SEGMENTAÇÃO E SIMILARIDADE ESTRUTURAL EM RADIOGRAFIAS DE TÓRAX INCERTAS

Leonardo de Souza Dunzer,Gilmario Barbosa dos Santos

INTRODUÇÃO

A radiografia de tórax é o exame de imagem mais solicitado na prática clínica mundial, sendo fundamental para o rastreio e o diagnóstico de patologias pulmonares. Sua interpretação, no entanto, é inerentemente subjetiva, dependendo da experiência do radiologista e das condições de aquisição da imagem (BRADY,2017). Bases de dados massivas, como o CheXpert, ampliaram o uso da Inteligência Artificial na triagem automatizada, mas introduziram também a classe de rótulos "incertos"(AZEVEDO; SANTOS, 2024), correspondente a casos em que o laudo clínico não permite confirmação binária da patologia. Diante disso, este trabalho teve como objetivo caracterizar a natureza estrutural dos casos rotulados como incertos no CheXpert, investigando se tais exames apresentam maior similaridade morfológica com referências saudáveis ou com referências patológicas.

DESENVOLVIMENTO

O estudo utilizou a base CheXpert, organizada em três conjuntos a partir da coluna No_Finding: Saudável (1.383 imagens), Doente (1.396 imagens) e Incerta (1.224 imagens). O pipeline metodológico compreendeu quatro etapas: (i) filtragem das radiografias em orientação frontal; (ii) segmentação pulmonar por meio da rede neural U-Net (RONNEBERGER et al., 2015), isolando o parênquima pulmonar e reduzindo interferências externas no diagnóstico auxiliado por computador (VAN GINNEKEN et al., 2006); (iii) padronização e normalização das imagens (escala de cinza, redimensionamento e normalização de intensidades); e (iv) cálculo do Índice de Similaridade Estrutural (SSIM) entre cada imagem incerta e as referências dos grupos saudável e doente, gerando as médias SSIM_saudável e SSIM_doente. A segmentação foi executada em resolução de 256×256 pixels e o cálculo do SSIM em 224×224 pixels, com o pipeline implementado em Python (TensorFlow, OpenCV, scikit-image, NumPy e Matplotlib) em ambiente com aceleração por GPU.

RESULTADOS

A análise de dispersão entre SSIM_saudável e SSIM_doente mostrou forte concentração dos pontos sobre a linha de identidade (Gráfico 1), indicando grau de similaridade praticamente idêntico entre os dois grupos de comparação. O coeficiente de correlação de Pearson confirmou essa tendência, atingindo $r = 0{,}9737$ ($p < 0{,}001$) sem segmentação e $r = 0{,}9653$ ($p < 0{,}001$) com segmentação, para a amostra de 1.224 radiografias incertas. Os boxplots da distribuição global revelaram medianas e amplitudes interquartílicas sobrepostas entre as classes, e a diferença absoluta concentrou-se em valores próximos de zero. A Tabela 1 apresenta o SSIM médio antes e após a segmentação pulmonar: observa-se que, embora a segmentação tenha removido elementos externos ao pulmão, ela não aumentou a separabilidade entre os grupos saudável e doente, evidenciando que a ambiguidade diagnóstica está associada a características intrínsecas do próprio tecido pulmonar, não capturadas por métricas globais de similaridade estrutural.

CONSIDERAÇÕES FINAIS

Em resumo, o trabalho teve como objetivo utilizar o SSIM como métrica para avaliar e entender a natureza dos casos tratados como incertos, em conjunto com ferramentas de segmentação de imagens

para trazer maior fidelidade à análise. Os objetivos propostos foram atingidos com êxito, e os resultados evidenciam uma forte sobreposição entre os valores de similaridade estrutural dos casos incertos em relação aos grupos saudável e doente, com uma correlação de Pearson superior a 0,96. Esse comportamento sugere que a incerteza diagnóstica registrada no CheXpert não expressa anormalidade nesses casos limítrofes. A etapa de segmentação reduziu interferências externas, mas não aumentou a separabilidade entre as classes, o que reforça a hipótese de que a ambiguidade diagnóstica está relacionada a características texturais de grão fino e a padrões incipientes, não capturados por avaliações globais. Como desdobramento, visto que métricas tradicionais apresentam limitações em imagens médicas, recomenda-se a exploração de outras métricas, como o LPIPS, que são mais sensíveis à semântica local e trazem uma nova visão aos resultados.

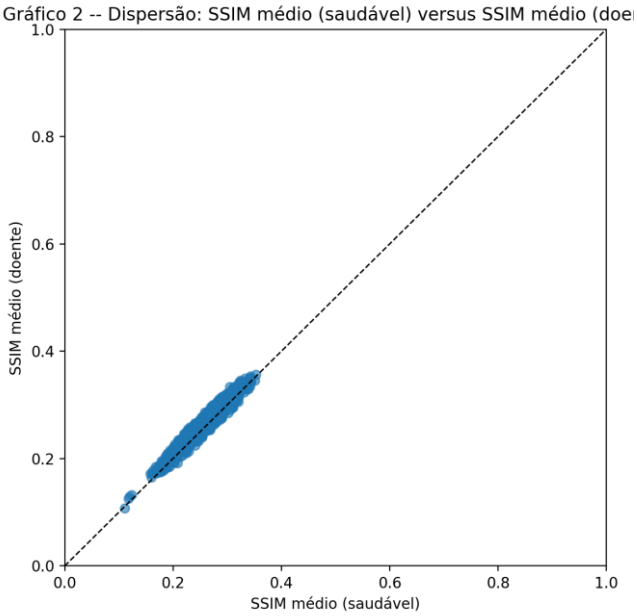
Palavras-chave: similaridade estrutural; segmentação semântica; U-Net; radiografia de tórax; inteligência artificial.

ANEXOS

Tabela 1. Comparação do SSIM médio antes e após a segmentação.

Comparação	Sem Segmentação	Com Segmentação
Incerto × Saudável	0,2626	0,5992
Incerto × Doente	0,2705	0,6093

Gráfico 1. Dispersão do SSIM médio (saudável) versus SSIM médio (doente) para os casos incertos



REFERÊNCIAS BIBLIOGRÁFICAS

AZEVEDO, A. A.; SANTOS, G. B. Characterization of Pneumonia Diagnostic Uncertainty: A Case Study on The CheXpert Dataset. Anais do Latinoware, pp. 1-10, 2024.

BRADY, A. P. Error and discrepancy in radiology: inevitable or avoidable? Insights into Imaging, 2017.
BREGER, A. et al. A study on the adequacy of common IQA measures for medical images. MICAD 2024, Springer, 2024.
IRVIN, J. et al. CheXpert: A large chest radiograph dataset with uncertainty labels and expert comparison. Proceedings of the AAAI Conference on Artificial Intelligence, 2019.
RONNEBERGER, O. et al. U-Net: Convolutional networks for biomedical image segmentation. Proceedings of MICCAI, 2015.
VAN GINNEKEN, B. et al. Segmentation of the lungs in chest radiographs as a first step in computer-aided diagnosis. Systems, Signals and Image Processing, 2006.
WANG, Z. et al. Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing, 2004.
ZHANG, R. et al. The Unreasonable Effectiveness of Deep Features as a Perceptual Metric. Proceedings of the CVPR, 2018.