

# Systematic Review on Machine Learning Applied to Bioacoustics using the PRISMA Method

## Revisão Sistemática sobre Machine Learning Aplicada a Bioacústica utilizando o Método PRISMA

Luiz E. R. Martins<sup>1</sup>, Virginia A. dos Santos<sup>1</sup>,  
Allan G. de Oliveira<sup>1</sup>, Thiago M. Ventura<sup>1</sup>, Nielsen Cassiano Simões<sup>1</sup>

<sup>1</sup>Instituto de Computação – Universidade Federal do Mato Grosso (UFMT)  
Cuiabá – MT – Brasil

{luiz.martins, virginia.santos}@sou.ufmt.br

{allan, thiago, nielsen}@ic.ufmt.br

**Abstract.** *Using the most appropriate technique is crucial in any Machine Learning process. In bioacoustics, due to the complexity of the data, many techniques have been applied and developed. In this context, this paper presents the systematic review carried out using the PRISMA methodology for Bioacoustics in environmental monitoring through bird vocalization. The review demonstrates that the Spiking Neural Network, Convolutional Neural Network and Residual Neural Network techniques have gained prominence in recent years.*

**Keywords:** *bioacoustic. neural networks. machine learning. classification. bird monitoring*

**Resumo.** *O uso da técnica mais adequada é crucial em qualquer processo Aprendizagem de Máquina. Na bioacústica, devido à complexidade dos dados muitas técnicas tem sido aplicadas e desenvolvidas. Nesse contexto, esse trabalho apresenta a revisão sistemática realizada utilizando a metodologia PRISMA para Bioacústica no monitoramento ambiental por meio de vocalização de pássaros. A revisão demonstrou que as técnicas de Spiking Neural Network, Convolutional Neural Network e Residual Neural Network ganharam destaque nos últimos anos.*

**Palavras-chave:** *bioacústica. redes neurais. aprendizado de máquina. classificação. monitoramento de pássaros*

### 1. Introdução

A preservação ambiental emergiu como uma preocupação global premente, uma vez que as transformações no meio ambiente possuem repercussões significativas na qualidade de vida das pessoas. Com crescimento da população mundial junto a expansão das áreas urbanas, surge a necessidade da extração de recursos naturais, adjunto das mudanças climáticas e eventos como incêndios florestais que têm alterado as condições naturais. Em resposta a esses desafios, urge uma série de abordagens para mitigar esses impactos.

Uma abordagem que auxilia no monitoramento das alterações no meio ambiente de maneira mais eficiente é a bioacústica, que utiliza o som como uma ferramenta para entender e gerenciar ecossistemas e populações de animais. O que torna essa técnica eficaz é a capacidade dos pesquisadores trabalharem de forma remota e não invasiva, minimizando o estresse causado aos animais [Kvsn et al. 2020]. Os pássaros são bons sentinelas, por serem observáveis e presentes em diferentes posições dos trópicos [Ferreira 2011] e, nesse contexto, o reconhecimento automatizado de eventos acústico de pássaros é considerando um importante componente tecnológico para atividades de monitoramento e conservação de biodiversidade [de Oliveira et al. 2020].

Entretanto, para essa abordagem, a quantidade de dados gerados para análise é diretamente proporcional à quantidade de tempo da observação/gravação e ao número de locais observados. Por isso, muitas técnicas computacionais têm sido estudadas, desenvolvidas e aplicadas para permitir a análise desses dados de maneira eficiente e escalável.

Diante disso, o objetivo deste trabalho foi revisar as técnicas de Inteligência Artificial (IA) aplicadas nos últimos anos ao contexto de bioacústica utilizando o método Prisma, que foi escolhido para avaliar a conformidade com os princípios estabelecidos para esse tipo de estudo. O método PRISMA facilita a avaliação da revisão, assegurando que os critérios e diretrizes específicos sejam seguidos de maneira rigorosa e precisa. Isso contribui para a qualidade e transparência no desenvolvimento e aplicação nos dados bioacústicos de estações de coletas no pantanal norte Mato-Grossense.

## **2. Metodologia**

De acordo com [Zhao et al. 2017], o processo do monitoramento acústico de forma automática pode ser resumido em 5 etapas: Gravação em campo; Pré-Processamento; Segmentação; Extração de Característica e Classificação. E para [Jing et al. 2016] há dois aspectos importantes, a representação do áudio e o aprendizado de máquina. Diante disso, para direcionar a pesquisa a qual esse trabalho se refere, foi proposta a realização da revisão de literatura a respeito dos métodos de aprendizagem de máquina recentemente utilizados para posterior estudo, aplicação e adequação, considerando também as etapas do processo de monitoramento acústico. Diante disso, este trabalho apresenta a revisão de literatura para direcionamento da pesquisa.

A revisão seguiu a metodologia PRISMA - Preferred Reporting Items for Systematic Reviews and Meta-Analyse, descrita em [Moher et al. 2009]. A seleção deste método reside na sua amplitude, uma vez que abrange uma considerável parcela dos aspectos de relevância dessa revisão sistemática. Além disso, apresenta facilidades na avaliação da referida revisão.

Na seleção dos trabalhos para análise, foi empregada a ferramenta de busca Google Scholar, com a seguinte consulta: "Bird monitoring"+ "Bioacoustics"+ "Artificial Intelligence". O processo de pesquisa de acordo com o método PRISMA foi subdividido em quatro fases distintas: identificação, triagem, elegibilidade e inclusão. Foram inicialmente coletados os artigos que atendiam aos critérios iniciais, os quais incluíam a necessidade de estarem em inglês, e terem sido publicados no período de 2018 a 2023. Após essa etapa, 44 artigos se enquadraram nesses parâmetros. Durante essa etapa, foram excluídos os trabalhos que não demonstraram inovação em seus métodos, assim como pré-prints. Após esse processo foi priorizado a inclusão daqueles que apresentavam informações acerca da

acurácia na classificação dos dados. Com o resultado dessa análise criteriosa, 11 trabalhos foram selecionados como elegíveis para prosseguir na pesquisa.

### 3. Resultados e Discussões

Após a filtragem da base de artigos baseado nos critérios definidos, 11 artigos foram selecionados e estão listados na Tabela 1. Os artigos selecionados foram agrupados de acordo com a técnica de machine learning utilizada, o que resultou em 9 grupos.

Em [Alghamdi et al. 2021], Redes Neurais Artificiais usando algoritmo de *back-propagation* foram utilizadas para classificação de áudio de pássaros que foram segmentados em vocalizações monossilábica e polissilábica, alcançando acurácia de 98%. Já [B. T. Padovese 2019] utilizou um Multilayer Perceptron para monitoramento do pássaro da amazônia brasileira, a *Rhodocorytha*, alcançando resultados de até 98% de acurácia.

**Tabela 1. Artigos selecionados após aplicação dos critérios segundo a metodologia PRISMA agrupados por modelo aplicado.**

Modelo	Citação
Artificial Neural Network	[Alghamdi et al. 2021]
Lightweight Neural Networks	[Yang et al. 2022]
Multilayer Perceptron	[B. T. Padovese 2019]
Residual Neural Network	[Xiao et al. 2022] e [Hu et al. 2023]
Convolutional Neural Network	[Hidayat et al. 2021] e [Maegawa et al. 2021]
Support Vector Machine (SVM)	[Weerasena et al. 2018]
Spiking Neural Network	[Mohanty et al. 2023]
Naive Bayes Classifier	[Tivarekar et al. 2018]
Recurrent Neural Network	[Xie et al. 2020]

Em [Yang et al. 2022] foi proposto um modelo de reconhecimento de pássaros usando a arquitetura Lightweight Neural Networks (LWN) que são um subconjunto das Redes de Peso Contínuo convencionais (CWN). Foi usado um conjunto de dados de Mel Spectrogram Min-Max Standardization, alcançando uma acurácia de 95,12%.

[Xiao et al. 2022] apresentam o AMResnet, um modelo que utiliza Residual Neural Network (RNN) para classificação de pássaros. Os dados testados foram dados de 20 pássaros gravados em ambiente real. O modelo proposto alcançou 92,6% de acurácia. Já [Hu et al. 2023] apresentam uma rede de fusão de recursos (MFF-ScSEnet) afim de evitar a perda de informações extraídas no processo de gerar o espectrograma. Foi utilizado o ResNet18, que aprimora as informações de ondulação sonora do espectrograma.

Em [Hidayat et al. 2021], foi utilizada a arquitetura Convolutional Neural Network (CNN) de quatro camadas formulada para classificar diferentes espécies de corujas da Indonésia com base em suas vocalizações e alcançou uma acurácia de 97,06%. Em [Maegawa et al. 2021], foi apresentada uma ferramenta para acompanhamento de uma espécie de pássaro com o uso de CNN e com acurácia geral de 97% para este sistema.

[Weerasena et al. 2018] utilizam Suport Vector Machine (SVM) para criação de

um classificador local, para funcionar junto aos dispositivos de captura em campo. O sistema foi validado com 93,5% de acurácia utilizando 214 arquivos de som de 5 espécies.

[Mohanty et al. 2023] apresentam o uso da Spiking Neural Network para classificação de dados de vocalização de pássaros em diferentes condições, alcançando 94% de acurácia num tempo de 4,16ms mais rápido do que as outras soluções comparadas no mesmo trabalho.

Por fim, Naive Bayes Classifier foram utilizadas em [Tivarekar et al. 2018] para classificação de cantos de pássaros na Índia. E Recurrent Neural Network em [Xie et al. 2020] foram utilizadas para um novo modelo de identificação individual de pássaros com acurácia de 97%.

#### 4. Conclusão

Este trabalho apresentou a revisão da literatura a respeito da Bioacústica aplicada ao monitoramento ambiental utilizando sons de pássaros, utilizando a metodologia PRISMA para direcionamento de pesquisas da área.

Após a aplicação da metodologia, 11 artigos foram selecionados. Há de se destacar o bom resultado alcançado por modelos já utilizados há muito tempo, como Multi-layer Perceptron em [B. T. Padovese 2019]. Nos últimos três anos, observa-se uma clara dominância das Residual Neural Network e Convolutional Neural Network com quatro trabalhos que se encaixaram nos critérios definidos, dois em cada pesquisa.

Com os resultados alcançados, pesquisas iniciais podem ser direcionadas para os modelos encontrados e posterior decisão a respeito do modelo mais adequado para os dados a serem trabalhados. Como trabalho futuro, sugere-se que a metodologia PRISMA seja aplicada considerando um conjunto maior de palavras-chave, buscando-se extrair ainda mais trabalhos importantes sobre monitoramento ambiental por meio da bioacústica utilizando sons de pássaros.

#### Referências

- Alghamdi, A., Mehtab, T., Iqbal, R., Leeza, M., Islam, N., Hamdi, M., and Shaikh, A. (2021). Automatic classification of monosyllabic and multisyllabic birds using pdhf. *Electronics*, 10(5).
- B. T. Padovese, L. R. P. (2019). Machine learning for identifying an endangered brazilian psittacidae species. *Journal of Environmental Informatics Letters*.
- de Oliveira, A. G., Ventura, T. M., Ganchev, T. D., Silva, L. N., Marques, M. I., and Schuchmann, K.-L. (2020). Speeding up training of automated bird recognizers by data reduction of audio features. *PeerJ*, 8:e8407.
- Ferreira, A. (2011). Assessment of heavy metals in egretta thula: case study: Coroa grande mangrove, sepetiba bay, rio de janeiro, brazil. *Brazilian Journal of Biology*, 71(1):77–82.
- Hidayat, A. A., Cenggoro, T. W., and Pardamean, B. (2021). Convolutional neural networks for scops owl sound classification. *Procedia Computer Science*, 179:81–87. 5th International Conference on Computer Science and Computational Intelligence 2020.

- Hu, S., Chu, Y., Wen, Z., Zhou, G., Sun, Y., and Chen, A. (2023). Deep learning bird song recognition based on mff-scenet. *Ecological Indicators*, 154:110844.
- Jing, L., Liu, B., Choi, J., Janin, A., Bernd, J., Mahoney, M. W., and Friedland, G. (2016). A discriminative and compact audio representation for event detection. In *Proceedings of the 24th ACM International Conference on Multimedia*, MM '16, page 57–61, New York, NY, USA. Association for Computing Machinery.
- Kvsn, R. R., Montgomery, J., Garg, S., and Charleston, M. (2020). Bioacoustics data analysis – a taxonomy, survey and open challenges. *IEEE Access*, 8:57684–57708.
- Maegawa, Y., Ushigome, Y., Suzuki, M., Taguchi, K., Kobayashi, K., Haga, C., and Matsui, T. (2021). A new survey method using convolutional neural networks for automatic classification of bird calls. *Ecological Informatics*, 61:101164.
- Mohanty, R., Bhuyan, H. K., Pani, S. K., Ravi, V., and Krichen, M. (2023). Bird species recognition using spiking neural network along with distance based fuzzy co-clustering. *International Journal of Speech Technology*, pages 1–14.
- Moher, D., Liberati, A., Tetzlaff, J., and Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the prisma statement. *BMJ*, 339.
- Tivarekar, R. P., Chavan, V. D., Shete, S. A., and Vartak, A. (2018). Audio based bird species recognition using naïve bayes algorithm. *International Journal of modern Trends in Engineering and Research (IJMTER)*.
- Weerasena, H., Jayawardhana, M., Egodage, D., Fernando, H., Sooriyaarachchi, S., Gamage, C., and Kottege, N. (2018). Continuous automatic bioacoustics monitoring of bird calls with local processing on node level. In *TENCON 2018 - 2018 IEEE Region 10 Conference*, pages 0235–0239.
- Xiao, H., Liu, D., Chen, K., and Zhu, M. (2022). Amresnet: An automatic recognition model of bird sounds in real environment. *Applied Acoustics*, 201:109121.
- Xie, J., Yang, J., Ding, C., and Li, W. (2020). High accuracy individual identification model of crested ibis (*nipponia nippon*) based on autoencoder with self-attention. *IEEE Access*, 8:41062–41070.
- Yang, F., Jiang, Y., and Xu, Y. (2022). Design of bird sound recognition model based on lightweight. *IEEE Access*, 10:85189–85198.
- Zhao, Z., hua Zhang, S., yong Xu, Z., Bellisario, K., hua Dai, N., Omrani, H., and Pijanowski, B. C. (2017). Automated bird acoustic event detection and robust species classification. *Ecological Informatics*, 39:99–108.