



CLASSIFICAÇÃO DE PÁSSAROS POR MEIO DE CANTOS E CHAMADOS: UMA ABORDAGEM INTEGRADA COM ESPECTROGRAMAS EM CONJUNTO COM DESCRITORES HANDCRAFTED E TRANSFER LEARNING

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

12/08/2024



INTRODUÇÃO

- Biodiversidade, sustentabilidade, equilíbrio dos ecossistemas
- Importância dos pássaros
 - 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
- Quantidade de especialistas reduzida
- Tempo de treinamento desses profissionais
- Demanda crescente por métodos mais eficientes



TRABALHOS ANTERIORES

- Classificação automática de espécies com base em características sonoras
 - Banco de dados diverso
 - Pré-processamento
 - Geração de espectrogramas
 - Descritores de características
 - Handcrafted (NANNI; GHIDONI; BRAHNAM, 2017)
 - TransferLearning (PAN; YANG, 2010)
 - Classificadores

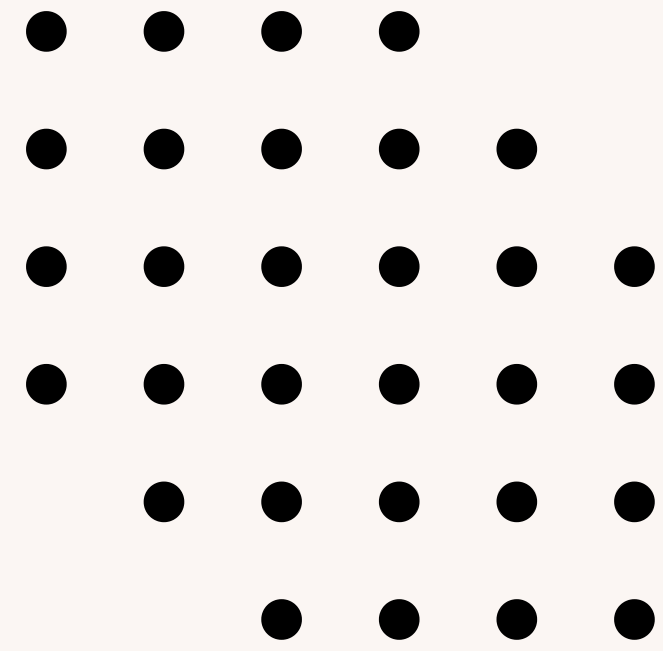


OBJETIVO

O objetivo principal deste trabalho foi desenvolver um modelo de aprendizado de máquina que seja capaz de classificar diversas espécies de pássaros com base em seus cantos e chamados.

DATASET

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



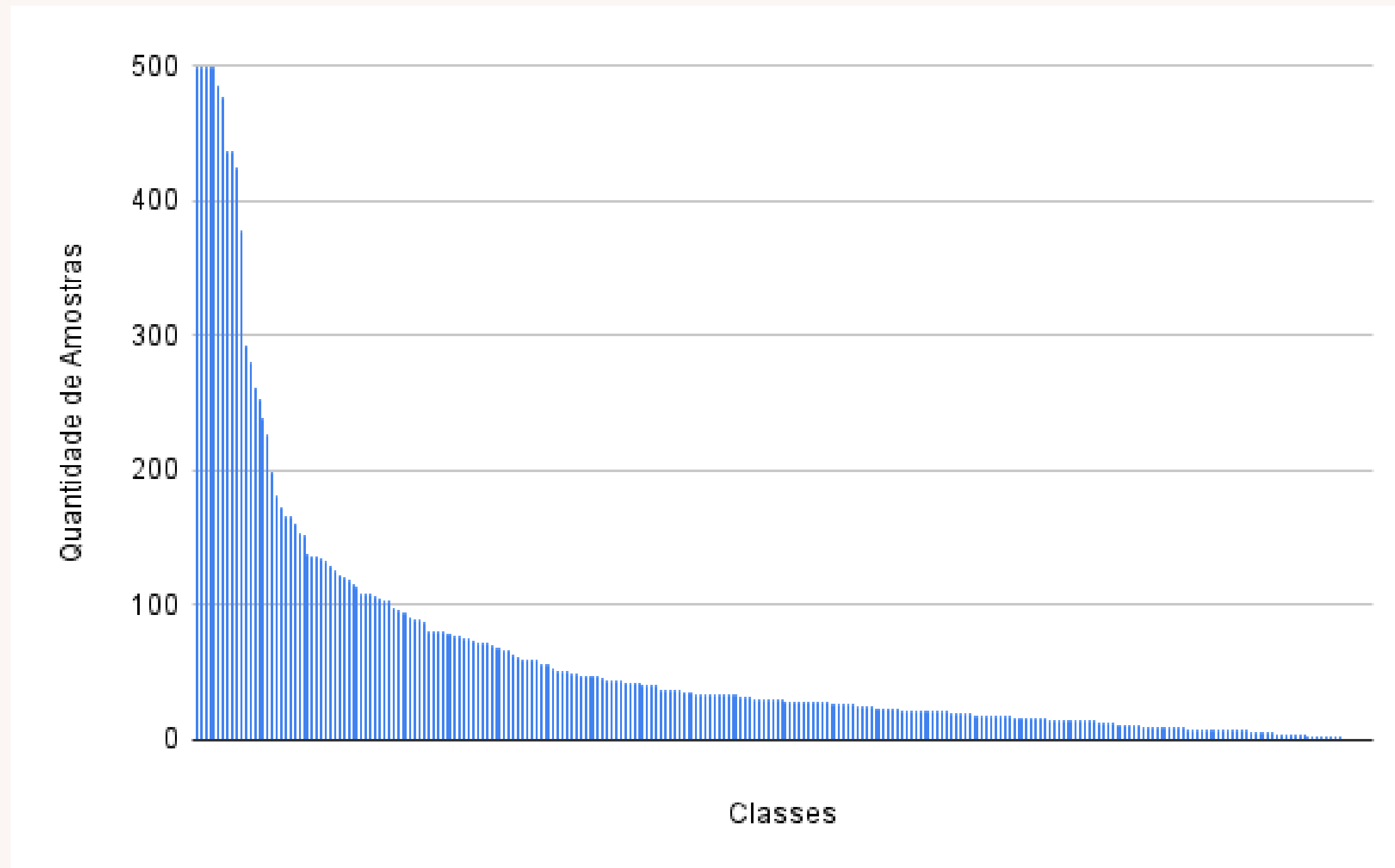


Figura 1 – Distribuição de exemplos em cada classe

MÉTODO PROPOSTO

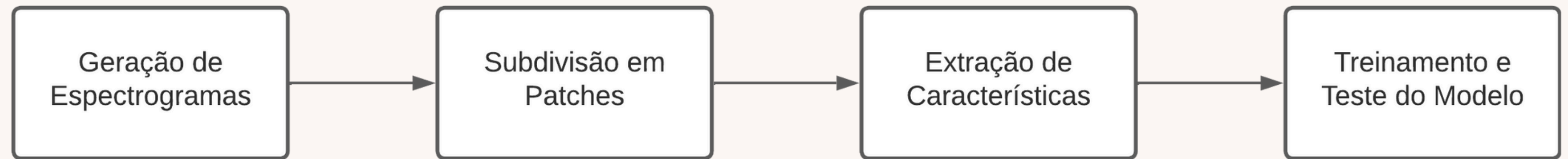
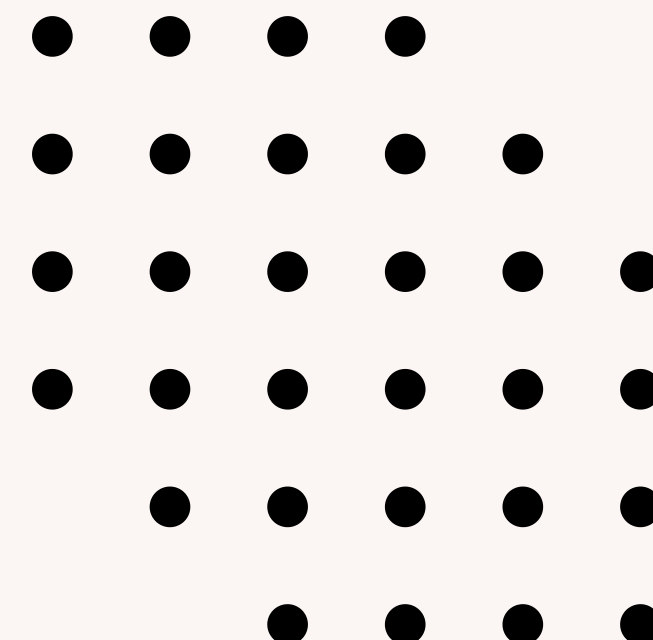


Figura 2 – Método Proposto

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



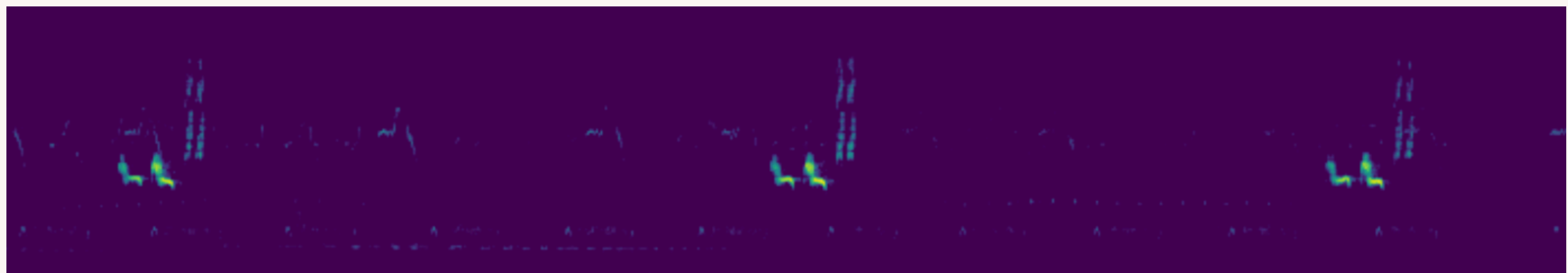
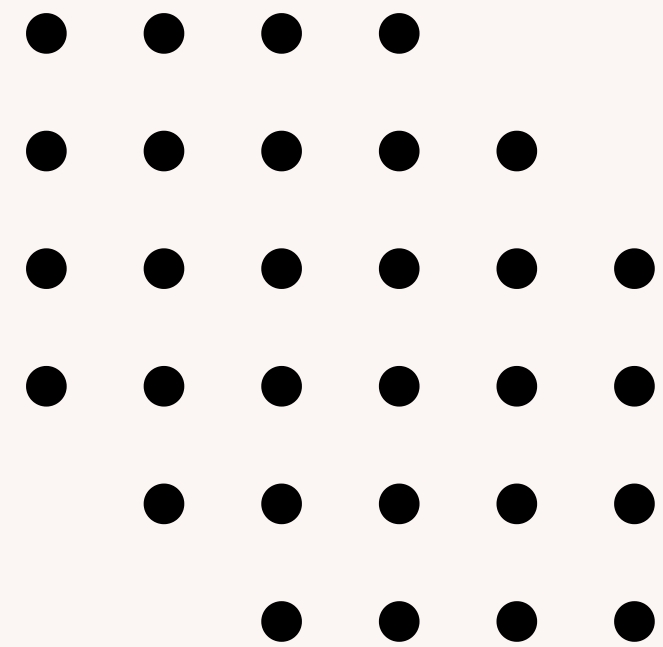


Figura 3 – Espectrograma da gravação XC128013.ogg

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):
 - Divisão em N unidades menores



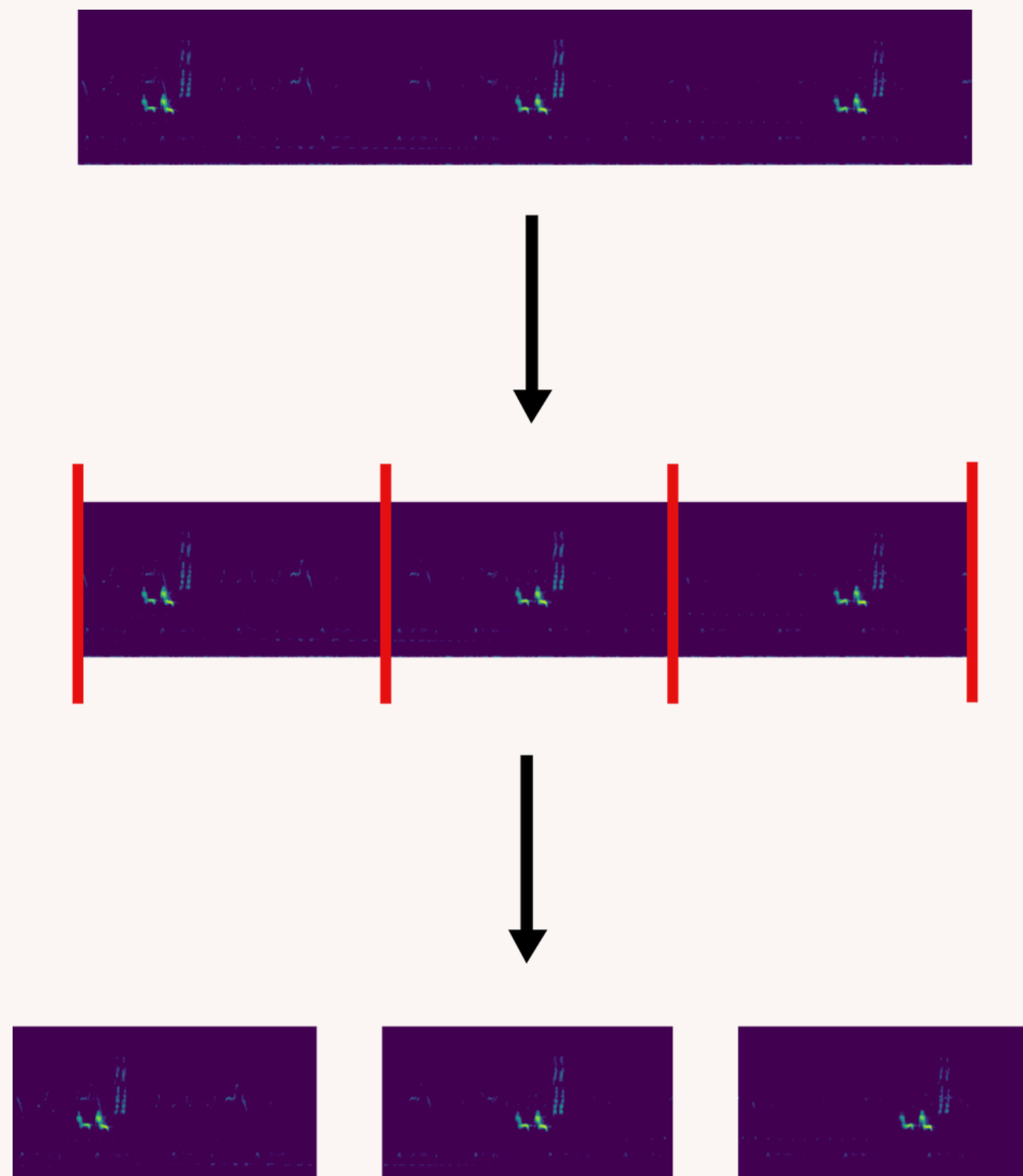
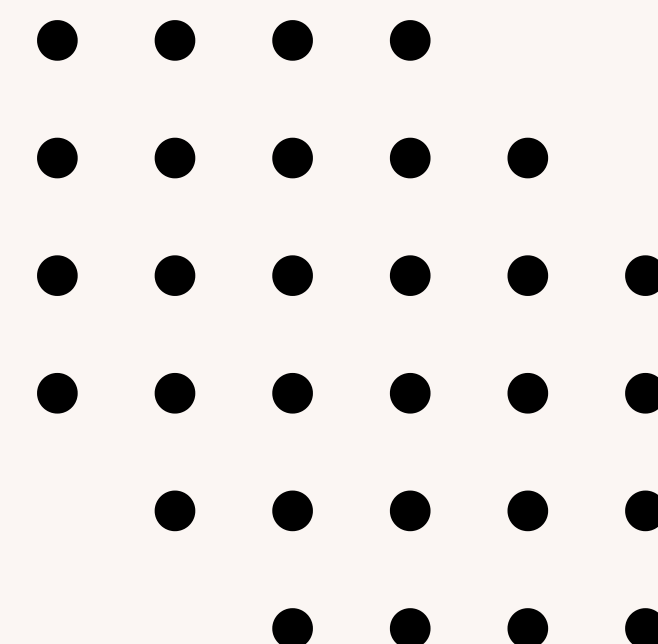


Figura 4 – Geração de 3 patches da gravação XC128013.ogg

MÉTODO PROPOSTO

3. EXTRAÇÃO DE CARACTERÍSTICAS

- Subdivisão da base de dados
 - Escolha do número de classes analisadas
 - Uso do Kernel Density Estimation - KDE
 - Filtro das classes com mais de 100 amostras
 - Aplicação do KDE sobre as classes (Figura 5 e 6)
 - Obtenção de um ponto de referência para cada classe
 - Aplicação do KDE sobre esses pontos de referência
 - Obtenção de um ponto de análise (Figura 7)
 - Cálculo da distância entre o ponto de análise e as classes
 - Ordenação crescente dessas distâncias
 - K espécies mais próximas



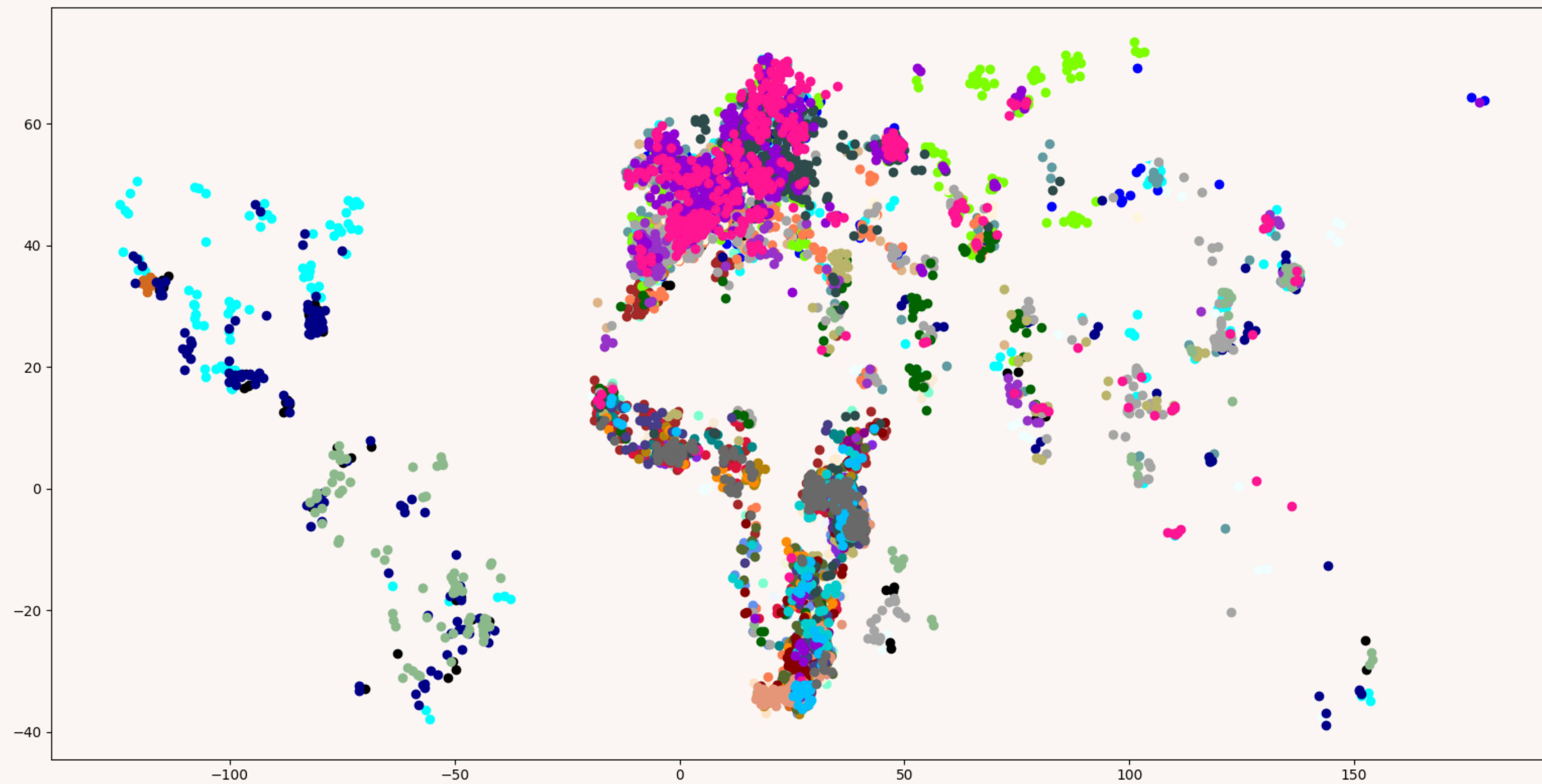


Figura 5 – Distribuição das amostras

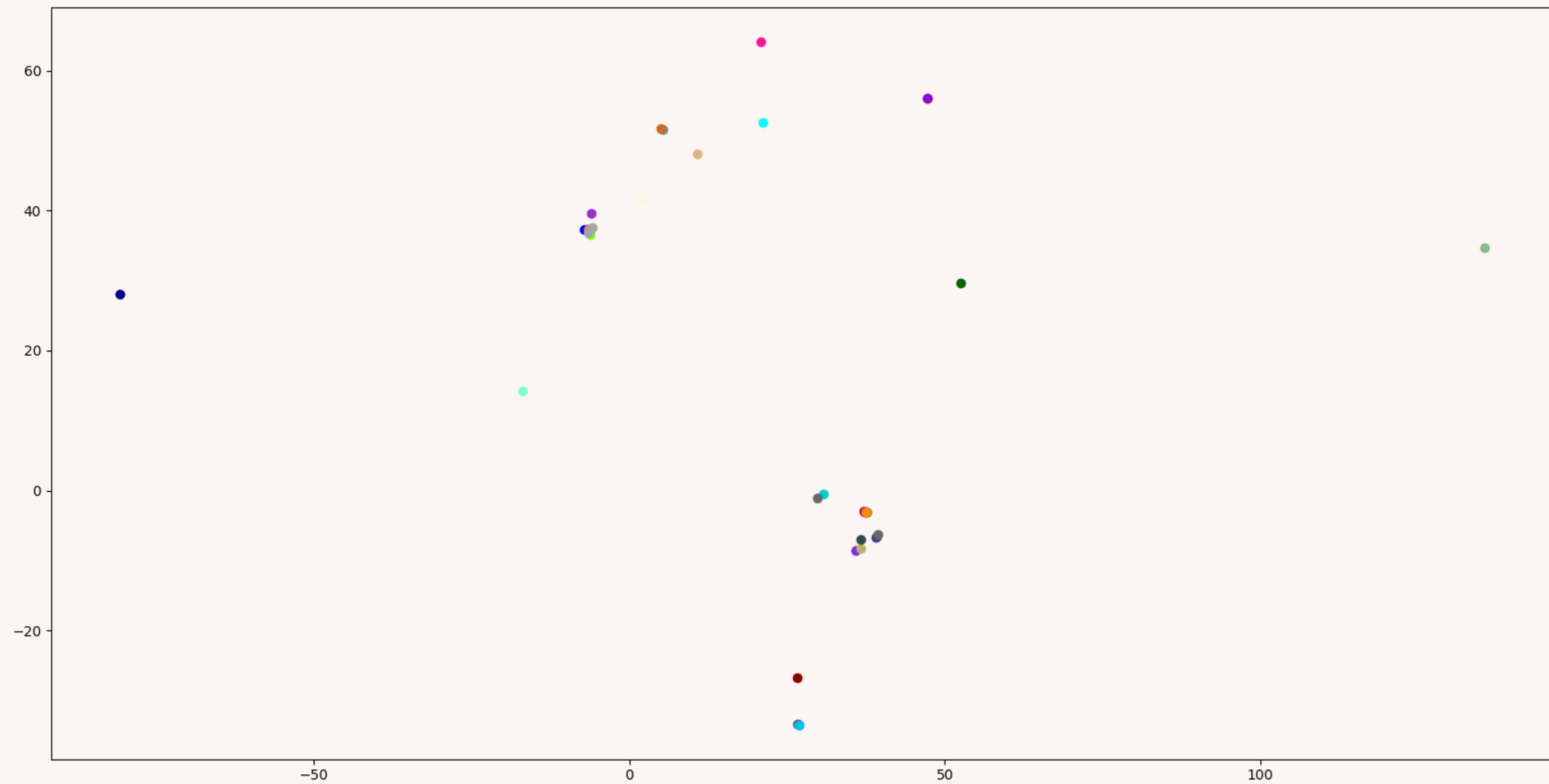


Figura 6 – Pontos de Referência de cada Espécie

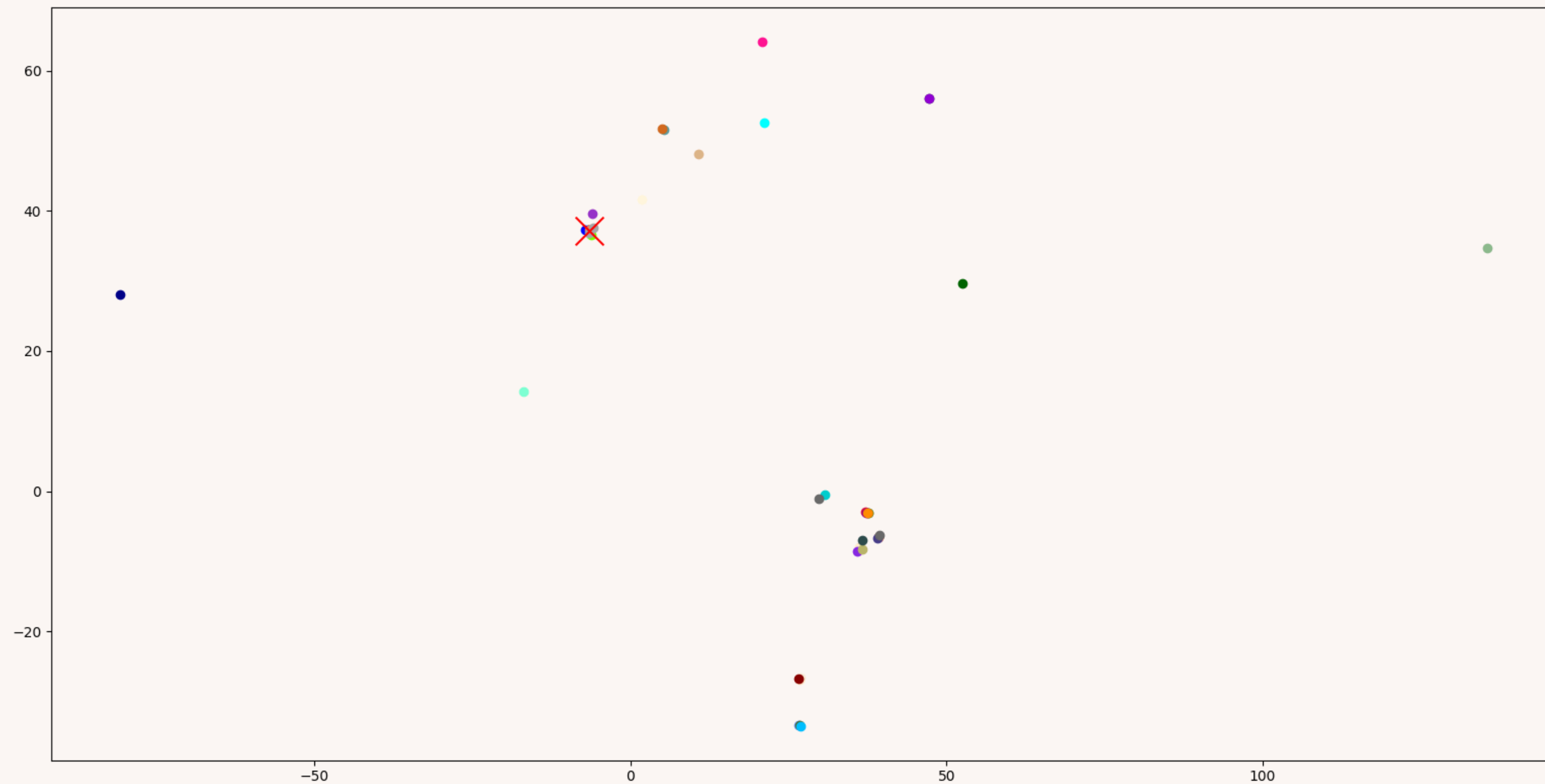


Figura 7 – Ponto de Referência em Análise

Nome Científico	2 Classes	3 Classes	5 Classes	10 Classes
Hirundo rustica	-	-	X	X
Milvus migrans	X	X	X	X
Bubulcus ibis	-	-	X	X
Delichon urbicum	-	-	-	X
Motacilla flava	-	-	-	X
Merops apiaster	-	X	X	X
Spatula querquedula	-	-	-	X
Upupa epops	-	-	-	X
Egretta garzetta	X	X	-	X
Cecropis daurica	-	-	-	X

Tabela 1 – Distribuição das dez Espécies mais próximas em cada subconjunto de classe

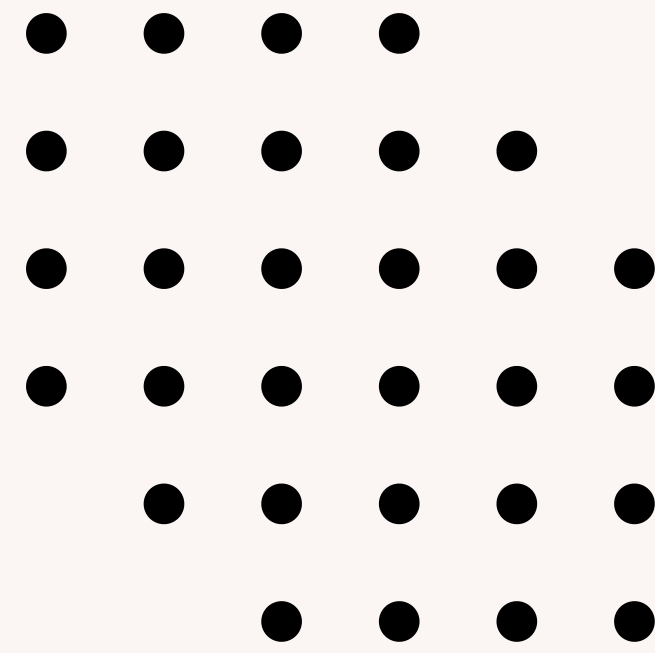
Nome Cientifico	Nome Comum	Localização	Nº amostras
Hirundo rustica	Barn Swallow	Sul da Espanha	500
Milvus migrans	Black Kite	Sul da Espanha	262
Bubulcus ibis	Cattle Egret	Sul da Espanha	166
Delichon urbicum	Common House-Martin	Sul da Espanha	425
Motacilla flava	Western Yellow Wagtail	Sul da Espanha	500
Merops apiaster	European Bee-eater	Sul da Espanha	437
Spatula querquedula	Garganey	Nordeste da Espanha	136
Upupa epops	Eurasian Hoopoe	Sul da Espanha	436
Egretta garzetta	Little Egret	Sul da Espanha	378
Cecropis daurica	Red-rumped Swallow	Centro-oeste da Espanha	227

Tabela 2 – Detalhes das dez Classes mais próximas

MÉTODO PROPOSTO

3. EXTRAÇÃO DE CARACTERÍSTICAS

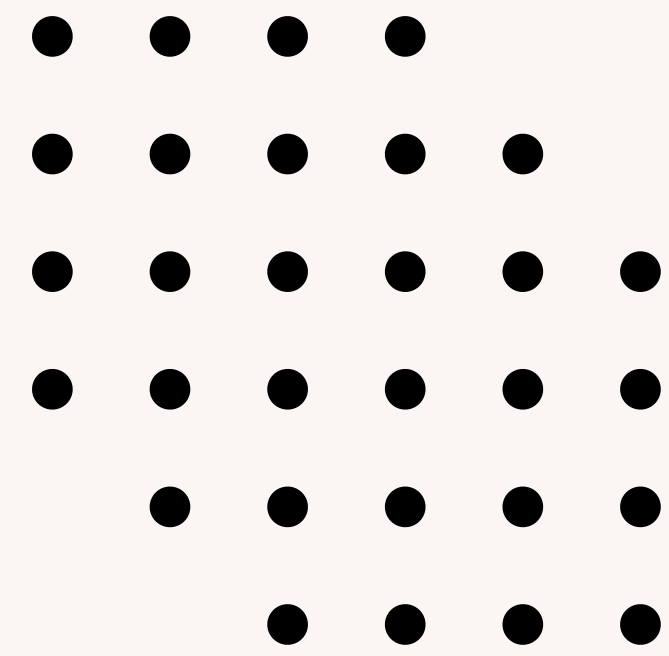
- Processamento dos patches
 - Descritores de Textura Handcrafted
 - Local Binary Patterns (LBP)
 - Filtros Gabor
 - Descritores Transfer Learning
 - VGG16
 - Resnet50
 - DenseNet
 - MobileNet



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
 - Encontrar os patches correspondentes
- Normalização dos dados
- Algoritmo de busca exaustiva para otimização de hiperparâmetros
- Treinamento do modelo
- Teste
- Sistema de votação entre patches de um mesmo áudio
- Média dos resultados





RESULTADOS E DISCUSSÃO

- Espectrogramas em escala Mel
- Subdivisão em um, três e cinco patches
- Extração de características
 - Subconjunto dos dados (2, 3, 5, e 10 classes)
 - LBP, Filtros Gabor, VGG16, Resnet50, DenseNet121 e MobileNet
- Treinamento e teste
 - Validação Cruzada
 - SVM

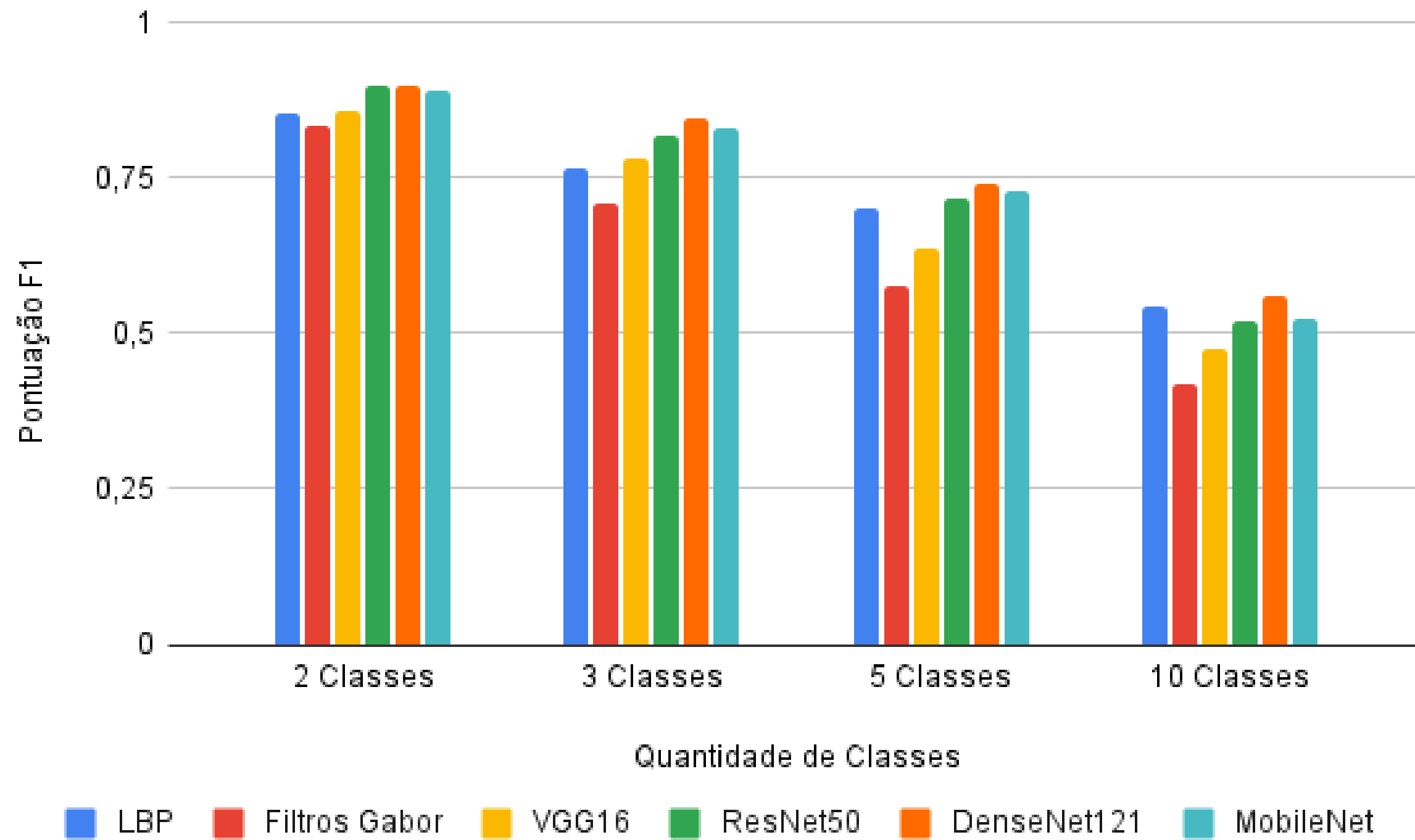


Figura 10 – Pontuação F1 para um patch

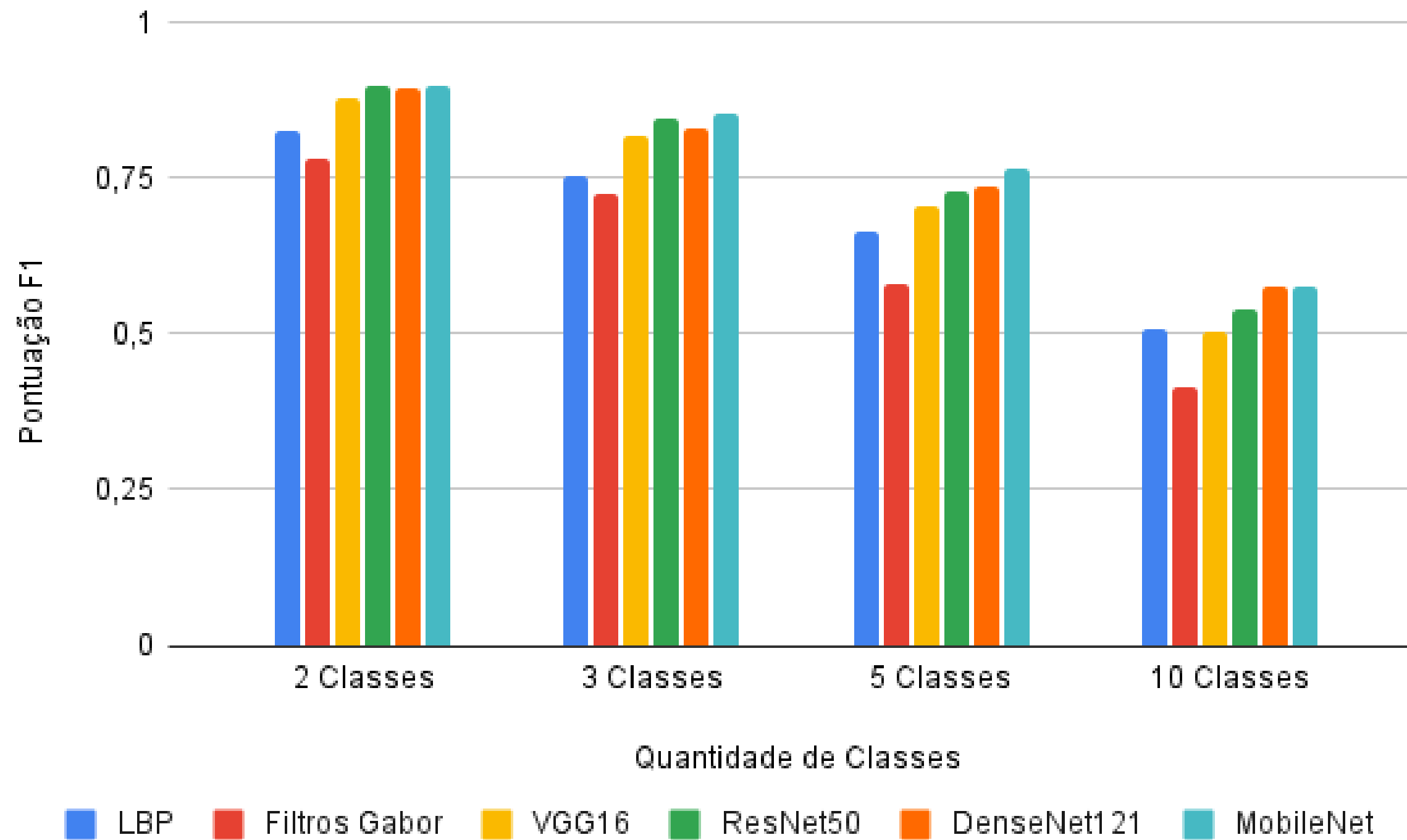


Figura 11 – Pontuação F1 para três patches

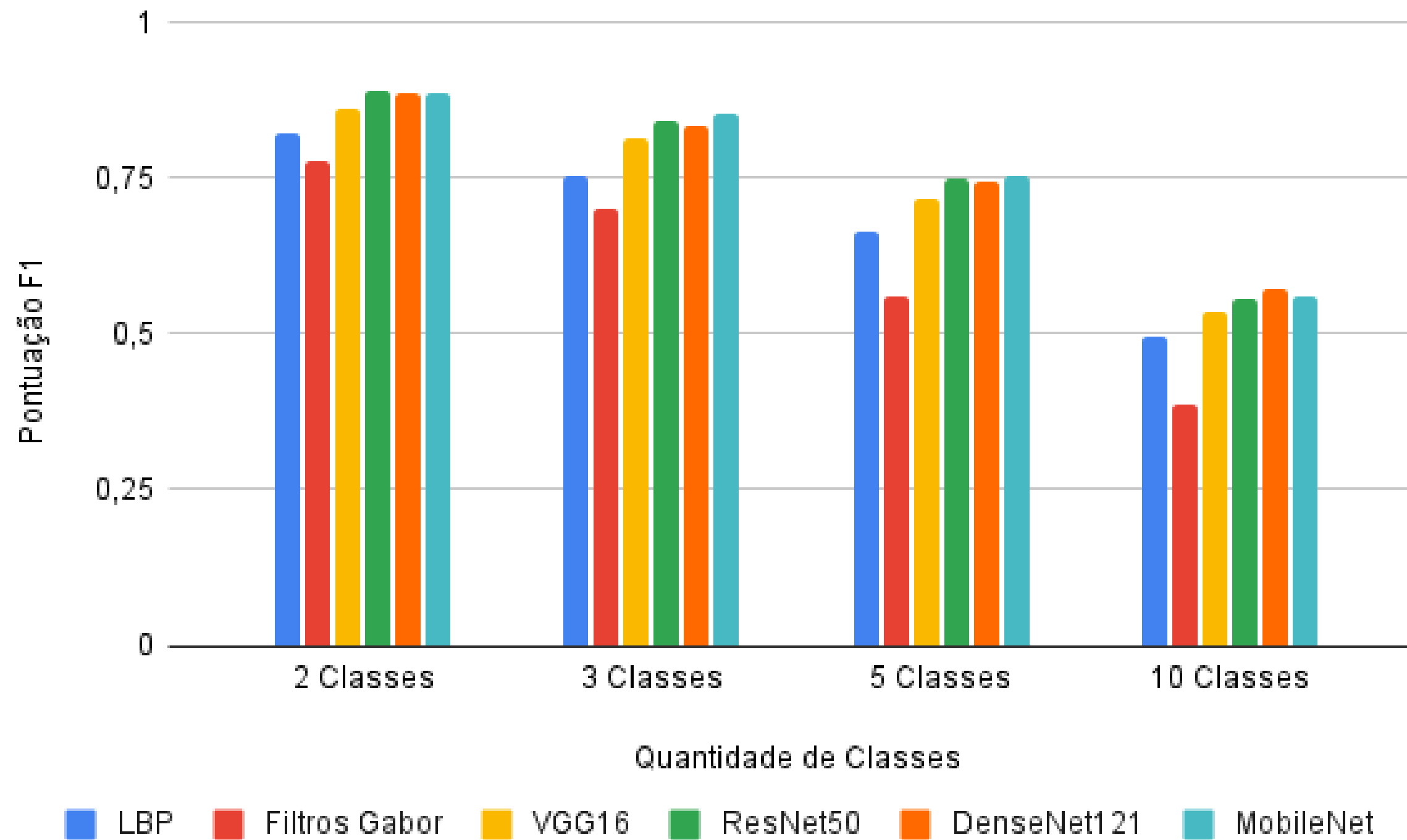
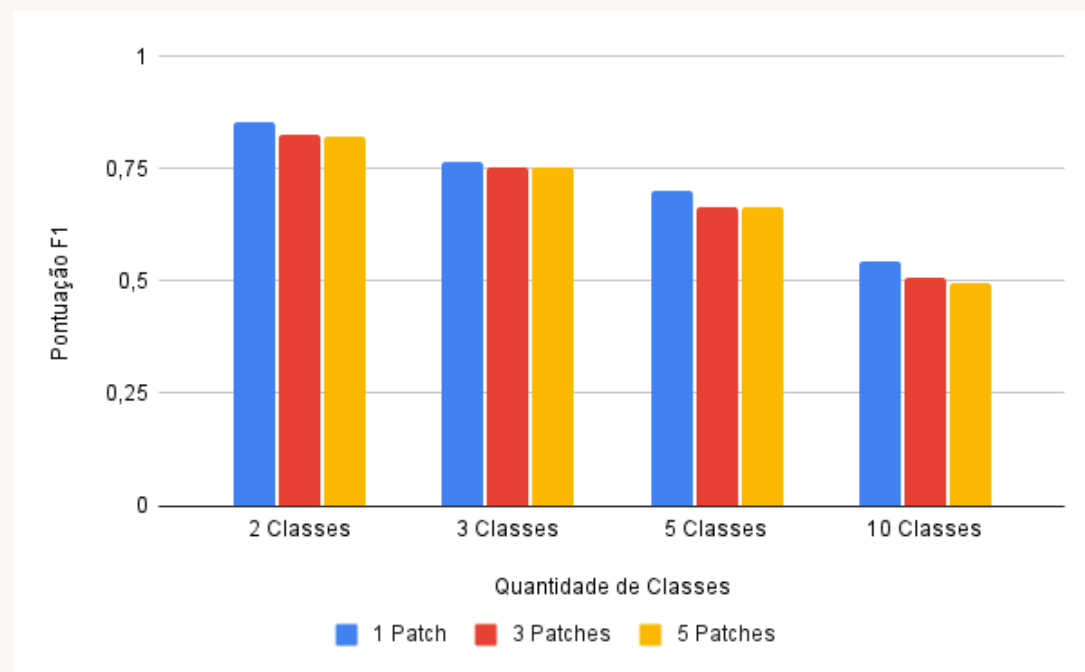


Figura 12 – Pontuação F1 para cinco patches

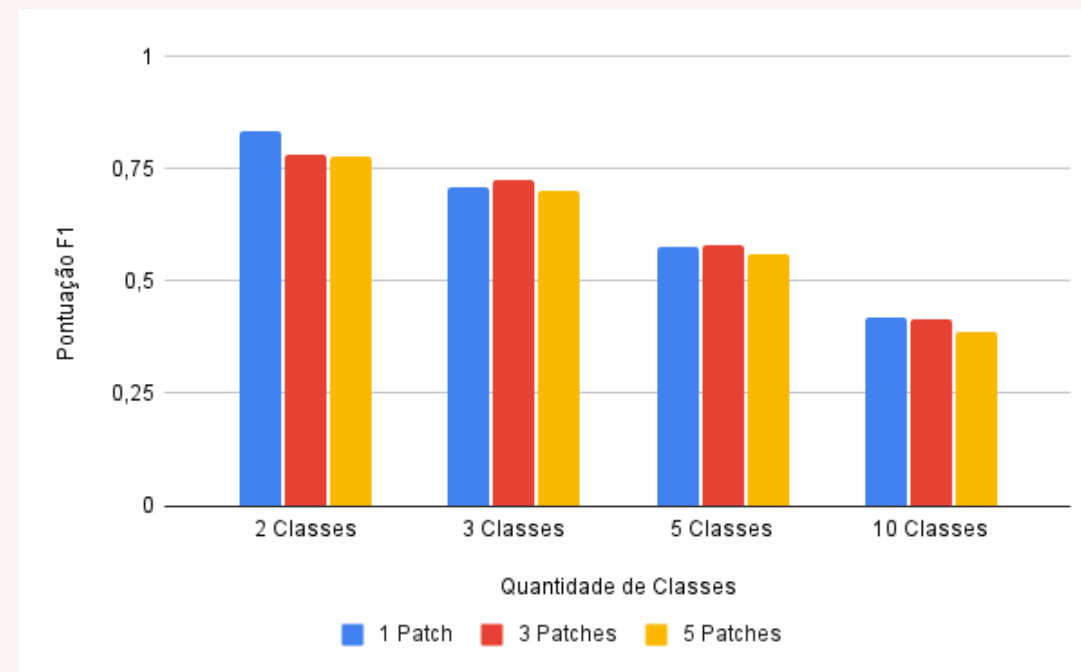


RESULTADOS E DISCUSSÃO

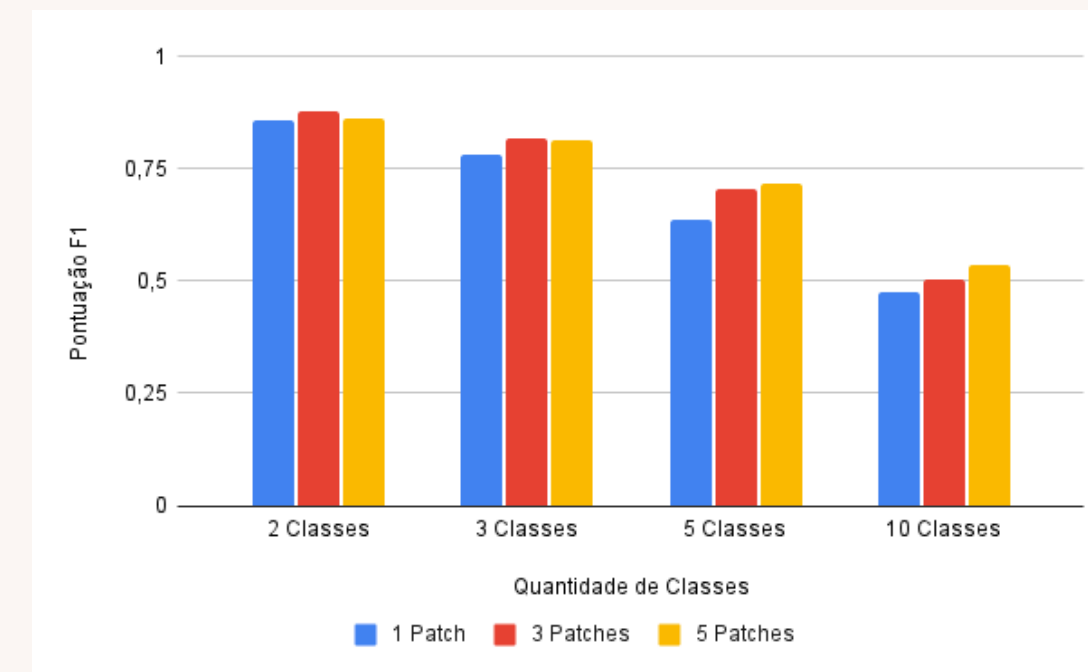
- Análise dos Patches
 - 1 Patches
 - Destaque para o DenseNet121 (0.8992), Resnet50 (0.8976) e MobileNet (0.8888)
 - Pior Resultado: Filtros Gabor (0.8342)
 - 3 Patches
 - Ponto de equilíbrio
 - Aumento significativo da pontuação F1 em relação a 1 patch, para todos os descritores
 - Destaque para Resnet50 (0.8994) e MobileNet (0.8989)
 - 5 Patches
 - Ponto de Saturação
 - Declínio da pontuação F1 em relação a 3 patches, para todos os descritores



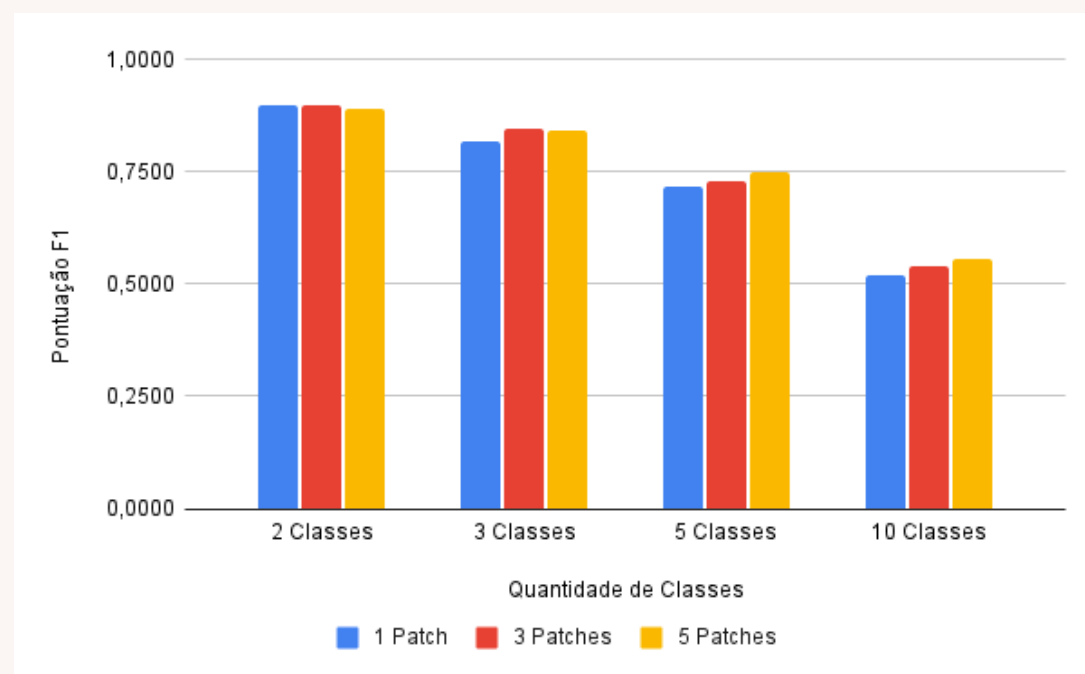
(a) LBP



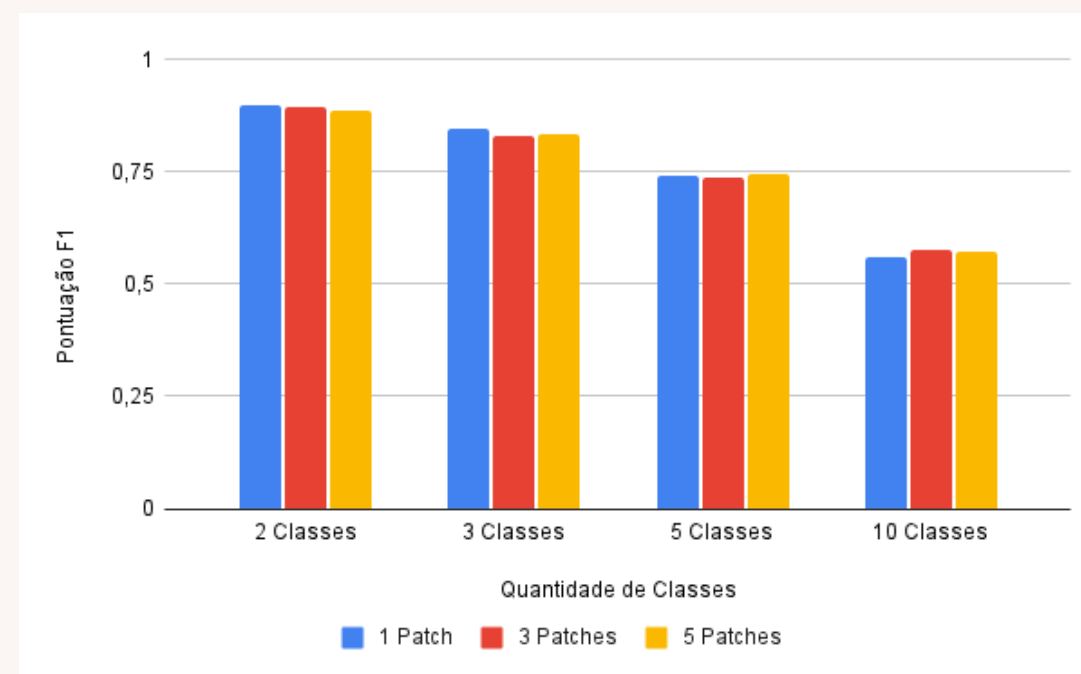
(b) Filtros Gabor



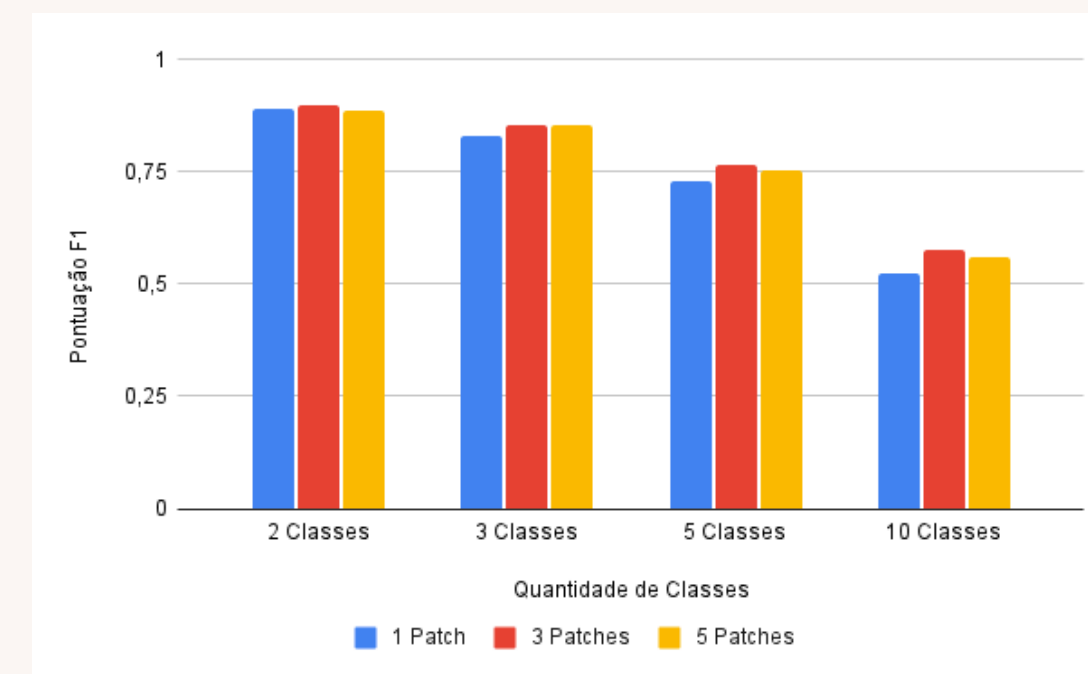
(c) VGG16



(d) ResNet50



(e) DenseNet121



(f) MobileNet

Figura 13 – Pontuação F1 em relação a quantidade de classes para cada descritor



RESULTADOS E DISCUSSÃO

- Análise em relação ao número de classes
 - Aumento do número de classes: diminuição da pontuação F1
 - Variação menos acentuada com DenseNet, Resnet50, MobileNet e LBP
 - Variação mais acentuada com Filtros Gabor



CONCLUSÃO

- Número de patches
 - Influência significativa no desempenho dos classificadores
 - O ponto de equilíbrio foi atingido com 3 patches
 - O ponto de saturação foi atingido com 5 patches
- Número de classes
 - Aumento da dificuldade do problema ao aumentar o numero de classes
- Handcrafted vs Transfer Learning
 - Vantagem do uso de descritores Transfer Learning para esse problema



TRABALHOS FUTUROS

- Pré-processamento
 - Áudios: Redução de ruídos
 - Espectrogramas: Filtros de realce
- Combinação de descritores
- Comparação do uso de variáveis de geolocalização no vetor de características

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.

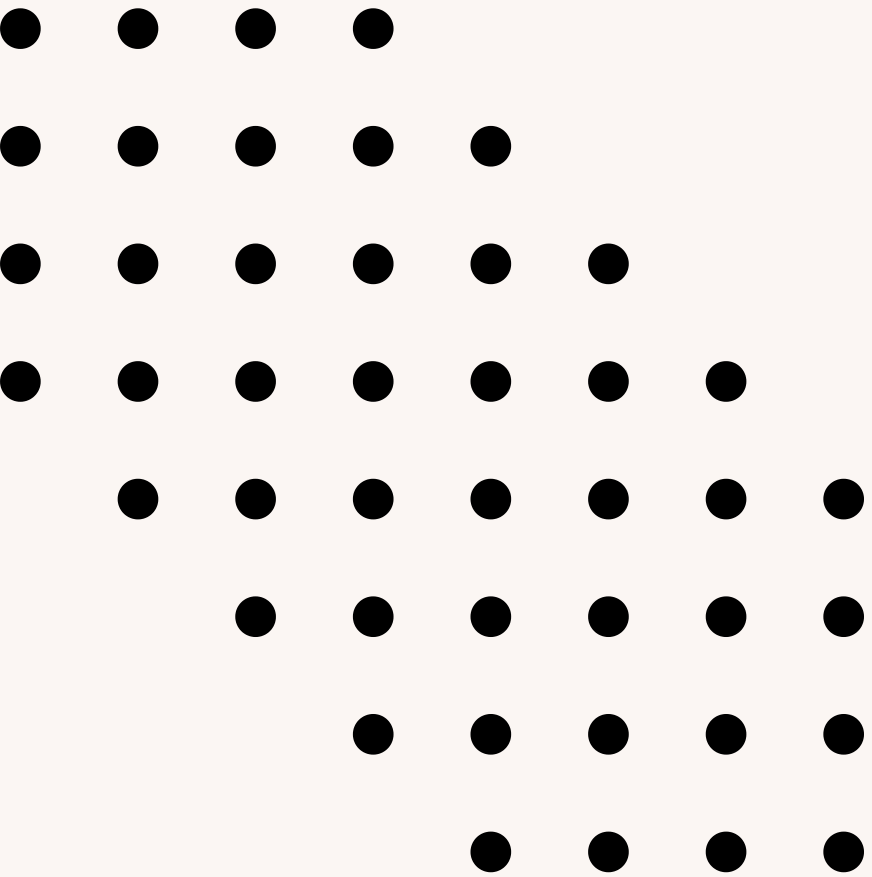
FIGUEIREDO, N. et al. A comparative study on filtering and classification of bird songs. In: Sound and Music Computing Conference – SMC. São Paulo, SP: SMC, 2018.

REFERÊNCIAS

NANNI, L.; GHIDONI, S.; BRAHNAM, S. Handcrafted vs. non-handcrafted features for computer vision classification. Pattern Recognition, v. 71, p. 158–172, 2017. ISSN 0031-3203. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0031320317302224>.

LIANG, H.; FU, W.; YI, F. A survey of recent advances in transfer learning. In: 2019 IEEE 19th International Conference on Communication Technology (ICCT). Xi'an, China: IEEE Computer Society Press., 2019. p. 1516–1523

KLINCK, H. et al. BirdCLEF 2023. Kaggle, 2023. Disponível em: <https://kaggle.com/competitions/birdclef-2023>.



Perguntas?

