

Inteligência artificial aplicada ao diagnóstico de câncer por exames de imagem

Artificial intelligence applied to cancer diagnosis through image exams

Inteligencia artificial aplicada al diagnóstico del cáncer a través de exámenes de imagen

Leticia Brum Croffi¹, Wagner Judice¹, Silvia Boschi¹, Silvia Martini¹.

RESUMO

Objetivo: Revisar a aplicação da Inteligência Artificial (IA) no diagnóstico de câncer através de exames de imagem, analisando tendências e avanços recentes na intersecção entre IA e medicina diagnóstica. **Métodos:** Foi realizada uma pesquisa bibliográfica em bases de dados eletrônicas, selecionando artigos científicos, revisões sistemáticas e meta-análises publicados entre 2017 e 2023. Os estudos incluíram algoritmos de IA, técnicas de aprendizado de máquina e redes neurais aplicadas a exames de imagem para detecção, classificação e diagnóstico de câncer. **Resultados:** A análise focou em modelos de reconhecimento de imagem para diagnóstico de câncer, priorizando métricas de sensibilidade e especificidade. Foram destacados estudos que compararam o desempenho de radiologistas com sistemas de IA mostrando que em alguns casos a IA superou os profissionais e, em outros, melhorou significativamente o desempenho dos radiologistas quando usada como assistência. **Considerações finais:** A IA mostrou-se uma ferramenta promissora no diagnóstico de câncer por imagem. A combinação de IA com dados clínicos pode melhorar as métricas de diagnóstico. Limitações incluem a qualidade e quantidade de imagens para treinamento, mas novas tecnologias como IA Generativa estão surgindo para superar esses desafios.

Palavras-chave: Câncer, Diagnóstico, Aprendizado profundo, Inteligência Artificial Generativa.

ABSTRACT

Objective: To review the application of Artificial Intelligence (AI) in cancer diagnosis through imaging tests, analyzing trends and recent advances at the intersection of AI and diagnostic medicine. **Methods:** A bibliographic search was conducted in electronic databases, selecting scientific articles, systematic reviews, and meta-analyses published between 2017 and 2023. The studies included AI algorithms, machine learning techniques, and neural networks applied to imaging tests for cancer detection, classification, and diagnosis. **Results:** The analysis focused on image recognition models for cancer diagnosis, prioritizing sensitivity and specificity metrics. Studies that compared the performance of radiologists with AI systems were highlighted, showing that in some cases AI outperformed the professionals and, in others, significantly improved the performance of radiologists when used as assistance. **Final considerations:** AI has proven to be a promising tool in cancer diagnosis through imaging. The combination of AI with clinical data may improve diagnostic metrics. Limitations include the quality and quantity of images for training, but new technologies such as Generative AI are emerging to overcome these challenges.

Keywords: Cancer, Diagnosis, Deep learning, Generative Artificial Intelligence.

¹ Universidade de Mogi das Cruzes (UMC), Mogi das Cruzes – SP.

RESUMEN

Objetivo: Revisar la aplicación de la Inteligencia Artificial (IA) en el diagnóstico de cáncer a través de exámenes de imagen, analizando tendencias y avances recientes en la intersección entre IA y medicina diagnóstica. **Métodos:** Se realizó una búsqueda bibliográfica en bases de datos electrónicas, seleccionando artículos científicos, revisiones sistemáticas y metaanálisis publicados entre 2017 y 2023. Los estudios incluyeron algoritmos de IA, técnicas de aprendizaje automático y redes neuronales aplicadas a exámenes de imagen para la detección, clasificación y diagnóstico de cáncer. **Resultados:** El análisis se centró en modelos de reconocimiento de imagen para el diagnóstico de cáncer, priorizando métricas de sensibilidad y especificidad. Se destacaron estudios que compararon el rendimiento de radiólogos con sistemas de IA, mostrando que en algunos casos la IA superó a los profesionales y, en otros, mejoró significativamente el rendimiento de los radiólogos cuando se utilizó como asistencia. **Consideraciones finales:** La IA se ha mostrado como una herramienta prometedora en el diagnóstico de cáncer por imagen. La combinación de IA con datos clínicos puede mejorar las métricas de diagnóstico. Las limitaciones incluyen la calidad y cantidad de imágenes para entrenamiento, pero nuevas tecnologías como la IA Generativa están surgiendo para superar estos desafíos.

Palabras clave: Cáncer, Diagnóstico, Aprendizaje profundo, Inteligencia Artificial Generativa.

INTRODUÇÃO

O termo Inteligência Artificial (IA) foi criado em uma conferência em Dartmouth de 1956 por John McCarthy e foi lá que toda a área de pesquisa se originou. Kaplan A e Haenlein M (2019) definiram a IA como "a capacidade de processar dados externos de forma sistemática e aprender com eles para alcançar metas e tarefas específicas" (PUPPE F, 1992).

A IA passa a ser aplicada na Medicina no início da década de 70, seu objetivo era aumentar a eficiência do diagnóstico e tratamento médico com este auxílio de sistemas de IA. A partir da década de 80, Liu R, et al. (2020) afirmam que a IA na Medicina poderia ser dividida aproximadamente em 4 estágios: a infância como primeiro estágio (1980) com árvores de decisão e o desenvolvimento das redes neurais; o segundo, a adolescência (1990) com os algoritmos das Máquina de Vetores de Suporte (Support Vector Machine - SVM); a maturidade no terceiro estágio (2000) quando o conceito de Deep Learning ou "Aprendizado Profundo" foi proposto, consistindo em uma rede neural com várias camadas; e por fim o quarto estágio, o período de maturação (2010) no qual o evento mais recente seria a publicação da Google da nova arquitetura Transformer de redes neurais com os mecanismos de atenção ou autoatenção em 2017, arquitetura essa que deu origem ao ChatGPT entre outros, mudando o atual paradigma da IA (VASWANI A, et al., 2017).

O reconhecimento de imagens é a tecnologia de processamento e análise de imagens por meio de computadores. É uma tecnologia importante no campo da IA de treinamento Supervisionado, que utiliza exames de imagens com diagnóstico médico ou laboratorial para treinar o modelo, o qual se baseia no Aprendizado Profundo. Segundo Masood A, et al. (2018), o processo de reconhecimento pode ser dividido em cinco etapas, sendo elas: (1) processamento de entrada, (2) pré-processamento de imagens, (3) extração de imagens, (4) construção de classificadores e (5) produção de saída ().

Atualmente, a tecnologia de reconhecimento de imagens pode ser aplicada em muitas áreas médicas e existem muitos equipamentos de exames que já a utilizam. O câncer cervical, por exemplo, é uma das quatro principais causas de morte em mulheres, e é causado principalmente por infecção pelo papilomavírus humano (HPV). As pacientes não apresentam sintomas óbvios nos estágios iniciais. Embora muitos tratamentos estejam disponíveis para pacientes com câncer cervical, como cirurgia, radioterapia e quimioterapia, o prognóstico das pacientes depende muito do diagnóstico precoce do câncer.

O Aprendizado Profundo, no reconhecimento de lesões em imagens do colo do útero, poderia ajudar os médicos no diagnóstico precoce do câncer cervical com uma taxa de precisão de aproximadamente 90%. Além disso, sistemas de reconhecimento de imagens poderiam auxiliar médicos com uma segunda opinião em situações mais complexas, e assim produzir diagnósticos mais precisos (LIN H, et al., 2021; LIU J, et al., 2018).

O Aprendizado Profundo desempenha um papel importante na aplicação da tecnologia de reconhecimento de imagens para identificar lesões. Kooi T, et al. (2017) utilizaram redes neurais convolucionais (Convolutional Neural Networks - CNNs) em um estudo para identificar lesões malignas na mama. Este método teve um desempenho superior ao sistema mais avançado em detecção assistida por computador, uma vez que a precisão de reconhecimento da CNN foi maior (LIU F, et al., 2018).

O reconhecimento de imagens para diagnóstico com o auxílio da IA só se torna possível por meio da imagem médica. A imagem médica é o processo de representação visual da estrutura e função de diferentes tecidos e órgãos do corpo humano para fins clínicos e para o estudo detalhado da anatomia e fisiologia normais e anormais do corpo na ciência médica. As técnicas de imagem médica são usadas para mostrar estruturas internas sob a pele e os ossos, bem como para diagnosticar anormalidades e tratar doenças. A imagem médica transformou-se em uma ciência na área da saúde.

É uma parte importante da imagem biológica e inclui a radiologia, que utiliza tecnologias de imagem como radiografia de raios X, tomografia computadorizada por raios X (CT), endoscopia, ressonância magnética (RM), espectroscopia de ressonância magnética (ERM), tomografia por emissão de pósitrons (PET), termografia, fotografia médica, imagem elétrica de fonte (ESI), mamografia digital, imagem tátil, imagem magnética de fonte (MSI), imagem óptica médica, tomografia computadorizada por emissão de fóton único (SPECT) e tomografia computadorizada por impedância elétrica (EIT) (HUSSAIN S, et al., 2022).

Mas, ainda existem limitações, devido aos diferentes equipamentos utilizados em diversos hospitais, a resolução das imagens adquiridas pode variar bastante, o que pode impactar a eficiência do modelo. Um modelo de Aprendizado Profundo também necessita de uma grande quantidade de imagens e requer muito poder de processamento. Problemas estes que a recente arquitetura do modelo Transformer desenvolvida pelo Google promete resolver ao melhorar a eficiência computacional e ao diminuir a necessidade na quantidade de imagens com seu mecanismo de atenção ou autoatenção (Vaswani A, et al., 2017). Este trabalho teve então como objetivo analisar as tendências e avanços recentes na intersecção entre IA e medicina diagnóstica.

MÉTODOS

Este trabalho consiste em uma revisão integrativa da aplicação da Inteligência Artificial (IA) no diagnóstico de câncer por meio de exames de imagem.

A pesquisa foi conduzida com o objetivo de analisar as tendências e avanços recentes na intersecção entre IA e medicina diagnóstica. A revisão bibliográfica foi realizada em bases de dados eletrônicas, incluindo PubMed, arXiv, IEEE Xplore, e Google Scholar.

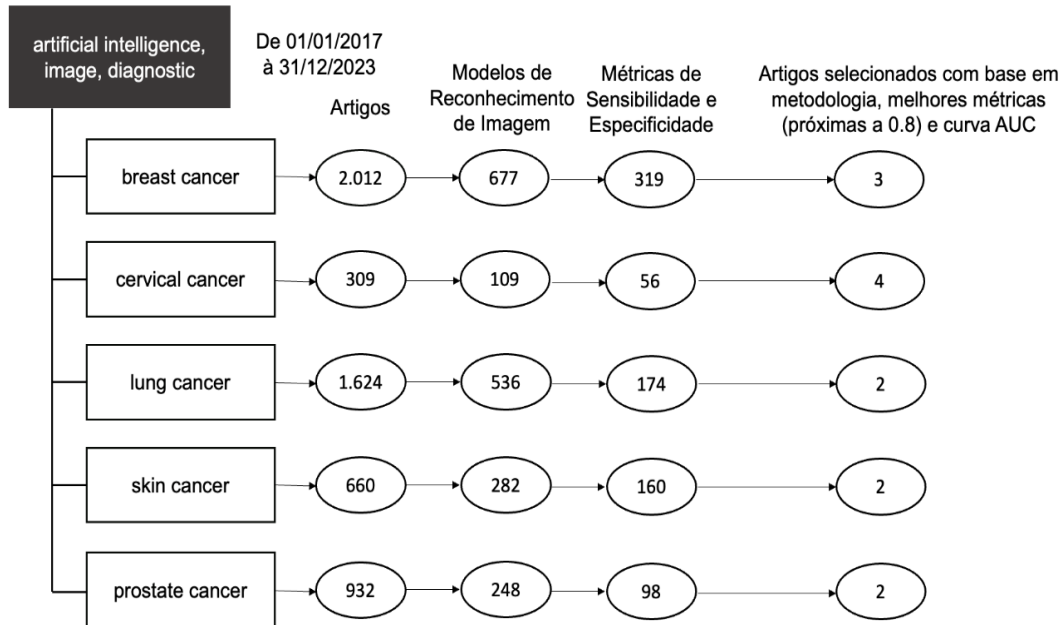
Foram considerados artigos científicos, revisões sistemáticas e meta-análises publicados entre 2017 e 2023 que abordavam o uso de IA em diagnósticos oncológicos por meio de imagens médicas. A escolha do intervalo de tempo se deve ao fato de que em 2017 foi quando a última competição da ImageNet foi realizada porque julgaram que o estado da arte com redes neurais convolucionais para reconhecimento de imagens já havia sido alcançado, também é o ano em que o artigo do algoritmo Transformer foi publicado provocando a mudança de paradigma (MISHKIN D, et al., 2017). Os artigos para análise incluíram estudos que exploraram algoritmos de IA, técnicas de aprendizado de máquina e redes neurais aplicadas a exames de imagem para detecção, classificação e diagnóstico de câncer.

Utilizou-se os descritores (palavras-chave): artificial intelligence e image e diagnostic alternando com os seguintes descritores breast cancer ou cervical cancer ou lung cancer ou skin cancer ou prostate cancer em inglês para a busca, pois as terminologias e tecnologias originais são em inglês. Estudos que não continham as métricas de sensibilidade e especificidade ou não utilizavam modelos de Deep Learning ou Machine Learning para reconhecimento de imagem foram excluídos do estudo.

A inclusão dos estudos se baseou em estudos que demonstraram resultados promissores ou inovações significativas na área. A análise dos dados concentrou-se na avaliação crítica das metodologias utilizadas nos

estudos selecionados, destacando as abordagens específicas de IA, conjunto de dados empregado, métricas de sensibilidade e especificidade para desempenho e conclusões relevantes para o diagnóstico de câncer por imagem. A **Figura 1** descreve o processo de seleção de artigos e os artigos encontrados.

Figura 1 – Artigos revisados e selecionados para a revisão.



Fonte: Croffi LB, et al., 2024.

A revisão está sujeita a limitações inerentes à disponibilidade de literatura e à constante evolução da tecnologia. Ademais, a interpretação dos resultados também depende da qualidade e metodologia dos estudos revisados.

RESULTADOS

Foram analisados resultados e métricas de modelos de reconhecimento de imagem desenvolvidos para diagnóstico de câncer em áreas específicas, selecionados pela maior incidência de casos e pela utilização de redes neurais de Aprendizado Profundo e das métricas de sensibilidade e especificidade. O Instituto Nacional do Câncer (INCA) lançou recentemente uma estimativa de 704 mil casos de câncer por ano no Brasil até 2025, sendo a neoplasia maligna de maior incidência a de pele não melanoma (31,3%), seguida de tumores malignos de mama (10,5%), próstata (10,2%), cólon e reto (6,5%), pulmão (4,6%) e estômago (3,1%).

No entanto, em regiões menos desenvolvidas, de menor IDH, o câncer cervical ocupa o terceiro lugar para mulheres no lugar do câncer colorretal, inclusive estudos já foram conduzidos analisando essa discrepância mostrando que em lugares menos desenvolvidos as mulheres eram mais resistentes a fazer o exame assim como não há muito pouca ou nenhuma infraestrutura (SANTOS MO, et al., 2023; ARBYN M, et al., 2020). As métricas de sensibilidade e especificidade foram priorizadas por serem consideradas altamente relevantes na área médica por fornecerem informações cruciais sobre o desempenho de testes diagnósticos, procedimentos de triagem e modelos preditivos em termos de sua capacidade de identificar corretamente as condições de saúde (JUNIOR GBV, et al., 2022).

Sensibilidade

Também conhecida como Recall. A sensibilidade de um teste médico refere-se à sua capacidade de identificar corretamente os indivíduos que têm a doença (verdadeiros positivos). Um teste com alta sensibilidade é importante para garantir que poucos diagnósticos críticos como câncer sejam perdidos (falsos negativos).

Especificidade

A especificidade de um teste é a sua capacidade de identificar corretamente os indivíduos que não têm a doença (verdadeiros negativos). Um teste com alta especificidade é importante para evitar o diagnóstico incorreto de pessoas saudáveis como doentes (falsos positivos). Isso é especialmente importante para evitar tratamentos desnecessários ou invasivos e a ansiedade associada a um diagnóstico incorreto.

AUC

Area Under the Curve, se refere à precisão do modelo. Um modelo cujas previsões estão 100% erradas tem uma AUC de 0, enquanto um modelo cujas previsões são 100% corretas tem uma AUC de 1. A curva AUC é derivada da curva ROC (Receiver Operating Characteristic) que utiliza a taxa de verdadeiro positivo e a taxa de falso positivo.

A **Tabela 1** ilustra uma síntese dos melhores resultados encontrados na pesquisa de artigos de modelos de reconhecimento de imagem para diagnóstico de câncer, assim como outras revisões sistemáticas. Os modelos listados foram selecionados com base em 2 métricas principais: sensibilidade e especificidade. Quanto mais próximo de 1 (ou 100%) melhor a assertividade do modelo. Por exemplo, uma sensibilidade de 0,90 indicaria que de cada 10 exames de imagem com diagnóstico positivo para o câncer, o algoritmo erraria 1 diagnóstico ao informar que o paciente não está com câncer, um falso negativo.

Já para uma especificidade de 0,90, de cada 10 exames de imagem com diagnóstico negativo, o modelo estaria errando ao diagnosticar o paciente informando que ele está com câncer, sem na realidade ele estar com câncer, um falso positivo.

Tabela 1 – Síntese das melhores métricas encontradas em modelos de reconhecimento de imagem para diagnóstico de câncer em artigos.

Neoplasia Maligna	Exame de Imagem	Modelo	Sensibilidade	Especificidade	AUC	Artigo
Pulmão	Tomografia Computadorizada	DFCNet	0,837	0,962	n/a	Maood A, et al., 2018
Pulmão	Tomografia Computadorizada	Radiomics + DCNN	0,940	0,930	0,950	Zhang X, et al., 2023
Não Melanoma	Imagem Dermatoscópica	EfficientNetB4	0,667	0,987	0,990	Huang HW, et al., 2021
Não Melanoma	Imagem Dermatoscópica	CNN	0,84	0,98	0,971	Kuo KM, et al., 2023
Mama	Mamografia	ResNet50	0,839	0,864	0,945	Lotter W, et al., 2021
Mama	Mamografia	CNN**	0,819	0,966	0,956	Salim M, et al., 2020
Mama	Mamografia	CNN**	0,914	0,860	0,959	Kim HE, et al., 2020
Colo do Útero	Citologia Cervical	VGG-19	0,897	0,892	0,950	Sornapudi S, et al., 2019
Colo do Útero	Citologia Cervical	ResNet50	0,992	0,997	0,996	Benyes YK, et al., 2022
Colo do Útero	Citologia Cervical	DPNet	0,907	0,800	0,925	Lin H, et al., 2021
Colo do Útero	Citologia Cervical	CNN**	0,957	0,847	0,940	Holmström O, et al., 2021
Próstata	Ressonância Magnética	MulmoUNet	0,727	0,955	n/a	Matsuoka Y, et al., 2023
Próstata	Ressonância Magnética	InceptionV1	0,815	0,820	0,889	Khosravi P, et al., 2021

Nota: ** modelos comerciais.

Fonte: Croffi LB, et al., 2024.

Alguns desses modelos comparam a performance do diagnóstico dado por um radiologista e o mesmo diagnóstico dado por um modelo de IA de forma independente. Os sistemas autônomos têm o potencial de filtrar exames provavelmente normais, diminuindo a carga de trabalho dos radiologistas. Já os sistemas de assistência são utilizados durante a interpretação dos exames.

As métricas de sensibilidade e especificidade podem ser aplicadas também a diagnósticos fornecidos por profissionais da saúde, técnicos e médicos, não apenas a sistemas de computador. Assim como o computador erra o diagnóstico, o profissional de saúde também pode errar. Alguns sinais de tumores em exames de imagem podem muitas vezes passarem despercebidos ao olho humano.

Ao mesmo tempo, o modelo de computador para reconhecimento de imagem sempre será tão bom quanto os médicos e profissionais de saúde que alimentam o modelo com exemplos de seus diagnósticos. Para melhorar o sistema, exames adicionais podem ser acrescentados como uma biópsia para confirmar diagnósticos. A partir desses exames adicionais o modelo passaria a procurar padrões para identificar problemas na imagem que podem ter passados despercebidos ao olho humano.

Tabela 2 – Comparação relativa entre profissionais de saúde e sistemas de IA nas melhores performances de diagnóstico.

Artigo	Exame de Imagem	Comparação
Kim HE, et al., 2020	Mamografia	Radiologistas x Sistema de IA: em média melhora de 16,13% na sensibilidade e 14,04% na especificidade para o sistema de IA autônomo
McKinney SM, et al., 2020	Mamografia	Radiologistas x Sistema de IA: nos EUA em média melhora de 9,4% na sensibilidade e 4,7% na especificidade para o sistema de IA autônomo
Lotter W, et al., 2021	Mamografia	Radiologistas x Sistema de IA: em média melhora de 14,2% na sensibilidade e 24% na especificidade para o sistema de IA autônomo

Fonte: Croffi LB, et al., 2024.

Como ilustra a **Tabela 2** em muitos casos os sistemas de IA melhoram a performance do diagnóstico de câncer. Kim HE, et al. (2020) relataram uma avaliação independente de um sistema de IA comercial (Lunit INSIGHT MMG) em um estudo de grande escala que incluiu 5 instituições na Coreia do Sul, EUA e Reino Unido, demonstrando alto desempenho diagnóstico.

O sistema de IA superou todos os radiologistas quando usado de forma autônoma. Além disso, quando usados simultaneamente, observou-se uma melhoria significativa no desempenho dos radiologistas com a ajuda da IA. Lotter W, et al. (2021) usaram imagens de diferentes estados dos EUA e um conjunto de dados de um hospital urbano chinês, com resultados promissores, onde o modelo alcançou um aumento absoluto na sensibilidade de 14,2% e na especificidade de 24% em comparação com a sensibilidade média alcançada pelo radiologista.

Salim M, et al. (2020) realizaram uma avaliação externa de três sistemas de IA comerciais para detecção de câncer de mama (não aprovados pela FDA na época e mantidos anônimos), tentando determinar se algum deles poderia ter desempenho igual ou superior ao dos radiologistas em diagnósticos de mamografia. Um dos sistemas mostrou-se superior aos radiologistas que realizaram a primeira leitura, modelo este que foi listado na tabela, com potencial para atuar como um leitor independente no futuro. A combinação desse sistema com os radiologistas da primeira leitura resultou nas avaliações mais precisas e alcançou as pontuações mais altas.

Holmström O, et al. (2021) conduziram um estudo em uma clínica rural no Kenya, um país onde o número de mortes entre mulheres por câncer cervical era o mais elevado, com a proposta de desenvolver um sistema barato e acessível para diagnóstico remoto. Assim como todos os algoritmos de Aprendizado Supervisionado, modelos de CNN necessitam de imagens anotadas para classificar corretamente. No caso do estudo o modelo treinado apresentou a mesma performance dos patologistas locais, os quais foram responsáveis por anotar

as imagens para o treinamento do modelo. As interpretações visuais para exames de Papanicolau são conhecidamente subjetivas, por conta disso houve diferenças com as avaliações dos patologistas remotos que analisaram as lâminas digitalizadas.

No caso só um exame de biópsia cervical poderia confirmar o diagnóstico. Já os modelos desenvolvidos por Zhang X, et al. (2023) e Matsuoka Y, et al. (2023) contam com imagens com diagnósticos confirmados por biópsia, o que torna suas classificações mais confiáveis. E pelo número de amostras ser limitado, o estudo também não é conclusivo. Alguns modelos como o modelo desenvolvido por Benyes YK, et al. (2022) podem demonstrar uma performance extremamente enviesada, produzindo bons resultados, mas sendo apenas eficiente para um único conjunto de dados específicos. Isso é comum para modelos treinados com poucas imagens e poucas variações.

DISCUSSÃO

Segundo Nasser M e Yusof UK (2023) em sua revisão sistemática para métodos de Deep Learning para diagnóstico de câncer de mama com 98 artigos selecionados de 1267 estudos, a CNN é a técnica de Aprendizado Profundo mais utilizada para diagnóstico de câncer de mama. As CNNs podem ser agrupadas em dois tipos: um modelo baseado em transferência de aprendizado e um modelo treinado do zero. Também segundo a revisão, os modelos mais eficazes são aqueles que utilizaram classificação binária ao invés de multiclasse.

Na maioria dos casos, isso se confirma. Ao comparar os modelos desenvolvidos para diagnóstico de câncer de pulmão de Masood A, et al. (2018) e Zhang X, et al. (2023), por exemplo. O primeiro conduziu uma classificação multiclasse para os 4 estágios de câncer do pulmão utilizando um conjunto de dados fornecido pelo hospital de Shanghai. O segundo tentava distinguir apenas dois problemas, um nódulo de tuberculose ou um câncer de pulmão com uma classificação binária, utilizando um conjunto de PET scans também de um hospital chinês. Para a métrica de sensibilidade menos casos verdadeiramente positivos para câncer foram perdidos no segundo modelo.

Mas, às vezes o modelo CNN também pode ser combinado com outras técnicas, não utilizando somente o reconhecimento de imagem para diagnóstico, mas também utilizando o histórico médico do paciente ou de diagnósticos. Essa combinação geralmente resulta em métricas melhores como no modelo de Lin H, et al. (2021), uma CNN treinada com 19 mil imagens com uma classificação adicional em estratificação de risco baseado em regras definidas com diagnósticos clínicos de citologistas.

Existem muitos modelos comerciais de diagnóstico de câncer já em utilização atualmente, muitos deles já atrelados às próprias máquinas de exames de imagem. Riveira-Martin M, et al. (2023) avaliaram um desses modelos comerciais, o Transpara da empresa ScreenPoint Medical para diagnóstico precoce de câncer de mama. Ao analisar determinados parâmetros chegaram à conclusão de que o sistema possui uma capacidade diagnóstica robusta e que sua precisão é independente do equipamento de mamografia utilizado, qualidade do feixe, rodada de triagem, idade da mulher, espessura da mama e compressão aplicada, densidade definida como PDV (Percent Dense tissue Volume – Porcentagem do Volume de tecido Denso) e dose glandular.

Nos exemplos selecionados há modelos clássicos como ResNet50, VGG-19 e InceptionV1 que são comumente utilizados na técnica de transferência de aprendizado, modelos que figuraram no top do ranking da ImageNet com as melhores métricas, mas também há novas arquiteturas que foram desenvolvidas posteriormente. Isso se deve ao fato do conhecimento da arquitetura de modelos de redes neurais estar disponível e pela quantidade crescente de artigos publicados, além da evolução da tecnologia e poder de processamento, que se tornam cada vez mais acessíveis.

As métricas de sensibilidade e especificidade selecionadas na **Tabela 1** para avaliar a qualidade dos modelos não possuem nenhuma diferença significativa entre elas ao analisá-las através de um teste estatístico. São métricas muito boas, que indicam um erro muito baixo de anos bem recentes. Porém, a média para a sensibilidade dos modelos selecionados foi de 0,859, já a média de especificidade foi de 0,907,

ligeiramente mais alta. Isso provavelmente ocorre porque em geral existem mais exemplos de imagens com diagnósticos saudáveis do que imagens com um diagnóstico de câncer, os modelos então tendem a classificar melhor para pessoas saudáveis. Se uma pessoa está saudável dificilmente o modelo errará ao diagnosticá-la com câncer, falsos positivos são mais raros. Ao mesmo tempo, a sensibilidade mais baixa implica que pacientes com câncer correm o risco de serem diagnosticados incorretamente.

Muitas limitações foram encontradas nos artigos pela falta de imagens de boa qualidade disponíveis para treinamento dos modelos. A técnica de Data Augmentation pode ser empregada para criar variações a partir das imagens disponíveis, mesmo assim a melhora que essa técnica pode garantir aos modelos é limitada. No entanto, com o surgimento da nova tecnologia de IA generativa combinada à arquitetura Transformer, novas soluções para esse problema estão sendo criadas.

Haveria então a possibilidade de gerar mais imagens com exemplos de câncer para treinar um modelo melhor e assim diminuir os falsos negativos. No projeto de Salvia ML, et al. (2022) por exemplo, se propuseram a desenvolver uma rede adversarial generativa convolucional profunda para gerar imagens hiperespectrais sintéticas de lesões epidérmicas, visando o diagnóstico de câncer de pele, e assim superar o desafio de conjuntos de dados insuficientes para treinar modelos. Obtiveram bons resultados iniciais como uma sensibilidade de 0,861.

É uma nova tecnologia que ainda está sendo validada, mas já há muitos estudos com resultados promissores. A partir da tecnologia da IA Generativa surge a possibilidade de se criar dados sintéticos que podem não só ajudar em modelos de reconhecimento de imagem para diagnóstico, mas também no desenvolvimento de novos medicamentos, ou até na telemedicina (Boyd BJ, 2019; Cascella M, et al., 2023).

CONSIDERAÇÕES FINAIS

A aplicação da Inteligência Artificial (IA) no diagnóstico de câncer por meio de exames de imagem tem demonstrado avanços significativos. Com a utilização de redes neurais de Aprendizado Profundo e métricas de sensibilidade e especificidade revelam-se como ferramentas valiosas para aprimorar a precisão diagnóstica. A evolução da tecnologia de IA, especialmente com o advento da arquitetura Transformer e a IA generativa, promete superar desafios relacionados à qualidade e quantidade de dados de treinamento, potencializando a criação de modelos mais eficientes e confiáveis. Estudos recentes indicam que sistemas de IA podem não apenas complementar, mas em alguns casos superar o desempenho de radiologistas e médicos, reduzindo a carga de trabalho e melhorando os resultados clínicos. Apesar das limitações atuais, como a necessidade de grandes conjuntos de dados e poder de processamento, a IA está se consolidando como uma tecnologia revolucionária na medicina diagnóstica, com potencial para transformar o diagnóstico e tratamento do câncer, bem como outras áreas da saúde.

REFERÊNCIAS

1. ARBYN M, et al. Estimates of incidence and mortality of cervical cancer in 2018: a worldwide analysis. *Lancet Glob Health*, 2020; 8(2): 191-203.
2. BENYES YK, et al. A Comparative Analysis of Deep Learning Models for Automated Cross-Preparation Diagnosis of Multi-Cell Liquid Pap Smear Images. *Diagnostics*, 2022; 12: 1838.
3. BOYD BJ, et al. Successful oral delivery of poorly water-soluble drugs both depends on the intraluminal behavior of drugs and of appropriate advanced drug delivery systems. *European Journal of Pharmaceutical Sciences*, 2019; 137: 104977.
4. CASCELLA M, et al. Utilizing an artificial intelligence framework (conditional generative adversarial network) to enhance telemedicine strategies for cancer pain management. *J Anesth Analg Crit Care*, 2023; 3(1): 19.
5. HOLMSTRÖM O, et al. Point-of-Care Digital Cytology with Artificial Intelligence for Cervical Cancer Screening in a Resource-Limited Setting. *JAMA Network Open*, 2021; 4(3): 211740.

6. HUANG HW, et al. Development of a light-weight deep learning model for cloud applications and remote diagnosis of skin cancers. *The Journal of Dermatology*, 2021; 48: 310-316.
7. HUSSAIN S, et al. Modern Diagnostic Imaging Technique Applications and Risk Factors in the Medical Field: A Review. *Biomed Res Int*, 2022; 5164970.
8. JUNIOR GBV, et al. Métricas utilizadas para avaliar a eficiência de classificadores em algoritmos inteligentes. *Centro de Pesquisas Avançadas em Qualidade de Vida*, 2022; 14: 2(1).
9. KAPLAN A e HAENLEIN M. Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 2019; 62(1): 15-25.
10. KHOSRAVI P, et al. A Deep Learning Approach to Diagnostic Classification of Prostate Cancer Using Pathology–Radiology Fusion. *Journal of Magnetic Resonance Imaging*, 2021; 4: 462–471.
11. KIM HE, et al. Changes in cancer detection and false-positive recall in mammography using artificial intelligence: a retrospective, multireader study. *Lancet Digit Heal*, 2020; 2(3): 138–48.
12. KOOI T, et al. Large scale deep learning for computer aided detection of mammographic lesions. *Med Image Anal*, 2017; 35: 303–312.
13. KUO KM, et al. The accuracy of artificial intelligence used for non-melanoma skin cancer diagnoses: a meta-analysis. *BMC Medical Informatics and Decision Making*, 2023; 23: 138.
14. LIN H, et al. Dual-path network with synergistic grouping loss and evidence driven risk stratification for whole slide cervical image analysis. *Medical Image Analysis*, 2021; 69: 101955.
15. LIU F, et al. Research progress of medical image recognition based on deep learning. *Chin J Biomed Eng*, 2018; 37(1): 86–94.
16. LIU J, et al. Application and research of deep learning in medical image recognition. *J Qingdao Univers (Nat Sci Edit)*, 2018; 31(1): 69-7480.
17. LIU R, et al. A review of medical artificial intelligence. *Global Health Journal*, 2020; 4(2).
18. LOTTER W, et al. Robust breast cancer detection in mammography and digital breast tomosynthesis using an annotation-efficient deep learning approach. *Nat Med*, 2021; 27(2): 244–9.
19. MASOOD A, et al. Computer-Assisted Decision Support System in Pulmonary Cancer detection and stage classification on CT images. *Journal of Biomedics Informatics*, 2018; 79: 117-128.
20. MATSUOKA Y, et al. Deep-learning prostate cancer detection and segmentation on biparametric versus multiparametric magnetic resonance imaging: Added value of dynamic contrast-enhanced imaging. *International Journal of Urology*, 2023; 1103-1111.
21. MCKINNEY SM, et al. International evaluation of an AI system for breast cancer screening. *Nature*, 2020; 577(7788): 89-94.
22. MISHKIN D, et al. Systematic evaluation of convolution neural networks advances on the Imagenet. *Computer Vision and Image Understanding*, 2017; 161: 11-19.
23. NASSER M e YUSOF UK. Deep Learning Based Methods for Breast Cancer Diagnosis: A Systematic Review and Future Direction. *Diagnostics*, 2023; 13: 161.
24. PUPPE F. Introduction to knowledge systems. *Artificial Intelligence in Medicine*, 1997; 9(2): 201-203.
25. RIVEIRA-MARTIN M, et al. Multi-vendor robustness analysis of a commercial artificial intelligence system for breast cancer detection. *Journal of Medical Imaging*, 2023; 10(5): 051807.
26. SALIM M, et al. External Evaluation of 3 Commercial Artificial Intelligence Algorithms for Independent Assessment of Screening Mammograms. *JAMA Oncol*, 2020; 6(10): 1581–8.
27. SANTOS MO, et al. Estimativa de Incidência de Câncer no Brasil, 2023-2025. *Revista Brasileira de Cancerologia*, 2023; 69(1): 213700.
28. SORNAPUDI S, et al. Comparing Deep Learning Models for Multi-cell Classification in Liquid- based Cervical Cytology Image. *AMIA Annu Symp Proc*, 2019; 820–827.
29. VASWANI A, et al. Attention Is All You Need. *Conference on Neural Information Processing Systems. NIPS*, 2017.
30. ZHANG X, et al. Deep learning PET/CT-based radiomics integrates clinical data: A feasibility study to distinguish between tuberculosis nodules and lung cancer. *Thoracic Cancer*, 2023; 1802-1811.