

APRENDIZADO PROFUNDO NA DETECÇÃO DE COVID-19 EM IMAGENS TORÁCICAS: REVISÃO CRÍTICA

Arthur Rodrigues Sobrinho¹; Evair Josino da Silva²

arthurrs601@gmail.com

Área Temática: Temas livres em Medicina

RESUMO

Introdução: A pandemia de COVID-19 impulsionou a busca por métodos diagnósticos complementares capazes de ampliar a rapidez da triagem e reduzir falhas de detecção, sobretudo em contextos de alta demanda assistencial. Nesse cenário, a análise automatizada de imagens torácicas por aprendizado profundo ganhou destaque como estratégia de apoio à decisão clínica. **Objetivo:** Analisar criticamente a aplicabilidade de modelos de aprendizado profundo para detecção de COVID-19 em imagens de tomografia computadorizada e radiografia de tórax, com ênfase no desempenho do Deep-CT-Net e do Deep-CXR-Net. **Metodologia:** Trata-se de uma revisão bibliográfica de caráter descritivo e abordagem qualitativa, realizada nas bases PubMed, SciELO e LILACS, considerando estudos publicados nos últimos 10 anos, nos idiomas português e inglês. A estratégia de busca foi construída com descritores em DeCS e MeSH relacionados a COVID-19, aprendizado profundo, tomografia computadorizada, radiografia de tórax e diagnóstico por imagem, combinados pelos operadores booleanos AND e OR. Foram incluídos artigos disponíveis na íntegra, com foco em modelos de inteligência artificial aplicados à detecção de COVID-19 em imagens torácicas; excluíram-se relatos de caso, estudos duplicados, textos incompletos e trabalhos sem descrição adequada de desempenho diagnóstico. **Resultados e Discussão:** A literatura indica que o Deep-CT-Net alcançou AUC de 0,886 e sensibilidade de 0,858 no CT-COV19, enquanto o Deep-CXR-Net apresentou AUC de 0,9839 e sensibilidade de 0,9824 em radiografias, sugerindo elevado potencial para triagem. Entretanto, a interpretação desses achados exige cautela, pois persistem limitações relacionadas à validação externa, ao risco de vieses de base de dados e à heterogeneidade dos cenários clínicos. **Conclusão:** Os modelos analisados apresentam desempenho promissor como ferramentas auxiliares de triagem, mas sua incorporação à prática clínica depende de validação multicêntrica, prospectiva e metodologicamente robusta.

Palavras-chave: COVID-19; Inteligência artificial; Diagnóstico por imagem; Aprendizado profundo.

1 INTRODUÇÃO

A COVID-19, causada pelo SARS-CoV-2, impôs um desafio diagnóstico expressivo aos sistemas de saúde, principalmente nos períodos de maior transmissão comunitária e sobrecarga assistencial. Embora a reação em cadeia da polimerase com transcrição reversa em tempo real (RT-PCR) tenha sido amplamente adotada como referência laboratorial, limitações relacionadas ao tempo de processamento, à disponibilidade e à sensibilidade em fases iniciais da infecção estimularam a busca por métodos complementares de avaliação (Fang *et al.*, 2020; Ai *et al.*, 2020).

No âmbito da imagem torácica, a tomografia computadorizada (TC) e a radiografia de tórax passaram a desempenhar papel relevante na investigação de pacientes com suspeita

clínica de acometimento pulmonar. Entre os achados mais frequentemente descritos na COVID-19 destacam-se opacidades em vidro fosco, consolidações de distribuição periférica e bilateral e espessamento septal, padrões que podem contribuir para a suspeição diagnóstica em contextos apropriados (Bernheim *et al.*, 2020; Shi *et al.*, 2020). Ainda assim, a interpretação dessas imagens depende de experiência profissional, sofre influência da variabilidade interobservador e, isoladamente, não substitui a confirmação etiológica.

Nesse contexto, os modelos de aprendizado profundo aplicados à análise de imagens médicas foram propostos como ferramentas de apoio capazes de acelerar a triagem, padronizar a leitura e ampliar a sensibilidade para padrões sutis de doença. Entre os sistemas publicados no período, destacam-se abordagens como o COVNET e modelos multinacionais baseados em TC, que evidenciaram o potencial da inteligência artificial, mas também a heterogeneidade metodológica das bases empregadas (Li *et al.*, 2020; Harmon *et al.*, 2020). Paralelamente, a literatura passou a demonstrar que muitos modelos desenvolvidos durante a pandemia apresentaram fragilidades metodológicas, como bases de dados pouco representativas, validação externa insuficiente e risco de superestimação do desempenho, o que limita sua translação imediata para a prática clínica (Roberts *et al.*, 2021).

Entre as propostas publicadas, destaca-se o estudo de Dialameh *et al.* (2022), que descreveu o dataset CT-COV19 e os modelos Deep-CT-Net e Deep-CXR-Net para detecção de COVID-19 em TC e radiografia de tórax, com arquitetura baseada em DenseNet-121 e mecanismo de atenção piramidal (Huang *et al.*, 2017; Dialameh *et al.*, 2022). Considerando a relevância desse trabalho e a necessidade de interpretação crítica de seus achados, o presente resumo expandido tem como objetivo analisar, à luz da literatura, o desempenho, as potencialidades e as limitações desses modelos de aprendizado profundo no diagnóstico assistido por imagem da COVID-19.

2 METODOLOGIA

Trata-se de uma revisão bibliográfica de caráter descritivo e abordagem qualitativa. A busca foi realizada nas bases PubMed, SciELO e LILACS, contemplando publicações dos últimos 10 anos, nos idiomas português e inglês. A estratégia foi construída a partir do tema e do título do trabalho, com descritores pesquisados em DeCS e MeSH, incluindo “COVID-19”, “deep learning”, “artificial intelligence”, “computed tomography”, “chest radiography” e “diagnostic imaging”, combinados pelos operadores booleanos AND e OR.

Foram incluídos artigos disponíveis na íntegra que abordassem modelos de aprendizado profundo aplicados à detecção de COVID-19 em imagens torácicas, com descrição metodológica mínima do conjunto de dados, da arquitetura proposta e de métricas

de desempenho, especialmente sensibilidade, precisão, F-measure, acurácia ou área sob a curva ROC (AUC). Também foram priorizados estudos com relevância clínica e metodológica, bem como artigos voltados à interpretação crítica da qualidade dos modelos desenvolvidos para radiografia e TC de tórax.

Foram excluídos relatos de caso, editoriais, textos incompletos, estudos duplicados, publicações fora do recorte temporal e trabalhos com descrição insuficiente do método ou sem dados que permitissem interpretação do desempenho diagnóstico. A análise dos dados foi conduzida de forma qualitativa, por síntese de evidências em saúde, com ênfase na relação entre desempenho algorítmico, utilidade clínica, interpretabilidade e limitações de validação externa. Nessa síntese, o estudo de Dialameh *et al.* (2022) foi considerado o eixo central de análise, sendo contextualizado por trabalhos sobre achados radiológicos da COVID-19 e sobre limitações metodológicas de modelos de inteligência artificial aplicados à imagem torácica.

3 RESULTADOS E DISCUSSÃO

A literatura selecionada mostra que os métodos de aprendizado profundo aplicados à imagem torácica alcançaram desempenho expressivo durante a pandemia, sobretudo quando treinados em bases de dados relativamente amplas e com desenho voltado à triagem. Nesse conjunto, o estudo de Dialameh *et al.* (2022) ocupa posição de destaque por reunir dois elementos centrais: a criação do CT-COV19, descrito como um banco público com mais de 13 mil imagens de TC provenientes de mais de 1.000 indivíduos, e o desenvolvimento de uma arquitetura baseada em DenseNet-121 associada a uma camada de atenção piramidal pixel a pixel, concebida para aumentar a detecção de padrões discretos de acometimento pulmonar.

Do ponto de vista metodológico, a proposta apresenta aspectos meritórios. O CT-COV19 foi estruturado com diversidade etária, de sexo e de origem regional, além de incluir imagens de COVID-19, outras pneumonias e pulmões normais, o que amplia, em tese, a heterogeneidade do treinamento. A incorporação da atenção piramidal sobre características de alto nível procura direcionar a rede para regiões mais informativas da imagem, favorecendo a identificação de sinais sutis. Essa escolha dialoga com a necessidade clínica de reduzir falsos negativos em cenários nos quais alterações precoces, como opacidades em vidro fosco ou consolidações discretas, podem passar despercebidas em leituras rápidas ou em contextos de alta demanda (Bernheim *et al.*, 2020; Shi *et al.*, 2020).

Nos resultados reportados, o Deep-CT-Net alcançou AUC de 0,886 no CT-COV19, com sensibilidade de 0,858, precisão de 0,720 e F-measure de 0,783, superando modelos comparativos como COVNET, Hybrid-3D, Unet++ e DL-system em sensibilidade e AUC no

mesmo cenário experimental (Dialameh *et al.*, 2022). Em um segundo benchmark, o COVID-CT, o modelo atingiu AUC de 0,899 e sensibilidade de 0,905; quando testado nesse conjunto independente sem ajuste fino, manteve AUC de 0,920 e F-measure de 0,801, sugerindo capacidade de generalização além da base de treinamento principal. Esses achados reforçam que a ênfase em sensibilidade pode ser clinicamente vantajosa para triagem, já que, em doenças infecciosas de elevada transmissibilidade, a redução de falsos negativos tem valor estratégico.

Entretanto, a interpretação desse desempenho precisa ser feita com prudência. Sensibilidade elevada não equivale, por si só, a utilidade clínica plena, sobretudo quando a precisão é mais modesta e quando o desempenho é calculado em bases públicas específicas, potencialmente sujeitas a vieses de seleção, diferenças de protocolo de aquisição e desequilíbrio entre classes. Roberts *et al.* (2021) alertam que muitos modelos de aprendizado de máquina para COVID-19 apresentaram limitações importantes, entre elas ausência de validação prospectiva, risco de vazamento de dados, inadequação na representação do espectro clínico e dificuldade de reprodução em cenários reais. Assim, embora os resultados de Dialameh *et al.* (2022) sejam superiores aos comparadores testados, eles não eliminam a necessidade de validação externa multicêntrica e prospectiva antes da adoção assistencial ampla.

A extensão do modelo para radiografias de tórax, denominada Deep-CXR-Net, também merece destaque. Com base em aprendizado por transferência, o sistema combinou redes DenseNet-121 previamente treinadas em bases voltadas a doenças pulmonares gerais com uma terceira rede ajustada especificamente para COVID-19. No dataset ieec8023, o modelo atingiu AUC de 0,9839, sensibilidade de 0,9824, precisão de 0,9872 e F-measure de 0,9848, superando os modelos comparativos avaliados pelos autores (Dialameh *et al.*, 2022). Em termos operacionais, esse achado é relevante porque a radiografia é mais acessível, menos onerosa e mais difundida do que a TC em muitos serviços, o que pode favorecer estratégias de triagem em larga escala.

Ainda assim, não é adequado concluir que a radiografia seja intrinsecamente superior à TC para detecção de COVID-19 apenas porque o Deep-CXR-Net apresentou métricas mais altas. Essa diferença pode decorrer da composição dos conjuntos de dados, do padrão de anotação, da complexidade distinta entre imagens bidimensionais e volumétricas e da própria distribuição dos casos utilizados no treinamento e no teste. Além disso, a literatura clínica demonstra que a TC pode apresentar maior sensibilidade em determinados contextos de suspeição, embora sua especificidade seja insuficiente para uso isolado como método

confirmatório (Fang *et al.*, 2020). Dessa forma, os achados devem ser lidos como evidência do potencial do algoritmo dentro de um cenário experimental específico, e não como prova definitiva de superioridade entre modalidades.

Outro ponto favorável do estudo é a preocupação com interpretabilidade. Os mapas de ativação de classe, descritos a partir da técnica Grad-CAM, sugerem que os modelos focalizaram regiões pulmonares anatomicamente compatíveis com áreas de opacidade e consolidação, o que aumenta a plausibilidade clínica da decisão algorítmica (Selvaraju *et al.*, 2017; Dialameh *et al.*, 2022). Além disso, a dispensa de etapas complexas de pré-processamento, como segmentação pulmonar manual, representa ganho operacional importante, pois reduz custo computacional e potencialmente favorece integração a fluxos de trabalho mais ágeis.

Apesar dessas virtudes, persistem limitações substantivas. O recorte temporal dos dados concentra-se no período inicial da pandemia, sem demonstração clara de desempenho diante de variantes posteriores, mudanças no perfil populacional ou transformações no padrão de encaminhamento para imagem. Também não há demonstração de impacto clínico direto, como redução de tempo diagnóstico, melhora de desfechos ou comparação com leitura radiológica especializada em ambiente real. Em síntese, os resultados sugerem que Deep-CT-Net e Deep-CXR-Net são ferramentas promissoras de apoio diagnóstico e triagem, mas a robustez metodológica exigida para incorporação clínica depende de validação adicional, transparência reprodutiva e análise contínua do risco de vieses.

4 CONCLUSÃO

A análise da literatura evidencia que os modelos Deep-CT-Net e Deep-CXR-Net apresentam desempenho promissor para detecção de COVID-19 em imagens torácicas, com destaque para a elevada sensibilidade e para o uso de mecanismos de atenção capazes de ampliar a identificação de padrões pulmonares discretos. O estudo de Dialameh *et al.* (2022) representa contribuição relevante ao combinar um banco público relativamente amplo com arquitetura baseada em DenseNet-121 e atenção piramidal, demonstrando resultados superiores aos de modelos comparativos em seus cenários experimentais.

Do ponto de vista clínico e científico, contudo, esses achados devem ser interpretados como evidência de potencialidade, e não como validação definitiva para uso autônomo na prática assistencial. A utilidade desses sistemas parece mais consistente como ferramenta auxiliar de triagem e apoio à decisão, sempre integrada à avaliação clínica, laboratorial e radiológica. Assim, conclui-se que a incorporação segura de modelos de aprendizado profundo no diagnóstico por imagem da COVID-19 requer validação externa

multicêntrica, estudos prospectivos e monitoramento rigoroso de vieses metodológicos, de modo a assegurar aplicabilidade real, reprodutibilidade e confiabilidade diagnóstica.

REFERÊNCIAS BIBLIOGRÁFICAS

AI, T. *et al.* Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. *Radiology*, [S. l.], v. 296, n. 2, p. E32-E40, 2020. DOI: 10.1148/radiol.2020200642.

BERNHEIM, A. *et al.* Chest CT findings in coronavirus disease-19 (COVID-19): relationship to duration of infection. *Radiology*, [S. l.], v. 295, n. 3, p. 685-691, 2020. DOI: 10.1148/radiol.2020200463.

DIALAMEH, M.; HAMZEH, A.; RAHMANI, H.; RADMARD, A. R.; DIALAMEH, S. Proposing a novel deep network for detecting COVID-19 based on chest images. *Scientific Reports*, [S. l.], v. 12, art. 3116, 2022. DOI: 10.1038/s41598-022-06802-7.

FANG, Y. *et al.* Sensitivity of chest CT for COVID-19: comparison to RT-PCR. *Radiology*, [S. l.], v. 296, n. 2, p. E115-E117, 2020. DOI: 10.1148/radiol.2020200432.

HARMON, S. A. *et al.* Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets. *Nature Communications*, [S. l.], v. 11, art. 4080, 2020. DOI: 10.1038/s41467-020-17971-2.

HUANG, G.; LIU, Z.; VAN DER MAATEN, L.; WEINBERGER, K. Q. Densely connected convolutional networks. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 2017, Honolulu. Proceedings [...]. Piscataway: IEEE, 2017. p. 4700-4708. DOI: 10.1109/CVPR.2017.243.

LI, L. *et al.* Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT. *Radiology*, [S. l.], v. 296, n. 2, p. E65-E71, 2020. DOI: 10.1148/radiol.2020200905.

ROBERTS, M. *et al.* Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. *Nature Machine Intelligence*, [S. l.], v. 3, n. 3, p. 199-217, 2021. DOI: 10.1038/s42256-021-00307-0.

SELVARAJU, R. R. *et al.* Grad-CAM: visual explanations from deep networks via gradient-based localization. In: IEEE INTERNATIONAL CONFERENCE ON COMPUTER VISION, 2017, Venice. Proceedings [...]. Piscataway: IEEE, 2017. p. 618-626. DOI: 10.1109/ICCV.2017.74.

SHI, H. *et al.* Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study. *The Lancet Infectious Diseases*, [S. l.], v. 20, n. 4, p. 425-434, 2020. DOI: 10.1016/S1473-3099(20)30086-4.