

1.1. Conjunto de dados:

Para a pesquisa, utilizou-se os dados disponibilizados pela Universidade de Oxford, através do projeto HumBug. O projeto HumBug 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 HumBug está disponível em um repositório do GitHub (<https://github.com/HumBugMosquito/HumBugDB>), que contém o artigo produzido pelo projeto, os notebooks, os dados e outras informações sobre a pesquisa. Além do GitHub, o projeto HumBug possui um site (<https://humbug.ox.ac.uk>) onde apresentam a história do projeto, a motivação e as atualizações do mesmo.

O conjunto de dados utilizados pela HumBug foi coletado da plataforma online Zooniverse (<https://www.zooniverse.org>), que possui áudios de mosquitos com duração de 1 e 2 segundos. O dataset está disponível no GitHub da HumBug (<https://github.com/HumBugMosquito/ZooniverseData>). Esse conjunto de áudios possui outros dois conjuntos, um conjunto contém áudios com duração de 1 segundo e o outro conjunto contém áudios com duração de 2 segundos. No presente projeto foi utilizado o mesmo dataset, porém, somente com áudios que possuem duração de 2 segundos (http://humbug.ac.uk/public/Zooniverse_audio_2sec.zip).

Os mosquitos catalogados pela HumBug foram divididos em quatro classes: A, B, C, D. Cada classe possui uma, ou mais, respectiva espécie de mosquito, conforme a Tabela 1.

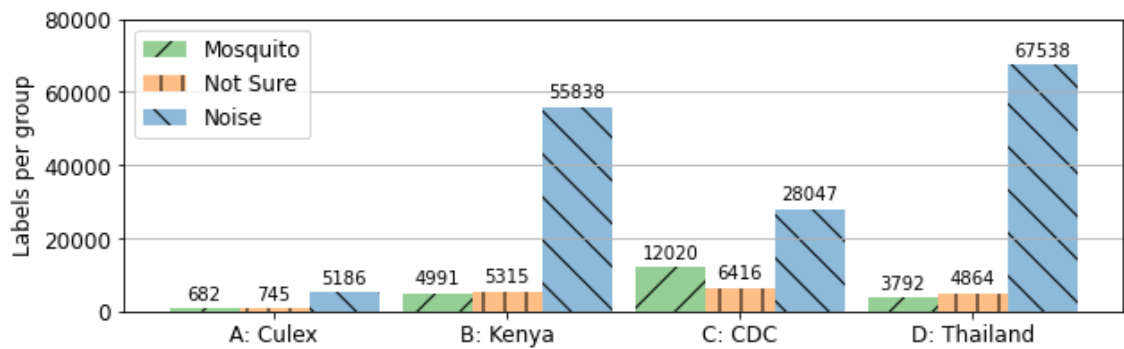
Tabela 1 – Classificação dos sons das espécies.

Classes	Espécie do Mosquito
A	Culex Quinquefasciatus
B	Anopholes Gambiae
C	Aedes Aegypti e Aedes Albopictus
D	Anopholes Barbirostris e Anopheles Macalatus

Fonte: O autor.

O Projeto HumBug também disponibilizou, em seu GitHub, a visualização gráfica da quantidade de áudios categorizados das 4 classes. As classificações possíveis são: mosquito, background e not_sure. Tem-se, ao total, 198.434 áudios de 2 segundos.

Figura 1 – Gráfico apresentando os áudios com duração de 2 segundos.



Fonte: HumBug (GitHub).

Conforme a legenda da Figura 1, nota-se as três categorias de áudios existentes – *mosquito*, *noise* e *not sure* –, como também as quatro classes A, B, C, D. Áudios classificados como mosquito possuem um som de mosquito, áudios classificados como noise não possuem som de mosquito – ruído, em tradução literal, são áudios que captaram o som ambiente. Os áudios classificados como not_sure são áudios que não conseguiram definir se possui ou não o som de um mosquito.

Percebe-se a diferença entre a quantidade de áudios classificados como noise – que pode ser entendido como áudios do tipo background – em comparação as outras duas classificações. Em todas as classes, a quantidade de áudio classificado como noise é superior a quantidade de áudios classificados como not_sure e mosquito. A quantidade de áudios classificados como mosquito é, aproximadamente, 11% dos áudios presentes – total de 198.434 áudios.

Além da imagem mostrada na Figura 1, também está disponível no GitHub o arquivo CSV (Comma-separated values) dos arquivos de áudio com duração de 2 segundos. Por meio dos respectivos atributos no cabeçalho: `audio_id`, `id`, `labeller_id`, `zooniverse_id`, `sound_type`, `path`, `subject_set`. Através da coluna `subject_set`, definida no cabeçalho do arquivo CSV, é possível distinguir qual a classe do mosquito presente naquele áudio. Como também é possível distinguir se um áudio é ou não de mosquito por meio da coluna `sound_type`, onde existem três possibilidades: *mosquito*, *background*, *not_sure*.

Figura 2 – Fragmento das informações contidas no CSV

```
audio_id,id,labeller_id,zooniverse_id,sound_type,path,subject_set
60453,160289,956,136608209,background,/Zoo_segment/sounds/0001_norm_length_2_interval_1_clip_1.wav,Culex Q
60454,65550,680,120066497,background,/Zoo_segment/sounds/0001_norm_length_2_interval_1_clip_10.wav,Culex Q
60454,132598,1107,128458354,background,/Zoo_segment/sounds/0001_norm_length_2_interval_1_clip_10.wav,Culex Q
60454,119446,1180,126931734,background,/Zoo_segment/sounds/0001_norm_length_2_interval_1_clip_10.wav,Culex Q
60454,115014,293,126443879,background,/Zoo_segment/sounds/0001_norm_length_2_interval_1_clip_10.wav,Culex Q
```

Fonte: o autor.

A classificação *background* pode ser interpretada como a classificação *noise*. Ambas as classificações possuem o mesmo significado: o áudio não é o som de um mosquito.

3.2. Metodologia

Foi utilizado no projeto um notebook Asus com as seguintes especificações de hardware: Processador AMD Ryzen 5 3500U, placa gráfica Radeon Vega 8, 8 gigabytes de memória RAM e um SSD NVME de 256 gigabytes. Em relação aos softwares utilizados: Jupyter Notebook.

Com os dados coletados (conforme o tópico 3.1), foi realizada, antes de qualquer etapa, a análise de quais são os mosquitos mais comuns no Brasil e quais desses mosquitos estão presentes nas quatro classes do projeto HumBug. Adicionar aqui algo que fundamente a escolha da próxima linha (referência) – Concluiu-se que as espécies presentes são: GRUPO A - *Culex Quinquefasciatus*; GRUPO B - *Anopheles Gambiae*; GRUPO C - *Aedes Aegypti*.

Em seguida, foram separados apenas os dois grupos que mais possuem áudios classificados como mosquito. Essa separação resultou em um total de 17.011 áudios classificados como mosquito, cerca de 8,57% dos áudios presentes no arquivo CSV.

Tabela 2 - Classes selecionadas e a quantidade de áudios

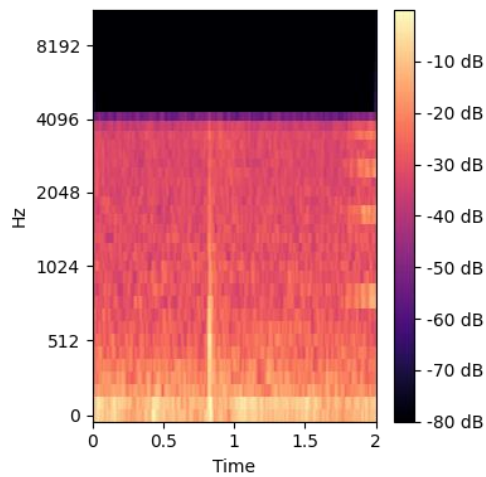
Classe	Quantidade de áudios
B: Kenya	3.766
C: CDC	7.163

Fonte: o autor

Após a separação dos 17.011 áudios, foi necessário balancear o conjunto de áudios devido a maior quantidade de dados presente no grupo C. Para isso, foi realizado o processo de subamostragem no grupo C, que passou a ter 3.766 áudios. Ficando assim, o conjunto de dados, balanceado, com 3.766 áudios para cada classe, totalizado 7.532 áudios.

Em seguida, foram extraídos os espectrogramas, em escala de mel, dos 7.352 áudios. N_bands = 40, n_frames = 40, hop_length = 10, win_length = 1024.

Figura 3 – Exemplo de espectrograma



Fonte: o autor

A partir desse ponto, em uma análise manual de todos os espectrogramas extraídos, foi possível encontrar 121 áudios que não seguiam o padrão dos outros espectrogramas. Ao analisar esses 121 áudios, foi concluído que não se tratavam de áudios de mosquito, mas sim de áudios que variavam entre vozes humanas, sons ambiente, ruídos altos, sons de outros animais. Dos 121 áudios analisados, 25 eram da classe C e 95 da classe B. Todos esses áudios foram removidos do dataset, assim como os seus respectivos espectrogramas. O conjunto de dados passou a ter 7.411 dados.

Tabela 3 – Conjunto de dados após os processos

Classe	Quantidade de áudios/espectrogramas
B: Kenya	3.672
C: CDC	3.739

Fonte: o autor.

Por fim, feita as análises manuais (exclusão de áudios que não possuíam sons de mosquito) e automatizadas (extração dos espectrogramas). Foi feita a divisão dos dados de treino, validação e teste, utilizando os espectrogramas obtidos anteriormente. A proporção de divisão foi de 60%, 20%, 20%.

Tabela 4 - Dados de treino, teste e validação da classe B

Classe B	Quantidade de espectrogramas
Treino	2.203
Teste	735
Validação	734

Fonte: o autor.

Tabela 5 - Dados de treino, teste e validação da classe C

Classe C	Quantidade de espectrogramas
Treino	2.243
Teste	749
Validação	747

Fonte: o autor.

```
Found 4446 images belonging to 2 classes.
Found 1481 images belonging to 2 classes.
Found 1484 images belonging to 2 classes.
```

Após a divisão de dados para treino, teste e validação, foi feito o carregamento desses dados por meio da biblioteca Keras para preparar os geradores de imagens, mais especificamente a ImageDataGenerator. O intuito em utilizar essa biblioteca está na aplicação de transformações nos espectrogramas, por exemplo: zoom, rotação, deslocamento, espelhamento.

Com as transformações já realizadas, os dados foram inseridos em uma rede neural convolucional básica, que possui uma camada de convolução, seguida de uma camada pooling, uma camada densa com 64 neurônios, e uma camada de saída com um neurônio. Segue os resultados obtidos por esse modelo, com 10 épocas:

Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 126, 126, 32)	896
max_pooling2d (MaxPooling2D)	(None, 63, 63, 32)	0
flatten (Flatten)	(None, 127008)	0
dense (Dense)	(None, 64)	8128576
dense_1 (Dense)	(None, 1)	65

=====
Total params: 8,129,537
Trainable params: 8,129,537
Non-trainable params: 0

```
Epoch 1/10
139/139 [=====] - 111s 790ms/step - loss: 0.6965 - accuracy: 0.8207 - val_loss: 0.1424 - val_accuracy: 0.9568
Epoch 2/10
139/139 [=====] - 106s 760ms/step - loss: 0.1417 - accuracy: 0.9600 - val_loss: 0.1126 - val_accuracy: 0.9662
Epoch 3/10
139/139 [=====] - 103s 738ms/step - loss: 0.1070 - accuracy: 0.9696 - val_loss: 0.0846 - val_accuracy: 0.9676
Epoch 4/10
139/139 [=====] - 104s 749ms/step - loss: 0.0843 - accuracy: 0.9750 - val_loss: 0.0538 - val_accuracy: 0.9723
Epoch 5/10
139/139 [=====] - 105s 751ms/step - loss: 0.0689 - accuracy: 0.9753 - val_loss: 0.0425 - val_accuracy: 0.9797
Epoch 6/10
139/139 [=====] - 106s 761ms/step - loss: 0.0617 - accuracy: 0.9768 - val_loss: 0.0326 - val_accuracy: 0.9872
Epoch 7/10
139/139 [=====] - 104s 749ms/step - loss: 0.0508 - accuracy: 0.9825 - val_loss: 0.0384 - val_accuracy: 0.9865
Epoch 8/10
139/139 [=====] - 104s 746ms/step - loss: 0.0427 - accuracy: 0.9861 - val_loss: 0.0186 - val_accuracy: 0.9946
Epoch 9/10
139/139 [=====] - 109s 782ms/step - loss: 0.0387 - accuracy: 0.9879 - val_loss: 0.0209 - val_accuracy: 0.9939
Epoch 10/10
139/139 [=====] - 107s 767ms/step - loss: 0.0380 - accuracy: 0.9876 - val_loss: 0.0154 - val_accuracy: 0.9946
```

```
47/47 [=====] - 15s 319ms/step - loss: 0.0125 - accuracy: 0.9960  
Precisão no conjunto de teste: 0.9959568977355957
```

Já, sua accuracy por meio da matriz de confusão foi de 0,4952.

REFERÊNCIAS

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

DOSHI, Ketan. Audio Deep Learning Made Simple: Sound Classification, Step-by-Step. [S. l.], 18 jun. 2021. Disponível em: <https://towardsdatascience.com>. Acesso em: 28 nov. 2022.

BIOSSEG UNIPRAG. Conheça as espécies de mosquitos no Brasil. [S. l.], 10 jun. 2018. Disponível em: <https://biosseg.com.br>. Acesso em: 20 out. 2022.

Fernandes MS, Cordeiro W and Recamonde-Mendoza M, Detecting Aedes aegypti mosquitoes through audio classification with convolutional neural networks. Comput Biol Med 129:104152 (2021).

<https://github.com/tomfran/urban-sound-classification> (não achei a publicação em nenhum lugar)

<https://vijayabhaskar96.medium.com/tutorial-image-classification-with-keras-flow-from-directory-and-generators-95f75ebe5720>