

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

FELIPE ARCHANJO DA CUNHA MENDES

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

CAMPO MOURÃO

2024

FELIPE ARCHANJO DA CUNHA MENDES

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

**BIRD CLASSIFICATION THROUGH SONGS AND CALLS: AN INTEGRATED
APPROACH WITH SPECTROGRAMS IN CONJUNCTION WITH HANDCRAFTED
DESCRIPTORS AND TRANSFER LEARNING**

Trabalho de conclusão de curso de graduação
apresentado como requisito parcial para a obtenção
do título de Bacharel em Ciência da Computação,
Universidade Tecnológica Federal do Paraná
(UTFPR).

Orientador(a): Juliano Henrique Foleis.

CAMPO MOURÃO

2024

FELIPE ARCHANJO DA CUNHA MENDES

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

Trabalho de Conclusão de Curso de Graduação
apresentado como requisito para obtenção do título
de Bacharel em Ciência da Computação,
Universidade Tecnológica Federal do Paraná
(UTFPR).

Data de aprovação: 12/agosto/2024

Juliano Henrique Foleis
Doutorado
Universidade Tecnológica Federal do Paraná

Rodrigo Campiolo
Doutorado
Universidade Tecnológica Federal do Paraná

Rodrigo Hubner
Doutorado
Universidade Tecnológica Federal do Paraná

CAMPO MOURÃO

2024

RESUMO

MENDES, Felipe Archanjo da Cunha. **Classificação de Pássaros por Meio de Cantos e Chamados:** Uma Abordagem Integrada com Espectrogramas em Conjunto com Descritores Handcrafted e Transfer Learning. 2024. Trabalho de conclusão de curso (Bacharelado em Ciência da Computação) – Universidade Tecnológica Federal do Paraná, Campo Mourão, 12/08/2024.

A classificação de espécies de pássaros é uma tarefa importante no monitoramento e conservação do meio ambiente. Neste estudo, o objetivo foi desenvolver um modelo de aprendizado de máquina destinado a classificar distintas espécies de pássaros com base em seus cantos e chamados. Inicialmente, espectrogramas foram gerados para cada gravação disponível no banco de dados. Posteriormente, adotou-se uma estratégia de divisão desses espectrogramas em segmentos menores, denominados *patches*, para investigar a influência dessa abordagem no desempenho do modelo. Diversas técnicas de extração de características foram exploradas, incluindo descritores *handcrafted*, como o *LBP* e Filtros Gabor, e descritores derivados de modelos de *Transfer Learning*, tais como *VGG16*, *ResNet50*, *DenseNet121* e *MobileNet*. Esse processo foi realizado utilizando conjuntos de duas, três, cinco e dez classes, com o objetivo de analisar o comportamento do resultado conforme o número de classes aumenta. Além disso, foram conduzidos experimentos utilizando o classificador *SVM* para avaliar a precisão e eficácia do modelo. Os resultados revelaram que os descritores baseados em *Transfer Learning* apresentaram desempenho superior, com destaque para o *Resnet50*, que obteve uma pontuação F1 de 0,8994 na classificação entre duas espécies de pássaros, com o espectrograma dividido em três *patches*. Outros modelos, como o *DenseNet121* e o *MobileNet*, também mostraram resultados competitivos na mesma configuração com dois pássaros: o *DenseNet121* atingiu uma pontuação F1 de 0,8992 utilizando apenas um *patch*, enquanto o *MobileNet* registrou uma pontuação de 0,8989 com três *patches*. Esses achados sugerem que a abordagem de divisão dos espectrogramas em *patches*, combinada com técnicas extração de características baseadas em *transfer learning*, podem ser eficazes para a classificação de espécies de pássaros baseada em seus cantos e chamados. Os resultados também mostraram que conforme o número de espécies aumenta, as taxas de classificação caem rapidamente.

Palavras-chave: classificação de espécies de pássaros; espectrogramas; aprendizagem de máquina; reconhecimento de padrões; descritores de características.

() Não autorizo a disponibilização de endereço de correio eletrônico para contato.

(X) Autorizo a disponibilização do seguinte correio eletrônico para contato:

felipecunhamendes@gmail.com

ABSTRACT

MENDES, Felipe Archanjo da Cunha. **Bird Classification Through Songs and Calls: An Integrated Approach with Spectrograms in Conjunction with Handcrafted Descriptors and Transfer Learning.** 2024. Trabalho de conclusão de curso (Bacharelado em Ciência da Computação) – Universidade Tecnológica Federal do Paraná, Campo Mourão, 12/08/2024. Título original: **Classificação de Pássaros por Meio de Cantos e Chamados: Uma Abordagem Integrada com Espectrogramas em Conjunto com Descritores Handcrafted e Transfer Learning.**

The classification of bird species is an important task in environmental monitoring and conservation. This study aimed to develop a machine-learning model to classify bird species based on their songs and calls. Initially, spectrograms were generated for each recording available in the database. Subsequently, a strategy of dividing these spectrograms into smaller segments, called patches, was adopted to investigate the influence of this approach on the model's performance. Various feature extraction techniques were explored, including handcrafted descriptors such as LBP and Gabor Filters, and descriptors derived from Transfer Learning models, such as *VGG16*, *ResNet50*, *DenseNet121*, and *MobileNet*. This process used sets with two, three, five, and ten classes to analyze the outcome behavior as this subset gradually increased. Additionally, experiments using the SVM classifier were conducted to evaluate the accuracy and effectiveness of the model. The results revealed that the descriptors based on Transfer Learning showed superior performance, with particular emphasis on *ResNet50*, which achieved an F1 score of 0.8994 in classifying two birds with the spectrogram divided into three patches. Other models, such as *DenseNet121* and *MobileNet*, also showed competitive results in the same configuration with two birds: *DenseNet121* achieved an F1 score of 0.8992 using just one patch. In contrast, *MobileNet* scored 0.8989 with three patches. These findings suggest that dividing the spectrograms into patches, combined with advanced feature extraction techniques, can effectively classify bird species based on their songs and calls. The results also show that classification performance drops sharply as the number of species increases.

Keywords: bird species classification; spectrograms; machine learning; pattern recognition; feature descriptors