

TB x-rAI: Apoio ao Diagnóstico de Tuberculose em Imagens Radiológicas através de Rede Neural Convolutacional e Aplicação Web

Isadora M.R. Cruz¹, Karina S. Machado¹

¹Centro de Ciências Computacionais (C3) – Universidade Federal do Rio Grande (FURG)

Rio Grande – RS – Brasil

isadoramataramoscruz@gmail.com, karinaecomp@gmail.com

Abstract. Tuberculosis (TB) is the second leading cause of death from an infectious agent worldwide, caused by the *Mycobacterium tuberculosis*. Chest X-ray imaging is the most commonly used screening examination for this diagnosis, despite the high inherent complexity and variability, especially in resource-limited settings. In this context, Machine Learning techniques, such as Convolutional Neural Networks (CNNs), emerge as valuable allies for faster diagnosis. Therefore, an adapted method based on an existing CNN model is proposed, aiming to detect TB more efficiently, along with a web application designed to serve as an additional decision-support tool for specialists.

Resumo. A Tuberculose (TB) é a segunda principal causa de morte por um agente infeccioso no mundo, causada pelo *Mycobacterium tuberculosis*. O raio-x torácico é o exame de triagem mais utilizado para esse diagnóstico, apesar da alta complexidade e variabilidade inerentes, principalmente em situações com recursos limitados. A partir desse cenário, técnicas de Machine Learning, como as Redes Neurais Convolucionais (CNNs) surgem como aliadas para um diagnóstico mais rápido. Assim, propõe-se um método adaptado de um modelo já existente de CNN, que seja capaz de detectar de forma mais eficaz a TB e uma aplicação web para servir como uma ferramenta de apoio adicional aos especialistas.

1. Introdução

A Tuberculose (TB), causada pelo *Mycobacterium tuberculosis*, segue sendo um desafio na saúde pública, já que representa a segunda principal causa de morte por um agente infeccioso no mundo [World Health Organization 2024]. Para realizar o diagnóstico da doença, o raio-x do tórax é o exame de triagem mais utilizado [Jaeger et al. 2013]. No entanto, a avaliação dessas radiografias é complexa e exige um alto conhecimento devido à variabilidade e sutileza nos padrões radiológicos, especialmente em cenários com recursos limitados [World Health Organization 2016].

Diante disso, a aplicação da inteligência artificial (IA), em especial o aprendizado de máquina (AM), proporcionou avanços em diversos segmentos, como no diagnóstico por meio de imagens médicas [Champendal et al. 2023]. A utilização do AM nesta área tem importante impacto visto que a capacidade de algoritmos, como as Redes Neurais Convolucionais (CNNs) em identificar padrões complexos para grandes conjuntos de dados pode auxiliar profissionais da área da saúde a tomarem decisões mais precisas e com maior rapidez, o que é essencial para um tratamento precoce e um melhor desfecho clínico para os pacientes.

Assim, este trabalho tem como objetivo propor uma solução baseada na arquitetura CNN proposta em [Pasa et al. 2019], para o diagnóstico de Tuberculose por meio de imagens de raio-x torácicas. Além disso, propõe-se também a criação do *website* TB x-rAI, uma interface amigável e intuitiva para os profissionais da área da saúde realizar o *upload* das imagens de raio-X e, utilizando o modelo proposto, predizer um diagnóstico de TB, auxiliando no diagnóstico.

2. Trabalhos Relacionados

Em [Pasa et al. 2019] foi proposto um modelo de CNN simples, rápido e mais eficiente do que em outros métodos tradicionais já existentes. Para isso, o *tbcnn* utilizou *saliency maps* e *grad-CAMs*, a fim de visualizar lesões sugestivas de tuberculose em radiografias. O estudo utilizou um conjunto de imagens provenientes de *datasets* públicos, *Montgomery Conty* nos *Estados Unidos* [Raddar 2020] e *Shenzhen* na *China* [Mengzhang 2024], além do *Belarus Tuberculosis Portal* [Frapa 2021], no qual foi obtida uma acurácia de 86% e AUC de 92%, combinando os três *datasets* para uma proporção de treino (80%), validação (10%) e teste (10%).

Em [Nafisah et al. 2021], os autores propõem um método utilizando o aprendizado profundo. Nesse, são empregadas redes de segmentação para extrair as regiões de interesse nos raios-x e IA explicável, na identificação de partes do pulmão infectadas por TB. Como resultados foram obtidos uma precisão média de 99,1% usando *EfficientNetB3* e área sob a curva (AUC) de 99%, para um conjunto de dados combinados (*Montgomery*, *Shenzhen* e *Belarus*).

Um estudo mais recente [Rajpurohit et al. 2025] propõe uma CNN com o aumento de dados generativos, sem sacrificar a transparência, melhorando a precisão e a interpretabilidade dos resultados através de mapas de ativação de classe e *LIME*, mecanismo utilizado para auxiliar na predição de uma rede neural, por meio da análise local de modelos complexos. Diante disso, o método utiliza um conjunto de 4200 imagens de raio-x torácicas, das quais apenas 700 são de TB-positivas. Realizando uma validação cruzada de cinco partições, o modelo alcançou uma acurácia de 95%.

3. Metodologia

O método proposto para a realização deste estudo concentra-se em duas etapas principais: (i) criação de uma arquitetura de Rede Neural baseada na de [Pasa et al. 2019], (ii) construção da aplicação *web* TB x-rAI, que permite a inserção de raios-x do pulmão e realiza predições (casos saudáveis e de TB) utilizando o modelo proposto.

Diferentemente da arquitetura original, o modelo proposto incorpora o mecanismo de atenção *Squeeze-and-Excitation* (SE), *ensemble learning* baseado na métrica AUC e emprega uma função de perda *Focal Loss*, além de aplicar outras técnicas de *data augmentation* e uso de *Early Stopping* durante o treinamento.

3.1. Base de imagens

O conjunto de dados utilizado é proveniente de imagens públicas de três *datasets* comumente utilizados na análise de raio-x torácicos [Frapa 2021], [Mengzhang 2024] e [Raddar 2020], respectivamente, *Belarus Tuberculosis Portal*, *Shenzhen* e *Montgomery Conty*. Essas bases de dados totalizaram 1104 imagens, sendo destas 662 do *dataset*

Shenzhen, com 336 TB-positivo e 326 TB-negativo, 138 de *Montgomery* com 80 imagens de TB-positivas e 58 TB-negativo e 304 de *Belarus* todas TB-positiva.

3.2. Pré-processamento

As imagens radiológicas passaram por uma padronização para estabelecer um formato comum a todas as amostras onde foram redimensionadas para 256x256 *pixels*. A seguir, foi realizada uma conversão para uma escala de cinza. Em sequência, foi utilizada a técnica de *data augmentation*, por meio do gerador *ImageDataGenerator*. Rotações, deslocamentos verticais e horizontais de até 15°, variação no brilho, zoom de 20%, espelhamento horizontal, distorções geométricas de 10% foram empregadas para simular as variações reais existentes em imagens torácicas, contribuindo para a redução do *overfitting* da rede e na melhora da generalização do modelo.

3.3. Arquitetura da rede

A arquitetura do método proposto (Figura 1) é composta por cinco blocos convolucionais residuais, na qual é aplicado um mecanismo de atenção *Squeeze-and-Excitation* (SE), que enfatiza regiões pulmonares importantes para o diagnóstico da TB. Cada bloco contém duas camadas convolucionais 3x3, seguidas de uma normalização em lote (*Batch Normalization*), função de ativação ReLU, uma conexão residual, com o intuito de preservar características relevantes nas imagens e, ao fim, uma operação de *max-pooling*. O modelo utiliza uma camada *GAP* (*Global Average Pooling*), seguida de *dropout* para reduzir o sobreajuste e uma camada densa, usando uma função de ativação *Sigmoid* para a classificação binária das imagens (0 - normal e 1 - tuberculose). A técnica de *ensemble learning* foi aplicada para melhorar a capacidade preditiva do modelo, no qual foram empregados três modelos com filtros iniciais e taxas de *dropout* distintas, (primeiro modelo filtro de 32 e *dropout* de 0,3, segundo modelo filtro de 64 e *dropout* de 0,4 e 128 e 0,5 para o terceiro). Além disso, uma função de perda *Binary Focal Cross Entropy* e pesos de classe inversamente proporcionais à frequência das amostras foram aplicados para amenizar o desbalanceamento entre as classes.

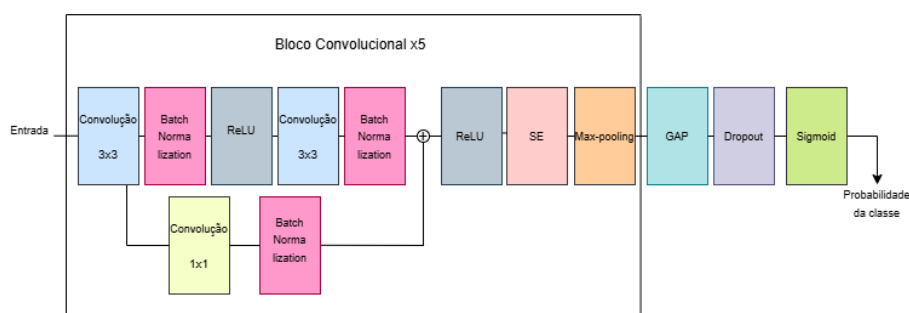


Figura 1. Arquitetura da CNN proposta.

3.4. Treinamento, Validação e Teste do Modelo

Nesta etapa, o conjunto de imagens foi dividido em três partes: treino (70%), validação (15%) e teste (15%). O treinamento foi realizado utilizando o otimizador Adam, com uma taxa de aprendizado de 10^{-4} , sendo determinados os valores de até 50 épocas para cada um dos três modelos e um *batch size* de 16. Técnicas como *Early Stopping* e

Reduce LR On Plateau foram aplicadas para reduzir o *overfitting* da rede e para ajustar a taxa de aprendizado obtida durante o treinamento. Na validação, a capacidade preditiva dos modelos foi monitorada ao longo das épocas, através das métricas acurácia, AUC, sensibilidade, precisão e função de perda. Nesse contexto, por meio da saída camada da rede que utiliza uma função de ativação *Sigmoid* são estimadas probabilidades no intervalo [0,1], na qual essas são aplicadas no cálculo da AUC. Cabe ressaltar que esses valores obtidos são empregados para definir os pesos do *ensemble*.

Ao fim, na etapa de testes, a abordagem do *ensemble learning* é adotada com a utilização de três modelos com taxas de *dropout* e filtros iniciais distintos, nos quais as saídas probabilísticas são combinadas através de uma média ponderada para se chegar a um único valor de predição. Para estabelecer o limiar de decisão do *ensemble*, foi aplicado o ajuste do *threshold*, sendo avaliados diferentes índices para a escolha de um valor ótimo. O valor final foi escolhido através da curva ROC, a partir do Índice de Youden, no qual foi encontrado o valor de 0,56.

3.5. Criação do *website*

Para a criação da interface *web* (TB x-rAI) foram utilizadas as tecnologias HTML, *Bootstrap*, *JavaScript* e o *microframework Flask* da linguagem *Python*, responsável pela publicação da plataforma e integração com o modelo desenvolvido. A ferramenta permitirá o *upload* de conjuntos de imagens de raio-x em que o resultado é gerado por meio da inferência considerando o modelo integrado ao *backend* da aplicação, a partir de imagens enviadas pelo usuário.

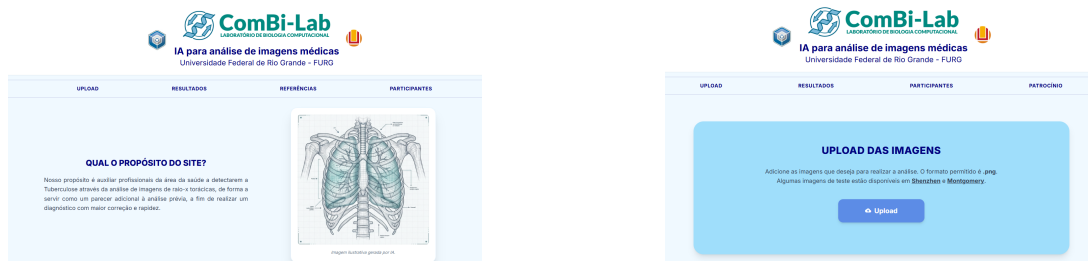


Figura 2. Tela principal do TB x-rAI (esquerda) e tela de *upload* das imagens, (direita).

4. Resultados Experimentais

4.1. Configuração Experimental

O treinamento do modelo foi realizado no ambiente virtual do *Google Colab*. Quanto às especificações do *hardware*, foi utilizada uma GPU Nvidia T4 Tensor Core. O desenvolvimento do modelo foi conduzido em linguagem *Python*, com a utilização das bibliotecas *TensorFlow*, *Keras*, *NumPy*, *Matplotlib* e *Scikit-Learn*.

4.2. Resultados do Modelo

O experimento realizado para o conjunto de teste demonstrou que o modelo proposto (Tabela 1, modelo completo) obteve uma acurácia de 84%, AUC de 89%, precisão de 95%, sensibilidade de 79% e F1-Score de 86%, para um *threshold* ótimo de 0,56. Em relação à análise das medidas de desempenho encontradas, a precisão de 95% reflete o baixo número de falsos positivos, essencial em cenários de diagnósticos médicos. Um

fator que se destaca como limitação é o baixo valor de sensibilidade (79%), resultando em um considerável número de falsos negativos, impactando negativamente o modelo. Com o objetivo de verificar a contribuição de cada uma das estratégias adotadas, bloco SE, *ensemble* e *data augmentation*, foi realizado um estudo de ablação, removendo, de maneira isolada, os componentes citados. A tabela 1 mostra os resultados obtidos.

Tabela 1. Comparação dos resultados obtidos em diferentes testes de ablação.

Estudo	Acc(%)	Auc(%)	Prec (%)	Sens. (%)	F1-Score(%)
Sem bloco SE	82	87	94	76	84
Sem ensemble learning	79	85	97	69	80
Sem data augmentation	85	91	97	79	87
Modelo completo	84	89	95	79	86

A partir dos resultados, nota-se que a utilização do *ensemble learning* foi relevante para uma melhor generalização do modelo e desempenho. Outro aspecto importante é que a ausência do bloco SE que também resultou em uma redução da performance do método, ao deixar de verificar regiões importantes no diagnóstico da TB. Entretanto, o teste sem *data augmentation* demonstrou resultados mais promissores em algumas das métricas, mostrando que variações na base de imagens podem ter introduzido ruído para a predição de casos de TB, influenciando negativamente o aprendizado. Comparando o modelo completo com a arquitetura proposta por [Pasa et al. 2019], semelhante ao método desenvolvido neste trabalho por utilizar a mesma base de dados de treinamento, diferenciando-se apenas no método de amostragem do treino/teste observa-se que os resultados obtidos pela CNN proposta mostram-se promissores, alcançando acurácia de 84% e AUC de 89%, valores próximos aos reportados por [Pasa et al. 2019] (acurácia de 86% e AUC de 92%).

4.3. Resultados para a plataforma TB x-rAI

A construção da plataforma TB x-rAI surge para realizar a integração do modelo de CNN ao *website* por meio do *microframework Flask*, no qual ele atua no *backend* da aplicação como uma arquitetura cliente/servidor, conectando o método proposto com o sistema *web* de forma a permitir o recebimento das imagens, o seu pré-processamento, análise e geração de resultados em tempo real. Destaca-se que a integração encontra-se em fase de implementação, com funcionamento pleno previsto para trabalhos futuros. Em relação à interface, o uso de JavaScript e demais ferramentas já exemplificadas anteriormente possibilitaram a criação de um *website* intuitivo e amigável para os profissionais da área da saúde, garantindo o envio das imagens e visualização dos resultados. Assim, o sistema possibilitará que especialistas analisem as radiografias de maneira rápida, servindo como suporte no diagnóstico da TB.

5. Conclusão

Este estudo apresentou um modelo de CNN Residual SE, juntamente com a interface *web* TB x-rAI, desenvolvidos para auxiliar no apoio ao diagnóstico da TB. Os resultados obtidos demonstraram o potencial do método proposto, devido principalmente ao uso de *ensemble learning*, ajuste do *threshold* e *Early Stopping*. O *website* TB x-rAI

encontra-se em estágio de desenvolvimento ainda não sendo possível a sua plena utilização. Como trabalhos futuros, propõe-se publicar online a TB x-rAI e adicionar novas funcionalidades. Com relação ao modelo proposto, como trabalhos futuros, busca-se melhorar as métricas de validação, incluir uma maior base de dados, realizar a validação do modelo com especialistas da área da saúde e estender a aplicabilidade desse modelo para outras doenças pulmonares.

6. Agradecimentos

O presente trabalho foi desenvolvido com apoio da CAPES, do CNPq processo número 406015/2025-2, da FAPERGS processos 25/2551-0002578-8, 22/2551-0000385-0 e 22/2551-0000390-7.

Referências

- Champendal, M., Müller, H., Prior, J., O., Reis, C., S. (2023). A scoping review of interpretability and explainability concerning artificial intelligence methods in medical imaging. *ScienceDirect*, 169:111159.
- Frapa (2021). A convolutional neural network for tuberculosis diagnosis from frontal chest x-rays, <https://www.github.com/frapa/tbcnn>. Acesso em: Janeiro de 2026.
- Jaeger, S., Karargyris, A., Candemir, S., Siegelman, J., Folio, L., Antani, S., Thoma, G. (2013). Automatic screening for tuberculosis in chest radiographs: a survey. *Quantitative Imaging in Medicine and Surgery*, 3(2), pages 89-99.
- Mengzhang, L. (2024). Awesome-Medical-Dataset, https://www.github.com/openmedlab/Awesome-Medical-Dataset/blob/main/resources/Shenzhen_chest_X-ray.md. Acesso em: Janeiro de 2026.
- Nafisah, S.I., Muhammad, G. (2021). Tuberculosis detection in chest radiograph using convolutional neural network architecture and explainable artificial intelligence. *National Library of Medicine*, 36, pages 111-131.
- Pasa, F., Golkov, V., Pfeiffer, F., Cremers, D., and Pfeiffer, D. (2019). Efficient deep network architectures for fast chest x-ray tuberculosis screening and visualization. *Scientific reports*, 9(1):6268.
- Raddar (2020). Tuberculosis Chest x-rays (Montgomery), <https://www.kaggle.com/datasets/raddar/tuberculosis-chest-xrays-montgomery>. Acesso em: Janeiro de 2026.
- Rajpurohit, S., S., Shirsath, V., A., Parvat, T., Patil, P., R., Nagarhalli, T., P., Save, A., M. (2025). Convolutional neural networks for automated tuberculosis detection in chest X-ray screening programs. *ScienceDirect*, 72, pages 115-122.
- World Health Organization (2016). *Chest radiography in tuberculosis detection: summary of current WHO recommendations and guidance on programmatic approaches*. World Health Organization.
- World Health Organization (2024). Global Tuberculosis Report. <https://www.who.int/teams/global-programme-on-tuberculosis-and-lung-health/tb-reports/global-tuberculosis-report-2024> Acesso em Abril de 2026.