

Alan Cesar de Oliveira Mathias

alan.mathias@ufpr.br

Universidade Federal do Paraná

Orcid: 0009-0004-6598-6904

Betwel Montanher

pg405266@uem.br

Universidade Estadual de Maringá

Orcid: 0009-0004-8487-6882

Carlos Roberto Beleti Junior

carlosbeleti@ufpr.br

Universidade Federal do Paraná

0000-0002-0158-8673

Robertino Mendes Santiago Junior

robertino@ufpr.br

Universidade Federal do Paraná

Orcid: 0000-0002-4194-568X

Rogério Ferreira da Silva

rogerio.ferreira@ufpr.br

Universidade Federal do Paraná

Orcid: 0000-0003-2151-226X

Rodrigo Clemente Thom de Souza

thom@ufpr.br

Universidade Federal do Paraná, Campus Jandaia do Sul

Universidade Estadual de Maringá, Programa de Pós-graduação em Ciência da Computação

Orcid: 0000-0003-2435-8528

**APRENDIZAGEM PROFUNDA APLICADA À CLASSIFICAÇÃO DE
SINAIS DE ÁUDIO DE MOSQUITOS**

**DEEP LEARNING APPLIED TO MOSQUITOES AUDIO SIGNAL
CLASSIFICATION**

Resumo: Este estudo visa classificar sinais sonoros emitidos durante o voo de mosquitos da espécie *Culex quinquefasciatus* e *Anopheles Gambiae*. Para alcançar esse objetivo, foram empregadas técnicas de *Deep Learning*, que consistem na capacidade de aprendizagem das máquinas por meio da entrada de dados e modelos de aprendizagem complexos chamados de redes neurais artificiais profundas, que processam informações em níveis abstratos e permitem que a máquina extraia padrões e características mais sofisticadas dos dados. Os dados de entrada usados para a tarefa de classificação foram os espectrogramas extraídos dos arquivos de áudio de ambas as classes. Os espectrogramas foram inseridos em seis modelos de redes neurais convolucionais diferentes. A maior acurácia alcançada entre os modelos durante o treino foi de 99,15% e de 98,87% no conjunto de testes, resultados obtidos pelo mesmo modelo.

Palavras-chave: Categorização; Espectrogramas; Mosquitos; Som; Voo.

Abstract: This study aims to classify sound signals emitted during the flight of mosquitoes of the species *Culex quinquefasciatus* and *Anopheles Gambiae*. To achieve this objective, Deep Learning techniques were employed, which involve machine learning through the input of data and complex learning models called deep artificial neural networks, which process information at abstract levels and allow the machine to extract more sophisticated patterns and features from the data. The input data used for the classification task consisted of spectrograms extracted from the audio files of both classes. These spectrograms were fed into six different convolutional neural network models. The highest accuracy achieved among the model's during training was 99.15%, and 98.87% on the test set, results obtained by the same model.

Keywords: Categorization; Spectrograms; Mosquitoes; Sound; Flight.

INTRODUÇÃO

Existe no Brasil uma enorme diversidade de mosquitos, sendo o país que possui a maior quantidade de espécies de mosquitos, com aproximadamente 447 espécies (FOLEY; LEOPOLDO; WILKERSON, 2007). Devido a essa incidência de mosquitos no país, foram realizadas tentativas de combate a algumas espécies de mosquitos, isso porque existem determinadas espécies que são vetores, isto é, transmissores de doenças mortais para o homem.

As principais espécies que transmitem doenças no Brasil são: *Culex quinquefasciatus*, *Aedes aegypti* e *Anopheles gambiae*. Os mosquitos *Culex quinquefasciatus* estão presentes em todo território brasileiro, localizando-se em maior quantidade nas áreas populacionais devido a sua preferência por sangue humano, dado que essa espécie também se alimenta de aves, porém, quando o homem está próximo, ele será o alvo preferencial do mosquito. (CONSOLI; OLIVEIRA, 1994).

A espécie *Culex quinquefasciatus* possui relação com a transmissão de doenças como a filariose (Elefantíase) e a arbovirose conhecida como vírus Oropouche – muito semelhante a dengue. Os outros subgêneros da espécie *Culex* não são tão perigosos quanto o *quinquefasciatus* (CONSOLI; OLIVEIRA, 1994).

Da mesma forma que o *Culex quinquefasciatus*, a espécie *Aedes aegypti* também é encontrada em maior quantidade nas áreas populacionais e preferem se alimentar do sangue humano, mesmo conseguindo se alimentar de outros animais. Essas duas espécies são as mais domésticas no território brasileiro. No entanto, o *Aedes aegypti* é vetor de arboviroses como a febre amarela e o único vetor da dengue (CONSOLI; OLIVEIRA, 1994).

A espécie *Anopholes gambiae*, oriunda do continente africano, já invadiu o nordeste brasileiro, entre 1930 e 1940, e causou uma das mais graves epidemias de malária no Brasil (CONSOLI; OLIVEIRA, 1994). Dessa forma, o *Anopholes gambiae* é uma das espécies mais mortais de mosquitos que existem no Brasil, junto com a *Culex quinquefasciatus* e *Aedes aegypti*, dado que ele é transmissor da malária, doença que em 2021, segundo o Ministério da Saúde, teve 139.112 casos de malária no Brasil, sendo 99% dos casos na região Amazônica do país.

Dessa forma, fica evidente a necessidade de tomar medidas preventivas contra os mosquitos. No entanto, uma pesquisa realizada por Brassolatti e Andrade (2002) revelou que, entre os 92 participantes do estudo, 55,8% eram incapazes de identificar corretamente o *Aedes aegypti*. Outros participantes conseguiram descrevê-lo apenas parcialmente, mesmo depois de receberem treinamento. Isso evidencia a dificuldade das pessoas em reconhecer esse mosquito, o que é preocupante, pois a identificação correta é crucial para a prevenção e controle de doenças transmitidas por esses insetos. Dessa maneira, a tarefa de identificar os mosquitos vetores de doenças graves é um desafio significativo.

Com base neste contexto, o presente trabalho abordará a aplicação de técnicas de *Deep Learning (DL)* para obter uma possível solução que consiga satisfazer a necessidade de identificação do mosquito por meio do som que cada espécie emite durante o voo.

DL é um subcampo da inteligência artificial que se concentra em aprendizado profundo de máquina com representações de dados de múltiplos níveis de abstração. Isto é, a *DL* permite que diversos modelos computacionais compostos de várias camadas de processamento aprendam as suas representações a partir de dados (LECUN; BENGIO; HINTON, 2015). A

utilização de *DL* combinada com as Redes Neurais Convolucionais (*Convolutional Neural Networks* – *CNNs*) é amplamente utilizada para diversos problemas, como no artigo "*Thermographic image-based diagnosis of failures in electrical motors using deep transfer learning*" (DAHMER DOS SANTOS *et al.*, 2023) que apresentou o uso de *CNNs* e Aprendizagem Profunda por Transferência (*Deep Transfer Learning* – *DTL*) para criar um modelo de diagnóstico de falhas em motores elétricos a partir de imagens termográficas.

O artigo "*Web pages from mockup design based on convolutional neural network and class activation mapping*" (CIZOTTO *et al.*, 2023) também faz o uso de *CNNs* para resolver um problema complexo, que é a segmentação e classificação de elementos em páginas *web*, validando o uso das *CNNs* para a tarefa que foi designada no estudo. Essas abordagens também podem ser realizadas de modo a buscar a otimização em uma tarefa, como apresentado no artigo "*Element-Wise Attention Layers: An Option for Optimization*" (BACOCINA; DE SOUZA, 2023), que propõe um novo método de mecanismo de atenção e, para realizar a tarefa, utilizou dois modelos de redes neurais e obteve resultados promissores para sua técnica de otimização.

Na literatura, *DL* é aplicada à classificação de espécies de mosquitos, usando gravações de voo e *CNNs*, que são ideais para reconhecer padrões em imagens, pois serão utilizados os espectrogramas dos arquivos de áudio. Para efetivar essa tarefa, a pesquisa foi realizada em 7 etapas. Na 1ª etapa foi realizada a coleta do conjunto de dados; na 2ª etapa somente os dados necessários para a pesquisa foram coletados; na 3ª etapa foram extraídas características dos dados e realizada a normalização dos mesmos; na 4ª etapa foi feita a extração dos espectrogramas dos dados; na 5ª etapa os dados foram preparados para o treinamento dos modelos; na 6ª etapa os dados foram inseridos nos modelos para treinamento e testes; na 7ª etapa foi feita a comparação entre os resultados de cada modelo durante os treinos e testes.

Por fim, espera-se como resultado desta pesquisa, a capacidade de identificar os mosquitos vetores de doenças mortais, utilizando as técnicas de *DL*, através do espectrograma extraído do zumbido (áudio) emitido por eles durante o voo.

MATERIAIS E MÉTODOS

Para a pesquisa, utilizou-se o conjunto de dados disponibilizados pela Universidade de Oxford, através do projeto *HumBugDB*, que realiza a captação e identificação de mosquitos, na

África subsaariana, com o objetivo de catalogar espécies de mosquitos usando a assinatura acústica (som) de seus tons de voo. Toda a pesquisa desenvolvida pela *HumBugDB* está disponível no site oficial¹

O conjunto de dados está disponível no *GitHub* do projeto², ou pode ser acessado diretamente pelo repositório online *Zenodo*³. Esse conjunto de dados possui 71.286 segundos (20 horas) de gravações de 36 espécies diferentes de mosquitos voando, conforme apresentado na Figura 1. Também está disponível nos links anteriores o arquivo *.csv* (*Comma-separated values*). A pesquisa consistiu em 7 etapas, que estão presentes na Figura 2.

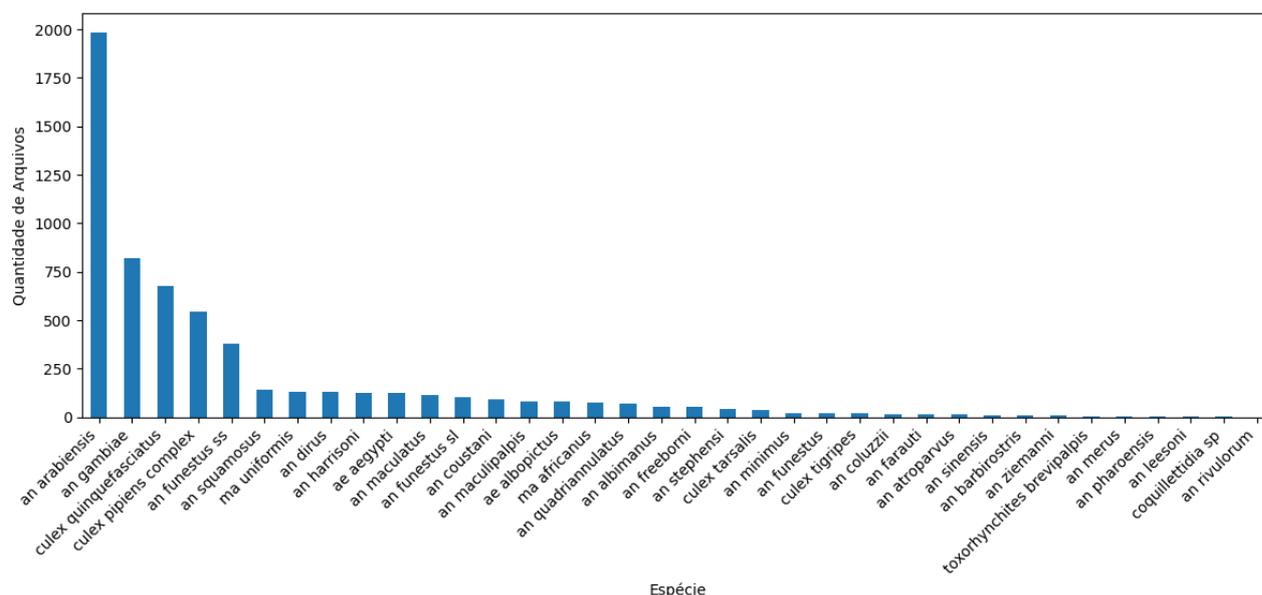


Figura 1. Quantidade de arquivos por espécie de mosquito

¹ Site do projeto HumBug: <https://humbug.ox.ac.uk>

² Repositório no GitHub do projeto HumBug: <https://github.com/HumBug-Mosquito/HumBugDB>

³ Repositório Zenodo do projeto HumBug: <https://zenodo.org/record/4904800>

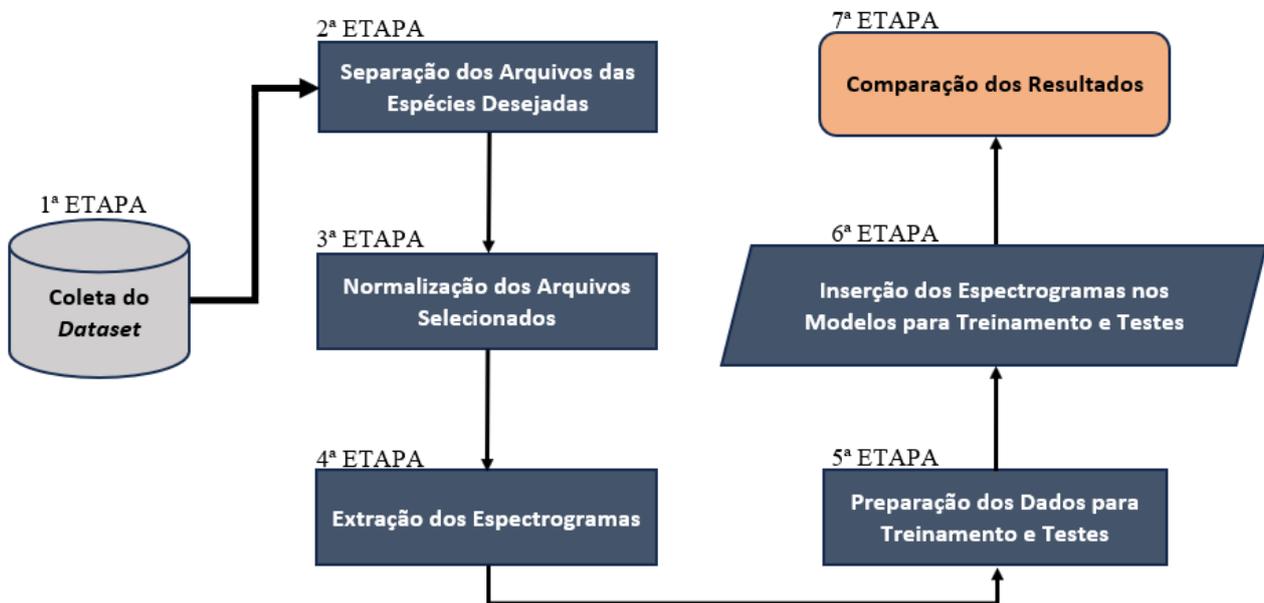


Figura 2. Fluxo seguido para realizar a pesquisa

Conforme a Figura 2, o conjunto de dados foi coletado e analisado na 1ª etapa. Foi verificado que o *dataset* (conjunto de dados) possui um total de 9.295 arquivos de áudio no formato *.WAV* rotulados com a sua respectiva espécie. Os arquivos de áudio que possuem o som de algum mosquito voando tem como rótulo a palavra “Mosquito”, enquanto arquivos sem som tem como rótulo a palavra “*Background*”. Foi analisado que 6.795 arquivos possuem sons de mosquito e 1.900 arquivos não possuem nenhum som de mosquito, isto é, o *dataset* é composto por, aproximadamente, 78% de sons de mosquitos.

Foi realizado, na segunda etapa, uma outra análise, mas dessa vez voltada para o contexto do nosso país, filtrando por arquivos de áudio rotulados como “Mosquito” e que fossem mosquitos comuns no Brasil. Existem 36 espécies de mosquitos no *dataset*, conforme apresentado na Figura 1 e, das 36 espécies, apenas 3 foram selecionadas para a pesquisa, a saber: *Culex quinquefasciatus* (678 áudios), *Anopholes gambiae* (818 áudios) e *Aedes aegypti* (123 áudios). A Figura 3 apresenta um gráfico de comparação da quantidade de arquivos entre as três espécies.

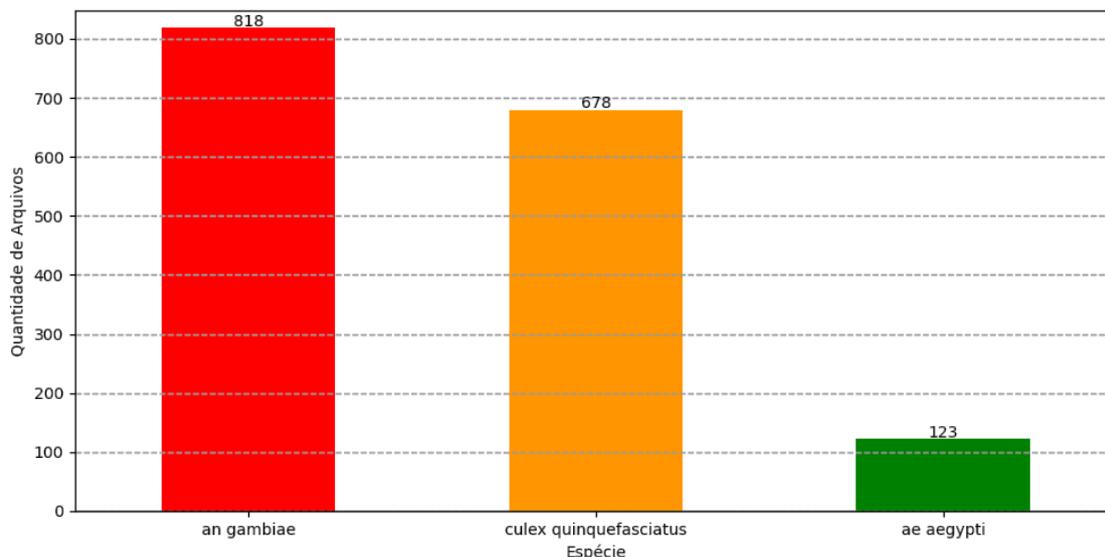


Figura 3. Quantidade de arquivos rotulados como “mosquito” das espécies selecionadas

Ainda na segunda etapa, entre as três espécies escolhidas, apenas as duas que mais possuem arquivos de áudios rotulados como “Mosquito” foram utilizadas na pesquisa, restando somente as espécies *Culex quinquefasciatus* e *Anopholes gambiae*. Após finalizar a segunda etapa, o conjunto de dados diminuiu aproximadamente 84% – passou de 9.295 áudios para 1.496 áudios, ou seja, utilizou-se apenas 16% dos áudios presentes no *dataset*.

A terceira etapa consistiu na extração de características pertinentes sobre os áudios, como taxa de amostragem das gravações e a duração de cada áudio. Foram encontrados dois tipos de taxa de amostragem nas gravações, 8 *kHz* e 44.1 *kHz*. Essa informação é importante pois, na literatura, as frequências que os mosquitos produzem som durante o voo, de acordo com Kiskin et al. (2018, apud Clements AN, 1999)⁴, estão entre 100 e 750 *Hz*. Isso significa que os áudios precisam ter uma taxa de amostragem que seja, pelo menos, o dobro de 100 a 750 *Hz*, portanto, devem ter taxas de amostragem de 200 a 1.5 *kHz*. Após obter as taxas de amostragem, foi necessário converter os áudios com 44.1 *kHz* para 8 *kHz*, o valor mais baixo obtido em todos os áudios.

A segunda informação extraída na terceira etapa foi a duração de todos os áudios de ambas as espécies. Para essa tarefa, foram consideradas as seguintes condições: áudios com 1

⁴ Clements AN (1999) The biology of mosquitoes, vol 2. CABI Publishing, Wallingford.

segundo de duração, áudios com duração de 2 segundo e áudios com duração superior a 2 segundos. As Tabelas 1 e 2 apresentam essas informações para cada espécie.

Tabela 1. Duração dos áudios da espécie *Culex quinquefasciatus*

Duração em segundos	Quantidade de áudios
Até 1 segundo	12
Entre 1 e 2 segundos	21
Maior que 2 segundos	645

Tabela 2. Duração dos áudios da espécie *Anopheles gambiae*

Duração em segundos	Quantidade de áudios
Até 1 segundo	132
Entre 1 e 2 segundos	196
Maior que 2 segundos	490

Para padronizar a duração dos áudios, primeiramente foi realizado a duplicação de áudios com 1 segundo de duração – que passaram a ter 2 segundos de duração. Em sequência, os áudios com duração superior a 2 segundos foram cortados em vários áudios com duração de 2 segundos. Dessa forma, todos os áudios do *dataset* passaram a ter a mesma duração. A Figura 4 apresenta o valor atualizado da quantidade de dados de áudio que cada espécie ficou após esse processo.

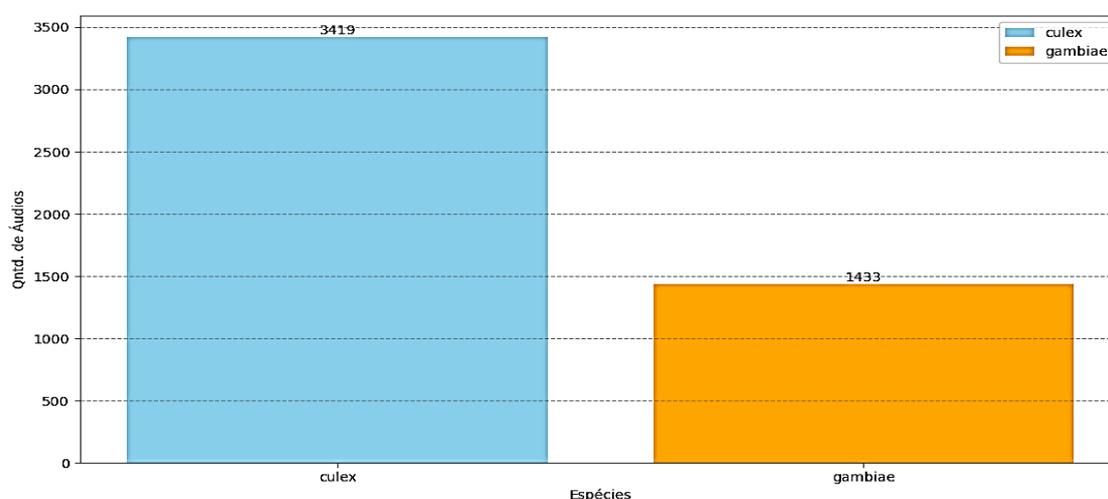


Figura 4. Nova quantidade de áudios após a padronização de 2 segundos

Devido a grande diferença entre a quantidade de dados de cada espécie, foi preciso novamente padronizar os dados. Dessa vez, para diminuir a lacuna entre as espécies, foi realizada a subamostragem (*undersampling*) da espécie *Culex quinquefasciatus*. *Undersampling* é uma técnica que consiste na redução da espécie majoritária de modo a ficar equilibrada com a espécie minoritária. Existem pontos positivos e negativos ao utilizar essa técnica, isso porque ela pode reduzir a quantidade de dados disponíveis para realizar o treinamento das redes neurais, no entanto, o *undersampling* pode também evitar um possível *overfitting* durante o treinamento (LIU; WU; ZHU, 2009). Após finalizada a técnica de *undersampling*, o conjunto de dados ficou equilibrado, isto é, ambas as espécies passaram a ter a mesma quantidade de arquivos de áudios (1.433 áudios).

Na 4ª etapa, foi utilizada a biblioteca librosa, que permite a utilização da Transformada Rápida de Fourier para extração dos espectrogramas de todos os arquivos de áudios de ambas as classes. Os parâmetros usados para a extração dos espectrogramas foram retirados da pesquisa de Fernandes, Cordeiro e Mendonza (2020).

A Figura 5 é um exemplo de espectrograma que foi extraído, estando armazenados no formato .png. Alguns espectrogramas, que não apresentavam nenhuma informação, foram retirados do *dataset*, o que explica a alteração na quantidade de dados na separação de treino e teste.

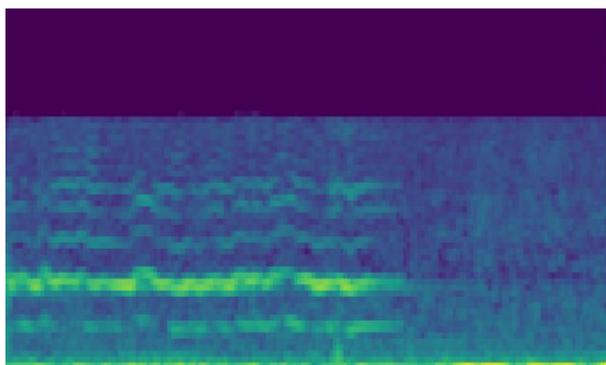


Figura 5. Espectrograma extraído de um áudio de 2 segundos que possui o som de um mosquito voando

Com os espectrogramas extraídos, foi feita a separação dos dados para treino e teste, conforme apresentado na 5ª etapa do fluxo mostrado na Figura 2. A separação foi realizada da seguinte maneira: 80% dos dados para treino e 20% dos dados para testes. A quantidade final, após a separação, está presente na Figura 6.

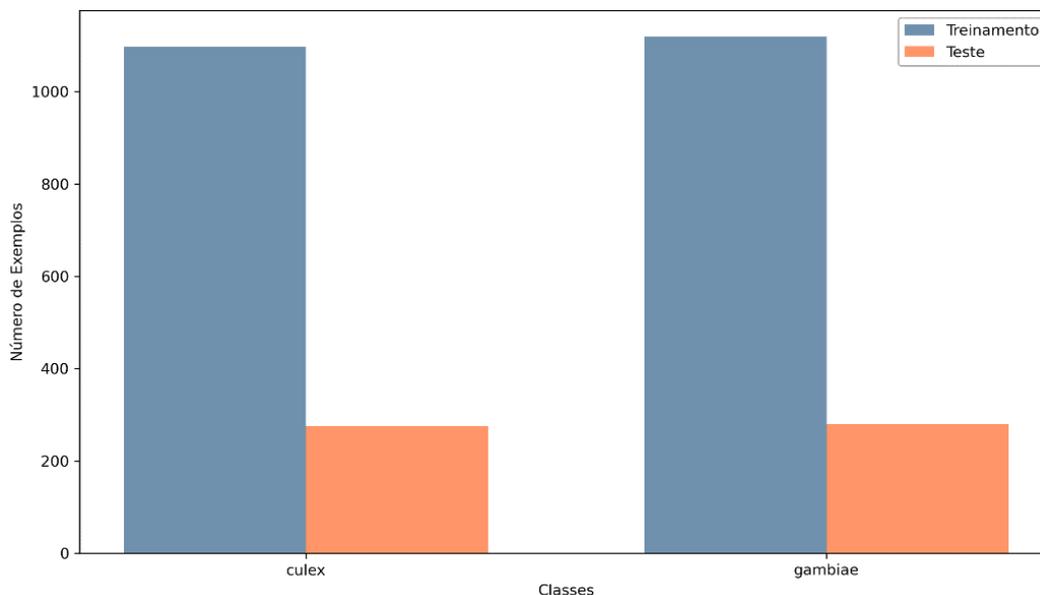


Figura 6. Distribuição de espectrogramas por classe nos conjuntos de treino e teste

A quantidade exata dos dados apresentados na Figura 6 é de 1.097 espectrogramas para a classe *Culex quinquefasciatus* e 1.120 espectrogramas para a classe *Anopheles gambiae* no conjunto de treinamento. No conjunto de teste a classe *Culex* possui 275 espectrogramas e a classe *Anopheles* possui 280 espectrogramas.

Conforme apresentado na Figura 2, a 6ª etapa tem o objetivo de treinar os modelos de redes neurais. A primeira rede neural treinada foi um modelo de classificação simples binário, apresentado na Figura 7. Esse modelo foi escolhido com base na pesquisa realizada por Fernandes, Cordeiro e Mendonza (2020). Na pesquisa em questão, a espécie classificada foi a *Aedes aegypti* e alcançaram uma acurácia de 97.65% com um modelo de classificação binário.

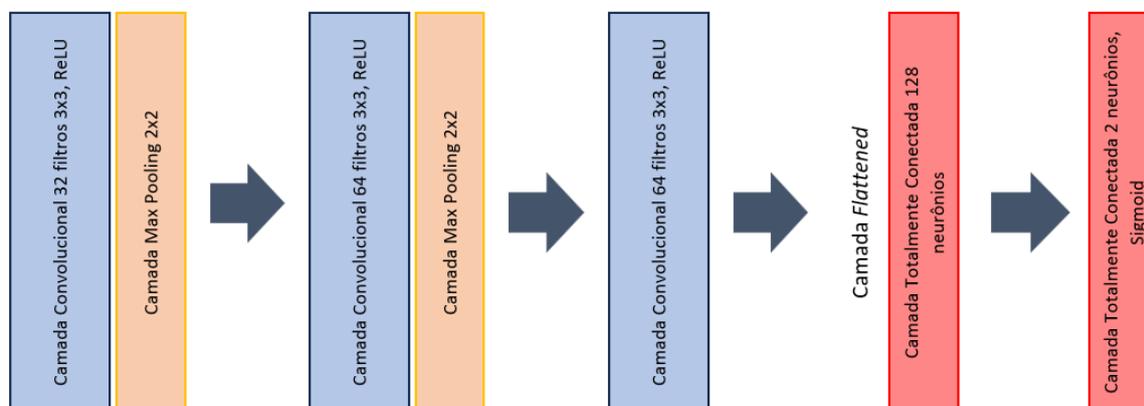


Figura 7. Modelo de classificação binaria

O modelo representado na Figura 7 passou por um treinamento de 10 épocas, utilizando imagens de entrada com dimensões de 60x60 pixels e três cores. Durante esse processo obtiveram-se as seguintes métricas na sua melhor época: acurácia (*accuracy*) de 99,15%, acurácia de validação de 98,88% (*val_accuracy*), perda (*loss*) de 0.036 e uma perda de validação (*val_loss*) de 0.035. A Figura 8 apresenta a evolução do modelo durante o seu treinamento.

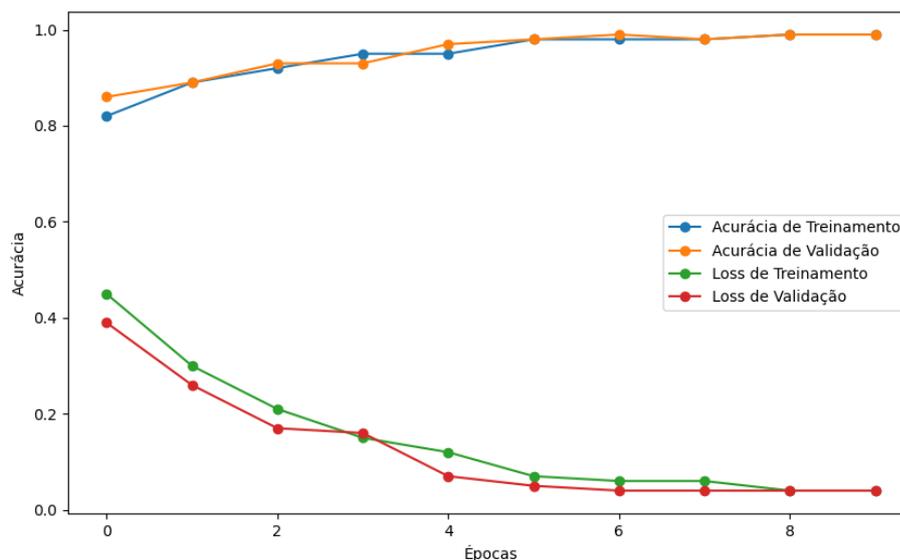


Figura 8. Métricas de treinamento do Modelo Binário

Após o treinamento, adicionalmente, realizou-se uma avaliação do desempenho do modelo através de uma matriz de confusão aplicada aos dados de teste. Os resultados revelaram uma taxa de acerto de aproximadamente 98,20%, como ilustrado na Figura 9.

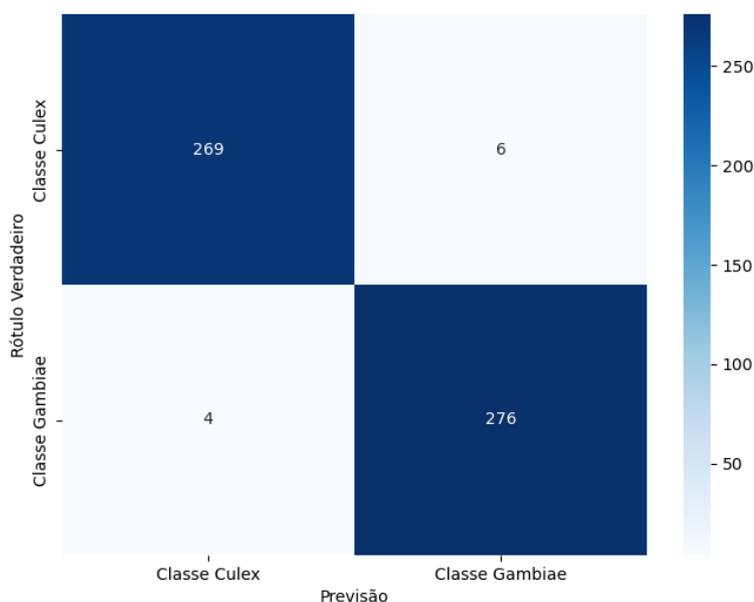


Figura 9. Matrix de confusão do Modelo Binário

O segundo modelo treinado foi o *VGG-16*, que foi customizado para resolver o problema de classificação binária. A modificação consistiu na mudança de sua camada *Flatten*, que foi substituída por uma camada totalmente conectada contendo 256 neurônios e, na camada de saída, que passou a ter apenas 2 neurônios.

Da mesma forma que o modelo binário, o *VGG-16* foi submetido a um treinamento que ocorreu por 10 épocas, com os seguintes resultados observados durante o ápice de seu desempenho durante o treinamento: acurácia de 91,14%, acurácia de validação de 90,99%, perda de 0.221 e perda de validação de 0.233. A Figura 10 apresenta a evolução do modelo *VGG-16* durante o seu treinamento.

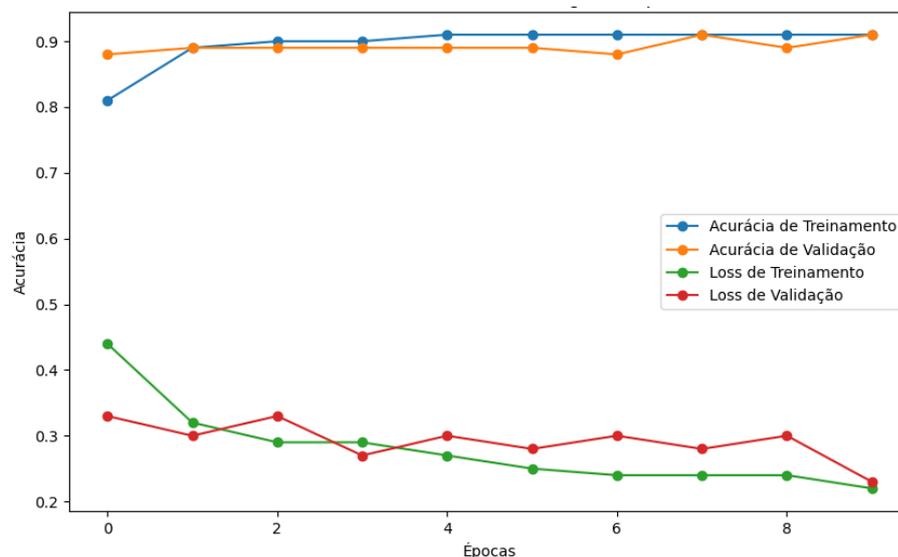


Figura 10. Métricas de treinamento do modelo *VGG-16*

O modelo *VGG-16* também foi avaliado por meio de uma matriz de confusão, utilizando os mesmos dados de teste empregados no modelo binário. Os resultados demonstraram uma taxa de acerto nas previsões de aproximadamente 91,18%, conforme apresentado na Figura 11.

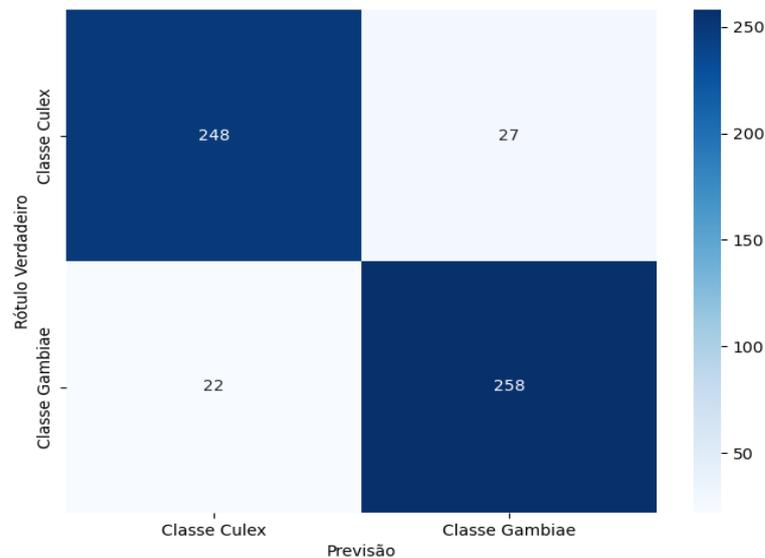


Figura 11. Matrix de confusão do modelo *VGG-16*

O terceiro modelo treinado foi o *VGG-19*, submetido ao mesmo conjunto de dados e configurações empregadas no modelo *VGG-16*. Isso incluiu a manutenção da camada *Flatten*. O melhor resultado alcançado pelo *VGG-19* foi o seguinte: uma acurácia de 91,26%, uma acurácia de validação de 91,89%, uma perda de 0.237 e uma perda de acurácia de 0.299. A Figura 12 apresenta a evolução do modelo *VGG-19* durante o seu treinamento.

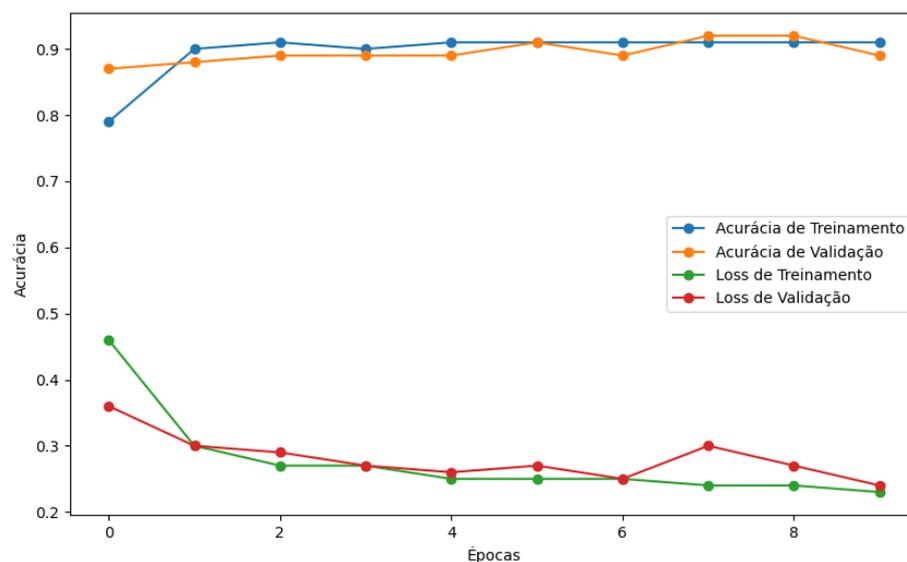


Figura 12. Métricas de treinamento do modelo *VGG-19*

Da mesma forma que os modelos anteriores, o *VGG-19* também foi avaliado através de uma matriz de confusão, com os mesmos dados de teste utilizados anteriormente, obtendo uma taxa de acerto de 91%, conforme ilustrado na Figura 13.

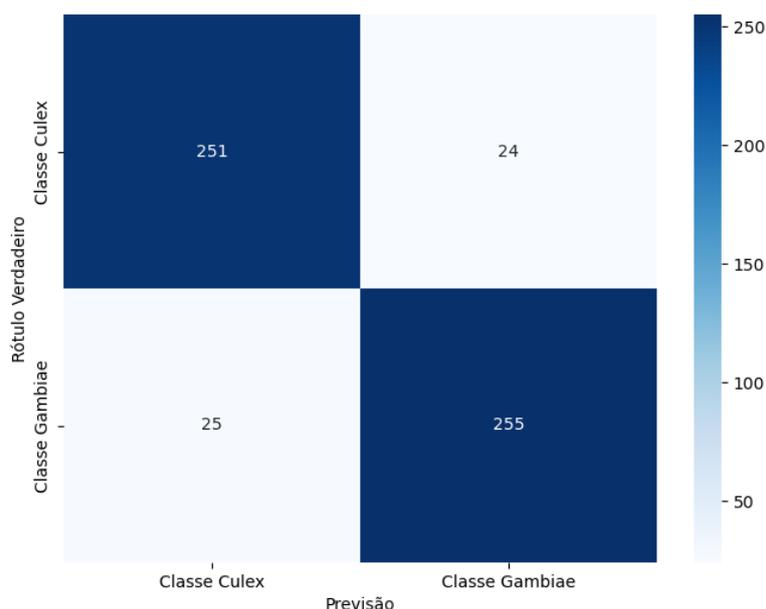


Figura 13. Matrix de confusão do modelo *VGG-19*

O quarto modelo treinado foi o *ResNet50*, submetido aos mesmos procedimentos de treinamento adotados nos modelos anteriores. O melhor desempenho registrado com o *ResNet50* foi o seguinte: precisão de 80,37%, precisão de validação de 88,06%, perda de 0.482 e uma perda de validação de 0.440. A Figura 14 apresenta a evolução do modelo *ResNet50* durante o seu treinamento.

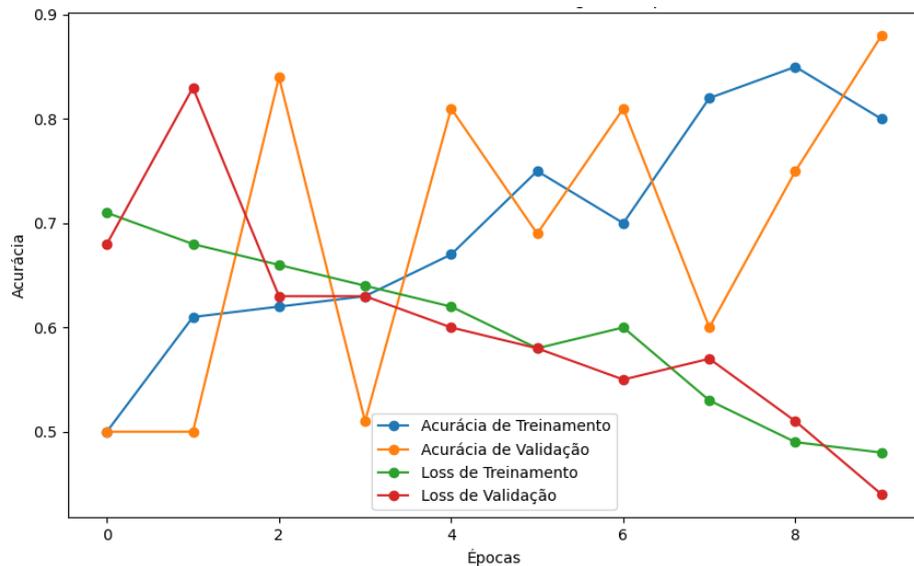


Figura 14. Métricas de treinamento do modelo *ResNet50*

O modelo *ResNet50*, após treinado, também foi submetido a uma matriz de confusão, da mesma forma que os modelos anteriores. Na matriz de confusão, o modelo obteve uma taxa de acerto de 88%, conforme apresentado na Figura 15.

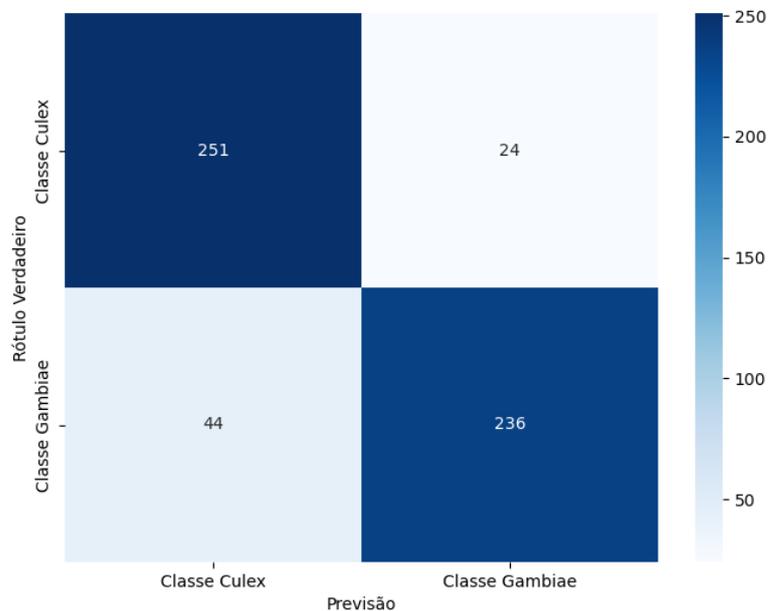


Figura 15. Matrix de confusão do modelo *ResNet50*

O quinto e último modelo treinado foi o *ResNet50v2*, que seguiu os mesmos procedimentos de treinamento adotados nos modelos anteriores. O resultado mais notável alcançado pelo *ResNet50v2* foi: precisão de 98,25%, precisão de validação de 92,34%, perda

de 0.047 e uma perda de validação de 0.288. A Figura 16 apresenta a evolução do modelo *ResNet50v2* durante o seu treinamento.

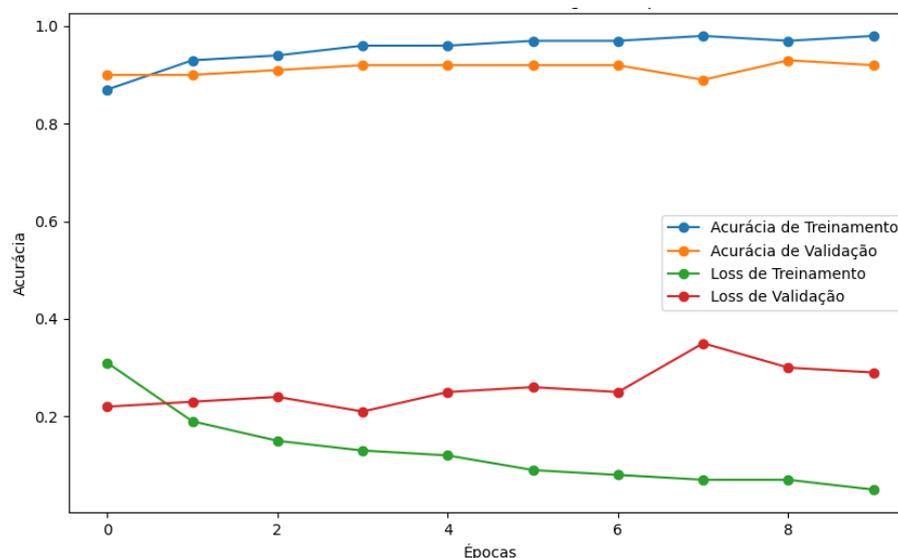


Figura 16. Métricas de treinamento do modelo *ResNet50v2*

Ao avaliar o desempenho do modelo *ResNet50v2*, por meio de uma matriz de confusão, utilizando os mesmos dados de teste empregados nos modelos apresentados anteriormente, observou-se uma taxa de acerto de 93%, conforme apresentado na Figura 17.

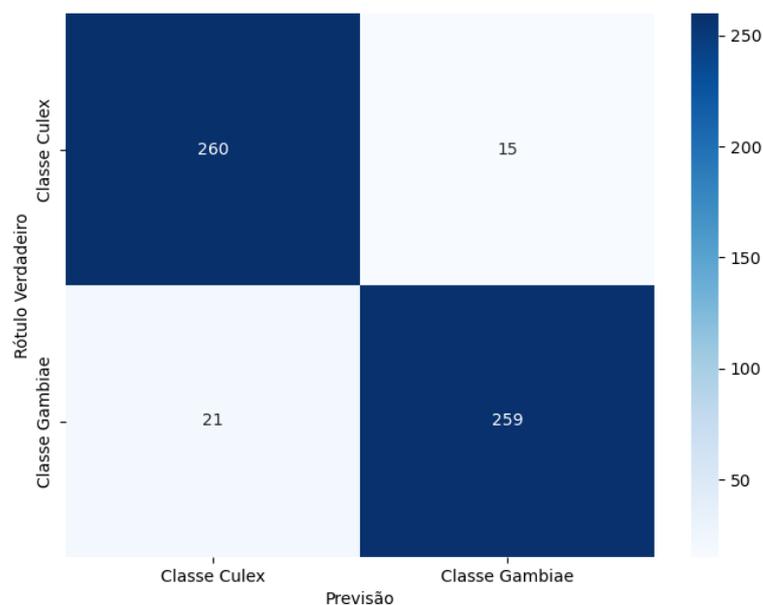


Figura 17. Matrix de confusão do modelo *ResNet50v2*

RESULTADOS E DISCUSSÃO

A sétima e derradeira etapa da pesquisa consistiu na comparação dos resultados dos modelos treinados, com o objetivo de determinar qual deles obteve o desempenho mais destacado em relação ao conjunto de dados empregado no estudo. Os resultados de cada modelo foram sintetizados e resumidos na Tabela 3, que fornece uma comparação tabular para simplificar a análise.

Tabela 3. Comparação entre os resultados dos modelos durante o treinamento

Modelo	Acurácia (%)	Val_acurácia (%)	Perda	Val_perda	Acurácia Matriz de Conf. (%)
Binário	99,15	98,87	0.036	0.035	98,20
<i>VGG-16</i>	91,14	90,99	0.221	0.233	91,18
<i>VGG-19</i>	91,26	91,89	0.237	0.237	91,00
<i>ResNet50</i>	80,37	88,06	0.482	0.482	87,74
<i>ResNet50v2</i>	98,25	92,34	0.047	0.047	93,51

Entre os cinco modelos treinados na pesquisa, o que obteve o melhor desempenho, tanto no treino quanto na matriz de confusão, foi o modelo binário proposto por Fernandes, Cordeiro e Mendonza (2020), que usaram esse mesmo modelo em um problema de classificação binária de sinais de áudio e alcançaram uma acurácia de 97,65%. O modelo binário teve uma acurácia de validação de 98,87% e 98,20% na matriz de confusão, com apenas 0.035 na perda de validação – o menor valor entre todos os modelos.

Já os modelos *VGG-16* e *VGG-19* alcançaram uma acurácia na validação superior a 90% e na matriz de confusão ambos obtiveram uma acurácia de 91,18%. Enquanto o modelo *ResNet50V2* obteve uma acurácia na validação de 92,34% e 93,51% na matriz de confusão, superando os dois modelos *VGG*. Por fim, o modelo *ResNet50* apresentou resultados inferiores quando comparado com os outros modelos, atingindo apenas 88,06% na acurácia de validação e 87,74% na matriz de confusão.

Apesar das informações utilizadas para comparar os modelos, algumas outras métricas poderiam ser levadas em consideração, como *F1_score*, *precision* e *recal*. Além disso, seria uma boa prática para esta pesquisa a aplicação de filtros para diminuir a presença de ruídos de fundo nos áudios, o que poderia aumentar a precisão de todos os modelos treinados. Por fim, a

validação cruzada é outra técnica que poderia ter sido aplicada na pesquisa para ter melhores resultados, assim como a utilização do *cepstrum* de frequência mel (*MFCC*) dos dados ao invés do espectrograma, dando a possibilidade de comparar os modelos treinados com espectrogramas e os modelos treinados com *MFCCs*.

CONCLUSÃO

Com base nos resultados obtidos, pode-se afirmar que a tarefa de classificação das espécies de mosquitos *Culex quinquefasciatus* e *Anopholes gambiae* apresentou acurácias de validação superiores a 90% em cinco dos seis modelos avaliados. O único modelo que não alcançou esse patamar foi o *ResNet50*. Entre os cinco modelos que obtiveram resultados superiores a 90%, o modelo binário se destacou como o mais adequado para a tarefa de classificação, considerando os resultados, os dados e a metodologia empregada na pesquisa.

Desse modo, o processo utilizado na realização da pesquisa demonstrou ser eficaz, pois alcançou o objetivo de classificação proposto, mesmo que tenham sido identificadas algumas oportunidades para a implementação de técnicas que poderiam aprimorar ainda mais os resultados finais.

REFERÊNCIAS

BACCHINA, Giovanni Araujo; THOM DE SOUZA, Rodrigo Clemente. Element-Wise Attention Layers: An Option for Optimization. Disponível em: <https://ssrn.com/abstract=4460607>

BRASIL. Ministério da Saúde. **Situação Epidemiológica da Malária**. Disponível em: <https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/m/malaria/situacao-epidemiologica-da-malaria>. Acesso em: 06 out. 2022.

BRASSOLATTI, Rejane Cristina; ANDRADE, Carlos Fernando S. Avaliação de uma intervenção educativa na prevenção da dengue. **Ciência & Saúde Coletiva**, v. 7, n. 2, p. 243-251, 2002. ISSN 1413-8123. DOI: 10.1590/S1413-81232002000200005.

CONSOLI, Rotraut A.G.B.; OLIVEIRA, Ricardo Lourenço de. **Principais mosquitos de importância sanitária no Brasil**. Rio de Janeiro: Editora Fiocruz, 1994. 228p.

CIZOTTO, André *et al.* Web pages from mockup design based on convolutional neural network and class activation mapping. **Multimedia Tools and Applications**, v. 82, mar. 2023. DOI: 10.1007/s11042-023-15108-3.

DAHMER DOS SANTOS, Luiz Fillipe *et al.* Thermographic image-based diagnosis of failures in electrical motors using deep transfer learning. **Engineering Applications of Artificial Intelligence**, v. 126, p. 107106, 2023. ISSN 0952-1976. DOI: <https://doi.org/10.1016/j.engappai.2023.107106>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0952197623012903>.

FERNANDES, Marcelo Schreiber; CORDEIRO, Weverton; RECAMONDE-MENDOZA, Mariana. Detecting *Aedes aegypti* mosquitoes through audio classification with convolutional neural networks. **Computers in Biology and Medicine**, v. 129, p. 104152, 2021. ISSN 0010-4825. DOI: <https://doi.org/10.1016/j.compbiomed.2020.104152>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0010482520304832>.

FOLEY, Desmond; RUEDA, Leopoldo; WILKERSON, Richard. Insight into Global Mosquito Biogeography from Country Species Records. **Journal of Medical Entomology**, v. 44, n. 8, p. 554-567, agosto de 2007. DOI: 10.1603/0022-2585(2007)44[554:IIGMBF]2.0.CO;2.

KISKIN, Ivan; ZILLI, Davide; LI, Yunpeng; SINKA, Marianne; WILLIS, Kathy; ROBERTS, Stephen. Bioacoustic detection with wavelet-conditioned convolutional neural networks. **Neural Computing and Applications**, v. 32, n. 4, p. 915-927, fevereiro de 2020. ISSN 1433-3058. DOI: 10.1007/s00521-018-3626-7.

LECUN, Yann; BENGIO, Yoshua; HINTON, Geoffrey. Deep learning. **Nature**, v. 521, n. 7553, p. 436-444, maio de 2015. ISSN 1476-4687. DOI: 10.1038/nature14539.

LIU, Xu-Ying; WU, Jianxin; ZHOU, Zhi-Hua. Exploratory Undersampling for Class-Imbalance Learning. **IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)**, v. 39, n. 2, p. 539-550, 2009. DOI: 10.1109/TSMCB.2008.2007853.