

Uso de Visão Computacional na Análise de Exames Radiológicos Torácicos

Rafael Augusto Vasconcelos da Silva
União Dinâmica de Faculdades Cataratas
Foz do Iguaçu, BR
ORCID: 0009-0002-3804-4511

Miguel Diogenes Matrakas
União Dinâmica de Faculdades Cataratas
Foz do Iguaçu, BR
ORCID: 0000-0002-7047-5437

Alessandra Bussador
União Dinâmica de Faculdades Cataratas
Foz do Iguaçu, BR
ORCID: 0000-0002-5900-9398

Abstract—This project explores the planned application of computer vision, specifically Convolutional Neural Networks (CNNs), in the analysis and classification of thoracic radiological examinations. The primary goal is to develop and train a model capable of differentiating between healthy and diseased X-rays, contributing to more accurate and timely diagnostics of pulmonary diseases. The research will employ the ChestX-ray14 dataset, comprising over 100,000 annotated images, as the primary data source. The project is currently in the development phase, with future work focusing on model implementation, evaluation.

Keywords—Computer Vision; Convolutional Neural Networks; Medical Image Analysis; Pulmonary Diseases.

Resumo—Este projeto explora a aplicação planejada de visão computacional, especificamente Redes Neurais Convolucionais (CNNs), na análise e classificação de exames radiológicos torácicos. O objetivo principal é desenvolver e treinar um modelo capaz de diferenciar entre radiografias saudáveis e doentes, contribuindo para diagnósticos mais precisos e oportunos de doenças pulmonares. A pesquisa utilizará o dataset ChestX-ray14, composto por mais de 100.000 imagens anotadas, como a principal fonte de dados. O projeto está atualmente na fase de desenvolvimento, com trabalhos futuros focados na implementação do protótipo e avaliação do modelo.

Palavras-chave—Visão Computacional; Redes Neurais Convolucionais; Análise de Imagens Médicas; Doenças Pulmonares.

I. INTRODUÇÃO

As doenças pulmonares, como o câncer de pulmão, a pneumonia e a tuberculose, estão entre as principais causas de mortalidade em todo o mundo, afetando milhões de pessoas anualmente [1]. Essas condições demandam diagnósticos rápidos e precisos, fundamentais para a implementação de tratamentos eficazes e a melhoria das taxas de sobrevivência. O diagnóstico por imagem, utilizando técnicas como radiografia e tomografia computadorizada (TC), é essencial para detectar essas doenças em seus estágios iniciais. No entanto, esses métodos tradicionais apresentam limitações significativas, como a variabilidade na interpretação humana, que pode levar a erros diagnósticos, especialmente em ambientes com alta carga de trabalho [2].

A aplicação de Redes Neurais Convolucionais (CNNs) em visão computacional oferece uma solução promissora para automatizar a análise de imagens radiológicas, reduzindo a subjetividade e aumentando a precisão dos diagnósticos [3]. A arquitetura Residual Network, em particular, é altamente eficaz para este tipo de tarefa devido à sua capacidade de treinar redes profundas sem sofrer de problemas como o degradação do gradiente, possibilitando a extração de características complexas das imagens [4].

O objetivo deste trabalho é desenvolver e treinar um modelo de visão computacional, utilizando a arquitetura ResNet implementada em TensorFlow, para a análise e classificação de exames radiológicos torácicos. O modelo será capaz de diferenciar radiografias de pacientes saudáveis daquelas que apresentam alguma doença. Para isso, será utilizado o conjunto de dados ChestX-ray14 [1], aplicando a Rede Residual (ResNet) na tarefa de classificação binária. O treinamento do modelo incluirá testes com diferentes técnicas de otimização e regularização [5], visando determinar quais produzem os melhores resultados. A performance do modelo será avaliada por meio de métricas como precisão, sensibilidade e especificidade [6].

II. REFERENCIAL TEÓRICO

A. Doenças Pulmonares e a Necessidade de Diagnóstico Preciso

A detecção precoce dessas condições é essencial para aumentar as chances de sobrevivência e melhorar os resultados do tratamento [1]. No entanto, o diagnóstico precoce ainda é um desafio significativo, devido à complexidade das doenças e à necessidade de técnicas de imagem que possam revelar anomalias em estágios iniciais [2].

B. Redes Neurais Convolucionais (CNNs)

As Redes Neurais Convolucionais (CNNs) são uma classe especializada de redes neurais projetadas para processar dados visuais. Elas são particularmente eficazes na análise de imagens

médicas devido à sua capacidade de aprender características hierárquicas, desde bordas simples até padrões complexos, o que é crucial para a detecção de anomalias em radiografias [5]. A arquitetura das CNNs, com suas camadas convolucionais, de pooling e de ativação, permite a extração eficiente de características relevantes, facilitando a classificação das imagens entre saudáveis e patológicas [3].

C. Aprendizagem Residual e Redes Residuais

A aprendizagem residual, introduzida por He et al. [4], é um avanço significativo no treinamento de redes neurais profundas. Nessa abordagem, as camadas da rede são reformuladas para aprender funções residuais em relação às entradas da camada [4], o que facilita a otimização de redes muito profundas. As ResNets, baseadas na aprendizagem residual, permitiram a construção de redes com até 152 camadas, mantendo a complexidade computacional relativamente baixa em comparação com outras arquiteturas, como as redes Visual Geometry Group (VGG) [7].

As ResNets foram testadas em diversos conjuntos de dados, como ImageNet [8] e CIFAR-10 [9], demonstrando excelente desempenho em tarefas de reconhecimento de imagens [4]. Em particular, mostraram que redes mais profundas, quando bem otimizadas, podem superar significativamente redes mais rasas, resolvendo o problema de degradação que geralmente ocorre com o aumento da profundidade da rede. A eficácia das ResNets em aumentar a precisão com o aumento da profundidade pode ser uma justificativa sólida para sua aplicação em tarefas complexas de análise de imagens médicas.

A ResNet-50 é uma das variantes mais conhecidas da família de Redes Residuais (ResNets) [4]. Essa rede possui 50 camadas de profundidade, organizadas em uma série de blocos residuais. O principal avanço das ResNets é a capacidade de treinar redes muito profundas, resolvendo o problema da degradação do gradiente que comumente ocorre em redes profundas tradicionais. Cada bloco residual é composto por camadas convolucionais, seguidas por uma função de ativação ReLU (Rectified Linear Unit) e uma conexão de atalho (skip connection) que permite que a entrada do bloco seja somada diretamente à saída [4]. Essa estrutura facilita o aprendizado de representações complexas e permite que a rede aprenda de forma mais eficaz, mesmo com grande profundidade.

D. Dataset ChestX-ray14

O ChestX-ray14 é um dos maiores e mais abrangentes conjuntos de dados públicos disponíveis para a pesquisa de doenças torácicas, contendo 112.120 imagens de raios-X de tórax em visão frontal, extraídas de mais de 30.000 pacientes únicos [1]. Este dataset inclui 14 categorias de doenças torácicas comuns, como atelectasia, cardiomegalia, derrame pleural, infiltração,

massa, nódulo, pneumonia, pneumotórax, consolidação, edema, enfisema, fibrose, espessamento pleural e hérnia. As etiquetas das doenças foram extraídas automaticamente dos relatórios radiológicos utilizando processamento de linguagem natural (NLP), o que permitiu a criação de um grande conjunto de dados com uma acurácia estimada superior a 90

O uso do ChestX-ray14 é motivado pela necessidade de grandes volumes de dados para treinar redes neurais profundas de maneira eficaz, especialmente em tarefas de classificação e localização de doenças torácicas. Este dataset fornece uma base sólida para o desenvolvimento de modelos de CNNs, permitindo a construção de sistemas de diagnóstico assistido por computador com potencial para melhorar significativamente a acurácia e a eficiência dos diagnósticos em ambientes clínicos [3].

E. Desafios e Limitações

Apesar de seu tamanho e abrangência, o ChestX-ray14 possui algumas limitações, como a ausência de caixas delimitadoras para a maioria das doenças e a possibilidade de erros nas etiquetas derivadas por NLP. Além disso, os relatórios radiológicos completos não estão disponíveis publicamente, o que limita a possibilidade de verificação manual das etiquetas [1]. Estes desafios destacam a necessidade de abordagens robustas para lidar com a incerteza e a variação nos dados ao utilizar este dataset.

F. Resultados Empíricos e Aplicabilidade Geral

Os resultados empíricos apresentados por He et al. [4] demonstram que as ResNets não apenas facilitam o treinamento de redes profundas, mas também melhoram significativamente a precisão dos modelos em tarefas de classificação de imagens. No conjunto de dados ImageNet [8], a ResNet com 152 camadas atingiu um erro de classificação top-5 de apenas 3,57%, um desempenho superior ao de redes mais rasas. Além disso, a versatilidade das ResNets foi comprovada em outras tarefas de visão computacional, como detecção e segmentação de objetos no conjunto de dados COCO [10], evidenciando sua capacidade de generalização para diferentes problemas.

A robustez das ResNets, combinada com sua capacidade de generalização, torna essa arquitetura particularmente adequada para aplicações em diagnósticos médicos, onde a precisão e a confiabilidade são cruciais [4]. A aplicação de ResNets na análise de radiografias torácicas pode, portanto, proporcionar melhorias significativas na detecção precoce de doenças pulmonares, contribuindo para melhores resultados clínicos [1], [3].

III. METODOLOGIA

Este projeto tem como objetivo desenvolver um modelo de Rede Neural Convolucional (CNN) utilizando a biblioteca

TensorFlow, especificamente aplicando a arquitetura ResNet [4], para a análise e classificação de exames radiológicos torácicos. A seguir, são descritas as etapas metodológicas em detalhes:

A. Coleta e Pré-processamento de Dados

As imagens radiológicas serão obtidas do dataset ChestX-ray14, que contém mais de 100.000 radiografias torácicas com diversas condições patológicas anotadas [1]. O pré-processamento dos dados inclui:

- **Redimensionamento:** As imagens serão redimensionadas para uma resolução padrão de 224x224 pixels, um formato compatível com a arquitetura ResNet [4].
- **Normalização:** Os valores dos pixels das imagens serão normalizados na faixa [0, 1] para melhorar a eficiência do treinamento [5].
- **Divisão de Dados:** O dataset será dividido em conjuntos de treinamento, validação e teste, permitindo a avaliação contínua da performance do modelo [3].

B. Arquitetura ResNet

O modelo CNN será baseado na arquitetura ResNet (Residual Network), conhecida por permitir o treinamento eficaz de redes muito profundas, superando problemas como a degradação do gradiente [4]. A ResNet utiliza blocos residuais, onde a entrada do bloco é adicionada à sua saída após a aplicação de convoluções e ativação. Essa estrutura facilita o aprendizado de funções complexas sem perda de informação durante o treinamento.

- **Blocos Residuais:** Cada bloco residual inclui camadas convolucionais seguidas de uma função de ativação ReLU e uma conexão de atalho que soma a entrada do bloco à sua saída [4]. Isso ajuda a mitigar problemas de degradação do gradiente, especialmente em redes profundas.
- **Camadas de Pooling e Fully Connected:** A rede finaliza com camadas de pooling para reduzir a dimensionalidade e camadas fully connected para a classificação das imagens em saudáveis ou anormais [4].

C. Algoritmo de Otimização: Gradiente Descendente Estocástico (SGD)

O treinamento do modelo será realizado utilizando o algoritmo de Gradiente Descendente Estocástico (SGD), uma técnica que ajusta os pesos da rede iterativamente para cada mini-batch de dados de entrada. O SGD permite uma convergência mais rápida em grandes datasets ao atualizar os pesos mais frequentemente do que o gradiente descendente padrão [5].

- **Atualização dos Pesos:** Em cada iteração, os pesos da rede serão ajustados na direção oposta ao gradiente calculado para minimizar a função de perda. Isso permite ao modelo melhorar sua performance a cada passo do treinamento [5].

D. Otimizações: Adam e Dropout

Para melhorar o desempenho e a generalização do modelo, serão aplicadas as seguintes otimizações:

- **Adam (Adaptive Moment Estimation):** Embora o SGD seja o algoritmo de base, a otimização será realizada com Adam, uma variante que ajusta a taxa de aprendizado de forma adaptativa para cada parâmetro, combinando os benefícios de Adaptive Gradient Algorithm (AdaGrad) e Root Mean Square Propagation (RMSProp) [5]. Adam mantém médias móveis dos gradientes e dos quadrados dos gradientes para estabilizar as atualizações, resultando em uma convergência mais rápida e eficiente.
- **Dropout:** Para prevenir overfitting, será utilizada a técnica de dropout, que "desativa" aleatoriamente uma fração dos neurônios durante o treinamento [5]. Isso força a rede a aprender representações mais robustas, sem depender excessivamente de neurônios específicos, o que melhora a capacidade de generalização do modelo.

E. Métricas de Avaliação

Para garantir que o modelo de Rede Neural Convolucional (CNN) desenvolvido seja eficaz na classificação de imagens radiológicas, será utilizado diversas métricas de avaliação [3]. Cada uma dessas métricas fornece uma perspectiva única sobre o desempenho do modelo:

- **Precisão (Accuracy):** A precisão é uma medida geral de quantas vezes o modelo acertou na classificação das imagens, comparado ao número total de imagens avaliadas. Se o modelo acertou muitas vezes, a precisão será alta. No contexto do seu projeto, a precisão mostrará quantas vezes o modelo identificou corretamente se uma imagem radiológica é saudável ou apresenta alguma anomalia. No entanto, a precisão pode não ser suficiente para avaliar o desempenho do modelo em casos onde há um desequilíbrio entre as classes de imagens saudáveis e anormais.
- **Sensibilidade (Recall):** A sensibilidade, também conhecida como recall, indica o quão bem o modelo consegue identificar as imagens que realmente possuem anomalias [3]. Em outras palavras, mede a capacidade do modelo de encontrar todos os casos de anomalia presentes no conjunto de dados.
- **Especificidade:** A especificidade mede a habilidade do modelo de reconhecer imagens saudáveis corretamente, ou

seja, de confirmar que uma imagem não apresenta anomalias quando realmente está saudável. Isso é crucial para evitar falsos alarmes, onde pacientes saudáveis poderiam ser erroneamente diagnosticados com uma condição que não possuem, levando a possíveis exames e tratamentos desnecessários.

- **AUC-ROC (Area Under the Curve - Receiver Operating Characteristic):** A AUC-ROC é uma métrica que avalia a capacidade geral do modelo de distinguir entre imagens saudáveis e anormais [3]. A AUC, ou área sob a curva ROC, oferece uma visão de como o modelo performa em diferentes situações, considerando tanto a sensibilidade quanto a especificidade em diferentes níveis de decisão. Um valor elevado de AUC significa que o modelo é bom em separar corretamente as imagens de pacientes saudáveis das imagens de pacientes com anomalias, independentemente do ponto de corte escolhido para decidir entre saudável e anormal.

F. Ferramentas e Tecnologias Utilizadas

- **TensorFlow 2.x:** A biblioteca principal para o desenvolvimento do modelo CNN, escolhida por sua robustez e escalabilidade [5].
- **Keras:** Utilizada como a API de alto nível para a construção do modelo, oferecendo uma interface simples e intuitiva [5].
- **Python:** A linguagem de programação utilizada, com suporte de bibliotecas adicionais como NumPy e Matplotlib [5].

IV. RESULTADOS ESPERADOS

Os resultados esperados deste projeto incluem a criação de um modelo eficiente de Rede Neural Convolutiva (CNN) baseado na arquitetura ResNet [4], para a classificação de imagens radiológicas torácicas. O modelo deve ser capaz de distinguir entre radiografias saudáveis e aquelas que apresentam patologias pulmonares, com altos índices de precisão, sensibilidade e especificidade [3].

A avaliação do modelo será realizada com base em métricas como a AUC-ROC, fornecendo uma visão completa sobre a capacidade do modelo em distinguir corretamente as imagens saudáveis das anormais [3]. Com isso, o projeto espera alcançar um modelo que ofereça suporte clínico para diagnósticos mais rápidos e precisos, potencialmente melhorando os resultados obtidos no processo de emissão de laudos aos pacientes [6].

V. CONCLUSÃO

Este artigo apresentou um estudo sobre o desenvolvimento de um sistema baseado em Redes Neurais Convolutivas (CNN) para a análise de exames radiológicos torácicos, com o objetivo

de auxiliar no diagnóstico de doenças pulmonares. O modelo proposto utiliza a arquitetura ResNet [4], implementada em TensorFlow, para a classificação de radiografias como saudáveis ou com alguma patologia pulmonar, com base no conjunto de dados ChestX-ray14 [1].

Embora o projeto ainda esteja na fase de desenvolvimento teórico, espera-se que a abordagem adotada contribua para a melhoria da precisão e da eficiência no diagnóstico médico assistido por computador. As etapas descritas, como a coleta e o pré-processamento de dados, a arquitetura da CNN e as técnicas de otimização [5], formam a base para a implementação futura e a realização de testes experimentais.

REFERÊNCIAS

- [1] X. Wang, Y. Peng, L. Lu, M. Bagheri, and R. M. Summers, "Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 3462–3471.
- [2] T. Bai, Y. Pang, and T. H. Meen, "Fusion high-resolution network for diagnosing chest x-ray images," *Electronics*, vol. 9, no. 1, p. 190, 2020.
- [3] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, and A. Y. Ng, "Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning," *arXiv preprint*, 2017. [Online]. Available: <https://arxiv.org/abs/1711.05225>
- [4] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015. [Online]. Available: <https://arxiv.org/abs/1512.03385>
- [5] M. Nielsen, *Neural Networks and Deep Learning*. Self-published, 2018. [Online]. Available: <http://neuralnetworksanddeeplearning.com/>
- [6] K. Hammoudi, H. Benhabiles, M. Melkemi, F. Dornaika, and I. Arganda-Carreras, "Deep learning on chest x-ray images to detect and evaluate pneumonia cases at the era of covid-19," *HAL Archives*, 2020. [Online]. Available: <https://hal.archives-ouvertes.fr/hal-025533605>
- [7] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014, acessado em: 22 setembro 2024. [Online]. Available: <https://arxiv.org/abs/1409.1556>
- [8] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. (2024) Imagenet: A large-scale hierarchical image database. Acessado em: 22 setembro 2024. [Online]. Available: <https://www.image-net.org/index.php>
- [9] A. Krizhevsky. (2024) Learning multiple layers of features from tiny images, cifar-10 and cifar-100 datasets. Acessado em: 22 setembro 2024. [Online]. Available: <https://www.cs.toronto.edu/~kriz/cifar.html>
- [10] C. Dataset. (2024) Coco - common objects in context. Acessado em: 22 setembro 2024. [Online]. Available: <https://cocodataset.org/#home>