



Aprimorando diagnósticos de tomografia computadorizada de cabeça: uma abordagem de IA para detecção de artefatos metálicos.

Palavras Chave: Tomografia computadorizada, Inteligência artificial, Detecção de anomalias

Hospital Israelita Albert Einstein (HIAE)

Faculdade de Engenharia Elétrica e de Computação (FEEC)

Autores

João Vitor Souza de Alcantara [IMECC]

Bruna Garbes Goncalves Pinto [HIAE]

Profa. Leticia Rittner (Orientadora)[FEEC]

1 Introdução

A Tomografia Computadorizada (TC) é uma técnica de aquisição de imagens amplamente utilizada em pesquisas médicas e diagnósticos clínicos, fornecendo informações detalhadas sobre as estruturas cerebrais [1]. No entanto, apesar de sua eficácia, a qualidade das imagens de TC podem ser comprometidas por diversos fatores, incluindo distorções na imagem que chamamos de *artefatos*. Os artefatos podem ter diferentes origens, como movimento de pacientes, problemas de hardware, peças metálicas, entre outros [2]. Esses artefatos levam a distorções e interferências nas imagens, prejudicando a precisão das análises e interpretações, e, conseqüentemente impactando nas tomadas de decisões clínicas e de pesquisa.

A detecção e correção de artefatos em imagens de Tomografia Computadorizada têm sido uma preocupação constante na prática clínica e em pesquisas médicas. Abordagens convencionais, como análise visual por radiologistas experientes e técnicas de pré-processamento de imagens, desempenham um papel crucial na identificação de artefatos. No entanto, essas abordagens têm limitações em termos de eficiência e precisão, especialmente com o crescente volume de dados gerados e a complexidade dos artefatos decorrentes de diferentes fontes, como peças metálicas e problemas de hardware. Portanto, existe uma necessidade premente de explorar soluções mais avançadas e automatizadas para melhorar a detecção e a interpretação de artefatos, a fim de garantir diagnósticos mais confiáveis e tomadas de decisões mais precisas [3].

Nesse contexto, o presente trabalho se propõe a investigar o uso de técnicas avançadas de processamento de imagens e aprendizado de máquina, com foco na aplicação de Inteligência Artificial [4], para enfrentar esse desafio e contribuir significativamente para a melhoria da qualidade dos dados de TC, viabilizando diagnósticos mais precisos e tratamentos mais efetivos. Ao enfrentar a problemática apresentada, espera-se que o modelo de IA desenvolvido possa proporcionar uma solução eficiente para identificar essas distorções, colaborando tanto com a prática clínica quanto pesquisas médicas.

2 Metodologia

Base de Dados

Nessa etapa do projeto, será utilizada uma base de dados de imagens de TC de cabeça proveniente do Hospital Israelita Albert Einstein (HIAE). A fonte de dados disponibilizou cerca de 8319 sujeitos previamente avaliados e diagnosticados por profissionais. No entanto, dentre os 8319 sujeitos, muitos apresentam erros nos laudos, colocando-os em classes erradas. Afim de evitar grandes erros e desbalanceamento de classes, optou-se por manter 1000 dados em cada classe (com e sem presença de artefatos)

Antes da inclusão na base de dados do projeto, foram aplicados critérios de seleção para garantir a qualidade e representatividade da amostra. Pacientes com artefatos metálicos em suas imagens foram cuidadosamente identificados e incluídos no conjunto de dados, tomando-se o cuidado para garantir um conjunto heterogêneo no que diz respeito a sexo, idade, tipos de artefatos e condições clínicas (Fig.1).

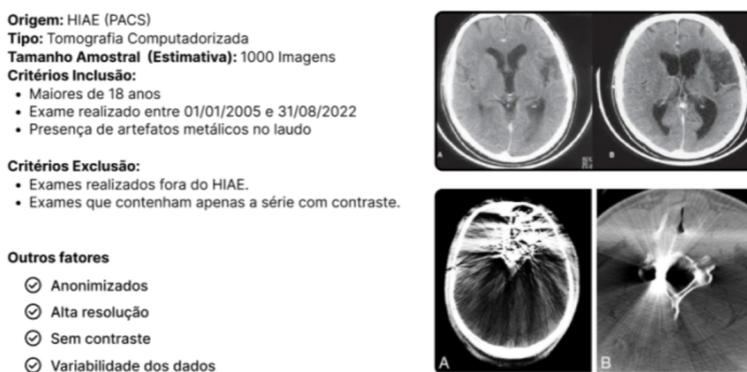


Figura 1: Informações sobre os dados utilizados no projeto;

Para utilizar as amostras, foi necessário realizar um pré-processamento, pois os modelos e arquiteturas de IA possuem padrões de entradas. As principais etapas de pré-processamento utilizadas são o redimensionamento das imagens e a normalização.

Implementação do Modelo: estudo e escolha de Arquiteturas

A classificação de imagens é um problema amplamente discutido no meio científico, assim, existem diversas arquiteturas e modelos já implementados que possuem bons resultados. Porém, a seleção adequada da arquitetura de rede neural a ser utilizada é um fator determinante para o sucesso deste tipo de projeto, pois afeta diretamente o desempenho do modelo a ser treinado. Para realizar essa seleção, foram adotados critérios que levam em consideração as características e desafios da tarefa.

Inicialmente, buscamos arquiteturas de redes neurais que demonstrem bons resultados em tarefas de classificação de imagens, especialmente em domínios médicos. Para isso, realizamos a pesquisa na literatura científica, analisando estudos e artigos que relatam o desempenho de diferentes arquiteturas em problemas similares de detecção de anomalias e artefatos em imagens médicas [5].

Além disso, consideramos a complexidade computacional e os recursos de treinamento necessários para cada arquitetura. Dada a natureza das imagens médicas em TC, é fundamental que o modelo seja eficiente em termos de tempo de treinamento e requisitos de memória, permitindo a aplicação prática em ambientes clínicos.

Outro critério importante a se considerar é o tamanho do conjunto de dados e o número de parâmetros da rede a ser treinada. Não adianta escolhermos uma rede com um número grande de parâmetros, se o conjunto de treinamento que temos é limitado.

Com base nesses critérios, identificamos quatro arquiteturas promissoras para a detecção de artefatos em imagens de TC:

- **EfficientNet**: arquitetura de Rede Neural Convolutiva (CNN) que apresenta um excelente equilíbrio entre desempenho e quantidade de parâmetros, tornando-a eficiente em ambientes computacionais restritos e em implementações clínicas que não possuem grande poder computacional;
- **ResNet**: reconhecida por sua capacidade de treinar redes profundas, permitindo a criação de modelos complexos e eficazes. Sua estrutura com blocos residuais apresenta aspectos relevantes para problemas com artefatos de TC, onde a detecção de características sutis é essencial [7];
- **DenseNet**: arquitetura CNN que se destaca por sua densa conectividade entre camadas. Enquanto em outras arquiteturas tradicionais as camadas se conectam sequencialmente, a DenseNet realiza uma abordagem única onde cada camada recebe entradas de todas as camadas anteriores. [8];
- **Vision Transformers (ViT)**: arquitetura recentemente introduzida que adapta os conceitos dos transformers, inicialmente desenvolvidos para Processamento de Linguagem Natural (PLN), para o processamento de imagens. O ViT possui potencial para capturar informações contextuais significativas em imagens médicas complexas, permitindo uma detecção mais precisa de artefatos [9].

O ViT é a principal abordagem a ser explorada nesse projeto, ele tem como principais inovações a remoção das camadas convolucionais tradicionais das CNNs, substituindo por blocos de *transformers* para processar diretamente os pixels. Os transformers consistem em um mecanismo de atenção que permite que a rede concentre em partes importantes da imagem, gerando representações significativas de características. Assim, a imagem é dividida em pequenas regiões (*patches*) retangulares e cada região é tratada como uma "palavra" nos *transformers*. A sequência dessas regiões é passada então para a arquitetura e podemos realizar uma classificação, que nesse caso, é binária (Fig.2).

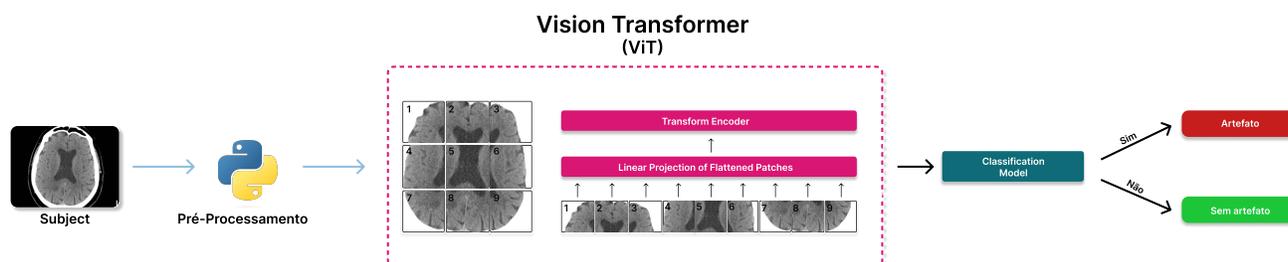


Figura 2: Ilustração de Pipeline para detecção de artefatos em imagens de (TC) utilizando ViT. Após o pré-processamento, as imagens são quebradas em regiões, tratadas como *palavras* pelo modelo.

O uso do ViT tem atraído crescente interesse na comunidade de pesquisa em visão computacional médica. Ele tem sido aplicado em tarefas como classificação e segmentação, no entanto, embora suas vantagens de precisão sejam altas, ele apresenta desafios em termos de requisitos computacionais e recursos de treinamento. Dessa forma, foram feitas análises e comparações entre os outros modelos, afim de encontrar o melhor resultado que contenha equilíbrio entre desempenho e custo.

Treinamento e Validação do Modelo

O conjunto de dados pré-processados foi dividido em 80% (1600) para treinamento e 20% (400) para validação, a fim de avaliar a capacidade de generalização e evitar o overfitting do modelo. O treinamento do modelo é conduzido por meio do ajuste dos pesos e parâmetros das redes com base em uma função de perda apropriada, como a Entropia Cruzada. O progresso do treinamento é monitorado para garantir a convergência e evitar problemas de otimização, além do monitoramento de custo e eficiência.

A fim de manter uma organização e comparação entre os resultados e performance dos diferentes modelos selecionados, foi desenvolvida uma interface de testes para facilitar a manutenção dos modelos e os testes. A interface permite selecionar um dos modelos previamente implementados e ajustar os hiperparâmetros para realizar novos testes e previsões.

As métricas de desempenho para a avaliação do modelo incluem a acurácia, precisão, F1-score e matriz de confusão. Foram calculadas tanto no conjunto de treinamento quanto no conjunto de validação, visando uma avaliação completa e imparcial. Após o treinamento e validação do modelo, os profissionais especializados desempenham um papel essencial, participando da validação visual das previsões do modelo, garantindo sua precisão e confiabilidade.

3 Resultados preliminares:

Os resultados preliminares sugerem que o modelo de IA desenvolvido tem o potencial significativo na detecção de artefatos em imagens de TC. No entanto, a classificação de artefatos em imagens de TC 3D apresenta desafios específicos em relação às aplicações em imagens 2D. Esses desafios incluem:

- A natureza 3D das imagens de TC apresenta informações adicionais que podem ser valiosas para a detecção de artefatos, porém torna o problema mais complexo. Uma das soluções pensadas é o uso de arquiteturas ViT baseadas em transformers 3D, que ajudam a coletar informações globais espaciais de forma efetiva.
- O volume dos dados, que é muito maior do que em imagens 2D, leva a demandas computacionais elevadas durante o treinamento e inferência. O uso de subvolumes menores (*patches*) das imagens 3D durante o treinamento foi adotado, ajudando a reduzir a carga computacional e torna o processo mais eficiente.
- Os dados obtidos através do HIAC apresentam anotações incorretas ou inconscientes, uma vez que os laudos são dados por diferentes profissionais, tornando difícil um padrão nas anotações. Assim, em alguns casos, é necessário a inspeção visual de alguns sujeitos para considerar sua permanência nos conjuntos.

Até o momento, apenas um modelo baseado em uma EfficientNetB2 foi implementado e treinado em um conjunto de dados de imagens de TC de base pública provenientes do The Cancer Imaging Archive (TCIA). O estudo contém 120 sujeitos, e foi obtida uma precisão de 0.89, recall de 0.9 E F1 Score de 0.93. Os resultados preliminares são encorajadores a respeito da possibilidade de classificar as imagens de TC e sugerem que o modelo tem o potencial de fazer uma contribuição significativa para a melhoria da qualidade dos dados de TC. No entanto, são necessários mais estudos para validar os resultados preliminares e avaliar o desempenho do modelo em um conjunto de dados mais abrangente.

4 Conclusão

O presente projeto de pesquisa tem como objetivo o desenvolvimento e avaliação de um modelo de Inteligência Artificial para a detecção de artefatos metálicos em imagens de Tomografia Computadorizada de cabeça. Ao longo do trabalho até o presente momento, foram exploradas diversas técnicas de processamento de imagens e modelos de aprendizado de máquina, incluindo o uso de arquiteturas como EfficientNet e Vision Transformer.

As próximas etapas do projeto incluem testes de diferentes arquiteturas (ViT, ResNet e DenseNet) para comparação e seleção e inserção do modelo treinado em um ambiente de teste com profissionais capacitados de verificar sua performance clínica e para validar sua eficiência em ambientes e pipelines reais.

Agradecimentos

Os autores gostariam de agradecer ao Hospital Israelita Albert Einstein (HIAE) pelo auxílio financeiro, bem como aos colaboradores do projeto: Artur José Marques Paulo; Camila Aquemi Silva; Joselisa Péres Queiroz de Paiva; Márcio Rodrigues da Cunha Reis; Rafael Maffei Loureiro; Regiane Maria Ribeiro de Carvalho; Tayran Milá Mendes Olegário.

Referências

- [1] Hussain S, Mubeen I, Ullah N, Shah SSUD, Khan BA, Zahoor M, Ullah R, Khan FA, Sultan MA. Modern Diagnostic Imaging Technique Applications and Risk Factors in the Medical Field: A Review. *Biomed Res Int.* 2022 Jun 6;2022:5164970. doi: 10.1155/2022/5164970.
- [2] Barrett JF, Keat N. Artifacts in CT: recognition and avoidance. *Radiographics.* 2004 Nov-Dec;24(6):1679-91. doi: 10.1148/rg.246045065.
- [3] Wei L, Rosen B, Vallières M, Chotchutipan T, Mierzwa M, Eisbruch A, El Naqa I. Automatic recognition and analysis of metal streak artifacts in head and neck computed tomography for radiomics modeling. *Phys Imaging Radiat Oncol.* 2019 Jun 6;10:49-54. doi: 10.1016/j.phro.2019.05.001.
- [4] Bruijne M. Machine learning approaches in medical image analysis: From detection to diagnosis. *Med Image Anal.* 2016 Oct;33:94-97. doi: 10.1016/j.media.2016.06.032. Epub 2016 Jun 23.
- [5] Cuevas-Rodriguez EO, Galvan-Tejada CE, Maeda-Gutiérrez V, Moreno-Chávez G, Galván-Tejada JI, Gamboa-Rosales H, Luna-García H, Moreno-Baez A, Celaya-Padilla JM. Comparative study of convolutional neural network architectures for gastrointestinal lesions classification. *PeerJ.* 2023 Mar 16;11:e14806. doi: 10.7717/peerj.14806.
- [6] Tan, Mingxing; LE, Quoc. Efficientnet: Rethinking model scaling for convolutional neural networks. In: *International conference on machine learning*. PMLR, 2019. p. 6105-6114.
- [7] He, K., Zhang, X., Ren, S., Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).
- [8] Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700-4708).
- [9] He, K., Gan, C., Li, Z., et al. (2022). Transformers in medical image analysis: A review. *Intelligent Medicine*.