

Inovações na Detecção de Ruídos Antropogênicos com Aprendizado de Máquina

Felipe R. Ahad¹, Josiel M. Figueiredo¹, Alvaro S. C. Junior¹, Allan G. Oliveira¹

¹Instituto de Computação, Universidade Federal de Mato Grosso, Brasil

{felipe.ahad, alvaro.junior1}@sou.ufmt.br, {josiel, allan}@ic.ufmt.br

Abstract. *The anthropogenic noise is one of the main challenges for wildlife conservation, masking vital signals and altering ecosystems. Passive acoustic monitoring, powered by machine learning, has become an essential tool for studying these impacts. This paper analyzes the evolution of anthropogenic noise detection techniques from the perspective of Joseph Schumpeter's theory of innovation. We argue that the transition from classical algorithms to convolutional neural nets represents a process of "creative destruction", where a new technology not only improves upon but replaces the previous one, redefining the frontiers of bioacoustics research.*

Resumo. *O ruído antropogênico é um dos principais desafios para a conservação da vida selvagem, mascarando sinais vitais e alterando ecossistemas. O monitoramento acústico passivo, impulsionado pelo aprendizado de máquina, tornou-se uma ferramenta essencial para estudar esses impactos. Este artigo analisa a evolução das técnicas de detecção de ruído antropogênico utilizando a teoria da inovação de Joseph Schumpeter. Argumentamos que a transição de algoritmos clássicos para redes neurais convolucionais representa um processo de "destruição criadora", onde uma nova tecnologia não apenas melhora, mas substitui a anterior, redefinindo as fronteiras da pesquisa em bioacústica.*

1. Introdução

O Monitoramento Acústico Passivo (MAP) é considerado uma ferramenta importante para estudar os efeitos da ação humana sobre determinados ambientes. Através de unidades de gravação autônomas instaladas em determinados locais, coleta-se sons por meses, produzindo uma quantidade vultosa de dados, tornando inviável a análise manual e criando um grande obstáculo para pesquisadores. Neste sentido, o aprendizado de máquina surge como um grande aliado. Quando treinados, esses algoritmos podem reconhecer padrões complexos nos dados, o que possibilita a automação das tarefas de identificação e classificação de ruídos de interesse, como o barulho de barcos, drones, motosserras, explosivos, disparos de arma de fogo, etc.

Este artigo tem como objetivo analisar as tecnologias sob a lente da teoria da inovação de Joseph Schumpeter, demonstrando que a mudança dos métodos manuais para soluções de aprendizado de máquina não é uma simples melhoria, mas um processo de "destruição criadora" que está mudando os métodos de análise, as perguntas de pesquisa e as aplicações práticas na área de ecologia acústica.

Portanto, o trabalho está organizado da seguinte forma: a Seção 2 apresenta a base teórica sobre ruídos humanos e o uso de técnicas de aprendizado de máquina aplicadas à análise de áudio. A Seção 3 descreve o método de pesquisa e as perguntas que norteiam este trabalho. A Seção 4 introduz as ideias centrais da teoria da inovação. A Seção 5 utiliza essa teoria para analisar as inovações de processo e produto, na detecção de ruídos antropogênicos, além das considerações sobre propriedade intelectual. A Seção 6 apresenta as respostas às questões que norteiam a pesquisa. A seção 7 aborda os possíveis impactos à validade do estudo. Por fim, a Seção 8 traz as considerações finais e trabalhos futuros.

2. Fundamentação Teórica

O ruído antropogênico advém de diversas fontes, como veículos automotores, trens, aeronaves, entre outros [Warren et al. 2006]. Seu impacto na vida selvagem é relevante, sobretudo pelo mascaramento acústico, que ocorre quando o ruído humano sobrepõe as frequências que os animais usam para se comunicar. Espécies de pássaros, anfíbios e mamíferos, podem precisar vocalizar mais alto ou em tons diferentes para competir com outras fontes [Higham et al. 2021; Dahlheim and Castellote 2016; Slabbekoorn et al. 2003]. Esses ruídos podem interferir em funções biológicas importantes de animais, podendo levá-los à distração, além de gerar custos fisiológicos, como a redução do sono e o aumento dos níveis de hormônios do estresse [Kok et al. 2023].

Os sons captados são submetidos à análise por computador, que, geralmente, inicia-se pela transformação do áudio em uma imagem, chamada de espectrograma. O espectrograma representa a energia do som em diferentes frequências ao longo do tempo [Lampert and O'Keefe 2010], e, a partir desta representação, modelos de aprendizado de máquina ou de aprendizado profundo podem ser treinados para reconhecer padrões existentes nesses espectrogramas.

Em trabalhos pioneiros no campo da bioacústica computacional, observa-se a utilização de algoritmos clássicos como *k-Nearest Neighbors* (kNN), *Support Vector Machine* (SVM), e *Random Forest* [Huang et. al. 2009; Stowell et al. 2014]. Todavia, essas técnicas demandam que um especialista defina e extraia, manualmente, as características do som, como no caso do *Mel-Frequency Cepstral Coefficients* (MFCC), que, por sua vez, apresenta um desempenho inferior em relação à outras alternativas, como o *Mel Spectrogram*, por ser considerado inadequado para arquiteturas de redes neurais convolucionais (CNNs) [Stowell 2022].

Paralelamente, outros estudos empregavam com sucesso os Modelos Ocultos de Markov (HMMs) para analisar dados bioacústicos: um modelo estatístico que, nesse campo, demonstrou ser capaz de lidar com variações nas vocalizações de espécies sob diversos níveis de ruído [Ren et al. 2009]. No entanto, a aplicação desses modelos era limitada por implementações de software que, além de custosas, eram tecnologicamente complexas para não especialistas envolvidos na conservação da vida selvagem. Buscando transpor essa barreira, Ranjard et al. [2017] desenvolveram o MatlabHTK, uma interface de software que simplifica a configuração e a utilização de HMMs, permitindo que pesquisadores sem conhecimento aprofundado na área pudessem implementar essa técnica para o processamento de sinais bioacústicos.

Por outro lado, em trabalhos mais recentes, tem se consolidado a utilização de ar-

quiteturas mais modernas de aprendizado profundo, em especial as CNNs [Stowell 2022]. Criadas originalmente para analisar imagens [LeCun et al. 1998], as CNNs funcionam muito bem com espectrogramas. Sua grande vantagem é que elas aprendem as características importantes sozinhas, eliminando a necessidade do trabalho manual [Stowell 2022]. Além disso, como demonstrado em estudos atuais sobre reconhecimento de fala, a implementação de mecanismos de atenção pode conferir às CNNs uma robustez ainda maior para lidar com ruídos no sinal acústico [Wijayasingha and Stankovic 2021].

3. Materiais e Métodos

A fim de sintetizar achados em tempo hábil, o presente estudo foi conduzido utilizando-se a metodologia *Rapid Review* (ou, Revisão Rápida da Literatura): um método em que o revisor adota uma estratégia de busca mais ampla e menos sofisticada, com uma avaliação de qualidade mais simples, limitando alguns estágios da revisão sistemática para encurtar o tempo deste processo [Grant and Booth 2009].

Para justificar a importância do tema, e, para mapear, a partir da teoria schumperiana, as inovações de técnicas utilizadas em pesquisas realizadas no campo da ecologia acústica, a análise foi norteada por três questões centrais, conforme Tabela 1.

	Questão	Justificativa
Q1	Quais estudos investigam os impactos dos ruídos gerados pela atividade humana sobre a vida selvagem?	Destacar a relevância do problema, estabelecendo o impacto do ruído antropogênico na vida selvagem e a necessidade de detectá-los.
Q2	Quais técnicas de aprendizado de máquina estavam sendo utilizadas em trabalhos desenvolvidos há cerca de quinze anos atrás na área ecologia acústica?	Mapear os estudos mais relevantes e apresenta-los cronologicamente, analisando sua eficácia na atualidade.
Q3	Quais técnicas de aprendizado de máquina estão sendo utilizadas em trabalhos desenvolvidos nos últimos cinco anos na área de ecologia acústica?	Apresentar o plano tecnológico atual que representa a "nova combinação" ou "destruição criadora".

Tabela 1: As questões que norteiam o estudo.

Das questões formuladas pelos autores, foi definido um conjunto contendo palavras-chave como "*anthropogenic sound*", "*anthropogenic noise*", "*effects*", "*wildlife*", "*nature*", "*bioacoustic*", "*machine learning*" e "*deep learning*", utilizando-se operadores booleanos (AND e OR) para estruturar as buscas.

As buscas foram conduzidas entre 03 de setembro de 2025 e 04 de outubro de 2025, e, para isso, optou-se pela utilização da plataforma Google Scholar considerando sua abrangência multidisciplinar e pela gratuidade no acesso, priorizando a seleção de: **i)** artigos com ênfase na detecção de sons biofônicos ou antropofônicos; **ii)** artigos publicados em inglês; **iii)** artigos publicados nos últimos cinco anos, exceto pela questão **Q2**, cujo objetivo é avaliar estudos clássicos neste campo.

Posteriormente, foram excluídos os artigos que: **a)** não apresentavam relação com o tema abordado; ou **b)** citavam o tema, sem que este fosse o objeto central do estudo. A aplicação dos critérios resultou na seleção dos trabalhos listados na Tabela 2.

Autor	Título
Aide et al. (2013)	Real-time bioacoustics monitoring and automated species identification
Dahlheim, M. & Castellote, M. (2016)	Changes in the acoustic behavior of gray whales <i>Eschrichtius robustus</i> in response to noise
Higham et al. (2021)	Traffic noise drives an immediate increase in call pitch in an urban frog
Huang et al. (2009)	Frog classification using machine learning techniques
Kok et al. (2023)	How chronic anthropogenic noise can affect wildlife communities
Ranjard et al. (2017)	Matlabhtk: a simple interface for bioacoustic analyses using hidden markov models
Ren et al. (2009)	A Framework for Bioacoustic Vocalization Analysis Using Hidden Markov Models
Slabbekoorn, H. & Peet, M. (2003)	Ecology: Birds sing at a higher pitch in urban noise
Stowell, D. (2022)	Computational bioacoustics with deep learning: a review and roadmap
Stowell et al. (2014)	Automatic large-scale classification of bird sounds is strongly improved by unsupervised feature learning

Tabela 2: Estudos selecionados.

Demais artigos utilizados para apoiar a escrita do presente estudo encontram-se na seção de referências bibliográficas.

4. A Inovação como Motor do Desenvolvimento Científico

Para analisar o impacto das técnicas de aprendizado de máquina no campo da ecologia acústica, utilizou-se a teoria da inovação de Joseph Schumpeter. Em sua obra ”A Teoria do Desenvolvimento Econômico” [1911], Schumpeter faz uma distinção clara entre invenção e inovação. Enquanto a invenção refere-se ao processo de criação de algo novo, a inovação ocorre quando há a aplicação bem-sucedida de uma invenção, isto é, quando um impacto real é gerado.

Para Schumpeter [1942], o desenvolvimento não é um processo lento e contínuo, mas que, na verdade, ocorre por meio de saltos, descrito por ele como ”destruição criadora”, isto é, quando uma nova tecnologia, produto ou método surge e torna o antigo obsoleto.

5. Inovações na Detecção de Ruídos Antropogênicos

5.1. A Automação da Escuta Ecológica como uma Inovação de Processo

A inovação, segundo Schumpeter [1911], não se restringe à criação de novos produtos, mas abrange, em sua essência, a transformação de processos produtivos. Historicamente,

os métodos tradicionais de identificação acústica dependiam da figura de um especialista capaz de ouvir, em campo, e classificar as espécies com base na sua vocalização. Entretanto, esse processo é limitado pela escassez de especialistas, e pode ser impactado pelo viés do observador, que advém da variação de expertise desses indivíduos no tocante à detecção de determinados grupos de espécies [Aide et al. 2013].

Embora os primeiros algoritmos de aprendizado de máquina representassem uma inovação ao permitir a automação da triagem dos dados, um grande salto veio com o uso de técnicas de aprendizado profundo, em especial pelo uso da arquitetura CNN. A capacidade de aprender as características do som diretamente dos espectrogramas [Stowell 2022] eliminou a necessidade de um especialista fazer esse trabalho manualmente. Isso não apenas melhorou a precisão dos resultados, mas também democratizou a técnica para pesquisadores de diferentes áreas.

Nesse sentido, a verdadeira inovação schumpeteriana não reside na invenção da arquitetura CNN, mas em sua aplicação como um novo processo que substitui a etapa crítica da engenharia manual de características. Ao fazer isso, a tecnologia não apenas otimiza, mas reestrutura fundamentalmente a prática da análise acústica, gerando um avanço em eficiência e acessibilidade.

5.2. Novos Sistemas e Ferramentas como uma Inovação de Produto

Como resultado da inovação de processos, foi possível alcançar o desenvolvimento de novos produtos para a ciência. Examinando trabalhos realizados por volta de quinze anos atrás, nota-se que a análise dos dados demandava um tempo muito maior para ser realizada, sendo esta uma consequência da produção massiva de dados bioacústicos devido à democratização do acesso aos dispositivos utilizados no monitoramento em larga escala, e na captura contínua de áudio [Stowell 2022].

Com esse "dilúvio de dados", a demanda pelo que Schumpeter [1934] chamaria de "novas combinações" tecnológicas e técnicas computacionais mais sofisticadas não era simples conveniência acadêmica. Em resposta a esse gargalo, ferramentas foram desenvolvidas para democratizar o acesso à análise de dados em larga escala, exemplo disso é a plataforma online, Arbimon¹, mantida pela organização *Rainforest Connection* (RFCx), que permite pesquisadores enviarem seus dados de áudio e utilizarem as ferramentas de aprendizado profundo para analisá-los, facilitando a colaboração em escala global.

Esses novos produtos exemplificam a inovação sob a ótica de Schumpeter ao transformarem radicalmente a capacidade de atuação da ciência, gerando valor e impacto concreto na ecologia acústica, que caracterizam uma genuína disruptão.

5.3. O Custo da Inovação: A Propriedade Intelectual e o Avanço Científico

Diante da onda de inovação tecnológica e da crescente disponibilização de novas ferramentas, levantam-se questões sobre propriedade intelectual. De um lado, têm-se a comunidade científica, a qual valoriza a colaboração e o código aberto, incentivando o compartilhamento de modelos e dados para acelerar pesquisas, como tem sido observado no movimento global de ciência aberta, o *Open Science*.

¹<https://arbimon.org/>

Por outro lado, com o avanço da tecnologia, surgem, também, oportunidades comerciais, motivadas pela exploração desse nicho de mercado. Uma empresa pode desenvolver e patentar um algoritmo que, por exemplo, tem maior eficiência na detecção de motosserras ou Veículos Aéreos Não Tripulados (VANTs), vendendo-o como um produto ou serviço. Nesse sentido, se o modelo ou os dados de treinamento forem protegidos por propriedade intelectual, cria-se um impedimento à replicação e a validação pela comunidade científica, criando uma barreira ao avanço acadêmico [Morandin et al. 2023].

Portanto, encontrar um equilíbrio entre a cultura de colaboração da ciência e os interesses comerciais será um desafio importante para garantir que essas tecnologias possam ser usadas para a conservação da natureza em todo o mundo. Nesse cenário, a inovação aberta desponta como uma via promissora para conciliar propriedade intelectual e compartilhamento de conhecimento, fortalecendo a cooperação entre diferentes atores e preservando o caráter coletivo do avanço científico.

6. Análise dos Resultados

Em resposta às questões de pesquisa delineadas na Tabela 1, esta seção analisa os resultados obtidos pela Revisão Rápida da Literatura. As subseções a seguir sintetizam os achados, abordando os impactos do ruído na vida selvagem e a transição das técnicas de aprendizado de máquina.

6.1. Q1: Como os ruídos gerados pela atividade humana impactam a vida selvagem?

A literatura selecionada evidencia que os ruídos antropogênicos causam distúrbios comportamentais e fisiológicos generalizados na vida selvagem. Estudos que abrangem diferentes classes e espécies de animais, documentam respostas adaptativas diretas: pássaros urbanos e anfíbios próximos a rodovias são forçados a elevar o tom de suas vocalizações para evitar o mascaramento acústico causado pelo tráfego. Em mamíferos, os impactos também são claros, alterando o comportamento acústico de baleias e, como aponta uma revisão recente, afetando comunidades inteiras ao induzir o estresse crônico e prejudicar funções vitais como o sono; achados que destacam a relevância do problema e a necessidade urgente de um monitoramento efetivo.

6.2. Q2: Quais técnicas de aprendizado de máquina estavam sendo utilizadas em trabalhos desenvolvidos há cerca de quinze anos atrás na área ecologia acústica?

Diante da análise de trabalhos anteriores à 2018, observou-se que algoritmos clássicos de aprendizado de máquina, como SVM e kNN, eram predominantes nas tarefas de classificação de ruídos. Outros estudos, também conduzidos neste período, utilizavam, ainda, o modelo estatístico HMM, por se apresentar como uma escolha robusta para a análise de vocalizações. Todavia, esses métodos possuíam uma limitação central: dependiam fortemente de um especialista para realizar a engenharia de características, ou seja, extrair manualmente atributos do áudio (como MFCCs e *Mel-spectrogram*), ou exigiam interfaces de software complexas que limitavam sua adoção, como no caso das HMMs.

6.3. Q3: Quais técnicas de aprendizado de máquina estão sendo utilizadas em trabalhos desenvolvidos nos últimos cinco anos na área de ecologia acústica?

Avançando para os trabalhos mais recentes, revela-se a transição tecnológica prevista pela teoria schumpeteriana da inovação. A revisão de Stowell [2022] sintetiza essa mudança, apontando que o aprendizado profundo tornou-se a abordagem dominante. Especificamente, as CNNs substituíram os métodos anteriores. A principal vantagem que impulsionou essa "destruição criadora" é a capacidade da arquitetura de aprender as características relevantes automaticamente, através da análise de espectrogramas de áudio. Isso não apenas aumentou a precisão, mas eliminou o gargalo da engenharia manual de características, tornando a análise mais acessível e eficiente.

7. Discussão

Os achados destacam o papel das inovações tecnológicas na detecção de ruídos antropogênicos, evidenciando como o aprendizado de máquina, em especial as redes neurais convolucionais, vem transformando a ecologia acústica. A partir de uma revisão rápida da literatura, foi possível identificar um processo de "destruição criadora", em que abordagens clássicas cedem lugar a soluções baseadas em aprendizado profundo.

A escolha pela Revisão Rápida da Literatura foi fundamentalmente para reunir os achados em tempo hábil. No entanto, cabe ressaltar que os resultados aqui sintetizados, mesmo advindo de pesquisas científicas, devem ser interpretados à luz de possíveis vieses, tendo em vista sua limitação em relação ao levantamento de dados, conforme apontado por Grant and Booth [2009].

7.1. Impactos à Validade Interna

A validade interna aborda a relação de substituição tecnológica, onde argumenta-se que as CNNs substituíram os métodos anteriores. Todavia, deve-se considerar a possibilidade da síntese ter negligenciado contradições. É possível que métodos clássicos (*e.g.*, SVM, HMM) ainda sejam superiores em nichos específicos, tornando a relação uma coexistência, e não uma substituição total.

7.2. Impactos à Validade Externa

Os autores limitaram a realização das buscas à plataforma Google Scholar, priorizando a seleção de artigos escritos em inglês. Isso pode ter excluído "literaturas cinzentas" ou trabalhos publicados em outros idiomas, onde métodos clássicos ainda podem ser amplamente utilizados. Portanto, a "destruição criadora" observada pode ser um fenômeno real apenas no subconjunto da literatura que foi indexada pela ferramenta de busca e acessível no idioma selecionado, não sendo generalizável para o campo da ecologia acústica em sua totalidade.

7.3. Impactos à Validade de Construto

A validade de construto refere-se à correta mensuração dos conceitos de "técnicas clássicas" (Q2) e "técnicas modernas" (Q3). A busca menos abrangente e a ausência de profundidade das pesquisas, não abrange todos os dados disponíveis, elucidando o construto de "destruição criadora" sem a devida carga de observações para sua validação.

7.4. Impactos à Validade por Conclusão

A validade da conclusão, de que a transição tecnológica se enquadra na teoria de Schumpeter, depende da qualidade das evidências. A Revisão Rápida da Literatura pode reduzir etapas de avaliação crítica e de síntese aprofundada [Grant and Booth 2009]. Logo, se os artigos sobre CNNs tiverem baixa qualidade, ou os clássicos alta, a força da conclusão fica comprometida, podendo haver uma sobrevalorização de pesquisas de menor qualidade.

8. Considerações Finais

Este artigo explorou como o uso de técnicas de aprendizado de máquina na detecção de ruídos antropogênicos representa mais do que um avanço técnico, mas um evento de inovação, como definido por Schumpeter. As novas tecnologias estão causando uma “destruição criadora” nos métodos de pesquisa em ecologia acústica, substituindo processos manuais por análises automáticas, rápidas e precisas.

As inovações de processo, como o uso de CNNs, e as inovações de produto, como os sistemas de monitoramento em tempo real, estão aumentando muito nossa capacidade de entender o impacto do ruído humano nos ecossistemas, possibilitando o monitoramento de mais lugares, por mais tempo e com um nível de detalhe que era impossível há uma década.

Para trabalhos futuros, objetiva-se a condução de uma Revisão Sistemática da Literatura (RSL) que valide e aprofunde os padrões de inovação elencados neste trabalho por meio de uma Revisão Rápida da Literatura. No tocante à análise das tecnologias, enquanto o estudo foca na transição para as CNNs, investigações futuras devem avaliar o impacto de arquiteturas emergentes, como *Transformers* [Vaswani et al. 2017] e abordagens de aprendizado auto-supervisionado, para determinar se representam a próxima onda de “destruição criadora” no domínio da bioacústica computacional.

Referências

- Aide, T. M., Corrada-Bravo, C., Campos-Cerqueira, M., Milan, C., Vega, G., and Alvarez, R. (2013). Real-time bioacoustics monitoring and automated species identification. *PeerJ*, 1:e103.
- Dahlheim, M. and Castellote, M. (2016). Changes in the acoustic behavior of gray whales *eschrichtius robustus* in response to noise. *Endangered Species Research*.
- Devane, D., C., H., Gartlehner, G., N.-S. B., Griebler, U., Affengruber, L., Saif-Ur-Rahman, K. M., and C., G. (2024). Key concepts in rapid reviews: an overview. *Journal of Clinical Epidemiology*, 175:111518. Epub 2024 Sep 6.
- Grant, M. J. and Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal*, 26(2):91–108.
- Higham, V., Deal, N. D. S., Chan, Y. K., Chanin, C., Davine, E., Gibbings, G., Keating, R., Kennedy, M., Reilly, N., Symons, T., Vran, K., and Chapple, D. G. (2021). Traffic noise drives an immediate increase in call pitch in an urban frog. *Journal of Zoology*.
- Huang, C.-J., Yang, Y.-J., Yang, D.-X., and Chen, Y.-J. (2009). Frog classification using machine learning techniques. *Expert Systems with Applications*, 36(2, Part 2):3737–3743.

- Kok, A. C. M., Berkhout, B. W., Carlson, N. V., Evans, N. P., Khan, N., Potvin, D. A., Radford, A. N., Sebire, M., Shafiei Sabet, S., Shannon, G., and Wascher, C. A. F. (2023). How chronic anthropogenic noise can affect wildlife communities. *Frontiers in Ecology and Evolution*, 11. Section Behavioral and Evolutionary Ecology.
- Lampert, T. A. and O'Keefe, S. E. (2010). A survey of spectrogram track detection algorithms. *Applied Acoustics*, 71:87–100.
- Lecun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324.
- Morandin, J. L. P. L., Silva, M. C. d., and Moura, A. M. M. d. (2023). As patentes e o desenvolvimento tecnológico no contexto da ciência aberta: perspectivas da influência do sigilo informacional e da pesquisa proprietária. *RDBCi: Revista Digital de Biblioteconomia e Ciência da Informação*, 21(00):e023019.
- Ranjard, L., Reed, B. S., Landers, T. J., Rayner, M. J., Friesen, M. R., Sagar, R. L., and Dunphy, B. J. (2017). Matlabhtk: a simple interface for bioacoustic analyses using hidden markov models. *Methods in Ecology and Evolution*, 8:615–621.
- Ren, Y., Johnson, M. T., Clemins, P. J., Darre, M., Glaeser, S. S., Osiejuk, T. S., and Out-Nyarko, E. (2009). A framework for bioacoustic vocalization analysis using hidden markov models. *Algorithms*, 2(4):1410–1428.
- Schumpeter, J. (1934). *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle*. Harvard University Press, Cambridge.
- Schumpeter, J. A. (1911). *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle*. Harvard University Press.
- Schumpeter, J. A. (1942). *Capitalism, Socialism and Democracy*. Harper Brothers.
- Slabbekoorn, H. and Peet, M. (2003). Ecology: Birds sing at a higher pitch in urban noise. *Nature*, 424(6946):267.
- Stowell, D. (2022). Computational bioacoustics with deep learning: a review and roadmap. *PeerJ*, 10:e13152.
- Stowell, D. and Plumbley, M. D. (2014). Automatic large-scale classification of bird sounds is strongly improved by unsupervised feature learning. *PeerJ*, 2:e488.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I. (2017). Attention is all you need.
- Wijayasingha, L. and Stankovic, J. A. (2021). Robustness to noise for speech emotion classification using cnns and attention mechanisms. *Smart Health*, 19:100165.
- Zhang, C., Sainath, T. N., Wu, Y., et al. (2018). Deep learning for environmentally robust speech recognition: An overview of recent developments. *IEEE Signal Processing Magazine*, 35(3):114–126.