



# AUXÍLIO A DIAGNÓSTICO POR IMAGENS MÉDICAS ATRAVÉS DO APRENDIZADO DE MÁQUINA COM TÉCNICA DE CLASSIFICAÇÃO UTILIZANDO APRENDIZADO PROFUNDO

*Wesley Vieira Godoy<sup>1</sup>, Marcos Monteiro Junior<sup>2</sup>, Maria Carolina Oliveira<sup>3</sup>*

<sup>1</sup> Acadêmico do Curso de Engenharia de Software, Campus Ponta Grossa-PR, Universidade Cesumar - UNICESUMAR. Bolsista PIBIC/ICETI - UniCesumar. w.g20@hotmail.com

<sup>2</sup> Campus Ponta Grossa-PR, Universidade Cesumar - UNICESUMAR. marcos.monteiro@unicesumar.edu.br

<sup>3</sup> Campus Ponta Grossa-PR, Universidade Cesumar - UNICESUMAR aulasmariacarolina@gmail.com

## RESUMO

A tuberculose é uma doença infecciosa causada pela bactéria *Mycobacterium tuberculosis*, que afeta milhões de pessoas em todo o mundo e representa um sério problema de saúde pública. O diagnóstico precoce e preciso da tuberculose é fundamental para melhorar os resultados dos pacientes e reduzir a propagação da doença. As radiografias de tórax têm sido amplamente utilizadas para o diagnóstico de tuberculose, pois podem revelar características típicas da infecção. No entanto, a interpretação das radiografias requer treinamento especializado e pode ser suscetível a erros, especialmente em regiões com escassez de recursos médicos. Com os recentes avanços em técnicas de aprendizado profundo, surgiu a possibilidade de desenvolver sistemas de inteligência artificial capazes de auxiliar no diagnóstico médico, incluindo o diagnóstico de tuberculose em radiografias de tórax. Neste contexto, este projeto científico tem como objetivo desenvolver um sistema de diagnóstico automatizado de tuberculose utilizando técnicas de aprendizado profundo, com o propósito de fornecer uma solução precisa e acessível para apoiar os profissionais de saúde no diagnóstico precoce da doença.

**PALAVRAS-CHAVE:** Aprendizagem de máquina; Aprendizado Profundo; Tuberculose.

## 1 INTRODUÇÃO

O Aprendizado de Máquina (AM) tem como propósito desenvolver algoritmos capazes de aprimorar sua performance através da análise de exemplos fornecidos (Mitchell, 1997). Esse processo requer uma vasta quantidade de dados para que o computador possa adquirir conhecimento e gerar hipóteses. Em essência, o AM busca aprender a partir de experiências passadas, permitindo que o algoritmo melhore suas habilidades e tomada de decisões à medida que mais dados são processados e analisados (FIA, 2021).

A aplicação de técnicas de AM tem revolucionado diversas áreas, e a medicina não é exceção. Em particular, o uso do Aprendizado Profundo (Deep Learning) tem se mostrado altamente promissor na área de diagnóstico por imagens médicas (Ludermir, 2021). O aprendizado profundo trouxe avanços significativos na análise de imagens médicas, possibilitando diagnósticos mais precisos e auxiliando os profissionais de saúde em suas decisões clínicas (Hongyu Zhu, 2018).

O diagnóstico precoce e preciso da tuberculose é um desafio significativo para os profissionais de saúde. A doença, causada pela bactéria *Mycobacterium Tuberculosis*, afeta principalmente os pulmões e pode ser assintomática em seus estágios iniciais, dificultando a detecção precoce e o tratamento efetivo. Além disso, a interpretação das imagens de radiologia de pulmões, como radiografias e tomografias computadorizadas, é complexa e pode exigir experiência clínica e conhecimento aprofundado em pneumologia. A interpretação de imagens de radiologia é uma tarefa desafiadora, especialmente em



casos de tuberculose pulmonar. Da mesma forma, os progressos nas pesquisas sobre diagnósticos automáticos de amostras histopatológicas estão gerando resultados que às vezes superam a análise feita por seres humanos (SABC, 2018).

Neste cenário, o uso do aprendizado de máquina com Aprendizado Profundo (AP) surge como uma ferramenta promissora para auxiliar os profissionais de saúde no diagnóstico de tuberculose. Ao treinar um modelo com uma grande quantidade de imagens radiológicas de pacientes diagnosticados com tuberculose e de indivíduos saudáveis, é possível capacitar a máquina para identificar padrões e características específicas associadas à doença. Essa capacidade pode levar a uma detecção mais precisa e precoce da tuberculose, permitindo um tratamento mais oportuno e melhorando o prognóstico dos pacientes. No entanto, a adoção de técnicas de inteligência artificial na área de imagens médicas enfrenta desafios singulares. Os diagnósticos nem sempre são definitivos, as classificações e conceitos nem sempre são consensuais e estão sujeitos a mudanças ao longo do tempo (Kohli M, Prevedello LM, Filice RW, et al. 2017).

No entanto, a aplicação do AP em diagnóstico por imagens médicas também apresenta desafios. A interpretação de imagens de radiologia envolve informações complexas e sutis que podem ser difíceis de serem capturadas por um modelo de aprendizado. Além disso, a falta de dados de treinamento rotulados e a necessidade de evitar erros de generalização do modelo podem ser obstáculos a serem superados. Por último, o aprendizado profundo apresenta um alto custo computacional, esse solucionado em parte, com uso de transferência de aprendizado com outras redes já treinadas e disponíveis em gratuitamente. No aprendizado por transferência são retiradas as camadas que não se deseja de uma rede treinada e coloca-se novas camadas de interesse.

Embora o uso do AP tenha mostrado resultados promissores, é importante abordar questões éticas e garantir que a tomada de decisão médica permaneça orientada por profissionais de saúde especializados.

Este trabalho tem como objetivo principal identificar qual rede de transferência de conhecimento o melhor desempenho em acurácia, para as imagens de doença pulmonar, especificamente a tuberculose.

Portanto, este projeto busca explorar como o AP pode ser aplicado de forma efetiva e segura no auxílio ao diagnóstico de tuberculose através de imagens de radiologia de pulmões. Ao enfrentar os desafios inerentes à aplicação dessa tecnologia, espera-se contribuir para o desenvolvimento de uma ferramenta promissora no âmbito da saúde pública, capaz de melhorar a qualidade do diagnóstico e, conseqüentemente, o tratamento dos pacientes com tuberculose (FIA, 2021).

## 2 MATERIAIS E MÉTODOS

Para esse trabalho será utilizada a base de dados *TB Chest Radiography Database* extraída do portal da PubMed Central (PMC). Essa base de dados possui 3500 imagens de radiografias de pulmões saudáveis e 700 imagens acometidos com tuberculose (Kaggle, 2022).

O trabalho utilizado como base de comparação foi proposto por T. Rahman (T. Rahman et al, 2020) que armazena código e base de dados públicos para utilização livre.

As bibliotecas necessárias para a implementação do trabalho são TensorFlow (Martín Abadi, 2015), Seaborn (Waskom, 2021), OpenCV (Bradski, G., 2000), Sklearn (Pedregosa et al., 2011).



Rahman, em seu trabalho, carregou os metadados de duas planilhas que contêm informações sobre as imagens de radiografia de tórax rotuladas como "Normal" e "Tuberculose". Os dados são combinados em um único *dataframe*. A rede para a transferência de aprendizado utilizada foi a ImageNet (DENG, J. et al., 2009).

No trabalho proposto, será utilizado o mesmo protocolo experimental de Rahman, descrito anteriormente, para efeitos comparativos. Serão utilizadas as redes GoogleLeNet (SZEGEDY, C. et al. 2015) e ResNet (HE et. Al., 2015). Será avaliado se existe ganho em acurácia quando aplicada as diferentes redes de aprendizado por transferência.

### 3 RESULTADOS E DISCUSSÕES

O modelo de Aprendizado Profundo, utilizando ImageNet, alcançou uma acurácia na classificação de tuberculose em imagens de radiologia de pulmões, com sensibilidade de 92% e especificidade de 88%. O uso do modelo pode auxiliar na detecção precoce da doença, melhorando o prognóstico dos pacientes.

Espera-se, com esse trabalho, atingir uma acurácia semelhante ou superior quando utilizadas outras redes pré-processadas de aprendizado profundo.

É fundamental destacar que os resultados sejam interpretados em conjunto com o conhecimento clínico, pois o modelo não substitui a expertise dos profissionais de saúde. A obtenção de uma base de dados bem anotada e a abordagem de questões éticas são aspectos importantes a serem considerados para o sucesso das aplicações clínicas do Aprendizado Profundo em diagnóstico por imagens médicas.

### 4 CONSIDERAÇÕES FINAIS

Concluiu-se que o uso do Aprendizado Profundo com técnicas de classificação representa uma abordagem promissora para auxiliar no diagnóstico de tuberculose por meio de imagens de radiologia de pulmões. Espera-se, que com outras redes de aprendizado seja possível melhorar o desempenho já obtido em trabalhos anteriores.

### REFERÊNCIAS

MITCHELL, T. Machine Learning. S. I.: McGraw Hill, 1997.

FIA Business School. Machine Learning: como funciona, benefícios, tipos e exemplos. Nov 2021. Disponível em: <https://fia.com.br/blog/machine-learning/>. Acesso em: 27 jun 2023.

LUDERMIR, Teresa. Inteligência Artificial e Aprendizado de Máquina: estado atual e tendências. Jan-Abr 2021. Disponível em: <https://doi.org/10.1590/s0103-4014.2021.35101.007>. Acesso em: 20 jun 2023.

ZHU, Hongyu, TBD: Benchmarking and Analyzing Deep Neural Network Training. Mar-Abr 2018. Disponível em: <https://doi.org/10.48550/arXiv.1803.06905>. Acesso em: 22 jun 2023.

SABC, Aspectos gerais da tuberculose: uma atualização sobre o agente etiológico e o tratamento. Mai-Nov 2018. Disponível em: <https://www.rbac.org.br/artigos/aspectos->



gerais-da-tuberculose-uma-atualizacao-sobre-o-agente-etilologico-e-o-tratamento/. Acesso em: 20 jun 2023.

Kohli M, Prevedello LM, Filice RW. Implementing machine learning in radiology practice and research. Abr 2017. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/28125274/>. Acesso em: 28 jun 2023.

Kaggle, Tuberculosis Classification DenseNet121 GradCAM. Out 2022. Disponível em: <https://www.kaggle.com/code/sanphats/tuberculosis-classification-densenet121-gradcam/notebook#Grad-cam-evaluation>. Acesso em: 10 jun 2023.

Rahman, T. et al., "Reliable Tuberculosis Detection Using Chest X-Ray With Deep Learning, Segmentation and Visualization," in IEEE Access, vol. 8, pp. 191586-191601. 2020. Disponível em: 10.1109/ACCESS.2020.3031384. Acesso em: 10 jun 2023.

Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng, Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Rafal Jozefowicz, Yangqing Jia, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dan Mané, Mike Schuster, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems. 2015. Disponível em: [tensorflow.org](https://tensorflow.org). Acesso em: 10 jun 2023.

Michael L. Waskom. Journal of Open Source Software. **Seaborn: statistical data visualization**, v. 6, pp. 3021. 2021. Disponível em: <https://doi.org/10.21105/joss.03021>. Acesso em: 10 jun 2023.

Bradski, G.. The OpenCV Library. **Journal of Software Tools**. 2000. Disponível em: [opencv.org](https://opencv.org). Acesso em: 10 jun 2023.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M. and Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E. . Journal of Machine Learning Research. 2011. **Scikit-learn: Machine Learning in Python**, v. 12, pp. 2825-2830.

Deng, J.. Imagenet: A large-scale hierarchical image database. **IEEE conference on computer vision and pattern recognition**, pp. 248–255, 2009.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D. & Rabinovich, A.. Going deeper with convolutions. **Proceedings of the IEEE conference on computer vision and pattern recognition**, pp. 1-9, 2015.

He, K., Zhang, X., Ren, S., & Sun, J.. Deep residual learning for image recognition. **Proceedings of the IEEE conference on computer vision and pattern recognition**, pp. 770-778, 2016.