

Anuracloud: Proposta de Plataforma em Nuvem para Detecção por Aprendizado de Máquina de Chamados Bioacústicos de Anuros

Renan H. Susuki¹, Allan G. de Oliveira¹, Thiago M. Ventura¹

¹Instituto de Computação – Universidade Federal de Mato Grosso (UFMT)
Av. Fernando Corrêa da Costa, nº 2367 – 78.060-900 – Cuiabá – MT – Brazil

renan.susuki@ufmt.br, {allan,thiago}@ic.ufmt.br

Abstract. *Passive acoustic monitoring of anurans, which are important ecological bioindicators, generates a data flood that makes manual screening by experts unfeasible due to high cost and time demands. To address this challenge, this work proposes a cloud-based methodology for collaborative bioacoustic monitoring. The approach leverages artificial intelligence, focusing on deep learning and convolutional neural networks (CNNs), to automate call detection and classification. The architecture integrates open-source tools and cloud services, enabling scalable audio storage, collaborative annotation of spectrograms by multiple simultaneous users, and model training on high-performance instances. Thus, the project aims to promote open science and data sharing.*

Resumo. *O Monitoramento Acústico Passivo de anuros, que são importantes bioindicadores ecológicos, gera uma “inundação de dados” que inviabiliza a triagem manual por especialistas devido ao alto custo e tempo exigidos. Para contornar esse desafio, este trabalho propõe uma metodologia baseada em uma plataforma em nuvem para o monitoramento bioacústico colaborativo. A abordagem utiliza inteligência artificial, com foco em aprendizado profundo e redes neurais convolucionais (CNNs), para automatizar a detecção e classificação de vocalizações. A arquitetura integra ferramentas de código aberto e serviços em nuvem, viabilizando o armazenamento escalável de áudios, a anotação colaborativa em espectrogramas por múltiplos usuários e o treinamento de modelos em instâncias de alta performance. Dessa forma, o projeto visa fomentar a ciência aberta e o compartilhamento de dados.*

1. Introdução

O monitoramento ambiental tem se beneficiado amplamente da bioacústica, uma abordagem não invasiva que envolve a gravação passiva de sons produzidos por animais em seus habitats naturais, viabilizando o monitoramento em larga escala [Kvsn et al. 2020]. Nesse contexto, o monitoramento de anuros (sapos, rãs e pererecas) é relevante, pois são considerados bioindicadores eficazes, ou seja, as dinâmicas de suas populações refletem diretamente a saúde e a qualidade dos habitats que ocupam, como pântanos e áreas alagadas [Strout et al. 2017].

Apesar dos benefícios, o avanço tecnológico na captura de áudio gerou um desafio substancial: o grande volume de dados capturados. Segundo [Nolasco et al. 2023], o monitoramento contínuo tem provocado uma “inundação de dados” (*data deluge*), onde

a taxa de captação de áudios excede as capacidades de gerenciamento e de avaliação manual humana. Ou seja, segundo os autores, analisar e anotar de forma precisa muitas horas de gravações exige um esforço demorado, custoso e dependente do raro tempo de especialistas.

Para lidar com a inviabilidade da análise manual, a computação tem oferecido soluções viáveis por meio de técnicas de Inteligência Artificial (IA). Conforme [Sharma et al. 2023], o uso de aprendizado de máquina (*machine learning*) e aprendizado profundo (*deep learning*), como as Redes Neurais Convolucionais (CNNs), permite automatizar a extração, a detecção e a classificação de sons da vida selvagem diretamente dos dados brutos e em meio a ruídos de fundo. Os autores apontam que essas técnicas computacionais diminuem o tempo e o esforço necessários para o processamento em grandes escalas.

Diante das necessidades tecnológicas e bioacústicas, este trabalho propõe a Anuracloud, uma metodologia via plataforma em nuvem para monitoramento ambiental de anuros, de forma aberta e colaborativa, resolvendo os problemas citados. A Anuracloud promove cooperação entre pesquisadores, com compartilhamento de *datasets*, anotação colaborativa para treinamento de modelos e classificação de espécies anotadas. Em geral, a plataforma em nuvem é fundamental, permitindo armazenamento de grandes volumes de dados bioacústicos, anotação de *datasets* e ambiente de alta performance para rodar modelos CNN.

2. Trabalhos relacionados

O trabalho de [Stowell 2022] faz uma revisão de literatura do estado da arte sobre uso de *deep learning* em bioacústica computacional. Os autores apontam que as redes neurais convolucionais (CNNs) são a arquitetura de rede neural amplamente utilizada em bioacústica computacional. Espectrogramas, particularmente mel-espectrogramas, são a representação de dados de entrada mais comum. Segundo os autores, a falta de ferramentas acessíveis ainda é um desafio na área. Muitos projetos de *deep learning* em bioacústica são disponibilizados apenas como *scripts* em Python, o que, embora favoreça a reprodutibilidade técnica, afasta a comunidade de zoólogos e conservacionistas devido à ausência de interfaces de usuário (UIs) amigáveis.

Além disso, conforme [Stowell 2022], destacam que a ciência aberta possibilita o progresso em bioacústica com *deep learning*, pois a publicação aberta de *datasets*, arquiteturas de redes neurais, pesos pré-treinados e códigos-fonte tem contribuído para esse avanço. Exemplo disso é o repositório Xeno-Canto (XC) de [Vellinga et al. 2026], que contribui de forma relevante para a divulgação de sons da natureza, melhorando a acessibilidade aos registros sonoros da vida selvagem, principalmente de aves.

O artigo de [Gibbons et al. 2023] descreve o NEAL (*Nature+Energy Audio Labeller*), ferramenta de código aberto em R e Shiny para anotação de vocalizações de aves. Permite carregar áudios, visualizar espectrogramas, delimitar e rotular vocalizações com espécies, tipos de chamadas e metadados, além de filtrar, exibir locais e exportar dados. Aplicada a mais de 1.700 horas de gravações em parques eólicos na Irlanda, a ferramenta identificou espécies, inclusive raras, gerando dados para modelos de *machine learning* em detecção automática e gestão de biodiversidade. Embora o NEAL demonstre o potencial de ferramentas para anotação bioacústica em aves, ainda há uma lacuna na adaptação e

validação de plataformas semelhantes especificamente para anuros.

O *workflow* desenvolvido por [Ruff et al. 2021], baseado em redes neurais convolucionais para a detecção de aves, reduz significativamente o esforço humano em comparação à revisão manual completa. Apesar dos autores disponibilizem uma interface gráfica para o usuário, o problema da inundação de dados não é totalmente resolvido, uma vez que a execução via aplicativo *desktop* mantém o gargalo de processamento e armazenamento no *hardware* local do usuário.

3. Proposta de materiais e métodos

A metodologia para a proposta de plataforma em nuvem voltada à detecção por aprendizado de máquina de dados bioacústicos de anuros atende às necessidades identificadas no contexto de monitoramento ambiental. Esses requisitos incluem o armazenamento escalável de grandes volumes de gravações de registros sonoros de anuros, a anotação colaborativa por múltiplos pesquisadores para superar limitações manuais e a capacidade de processar modelos de *deep learning* com GPUs para análise precisa de chamados de anuros em espectrogramas.

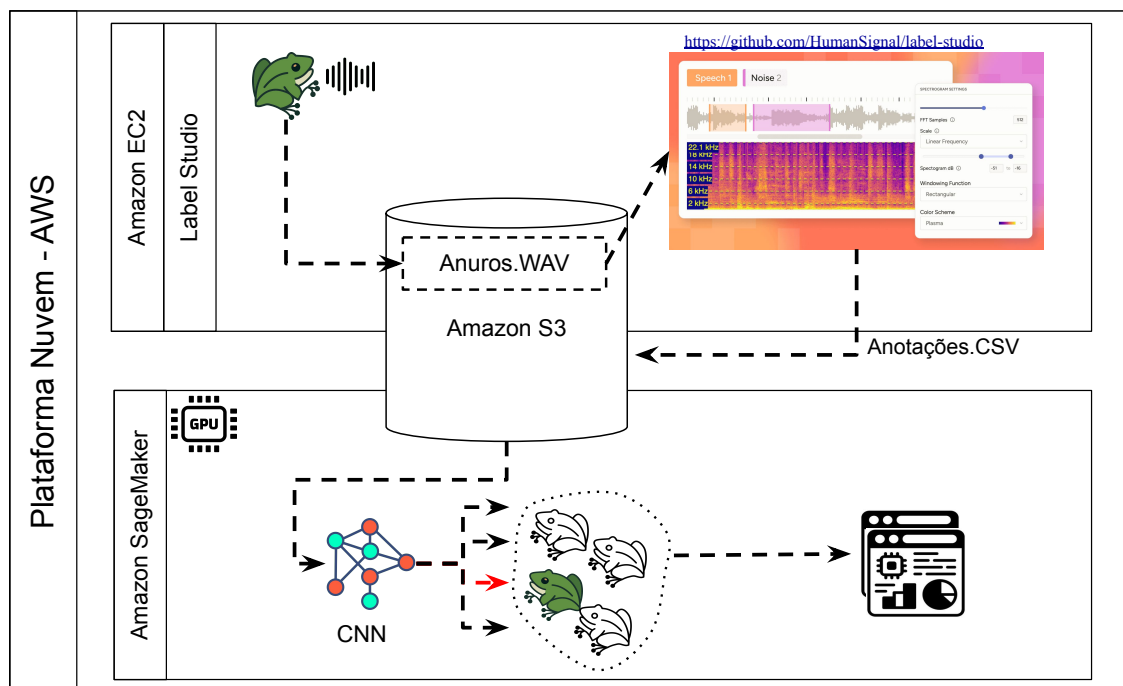


Figura 1. Pipeline do Anuracloud.

Para atender a essas demandas, a Figura 1 ilustra o *pipeline* no qual se propõe a integração de serviços da Amazon Web Services (AWS) com o Label Studio (*labelstudio.io*), ferramenta *web open-source* de anotação de dados que suporta múltiplos formatos, incluindo áudio e imagens. O Label Studio será hospedado no Amazon EC2 e integrado ao Amazon S3 (utilizado como solução de armazenamento em nuvem), permitindo que os usuários carreguem arquivos de áudio bioacústico, como o formato .WAV.

Por meio desse sistema, os usuários poderão colaborar na análise e na anotação dos dados, com geração de espectrogramas e delimitação de segmentos de vocalizações

de anuros, o que inclui a rotulagem de espécies e o registro de metadados ambientais para a construção de *datasets*. Por fim, as anotações serão salvas nos formatos JSON ou CSV para serem utilizadas na etapa de treinamento dos modelos.

O treinamento de modelos ocorre no Amazon SageMaker, utilizando instâncias GPU escaláveis para *pipelines* de *deep learning* que consomem dados do S3, geram mel-espectrogramas e experimentam CNNs, registrando métricas. A solução também acomoda outros modelos de aprendizado de máquina além das CNNs, incluindo abordagens tradicionais como SVM e Random Forest, bem como extração de características acústicas clássicas (MFCC, *zero-crossing rate*) para classificação de chamados de anuros.

Além disso, busca-se promover a ciência aberta, com compartilhamento de *datasets* e modelos treinados em repositórios públicos, facilitando a reprodutibilidade de experimentos e a colaboração interdisciplinar entre ecólogos e especialistas em IA.

4. Resultados esperados e considerações

Espera-se que a implementação da metodologia proposta resulte no desenvolvimento funcional da Anuracloud, estabelecendo uma infraestrutura em nuvem escalável que integre serviços de armazenamento e processamento da AWS com a ferramenta *open source* Label Studio. Como resultado direto, a plataforma deverá viabilizar o armazenamento eficiente de grandes volumes de dados bioacústicos no Amazon S3, solucionando o gargalo de gerenciamento causado pela “inundação de dados” (*data deluge*) típica do monitoramento acústico passivo contínuo.

A partir do uso da plataforma, projeta-se também uma otimização significativa no processo de triagem acústica por meio da anotação colaborativa. Espera-se que múltiplos pesquisadores possam atuar na geração de espectrogramas, na delimitação de vocalizações e na rotulação de espécies de anuros, exportando arquivos estruturados (em JSON ou CSV) que servirão de base para o treinamento automatizado. Consequentemente, prevê-se que o uso do Amazon SageMaker com instâncias GPU escaláveis permita a execução e a experimentação fluida de modelos de *deep learning*, como as Redes Neurais Convolucionais (CNNs), sobre mel-espectrogramas para a identificação precisa das espécies.

Por fim, destaca-se que a Anuracloud tem a possibilidade de contribuição de impulsionar a adoção de práticas de ciência aberta na bioacústica de anuros. Ao facilitar o compartilhamento de conjuntos de dados e modelos preditivos, a plataforma pretende reduzir as barreiras tecnológicas e de infraestrutura computacional para ecólogos, inclusive aqueles em regiões remotas.

Dessa forma, a proposta preenche uma lacuna importante de ferramentas adaptadas especificamente para anuros, fornecendo um recurso tecnológico para o monitoramento desses animais que são bioindicadores essenciais da saúde e da qualidade dos ecossistemas.

5. Declaração de uso de Inteligência Artificial e ausência de interesses

A inteligência artificial foi utilizada neste trabalho de forma assistida, restrita à correção gramatical, sínteses de autores das referências bibliográficas e auxílio na redação, sem geração autônoma ou não controlada de informações.

Declara-se também que esta pesquisa não recebeu financiamento externo nem apresenta conflitos de interesses comerciais associados às soluções tecnológicas propostas neste trabalho.

Referências

- Gibbons, A., Donohue, I., Gorman, C., King, E., and Parnell, A. (2023). NEAL: an open-source tool for audio annotation. *PeerJ*, 11:e15913.
- 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.
- Nolasco, I., Singh, S., Morfi, V., Lostanlen, V., Strandburg-Peshkin, A., Vidaña-Vila, E., Gill, L., Pamuła, H., Whitehead, H., Kiskin, I., Jensen, F. H., Morford, J., Emmerson, M. G., Versace, E., Grout, E., Liu, H., Ghani, B., and Stowell, D. (2023). Learning to detect an animal sound from five examples. *Ecological Informatics*, 77:102258.
- Ruff, Z. J., Lesmeister, D. B., Appel, C. L., and Sullivan, C. M. (2021). Workflow and convolutional neural network for automated identification of animal sounds. *Ecological Indicators*, 124:107419.
- Sharma, S., Sato, K., and Gautam, B. P. (2023). A Methodological Literature Review of Acoustic Wildlife Monitoring Using Artificial Intelligence Tools and Techniques. *Sustainability*, 15(9):7128.
- Stowell, D. (2022). Computational bioacoustics with deep learning: a review and roadmap. *PeerJ*, 10:e13152.
- Strout, J., Rogan, B., Seyednezhad, S. M., Smart, K., Bush, M., and Ribeiro, E. (2017). Anuran call classification with deep learning. In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2662–2665, New Orleans, LA. IEEE.
- Vellinga, W.-P., Creuwels, J., and Raes, N. (2026). Xeno-canto - Bird sounds from around the world.