



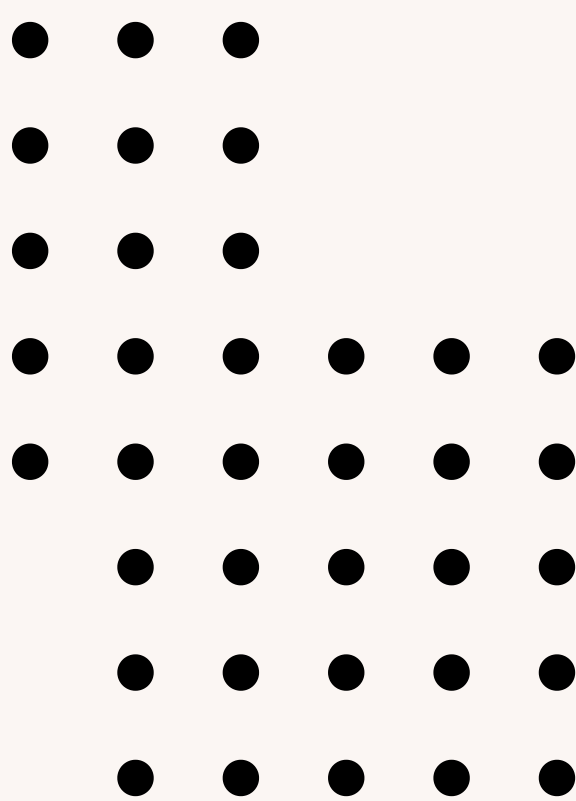
CLASSIFICAÇÃO DE ESPÉCIES DE AVES POR MEIO DA EXTRAÇÃO DE CARACTERÍSTICAS DE SEUS MEL-ESPECTROGRAMAS

Felipe Archanjo da Cunha Mendes (Orientado)
Prof. Dr. Juliano Henrique Foleis (Orientador)

23/11/2023



INTRODUÇÃO

- Biodiversidade, sustentabilidade, equilíbrio dos ecossistemas
 - Importância das aves
 - Polinização (CARPENTER, 1978)
 - Controle de Pragas (HOLMES, 1990)
 - Regulação do ciclo de nutrientes (SNOW, 1971)
- 



PROBLEMA

- Classificação correta de espécies (FIGUEIREDO, 2018)
 - Desafio complexo
 - Experiência por parte dos ornitólogos
 - características morfológicas ou sonoras
- Abordagem manual é demorada, propensa a erros e dependente do conhecimento e habilidades dos especialistas
- Demanda crescente por métodos mais eficientes



TRABALHOS ANTERIORES

- Classificação automática de espécies com base em características sonoras
 - Banco de dados robusto
 - Pré-processamento
 - Geração de espectrogramas
 - Descritores de características
 - Handcrafted (LUCIO; COSTA, 2016)
 - TransferLearning (INCZE et al., 2018)
 - Classificadores

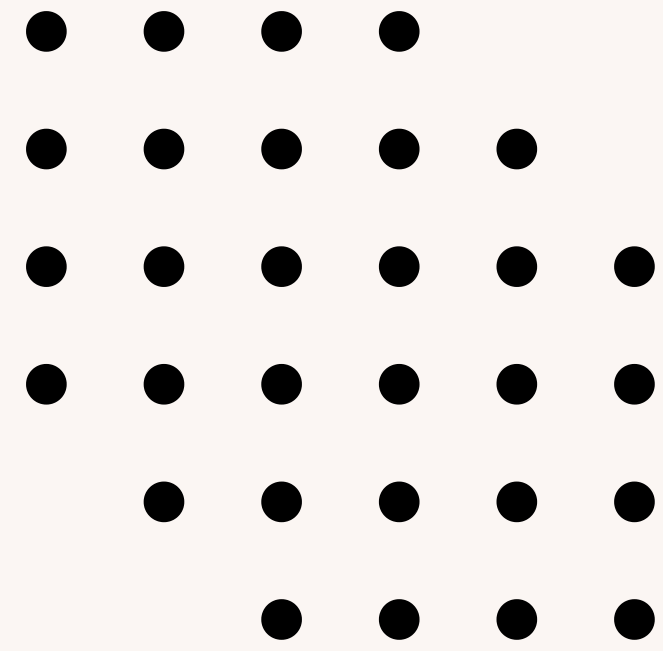


OBJETIVO

- Desenvolver um modelo de aprendizado de máquina, visando a classificação precisa de diversas espécies de aves por meio da gravação de seus cantos e chamados.

DATASET

- Kaggle
- BirdCLEF 2023 (BIRDCLEF, 2023)
- Desbalanceado (Figura 1)
- 16900 gravações
- 264 diferentes espécies



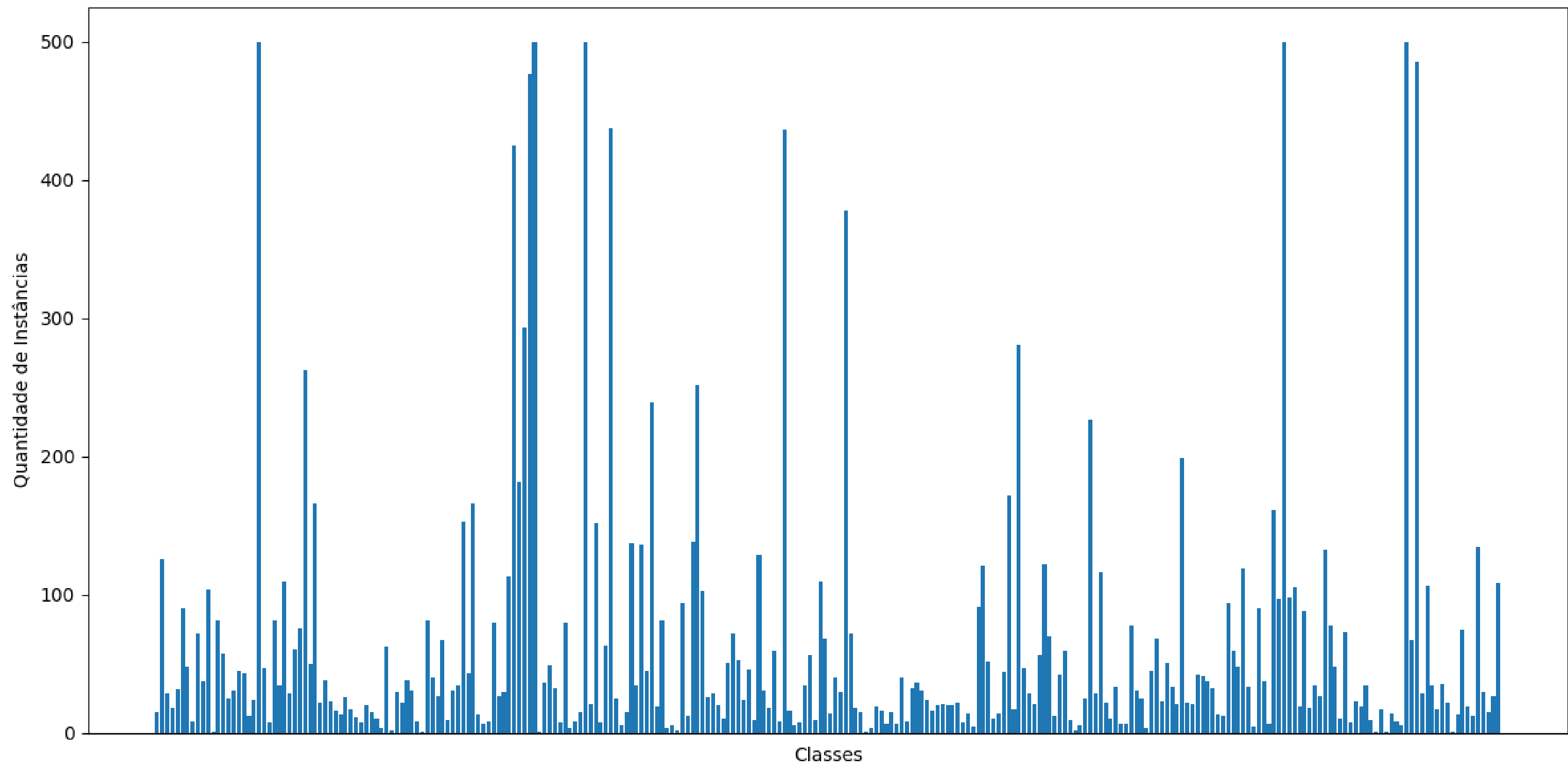


Figura 1 – Distribuição de exemplos em cada classe

MÉTODO PROPOSTO

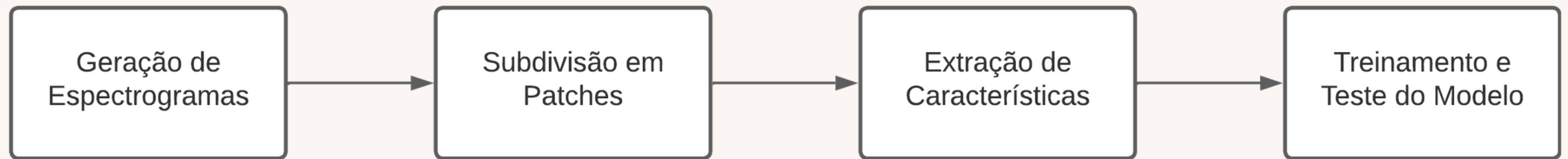
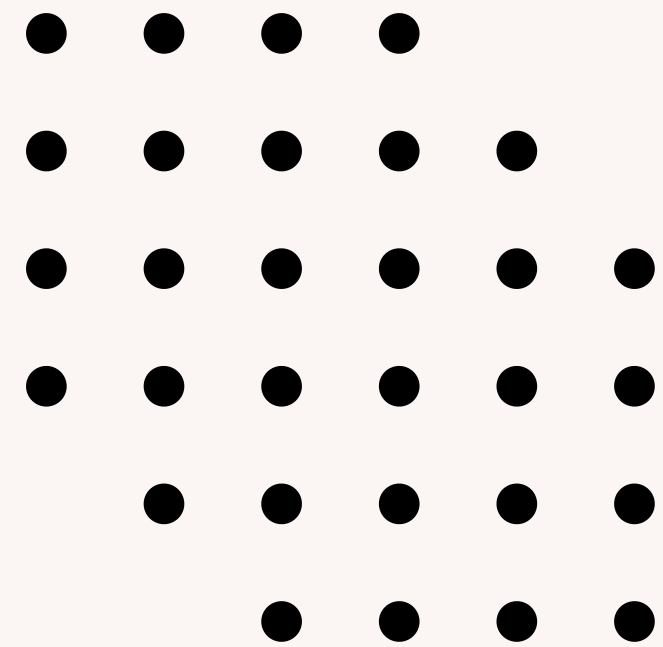


Figura 2 – Método Proposto

MÉTODO PROPOSTO

1. GERAÇÃO DE ESPECTROGRAMAS

- Cálculo da STFT
 - 2048 amostras
 - Sobreposição de 50% entre janelas subsequentes
- Conversão do espectrograma em escala mel (Figura 3)
 - Destacar as faixas de frequência que correspondem à sensibilidade auditiva humana
 - Reproduzir a percepção auditiva humana



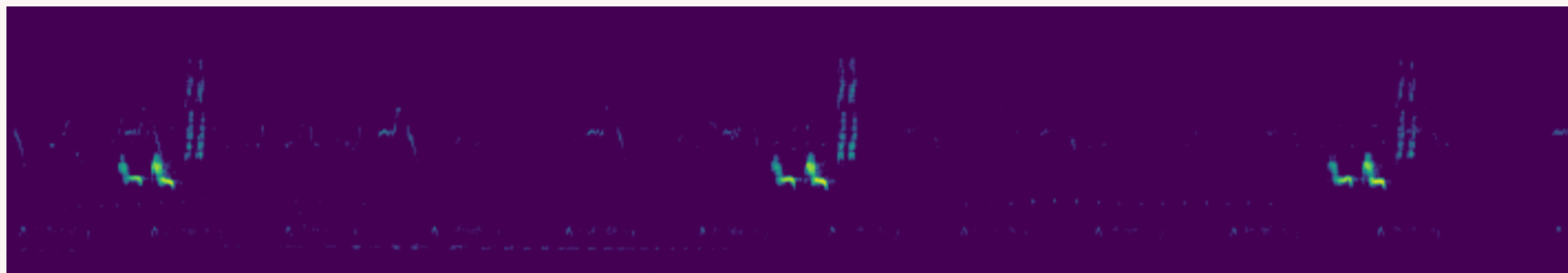
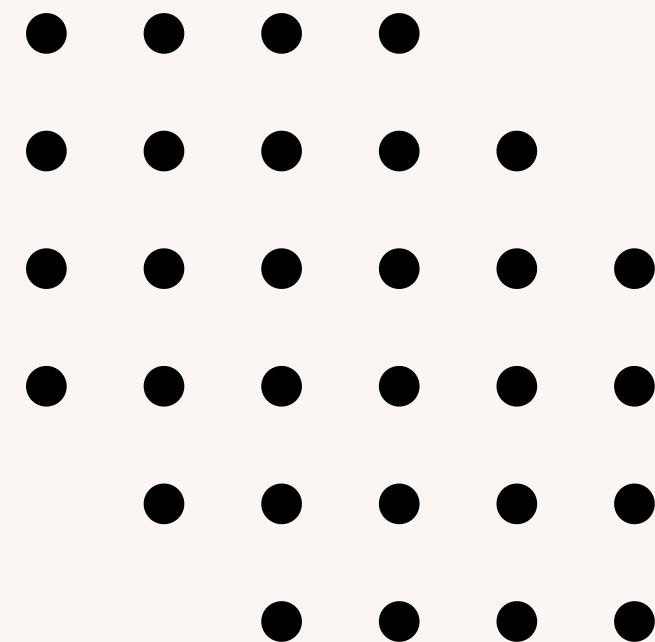


Figura 3 – Espectrograma da gravação XC128013.ogg

MÉTODO PROPOSTO

2. SUBDIVISÃO EM PATCHES

- Espectrogramas possuem tamanhos variados
 - Comprimento do áudio
 - Natureza não uniforme do áudio
 - Variedade de sons em vários momentos
- Padronização (Figura 4):
 - Segmentação em unidades menores, com N colunas
 - Padding no ultimo patch



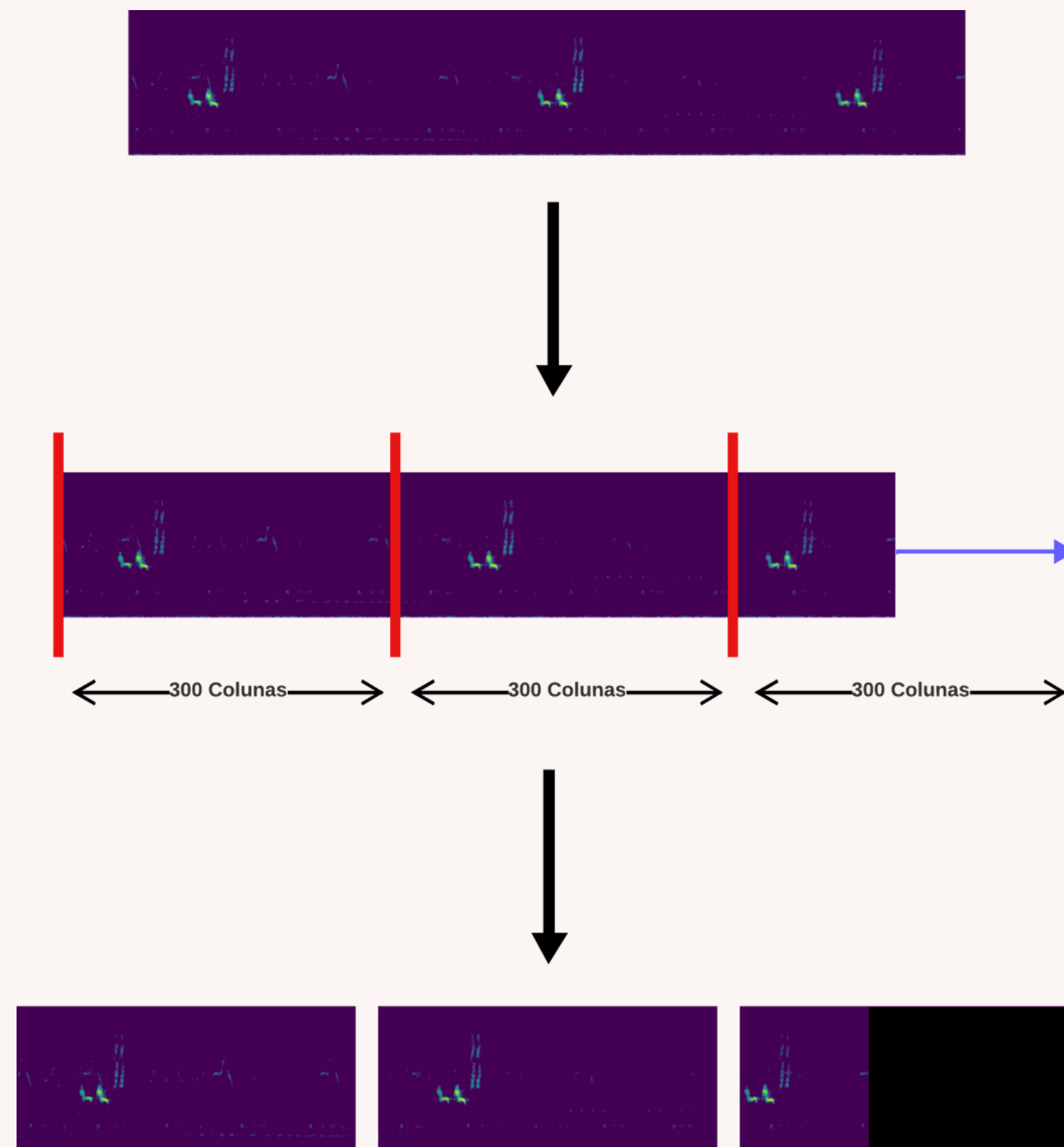
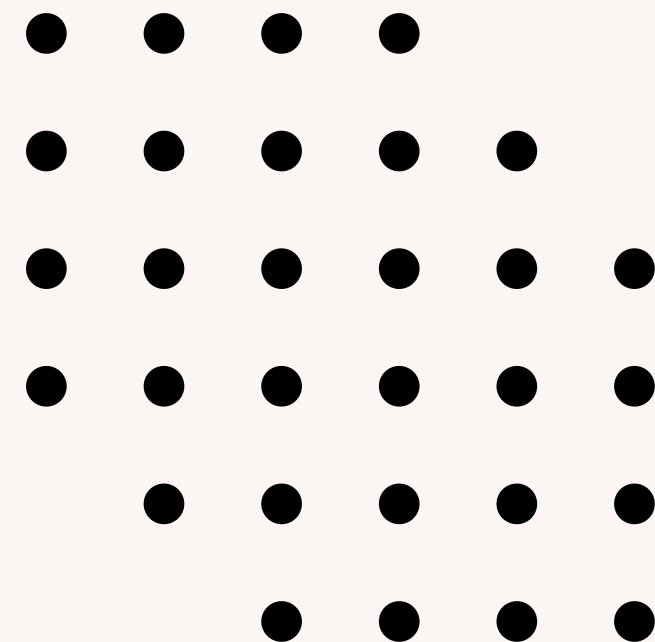


Figura 4 – Geração de patches com 300 colunas da gravação XC128013.ogg

MÉTODO PROPOSTO

3. EXTRAÇÃO DE CARACTERÍSTICAS

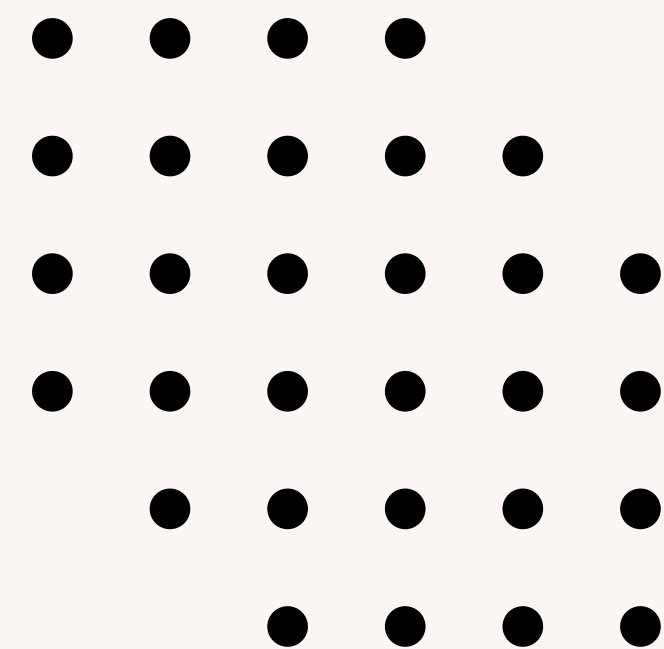
- Subconjunto de classes
 - Simular um ambiente real
- Descritores de Textura Handcrafted
 - Local Binary Patterns (LBP)
 - Filtros Gabor
- Descritores Transfer Learning
 - VGG16
 - Resnet50
 - DenseNet
 - MobileNet



MÉTODO PROPOSTO

4. TREINAMENTO E TESTE DO MODELO

- Separação dos dados em folds para cada split
 - Obter os nomes dos áudios utilizados para treino e teste
 - Obter os patches correspondentes
- Algoritmo de busca exaustiva para otimização de hiperparâmetros
- Treinamento do modelo
- Teste
- Sistema de votação em relação aos patches de dado arquivo de áudio
- Média dos resultados





RESULTADOS PRELIMINARES

- Espectrogramas em escala Mel
- Subdivisão em patches de 300 colunas
- Extração de características
 - LBP
 - Raio: 2
 - Pontos: 8
 - VGG16
- Treinamento e teste
 - Apenas um split
 - Subconjunto dos dados (2, 3, 5, e 10 classes)
 - KNN
 - Hiperparâmetro K de 1 a 20
 - Resultados na Tabela 1

| KNN | | | | |
|--------|-----------|-----------|-----------|------------|
| Método | 2 Classes | 3 Classes | 5 Classes | 10 Classes |
| LBP | 82% | 68% | 58% | 43% |
| VGG16 | 86% | 69% | 63% | 47% |

Tabela 1 – Pontuação F1 obtida com os descritores LBP e VGG16



RESULTADOS PRELIMINARES

- Influencia do número de classes
 - Pontuação F1 tende a diminuir ao aumentar o número de classes
 - Desbalanceamento das classes
 - Presença de ruídos semelhantes
- VGG com pequena vantagem
 - significância estatística dessa diferença apenas com validação cruzada em k folds
- Potencial para alcançar resultados satisfatórios

| 2024 | | | | | |
|--|-----|-----|-----|-----|-----|
| Atividades | MAR | ABR | MAI | JUN | JUL |
| Estudo sobre outros descritores de características | x | x | | | |
| Geração dos espectrogramas | x | | | | |
| Geração dos patches | x | | | | |
| Extração de Características com o LBP | x | | | | |
| Extração de Características com os filtros Gabor | x | | | | |
| Extração de Características com o VGG16 | | x | | | |
| Extração de Características com o RESNET50 | | x | | | |
| Extração de Características com o DenseNet | | x | | | |
| Extração de Características com o MobileNet | | x | | | |
| Treinamento e Teste do modelo usando SVM | x | x | x | x | |
| Escrita do TCC | x | x | x | x | x |
| Defesa do TCC | | | | | x |

Tabela 2 – Cronograma

REFERÊNCIAS

CARPENTER, F. L. A spectrum of nectar-eater communities. *American Zoologist*, Oxford University Press, v. 18, n. 4, p. 809–819, 1978. ISSN 00031569. Disponível em: <http://www.jstor.org/stable/3882538>.

HOLMES, R. T. Ecological and evolutionary impacts of bird predation on forest insects an overview. *Studies in Avian Biology*, v. 13, p. 6–13, 1990.

SNOW, D. W. Evolutionary aspects of fruit-eating by birds. *Ibis*, v. 113, p. 194–202, 1971.

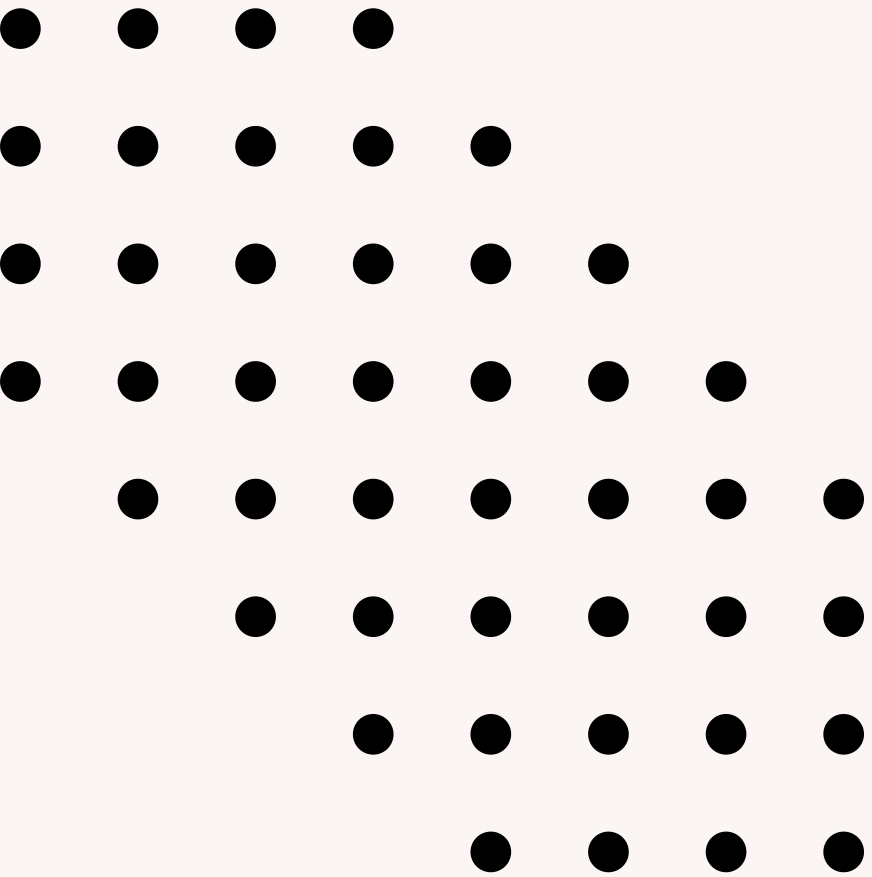
FIGUEIREDO, N. et al. A comparative study on filtering and classification of bird songs. In: *Sound and Music Computing Conference - SMC*. [S.l.]: SMC, 2018.

REFERÊNCIAS

LUCIO, D. R.; COSTA, Y. M. e Gomes da. Bird species classification using visual and acoustic features extracted from audio signal. In: 2016 35th International Conference of the Chilean Computer Science Society (SCCC). [S.l.: s.n.], 2016. p. 1–12.

INCZE, A. et al. Bird sound recognition using a convolutional neural network. In: 2018 IEEE 16th International Symposium on Intelligent Systems and Informatics (SISY). [S.l.: s.n.], 2018. p. 000295–000300.

BIRDCLEF. 2023. Cornell Lab of Ornithology. Disponível em: <https://www.kaggle.com/competitions/birdclef-2023/data>. Acesso em: 07 out. 2023.



Perguntas?

