

Marcos Antônio Gonçalves Barros Brito
Orientador: Alcilene Dalília de Sousa

Análise da Intensidade das Colmeias Utilizando Processamento de Áudio e Aprendizado de Máquina

Picos - PI
5 de junho de 2025

Marcos Antônio Gonçalves Barros Brito
Orientador: Alcilene Dalília de Sousa

Análise da Intensidade das Colmeias Utilizando Processamento de Áudio e Aprendizado de Máquina

Modelo de Trabalho de Conclusão de Curso em Bacharelado em Sistemas de Informação na Universidade Federal do Piauí. Este modelo está em conformidade com as normas ABNT.

Universidade Federal do Piauí
Campus Senador Heuvídio Nunes de Barros
Bacharelado em Sistemas de Informação

Picos - PI
5 de junho de 2025

FICHA CATALOGRÁFICA
Serviço de Processamento Técnico da Universidade Federal do Piauí
Biblioteca José Albano de Macêdo

B862a Brito, Marcos Antônio Gonçalves Barros.

Análise da intensidade das colmeias utilizando processamento de áudio e aprendizado de máquina / Marcos Antônio Gonçalves Barros Brito – 2025.
32 f.

1 Arquivo em PDF.

Indexado no catálogo *online* da biblioteca José Albano de Macêdo-CSHNB
Aberto a pesquisadores, com restrições da Biblioteca.

Trabalho de Conclusão de Curso (Graduação) – Universidade Federal do
Piauí, Curso de Bacharelado em Sistemas de Informação, Picos, 2025.
“Orientador: Alcilene Dalília de Sousa”.

1. Sistemas informacionais. 2. Aprendizado de máquinas. 3.
Monitoramento de colmeias I. Brito, Marcos Antônio Gonçalves Barros. II.
Sousa, Alcilene Dalília de. III. Título.

CDD 005.7

Elaborada por Maria Letícia Cristina Alcântara Gomes
Bibliotecária CRB nº 03/1835

Análise da Intensidade das Colmeias Utilizando Processamento de Áudio e Aprendizado de
Máquina

Marcos Antônio Gonçalves Barros Brito

Monografia aprovada como exigência parcial para obtenção do grau de Bacharel em Sistemas
de Informação.

Data de Aprovação

Picos – PI, 26 de junho de 2025

Documento assinado digitalmente
 **ALCILENE DALILIA DE SOUSA**
Data: 02/07/2025 21:17:56-0300
Verifique em <https://validar.iti.gov.br>

Prof^a. Dra. Alcilene Dalília de Sousa - Orientadora

Documento assinado digitalmente
 **FLAVIO HENRIQUE DUARTE DE ARAUJO**
Data: 01/07/2025 16:02:15-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Flávio Henrique Duarte De Araújo

Documento assinado digitalmente
 **ANTONIO OSEAS DE CARVALHO FILHO**
Data: 01/07/2025 15:16:36-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Antonio Oseas de Carvalho Filho

Agradecimentos

Primeiramente, agradeço à minha avó, Adalgisa Barros Brito, pelo amor incondicional, pela força que sempre me deu e pelas orações que me acompanharam em cada etapa dessa jornada, sendo meu sol e minhas estrelas. Agradeço também à professora orientadora, Alcilene Dalilia, pela paciência, pelas críticas construtivas e pela dedicação em guiar este trabalho. Aos demais professores do curso, minha gratidão pelo apoio fundamental em todas as etapas da minha formação. Aos meus amigos, reconheço o companheirismo, as risadas que aliviaram o estresse e o incentivo constante. Vocês foram essenciais para manter minha sanidade nos períodos mais intensos. Por fim, agradeço à minha família pelo apoio emocional e a todos que, direta ou indiretamente, contribuíram para esta conquista.

O que não é bom para a colmeia, não pode ser bom para a abelha.

Marco Aurélio

Resumo

A apicultura é essencial para a agricultura, a economia e o meio ambiente, pois além de gerar mel, cera e outros derivados, as abelhas realizam a polinização de inúmeras culturas agrícolas e plantas silvestres. Portanto, o manejo eficiente das colmeias não só é crucial para garantir a saúde das abelhas e a produtividade dos apiários, mas também para promover a diversidade biológica. Um dos principais desafios enfrentados pelos apicultores é a avaliação precisa e não invasiva da intensidade e saúde das colmeias. As técnicas convencionais de avaliação podem ser trabalhosas e invasivas, levando ao estresse das abelhas e à perturbação do ambiente da colmeia, ademais são limitadas em sua capacidade de detectar problemas sutis, especialmente durante períodos críticos, como ocorre durante períodos de escassez de alimentos. Além disso, a necessidade de inspeções frequentes em grandes apiários torna o processo ainda mais desafiador e trabalhoso, o que pode afetar negativamente a eficiência da produção e a saúde das abelhas. Para contornar essas limitações, propomos empregar o modelo AST (*Audio Spectrogram Transformer*) da HuggingFace combinado a características espectrais, classificando gravações de áudio das colmeias como fortes ou fracas. Utilizou-se um conjunto de 126 gravações rotuladas, segmentadas em janelas de 15 s e convertidas em espectrogramas log-Mel; desse total, 398 amostras foram destinadas ao treinamento e 100 aos testes. Diversos hiperparâmetros foram avaliados, e a configuração ótima — taxa de aprendizado de 0,001, warm-up de 50% e agendamento cosseno — gerou 94,4% de acurácia e *eval loss* de 0,119, superando métodos clássicos como MFCC + Random Forest.

Palavras-chaves: Apicultura, Monitoramento Acústico, Aprendizado de Máquina, AST, Classificação de Colmeias.

Abstract

Beekeeping is essential to agriculture, the economy, and the environment: in addition to producing honey, wax, and other derivatives, bees pollinate countless agricultural crops and wild plants. Efficient hive management is therefore crucial not only to ensure bee health and apiary productivity, but also to foster biodiversity. One of the main challenges beekeepers face is obtaining precise, non-invasive assessments of colony strength and health. Conventional evaluation methods can be labor-intensive and invasive, stressing the bees and disturbing the hive environment, and they are limited in detecting subtle problems—especially during critical periods such as food shortages. Frequent inspections in large apiaries further compound the workload, potentially reducing production efficiency and compromising bee well-being. To overcome these limitations, we propose using HuggingFace’s AST (Audio Spectrogram Transformer) model combined with spectral features to classify hive audio recordings as strong or weak. A dataset of 126 labeled recordings was segmented into 15-second windows and converted to log-Mel spectrograms; 398 samples were allocated for training and 100 for testing. After evaluating several hyperparameter configurations, the best setup—a learning rate of 0.001, 50% warm-up, and a cosine scheduler—achieved 94.4% accuracy and an evaluation loss of 0.119, outperforming classical approaches such as MFCC + Random Forest.

Lista de ilustrações

Figura 1 – Fluxograma representativo do método proposto que está dividido em 4 etapas principais	22
Figura 2 – Gráfico de forma de onda de uma amostra de áudio.	23
Figura 3 – Espectrograma log-Mel gerado a partir de um segmento de áudio da colmeia.	24
Figura 4 – Evolução da <i>eval_accuracy</i> do run awesome-perch ao longo dos passos de treinamento. Observa-se que o maior valor (94,4%) ocorre no passo 110, sendo este o checkpoint adotado como modelo otimizado. . .	26
Figura 5 – Comportamento da <i>eval_loss</i> . O menor valor (0,119) coincide com o pico de acurácia na Figura 4, indicando bom equilíbrio entre erro e capacidade de generalização.	27
Figura 6 – Curva de <i>train_loss</i> . A queda pronunciada após o passo 380 confirma a convergência do modelo sem indícios de sobreajuste até o checkpoint escolhido.	28

Lista de tabelas

Tabela 1 – Trabalhos relacionados.	21
Tabela 2 – Desempenho dos cinco checkpoints comparados.	25
Tabela 3 – Hiperparâmetros detalhados dos cinco modelos selecionados.	26

Lista de abreviaturas e siglas

Ac	Acurácia
API	Application Programming Interface
AST	<i>Audio Spectrogram Transformer</i>
CHCl ₃	Triclorometano (clorofórmio)
CNN	<i>Convolutional Neural Network</i>
DNN	<i>Deep Neural Network</i>
Eval Loss	Perda de validação (<i>Evaluation Loss</i>)
FN	<i>False Negative</i> (Falso-negativo)
FP	<i>False Positive</i> (Falso-positivo)
IoT	<i>Internet of Things</i>
kHz	Quilo-hertz (10 ³ Hz)
LR	<i>Learning Rate</i> (Taxa de aprendizado)
MFCC	<i>Mel Frequency Cepstral Coefficient</i>
MP3	<i>Moving Picture Experts Group – Layer III</i> (formato de áudio)
S3	<i>Amazon Simple Storage Service</i>
SVM	<i>Support Vector Machine</i>
TP	<i>True Positive</i> (Verdadeiro-positivo)
TN	<i>True Negative</i> (Verdadeiro-negativo)
WD	<i>Weight Decay</i>

Lista de símbolos

Ac	Acurácia, definida pela Eq. eq:accuracy
FN	Número de falsos-negativos
FP	Número de falsos-positivos
$Eval Loss$	Função de perda calculada sobre o conjunto de validação
LR	Taxa de aprendizado utilizada no treinamento
$Loss$	Função de perda calculada sobre o conjunto de treinamento
VN	Número de verdadeiros-negativos
VP	Número de verdadeiros-positivos

Sumário

1	Introdução	13
1.1	Objetivos Geral e Específicos	14
1.1.1	Objetivos Específicos	14
1.2	Estrutura do Trabalho	14
2	Referencial Teórico	15
2.1	Monitoramento Acústico de Colmeias	15
2.2	Aprendizado de Máquina em Apicultura	16
2.3	Modelo AST (<i>Audio Spectrogram Transformer</i>)	17
2.4	Métricas de Avaliação	18
3	Trabalhos relacionados	19
4	Metodologia Proposta	22
4.1	Aquisição dos dados	22
4.2	Pré-processamento	23
4.3	Extração de características	23
4.4	Classificação	24
5	Resultados	25
6	Conclusão	29
	Referências	30

1 Introdução

A função das abelhas vai muito além da produção de mel e própolis. As abelhas desempenham um papel essencial na polinização, contribuindo diretamente para a produtividade de diversas culturas e gerando para a economia brasileira aproximadamente U\$ 12 bilhões com o serviço de polinização (VIEIRA; ANDRADE; RIBEIRO, 2021). Nesse contexto, a saúde e o vigor das colmeias são elementos fundamentais para garantir a continuidade desse ciclo produtivo. Colmeias saudáveis refletem uma apicultura forte e eficiente, enquanto colmeias fragilizadas exigem intervenções adequadas para restabelecer sua funcionalidade (GORROI; FREITAS; ASSIS, 2020).

Entretanto, o manejo apícola, embora necessário, pode trazer efeitos colaterais. As inspeções físicas frequentes — realizadas para verificar a presença da rainha, o padrão de postura e a presença de predadores — podem causar estresse às colônias, especialmente em períodos de escassez alimentar, aumentando o risco de colapsos e perdas (COSTA et al., 2016). Por outro lado, a ausência de monitoramento pode permitir que problemas se agravem silenciosamente, comprometendo a saúde da colmeia e impactando negativamente a produção.

Com o avanço da tecnologia, estratégias de monitoramento remoto têm ganhado destaque. Dispositivos baseados em sensores, câmeras e soluções de Internet das Coisas (IoT) vêm sendo utilizados para reduzir a necessidade de intervenções manuais. No entanto, essas soluções muitas vezes envolvem altos custos de implementação e, em muitos casos, não oferecem dados suficientemente detalhados para uma tomada de decisão eficaz (LIAKOS et al., 2018).

Diante desse cenário, métodos alternativos têm sido explorados. O monitoramento acústico das colmeias surge como uma abordagem promissora, capaz de captar padrões sonoros que refletem o comportamento e o estado interno das colônias. Estudos recentes demonstram que é possível identificar eventos relevantes, como ausência da rainha, comportamento de enxameação e alterações na atividade das abelhas, por meio da análise dos sons produzidos no interior das colmeias (SHOSTAK; PRODEUS, 2019; KIM; OH; HEO, 2021).

O uso de sinais sonoros como fonte de informação representa uma alternativa acessível, menos invasiva e potencialmente automatizável para o monitoramento apícola. A identificação precoce de padrões acústicos associados a alterações no comportamento das colônias pode contribuir para uma gestão mais eficiente, reduzindo perdas e promovendo a sustentabilidade da apicultura.

1.1 Objetivos Geral e Específicos

O objetivo geral desta pesquisa é analisar e implementar uma solução baseada em aprendizado de máquina para o monitoramento não invasivo da intensidade das colmeias de abelhas, utilizando dados de áudio processados e características espectrais. Para alcançar esse objetivo geral, os objetivos específicos deste trabalho são :

1.1.1 Objetivos Específicos

- Identificar e extrair características espectrais relevantes dos sinais de áudio das colmeias, que possam ser utilizadas para inferir seu estado de intensidade.
- Adaptar e treinar o modelo AST (*Audio Spectrogram Transformer*) da HuggingFace para o contexto específico da classificação de colmeias.
- Avaliar o desempenho do modelo treinado em um conjunto de dados de validação, utilizando métricas apropriadas para medir sua eficácia na classificação de colmeias.

1.2 Estrutura do Trabalho

Este trabalho está organizado em cinco capítulos, além desta introdução. Na Seção 2 é apresentado o referencial teórico com os principais conceitos relacionados ao monitoramento acústico de colmeias, aprendizado de máquina e modelo AST. A Seção 3 aborda os principais trabalhos relacionados ao tema, destacando suas contribuições e lacunas. A metodologia proposta é detalhada na Seção 4, incluindo as etapas de coleta, pré-processamento, extração de características, treinamento e validação do modelo. A Seção 5 apresenta e discute os resultados obtidos com a metodologia aplicada, enquanto na Seção 6 são apresentadas as conclusões obtidas e perspectivas para trabalhos futuros.

2 Referencial Teórico

Nesta Seção, serão apresentados conceitos fundamentais para a compreensão do projeto, incluindo monitoramento acústico de colmeias, aprendizado de máquina e técnicas relacionadas.

2.1 Monitoramento Acústico de Colmeias

O monitoramento acústico de colmeias é uma prática inovadora que se baseia na análise de sons para avaliar a saúde e o comportamento das abelhas. Esta técnica não invasiva permite aos apicultores e cientistas obter informações valiosas sobre o estado interno da colmeia sem perturbar as abelhas (TERENZI; CECCHI; SPINSANTE, 2020; QANDOOUR et al., 2014). A bioacústica, que é o estudo de como os seres vivos produzem e percebem o som, tem sido aplicada ao comportamento das abelhas para identificar padrões sonoros que indicam diferentes atividades e condições dentro da colmeia. Por exemplo, a frequência e intensidade dos zumbidos podem sinalizar a presença da rainha ou alertar sobre o estado de saúde das abelhas (QANDOOUR et al., 2014).

A análise acústica avançada pode detectar mudanças sutis nos sons das colmeias que podem ser indicativos de estresse, doença ou outros fatores ambientais adversos (SHARIF et al., 2020; TERENZI; CECCHI; SPINSANTE, 2020). Estudos internacionais têm explorado a relação entre os sons das colmeias e a saúde das abelhas, fornecendo uma base de conhecimento para o desenvolvimento de ferramentas de monitoramento acústico (SHARIF; DI; YU, 2023). Essas ferramentas são projetadas para capturar e analisar dados acústicos de forma contínua, oferecendo aos apicultores uma maneira de monitorar suas colmeias com maior precisão e menos intervenção.

Ademais, o uso de inteligência artificial e aprendizado de máquina para analisar esses dados permite uma interpretação mais precisa e rápida dos sinais, possibilitando intervenções tempestivas que podem salvar colônias de abelhas e melhorar a produtividade apícola. Os dados acústicos podem também ser utilizados para fins de pesquisa, ajudando cientistas a compreender melhor os comportamentos complexos das abelhas e as dinâmicas dentro das colmeias. O monitoramento contínuo e a análise de longo prazo podem revelar padrões sazonais e efeitos de fatores externos, como pesticidas e mudanças climáticas, sobre as abelhas (QANDOOUR et al., 2014). Dessa forma, o monitoramento acústico se apresenta como uma ferramenta multifacetada, essencial tanto para a prática apícola quanto para a pesquisa científica.

2.2 Aprendizado de Máquina em Apicultura

O aprendizado de máquina tem se mostrado uma tecnologia transformadora em muitos campos, incluindo a apicultura. A capacidade de processar e aprender com grandes conjuntos de dados torna o aprendizado de máquina ideal para interpretar os complexos sinais acústicos das colmeias (SOARES et al., 2022). Algoritmos de aprendizado de máquina podem ser treinados para reconhecer padrões específicos nos dados acústicos que correspondem a diferentes estados ou comportamento das abelhas (NOLASCO et al., 2019; SOARES et al., 2022). Isso permite uma detecção precoce de problemas e uma resposta mais rápida por parte dos apicultores, potencialmente salvando colônias de eventos adversos (ZGANK, 2019).

Utilizando dados históricos e em tempo real das colmeias, modelos preditivos podem ser treinados para antecipar eventos críticos, como surtos de doenças ou mudanças comportamentais significativas nas abelhas (NOLASCO et al., 2019; RUSTAM et al., 2024). Esses modelos podem ajudar os apicultores a tomar decisões proativas, minimizando perdas e maximizando a produção. Além disso, a integração do aprendizado de máquina com outras tecnologias, como IoT (Internet das Coisas), permite a criação de colmeias inteligentes, onde sensores coletam dados continuamente e algoritmos de aprendizado de máquina processam essas informações para fornecer *insights* detalhados sobre o bem-estar das abelhas e a eficácia das práticas de apicultura (RUSTAM et al., 2024).

Pesquisas recentes reforçam esse panorama ao mostrarem que a fusão de dados multissensoriais como temperatura interna, peso da colmeia e variáveis climáticas associada a registros de inspeção pode ser explorada por algoritmos supervisionados para inferir automaticamente o estado de saúde das colônias, potencializando práticas de manejo preventivo (BRAGA et al., 2020). Complementarmente, revisões sobre monitoramento vibroacústico indicam que sinais de som e vibração captados por microfones e acelerômetros constituem fontes ricas para detectar a presença da rainha e antecipar a enxameação, embora persistam desafios de padronização de atributos e de construção de modelos mais generalizáveis (UTHOFF; HOMSI; BERGEN, 2023).

O uso de aprendizado de máquina também pode facilitar a personalização das práticas de manejo apícola, adaptando intervenções às necessidades específicas de cada colmeia com base em dados em tempo real (RUSTAM et al., 2024). Por exemplo, modelos de aprendizado de máquina podem ser usados para ajustar a alimentação suplementar ou identificar a necessidade de tratamentos contra parasitas de maneira mais precisa do que os métodos tradicionais. Isso pode levar ao desenvolvimento de estratégias de manejo mais eficazes e sustentáveis, que considerem as particularidades de cada ambiente apícola.

2.3 Modelo AST (*Audio Spectrogram Transformer*)

O modelo AST (*Audio Spectrogram Transformer*) é uma das mais recentes inovações no campo do aprendizado de máquina aplicado ao processamento de áudio (GONG; CHUNG; GLASS, 2021). Utilizando uma abordagem baseada em *transformers*, o AST é capaz de analisar espectrogramas de áudio, que são representações visuais da intensidade do som ao longo do tempo como demonstrado em Verma e Berger (2021). Isso permite que o modelo identifique padrões complexos e nuances nos dados acústicos que seriam difíceis de detectar por outros meios. A eficácia do AST em tarefas de classificação de áudio foi demonstrada em diversos estudos (GONG; CHUNG; GLASS, 2021; ZHANG et al., 2022), estabelecendo-o como uma ferramenta promissora para o monitoramento acústico de colmeias.

A implementação do modelo AST em sistemas de monitoramento de colmeias pode levar a avanços significativos na capacidade dos apicultores de entender e responder às necessidades de suas colônias. Com sua capacidade de processar e aprender com grandes volumes de dados como demonstrado em (GONG; CHUNG; GLASS, 2021), o AST pode ajudar a identificar os primeiros sinais de problemas dentro da colmeia, permitindo intervenções mais informadas e oportunas. Isso representa um potencial enorme para melhorar a saúde das colmeias e a sustentabilidade da apicultura como um todo.

O modelo AST por ser um modelo puramente baseado em *transformers* e livre de convoluções, o que o diferencia dos tradicionais modelos híbridos que combinam CNNs com mecanismos de atenção (GONG; CHUNG; GLASS, 2021). Experimentos realizados demonstram que o AST alcança resultados de estado da arte em diversas bases de dados desafiadoras de classificação de áudio, como AudioSet, ESC-50 e Speech Commands, com métricas superiores às dos modelos CNN+atenção tradicionais, além de apresentar convergência mais rápida e arquitetura mais simples, com menos parâmetros (GONG; CHUNG; GLASS, 2021). Isso reforça o potencial do AST como um classificador genérico de áudio para diferentes durações e tipos de sinais sonoros.

Além disso, o AST pode ser integrado a plataformas para permitir o monitoramento remoto e contínuo das colmeias. Sensores de áudio conectados podem transmitir dados em tempo real para serem analisados pelo modelo, fornecendo atualizações instantâneas sobre a saúde das abelhas e permitindo ações imediatas quando necessário (GONG; CHUNG; GLASS, 2021). A flexibilidade do AST também permite sua adaptação para outras aplicações na apicultura, como a análise de fatores ambientais que afetam as abelhas e a identificação de padrões sazonais nos comportamentos das colônias. Estudos futuros podem explorar ainda mais essas possibilidades, ampliando o impacto dessa tecnologia no campo da apicultura (ZHANG et al., 2022).

2.4 Métricas de Avaliação

Em tarefas de classificação a escolha de métricas adequadas é fundamental para estimar o quão distante o modelo está do desempenho ideal e garantir sua capacidade de generalização (SAMMUT; WEBB, 2011). Neste trabalho, após o treinamento do modelo AST (*Audio Spectrogram Transformer*), emprega-se um conjunto de validação independente para estimar três métricas principais: Acurácia, *Loss* de treinamento e *Eval Loss*. A análise cuidadosa desses indicadores com o AST (GONG; CHUNG; GLASS, 2021), permite detectar sobreajuste (*overfitting*) e orientar melhorias no ajuste de hiperparâmetros.

- **Acurácia** — indica a proporção global de previsões corretas, independentemente da classe. É definida por

$$\text{Acurácia} = \frac{VP + VN}{VP + FP + VN + FN}, \quad (2.1)$$

onde VP , VN , FP e FN correspondem, respectivamente, a verdadeiros-positivos, verdadeiros-negativos, falsos-positivos e falsos-negativos (SAMMUT; WEBB, 2011).

- **Loss** — valor da função de perda (entropia cruzada binária) calculada em cada época durante o treinamento; quanto menor o *loss*, mais próximo o modelo está de reproduzir os rótulos corretos (GOODFELLOW et al., 2016).
- **Eval Loss** — mesma função de perda, porém computada no conjunto de validação. Serve para aferir a capacidade de generalização: aumento consistente do *Eval Loss* enquanto o *loss* de treinamento diminui sinaliza sobreajuste.

Após o término do treinamento de cada modelo candidato, as métricas coletadas (*Loss*, *Eval Loss* e Acurácia) serão comparados. A escolha do modelo final basear-se-á na melhor combinação de acurácia elevada e menor *Eval Loss*. Em seguida, esses resultados serão comparados com os relatados na literatura em especial trabalhos que empregam técnicas tradicionais de análise acústica a fim de comparar diferentes abordagens. Tal comparação fornecerá um referencial objetivo sobre a efetividade do AST no contexto de monitoramento de colmeias, fundamentando discussões sobre melhorias futuras e extensões deste estudo.

3 Trabalhos relacionados

A busca por métodos não invasivos para o monitoramento de colmeias tem sido objeto de estudo em diversos trabalhos, sendo a análise e o processamento de áudio utilizados em aplicações diversas: contagem da saída e chegada das abelhas na colmeia (HEISE et al., 2020), detecção da abelha rainha (RUVINGA et al., 2021), enxameação (ZGANK, 2021), ritmo circadiano (KIM; OH; HEO, 2021), presença de poluentes no ar (SHARIF et al., 2020) e estimativa do período de maior atividade forrageira das abelhas (SHOSTAK; PRODEUS, 2019). Além disso, Abdollahi, Giovenazzo e Falk (2022) identificaram cerca de 60 trabalhos que também exploraram o processamento de áudio para o monitoramento de colmeias. Esses estudos evidenciam a versatilidade do processamento de áudio e sua capacidade de fornecer informações valiosas sobre o comportamento e a saúde das colônias de abelhas, destacando a relevância de desenvolver métodos cada vez mais precisos e eficientes.

O estudo conduzido por Shostak e Prodeus (2019) identificou o período ideal para a colheita de mel. Este é um aspecto crucial na apicultura, pois a colheita prematura ou tardia pode afetar a qualidade do mel e a saúde da colmeia. Para realizar essa tarefa complexa, os pesquisadores utilizaram uma abordagem baseada na análise acústica. Eles extraíram a densidade espectral dos áudios das colmeias, uma característica que fornece informações valiosas sobre a intensidade e a distribuição das frequências presentes no sinal de áudio. Esta informação foi então usada para estimar se a colmeia estava pronta para a coleta do mel. A prontidão para a colheita é um indicador do estado de saúde e produtividade da colmeia, e sua determinação precisa pode ajudar a maximizar a produção de mel. A validação do modelo foi realizada por meio da probabilidade de classificação correta, uma métrica que mede a precisão do modelo em prever a classe correta. O modelo alcançou um resultado de 96%.

Durante as investigações realizadas por Bromenshenk et al. (2009), foi observado que o som emitido pelas abelhas não só permite detectar a presença de poluentes, como também, permite identificar o tipo de poluente presente. Pensando nisso Sharif et al. (2020) analisaram o desempenho de três diferentes conjuntos de características para realizar a detecção de poluentes dentro da colmeia, especificamente a detecção do composto orgânico *Triclorometano* ($CHCl_3$). Os conjuntos de características utilizados pelos autores foram *MFCCs*, do inglês *Mel Frequency Cepstral Coefficients* e índices de paisagem sonora. As características extraídas foram enviadas para o classificador *Random Forest* que alcançou 91,66% de acurácia com os índices de paisagem sonora e 80% de acurácia com os *MFCCs*. No entanto, os autores ressaltaram o fato de se tratar de um cenário de classificação binária, indicando a necessidade de mais pesquisas para validar esses achados em contextos mais complexos.

Um estudo relevante nesse contexto é o de [Zgank \(2021\)](#), que explora a classificação da atividade de enxames de abelhas a partir de sinais de áudio utilizando redes neurais profundas (DNN). Este trabalho propõe uma técnica eficiente de modelagem acústica para sons de abelhas, seja não comprimidos ou comprimidos de forma lossy, abordando também a transmissão de dados de áudio em sistemas de IoT. Os resultados demonstraram uma precisão de 94,09% na classificação da atividade dos enxames, embora a compressão de áudio MP3 lossy tenha causado uma diminuição na precisão de até 11,51%, dependendo da taxa de bits e dos modelos utilizados. Este estudo destaca a importância de considerar os impactos da compressão de áudio em aplicações práticas de monitoramento de colmeias, contribuindo significativamente para o avanço das técnicas de monitoramento acústico de abelhas.

O artigo [Libal e Biernacki \(2024\)](#) também se insere nesse contexto, explorando a utilização de técnicas de aprendizado de máquina para classificar automaticamente abelhas com base em sinais de áudio. Este estudo destaca a importância da seleção de características MFCC pelo método LASSO, demonstrando que a análise acústica pode rastrear a frequência de abelhas operárias e zangões que entram e saem de uma colmeia. A implementação desse sistema de classificação automática pode auxiliar os apicultores na prevenção de enxames e na gestão mais eficiente de suas colônias, com resultados que incluem uma precisão de classificação ligeiramente acima de 95%. A seleção de características pelo LASSO resultou em um desempenho de detecção aprimorado, especialmente na detecção de zangões, comprovando a eficácia dessa abordagem na prática apícola. Além disso, o estudo ressalta a possibilidade de integração dessa tecnologia com sistemas de monitoramento IoT, potencializando ainda mais sua aplicação no campo.

A Tabela 1 apresenta uma visão geral dos trabalhos relacionados que foram analisados neste estudo. Cada trabalho é descrito em termos da aplicação específica e das técnicas de predição utilizadas. O estudo conduzido por ([SHOSTAK; PRODEUS, 2019](#)) identificou o período ideal para a colheita de mel, utilizando uma abordagem baseada na análise acústica. ([SHARIF et al., 2020](#)) concentrou-se na detecção de poluentes dentro da colmeia, utilizando conjuntos de características como MFCCs e índices de paisagem sonora, e o classificador *Random Forest*. No estudo de ([ZGANK, 2021](#)), a classificação da atividade de enxames de abelhas a partir de sinais de áudio foi explorada, empregando redes neurais profundas (DNN) e considerando a compressão de áudio MP3 lossy. O artigo de ([LIBAL; BIERNACKI, 2024](#)) investigou a classificação automática de abelhas com base em sinais de áudio, destacando a importância da seleção de características MFCC pelo método LASSO.

Apesar dos avanços observados nos trabalhos relacionados, muitos estudos focam em tarefas específicas como identificação da rainha, detecção de poluentes ou atividade forrageira, sem abordar diretamente a classificação da intensidade geral da colmeia. Além disso, diversas abordagens utilizam técnicas tradicionais de extração de características,

Tabela 1 – Trabalhos relacionados.

Trabalho	Aplicação	Técnicas de predição utilizadas	Resultados
(SHOSTAK; PRODEUS, 2019)	Identificação do período ideal para colheita de mel	Análise acústica da densidade espectral	Precisão de classificação de 96%
(SHARIF et al., 2020)	Detecção de poluentes ($CHCl_3$)	MFCCs e índices de paisagem sonora Classificador <i>Random Forest</i>	Acurácia de 91,66% (paisagem sonora) Acurácia de 80% (MFCCs)
(ZGANK, 2021)	Classificação da atividade de enxames de abelhas	Redes neurais profundas (DNN); análise de impactos de compressão MP3 lossy	Precisão de 94,09% (sem compressão) Queda de até 11,51% com MP3
(LIBAL; BIER-NACKI, 2024)	Classificação automática de abelhas (operárias × zangões)	Seleção de características MFCC via LASSO	Precisão ligeiramente acima de 95%

como MFCCs, associadas a classificadores convencionais, o que pode limitar a capacidade de capturar padrões mais complexos nos sinais acústicos. Por fim, destaca-se que, até o momento, não foram encontrados estudos que empreguem o modelo *Audio Spectrogram Transformer* (AST) para a classificação de sons de abelhas. Essas lacunas motivaram o desenvolvimento da presente proposta, que visa aplicar uma abordagem mais robusta e generalizável para a classificação da intensidade sonora das colmeias, uma abordagem baseada no modelo AST, que permite o aprendizado automático de representações espectrais a partir de espectrogramas log-Mel padronizados, visando uma classificação mais robusta e precisa da intensidade sonora das colmeias.

4 Metodologia Proposta

Nesta Seção, descrevemos a metodologia utilizada para classificar a intensidade sonora das colmeias. Os estágios envolvidos neste processo estão ilustrados na Figura 1. O método proposto foi dividido em quatro etapas principais: 1) Aquisição dos Dados; 2) Pré-processamento; 3) Extração de Características; e 4) Classificação.

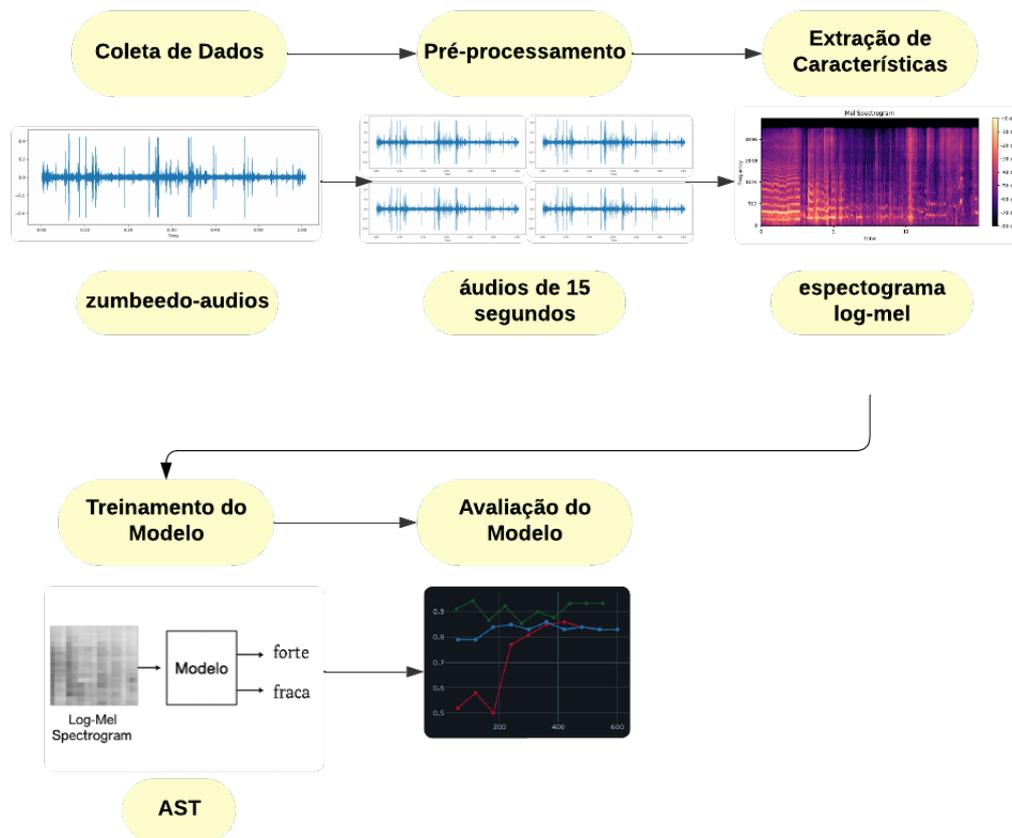


Figura 1 – Fluxograma representativo do método proposto que está dividido em 4 etapas principais

4.1 Aquisição dos dados

O estudo teve início com um arquivo de anotações que listava 126 gravações em formato *.wav, como mostra a Figura 2. Cada linha desse arquivo registra o nome do áudio e o rótulo *Intensidade*, que pode assumir dois valores “forte” e “fraca”. Depois de obtidos, os arquivos foram organizados localmente no diretório; diretórios criados em execuções anteriores foram removidos de antemão de modo a evitar conflitos. Durante o processo, possíveis duplicatas ou ausências foram registradas, garantindo a integridade entre os áudios baixados e seus metadados. Ao término dessa etapa, todos os áudios anotados haviam

tido transferidos com sucesso, assegurando correspondência completa entre gravações e rótulos.

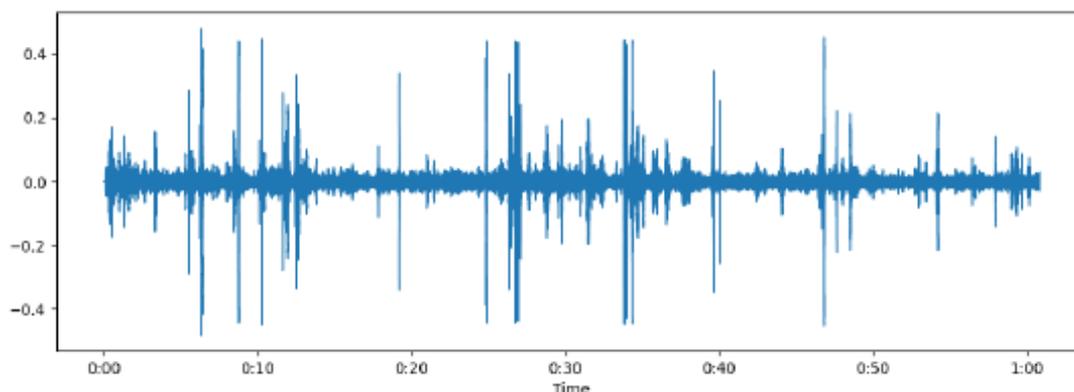


Figura 2 – Gráfico de forma de onda de uma amostra de áudio.

4.2 Pré-processamento

As gravações originais possuem taxa de amostragem de 11 kHz, profundidade de 16 bits e amplitudes variando de $-32\,768$ a $+32\,767$. Para padronizar a duração das entradas, cada arquivo foi segmentado em uma única janela de 15 s (165 000 amostras), mantendo-se a frequência de 11 kHz. Trechos residuais menores que 15 s foram descartados. Ao final, obtiveram-se 498 segmentos – um para cada gravação original –, quantidade considerada suficiente para treinar o modelo.

O conjunto final de segmentos de áudio foi dividido de maneira estratificada em dois subconjuntos: 80 % para treinamento e 20 % para teste (VABALAS et al., 2019). Essa estratégia garantiu que as proporções das duas classes fossem preservadas em ambas as partições, minimizando o viés de amostragem. Com a divisão realizada, os rótulos categóricos foram convertidos para inteiros 0 e 1, viabilizando o uso direto da perda de entropia cruzada. No conjunto de treinamento, com 398 amostras, a distribuição dos rótulos foi equilibrada, com 199 amostras para cada classe (0 e 1). Já no conjunto de teste, composto por 100 amostras, a contagem de rótulos também foi equilibrada, com 50 amostras para cada classe. Essa divisão cuidadosa assegura que o modelo seja avaliado de maneira justa, com uma representação equilibrada de cada classe tanto no treinamento quanto no teste.

4.3 Extração de características

Com os segmentos prontos, aplicou-se o *pipeline* de extração de atributos do Audio Spectrogram Transformer (AST) pré-treinado MIT/ast-finetuned-audioset-10-10-0.4593 da Hugging Face. Primeiramente, cada sinal foi remapeado para 16 kHz, frequência exigida

pelo modelo (GONG; CHUNG; GLASS, 2021). Em seguida, o `AutoFeatureExtractor` converteu cada amostra em um espectrograma *log-mel* como mostra a Figura 3. Para assegurar dimensões uniformes, todos os espectrogramas foram *padronizados* para 10 s por meio de preenchimento ou truncamento. Por fim, os tensores resultantes foram serializados, possibilitando reaproveitamento em experimentos futuros.

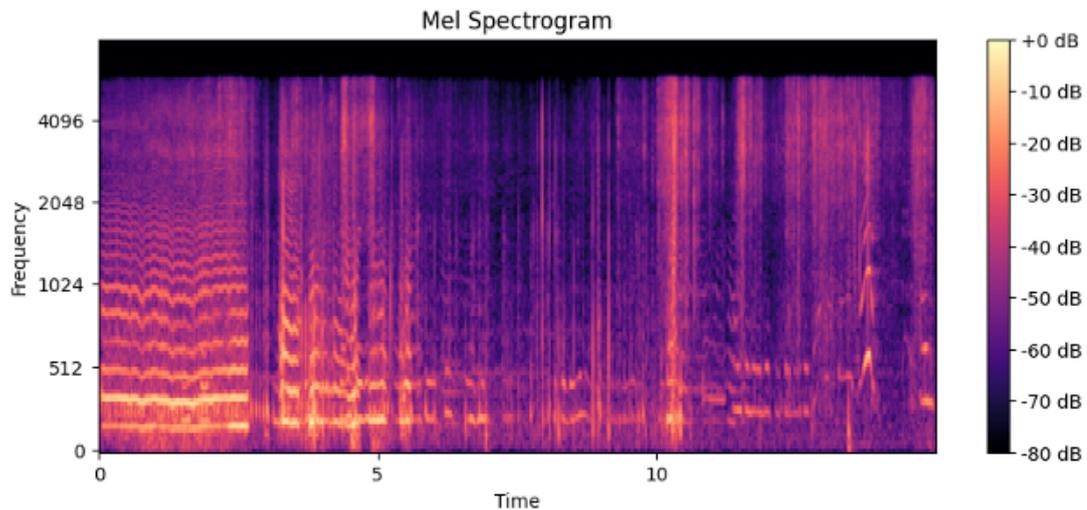


Figura 3 – Espectrograma log-Mel gerado a partir de um segmento de áudio da colmeia.

4.4 Classificação

O modelo foi treinado utilizando o *Audio Spectrogram Transformer* (AST), com experimentos realizados para avaliar o impacto de diferentes configurações de hiperparâmetros. Durante o treinamento, os parâmetros fundamentais foram ajustados para otimizar a performance do modelo, incluindo o número de épocas de treinamento fixado em 50, tamanho de lote de 4 amostras por dispositivo, passos de acumulação de gradiente variando entre 4 e 8, taxas de aprendizado testadas em 0.001 e 0.01, e proporções de warmup entre 0.0 e 0.5, valores que se mostraram suficiente com base nos testes.

O desempenho foi avaliado através de múltiplas métricas, sendo a acurácia nos conjuntos de treino e validação o principal indicador de qualidade do modelo. A função de perda durante o treinamento (train loss) e a perda de avaliação (eval loss) foram monitoradas para detectar possíveis casos de overfitting. Os treinamentos foram salvos periodicamente, permitindo a avaliação comparativa tanto dos parâmetros configurados quanto das métricas coletadas.

5 Resultados

Foram conduzidos onze processos independentes de ajuste fino do *Audio Spectrogram Transformer* (AST). A partir dessas execuções, escolheram-se os **cinco checkpoints** com melhor compromisso entre acurácia e *Eval Loss*, de modo a representar casos de alto, médio e intermediário desempenho sem sobrecarregar a discussão com evidências redundantes. Todos os modelos foram treinados durante cinquenta épocas, número que se mostrou suficiente com base em resultados e curvas de desempenho observadas em treinamentos anteriores, mantendo-se fixos o lote efetivo de trinta e duas amostras; variaram, contudo, o percentual de aquecimento inicial (*warm-up*) e a periodicidade das avaliações.

Os números resumidos na Tabela 2 revelam clara superioridade do modelo **awesome-perch**, o qual atingiu 94,4% de acurácia, aliada ao menor *Eval Loss* (0,119). Tal desempenho justificou sua adoção como *modelo otimizado* para as análises qualitativas subsequentes. O segundo colocado, **debonair-fish** (87% de acerto), não conseguiu ultrapassar a configuração de referência. Já **dapper-horse** e **abundant-grouse**, ambos com 85% de acerto. Por fim, **intelligent-doe**, com 81%, serve de *baseline* interno para contrastar as demais abordagens. Os resultados obtidos com o modelo corroboram as discussões da literatura (FERREIRA et al., 2025), que destacam a importância do pré-treinamento e ajuste de hiperparâmetros para alcançar melhorias significativas no reconhecimento acústico de espécies de abelhas.

Tabela 2 – Desempenho dos cinco checkpoints comparados.

Run (Nome)	Ac (%)	Eval Loss	Train Loss
awesome-perch	94.4	0.119	0.294
debonair-fish	87.0	0.389	0.449
dapper-horse	85.0	0.395	0.472
abundant-grouse	85.0	0.417	0.463
intelligent-doe	81.0	0.434	0.374

A análise dos hiperparâmetros, apresentada na Tabela 3, confirma que todos os modelos de topo partilharam taxa de aprendizado $LR = 0,001$, ao passo que execuções descartadas com valores dez vezes maiores sofreram perda de desempenho acentuada que corrobora a hipótese de que taxas de aprendizado moderadas associadas a *warm-up* prolongado evitam o colapso entrópico da atenção (ZHAI et al., 2023). Observa-se, ademais, que proporções de *warm-up* de 50% (*awesome-perch* e *abundant-grouse*) amorteceram as oscilações iniciais do gradiente e reduziram a perda de validação em até 24% relativamente às execuções sem aquecimento. Frequências de avaliação entre 55 e 60 passos mostraram-se suficientes para detectar o ponto ótimo de generalização, enquanto intervalos mais longos não agregaram benefícios mensuráveis.

Tabela 3 – Hiperparâmetros detalhados dos cinco modelos selecionados.

Run (Nome)	Eval Steps	Warm-up	LR	Scheduler
awesome-perch	55	0.50	0.001	cosine
debonair-fish	60	0.00	0.001	cosine
dapper-horse	120	0.00	0.001	cosine
abundant-grouse	60	0.50	0.001	cosine
intelligent-doe	60	0.00	0.001	cosine

Em conjunto, esses achados indicam que a combinação de $LR = 0,001$, *warm-up* de 0,5, agendamento cosseno e avaliações frequentes constitui o melhor ponto de equilíbrio entre desempenho e robustez quando se trabalha com corpora acústicos de colmeias. Tal configuração passa a servir de referência para extensões futuras do estudo, como a incorporação de *data augmentation* espectral, a ampliação da diversidade de gravações e a experimentação de funções de perda resistentes a desbalanceamento.

A Figura 4 descreve a trajetória da acurácia de validação do run **awesome-perch**. Logo no início do ajuste fino observa-se um crescimento rápido, culminando em 94,4% no passo 110. Esse pico não volta a ser superado nas iterações seguintes, embora a curva apresente oscilações que sugerem pequenas recuperações de desempenho entre os passos 320 e 520. Portanto, o passo 110 foi selecionado como *checkpoint ótimo*, pois representa o ponto em que o modelo atinge sua melhor capacidade de generalização antes que novas atualizações de pesos comecem a introduzir ruído.

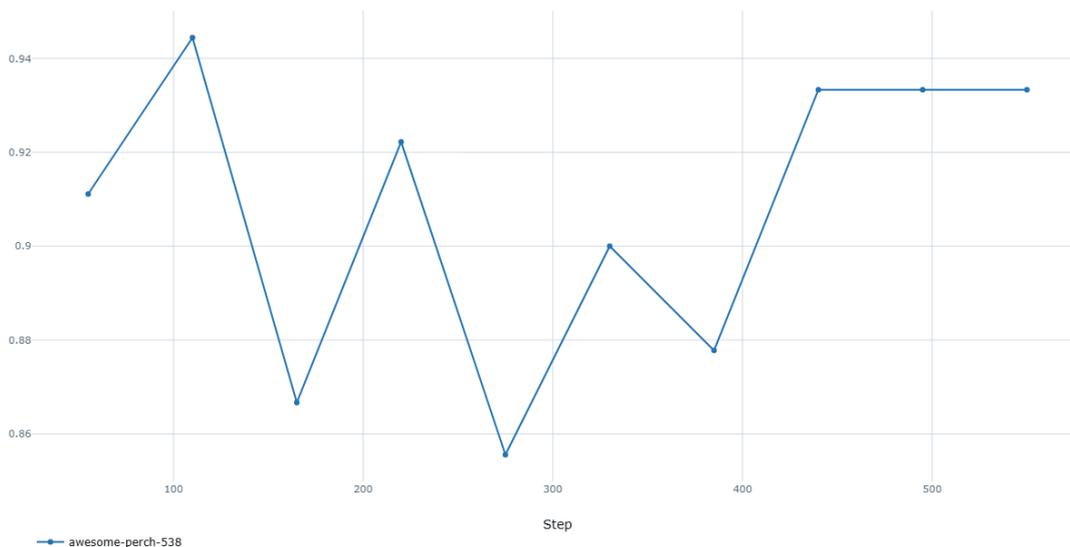


Figura 4 – Evolução da *eval_accuracy* do run **awesome-perch** ao longo dos passos de treinamento. Observa-se que o maior valor (94,4%) ocorre no passo 110, sendo este o checkpoint adotado como modelo otimizado.

O comportamento da *eval_loss*, mostrado na Figura 5, reforça essa escolha. O valor mínimo (0.119) ocorre exatamente no mesmo passo em que a acurácia atinge o seu máximo, indicando que o modelo entrega simultaneamente maior precisão e menor erro na

validação. A partir desse ponto, a *eval_loss* cresce progressivamente — mesmo quando a acurácia oscila para cima —, sinalizando que novas iterações trazem ganhos apenas aparentes, fruto de flutuações estatísticas que não se traduzem em redução do erro médio.

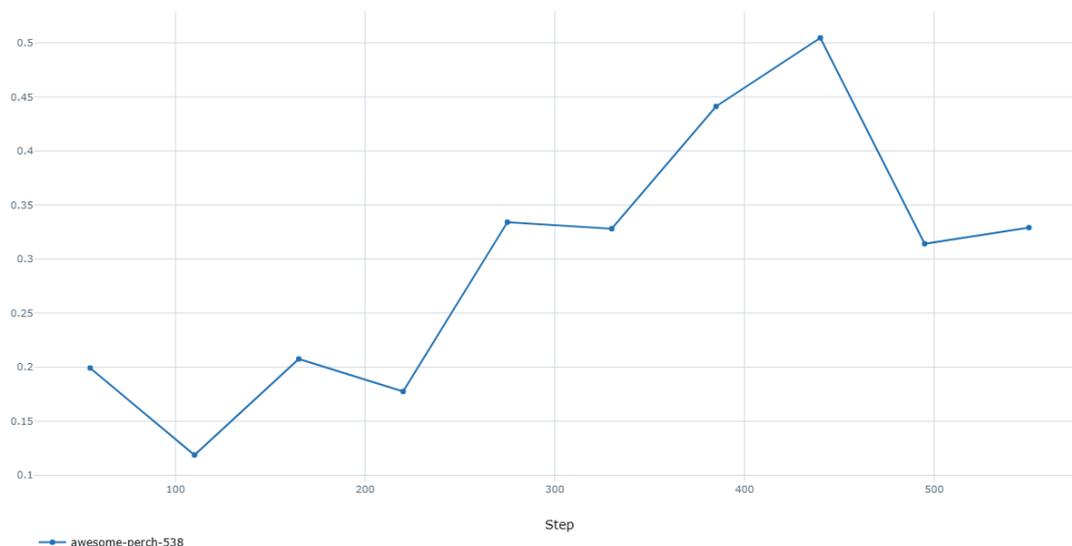


Figura 5 – Comportamento da *eval_loss*. O menor valor (0,119) coincide com o pico de acurácia na Figura 4, indicando bom equilíbrio entre erro e capacidade de generalização.

Já a Figura 6 exibe a queda constante da *train_loss*. Entre os passos 380 e 520 a curva praticamente encosta em zero, compatível com o cenário em que o modelo memoriza o conjunto de treinamento. No entanto, como a *eval_loss* não acompanha essa redução depois do passo 110, fica caracterizado que o benefício adicional em treinamento não se converte em melhorias reais sobre dados não vistos. Assim, o checkpoint selecionado preserva o balanço ideal entre erro de treinamento e erro de validação, evitando o sobreajuste.

A sobreposição dos três gráficos evidencia que a combinação $LR = 0,001$ com 50 de aquecimento inicial (*warm-up*) foi capaz de posicionar o AST em sua região ótima de convergência já na centésima etapa. Desse modo, a escolha do checkpoint 110 não apenas maximiza a acurácia, mas também confere ao modelo estabilidade e robustez para servir de referência às análises qualitativas futuras.

Apesar do desempenho elevado obtido com o *Audio Spectrogram Transformer* (AST), algumas limitações do método devem ser reconhecidas. O modelo exige entradas com comprimento fixo (10 segundos), o que obriga truncamento ou preenchimento de gravações, podendo comprometer a integridade temporal, sobretudo quando os eventos de interesse não estão centralizados. Além disso, embora pré-treinado com um grande conjunto de dados (AudioSet), o AST tende a demandar volumes substanciais de dados para generalizar bem. Como o corpus aqui utilizado é menor, há risco de sobreajuste em execuções prolongadas, especialmente sem *warm-up* ou regularização espectral adequada.

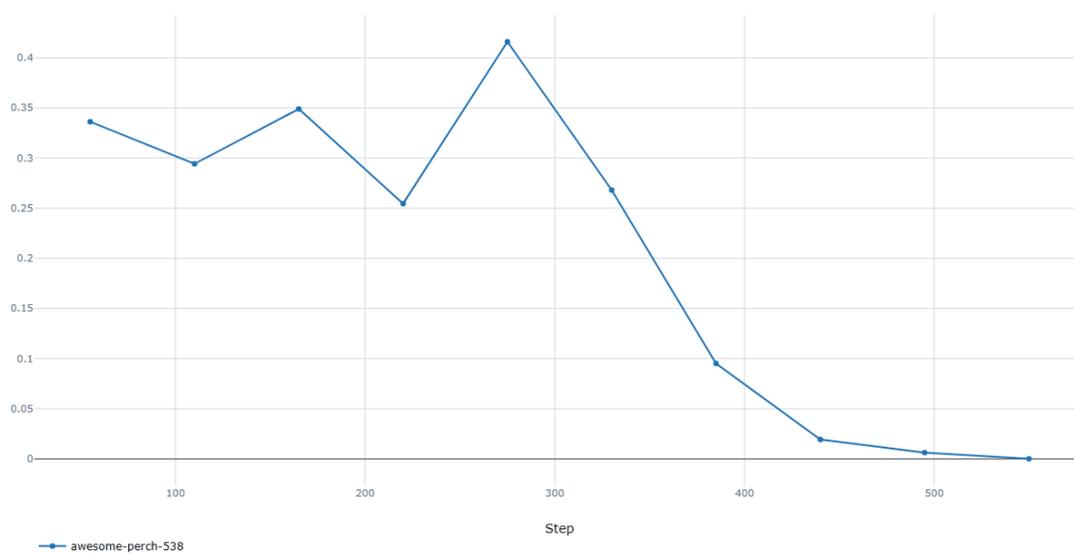


Figura 6 – Curva de *train_loss*. A queda pronunciada após o passo 380 confirma a convergência do modelo sem indícios de sobreajuste até o checkpoint escolhido.

6 Conclusão

Neste trabalho apresentamos uma abordagem de monitoramento não invasivo da intensidade das colmeias baseada no processamento acústico de gravações e no modelo *Audio Spectrogram Transformer* (AST). Concluiu-se que a conversão dos áudios em espectrogramas, aliada ao fine-tuning do AST, melhora significativamente a precisão na discriminação entre colmeias fortes e fracas, alcançando acurácia superior a 94 %. A principal vantagem de empregar o AST reside na capacidade do modelo em aprender automaticamente padrões espectrais relevantes, eliminando a necessidade de etapas manuais de extração de atributos. Como limitação, destaca-se o tamanho ainda reduzido do conjunto de gravações e a influência de ruídos ambientais, fatores que podem restringir a generalização dos resultados para apiários com condições acústicas distintas.

Como trabalhos futuros, propõe-se a ampliação da base de dados com novas gravações coletadas ao longo de diferentes estações do ano, em condições climáticas e acústicas variadas, a fim de aumentar a diversidade e representatividade do conjunto de treinamento. Pretende-se também investigar estratégias de *data augmentation* específicas para sinais acústicos de colmeias, como perturbações de frequência, ruído de fundo controlado e estiramento temporal, com o objetivo de melhorar a robustez do modelo frente a variações naturais do ambiente. Por fim, planeja-se realizar a implantação de um protótipo mobile funcional de monitoramento em tempo real, visando a aplicação prática da solução em campo.

Referências

- ABDOLLAHI, M.; GIOVENAZZO, P.; FALK, T. H. Automated beehive acoustics monitoring: A comprehensive review of the literature and recommendations for future work. *Applied Sciences*, MDPI, v. 12, n. 8, p. 3920, 2022. Citado na página 19.
- BRAGA, A. R. et al. A method for mining combined data from in-hive sensors, weather and apiary inspections to forecast the health status of honey bee colonies. *Computers and Electronics in Agriculture*, Elsevier, v. 169, p. 105161, 2020. Citado na página 16.
- BROMENSHENK, J. J. et al. *Honey bee acoustic recording and analysis system for monitoring hive health*. [S.l.]: Google Patents, 2009. US Patent 7,549,907. Citado na página 19.
- COSTA, R. O. et al. Análise hierárquica dos problemas existentes na produção de mel do estado da paraíba. *Revista Verde de Agroecologia e Desenvolvimento Sustentável*, Grupo Verde de Agroecologia e Abelhas, v. 11, n. 2, p. 24–28, 2016. Citado na página 13.
- FERREIRA, A. I. S. et al. Transformer models improve the acoustic recognition of buzz-pollinating bee species. *Ecological Informatics*, Elsevier, p. 103010, 2025. Citado na página 25.
- GONG, Y.; CHUNG, Y.-A.; GLASS, J. Ast: Audio spectrogram transformer. *arXiv preprint arXiv:2104.01778*, 2021. Citado 3 vezes nas páginas 17, 18 e 24.
- GOODFELLOW, I. et al. *Deep learning*. [S.l.]: MIT press Cambridge, 2016. Citado na página 18.
- GORROI, G.; FREITAS, L. P. V. d.; ASSIS, D. C. S. d. Apicultura: o manejo das abelhas do gênero apis. *Cad. técn. Vet. Zoot.*, p. 9–36, 2020. Citado na página 13.
- HEISE, D. et al. Bumble bee traffic monitoring using acoustics. In: IEEE. *2020 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*. [S.l.], 2020. p. 1–6. Citado na página 19.
- KIM, J.; OH, J.; HEO, T.-Y. Acoustic scene classification and visualization of beehive sounds using machine learning algorithms and grad-cam. *Mathematical Problems in Engineering*, Hindawi Limited, v. 2021, p. 1–13, 2021. Citado 2 vezes nas páginas 13 e 19.
- LIAKOS, K. G. et al. Machine learning in agriculture: A review. *Sensors*, MDPI, v. 18, n. 8, p. 2674, 2018. Citado na página 13.
- LIBAL, U.; BIERNACKI, P. Mfcc selection by lasso for honey bee classification. *Applied Sciences*, MDPI, v. 14, n. 2, p. 913, 2024. Citado 2 vezes nas páginas 20 e 21.
- NOLASCO, I. et al. Audio-based identification of beehive states. In: IEEE. *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. [S.l.], 2019. p. 8256–8260. Citado na página 16.
- QANDOUR, A. et al. Remote beehive monitoring using acoustic signals. Australian Acoustical Society, 2014. Citado na página 15.

- RUSTAM, F. et al. Bee detection in bee hives using selective features from acoustic data. *Multimedia Tools and Applications*, Springer, v. 83, n. 8, p. 23269–23296, 2024. Citado na página 16.
- RUVINGA, S. et al. Use of lstm networks to identify “queenlessness” in honeybee hives from audio signals. In: IEEE. *2021 17th International Conference on Intelligent Environments (IE)*. [S.l.], 2021. p. 1–4. Citado na página 19.
- SAMMUT, C.; WEBB, G. I. *Encyclopedia of machine learning*. [S.l.]: Springer Science & Business Media, 2011. Citado na página 18.
- SHARIF, M. Z.; DI, N.; YU, B. Honeybee (apis spp.)(hymenoptera: Apidae) colony monitoring using acoustic signals from the beehive: An assessment by global experts and our feedback. *Agriculture*, MDPI, v. 13, n. 4, p. 769, 2023. Citado na página 15.
- SHARIF, M. Z. et al. Soundscape indices: new features for classifying beehive audio samples. *Sociobiology*, v. 67, n. 4, p. 566–571, 2020. Citado 4 vezes nas páginas 15, 19, 20 e 21.
- SHOSTAK, S.; PRODEUS, A. Classification of the bee colony condition using spectral features. In: IEEE. *2019 IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PIC S&T)*. [S.l.], 2019. p. 737–740. Citado 4 vezes nas páginas 13, 19, 20 e 21.
- SOARES, B. S. et al. Mfcc-based descriptor for bee queen presence detection. *Expert Systems with Applications*, Elsevier, v. 201, p. 117104, 2022. Citado na página 16.
- TERENZI, A.; CECCHI, S.; SPINSANTE, S. On the importance of the sound emitted by honey bee hives. *Veterinary Sciences*, MDPI, v. 7, n. 4, p. 168, 2020. Citado na página 15.
- UTHOFF, C.; HOMSI, M. N.; BERGEN, M. V. Acoustic and vibration monitoring of honeybee colonies for beekeeping-relevant aspects of presence of queen bee and swarming. *Computers and electronics in agriculture*, Elsevier, v. 205, p. 107589, 2023. Citado na página 16.
- VABALAS, A. et al. Machine learning algorithm validation with a limited sample size. *PloS one*, Public Library of Science San Francisco, CA USA, v. 14, n. 11, p. e0224365, 2019. Citado na página 23.
- VERMA, P.; BERGER, J. Audio transformers: Transformer architectures for large scale audio understanding. adieu convolutions. *arXiv preprint arXiv:2105.00335*, 2021. Citado na página 17.
- VIEIRA, F. R.; ANDRADE, D. C.; RIBEIRO, F. L. A polinização por abelhas sob a perspectiva da abordagem de serviços ecossistêmicos (ase). *Revista Ibero-Americana de Ciências Ambientais*, v. 12, n. 4, p. 544–560, 2021. Citado na página 13.
- ZGANK, A. Bee swarm activity acoustic classification for an iot-based farm service. *Sensors*, MDPI, v. 20, n. 1, p. 21, 2019. Citado na página 16.
- ZGANK, A. Iot-based bee swarm activity acoustic classification using deep neural networks. *Sensors*, MDPI, v. 21, n. 3, p. 676, 2021. Citado 3 vezes nas páginas 19, 20 e 21.

ZHAI, S. et al. Stabilizing transformer training by preventing attention entropy collapse. In: PMLR. *International Conference on Machine Learning*. [S.l.], 2023. p. 40770–40803. Citado na página 25.

ZHANG, Y. et al. Spectrogram transformers for audio classification. In: IEEE. *2022 IEEE International Conference on Imaging Systems and Techniques (IST)*. [S.l.], 2022. p. 1–6. Citado na página 17.



TERMO DE AUTORIZAÇÃO PARA PUBLICAÇÃO DIGITAL NA BIBLIOTECA “JOSÉ ALBANO DE MACEDO”

Identificação do Tipo de Documento

- () Tese
() Dissertação
(x) Monografia
() Artigo

Eu, **Marcos Antônio Gonçalves Barros Brito**, autorizo com base na Lei Federal nº 9.610 de 19 de Fevereiro de 1998 e na Lei nº 10.973 de 02 de dezembro de 2004, a biblioteca da Universidade Federal do Piauí a divulgar, gratuitamente, sem ressarcimento de direitos autorais, o texto integral da publicação **Análise da Intensidade das Colmeias Utilizando Processamento de Áudio e Aprendizado de Máquina** de minha autoria, em formato PDF, para fins de leitura e/ou impressão, pela internet a título de divulgação da produção científica gerada pela Universidade.

Picos-PI __08__ de _____ Julho _____ de 2025__.



Documento assinado digitalmente
MARCOS ANTONIO GONCALVES BARROS BRITO
Data: 08/07/2025 20:31:03-0300
Verifique em <https://validar.iti.gov.br>

Assinatura