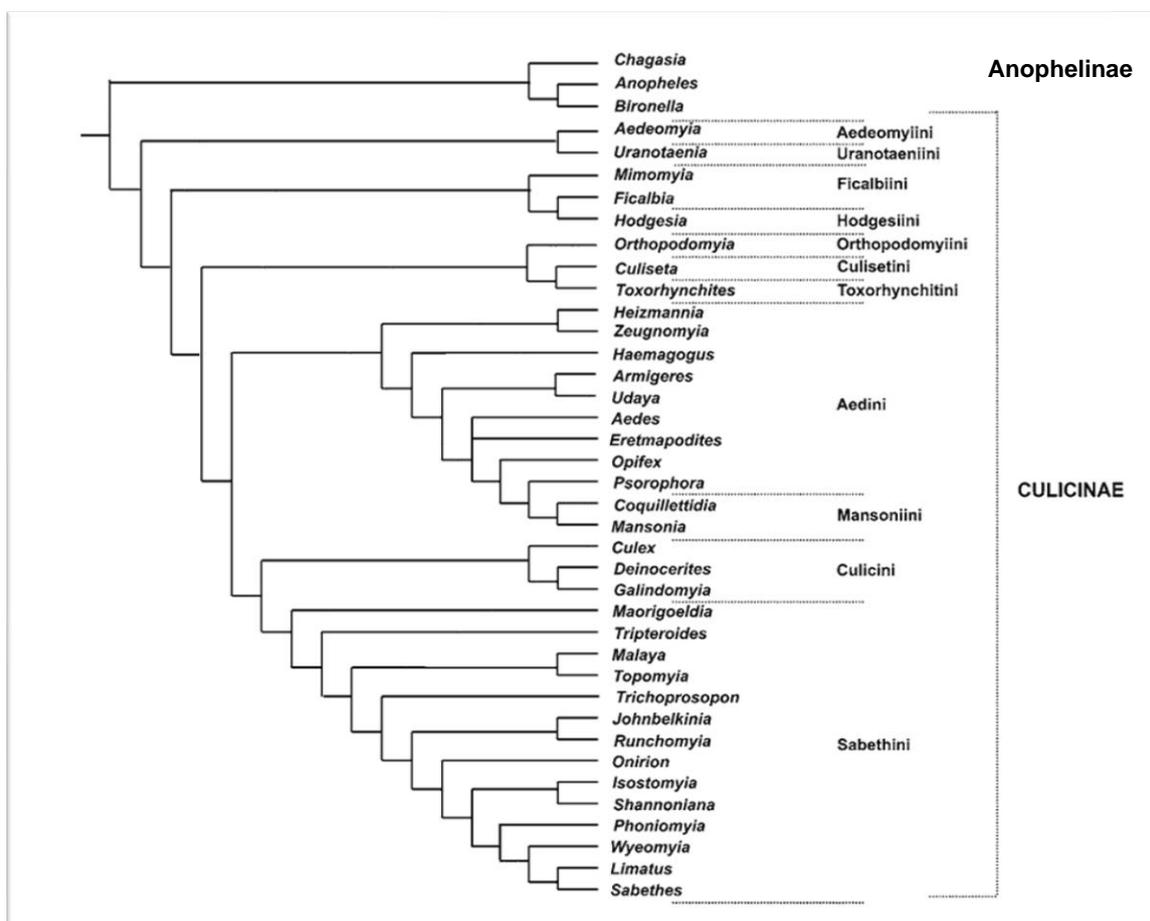


Figura 3 - Filogenia de gêneros de mosquitos com base na morfologia segundo HARBACH & KITCHING.,1998 (HARBACH.,2007)

1.2.3 Bionomia e importância médica

Os mosquitos são insetos holometábolos e suas larvas ocupam uma ampla gama de habitats aquáticos. A maioria das larvas se alimenta de material

particulado em suspensão e microrganismos que extraem da água com escovas bucais filamentosas. Outras larvas são predadoras obrigatórias ou facultativas que capturam e se alimentam principalmente dos estágios imaturos de outros mosquitos por meio de escovas bucais modificadas ou mandíbulas ou maxilas de preensão. Ocasionalmente, as larvas recorrem ao canibalismo quando a comida é escassa. Algumas espécies obtêm oxigênio diretamente da superfície da água, enquanto outras utilizam plantas aquáticas ou estruturas especializadas para respiração (HAWKES & HOPKINS., 2022; HARBACH., 2024).

Como os mosquitos são insetos delicados e são encontrados em regiões tropicais. Muitas espécies vivem a poucos metros do solo, enquanto algumas espécies florestais ocorrem principalmente no dossel da floresta. Os machos alimentam-se exclusivamente de líquidos vegetais. Já as fêmeas além da alimentação baseadas em líquidos vegetais, a maioria delas requerem sangue para maturação dos seus ovos (HARBACH., 2024; HAWKES & HOPKINS., 2022). Embora os vertebrados homeotérmicos sejam os hospedeiros mais comuns, algumas espécies alimentam-se de vertebrados pecilotérmicos, como anfíbios e répteis, os quais são importantes fontes de sangue para muitos mosquitos dos gêneros *Mimomyia*, *Uranotaenia* e *Deinocerites* (BORKENT & BELTON., 2006; TAMASHIRO et al., 2011; CAMP et al., 2018). Além de invertebrados como minhocas e sanguessugas (REEVES., 2018). O tempo de voo e a atividade alimentar são específicos para a maioria das espécies. Algumas espécies são ativas à noite ou durante o crepúsculo, enquanto outras são ativas durante o dia. Os culicídeos podem ser classificados como: diurnos, crepusculares ou noturnos (SANT'ANA & LOZOVEI., 2001; GOMES et al., 2010; BONA; NAVARRO-SILVA, 2008)

A reprodução dos mosquitos envolve uma variedade de comportamentos importantes para o sucesso reprodutivo. Geralmente, o acasalamento ocorre em enxames, onde machos e fêmeas se reúnem no ar. Essas interações são breves e podem resultar em acasalamento bem-sucedido ou na rejeição pela fêmea. Os machos localizam as fêmeas utilizando a frequência de batimentos da asa das mesmas. Os sistemas de acasalamento são, em sua maioria, políginos,

permitindo que os machos copulem com várias fêmeas, sendo comum a competição entre eles (YUVAL.,2008; CATOR, WYER, HARRINGTON., 2021). Além disso, as fêmeas podem demonstrar preferência na escolha do parceiro, influenciando quais machos têm sucesso no acasalamento, a maioria das fêmeas acasalam apenas uma vez durante sua vida, armazenando o esperma do macho para fertilizar todos os óvulos subsequentes que produzirem (CATOR, WYER, HARRINGTON., 2021). As fêmeas dos mosquitos precisam de sangue para nutrir e desenvolver seus ovos. A relação entre o tipo de sangue do hospedeiro e a reprodução do mosquito é complexa e não é totalmente compreendida. Vários constituintes do sangue, como proteínas e aminoácidos, desempenham papéis importantes nesse processo. Por exemplo, proteínas como a albumina sérica são essenciais para o desenvolvimento do ovo e seus efeitos podem variar dependendo da espécie hospedeira da qual o sangue é derivado (DE SWART et al., 2023).

A dependência da alimentação sanguínea pelas fêmeas é a característica que permite que os mosquitos transmitam patógenos de hospedeiros infectados para hospedeiros saudáveis. Os patógenos são adquiridos pelos mosquitos acidentalmente durante o repasto. Esses vírus e parasitas se replicam e se acumulam nas glândulas salivares do mosquito e então são inoculados em novos hospedeiros junto com a saliva durante a alimentação. Esse hábito hematófago das fêmeas de muitas espécies de mosquitos incomoda humanos e outros animais além de serem importantes vetores de relevância médica e veterinária, transmitindo patógenos como arbovírus, helmintos e protozoários. Esses insetos estão associados a níveis elevados de morbidade e mortalidade humana, sendo responsáveis pela disseminação de doenças como dengue, febre amarela, malária e filariose (HAWKES & HOPKINS., 2022).

1.2.4 Mosquitos vetores da febre amarela

Os mosquitos dos gêneros *Haemagogus* e *Sabethes* desempenham funções distintas como vetores da febre amarela (FA), sendo os do gênero

Haemagogus considerados os vetores principais, enquanto os do gênero *Sabethes* atuam como vetores secundários (HANLEY et al., 2013; ABREU et al., 2019; OLIVEIRA et al., 2023). No Brasil, as espécies mais relevantes para a transmissão da FA incluem *Hg. janthinomys*, *Hg. leucocelaenus* e *Sa. chloropterus* (VASCONCELOS et al., 1997; MONDE., 2002; CARDOSO et al., 2010; OBARA et al., 2012; OLIVEIRA et al., 2023). *Ae. albopictus* também é apontado como potencial vetor da doença, podendo ser uma “ponte” para reurbanização da febre amarela devido à sua ampla distribuição geográfica e alta plasticidade ecológica (PEREIRA DOS SANTOS et al., 2018; DAMASCENO-CALDEIRA et al., 2023). Além dessas, também há registros de infecção natural pelo vírus YFV em outras espécies: *Aedes serratus*, *Ae. aureolineatus*, *Ae. fulvus*, *Ae. scapularis*, *Ae. taeniorhynchus*, *Psorophora ferox*, *Ps. albipes*, *Sa. soperi* e *Sa. albiprivus*, *Sa. identicus* e *Shannoniana fluviatilis* (DÉGALLIER et al., 1992; VASCONCELOS., 2003; CARDOSO et al., 2011; MORENO et al., 2011; ABREU et al., 2019; DE OLIVEIRA et al., 2021; CUNHA et al., 2020; CANO et al., 2022).

Em relação as espécies utilizadas nesse estudo, *Hg. leucocelaenus* (Figura 4) é um dos principais vetores da febre amarela no Brasil apresenta atividade diurna e acrodendrófila. Apesar da sua preferência por ambientes florestais *Hg. leucocelaenus* pode ser encontrada fora das florestas, em bosques ou áreas adjacentes (VASCONCELOS et al., 2010) e é facilmente encontrada próxima aos arredores de áreas periurbanas e demonstra tendência à domiciliação devido a sua maior capacidade de adaptação à ambientes modificados (CARDOSO et al., 2010, MARCONDES & ALENCAR., 2010; CAMARGO-NEVES et al., 2005).

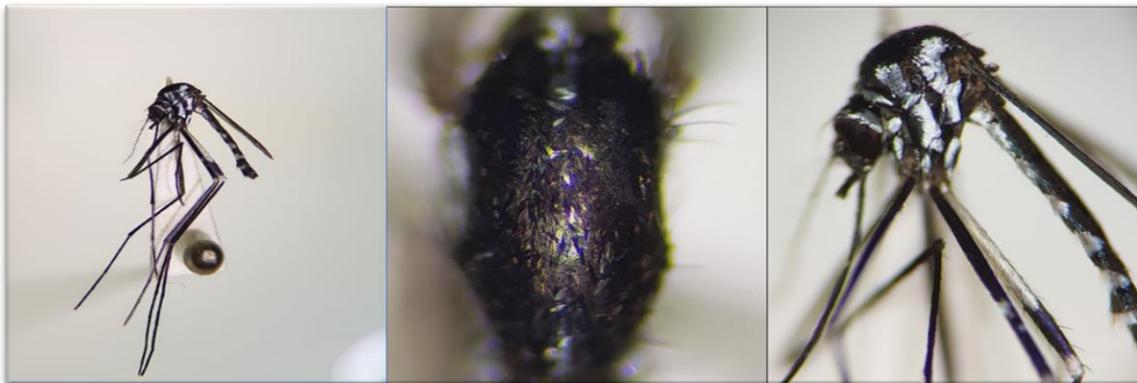


Figura 4 - Aspectos gerais de *Haemagogus leucocelaenus*.

Sabethes albiprivus (Figura 5), é um vetor secundário da FA no Brasil. Mosquitos dessa espécie foram encontrados infectados naturalmente com o vírus da FA pela primeira vez em 2021 (OLIVEIRA et al., 2023). Na Argentina a espécie é incriminada como importante vetor no ciclo de transmissão da doença (CANO et al., 2022). Os mosquitos Sabethini são diurnos, tímidos e algumas espécies são acrodendrófilas. Esses mosquitos se alimentam preferencialmente de aves e primatas não humanos em vegetação preservada. Em bordas de florestas podem alimentar-se de humanos ou outros animais no solo. (FORATTINI., 2002; MARASSÁ et al., 2009; GOMES et al., 2010; NAVARRO-SILVA, 2008)

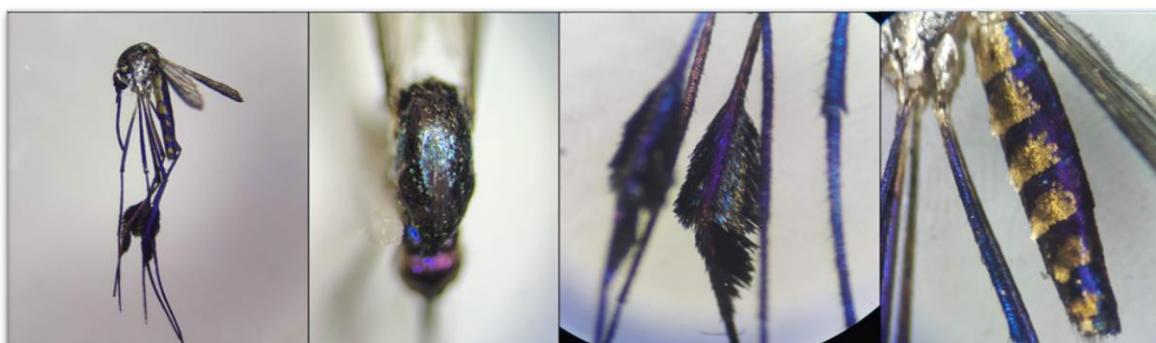


Figura 5- Aspectos gerais de *Sabethes albiprivus*.

As espécies *Aedes scapularis* (Figura 6) e *Ae. serratus* (Figura 7) foram encontrados naturalmente infectados com o vírus da FA, mas ainda não foram realizados estudos detalhados sobre a competência dessas espécies para

transmitir FA (CUNHA et al., 2020; CARDOSO et al., 2010). *Ae. scapularis* foi incriminado como principal vetor da epidemia de encefalite por vírus Rocio no Vale do Ribeira / São Paulo na década de 1970 (FORATTINI et al., 1981). Também há registro da espécie *Ae. serratus* naturalmente infectado com o vírus Oropouche na Região Amazônica (VASCONCELOS et al., 1997); e com o vírus trocara no Pará e Peru (TRAVASSOS-DA-ROSA et al., 2001), com vírus Aura no Pará e na Província de Misiones na Argentina (TRAVASSOS-DA-ROSA et al., 1998; SABATINI., 1998); *Ae. serratus* também é considerado vetor secundário do vírus Ilhéus (VASCONCELOS et al., 1997). CUNHA et al. (2020) e Cardoso et al. (2010) destacam a necessidade estudar a competência vetorial de *Ae. scapularis* e *Ae. serratus*, para desvendar o papel dessas espécies na transmissão da FA nas Américas, já que as duas espécies já foram encontradas naturalmente infectadas com o vírus da FA.



Figura 6 - Aspectos gerais de *Aedes scapularis*.



Figura 7- Aspectos gerais *Aedes serratus*.

1.3 Identificação de culicídeos

Após investigações realizadas no final do século XIX e início do século XX, descobriu-se que a malária e a febre amarela eram transmitidas por mosquitos. Essa relação entre mosquitos e doenças tropicais impulsionou a necessidade de descrever e classificar esses insetos (HARBACH., 2007). Em 1901, o entomologista britânico Frederick Vincent Theobald publicou o primeiro volume da primeira chave taxonômica para identificação de mosquitos, *A Monograph of the Culicidae, Or Mosquitoes* (Figura 8). Nessa obra, Theobald descreve centenas de espécies de mosquitos, baseando-se em espécimes da coleção do Museu de História Natural Britânico, com o objetivo de auxiliar médicos envolvidos nas investigações da transmissão de malária e febre amarela na identificação precisa das espécies incriminadas como vetores dessas doenças. A publicação da obra ocorreu em cinco volumes, entre 1901 e 1910, sendo considerada a primeira chave dicotômica global publicada.

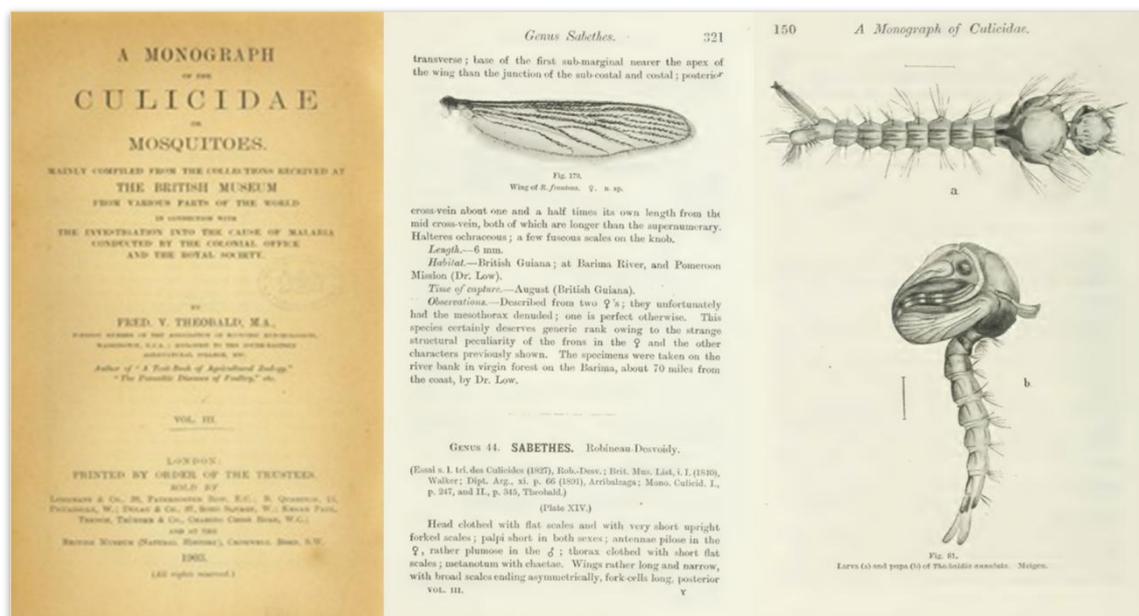


Figura 8 - Primeira chave taxonômica para identificação de mosquitos, *A Monograph of the Culicidae, Or Mosquitoes* de Frederick Vincent Theobald.

A identificação de mosquitos é realizada principalmente por meio de chaves dicotômicas clássicas, as quais requerem tempo e conhecimento especializado do profissional responsável pela tarefa de identificação. No Brasil, três bibliografias são amplamente utilizadas como referência para essa finalidade: *Neotropical Culicidae* (LANE., 1953), uma referência clássica para identificação de mosquitos na região neotropical, com ampla cobertura de espécies encontradas no território brasileiro, contendo descrições detalhadas, ilustrações e chaves de identificação; *Culicidologia Médica, Volume 2* (FORATTINI., 2002), que aborda não apenas a identificação das espécies da região neotropical, mas também aspectos de sua biologia e epidemiologia; e *Principais Mosquitos de Importância Sanitária no Brasil* (CONSOLI & LOURENÇO., 1994), que fornece informações detalhadas sobre as principais espécies de mosquitos vetores de doenças no Brasil, abrangendo aspectos biológicos e ecológicos das formas adultas e imaturas, além de tratar da morfologia, métodos de captura, identificação e montagem de espécimes (Figura 9).

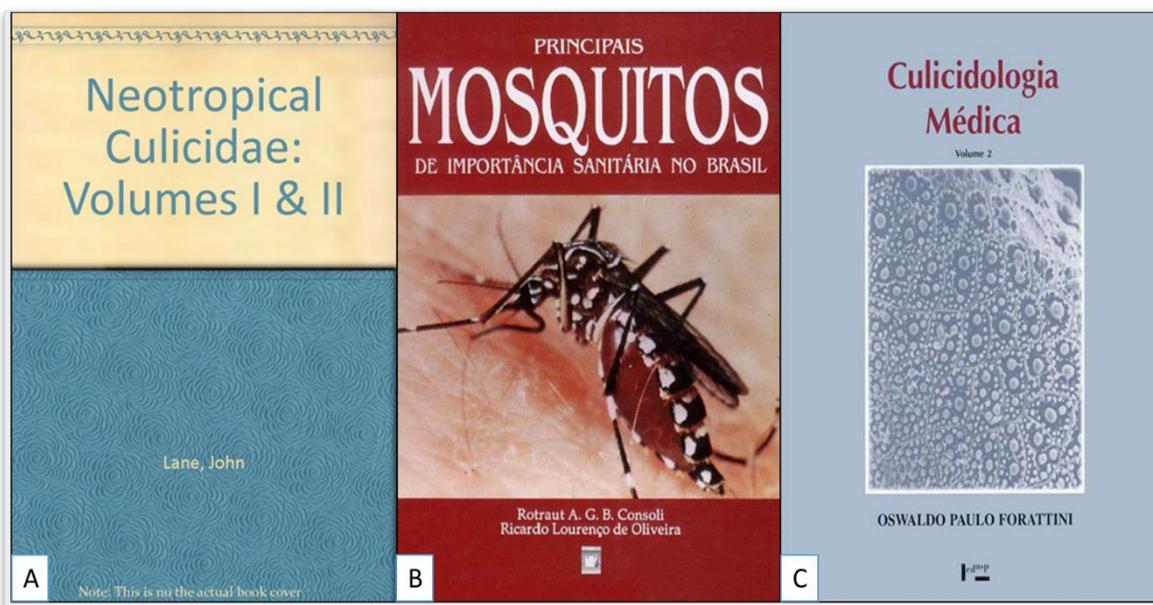


Figura 9 - Bibliografias utilizadas para identificação de mosquitos a partir de chaves dicotômicas. A) *Neotropical Culicidae* (LANE., 1953); B) *Culicidologia Médica, Volume 2* (FORATTINI., 2002); C) *Principais Mosquitos de Importância Sanitária no Brasil* (CONSOLI & LOURENÇO., 1994).

Outras técnicas de identificação de mosquitos podem ser utilizadas, como morfometria geométrica de asas e biologia molecular. Ambas possuem vantagens e desvantagens. A morfometria geométrica tem com vantagem não requerem equipamentos laboratoriais caros, fornece ferramentas objetivas para discriminação de formas e reconhecimento de padrões, que podem ser particularmente úteis na distinção entre espécies intimamente relacionadas. Contudo, existem barreiras metodológicas que impedem a adoção generalizada da morfometria. Por exemplo, não há consenso sobre as melhores práticas para marcar pontos de referência nas asas, o que pode introduzir viés na análise (LORENZ & SUSDESK., 2013). Já a taxonomia a partir de biologia molecular fornece identificação precisa das espécies por meio da análise de material genético, sendo útil para identificação de espécies crípticas que são morfologicamente semelhantes. Porém técnicas moleculares requerem profissionais treinados, equipamentos e reagentes caros, tornando-as menos acessíveis, particularmente em ambientes com recursos limitados (LORENZ & SUSDESK., 2013).

Com os avanços tecnológicos e a praticidade proporcionada pelos dispositivos móveis, surgem alternativas ao uso das chaves dicotômicas e pictóricas tradicionais. O desenvolvimento de métodos interativos de identificação, baseados em dicotomias e politomias, já vem sendo utilizado para diversos organismos. As chaves eletrônicas e pictóricas são uma alternativa que vem substituindo as chaves de identificação impressas, com vantagens como facilidade de atualização e possibilidade de inclusão em smartphones (GURGEL-GONÇALVES et al., 2021; WALTER & WINTERTON., 2007). Ferramentas de desenvolvimento web, como a plataforma online Xper, têm sido empregadas para a elaboração de chaves eletrônicas destinadas à identificação de mosquitos (NEVES et al., 2024) (Figura 10). Essa ferramenta online é capaz de integrar uma grande quantidade de informações, considerando caracteres morfológicos e fotografias de indivíduos de diversos táxons, além de não exigir conhecimentos avançados em desenvolvimento web para sua utilização (UNG et al., 2010; VIGNES-LEBBE et al., 2016). Embora essas novas ferramentas tornem a identificação mais acessível, elas ainda demandam profissionais devidamente treinados para reconhecer e distinguir as características

morfológicas apresentadas. Uma abordagem promissora é o desenvolvimento de sistemas automatizados para identificação de espécies de mosquitos, utilizando processamento e análise de imagens por meio de algoritmos baseados em inteligência artificial.

Sabethes Pictures : hide Settings

20 Descriptors History (0)

Upper proepisternal setae

Present (18)

Absent (16)

Midtarsi

Foretarsi

Lower mesokatepisternal seta

34 Remaining taxa Among 34

- Sabethes (Sabethes) albiprivus (Morphotype A)
- Sabethes (Sabethes) albiprivus (Morphotype B)
- Sabethes (Sabethes) amazonicus
- Sabethes (Peytonulus) aurescens
- Sabethes (Sabethes) batesi
- Sabethes (Sabethes) belisarioi
- Sabethes (Sabethes) bipartipes
- Sabethes (Sabethoides) chloropterus
- Sabethes (Sabethes) cerqueirai
- Sabethes (Sabethoides) conditus
- Sabethes (Sabethes) cyaneus
- Sabethes (Peytonulus) fabricii
- Sabethes (Sabethes) forattinii
- Sabethes (Sabethoides) glaucodaemon
- Sabethes (Sabethes) gymnothorax
- Sabethes (Peytonulus) hadrognathus
- Sabethes (Peytonulus) identicus
- Sabethes (Sabethinus) idiogenes
- Sabethes (Peytonulus) ignotus
- Sabethes (Sabethinus) intermedius
- Sabethes (Sabethes) lanei
- Sabethes (Sabethinus) melanonymphe
- Sabethes (Sabethes) ortizi

xper3 Documentation Database : Sabethes Authors : Neves, M.S.A.S., Motta, M.A., Maciel-de-Freitas, R., Xavier, A.S., Lourenço-de-Oliveira, R., Silva-do-Nascimento, T.F.

Figura 10 - Chave eletrônica online na plataforma *Xper* para identificação de fêmeas do gênero *Sabethes*. Fonte: NEVES et al., 2024.

1.4 Aprendizado de máquina

Uma das primeiras publicações alusivas à inteligência artificial (IA) foi o trabalho publicado por MCCULLOCH & PITTTS em 1943. Seu estudo descreve um modelo de computador utilizado para aprender com base em um processo comparável aos neurônios do cérebro humano. O modelo descrito em sua publicação foi denominado neurônio MCP e funcionava recebendo entradas nas quais as variáveis ou valores só podem assumir dois estados possíveis: verdadeiro ou falso. Modelos matemáticos computacionais capazes de prever resultados após o treinamento baseado em bancos de dados foram relatados pela primeira vez em 1950. Nesse estudo, o autor levantou a questão de saber se as máquinas poderiam imitar a inteligência humana e introduziu um teste para avaliar a inteligência do modelo, conhecido como Teste de Turing. Este teste envolve um interrogador humano cego questionando um entrevistado humano e um entrevistado máquina. A tarefa do interrogador é identificar qual entrevistado é a máquina. Se o interrogador não for capaz de discernir as respostas da máquina das respostas humanas, considera-se que a máquina tenha passado no Teste de Turing (CICEKLI, AKMAN, SAYGIN., 2000).

O aprendizado de máquina (Machine Learning - ML), um subcampo da inteligência artificial (IA), tem como objetivo desenvolver modelos de algoritmos capazes de aprender tarefas complexas a partir de dados, sem que sejam programados para isso. O sistema de aprendizado de máquina comporta-se de maneira semelhante ao cérebro humano, com seu desempenho melhorando com a experiência quanto mais dados, melhor o desempenho. O sucesso de um modelo de ML depende de um banco de dados de qualidade. Os primeiros modelos de IA tentaram simular a função de um único neurônio, com funções simples de entrada e saída. Com o tempo, esses modelos tornaram-se mais sofisticados com a adição de novas funções, originando as redes neurais profundas (Deep Learning - DL) (Figura 11) (MUTHUKRISHNAN et al., 2020).

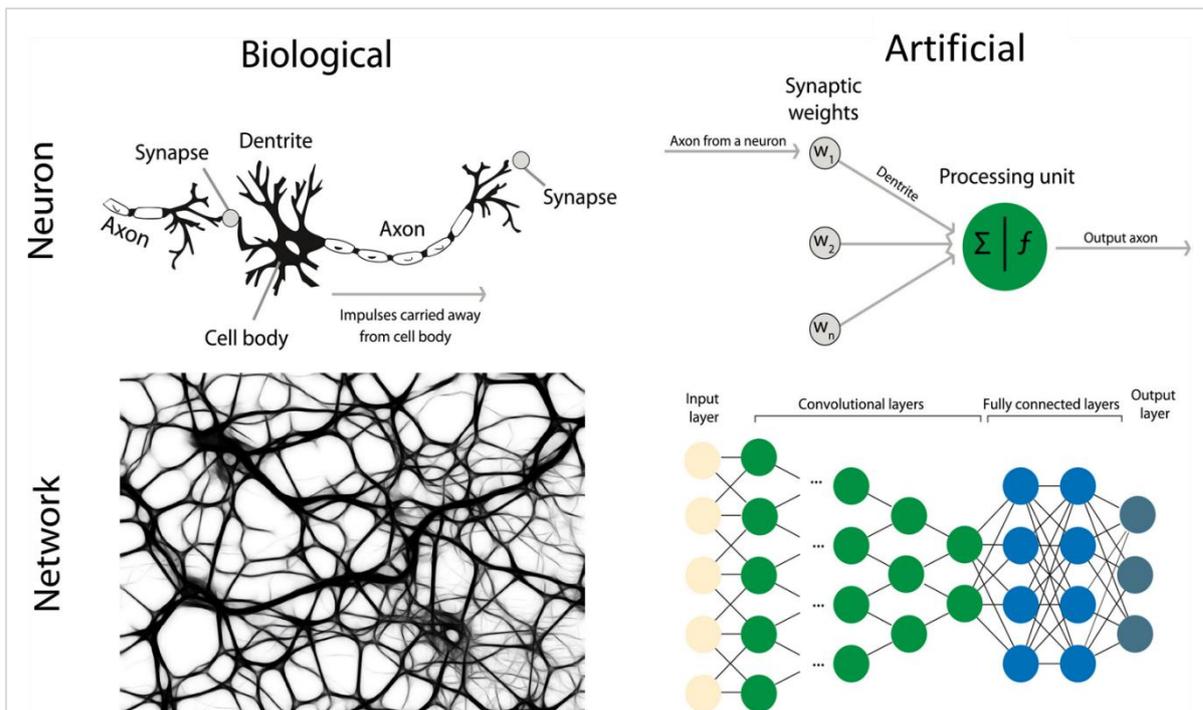


Figura 11 - Esquema ilustrativo comparativo entre redes neurais biológicas e redes neurais artificiais. Fonte: WÄLDCHEN & MÄDE., 2018.

Métodos de ML podem ser classificados em três tipos de aprendizado: (1) aprendizado supervisionado, no qual o modelo é treinado utilizando um conjunto de dados rotulado, ou seja, para cada entrada no conjunto de treinamento, existe uma saída conhecida (rótulo). Exemplo: classificar imagens de lesões como benignas ou malignas; (2) aprendizado não supervisionado, onde o modelo é treinado com um conjunto de dados que não possui rótulos, e o modelo busca identificar estruturas ou padrões ocultos nos dados, sem a ajuda de saídas pré-definidas; (3) aprendizado por reforço, no qual o modelo é treinado para atingir uma meta em um ambiente incerto, aprendendo com tentativas e erros para cumprir um objetivo e receber uma recompensa. Esse tipo de aprendizado de máquina é muito utilizado no desenvolvimento de jogos e na robótica (LECUN et al., 2015; DEO., 2015; LE et al., 2020).

Os algoritmos de aprendizado de máquina foram, desde o início, projetados e utilizados para analisar conjuntos de dados médicos. Atualmente, o aprendizado de máquina fornece várias ferramentas indispensáveis para análise inteligente de dados (KONONENKO., 2001), revolucionando muitos setores e executando tarefas que, normalmente, exigiriam inteligência humana. Tarefas

complexas, como detecção rápida de doenças e diagnóstico automatizado, são realizadas com a ajuda da IA (MIOTTO et al., 2018). A IA tem sido aplicada para diagnóstico precoce da doença de Alzheimer a partir de ressonâncias magnéticas cerebrais (LIU et al., 2014), avaliação de lesões para diagnóstico de câncer de pele e leishmaniose cutânea (ESTEVA et al., 2017; LEAL et al., 2023), identificação de retinopatia diabética (GULSHAN et al., 2016) e monitoramento da saúde geral (PHAM et al., 2016). No domínio das doenças tropicais transmitidas por vetores, os modelos de DL têm potencial para ajudar na detecção e identificação de vetores, na classificação e previsão de cenários epidemiológicos, otimizando o monitoramento e controle de vetores (KAUR et al., 2022, NAYAK et al., 2022; RAIZADA., 2021).

Para identificação de imagens, métodos de DL baseados em Redes Neurais Convolucionais (CNN - Convolutional Neural Networks) são os mais eficientes. As CNNs simulam o processo de aprendizagem humana para a classificação de imagens, extraem características importantes dos dados automaticamente, sem supervisão humana, e são eficazes na identificação automatizada de objetos a partir de imagens (LECUN et al., 2015; LE et al., 2020). Os modelos de DL comumente utilizados para identificação de imagens são classificados em duas categorias, com base nas tarefas de pesquisa: classificação de imagens e detecção de objetos. A classificação de imagens visa reconhecer a imagem de entrada e gerar a categoria à qual a imagem pertence como os modelos de classificação de imagens AlexNet, VGGNet (rede de grupos de geometria visual), ResNet (rede residual profunda) e MobileNet. Já as redes de detecção de objetos têm como tarefa identificar vários objetos, determinar suas posições na imagem e atribuir as categorias correspondentes (LECUN et al., 2015; AZIZ et al., 2020) (Figura 12). Exemplos de modelos de detecção de objetos incluem a rede You Only Look Once (YOLO) e o Single Shot Multibox Detector (SSD), que são muito eficazes para classificação de imagens em tempo real.

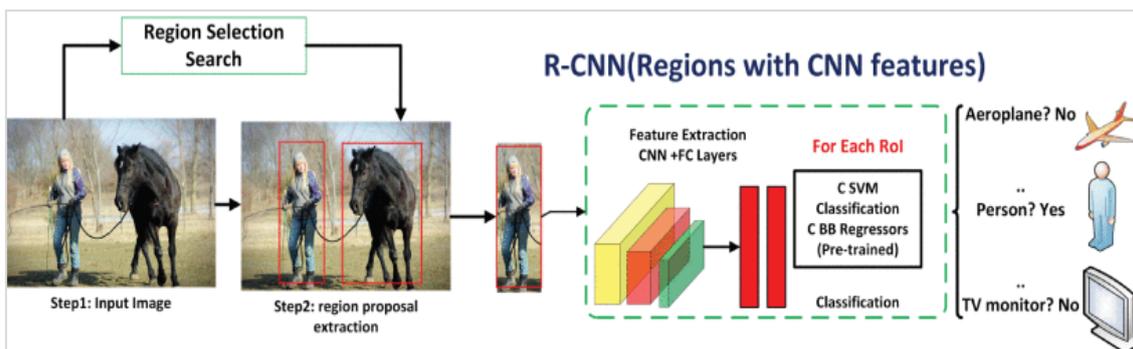


Figura 12 - Arquitetura de uma rede de detecção de objetos. 1) Imagem de entrada; 2) reconhecimento de objetos e isolamento do mesmo; 3) Extração das características de cada objeto, 4) classificação de cada objeto. Fonte: AZIZ et al., 2020.

1.5 Identificação automatizada de insetos e outros invertebrados

A identificação taxonômica por meio de métodos de aprendizagem profunda a partir de CNNs vem se mostrando promissora. A técnica já foi aplicada com sucesso em vários grupos taxonômicos como: mamíferos, peixes e plantas (GOMEZ VILLA, SALAZAR, VARGAS., 2017; QIN et al., 2016; CHENG et al., 2017). Para o grupo de invertebrados, a identificação automatizada vem sendo utilizada para identificação de polinizadores, invertebrados aquáticos, insetos de interesse médico, forense e agrícola (VALAN et al, 2019). Esses métodos objetivam simplificar e aprimorar a precisão do processo de identificação desses organismos.

SPIESMAN et al. (2021) compararam quatro modelos de CNNs para identificação automatizada de abelhas da América do Norte. Entre os modelos, o InceptionV3 apresentou a melhor acurácia (91,6%). Com os resultados do estudo, foi desenvolvido o aplicativo BeeMachine (Figura 13), uma ferramenta web que permite que qualquer pessoa identifique abelhas com suas próprias imagens. Posteriormente, SPIESMAN et al. (2024) treinaram um outro modelo de classificação de imagens, o EfficientNetV2L, para identificar imagens de 20 espécies de abelhas pequenas e difíceis de identificar, utilizando um conjunto de imagens de asas e outro de espécimes fixados. A acurácia média da rede utilizada foi de 94,9% para imagens de espécimes fixados e de 98,1% para imagens de asas. STARK et al. (2023) utilizaram três modelos leves de CNNs: YOLOv5nano, YOLOv5small e YOLOv7tiny, o modelo YOLO é um tipo de CNN

de detecção de objetos e foi utilizada para o reconhecimento e classificação de visitantes florais em tempo real, alcançando acurácias variando de 93% a 97% para identificar artrópodes das ordens Hymenoptera, Diptera, Coleoptera, Lepidoptera e Arachnida. Os resultados desses estudos mostram que a utilização de CNNs para identificação de abelhas e outros artrópodes é uma ferramenta valiosa para o monitoramento e a conservação de polinizadores.

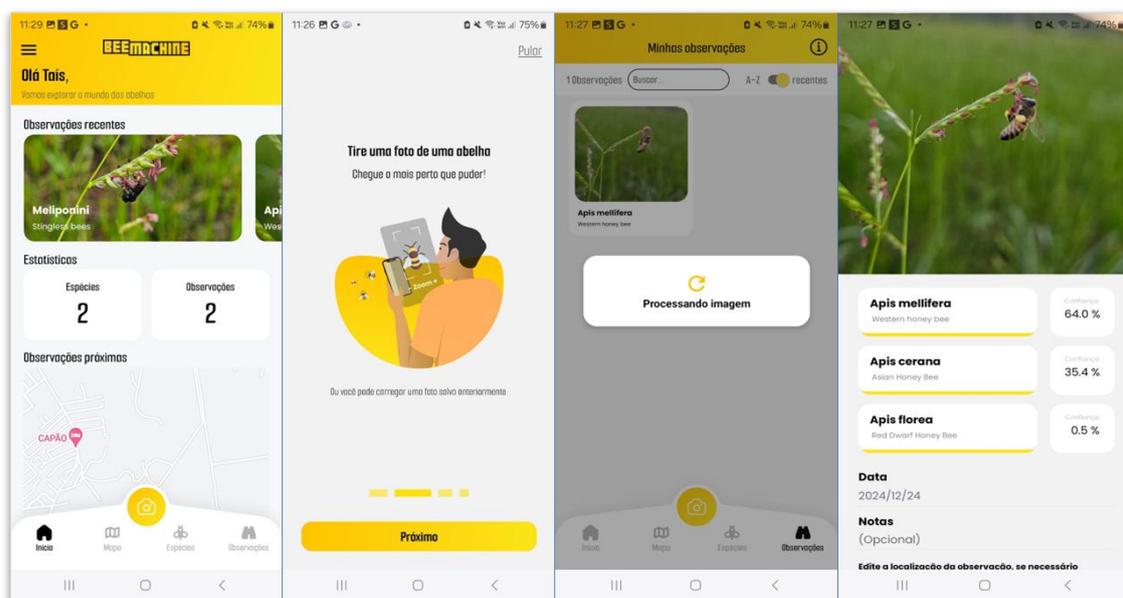


Figura 13 - Bee Machine: Aplicativo móvel para identificação de abelhas da América do Norte. Fonte: SPIESMAN et al., 2021. (<https://beemachine.ai/>)

Estudos para identificação automatizada de insetos aquáticos bioindicadores também já foram realizados. SIMOVIĆ et al. (2024) montaram um dos maiores bancos de dados de imagens de insetos aquáticos, treinando uma rede para identificação de 90 espécies utilizadas para avaliar a saúde ambiental de ambientes aquáticos, com acurácia de 98,7%. JABALLAH et al. (2021) propuseram um novo modelo de CNN, apelidado de MacroNet, uma rede melhorada baseada no modelo RetinaNet para detecção de objetos. Tanto a rede RetinaNet quanto a MacroNet foram avaliadas para identificação de insetos aquáticos vivos, com acurácias médias de 4,3% e 74,93%, respectivamente. Melhoramentos na rede MacroNet produziram resultados melhores de acurácia para identificação dos insetos variando de 70 a 100% (JABALLAH et al., 2021).

Os resultados dos estudos são promissores para o monitoramento de baixo custo e não letal de bioindicadores de água doce.

Para otimizar a identificação de moscas de interesse forense, APASRAWIROTE et al. (2022) avaliaram o desempenho de quatro modelos de CNNs, incluindo ResNet-101, Densenet161, VGG19_bn e AlexNet, para a identificação de larvas de quatro espécies de moscas: *Chrysomya megacephala* (Diptera: Calliphoridae), *Chrysomya (Achoetandrus) rufifacies* (Diptera: Calliphoridae), *Lucilia cuprina* (Diptera: Calliphoridae) e *Musca domestica* (Diptera: Muscidae), com base em imagens de espiráculos posteriores. Todas as redes tiveram bons resultados de acurácia, mas o AlexNet demonstrou melhor desempenho e velocidade, sendo escolhido para o desenvolvimento do aplicativo web "The Fly". GOHE et al. (2024) identificaram cinco gêneros de moscas forenses — *Chrysomya*, *Lucilia*, *Sarcophaga*, *Rhiniinae* e *Stomorphina* — utilizando dois modelos de CNN, MobileNetV3-Large e VGG19. Ambos os modelos alcançaram uma precisão muito alta de 99,39% e 99,79%. No geral, os resultados destes estudos destacam o potencial dos modelos de aprendizado profundo para a classificação de espécies de moscas de interesse forense.

Para os vetores de leishmanioses, estudos de identificação utilizando CNNs ainda são escassos. CANNET et al. (2023) desenvolveram dois modelos de CNN, DarkNet-9 e DarkNet-14, para a identificação de fotos de asas de 17 espécies de flebotomíneos, obtendo uma taxa de identificação de 99,8% para a tarefa de separar flebotomíneos de não flebotomíneos. Para a tarefa de classificação a nível de gênero, a acurácia foi superior a 90%, e para a tarefa de classificação a nível de espécie, a taxa de identificação foi superior a 77%. FRAIWAN et al. (2024) avaliou quatro modelos de CNNs para identificação de flebotomíneos utilizando imagens da região da faríngea e do aparelho genital, obtendo médias de acurácias de 95% para identificação a partir de genitálias e 86.7% para identificação com imagem da região da faringe.

Para os vetores da doença de Chagas, um estudo pioneiro alcançou sucesso de identificação de 87,8% e 80,3% para as espécies de triatomíneos brasileiros e mexicanos, respectivamente (GURGEL-GONÇALVES et al., 2017). Posteriormente, um estudo utilizando o mesmo conjunto de dados, com DL e incorporando informações de distribuição geográfica para reduzir o número de

espécies nas análises, melhorou as taxas de identificação para 98,9% para as espécies brasileiras e 95,8% para as mexicanas (KHALIGHIFAR et al., 2019). PARSONS & BANITAAN (2021) também conduziram uma análise comparativa das metodologias de ML e DL para fins de identificação, alcançando acurácias de 88,60% e 89,20% para os barbeiros brasileiros e mexicanos, respectivamente. ABDELGHANI et al. (2021) alcançaram ótimos resultados de classificação utilizando o modelo VGG-16, com uma taxa de precisão de 99,45% para distinguir barbeiros de não barbeiros e 96% para diferenciar entre várias espécies de barbeiros. COCHERO et al. (2021) avaliaram três modelos leves adequados para aplicativos móveis para a identificação de barbeiros, alcançando uma taxa de identificação de 94,3%. MIRANDA et al. (2024) avaliou o desempenho de cinco algoritmos de aprendizado de máquina na identificação de 13 espécies de triatomíneos utilizando imagens de diferentes ângulos do inseto, obtendo acurácias entre 69% e 97%. Também foi observado que o ângulo não influencia no desempenho de identificação. Posteriormente, MIRANDA et al. (2024b) utilizou a rede AlexNet para a identificação de 65 espécies de triatomíneos a partir de imagens dorsais de insetos adultos obteve acurácias entre 93 e 95%.

AKBARIAN et al. (2022) utilizaram métodos de DL para a classificação de carrapatos. O modelo Inception-ResNet foi treinado, alcançando uma precisão de 92% para a tarefa de classificação binária, separando *Ixodes scapularis* de outros carrapatos. OMODIOR et al. (2021) desenvolveram um modelo de CNN personalizado e leve para a classificação de imagens de carrapatos e compararam-no com a rede ResNet-50. As precisões de teste foram de 80% e 75%, respectivamente, para a identificação de *Ixodes scapularis*, *Dermacentor variabilis*, *Amblyomma americanum* e *Haemaphysalis* sp. Em outro estudo, pesquisadores compararam cinco modelos de CNNs para a identificação de *Amblyomma americanum*, *Dermacentor variabilis* e *Ixodes scapularis*, todos alcançando precisões superiores a 98%. Entre eles, o InceptionV3 obteve a maior precisão, com 99,5%, enquanto o MobileNetV2 obteve a menor, com 98,73%.

No contexto de gastrópodes de interesse médico, QASMI et al. (2024) utilizaram uma abordagem de aprendizado supervisionado (Random Forest) e

uma rede neural ML (XGBoost) para identificar espécies de caramujos marinhos do gênero *Conus sp.*, que são venenosos, obtendo acurácia de 95,7%. TALLAM et al. (2021) aplicaram métodos de DL para a identificação de caramujos hospedeiros de *Schistosoma sp.* Este estudo comparou sete modelos de CNN, com o modelo VGG16 tendo o melhor desempenho, com uma precisão de 99,60%. SHI et al. (2021) avaliaram os modelos MobileNetV2, ResNet50 e Inception-ResNet-V2 para a tarefa de classificação binária de identificação de *Oncomelania hupensis* e não-*Oncomelania hupensis*. Todos os modelos alcançaram precisões de mais de 92%, sendo o Inception-ResNet-V2 o modelo com maior acurácia: 96,13%

Em conjunto, todos os resultados obtidos nos estudos aqui mencionados mostram que as CNNs são ferramentas promissoras tanto para estudos de monitoramento, como para estudos taxonômicos e ecológicos. Auxiliado os profissionais na tarefa de identificação, otimizando o tempo de ação quando for necessário intervenções, como por exemplo o controle de vetores.

1.6 Identificação automatizada de Culicídeos

Os mosquitos podem ser identificados e classificados por meio de algoritmos de aprendizado profundo, os quais podem ser treinados utilizando imagens, assinaturas acústicas dos batimentos das asas e assinaturas pseudoacústicas, como dados de frequência provenientes de sensores ópticos (OKAYASU et al., 2019; PARK et al., 2020; MOTTA et al., 2019, 2020; ONG et al., 2020, 2022; ZHAO et al., 2022; ARISTA-JALIFE et al., 2020; ORTIZ et al., 2018; ASMAI et al., 2019; GOODWIN et al., 2021; GUPTA et al., 2024; MONTALBO et al., 2023; KISKIN et al., 2020; KHALIGHIFAR et al., 2021; WEI et al., 2022; FERNANDES et al., 2021) (Tabela 1)

Tabela 1 - Estudos que aplicaram CNNs para a identificação de espécies de mosquitos, destacando as redes utilizadas, as espécies identificadas e médias gerais de acurácia.

Referências	Rede Utilizada	Espécies identificadas	Acurácia média (%)
OKAYASU et al., 2019	AlexNet, VGGNet, ResNet	<i>Aedes albopictus</i> , <i>Anopheles stefensi</i> e <i>Culex pipiens pallens</i>	95.5
PARK et al., 2020	VGG-16, ResNet-50, SqueezeNet	<i>Aedes albopictus</i> , <i>Aedes vexans</i> , <i>Anopheles sp.</i> , <i>Culex pipiens</i> , <i>Culex tritaeniorhynchus</i> , <i>Aedes dorsalis</i> , <i>Aedes koreicus</i> e <i>Culex inatomi</i>	94.3
MOTTA et al., 2019	LeNet, AlexNet, GoogleNet	<i>Aedes aegypti</i> , <i>Aedes albopictus</i> e <i>Culex quinquefasciatus</i>	76.2
MOTTA et al., 2020	ResNet50, VGG16, XceptionV3, DenseNet201	<i>Aedes aegypti</i> , <i>Aedes albopictus</i> e <i>Culex quinquefasciatus</i>	93.5
ONG et al., 2021	MobileNet	<i>Aedes aegypti</i> e <i>Aedes albopictus</i>	98.06
ONG et al., 2022	DenseNet201, InceptionV3, InceptionResNetV2, MobileNetV2, Xception e VGG19	<i>Aedes aegypti</i> e <i>Aedes albopictus</i>	77
ZHAO et al., 2022	Detection Transformer, Vision Transformer, Swin Transformer, YOLOv5, Mask R-CNN e DenseNet	17 espécies (3 subespécies do complexo <i>Culex pipiens</i>)	80
ARISTA-JALIFE et al., 2020	Dual-Source DNN ,Shallow-NN, Slim-DNN, Robust-DNN	<i>Aedes aegypti</i> e <i>Aedes albopictus</i>	86
ORTIZ et al., 2018	AlexNet	<i>Aedes spp.</i>	97
ASMAI et al., 2019	Inception V3, ResNet-50, VGG16, VGG19	<i>Aedes spp.</i>	82.3
GOODWIN et al., 2021	Xception	67 espécies	88
GUPTA et al., 2024	Xception	53 espécies	93
MONTALBO et al., 2023	EfficientNetB7, DenseNet201, DenseNet121, EfficientNetB0, ResNet50V2, ResNet152V2, MobileNetV2, LeNet5, InceptionV3, NASNetLarge, Xception, InceptionResNetV2 e ONASNetMobile	<i>Aedes albopictus</i> , <i>Aedes vexans</i> , <i>Anopheles sinensis</i> , <i>Culex pipiens</i> e <i>Culex tritaeniorhynchus</i>	88
KISKIN et al., 2020	CNN autoral	<i>Culex quinquefasciatus</i>	92.2
KHALIGHIFAR et al., 2021	Inception V3	13 espécies	85
WEI et al., 2022	ResNet-18	<i>Culex quinquefasciatus</i> , <i>Anopheles arabiensis</i> , <i>Anopheles gambiae</i> , <i>Aedes albopictus</i> , <i>Aedes aegypti</i> e <i>Culex pipiens</i>	94
FERNANDES et al., 2021	CNN autoral	20 espécies de mosquitos	87,5

Na identificação de mosquitos, OKAYASU et al. (2019) foram os primeiros a avaliar o uso de identificação automatizada baseada em CNNs. Nesse estudo, foi utilizado um conjunto de dados de imagens de mosquitos dos gêneros *Aedes*, *Culex* e *Anopheles*. Foram comparadas técnicas baseadas em ML e CNNs, incluindo AlexNet, VGGNet e ResNet. As técnicas de DL apresentaram maior desempenho, com uma acurácia de 95,5%, em contraste com as técnicas de data mining, que alcançaram 82,4%. Os resultados também demonstraram que o desempenho das técnicas de DL é inferior às técnicas tradicionais baseadas em ML para identificação de mosquitos quando não treinadas com grandes conjuntos de imagens.

Em uma análise utilizando um banco de dados com cerca de 3.600 imagens de mosquitos fotografados em várias posições e condições de deformação, foram avaliadas três CNNs: VGG-16, ResNet-50 e SqueezeNet. Essas redes alcançaram acurácias de 97%, 96% e 90%, respectivamente, para identificar oito espécies de mosquitos (*Aedes albopictus*, *Aedes vexans*, *Anopheles* spp., *Culex pipiens*, *Culex tritaeniorhynchus*, *Aedes dorsalis*, *Aedes koreicus* e *Culex inatomi*), apesar da alta similaridade entre espécies e das variações intraespecíficas. Os resultados indicaram que mais de 90% de acurácia pode ser alcançada quando são aplicadas técnicas de ajustes finos (adaptar uma rede pré-treinada para um novo conjunto de dados) e aumento de dados (técnica para aumentar a quantidade de dados disponíveis para treinamento de modelos de DL sem precisar coletar novos dados). Nesse estudo, também foi observado que as CNNs utilizam as mesmas áreas e características morfológicas empregadas por taxonomistas para identificação de mosquitos, como, por exemplo, as pernas e a coloração (PARK et al., 2020). ADHANE et al. (2021) e ZHAO et al. (2022) também obtiveram resultados semelhantes ao utilizarem técnicas de ativação de mapas de calor, indicando que as áreas nas quais o modelo se concentra para a classificação da imagem correspondem às mesmas utilizadas na identificação taxonômica por métodos tradicionais, como chaves dicotômicas e eletrônicas.

MOTTA et al. (2019) identificaram mosquitos adultos das espécies *Aedes aegypti*, *Aedes albopictus* e *Culex quinquefasciatus* capturados em campo, utilizando as redes LeNet, AlexNet e GoogleNet, e obtiveram uma precisão

máxima de 76,2%. Em um estudo posterior, os mesmos autores, ao realizarem ajustes para a otimização de hiperparâmetros em modelos baseados em CNNs com arquiteturas complexas, aumentaram o desempenho para 93,5% de acurácia para as três espécies de mosquitos (MOTTA et al., 2020).

Embora alguns estudos tenham utilizado mosquitos capturados em campo e em diferentes posições e deformidades (MOTTA et al., 2019, 2020; PARK et al., 2020; ONG et al., 2021; PATAKI et al., 2021), ONG et al. (2022) avaliaram a identificação de imagens de mosquitos íntegros e danificados em pele humana enviadas pela comunidade. Nesse caso, foi comparado o desempenho de classificação de várias CNNs, sendo que, entre todas, a rede Xception exibiu o melhor desempenho, com uma precisão média de 78,12%.

ZHAO et al. (2022), diferentemente de outros pesquisadores que usaram CNNs, desenvolveram um modelo de identificação denominado Swin MSI, baseado no Swin Transformer, para a classificação de mosquitos. A acurácia atingiu 99,04%, superando os modelos de CNNs comparados no estudo. Metodologias baseadas em DL também têm sido aplicadas à identificação de larvas de mosquitos. CNNs identificaram larvas de *Aedes aegypti* e *Aedes albopictus* com acurácias acima de 90% (ORTIZ et al., 2018; ARISTA-JALIFE et al., 2020). ASMAI et al. (2019) investigaram a detecção de larvas de *Aedes* utilizando diferentes arquiteturas de CNN, como Inception V3, ResNet-50, VGG16 e VGG19, concluindo que a ResNet-50 apresentou o melhor desempenho.

Um outro estudo utilizando uma base de dados com 2696 imagens de 67 espécies de mosquitos, os pesquisadores propuseram um modelo de CNNs para classificação de espécies mosquitos, obtendo acurácias entre 86 e 97% (GOODWIN et al., 2021). Utilizando os mesmos algoritmos utilizados no estudo citando anteriormente, porém dessa vez melhorados para a tarefa de identificar mosquitos, criaram um dispositivo IDX (Identification-X) (Figura 14), que é um sistema óptico de alta resolução integrado a algoritmos de visão computacional para fornecer identificação precisa de mosquitos adultos e espécies de

carrapatos. Para mosquitos o dispositivo teve acurácia de identificação com médias entre 80 e 93% (GUPTA et al., 2024).

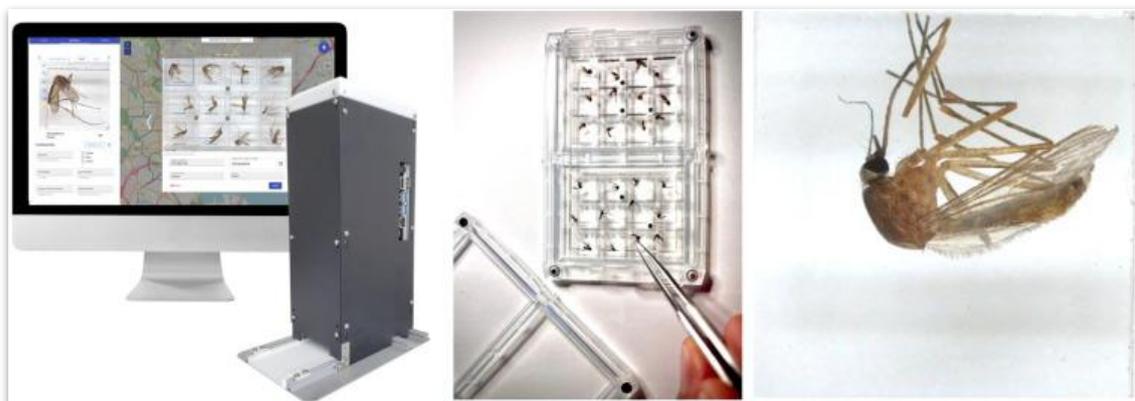


Figura 14. IDX um microscópio digital de captura imagens e identificação automatizada de mosquitos e carrapatos. Fonte: GUPTA et al., 2024.

Apesar de muitos estudos terem alcançado um bom desempenho em termos de acurácia, esses modelos geralmente demandam alto poder computacional e recursos de armazenamento para seu treinamento. Utilizando modelos leves que podem ser implementados em aplicativos web para identificação em tempo real, ONG et al. (2021) avaliaram o MobileNet para identificar mosquitos, alcançando uma precisão máxima de 98,06%. MONTALBO et al. (2023) utilizaram o EfficientNet B0, que possui o menor número de parâmetros entre as variantes do EfficientNet. O modelo leve proposto tem baixa complexidade e alcançou uma acurácia de identificação de 99,22%. Além disso, consome apenas 437 KB de espaço em disco, reduzindo consideravelmente a demanda por recursos computacionais.

Diferente dos estudos anteriores que utilizaram imagens para classificação, KISKIN et al. (2020) empregaram CNNs para identificar a presença de *Culex quinquefasciatus* a partir de bioacústica, alcançando uma acurácia de 92,2%. Da mesma forma, KHALIGHIFAR et al. (2021) avaliaram o modelo Inception v3 para identificar espécies de mosquitos utilizando espectrogramas, permitindo a discriminação de espécies de mosquitos. WEI et al. (2022) combinaram o ResNet-18 com um mecanismo de auto atenção que permite que o modelo se concentre em diferentes partes de uma sequência de entrada para construir a rede WbNet, alcançando precisões de 89,9% e 98,9% para dois

conjuntos de dados de áudio de mosquitos WINGBEATS e ABUZZ, respectivamente. FERNANDES et al. (2021) construíram um modelo CNN para classificar e identificar 20 espécies de mosquitos, alcançando uma precisão de classificação binária de 97,65% e uma precisão de modelo multiclasse de 78,12%.

Todos esses estudos demonstram que o uso das CNNs não apenas acelera o processo de identificação de mosquitos, mas também permite o monitoramento em tempo real das populações de mosquitos quando utilizados dados baseados em ciência cidadã (PATAKI et al., 2021; KHALIGHIFAR et al., 2021). Embora a utilização das CNNs para identificação de mosquitos vetores ainda esteja em teste e desenvolvimento, estudos contínuos possibilitarão a criação de sistemas inteligentes e eficazes para a vigilância e o controle de vetores, resultando em processos de intervenções mais rápidas.

2 JUSTIFICATIVA

A identificação precisa e rápida de mosquitos vetores é fundamental para o controle de doenças transmitidas por esses insetos. Tradicionalmente, as chaves dicotômicas impressas têm sido as principais ferramentas para identificação taxonômica. No entanto, chaves dicotômicas se tornam obsoletas e apresentam limitações quando as características diagnósticas são ambíguas, dificultando a identificação correta. O uso dessas chaves também exige conhecimento especializado, necessitando apoio de taxonomistas experientes. Embora métodos alternativos como a taxonomia molecular sejam eficazes, envolvem custos elevados e infraestrutura laboratorial complexa, tornando inviável para os laboratórios de referência entomológica adotar esses métodos moleculares, especialmente no contexto de recursos limitados.

Nesse cenário, a inteligência artificial surge como uma solução inovadora e acessível para superar essas limitações. Através de algoritmos de aprendizado de máquina é possível treinar modelos para reconhecer padrões em imagens de mosquitos, permitindo a identificação precisa e automatizada de espécies. Essa alternativa oferece vantagens, como a possibilidade de análise de grandes

volumes de dados em tempo real sem a necessidade de intervenções manuais constantes e com custos mais baixos em comparação com técnicas moleculares. Modelos de identificação automatizada já foram avaliados com sucesso para identificação de mosquitos urbanos, mas não têm sido aplicados para identificação de mosquitos silvestres, como os dos gêneros *Sabethes* e *Haemagogus*, vetores da febre amarela silvestre. A automatização do processo de identificação agiliza a tomada de decisões, possibilitando intervenções mais rápidas e eficazes para o controle de vetores diminuindo a propagação e os impactos das doenças transmitidas por mosquitos na sociedade.

3 OBJETIVO

Identificar espécies de mosquitos vetores do vírus da febre amarela silvestre aplicando uma rede de aprendizagem de máquina (AlexNet).

3.1 Objetivos específicos

- 1) Organizar um banco de imagens de mosquitos vetores de febre amarela e outros arbovírus.
- 2) Avaliar o desempenho da rede (AlexNet) para classificar 4 espécies de mosquitos a partir de imagens de diferentes partes do corpo.

4 METODOLOGIA

4.1 Banco de imagens

Para a construção do banco de imagens, foram selecionadas quatro espécies de mosquitos disponíveis no Laboratório de Parasitologia Médica e Biologia de Vetores da Universidade de Brasília. As espécies escolhidas são

vetoras ou já foram encontradas naturalmente infectadas com o vírus da febre amarela, sendo elas: *Sa. albiprivus* (n=100), *Hg. leucocelaenus* (n=98), *Ae. serratus* (n=100) e *Ae. scapularis* (n=100). Os mosquitos foram previamente fixados, sendo realizada a montagem dos espécimes em triângulos de cartolina inseridos em alfinetes. Os mosquitos foram fixados ao triângulo pela pleura torácica utilizando esmalte incolor, de forma a posicionar as pernas voltadas para o alfinete e a pleura superior livre para observação, conforme descrito por CONSOLI & OLIVEIRA (1994).

As fotografias foram capturadas usando uma câmera de celular (Samsung Galaxy S8, modelo SM-G950F, 12 mp, sensor 1/2.5', abertura f/1.7) acoplada a um estereoscópio (Zeiss Stemi 508, ampliação de 1-10x). Foram tiradas fotografias do corpo inteiro (n= 565), pronoto (n=484) e vista lateral do tórax (n=294) (Figura 15), totalizando 1.343 fotos. As imagens foram cortadas em formato quadrado (proporção 1:1) no software computacional Photoscape, selecionando apenas o mosquito na imagem, e posteriormente redimensionadas para 227 × 227 pixels usando o software computacional MATLAB. Todas as imagens foram organizadas em pastas de acordo com espécie e no experimento a ser realizado.

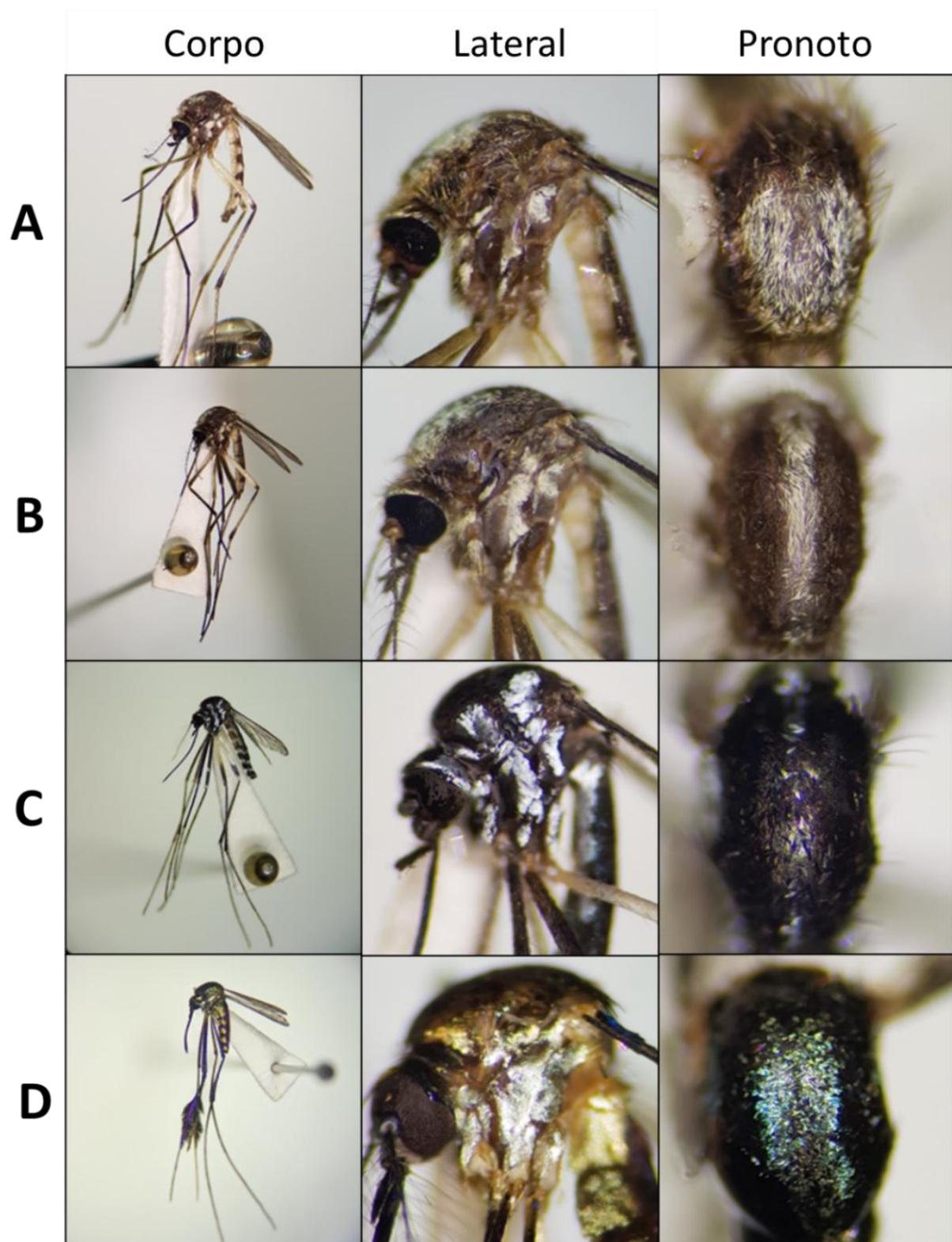


Figura 15 - Exemplos de imagens usadas para treinar a rede AlexNet. Mosquitos fotografados com estereomicroscópio e celular. (A) *Aedes scapularis*, (B) *Aedes serratus* (C) *Haemagogus leucocelaenus*, (D) *Sabethes albiprivus*.

4.2 Algoritmo

A rede AlexNet, proposta por Krizhevsky e colaboradores em 2012, é uma rede neural convolucional pioneira que revolucionou a forma como as máquinas "enxergam" o mundo, ao vencer a competição ImageNet Large Scale Visual Recognition em 2012. Sua arquitetura é composta por cinco camadas convolucionais, seguidas por três camadas totalmente conectadas, totalizando aproximadamente 60 milhões de parâmetros, sendo uma rede pré treinada com um grande conjunto de dados composto por 1,2 milhões de imagens, distribuídas em mais de 1.000 classes do banco de dados ImageNet.

A entrada da rede consiste em imagens RGB com resolução de 227×227 pixels, que são processadas em mini-lotes durante o treinamento. A rede opera com 5 cinco camadas convolucionais que extraem características das imagens usando filtros numéricos: A primeira camada convolucional é capaz de extrair padrões simples, como bordas e texturas, utilizando filtros 11×11 . A segunda camada convolucional refina as características extraídas, utilizando filtros menores de 5×5 . As camadas convolucionais seguintes (3, 4 e 5) extraem padrões mais complexos, como formas e componentes de objetos, com filtros de 3×3 . Nessas camadas é utilizado a técnica de normalização de resposta local (Local Response Normalization) que melhora a normalização em pequenas regiões da imagem, destacando as características mais importantes e suprimindo as menos relevantes. Entre as camadas também é utilizado operações de agrupamento máximo (max pooling). Essas operações funcionam como uma espécie de "resumo" das informações mais importantes em uma região da imagem. Mantendo as informações mais relevantes, reduzindo o processamento computacional. Após cada camada também é usado a função de ativação ReLU (Rectified Linear Unit) que acelera o processamento, introduzindo não linearidade na rede, tornando-a capaz de aprender funções mais complexas.

Além das cinco camadas convolucionais, a rede opera com mais três camadas totalmente conectadas que formam a parte final da arquitetura da rede, sendo elas responsáveis por combinar as características extraídas para realizar a classificação. Cada camada contém milhares de neurônios que processam as

representações aprendidas pelas camadas anteriores. Nessas camadas é aplicado a técnica de *Dropout* para reduzir o sobreajuste (overfitting), um problema comum onde o modelo se adapta demais aos dados de treinamento. A técnica desativa aleatoriamente um conjunto de neurônios da rede durante cada interação de treinamento. Forçando a rede a aprender representações complexas sem depender de nenhum neurônio específico. Na última camada, a função *Softmax* é utilizada para converter as ativações da rede em probabilidades para cada classe, permitindo a escolha da classe com maior probabilidade como a predição final (KRIZHEVSKY; SUTSKEVER; HINTON, 2012).

4.3 Treinamento e teste do algoritmo

A implementação da rede AlexNet foi realizada no ambiente computacional MATLAB (www.mathworks.com). A rede foi configurada para utilizar, do total de imagens, 80% das imagens para treinamento, 10% para validação interna e 10% para teste. A rede selecionou aleatoriamente o conjunto de imagens para cada etapa para garantir a imparcialidade dos resultados. Nas configurações do modelo, o número máximo de épocas (número de vezes que o modelo analisará todo o conjunto de dados de treinamento) foi definido como 50, com base em resultados de testes preliminares que indicaram essa configuração como adequada para a tarefa. A otimização foi realizada por meio do algoritmo de descida de gradiente estocástica com um otimizador de momento, utilizando parâmetros padrão. A única exceção foi a taxa de aprendizado inicial, ajustada para 0,001, visando um equilíbrio entre estabilidade e velocidade de aprendizado.

Foram conduzidos quatro experimentos distintos, cada um envolvendo quatro classes, com dez pseudo-replicatas (ou seja, 10 repetições de treinamento e teste para cada experimento).

1. Primeiro experimento: Utilizou o conjunto completo de imagens, abrangendo corpo inteiro, pronoto e vista lateral do tórax.

2. Experimentos subsequentes: Cada um foi realizado com imagens separadas de uma parte específica do corpo, como apenas o corpo inteiro, apenas o pronoto ou apenas a lateral do tórax.

4.4 Análise de dados

O desempenho da rede AlexNet foi avaliado por meio da análise de matrizes de confusão, bem como do cálculo da acurácia geral e da acurácia específica. A acurácia geral refere-se à probabilidade de identificar corretamente as espécies de forma independente, enquanto a acurácia específica, que pode ser interpretada como sensibilidade, representa a probabilidade de identificar corretamente cada espécie individualmente, dependendo do contexto de análise do algoritmo em cada experimento. Essa abordagem permitiu uma avaliação detalhada e quantitativa do modelo, destacando sua capacidade de identificação precisa, tanto em termos gerais quanto para cada espécie analisada nos experimentos. As métricas de desempenho foram calculadas utilizando o pacote 'Hmisc' no software estatístico R (versão 4.7.1), integrado à interface RStudio (versão 2024.03.1.446) (RSTUDIO TEAM., 2024; R CORE TEAM., 2024; HARRELL., 2014). Para garantir a precisão dos resultados, as análises incluíram o cálculo da média e do intervalo de confiança (binomial de Wilson com score de 95%) para cada métrica. As seguintes equações foram empregadas para determinar as métricas de desempenho da rede.

$$\text{Acurácia geral} = \frac{VP + VN}{VP + FP + VN + FN}$$

$$\text{Acurácia específica} = \frac{VP}{TP + FN}$$

VP = Verdadeiros positivos; VN = Verdadeiros negativos; FP = Falsos positivos;
FN = Falsos negativos

5 RESULTADOS

5.1 Matriz de confusão

Para compreender o desempenho do algoritmo AlexNet, a **Figura 16** apresenta um conjunto de matrizes de confusão, destacando os padrões de identificação correta e incorreta das imagens de mosquitos para as 10 pseudo-replicatas realizadas. Essas matrizes permitem visualizar de forma clara como o modelo classificou as diferentes espécies e onde ocorreram erros, oferecendo insights detalhados sobre sua precisão e possíveis limitações.

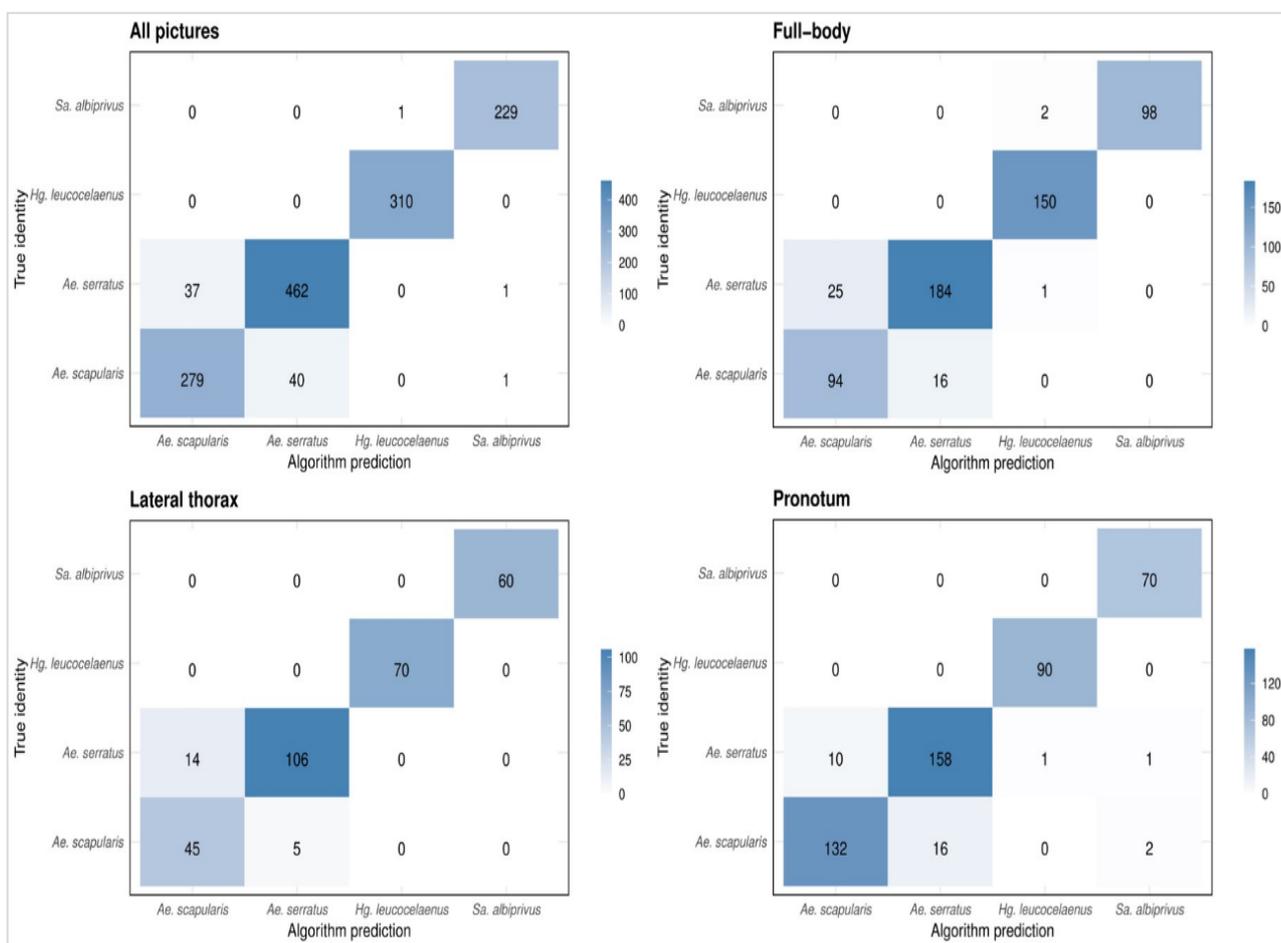


Figura 16 - Matrizes de confusão ilustrando os acertos e erros de classificação para os 10% das imagens utilizadas no conjunto de teste em cada um dos quatro experimentos. Para cada experimento foi realizado 10 pseudo-replicatas.

5.2 Acurácia geral

O desempenho do algoritmo foi avaliado pelo processamento de imagens de teste após o treinamento da rede AlexNet. Quatro experimentos foram conduzidos, nos quais um dos experimentos foi utilizado todas as 1.343 imagens e experimentos com conjuntos de fotos separadamente (corpo, lateral do tórax e pronoto). Foi calculado a média e intervalo de confiança para acurácia geral. A rede AlexNet demonstrou maior desempenho na identificação de mosquitos quando todas as imagens foram analisadas juntas, fornecendo ao algoritmo informações mais completas sobre a morfologia dos mosquitos, alcançando acurácia média de 94% (IC de 95%, 90-97). Para os experimentos utilizado os conjuntos de fotos separadamente rede alcançou acurácia média de 92,3% (IC de 95%, 83-97) para identificar mosquitos a partir das imagens do corpo, para o conjunto de fotos com a vista lateral do torác a acurácia média foi de 93,7% (IC de 95%, 79-98). Finalmente, para o conjunto de imagens do pronoto, a acurácia média geral foi de 93,8% (IC de 95%,83-98). Os menores intervalos de confiança foram observados no experimento contendo todo o conjunto de imagens (Figura 17).

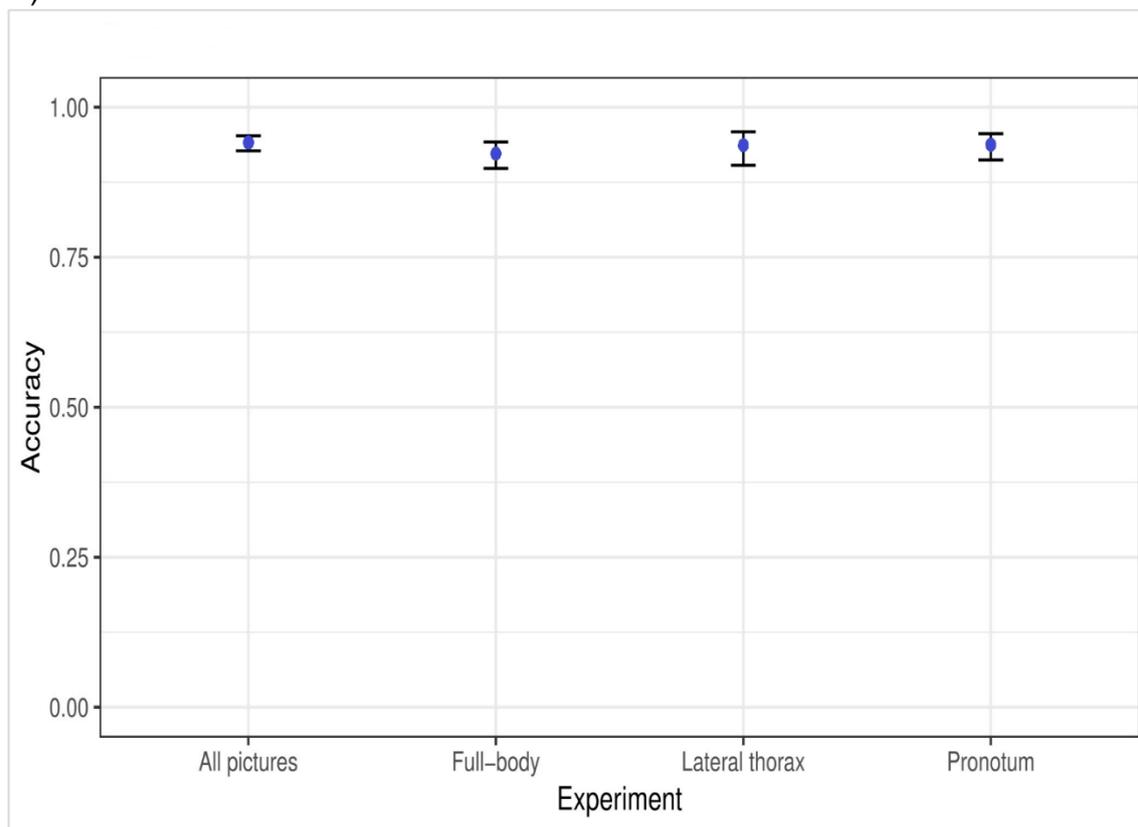


Figura 17 - Acurácia média geral (pontos) e intervalos de confiança de 95% (barras) do desempenho da rede AlexNet para identificação de mosquitos vetores da febre amarela.

5.3 Acurácia específica

Em seguida, avaliamos o desempenho do algoritmo para cada espécie com base nos três conjuntos de imagens. O melhor desempenho da rede foi observado quando as imagens dos três conjuntos foram combinadas no mesmo experimento. O desempenho da rede AlexNet na identificação de mosquitos do gênero *Aedes* foi menor, com os maiores intervalos de confiança observados para as duas espécies do gênero, em comparação com as outras (IC de 95%, 72-98) (Figura 18).

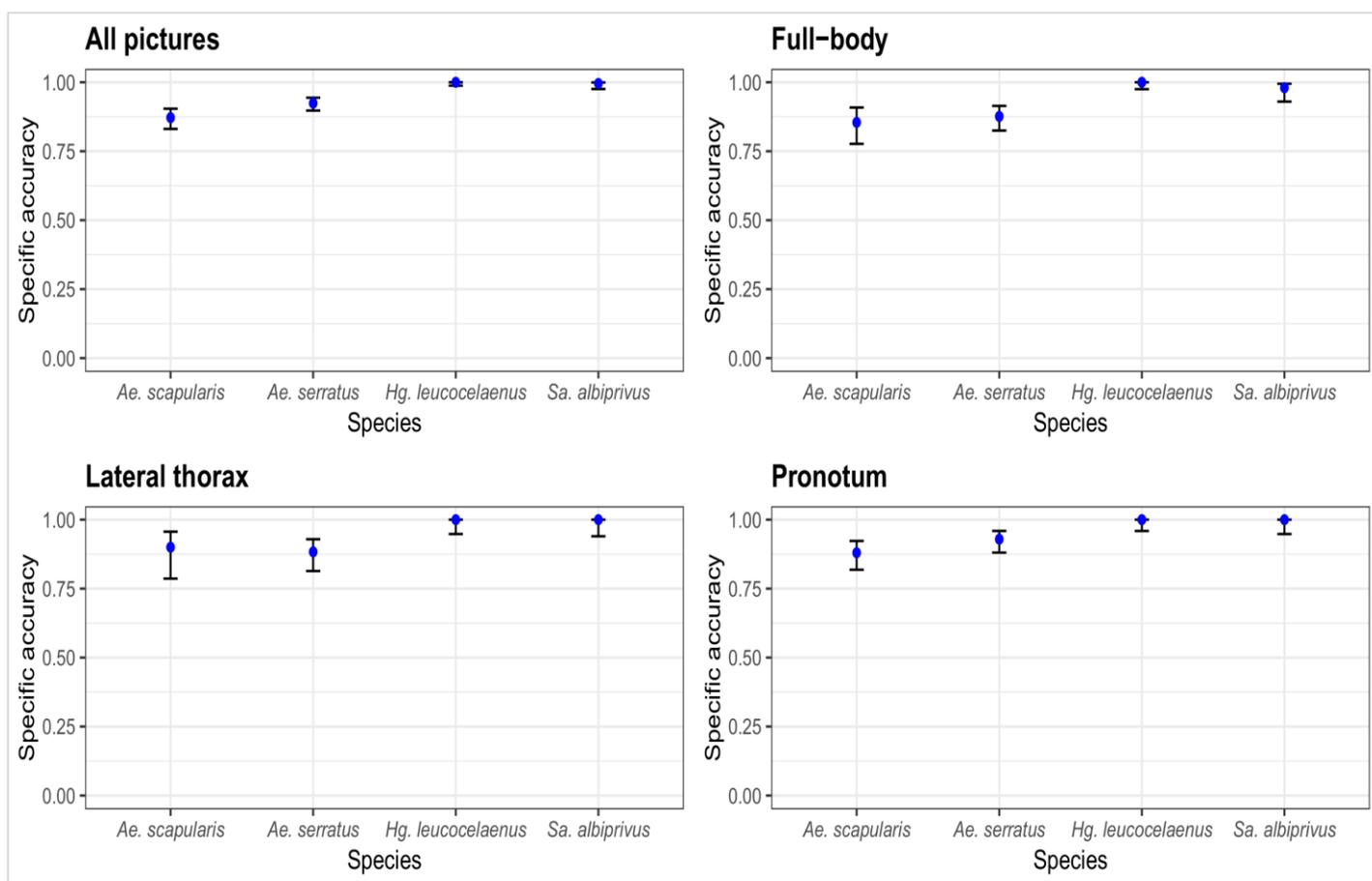


Figura 18 - Acurácia média do algoritmo AlexNet na identificação interespecífica de espécies. Os experimentos levaram em consideração a identificação a partir de imagens de regiões isoladas de partes do corpo, com intervalos de confiança de 95%.

No experimento em que todas as 1.343 imagens estavam disponíveis, a rede identificou incorretamente pelo menos uma imagem de *Ae. scapularis* como *Ae. serratus* ou vice e versa. *Ae. serratus* também foi identificado incorretamente como *Hg. leucocelaenus*. No experimento com imagens da vista lateral do tórax, apenas mosquitos do gênero *Aedes* foram identificados incorretamente, com as duas espécies de *Aedes* sendo trocadas. No experimento com imagens do pronoto, ocorreram várias identificações incorretas. *Ae. serratus* foi identificado incorretamente como *Sa. albiprivus* ou *Hg. leucocelaenus*, enquanto *Ae. scapularis* foi identificado como *Sa. albiprivus*. Vale ressaltar que muitos dos espécimes fotografados não estavam bem preservados, o que pode ter contribuído para identificação incorretas (Figura 19).



Figura 19- Exemplos de imagens identificadas incorretamente, devido a espécimes danificados. (A e H) *Ae. serratus* foi confundido com *Ae. scapularis*. (D e E) *Ae. scapularis* foi confundido com *Ae. serratus*. (F) *Ae. serratus* foi confundido com *Sa. albiprivus*. (G) *Sa. albiprivus* foi confundido com *Hg. leucocelaenus*. (I) *Ae. scapularis* foi identificado erroneamente como *Hg. leucocelaenus*.

6 DISCUSSÃO

Este estudo, teve como objetivo principal identificar quatro espécies de mosquitos que transmitem febre amarela ou outros arbovírus aplicando a rede neural convolucional AlexNet. Além disso, investigou-se também se o desempenho da rede na classificação dos mosquitos muda de acordo com as regiões do corpo mostradas nas imagens para treinamento e teste. Os resultados demonstraram que o modelo de rede AlexNet pode identificar com precisão fotos de mosquitos do gênero *Aedes*, *Sabethes* e *Haemagogus* com mais de 90% de acurácia. Contudo, as análises não evidenciaram diferenças significativas no desempenho do modelo quando utilizado imagens diferentes regiões isoladas do corpo dos mosquitos.

LORENZ et al. (2015) classificaram mosquitos com base nas características morfométricas de suas asas usando redes neurais, alcançando acurácias variando de 86% a 100%. SAUER et al. (2024) desenvolveram um modelo de CNN que foi treinada para diferenciar entre mosquitos *Aedes* e não-*Aedes* e classificar as sete espécies de *Aedes* com base em imagens de asas em tons de cinza e RGB. A CNN de melhor desempenho alcançou acurácia de 99% para discriminar espécies de mosquitos *Aedes* e não *Aedes*, 90% de acurácia pra identificar espécies a partir de imagens em tons de cinza e 91 % de acurácia para imagens RGB. Apesar desses estudos concluírem que imagens de asas são suficientes para identificar espécies de mosquitos com CNNs. No entanto, um sistema de identificação baseado apenas em uma estrutura corporal como a asa é muito frágil, pois se a estrutura não estiver presente na foto analisada, a identificação fica comprometida. Portanto, um bom sistema de identificação deve funcionar com qualquer parte do corpo do inseto. MOTTA et al (2019) usaram três redes pré-treinadas para identificar mosquitos urbanos (*Aedes* e *Culex*) alcançando uma precisão de 76,2% para o GoogleNet, 52,4% para o LeNet e 51,2% para o AlexNet. Posteriormente os mesmos autores otimizando os hiperparâmetros das redes CNN e obtiveram 97,3% de precisão na distinção entre os mosquitos do gênero *Aedes* e *Culex* (MOTTA et al., 2020).

PARK et al. (2020) utilizando um banco de dados com cerca de 3.600 imagens de mosquitos fotografados em várias posições e condições de

deformação, avaliaram três CNNs: VGG-16, ResNet-50 e SqueezeNet. Essas redes alcançaram acurácias de 97%, 96% e 90%, respectivamente, para identificar oito espécies de mosquitos do gênero *Aedes*, *Anopheles*, *Culex*, apesar da alta similaridade entre espécies e das variações intraespécies. Os resultados indicaram que mais de 90% de acurácia pode ser alcançada quando são aplicadas técnicas de ajustes finos e aumento de dados. O mesmo resultado em relação ao aumento de dados foi encontrado por OKAYASU et al. (2019), chegando a melhores acurácias para identificar *Ae. albopictus*, *Anopheles stefensi* e *Cx. pipiens pallens* usando AlexNet (92,3%) e Resnet (95,5).

KITTICHAJ et al. (2021) desenvolveram um modelo utilizando duas CNNs do tipo YOLO v3 integradas capazes de classificar e localizar imagens simultaneamente, o modelo foi treinado para identificar espécies de mosquitos capturados em campo. O modelo identificou *Ae. aegypti*, *Ae. albopictus* e *Cx. quinquefasciatus* com acurácias médias de 98-100%. PORA et al. (2023) aprimoraram a arquitetura VGG16 para torná-la mais eficiente para lidar com imagens de mosquitos aplicando a técnica de dropout adaptando a rede para lidar com entradas de múltiplas visões. Esses ajustes visam garantir um modelo funcione bem em cenários reais e diversos. No nosso estudo a rede AlexNet identificou *Ae. scapularis*, *Ae. serratus*, *Ha. leucocelaenus* e *Sa. albiprivus* com acurácia média de 94% utilizando o conjunto total de imagens. Em comparação com estudos anteriores que utilizaram redes neurais para a identificação de mosquitos, nossos resultados são consistentes com a maioria dos resultados observados por outros autores.

Em nosso estudo, não encontramos uma diferença significativa no desempenho do AlexNet na identificação de mosquitos com base em diferentes regiões do corpo. Em outros estudos utilizando técnicas de ativação de mapas de calor, já foram observados que as áreas que as CNNs se concentram para a classificação da imagem correspondem às mesmas utilizadas na identificação taxonômica por métodos tradicionais, conseguindo detectar diferenças morfológicas em várias regiões do corpo, como cabeça, pernas e tórax (PARK et al., 2020; ADHANE et al., 2021; ZHAO et al., 2022). Mostrando que os modelos de CNNs aprendem a partir das mesmas características utilizadas por taxonomistas. Por exemplo, *Ae. scapularis* pode ser diferenciado dos demais

Aedes pelo seu abdômen serrilhado, probóscide maior que o fêmur anterior e o mesonoto com escamas claras formando um círculo. *Ae. serratus* é identificado por seu abdômen serrilhado, probóscide semelhante ou menor que o fêmur anterior e mesonoto que pode ou não ter uma faixa longitudinal de escamas claras (SEGURA & CASTRO, 2007). As duas espécies são muito semelhantes porque pertencem ao subgênero *Ochlerotatus* (Lynch Arribálzaga, 1891), onde a maioria das espécies deste grupo são indistinguíveis com base em caracteres morfológicos e fazem parte de complexos (REINERT et al 2009). *Sa. albiprivus* possuem remos nas pernas medianas com escamas azuladas, abdômen com escamas douradas formando quadriláteros e uma probóscide muito menor que o fêmur anterior. *Sa. albiprivus* e *Hg.leucocelaenus* são duas espécies com características morfológicas bem diferentes. *Hg. leucocelaenus* pode ser distinguido de *Sa. albiprivus* por sua cor predominantemente opaca e escura e pleura com duas linhas verticais de escamas prateadas (SEGURA & CASTRO., 2007). Outros estudos mostram que a precisão das CNNs na identificação de outros insetos não é significativamente afetada pela região do corpo mostrada na imagem (MIRANDA et al., 2024). Nossos achados mostram que as características morfológicas usadas para a identificação dos mosquitos incluídos neste estudo estão presentes em múltiplas regiões do corpo e, portanto, qualquer uma das regiões do corpo aqui estudadas permitiu que a rede AlexNet identificasse com precisão as espécies de mosquitos.

Redes neurais de aprendizado profundo consistem em múltiplas camadas convolucionais que extraem características hierárquicas dos dados, e conjuntos de dados com maior quantidade e diversidade de imagens favorecem o treinamento, permitindo que a rede aprenda padrões mais representativos e complexos (KRIZHEVSKY; SUTSKEVER; HINTON., 2012). Além disso, muitos estudos indicam que um banco de imagens maior melhora o desempenho do algoritmo (PARK et al., 2020; PATAKI et al., 2021; MOTTA et al., 2020, 2021; MIRANDA et al. 2024; GURGEL-GONÇALVES et al., 2017, KHALIGHIFAR et al., 2019). Embora um banco de dados com milhares de imagens seja sempre desejado, usando um banco de dados com apenas 1.343 imagens, alcançamos taxas de acurácias semelhantes às que usam bancos de dados 10x maiores que o nosso (MOTTA et al., 2020; OKAYASU et al., 2019). A precisão do AlexNet

para identificar mosquitos *Sabethes* e *Haemogogus* foi semelhante à precisão obtida com outras CNNs usadas para identificar outros gêneros (MOTTA et al., 2020; OKAYASU et al., 2019). No entanto, a precisão do AlexNet na identificação de *Ae. serratus* e *Ae. scapularis* ficou abaixo de 90% e, portanto, abaixo do ideal quando comparado com o desempenho de outras CNNs (VGG-16, ResNet-50, SqueezeNet) que aplicam técnicas de aumento de dados e ajuste fino para identificar *Ae. aegypti* (MOTTA et al., 2019, 2020), *Ae. albopictus* e *Ae. vexas* (PARK et al., 2020)

Esses valores de precisão podem ser devidos a diferenças na arquitetura e treinamento do algoritmo (PARK et al., 2020). O AlexNet é relativamente superficial em comparação com algoritmos mais profundos, como ResNet ou DenseNet, que podem capturar recursos hierárquicos complexos nas imagens. Embora o AlexNet tenha sido projetado para aproveitar a aceleração da GPU, sua eficiência computacional é menor em comparação com outros algoritmos, como o MobileNet. Finalmente, os recursos aprendidos pelo AlexNet podem não ser tão discriminativos quanto aqueles aprendidos por algoritmos treinados com grandes conjuntos de dados. Apesar das limitações do AlexNet quando comparado a outros algoritmos, nossos resultados mostram alta precisão para identificação de mosquitos da febre amarela. Nossa perspectiva é aplicar outros algoritmos (ResNet, DenseNet ou MobileNet) para testar sua eficiência na identificação de *Ae. serratus*, *Ae. scapularis* e outras espécies de mosquitos morfológicamente similares. Além disso, a otimização dos hiperparâmetros da CNN para aumentar a precisão da identificação dos mosquitos (MOTTA et al., 2020).

O menor desempenho do algoritmo em alguns casos pode ter sido influenciado pelo estado de preservação dos espécimes. A análise das imagens identificadas incorretamente em todos os experimentos mostrou que os espécimes fotografados não estavam bem preservados, especialmente na área do pronoto, onde faltavam cerdas e escamas, assim como as pernas. Devido ao seu tamanho e à presença de escamas e cerdas, os mosquitos são facilmente danificados durante a captura, congelamento e secagem, resultando na perda de características morfológicas críticas necessárias para a identificação adequada. O estado de preservação dos mosquitos foi um fator limitante neste

trabalho pois a boa preservação dos espécimes é importante para o desempenho ideal do algoritmo. Além disso, a rede não foi avaliada com imagens capturadas por diferentes dispositivos fotográficos. Portanto, não é possível dizer se imagens tiradas por câmeras com especificações diferentes têm efeito no desempenho da rede AlexNet.

Neste estudo, identificamos com sucesso quatro espécies de mosquitos vetores de febre amarela e outras arboviroses aplicando a rede neural convolucional AlexNet com acurácias superiores a 90%. Esses resultados indicam o grande potencial da inteligência artificial na vigilância entomológica permitindo rápida identificação e resposta a situações que necessitam de vigilância e controle reduzindo os impactos das doenças transmitidas por mosquitos.

7 CONCLUSÃO

O objetivo deste estudo foi identificar quatro espécies de mosquitos responsáveis pela transmissão de febre amarela e outros arbovírus, utilizando a rede neural convolucional AlexNet. Os resultados indicaram que o modelo AlexNet é capaz de identificar imagens de 4 espécies mosquitos dos gêneros *Aedes* (2), *Sabethes* (1) e *Haemagogus* (1) com acurácias superiores a 90%, independentemente da região do corpo do mosquito apresentada nas imagens. Esses achados reforçam a viabilidade do uso de redes neurais convolucionais na identificação de mosquitos vetores de doenças tropicais. Essa descoberta abre novas perspectivas para o uso de inteligência artificial na vigilância entomológica, facilitando a identificação rápida e precisa de mosquitos vetores. A aplicação dessa tecnologia em programas de saúde pública pode otimizar o monitoramento de vetores, agilizando a tomada de decisões e a implementação de medidas de controle mais eficazes. Adicionalmente, a disponibilização de ferramentas de identificação baseadas em redes neurais convolucionais pode auxiliar a população a reconhecer os principais vetores e a colaborar ativamente nas ações de combate às arboviroses.

REFERÊNCIAS

ABDELGHANI, B. A.; BANITAAN, S.; MALEKI, M.; MAZEN, A. Kissing bugs identification using convolutional neural network. **IEEE Access**, v. 9, p. 140539–140548, 2021.

ABREU, F. V. S. de et al. Ecological and environmental factors affecting transmission of sylvatic yellow fever in the 2017–2019 outbreak in the Atlantic Forest, Brazil. **Parasites & Vectors**, v. 15, n. 1, p. 1–18, 2022.

ABREU, F. V. S. de et al. *Haemagogus leucocelaenus* and *Haemagogus janthinomys* are the primary vectors in the major yellow fever outbreak in Brazil, 2016–2018. **Emerging Microbes & Infections**, v. 8, n. 1, p. 218–231, 2019.

ADHANE, G.; DEHSHIBI, M. M.; MASIP, D. A deep convolutional neural network for classification of *Aedes albopictus* mosquitoes. **IEEE Access**, v. 9, p. 72681–72690, 2021.

AKBARIAN, S. et al. A computer vision approach to identifying ticks related to Lyme disease. **IEEE Journal of Translational Engineering in Health and Medicine**, v. 10, 2022.

ANDRADE, M. D. S. et al. Real-time genomic surveillance during the 2021 re-emergence of the yellow fever virus in Rio Grande do Sul State, Brazil. **Viruses**, v. 13, n. 10, p. 1976, 2021.

ANDRADE, M. S.; CAMPOS, F. S.; OLIVEIRA, C. H. D.; OLIVEIRA, R. S.; CAMPOS, A. A. S.; ALMEIDA, M. A. B. D.; ABREU, F. V. S. D. Fast surveillance response reveals the introduction of a new yellow fever virus sub-lineage in 2021, in Minas Gerais, Brazil. **Memórias do Instituto Oswaldo Cruz**, Rio de Janeiro, v. 117, p. e220127, 2022.

APASRAWIROTE, D. et al. Assessment of deep convolutional neural network models for species identification of forensically important fly maggots based on images of posterior spiracles. **Scientific Reports**, v. 12, n. 1, p. 4753, 2022.

ARISTA-JALIFE, A. et al. Aedes mosquito detection in its larval stage using deep neural networks. **Knowledge-Based Systems**, v. 189, p. 104841, 2020.

ASMAI, S. et al. Mosquito larvae detection using deep learning. **International Journal of Innovative Technology and Exploring Engineering**, v. 8, n. 12, p. 804–809, 2019.

AZIZ, L. et al. Exploring deep learning-based architecture, strategies, applications and current trends in generic object detection: A comprehensive review. **IEEE Access**, v. 8, p. 170461–170495, 2020.

BORKENT, A.; BELTON, P. Attraction of female *Uranotaenia lowii* (Diptera: Culicidae) to frog calls in Costa Rica. **Canadian Entomologist**, v. 138, p. 91-94, 2006.

BRASIL. Ministério da Saúde. Nota informativa conjunta: alerta sobre febre amarela. **Brasília: Ministério da Saúde**, 2024. Disponível em: <https://www.gov.br/saude/pt-br/centrais-de-conteudo/publicacoes/estudos-e-notas-informativas/2024/nota-informativa-conjunta-alerta-sobre-febre-amarela.pdf>. Acesso em: 05 de jan de 2025.

CAMARGO-NEVES, V. L. F. D. et al. Entomological investigation of a sylvatic yellow fever area in São Paulo State, Brazil. **Cadernos de Saúde Pública**, v. 21, n. 4, p. 1278–1286, 2005.

CANNET, A. et al. Species identification of phlebotomine sandflies using deep learning and wing interferential pattern (WIP). **Scientific Reports**, v. 13, n. 1, p. 21389, 2023.

CANO, M. E. et al. Categorization by score of mosquito species (Diptera: Culicidae) related to yellow fever epizootics in Argentina. **Journal of Medical Entomology**, v. 59, n. 5, p. 1384–1388, 2022.

CARDOSO, J. C. et al. Yellow fever virus in *Haemagogus leucocelaenus* and *Aedes serratus* mosquitoes, Southern Brazil, 2008. **Emerging Infectious Diseases**, v. 16, n. 12, p. 1918–1924, 2010.

CATOR, L. J.; WYER, C. A. S.; HARRINGTON, L. C. Mosquito sexual selection and reproductive control programs. **Trends in Parasitology**, [S.l.], v. 37, n. 4, p. 330-339, 2021.

CHENG, X. et al. Pest identification via deep residual learning in complex background. **Computers and Electronics in Agriculture**, v. 141, p. 351–356, 2017.

CICEKLI, V.; AKMAN, A. P.; SAYGIN, I. et al. Turing test: 50 years later. **Minds and Machines**, v. 10, p. 463–518, 2000.

COCHERO, J.; PATTORI, L.; BALSALOBRE, A.; CECCARELLI, S.; MARTI, G. A convolutional neural network to recognize Chagas disease vectors using mobile phone images. **Eco Inform**, v. 68, p. 101587, 2022.

CONSOLI, R. A. G. B.; OLIVEIRA, R. L. Principais mosquitos de importância sanitária no Brasil. 1. ed. Rio de Janeiro: **Editora Fiocruz**, 1994.

CUNHA, M. S.; et al. Genomic evidence of yellow fever virus in *Aedes scapularis*, southeastern Brazil, 2016. **Acta Tropica**, v. 205, p. 105390, 2020.

DE SWART, M. M.; BALVERS, C.; VERHULST, N. O.; KOENRAADT, C. J. M. Effects of host blood on mosquito reproduction. **Trends in Parasitology**, v. 39, n. 7, p. 575-587, 2023.

DAMASCENO-CALDEIRA, Rossela; et al. Vector competence of *Aedes albopictus* for yellow fever virus: Risk of reemergence of urban yellow fever in Brazil. **Viruses**, v. 15, n. 4, p. 1019, 2023.

DEO, R. C. Machine learning in medicine. **Circulation**, v. 132, n. 20, p. 1920–1930, 2015.

DE OLIVEIRA, C. H. et al. Yellow fever virus maintained by *Sabethes* mosquitoes during the dry season in Cerrado, a semiarid region of Brazil, in 2021. **Viruses**, v. 15, n. 3, p. 757, 2023.

ESTEVA, A ; et al. Dermatologist-level classification of skin cancer with deep neural networks. **Nature**, v. 542, n. 7639, p. 115–118, 2017.

FRAIWAN, Mohammad. Efficient Convolutional Neural Network Model for the Taxonomy and Sex Identification of Three Phlebotomine Sandfly Species (Diptera, Psychodidae, and Phlebotominae). **Animals**, v. 14, n. 24, p. 3712, 2024.

FERNANDES, M. S.; CORDEIRO, W.; RECAMONDE-MENDOZA, M. Detecting *Aedes aegypti* mosquitoes through audio classification with convolutional neural networks. **Comput Biol Med**, v. 129, p. 104152, 2021.

FORATTINI, O. P.; GOMES, A. de C.; SANTOS, J. L. F.; GALATI, E. A. B.; RABELLO, E. X.; NATAL, D. Observações sobre atividade de mosquitos Culicidae em mata residual no Vale do Ribeira, São Paulo, Brasil. **Rev. Saúde Pública**, v. 15, p. 557–586, 1981.

FORATTINI, O. P. Culicidologia médica. São Paulo: **Edusp**, 2002.

GARCIA-OLIVEIRA, G; et al. Yellow alert: Persistent Yellow Fever Virus Circulation among Non-Human Primates in Urban Areas of Minas Gerais State, Brazil (2021–2023). **Viruses**, v. 16, n. 1, p. 31, 2023.

GIOVANETTI, Marta; et al. Genomic epidemiology unveils the dynamics and spatial corridor behind the Yellow Fever virus outbreak in Southern Brazil. **Science Advances**, v. 9, n. 35, p. 9204, 2023.

GIOVANETTI, Marta; et al. Yellow fever virus reemergence and spread in Southeast Brazil, 2016–2019. **Journal of Virology**, v. 94, n. 1, p. 1–12, 2019.

GOHE, A. K; et al. Classifying forensically important flies using deep learning to support pathologists and rescue teams during forensic investigations. **PloS One**, v. 19, n. 12, p. e0314533, 2024.

GOODWIN, A. et al. Mosquito species identification using convolutional neural networks with a multitiered ensemble model for novel species detection. **Scientific reports**, v. 11, n. 1, p. 13656, 2021.

GOMES, A. C.; TORRES, M. A. N.; PAULA, M. B.; FERNANDES, A.; MARASSÁ, A. M.; CONSALES, C. A.; FONSECA, D. F. Ecologia de *Haemagogus* e *Sabethes* (Diptera: Culicidae) em áreas epizoóticas do vírus da febre amarela, Rio Grande do Sul, Brasil. **Epidemiologia e Serviços de Saúde**, v. 19, p. 101-113, 2010.

GOMEZ VILLA, A.; SALAZAR, A.; VARGAS, F. Towards automatic wild animal monitoring: Identification of animal species in camera-trap images using very deep convolutional neural networks. **Ecological Informatics**, v. 41, p. 24–32, 2017.

GULSHAN, V; et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. **JAMA**, v. 316, n. 22, p. 2402–2410, 2016.

GUPTA, K. A; et al. Mosquito species identification accuracy of early deployed algorithms in IDX, A vector identification tool. **Acta Tropica**, v. 260, p. 107392, 2024.

GURGEL-GONÇALVES, R.; KOMP, E.; CAMPBELL, L. P.; et al. Automated identification of insect vectors of Chagas disease in Brazil and Mexico: The Virtual Vector Lab. **PeerJ**, v. 5, p. e3040, 2017.

GURGEL-GONÇALVES, R; et al. TRIATODEX, an electronic identification key to the Triatominae (Hemiptera: Reduviidae), vectors of Chagas disease: Development, description, and performance. **PloS One**, v. 16, n. 4, p. e0248628, 2021.

HANLEY, K; et al. Fever versus fever: The role of host and vector susceptibility and interspecific competition in shaping the current and future distributions of the sylvatic cycles of dengue virus and yellow fever virus. *Infection*, **Genetics and Evolution**, v. 19, p. 292–311, 2013.

HARBACH, R. E. **Mosquito taxonomic inventory**. Disponível em: <http://mosquito-taxonomic-inventory.info/simpletaxonomy/term/6231>. Acesso em: 24 fev. 2025.

HARBACH, Ralph E. The Culicidae (Diptera): A review of taxonomy, classification and phylogeny. **Zootaxa**, v. 1668, n. 1, p. 591–638, 2007.

HARBACH, R. E.; KITCHING, Ian J. Phylogeny and classification of the Culicidae (Diptera). **Systematic Entomology**, v. 23, n. 4, p. 327–370, 1998.

HARRELL, F. E. Hmisc: A package of miscellaneous R functions. 2014. Disponível em: <http://biostat.mc.vanderbilt.edu/Hmisc>.

HAWKES, F. M.; HOPKINS, R. J. The mosquito: an introduction. In: HALL, M.; TAMİR, D. (ed.). *Mosquitopia: the place of pests in a healthy world* [Internet]. New

York: **Routledge**, 2022. Chapter 2. Disponível em: <https://www.ncbi.nlm.nih.gov/books/NBK585164/>. Acesso em: 21 jan. 2025.

JABALLAH, Sami; et al. A deep learning approach to detect and identify live freshwater macroinvertebrates. *Aquatic Ecology*, v. 57, n. 4, p. 933–949, 2023.

KAUR, Inderpreet; SANDHU, Amanpreet Kaur; KUMAR, Yogesh. Artificial intelligence techniques for predictive modeling of vector-borne diseases and its pathogens: A systematic review. **Archives of Computational Methods in Engineering**, v. 29, n. 6, p. 3741–3771, 2022.

KHALIGHIFAR, A.; JIMÉNEZ-GARCÍA, D.; CAMPBELL, L. P.; et al. Application of deep learning to community-science-based mosquito monitoring and detection of novel species. **Journal of Medical Entomology**, v. 59, p. 355–362, 2021.

KHALIGHIFAR, A.; KOMP, E.; RAMSEY, J. M.; GURGEL-GONÇALVES, R.; PETERSON, A. T. Deep Learning Algorithms Improve Automated Identification of Chagas Disease Vectors. **Journal of Medical Entomology**, v. 56, n. 2, p. 461–468, 2019.

KISKIN, I.; ZILLI, D.; LI, Y.; SINKA, M.; WILLIS, K.; ROBERTS, S. Bioacoustic detection with wavelet-conditioned convolutional neural networks. **Neural Computing and Applications**, v. 32, p. 915–927, 2020.

KITTICHAJ, V.; PENGSAKUL, T.; CHUMCHUEN, K.; SAMUNG, Y.; SRIWICHAJ, P.; PHATTHAMOLRAT, N.; et al. Deep learning approaches for challenging species and gender identification of mosquito vectors. **Scientific Reports**, v. 11, n. 1, p. 19375, 2021.

KONONENKO, I. Machine learning for medical diagnosis: history, state of the art and perspective. **Artificial Intelligence in Medicine**, v. 23, n. 1, p. 89–109, 2001.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. ImageNet classification with deep convolutional neural networks. **Advances in Neural Information Processing Systems**, v. 25, p. 1106–1114, 2012.

LANE, J. **Neotropical Culicidae**. Volumes I & II. 1953.

LE, W. T.; et al. Overview of machine learning: part 2: deep learning for medical image analysis. **Neuroimaging Clinics of North America**, v. 30, n. 4, p. 417–431, 2020.

LECUN, Y; BENGIO, Y; HINTON, G. Deep learning. **Nature**, v. 521, n. 7553, p. 436–444, 2015.

LIU, S; et al. Early diagnosis of Alzheimer's disease with deep learning. In: 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI). **IEEE**, 2014. p. 1015–1018.

LORENZ, C.; FERRAUDO, A. S.; SUESDEK, L. Artificial Neural Network applied as a methodology of mosquito species identification. **Acta Tropica**, v. 152, p. 165–169, 2015.

LORENZ, C; SUESDEK, L. Evaluation of chemical preparation on insect wing shape for geometric morphometrics. **The American Journal of Tropical Medicine and Hygiene**, v. 89, n. 5, p. 928, 2013.

MARCONDES, C. B.; ALENCAR, J. Revisión de los mosquitos del género *Haemagogus* Williston (Diptera: Culicidae) de Brasil. **Revista Biomédica**, v. 21, n. 3, p. 221–238, 2010.

MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **The Bulletin of Mathematical Biophysics**, v. 5, p. 115–133, 1943.

MIOTTO, R; et al. Deep learning for healthcare: review, opportunities and challenges. **Briefings in Bioinformatics**, v. 19, n. 6, p. 1236–1246, 2018.

MONDET, B. et al. Isolation of yellow fever virus from nulliparous *Haemagogus (Haemagogus) janthinomys* in eastern Amazonia. **Vector Borne and Zoonotic Diseases**, v. 2, n. 1, p. 47–50, 2002. DOI:

MONTALBO, F. J. P. Automating mosquito taxonomy by compressing and enhancing a feature-fused EfficientNet with knowledge distillation and a novel residual skip block. **MethodsX**, v. 10, p. 102072, 2023.

MIRANDA, V. L 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. **Ecological Informatics**, v. 79, p. 102430, 2024.

MIRANDA, V. L. et al. Automated identification of Chagas disease vectors using AlexNet pre-trained convolutional neural networks. **Medical and Veterinary Entomology**, 2024b.

MORENO, E. S. et al. Reemergence of yellow fever: detection of transmission in the State of São Paulo, Brazil, 2008. **Revista da Sociedade Brasileira de Medicina Tropical**, v. 44, n. 3, p. 290–296, 2011.

MOTTA, D.; BANDEIRA SANTOS, A. Á.; SOUZA MACHADO, B. A.; VICENTE RIBEIRO-FILHO, O. G.; ARRIAGA CAMARGO, L. O.; VALDENEGRO-TORO, M. A.; et al. Optimization of convolutional neural network hyperparameters for automatic classification of adult mosquitoes. **PLoS ONE**, v. 15, n. 8, p. e0237151, 2020.

MOTTA, D.; SANTOS, A. Á. B.; WINKLER, I.; MACHADO, B. A. S.; PEREIRA, D. A. D. I.; CAVALCANTI, A. M.; et al. Application of convolutional neural

networks for classification of adult mosquitoes in the field. **PLoS ONE**, v. 14, n. 5, p. e0216427, 2019.

MUTHUKRISHNAN, N.; MALEKI, F.; OVENS, K.; REINHOLD, C.; FORGHANI, B.; FORGHANI, R. Brief history of artificial intelligence. **Neuroimaging Clinics of North America**, v. 30, n. 4, p. 393–399, 2020.

NAYAK, B.; KHUNTIA, B.; MURMU, L. K.; et al. Artificial intelligence (AI): a new window to revamp the vector-borne disease control. **Parasitology Research**, v. 122, p. 369–379, 2023.

NAVARRO, J. C.; MACHADO-ALLISON, C. Aspectos ecológicos de *Sabethes chloropterus* (Humboldt) (Diptera: Culicidae) em um bosque úmido do Edo. Miranda, Venezuela. **Boletín Entomológico Venezolano**, v. 10, p. 91-104, 1995.

NEVES, M. S. A. S.; MOTTA, M. A.; MACIEL-DE-FREITAS, R.; XAVIER, A. D. S.; LOURENÇO-DE-OLIVEIRA, R.; SILVA-DO-NASCIMENTO, T. F. Illustrated identification key to females of the genus *Sabethes* Robineau-Desvoidy recorded from Brazil (Diptera: Culicidae), in dichotomous and interactive formats, including an updated list of species and new records for the states. **Zootaxa**, v. 5406, n. 2, p. 253–287, 2024.

ONG, S. Q.; AHMAD, H.; NAIR, G.; ISAWASAN, P.; MAJID, A. H. A. Implementation of a deep learning model for automated classification of *Aedes aegypti* (Linnaeus) and *Aedes albopictus* (Skuse) in real time. **Scientific Reports**, v. 11, p. 9908, 2021.

ONG, S. Q.; NAIR, G.; YUSOF, U. K.; AHMAD, H. Community-based mosquito surveillance: an automatic mosquito-on-human-skin recognition system with a deep learning algorithm. **Pest Management Science**, v. 78, p. 4092–4104, 2022.

OBARA, M.T; et al. Infecção natural de *Haemagogus janthinomys* e *Haemagogus leucocelaenus* pelo vírus da febre amarela no Distrito Federal, Brasil, 2007-2008. **Epidemiologia e Serviços de Saúde**, v. 21, n. 3, p. 457–463, 2012.

OKAYASU, K.; YOSHIDA, K.; FUCHIDA, M.; NAKAMURA, A. Vision-based classification of mosquito species: Comparison of conventional and deep learning methods. **Applied Sciences**, v. 9, n. 18, p. 3818, 2019.

OLIVEIRA, Cirilo et al. Yellow Fever Virus Maintained by *Sabethes* Mosquitoes during the Dry Season in Cerrado, a Semi-arid Region of Brazil, in 2021. **Viruses**, v. 15, n. 3, p. 757, 2023.

OLIVEIRA, C. H. de; ANDRADE, M. S.; CAMPOS, F. S.; CARDOSO, J. da C.; GONÇALVES-DOS-SANTOS, M. E.; OLIVEIRA, R. S.; ABREU, F. V. S. de. Yellow fever virus maintained by *Sabethes* mosquitoes during the dry season in

Cerrado, a semiarid region of Brazil, in 2021. **Viruses**, Basel, v. 15, n. 3, p. 757, 2023.

ORTIZ, A. S.; MIYATAKE, M. N.; TÜNNERMANN, H.; TERAMOTO, T.; SHOUNO, H. Mosquito larva classification based on a convolution neural network. In: Proceedings of the International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA), Cholula, **Mexico: The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp)**, 2018. p. 320–325.

PARK, J.; KIM, D. I.; CHOI, B.; KANG, W.; KWON, H. W. Classification and Morphological Analysis of Vector Mosquitoes using Deep Convolutional Neural Networks. **Scientific Reports**, v. 10, n. 1, p. 3144, 2020.

PARSONS, Z.; BANITAAN, S. Automatic identification of Chagas disease vectors using data mining and deep learning techniques. **Ecological Informatics**, v. 62, p. 101270, 2021.

PATAKI, B. A. et al. Deep learning identification for citizen science surveillance of tiger mosquitoes. **Scientific Reports**, v. 11, p. 4718, 2021.

PEREIRA DOS SANTOS, et al. Potential of *Aedes albopictus* as a bridge vector for enzootic pathogens at the urban-forest interface in Brazil. **Emerging Microbes & Infections**, v. 7, n. 1, p. 1-8, 2018.

PHAM, T. et al. Deepcare: A deep dynamic memory model for predictive medicine. In: Advances in Knowledge Discovery and Data Mining: 20th Pacific-Asia Conference, PAKDD 2016, Auckland, New Zealand, April 19-22, 2016, Proceedings, Part II. **Springer International Publishing**, 2016. p. 30-41.

PORA, W. et al. Enhancement of VGG16 model with multi-view and spatial dropout for classification of mosquito vectors. **PLoS One**, v. 18, n. 7, p. e0253793, 2023.

POSSAS, C. et al. Yellow fever outbreak in Brazil: the puzzle of rapid viral spread and challenges for immunisation. **Memórias do Instituto Oswaldo Cruz**, v. 113, n. 10, p. e180278, 2018.

QASMI, N.; BIBI, R.; RASHID, S. Recognition of Conus species using a combined approach of supervised learning and deep learning-based feature extraction. **PLoS One**, v. 19, n. 12, p. e0313329, 2024.

QIN, H. et al. DeepFish: Accurate underwater live fish recognition with a deep architecture. **Neurocomputing**, v. 187, p. 49–58, 2016.

R DEVELOPMENT CORE TEAM. R: a language and environment for statistical computing. Version 4.4.0. **Vienna: R Foundation for Statistical Computing**, 2024. Disponível em : <https://www.R-project.org>.

REEVES, L. E. et al. Identification of *Uranotaenia sapphirina* as a specialist of annelids broadens known mosquito host use patterns. **Communications biology**, v. 1, n. 1, p. 92, 2018.

RAHMAN,R; UDDIN, J. A deep learning-based dengue mosquito detection method using faster R-CNN and image processing techniques. **Annals of Emerging Technologies in Computing (AETiC)**, v. 5, n. 3, 2021

RAIZADA, S.; MALA, S.; SHANKAR, A. Vector-borne disease outbreak prediction using machine learning techniques. In: **Advanced Deep Learning for Engineers and Scientists: A Practical Approach**. p. 227-241, 2021.

REINERT, J. F.; HARBACH, R. E.; KITCHING, I. J. Phylogeny and classification of tribe Aedini (Diptera: Culicidae). **Zoological Journal of the Linnean Society**, v. 157, p. 700–794, 2009.

RSTUDIO. Integrated Development for R. Version 4.4.0. **Vienna: R Foundation for Statistical Computing**, 2024. Disponível: <https://posit.co/products/open-source/rstudio/>.

SABATTINI, M. S.; AVILÉS, G.; MONATH, T. P. Historical, epidemiological, and ecological aspects of arboviruses in Argentina: *Togaviridae*, *Alphavirus*. In: An overview of arbovirology in Brazil and neighbouring countries. **Belém: Instituto Evandro Chagas**, 1998. p. 135–153.

SACCHETTO, L. et al. Neighbor danger: Yellow fever virus epizootics in urban and urban-rural transition areas of Minas Gerais state, during 2017-2018 yellow fever outbreaks in Brazil. **PLoS Neglected Tropical Diseases**, v. 14, n. 10, p. e0008658, 2020.

SÃO PAULO (Estado). Secretaria de Estado da Saúde. *Boletim Epidemiológico de Febre Amarela – 01/2025*. **São Paulo: Secretaria de Estado da Saúde**, 2025. Disponível em: <[Boletim Epidemiológico 01_2025_17.02.25](#)>. Acesso em: 18 fev 2025.

SANT'ANA, A. L.; LOZOVEI, A. L. Influência do ciclo lunar na captura de *Aedes scapularis* (Diptera, Culicidae) na Mata Atlântica do Paraná. **Iheringia, Série Zoologia**, v. 90, p. 175-182, 2001.

SAUER, F. G. et al. A convolutional neural network to identify mosquito species (Diptera: Culicidae) of the genus *Aedes* by wing images. **Scientific Reports**, v. 14, p. 14147, 2024.

SEGURA, M. N. O.; CASTRO, F. C. Atlas de culicídeos na Amazônia brasileira: características específicas de insetos hematófagos da família Culicidae. **Belém: Instituto Evandro Chagas**, 2007.

SHI, L. et al. Establishment of a deep learning-based visual model for intelligent recognition of *Oncomelania hupensis*. **Chinese Journal of Schistosomiasis Control**, v. 33, p. 445–451, 2021.

SILVA, N. I. O. et al. Recent sylvatic yellow fever virus transmission in Brazil: the news from an old disease. **Virology Journal**, v. 17, p. 1–12, 2020.

SIMOVIĆ, Predrag et al. Automated identification of aquatic insects: A case study using deep learning and computer vision techniques. **Science of The Total Environment**, v. 935, p. 172877, 2024.

SPIESMAN, B. J.; GRATTON, C.; GRATTON, E.; HINES, H. Deep learning for identifying bee species from images of wings and pinned specimens. **PLoS One**, v. 19, n. 5, p. e0303383, 2024.

SPIESMAN, B. J. et al. Assessing the potential for deep learning and computer vision to identify bumble bee species from images. **Scientific Reports**, v. 11, n. 1, p. 7580, 2021.

STARK, Thomas et al. YOLO object detection models can locate and classify broad groups of flower-visiting arthropods in images. **Scientific Reports**, v. 13, n. 1, p. 16364, 2023.

TALLAM, K. et al. Identification of snails and *Schistosoma* of medical importance via convolutional neural networks. **The American Journal of Tropical Medicine and Hygiene**, v. 105, p. 295, 2021.

TRAVASSOS-DA-ROSA, A. P. et al. Trocara virus: a newly recognized *Alphavirus* (*Togaviridae*) isolated from mosquitoes in the Amazon Basin. **The American Journal of Tropical Medicine and Hygiene**, v. 64, n. 1, p. 93-97, 2001.

TRAVASSOS-DA-ROSA, A. P. A.; VASCONCELOS, P. F. C.; JORGE, F. S. An overview of arbovirology in Brazil and neighbouring countries. **Belém: Instituto Evandro Chagas**, 1998.

TUELLS, Jose et al. The perennial threat of yellow fever. **Archives of Medical Research**, 2022.

UNG, V. et al. Xper2: introducing e-taxonomy. **Bioinformatics**, v. 26, n. 5, p. 703-704, 2010.

VALAN, M. et al. Automated taxonomic identification of insects with expert-level accuracy using effective feature transfer from convolutional networks. **Systematic Biology**, v. 68, n. 6, p. 876-895, 2019.

VASCONCELOS, P. F. C. et al. An epidemic of sylvatic yellow fever in the southeast region of Maranhão State, Brazil, 1993–1994: epidemiologic and

entomologic findings. **The American Journal of Tropical Medicine and Hygiene**, v. 57, n. 2, p. 132-137, 1997.

VASCONCELOS, P. F. C. et al. Febre amarela: reflexões sobre a doença, as perspectivas para o século XXI e o risco da reurbanização. **Revista Brasileira de Epidemiologia**, v. 5, n. 3, p. 244–258, 2002.

VIGNES-LEBBE, R.; CHESSELET, P.; DIEP THI, M. H. Xper3: new tools for collaborating, training and transmitting knowledge on botanical phenotypes. In: RAKOTOARISOA, N. R.; BLACKMORE, S.; RIERA, B. (org.). *Botanists of the twenty-first century: roles, challenges and opportunities*. Paris: **UNESCO**, 2016. p. 228–239.

WÄLDCHEN, J ; MÄDER, P. Machine learning for image-based species identification. **Methods in Ecology and Evolution**, v. 9, n. 11, p. 2216-2225, 2018.

WALTER, D. E.; WINTERTON, S. Keys and the crisis in taxonomy: extinction or reinvention? *Annual Review of Entomology*, v. 52, p. 193-208, 2007.

WILKERSON, R. C. et al. Making mosquito taxonomy useful: a stable classification of tribe Aedini that balances utility with current knowledge of evolutionary relationships. **PLoS One**, v. 10, n. 3, p. e0133602, 2015.

WEI, X; HOSSAIN, M. Z; AHMED, K. A. A ResNet attention model for classifying mosquitoes from wing-beating sounds. **Scientific Reports**, v. 12, n. 1, p. 10334, 2022.

YUVAL, B. Mating systems of blood-feeding flies. **Annual Review of Entomology**, v. 51, n. 1, p. 413-440, 2006.

ZHAO, D. Z. et al. A Swin transformer-based model for mosquito species identification. **Scientific Reports**, v. 12, p. 18664, 2022.

APÊNDICES – acurácias observadas para cada pseudo-replicata.

Experiment	Species	(pseudo)-replicate	Accuracy				
			"N"	Correct	Proportion	CI lower	CI upper
All pictures	<i>Aedes scapularis</i>	AN01	32	30	0,938	0,799	0,983
		AN02	32	23	0,719	0,546	0,844
		AN03	32	30	0,938	0,799	0,983
		AN04	32	27	0,844	0,682	0,931
		AN05	32	30	0,938	0,799	0,983
		AN06	32	29	0,906	0,758	0,968
		AN07	32	27	0,844	0,682	0,931
		AN08	32	28	0,875	0,719	0,950
		AN09	32	26	0,813	0,647	0,911
		AN10	32	29	0,906	0,758	0,968
		Sub-total	320	279	0,872	0,831	0,904
	<i>Aedes serratus</i>	AN01	50	45	0,900	0,786	0,957
		AN02	50	48	0,960	0,865	0,989
		AN03	50	43	0,860	0,738	0,930
		AN04	50	47	0,940	0,838	0,979
		AN05	50	41	0,820	0,692	0,902
		AN06	50	47	0,940	0,838	0,979
		AN07	50	50	1,000	0,929	1,000
		AN08	50	47	0,940	0,838	0,979
		AN09	50	45	0,900	0,786	0,957
		AN10	50	49	0,980	0,895	0,996
		Sub-total	500	462	0,924	0,897	0,944
	<i>Haemagogus leucocelaenus</i>	AN01	31	31	1,000	0,890	1,000
		AN02	31	31	1,000	0,890	1,000
		AN03	31	31	1,000	0,890	1,000
		AN04	31	31	1,000	0,890	1,000
		AN05	31	31	1,000	0,890	1,000
		AN06	31	31	1,000	0,890	1,000
		AN07	31	31	1,000	0,890	1,000
		AN08	31	31	1,000	0,890	1,000
		AN09	31	31	1,000	0,890	1,000
		AN10	31	31	1,000	0,890	1,000
		Sub-total	310	310	1,000	0,988	1,000
	<i>Sabethes albiprivus</i>	AN01	23	23	1,000	0,857	1,000
		AN02	23	23	1,000	0,857	1,000
		AN03	23	23	1,000	0,857	1,000
		AN04	23	23	1,000	0,857	1,000
		AN05	23	23	1,000	0,857	1,000
		AN06	23	23	1,000	0,857	1,000
		AN07	23	23	1,000	0,857	1,000
		AN08	23	23	1,000	0,857	1,000
		AN09	23	23	1,000	0,857	1,000
		AN10	23	22	0,957	0,790	0,992
		Sub-total	230	229	0,996	0,976	0,999
		Sub-total	1360	1280	0,941	0,927	0,952
Full-body	<i>Aedes scapularis</i>	AN01	11	11	1,000	0,741	1,000
		AN02	11	9	0,818	0,523	0,949
		AN03	11	11	1,000	0,741	1,000
		AN04	11	9	0,818	0,523	0,949
		AN05	11	10	0,909	0,623	0,984
		AN06	11	8	0,727	0,434	0,903
		AN07	11	9	0,818	0,523	0,949
		AN08	11	9	0,818	0,523	0,949
		AN09	11	10	0,909	0,623	0,984
		AN10	11	8	0,727	0,434	0,903
		Sub-total	110	94	0,855	0,777	0,908
	<i>Aedes serratus</i>	AN01	210	184	0,876	0,825	0,914
		AN02	21	20	0,952	0,773	0,992
		AN03	21	18	0,857	0,654	0,950

		AN04	21	17	0,810	0,600	0,923
		AN05	21	18	0,857	0,654	0,950
		AN06	21	21	1,000	0,845	1,000
		AN07	21	17	0,810	0,600	0,923
		AN08	21	18	0,857	0,654	0,950
		AN09	21	19	0,905	0,711	0,973
		AN10	21	18	0,857	0,654	0,950
		Sub-total	21	18	0,857	0,654	0,950
	<i>Haemagogus leucocelaenus</i>	AN01	15	15	1,000	0,796	1,000
		AN02	15	15	1,000	0,796	1,000
		AN03	15	15	1,000	0,796	1,000
		AN04	15	15	1,000	0,796	1,000
		AN05	15	15	1,000	0,796	1,000
		AN06	15	15	1,000	0,796	1,000
		AN07	15	15	1,000	0,796	1,000
		AN08	15	15	1,000	0,796	1,000
		AN09	15	15	1,000	0,796	1,000
		AN10	15	15	1,000	0,796	1,000
		Sub-total	150	150	1,000	0,975	1,000
	<i>Sabethes albiprivus</i>	AN01	10	10	1,000	0,722	1,000
		AN02	10	10	1,000	0,722	1,000
		AN03	10	10	1,000	0,722	1,000
		AN04	10	10	1,000	0,722	1,000
		AN05	10	9	0,900	0,596	0,982
		AN06	10	10	1,000	0,722	1,000
		AN07	10	10	1,000	0,722	1,000
		AN08	10	10	1,000	0,722	1,000
		AN09	10	9	0,900	0,596	0,982
		AN10	10	10	1,000	0,722	1,000
		Sub-total	100	98	0,980	0,930	0,994
	Sub-total		570	526	0,923	0,898	0,942
Lateral thorax	<i>Aedes scapularis</i>	AN01	5	4	0,800	0,376	0,964
		AN02	5	5	1,000	0,566	1,000
		AN03	5	5	1,000	0,566	1,000
		AN04	5	3	0,600	0,231	0,882
		AN05	5	3	0,600	0,231	0,882
		AN06	5	5	1,000	0,566	1,000
		AN07	5	5	1,000	0,566	1,000
		AN08	5	5	1,000	0,566	1,000
		AN09	5	5	1,000	0,566	1,000
		AN10	5	5	1,000	0,566	1,000
		Sub-total	50	45	0,900	0,786	0,957
	<i>Aedes serratus</i>	AN01	12	11	0,917	0,646	0,985
		AN02	12	10	0,833	0,552	0,953
		AN03	12	11	0,917	0,646	0,985
		AN04	12	10	0,833	0,552	0,953
		AN05	12	12	1,000	0,758	1,000
		AN06	12	10	0,833	0,552	0,953
		AN07	12	11	0,917	0,646	0,985
		AN08	12	8	0,667	0,391	0,862
		AN09	12	11	0,917	0,646	0,985
		AN10	12	12	1,000	0,758	1,000
		Sub-total	120	106	0,883	0,814	0,929
	<i>Haemagogus leucocelaenus</i>	AN01	7	7	1,000	0,646	1,000
		AN02	7	7	1,000	0,646	1,000
		AN03	7	7	1,000	0,646	1,000
		AN04	7	7	1,000	0,646	1,000
		AN05	7	7	1,000	0,646	1,000
		AN06	7	7	1,000	0,646	1,000
		AN07	7	7	1,000	0,646	1,000
		AN08	7	7	1,000	0,646	1,000
		AN09	7	7	1,000	0,646	1,000
		AN10	7	7	1,000	0,646	1,000
		Sub-total	70	70	1,000	0,948	1,000
	<i>Sabethes albiprivus</i>	AN01	6	6	1,000	0,610	1,000

		AN02	6	6	1,000	0,610	1,000
		AN03	6	6	1,000	0,610	1,000
		AN04	6	6	1,000	0,610	1,000
		AN05	6	6	1,000	0,610	1,000
		AN06	6	6	1,000	0,610	1,000
		AN07	6	6	1,000	0,610	1,000
		AN08	6	6	1,000	0,610	1,000
		AN09	6	6	1,000	0,610	1,000
		AN10	6	6	1,000	0,610	1,000
		Sub-total	60	60	1,000	0,940	1,000
	Sub-total		300	281	0,937	0,903	0,959
Pronotum	<i>Aedes scapularis</i>	AN01	15	13	0,867	0,621	0,963
		AN02	15	14	0,933	0,702	0,988
		AN03	15	15	1,000	0,796	1,000
		AN04	15	12	0,800	0,548	0,930
		AN05	15	10	0,667	0,417	0,848
		AN06	15	13	0,867	0,621	0,963
		AN07	15	13	0,867	0,621	0,963
		AN08	15	15	1,000	0,796	1,000
		AN09	15	14	0,933	0,702	0,988
		AN10	15	13	0,867	0,621	0,963
		Sub-total	150	132	0,880	0,818	0,923
	<i>Aedes serratus</i>	AN01	17	16	0,941	0,730	0,990
		AN02	17	15	0,882	0,657	0,967
		AN03	17	15	0,882	0,657	0,967
		AN04	17	17	1,000	0,816	1,000
		AN05	17	16	0,941	0,730	0,990
		AN06	17	15	0,882	0,657	0,967
		AN07	17	16	0,941	0,730	0,990
		AN08	17	16	0,941	0,730	0,990
		AN09	17	15	0,882	0,657	0,967
		AN10	17	17	1,000	0,816	1,000
		Sub-total	170	158	0,929	0,881	0,959
	<i>Haemagogus leucocelaenus</i>	AN01	9	9	1,000	0,701	1,000
		AN02	9	9	1,000	0,701	1,000
		AN03	9	9	1,000	0,701	1,000
		AN04	9	9	1,000	0,701	1,000
		AN05	9	9	1,000	0,701	1,000
		AN06	9	9	1,000	0,701	1,000
		AN07	9	9	1,000	0,701	1,000
		AN08	9	9	1,000	0,701	1,000
		AN09	9	9	1,000	0,701	1,000
		AN10	9	9	1,000	0,701	1,000
		Sub-total	90	90	1,000	0,959	1,000
	<i>Sabethes albiprivus</i>	AN01	7	7	1,000	0,646	1,000
		AN02	7	7	1,000	0,646	1,000
		AN03	7	7	1,000	0,646	1,000
		AN04	7	7	1,000	0,646	1,000
		AN05	7	7	1,000	0,646	1,000
		AN06	7	7	1,000	0,646	1,000
		AN07	7	7	1,000	0,646	1,000
		AN08	7	7	1,000	0,646	1,000
		AN09	7	7	1,000	0,646	1,000
		AN10	7	7	1,000	0,646	1,000
		Sub-total	70	70	1,000	0,948	1,000
	Sub-total		480	450	0,938	0,912	0,956
Grand-total			2710	2537	0,936	0,926	0,945

ANEXO A – Artigo Publicado

AI-driven convolutional neural networks for accurate identification of yellow fever vectors

Revista *Parasites & Vectors*

RESEARCH

Open Access



AI-driven convolutional neural networks for accurate identification of yellow fever vectors

Taís Oliveira de Araújo^{1,2†}, Vinicius Lima de Miranda^{2†} and Rodrigo Gurgel-Gonçalves^{1,2*}

Abstract

Background Identifying mosquito vectors is crucial for controlling diseases. Automated identification studies using the convolutional neural network (CNN) have been conducted for some urban mosquito vectors but not yet for sylvatic mosquito vectors that transmit the yellow fever. We evaluated the ability of the AlexNet CNN to identify four mosquito species: *Aedes serratus*, *Aedes scapularis*, *Haemagogus leucocelaenus* and *Sabethes albiprivus* and whether there is variation in AlexNet's ability to classify mosquitoes based on pictures of four different body regions.

Methods The specimens were photographed using a cell phone connected to a stereoscope. Photographs were taken of the full-body, pronotum and lateral view of the thorax, which were pre-processed to train the AlexNet algorithm. The evaluation was based on the confusion matrix, the accuracy (ten pseudo-replicates) and the confidence interval for each experiment.

Results Our study found that the AlexNet can accurately identify mosquito pictures of the genus *Aedes*, *Sabethes* and *Haemagogus* with over 90% accuracy. Furthermore, the algorithm performance did not change according to the body regions submitted. It is worth noting that the state of preservation of the mosquitoes, which were often damaged, may have affected the network's ability to differentiate between these species and thus accuracy rates could have been even higher.

Conclusions Our results support the idea of applying CNNs for artificial intelligence (AI)-driven identification of mosquito vectors of tropical diseases. This approach can potentially be used in the surveillance of yellow fever vectors by health services and the population as well.

Keywords Deep learning, Machine learning, Culicidae

[†]Taís Oliveira de Araújo and Vinicius Lima de Miranda have Contributed equally to this work.

*Correspondence:
Rodrigo Gurgel-Gonçalves
gurgelrg@hotmail.com

¹ Programa de Pós-Graduação em Medicina Tropical, Faculdade de Medicina, Universidade de Brasília, Brasília, DF, Brasil

² Laboratório de Parasitologia Médica e Biologia de Vetores, Faculdade de Medicina, Universidade de Brasília, Brasília, DF, Brasil

Background

Mosquito-borne diseases are a major public health concern. More than half of the world's population is exposed to arboviruses such as yellow fever, dengue and Zika. Currently, 47 countries, including 34 in Africa and 13 in Central and South America, are endemic or have endemic regions for yellow fever [1, 2]. These pathogens are mainly transmitted by mosquito bites, with *Aedes aegypti* and *Aedes albopictus* species being the most important urban vectors of these arboviruses [3, 4]. Sylvatic mosquito species from the *Sabethes* and *Haemagogus* genera



have been found to be naturally infected with the yellow fever virus and other arboviruses, for which no specific therapeutic agents exist [5–11].

One approach to prevent the spread of these diseases is by controlling the spread of the virus' vector. Traditional mosquito surveillance relies on catches and species identification, which require regular manual inspection and dedicated personnel, making large-scale monitoring difficult and expensive. Additionally, identifying mosquitoes is a difficult task that demands specialized knowledge due to the vast range of morphological characteristics found in the Culicidae family, which includes all mosquitoes [12]. New approaches that rely on smartphones and the internet can enable new community and digital observatories with the task of species identification. These observatories allow individuals to submit photos of mosquitoes they come across. In addition, smartphone cameras are improving macro photography, which can produce high-quality images of small organisms such as mosquitoes. However, manually inspecting each picture is not a feasible long-term solution due to the large volume of pictures and professionals needed for the task [13–15].

Deep learning methods based on a convolutional neural networks (CNNs) have shown promise for mosquito identification [13, 14, 16–19]. CNNs simulate the human learning process for classifying pictures and extract important features from data automated without the need of human supervision [20]. AlexNet is a CNN that was pre-trained on 1.2 million images of objects, animals and plants available in the ImageNet database [21]. AlexNet has been successfully used for the identification of insect vectors [22]. Automated identification studies have been conducted for some urban mosquito vectors [13, 14, 16–19] but not yet for sylvatic mosquito vectors that transmit the yellow fever and other arboviruses [6, 11].

Therefore, considering the lack of specialised tools for the automated identification of wild yellow fever vectors using image processing, the aim of this study is to test AlexNet's ability to identify the species *Aedes serratus*, *Aedes scapularis*, *Haemagogus leucocelaenus* and *Sabethes albiprivus*. These four mosquito species were chosen because they are confirmed vectors of the yellow fever virus (*Hg. leucocelaenus* and *Sa. albiprivus*) or because they are suspected vectors of the virus (*Ae. serratus* and *Ae. scapularis*). *Hg. leucocelaenus* was identified as a primary vector in major yellow fever outbreaks in Brazil, exhibiting a broad geographical distribution within the country. The species is found in humid and well-preserved forests and is tolerant in anthropogenic environments, which increases the human exposure to infected mosquitoes [8, 10, 23]. *Sa. albiprivus* is widely

distributed in South American forests and has competence in transmitting yellow fever virus being the most important species in Argentina [11]. *Ae. serratus* and *Ae. scapularis* have been identified as potential vectors of the yellow fever, as they have been found naturally infected with the virus. However, their vectorial competence and role in the sylvatic cycle of the disease remain poorly understood [5, 6]. We also asked whether there is variation in AlexNet's ability to classify mosquitoes based on images of three different body regions. Our findings indicate that the AlexNet network can accurately identify yellow fever vectors with over 90% accuracy for the four body regions analysed, presenting great potential for the development of an application to facilitate surveillance of these vectors.

Methods

Picture database

To build the picture database, four mosquito species were selected available at the Laboratory of Medical Parasitology and Vector Biology at the University of Brasilia. Among them, *Sa. albiprivus* ($n=100$) and *Hg. leucocelaenus* ($n=98$) were vectors of the wild yellow fever virus, while *Ae. serratus* ($n=100$) and *Ae. scapularis* ($n=100$) were vectors of other arboviruses. A total of 565 full-body pictures were captured, including 294 of the lateral thorax and 484 of the pronotum, resulting in 1343 pictures. The Culicidae specimens were mounted on a cardboard triangle using a pin. The mosquitoes were attached to the triangle by the thoracic pleura, leaving the legs facing the pin and the upper pleura free for observation.

Pictures were captured using a cell phone camera (Samsung Galaxy S8, model SM-G950F, 12 mp, sensor 1/2.5, aperture size $f/1.7$) attached to a stereoscope (Zeiss Stemi 508, 1–10 \times magnification). Photographs of the full-body, pronotum and lateral view of the thorax were taken (Fig. 1), cropped in a square format (proportion 1:1) in the computer software Photoscape, selecting only the mosquito, and subsequently resized to 227 \times 227 pixels using the software computational MATLAB. They were then organized into folders based on the species and experiment to be conducted.

Algorithm

The AlexNet architecture, comprising five convolutional layers and three fully connected layers, has 60 million parameters. This network utilises 227 \times 227 pixels RGB images as input. The model was pre-trained with 1.2 million images of varying resolutions, representing over 1000 classes from the ImageNet database [21]. The AlexNet architecture employs convolutional layers to extract features from images, including edges, textures and shapes. The maximum pooling layers reduce

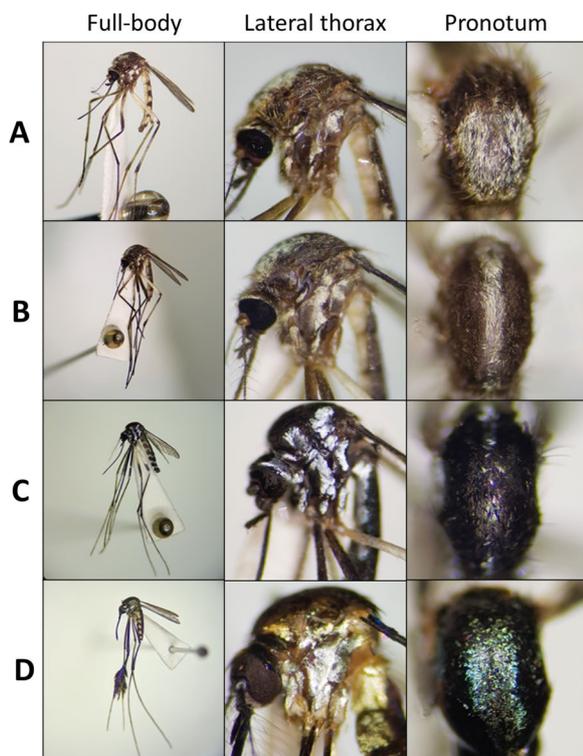


Fig. 1 Examples of pictures used to train the AlexNet network. Mosquitoes photographed with a stereomicroscope and cell phone: *Ae. scapularis* (A), *Ae. serratus* (B), *Hg. leucocelaenus* (C) and *Sa. albiprivus* (D)

the dimensionality of the image while retaining important features. The rectified linear unit (ReLU) function introduces non-linearity to the network, enabling it to learn more complex features. Local response normalisation normalises the outputs of the convolutional layers, improving generalisation. The final classification is performed by three fully connected layers. The dropout technique, which randomly deactivates neurons during training, prevents overfitting. The Softmax function converts the network outputs into probabilities, selecting the highest probability as the final prediction [21].

Algorithm training and testing

The AlexNet algorithm, which has already been used to identify insect vectors [16, 22], was implemented in MATLAB. The image database was divided into three sections for the experiments: 80% for training, 10% for internal validation and 10% for testing (Additional file 2: Dataset S1). The algorithm randomly selected the images for each stage. In the algorithm settings, the maximum number of epochs was set to 50 following preliminary testing. Stochastic gradient descent with a moment optimiser was employed, with default settings aside from the initial learning rate, which was set to 0.001. Four

experiments were conducted with four classes in ten pseudo-replicates each (ten repetitions of training and testing using the same set of pictures). The first experiment utilised all images, including full-body, pronotum and lateral thorax. The other experiments employed separate images of each body part.

Data analysis

The network's performance was evaluated based on observations from confusion matrices/heat maps, the mean and confidence interval of the general accuracy (probability of correctly identifying the species independently) and the specific accuracy (probability of correctly identifying each species specific) that can be interpreted as sensitivity, depending on the context of the analysis of the algorithm for each experiment. To achieve this, we employed the 'Hmisc' package within the computational software R, version 4.7.1, in conjunction with the RStudio interface, version 2023.03.1.446 [24–26]. We use the following equations to calculate AlexNet performance metrics:

$$\text{General accuracy} = \frac{TP + TN}{TP + FP + TN + FN}$$

$$\text{Specific accuracy} = \frac{TP}{TP + FN}$$

TP, true positives;
TN, true negatives;
FP, false positives;
FN, false negatives.

Results

Confusion matrices

To gain an understanding of the performance of the AlexNet algorithm, Fig. 2 presents a series of confusion matrices, which illustrate the patterns of correct and incorrect identification of mosquito images across the ten pseudo-replicates.

General accuracy of the algorithm

The performance of the algorithm was evaluated by processing test pictures after training the AlexNet algorithm. Four experiments were conducted, in which each group (full-body, lateral thorax and pronotum) was submitted to the AlexNet algorithm. Additionally, all 1343 photos corresponding to the three groups were submitted together. The average confidence interval for general accuracy was then calculated. The AlexNet algorithm demonstrated the highest general accuracy in identifying mosquitoes when all pictures were analysed together, providing the algorithm with

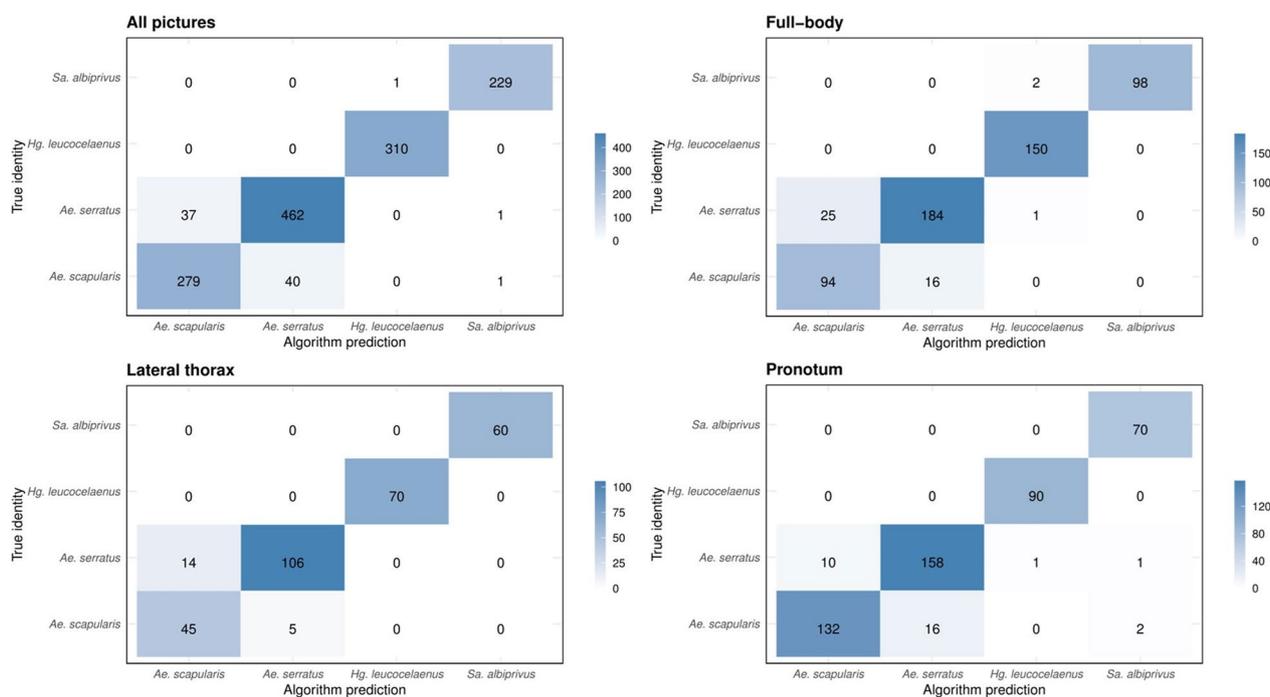


Fig. 2 Confusion matrices showing the classification hits and misses of the 10% of the test images for the four experiments. Each run was pseudo-replicated 10x. The blue scale indicates the number of correct predictions

the most comprehensive information about the mosquito's body. Each run was repeated ten times, achieving an average general accuracy of 94% (95% confidence interval (CI) 90–97). The run with the total body group achieved an average general accuracy of 92.3% (95% CI 83–97), while with the lateral view group the value

was 93.7% (95% CI 79–98). Finally, with the pronotum group, the average general accuracy was 93.8% (95% CI 83–98). The smallest confidence intervals were observed with pictures of all parts of the body (Fig. 3; see Additional file 1: Table S1).

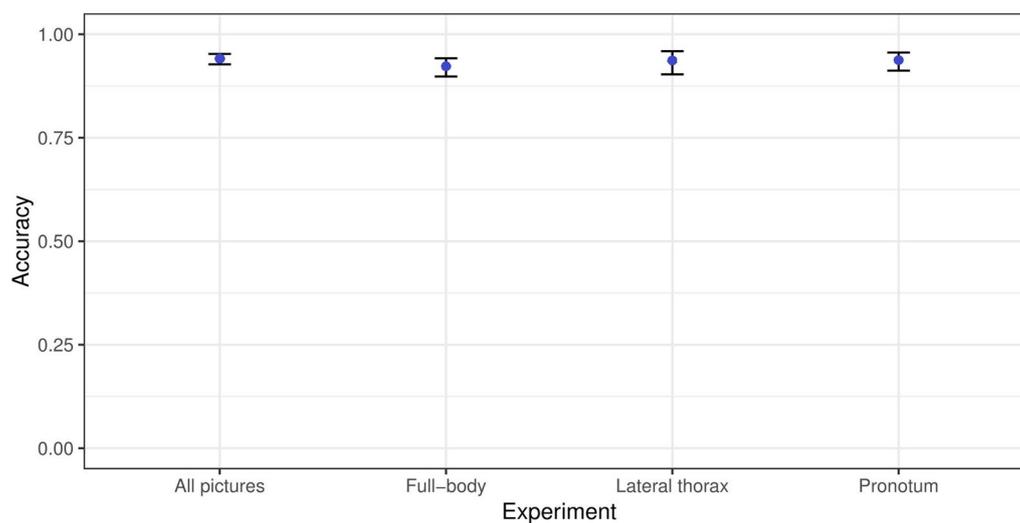


Fig. 3 Average accuracy of the AlexNet algorithm in identifying mosquito vectors of yellow fever and other arboviruses. The experiments took into account identification from images of isolated body parts regions, with 95% confidence intervals. Each run was pseudo-replicated 10x

Algorithm-specific accuracy

We then evaluated the algorithm's performance for each species based in the three groups of pictures. The algorithm's best performance was observed when pictures from the three groups were combined in the same experiment. The performance of the AlexNet algorithm in identifying mosquitoes of the genus *Aedes* was poor, with the highest confidence intervals observed for the two species in this genus, compared with the others (95% CI 72–98) (Fig. 4). The algorithm misidentified at least one picture of *Ae. scapularis* as *Ae. serratus* or vice versa when presented with pictures of all three groups. In the run where all 1343 pictures were available, *Ae. serratus* was misidentified as *Hg. leucocelaenus*. In the run with pictures of the lateral view, only mosquitoes of the genus *Aedes* were misidentified, with the two *Aedes* species being swapped. In the run with the pronotum, several misidentifications occurred. *Ae. serratus* was incorrectly identified as either *Sa. albiprivus* or *Hg. leucocelaenus*, while *Ae. scapularis* was identified as *Sa. albiprivus* (Fig. 2). It is worth noting that many of the photographed specimens were not well-preserved, which may have contributed to the high rate of misidentification (Fig. 5).

Discussion

In this study, we aim to identify four mosquito species that transmit yellow fever or other arboviruses by using a CNN (AlexNet). We also wanted to investigate whether the algorithm performance in classifying the mosquitoes

changes according to the body regions shown on the pictures submitted. Our study found that the AlexNet can accurately identify mosquito pictures of the genus *Aedes*, *Sabethes* and *Haemagogus* with over 90% accuracy. Furthermore, the algorithm performance did not change according to the body regions shown.

Lorenz et al. [27] classified mosquitoes based on morphometric characteristics of their wings using neural networks, achieving accuracies ranging from 86% to 100%. However, an identification system based only on a body structure such as the wing is more fragile because if the structure is not present in the analysed photo, the identification is compromised. Therefore, a good identification system should work with any part of the insect's body. Sauer et al. [28] showed that best-performing CNN configuration achieved a precision of 99% to discriminate between *Aedes* and non-*Aedes* mosquito species; the mean precision to predict the *Aedes* species was 91% for RGB pictures. Motta et al. [16] used three pre-trained networks to identify urban mosquitoes (*Aedes* and *Culex*), achieving an accuracy of 76.2% for the GoogLeNet, 52.4% for LeNet and 51.2% for AlexNet. Okayasu et al. [29] showed better accuracy results (92.3%) with the identification of *Ae. albopictus*, *Anopheles stefensi* and *Cx. pipiens pallens* using AlexNet based on data augmentation and 12,000 training pictures. More recently, Motta et al. [17] optimized the CNN hyperparameters and obtained 97.3% accuracy in distinguishing between the mosquitoes of the genus *Aedes* and the *Culex*

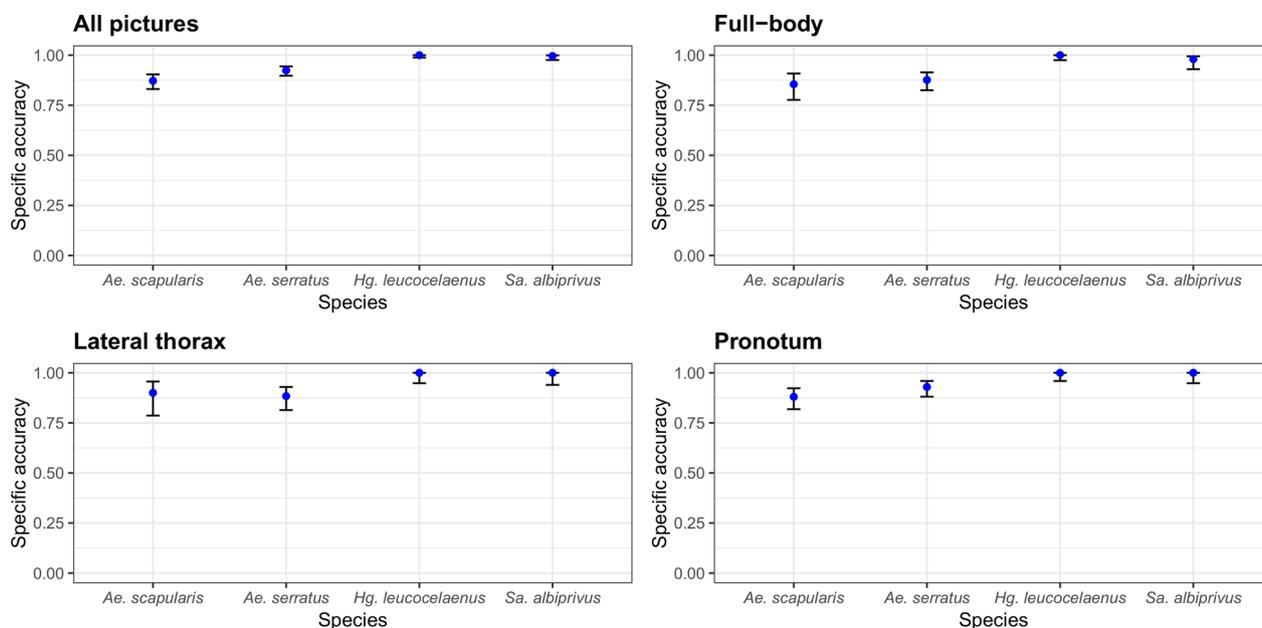


Fig. 4 Average accuracy of the AlexNet algorithm in interspecies identification. The experiments took into account identification from images of isolated body parts regions, with 95% confidence intervals. Each run was pseudo-replicated 10x

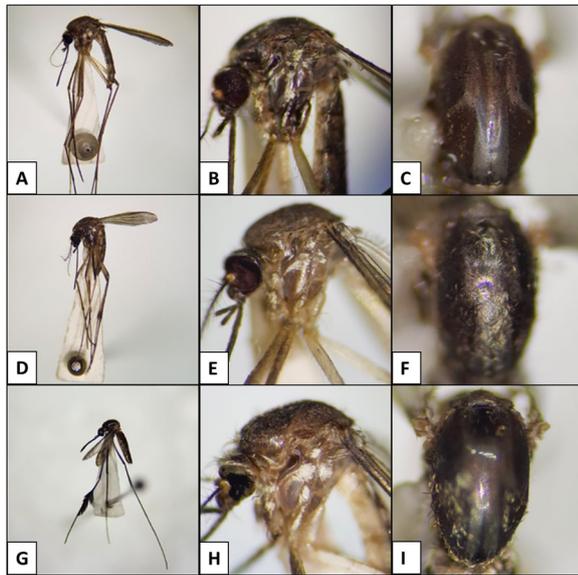


Fig. 5 Examples of incorrectly identified images, due to damaged specimens. **A–C** and **H** *Ae. serratus* was mistaken for *Ae. scapularis*. **D** and **E** *Ae. scapularis* was mistaken for *Ae. serratus*. **F** *Ae. serratus* was mistaken for *Sa. albiprivus*. **G** *Sa. albiprivus* was mistaken for *Hg. leucocelaenus*. **I** *Ae. scapularis* was erroneously identified as *Hg. leucocelaenus*

mosquitoes. Similarly, Goodwin et al. [30] and Park et al. [13] achieved 97% accuracy rates for mosquito identification (*Anopheles*, *Culex*, *Psorophora* and *Aedes* species) using deep learning neural networks. These networks rely on morphological features like those used by taxonomists [13]. Kittichai et al. [15] using two YOLO v3 model identified *Ae. aegypti*, *Ae. albopictus* and *Cx. quinquefasciatus* at a mean average accuracy of 98–100%. A recent study has shown that the accuracy and robustness of the CNN may reach 99% accuracy by incorporating spatial dropout layers to regularize the network and by modifying its structure to incorporate multi-view inputs [31]. Concatenating two YOLO v3 model exhibited the optimal performance in identifying mosquitoes, with mean average accuracy of 99%. The Swin MSI successfully identified 100% subspecies-level in *Culex pipiens* complex. Based on pictures of all body regions, AlexNet identified *Ae. scapularis*, *Ae. serratus*, *Hg. leucocelaenus* and *Sa. albiprivus* at 94% accuracy on average (Fig. 3). Compared with previous studies that have used neural networks for mosquito identification, our accuracy rate of 94% is aligned with most results obtained by others.

In our study, we did not find a significant difference in AlexNet performance in identifying mosquitoes based on different body regions. In fact, other studies have shown that CNNs are able to detect morphological differences in various body regions of *Aedes* mosquitoes [13], some

of which are consistent with the most used dichotomous keys [32]. Such results reveal that deep learning models learn the distinctive morphological features of mosquitoes body areas; these are the same ones used by taxonomists. For instance, *Ae. scapularis* can be identified by its serrated abdomen, a proboscis that is larger than the anterior femur and the mesonotum with white scales forming a circle. *Ae. serratus* is identified by its serrated abdomen, a proboscis that is similar to or smaller than the anterior femur and a mesonotum that may or may not have a longitudinal stripe of white scales [7]. The two species are very similar because they belong to the subgenus *Ochlerotatus* (Lynch Arribáizaga, 1891), where most of the species in this group are indistinguishable based on morphological characters [33]. *Sa. albiprivus* has medium-sized legs with bluish scales, a golden-scaled abdomen that forms quadrilaterals and a proboscis that is much smaller than the anterior femur. *Sa. albiprivus* and *Hg. leucocelaenus* are two species with different morphological characteristics. *Sa. albiprivus* can be distinguished from *Hg. leucocelaenus* by its predominantly dull, dark colour and pleura with two vertical lines of silvery scales [7]. Other studies show that the accuracy of CNNs in identifying other insects is not significantly affected by the body region shown on the picture [22]. Our findings show that the morphological characteristics used for the identification of the mosquitoes included in this study are present in multiple regions of the body and, therefore, any of the body regions here studied allowed the AlexNet to accurately identify the mosquito species.

Deep learning neural networks consist of multiple convolutional layers, and databases with more pictures are more conducive to learning [21]. Additionally, many studies indicate that a larger picture bank improves the algorithm's performance [13, 14, 16, 17, 22, 34, 35]. Even though a database with thousands of pictures is always desired, using a database with only 1343 pictures, we reached accuracy rates similar to those using databases 10× bigger than ours [17, 29]. AlexNet accuracy to identify *Sabethes* and *Haemagogus* mosquitoes was similar to the accuracy obtained with other CNNs used to identify other genera [17, 29]. However, the accuracy of the AlexNet in identifying *Ae. serratus* and *Ae. scapularis* was below 90% and, thus, suboptimal when compared with the performance of other CNNs (VGG-16, ResNet-50, SqueezeNet) that apply data augmentation and fine-tuning techniques to identify *Ae. aegypti* [16, 17], *Ae. albopictus* and *Ae. vexans* [13].

These accuracy values may be due to differences in algorithm architecture and training [13]. AlexNet is relatively shallow compared with deeper algorithms such as ResNet or DenseNet, which can capture complex hierarchical features in the pictures. Although AlexNet

was designed to take advantage of GPU acceleration, its computational efficiency is suboptimal compared to other algorithms such as MobileNet. Finally, the features learned by AlexNet may not be as discriminative as those learned by algorithms trained on large datasets. Despite the limitations of AlexNet when compared with other algorithms, our results show high accuracy for identification of yellow fever mosquitoes. Our perspective is to apply other algorithms (ResNet, DenseNet or MobileNet) to test their efficiency in identifying *Ae. serratus*, *Ae. scapularis* and other sibling mosquito species. Moreover, optimization of CNN hyperparameters increase the accuracy of mosquito identification [17]. The lower performance of the algorithm, in some cases, may have been influenced by the state of preservation of the specimens. Analysis of the misidentified pictures in all experiments showed that the photographed specimens were not well preserved, especially in the pronotum area, where bristles and scales were missing, as well as the legs. Due to their size and the presence of scales and bristles, mosquitoes are easily damaged during capture, freezing and drying, resulting in the loss of critical morphological features necessary for proper identification. The state of preservation of the mosquitoes was a limiting factor in this work, and good preservation of specimens is important for optimal algorithm performance. Furthermore, the network has not been evaluated with images captured by different photographic devices. Therefore, it is not possible to say whether images taken by cameras with different specifications influence AlexNet performance.

Conclusions

We aim to identify four mosquito species that transmit yellow fever using AlexNet. We found that the AlexNet CNN can identify mosquito pictures of the genus *Aedes*, *Sabethes* and *Haemagogus* with over 90% accuracy, regardless of the body region being shown. Our results support the idea of applying CNNs for AI-driven identification of mosquito vectors of tropical diseases. This approach can potentially be used in the surveillance of yellow fever vectors by health services and the population as well.

Abbreviations

CI Confidence interval
CNN Convolutional neural network

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13071-024-06406-2>.

Additional file 1: Table S1. Accuracy observed for each experiment.

Additional file 2: Database S1.

Acknowledgements

We thank Marcia Triunfol at Publicase for reviewing this manuscript and Marcos Takashi Obara at the University of Brasília for your helpful comments.

Author contributions

T.O.A., R.G.-G. and V.L.M. conceived the study. R.G.-G. raised funds and administered the project. R.G.-G. and V.L.M. contributed to the design of trial methods. T.O.A., R.G.-G. and V.L.M. performed research. R.G.-G. supervised students involved in laboratory research. T.O.A. obtained the databases. R.G.-G. and V.L.M. curated the dataset and analysed the data. R.G.-G. and T.O.A. drafted the first version of the manuscript. All authors contributed to the interpretation of results, read and commented on manuscript drafts and approved the final version.

Funding

T.O.A. received specific funding of Coordenação de Aperfeiçoamento de Pessoal de Nível Superior <https://www.capes.gov.br/> Award Number: finance code 001. RGG and VLM received specific funding of Conselho Nacional de Desenvolvimento Científico e Tecnológico <http://www.cnpq.br/>. The funding sources of this study had no role in the study design, data collection, data analysis, data interpretation, writing of the report or in the decision to submit the paper for publication.

Availability of data and materials

Data supporting the conclusions of this article are included in the article and its additional files.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 24 May 2024 Accepted: 13 July 2024

Published online: 02 August 2024

References

- World Health Organization. World Health Day 2014: preventing vector-borne diseases. Geneva: World Health Organization; 2014.
- World Health Organization. Yellow fever. Geneva: World Health Organization; 2023.
- Donalisio MR, Freitas ARR, Von ZAPB. Arboviruses emerging in Brazil: challenges for clinic and implications for public health. *Rev Saude Publica*. 2017;51:30. <https://doi.org/10.1590/s1518-8787.2017051006889>.
- Ahebwa A, Hii J, Neoh KB, Chareonviriyaphap T. *Aedes aegypti* and *Ae. albopictus* (Diptera: Culicidae) ecology, biology, behaviour, and implications on arbovirus transmission in Thailand: Review. *One Health*. 2023;16:100555. <https://doi.org/10.1016/j.onehlt.2023.100555>.
- Cunha MS, Faria NR, Caleiro GS, Candido DS, Hill SC, Claro IM, et al. Genomic evidence of yellow fever virus in *Aedes scapularis*, southeastern Brazil, 2016. *Acta Trop*. 2020;205:105390.
- da Cardoso JC, de Almeida MAB, dos Santos E, da Fonseca DF, Sallum MAM, Noll CA, et al. Yellow fever virus in *Haemagogus leucocelaenus* and *Aedes serratus* mosquitoes, Southern Brazil, 2008. *Emerg Infect Dis*. 2010;16:1918–24.
- Segura MNO, Castro FC. Atlas de culicídeos na Amazônia brasileira: características específicas de insetos hematófagos da família Culicidae. Belém: Instituto Evandro Chagas; 2007.
- Vasconcelos PFC, Rosa APAT, Rodrigues SG, Rosa EST, Monteiro HAO, Cruz ACR, et al. Yellow fever in Pará State, Amazon region of Brazil, 1998–1999: entomologic and epidemiologic findings. *Emerg Infect Dis*. 2001;7:565.

9. Vasconcelos PFC, Rodrigues SG, Degallier N, Moraes MAP, Travassos Da Rosa JFS, Travassos Da Rosa ES, et al. An epidemic of sylvatic yellow fever in the southeast region of Maranhao State, Brazil, 1993–1994: epidemiologic and entomologic findings. *Am J Trop Med Hyg.* 1997;57:132.
10. Vasconcelos PFC, Travassos da Rosa APA, Pinheiro FP, Shope RE, Travassos da Rosa JFS, Rodrigues SG, et al. Arboviruses pathogenic for man in Brazil. In: Travassos da Rosa APA, Amelia PA, Vasconcelos PFC, Pedro FC, Travassos da Rosa JFS, editors., et al., *An Overview of Arbovirology in Brazil and Neighbouring Countries*. Belem: Instituto Evandro Chagas; 1998.
11. Cano ME, Marti GA, Alencar J, Silva SOF, Micieli MV. Categorization by score of Mosquito Species (Diptera: Culicidae) Related to yellow fever Epizootics in Argentina. *J Med Entomol.* 2022;59:1384–8.
12. Wilkerson RC, Linton YM, Fonseca DM, Schultz TR, Price DC, Strickman DA. Making mosquito taxonomy useful: a stable classification of tribe Aedini that balances utility with current knowledge of evolutionary relationships. *PLoS ONE.* 2015;10:e0133602.
13. 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;10:3144.
14. Pataki BA, Garriga J, Eritja R, Palmer JRB, Bartumeus F, Csabai I. Deep learning identification for citizen science surveillance of tiger mosquitoes. *Sci Rep.* 2021;11:206.
15. Kittichai V, Pengsakul T, Chumchuen K, Samung Y, Sriwichai P, Phattamolrat N, et al. Deep learning approaches for challenging species and gender identification of mosquito vectors. *Sci Rep.* 2021;11:19375.
16. 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:e0216427.
17. Motta D, Bandeira Santos AÁ, Souza Machado BA, Vicente Ribeiro-Filho OG, Arriaga Camargo LO, Valdenegro-Toro MA, et al. Optimization of convolutional neural network hyperparameters for automatic classification of adult mosquitoes. *PLoS ONE.* 2020;15:e0237151.
18. Kittichai V, Kaewthamasorn M, Samung Y, Jomtarak R, Naing KM, Tongloy T, et al. Automatic identification of medically important mosquitoes using embedded learning approach-based image-retrieval system. *Sci Rep.* 2023;13:13072.
19. Zhao OS, Kolluri N, Anand A, Chu N, Bhavaraju R, Ojha A, et al. Convolutional neural networks to automate the screening of malaria in low-resource countries. *PeerJ.* 2020;8:e8965.
20. Valan M, Makonyi K, Maki A, Vondráček D, Ronquist F. Automated taxonomic identification of insects with expert-level accuracy using effective feature transfer from convolutional networks. *Syst Biol.* 2019;68:876–95.
21. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Adv Neural Inf Process Sys.* 2012;25:1106–14.
22. de 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;79:101660.
23. Lira-Vieira AR, Gurgel-Gonçalves R, Moreira IM, Yoshizawa MAC, Coutinho ML, Prado PS, et al. Ecological aspects of mosquitoes (Diptera: Culicidae) in the gallery forest of Brasília National Park, Brazil, with an emphasis on potential vectors of yellow fever. *Rev Soc Bras Med Trop.* 2013;46:566–74.
24. Harrell FE. *Hmisc: a package of miscellaneous R functions.* 2014. <http://biostat.mc.vanderbilt.edu/Hmisc>. Accessed 15 Apr 2024.
25. RS Team. *RStudio: integrated Development for R.* RStudio, Version 4.4.0. Vienna: R Foundation for Statistical Computing; 2024.
26. R Development Core Team. *R: a language and environment for statistical computing.* Version 4.4.0. Vienna: R Foundation for Statistical Computing; 2024.
27. Lorenz C, Ferraudo AS, Suesdek L. Artificial Neural Network applied as a methodology of mosquito species identification. *Acta Trop.* 2015;152:165–9.
28. Sauer FG, Werny M, Nolte K, Villacañas de Castro C, Becker N, Kiel E, et al. A convolutional neural network to identify mosquito species (Diptera: Culicidae) of the genus *Aedes* by wing images. *Sci Rep.* 2024;14:14147.
29. Okayasu K, Yoshida K, Fuchida M, Nakamura A. Vision-based classification of mosquito species: comparison of conventional and deep learning methods. *Appl Sci.* 2019;9:3818.
30. Goodwin A, Padmanabhan S, Hira S, Glancey M, Slinowsky M, Immidisetti R, et al. Mosquito species identification using convolutional neural networks with a multitiered ensemble model for novel species detection. *Sci Rep.* 2021;11:14394.
31. Pora W, Kasamsunran N, Tharawatcharasart K, Ampol R, Siriyasatien P, Jariyapan N. Enhancement of VGG16 model with multi-view and spatial dropout for classification of mosquito vectors. *PLoS ONE.* 2023;18:e0253793.
32. Consoli RAGB, de Oliveira RL. *Principais mosquitos de importância sanitária no Brasil.* 1st ed. Rio de Janeiro: Editora Fiocruz; 1994.
33. Reinert JF, Harbach RE, Kitching IJ. Phylogeny and classification of tribe Aedini (Diptera: Culicidae). *Zool J Linn Soc.* 2009;157:700–94.
34. 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.
35. 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;56:461–8.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.