Secretaria de Estado da Saúde
Coordenadoria de Controle de Doenças
Instituto Adolfo Lutz

Curso de Especialização

Vigilância Laboratorial em Saúde Pública

Ademir Pedroso

INTELIGÊNCIA ARTIFICIAL E SAÚDE: UMA REVISÃO SOBRE O DESEMPENHO DE TÉCNICAS DE APRENDIZAGEM DE MÁQUINA NA IDENTIFICAÇÃO DE LESÕES CERVICAIS

São Paulo

2024

Curso de Especialização
Vigilância Laboratorial em Saúde Pública
Instituto Adolfo Lutz

Ademir Pedroso

INTELIGÊNCIA ARTIFICIAL E SAÚDE: UMA REVISÃO SOBRE O DESEMPENHO DE TÉCNICAS DE APRENDIZAGEM DE MÁQUINA NA IDENTIFICAÇÃO DE LESÕES CERVICAIS

Trabalho de conclusão de curso de especialização apresentado ao Instituto Adolfo Lutz - Unidade do Centro de Formação de Recursos Humanos para o SUS/SP - Doutor Antônio Guilherme de Souza como requisito parcial para obtenção do título de Especialista em Vigilância Laboratorial em Saúde Pública.

Orientadora: M.a. Daniela Etlinger-Colonelli

São Paulo

2024

Inteligência artificial e saúde: uma revisão sobre o desempenho de técnicas de aprendizagem de máquina na identificação de lesões cervicais

Ademir Pedroso[1]*; Daniela Etlinger-Colonelli[2]

RESUMO

O câncer do colo do útero é uma das principais causas de mortalidade entre mulheres no Brasil, sendo o papilomavírus humano (HPV) o principal fator associado. A detecção precoce, especialmente pela citologia cervical (exame de Papanicolaou), é essencial para prevenir a progressão de lesões malignas. No entanto, esse método tradicional enfrenta limitações como a possibilidade de exames falsos positivos e negativos, além dos desafios na interpretação manual das lâminas citológicas. Nesse contexto, a inteligência artificial (IA) surge como uma ferramenta promissora para automatizar e aumentar a precisão no diagnóstico de alterações celulares. O objetivo deste trabalho foi avaliar a eficácia dessas tecnologias na detecção de lesões cervicais em exames de citologia. O estudo revisou artigos originais publicados entre 2019 e 2024, nas bases de dados LILACS e PUBMED, foram combinados os descritores; "*machine learning*", "*artificial intelligence*" e "papanicolaou", e para maior rigor na seleção dos artigos foi utilizada a metodologia PRISMA. Os resultados indicam que os sistemas baseados em IA, alcançam alta precisão e sensibilidade, com alguns estudos chegando a 99,7% de precisão. Contudo, persistem desafios importantes, como a necessidade de conjuntos de dados maiores e mais diversificados, além de melhorias na segmentação de imagens citológicas. Conclui-se que, embora a IA mostre grande potencial para complementar os métodos tradicionais, ainda são necessários avanços significativos na qualidade dos dados e na generalização dos modelos para garantir sua ampla adoção em ambientes clínicos.

Palavras-chave: Teste de Papanicolaou; Neoplasias do Colo Uterino; Redes Neurais (Computação); Inteligência Artificial; Aprendizado de Máquina.

[1]Discente do Curso de Especialização em Vigilância Laboratorial em Saúde Pública, Centro de Patologia, Núcleo de Anatomia Patológica, Instituto Adolfo Lutz, São Paulo, biomedicoademir@gmail.com

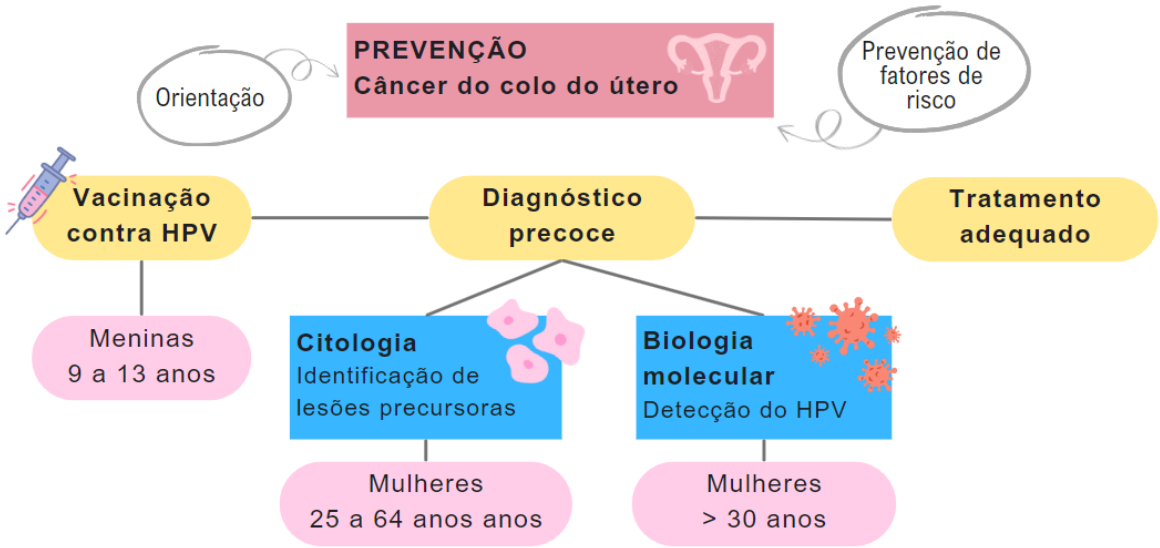
[2]Biomédica e pesquisadora científica VI, Centro de Patologia, Núcleo de Anatomia Patológica, Instituto Adolfo Lutz, São Paulo, dani_etlinger@yahoo.com.br

INTRODUÇÃO

O câncer do colo do útero é o terceiro tipo de câncer que mais afeta mulheres no Brasil, com alta taxa de mortalidade, especialmente na faixa etária entre os 30 e 50 anos. É um tipo de tumor maligno que se desenvolve a partir de células do colo uterino, apresenta desenvolvimento lento, com alterações celulares chamadas de lesões precursoras, que podem evoluir para um tumor maligno ao longo do tempo. (HOFF, 2013).

A principal causa é a infecção pelo papilomavírus humano (HPV), principalmente os subtipos 16 e 18, responsáveis por 70% dos casos de câncer cervical (LIMA et al., 2024). A prevenção primária contra o HPV envolve a vacinação de meninas entre 9 e 13 anos. Além da vacinação, o diagnóstico precoce da doença e o acesso ao tratamento adequado são fundamentais para reduzir a mortalidade. Outras medidas preventivas incluem orientação sobre sexo seguro, incentivo ao uso de preservativos e prevenção de fatores de risco (Figura 1) (OPAS; OMS, 2024).

Figura 1. Principais estratégias de prevenção, diagnóstico e tratamento de lesões precursoras do câncer de colo do útero.



Fonte: Elaborado pelo autor (2024)

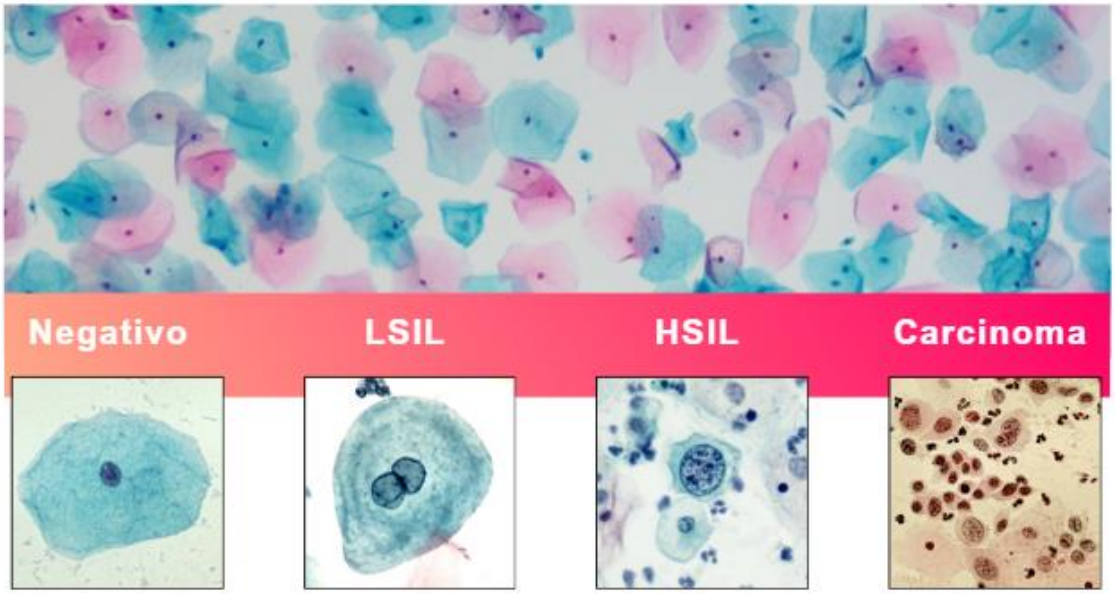
O diagnóstico precoce é fundamental para aumentar as chances de cura e reduzir a mortalidade. Atualmente, duas principais estratégias podem ser adotadas para a detecção precoce: a biologia molecular e a citologia. Os testes moleculares

para detecção do HPV são ferramentas avançadas e eficazes para o rastreamento do câncer do colo do útero. Esses testes identificam a presença do DNA viral, especialmente os tipos de alto risco oncogênico, que são os principais responsáveis pelo câncer cervical. Os testes moleculares são mais sensíveis e específicos do que os exames citopatológicos tradicionais, eles conseguem detectar o HPV mesmo em estágios iniciais, permitindo diagnóstico precoce (INCA, 2024). Os testes moleculares possuem muitas vantagens, sendo sua maior desvantagem o custo, superior à citologia (BRASIL, 2023).

A citologia cervical, popularmente conhecida como exame de Papanicolaou, possibilita a detecção precoce de alterações celulares a partir da avaliação morfológica das células esfoliadas do colo do útero. É um procedimento ginecológico essencial para a prevenção do câncer do colo uterino, que consiste na coleta de células da ectocervice e endocervice para análise microscópica, com o objetivo de detectar alterações celulares que sugerem lesões precursoras ou até mesmo câncer (AZEVEDO E SILVA et al., 2023). Na citologia convencional, após a coleta do material, as células são fixadas em uma lâmina de vidro, coradas e observadas ao microscópio, a fim de encontar possíveis alterações celulares (Figura 2). Essas alterações podem ser indicativas de infecção pelo HPV, sendo essencial identificá-las em estágio inicial para permitir o monitoramento ou tratamento adequado, antes que evoluam para o câncer (NAYAR; WILBUR, 2018).

A análise microscópica das células nos permite avaliar critérios citomorfológicos, como: hipercromasia nuclear, que indica um aumento na coloração do núcleo devido à maior concentração de cromatina; aumento nuclear, em que o núcleo da célula se apresenta maior que o normal; contorno nuclear irregular, no qual o núcleo perde sua forma regular; distribuição anormal da cromatina, em que o material genético no núcleo aparece desorganizado; aumento da relação núcleo-citoplasma, onde o núcleo ocupa uma porção maior da célula; e multinucleação, que é a presença de mais de um núcleo em uma mesma célula. Para a interpretação deste conjunto de características é utilizado o Sistema Bethesda, que propõe classificar as alterações com base no tipo de epitélio envolvido, no grau de anormalidades e no estágio de maturidade celular (NAYAR; WILBUR, 2018).

Figura 2. Representação de células escamosas esfoliadas do colo uterino e coradas com a coloração de Papanicolaou. Na imagem superior, as células observadas em menor aumento (100x) e nas imagens inferiores, células (aumento de 400x) que representam as diferentes categorias diagnósticas descritas na figura.



Fonte: Elaborado pelo autor (2024)

Embora o exame de Papanicolaou seja utilizado desde a década de 1940, ele possui limitações inerentes ao seu método. Podem ocorrer erros que resultam em diagnósticos falsos positivos ou falsos negativos, causados por problemas em diferentes etapas do processo, desde a coleta do material citológico até a interpretação da lesão. Tais equívocos podem ter um impacto significativo na conduta clínica, levando a dois cenários prejudiciais: a paciente pode não receber o acompanhamento adequado, permitindo que o câncer de colo do útero se desenvolva de forma silenciosa, ou, por outro lado, pode ser submetido a intervenções desnecessárias devido a uma classificação equivocada da lesão pelo profissional (DINIZ et al., 2021).

Examinar as lâminas ao microscópio ainda é uma tarefa desafiadora, manual e trabalhosa, mesmo para citopatologistas experientes. Isso se deve à necessidade de revisar inúmeras células em uma única lâmina. Aglomerados de células são frequentes ou alterações importantes são mascaradas por muco, por sangue, dificultando a visualização das imagens do exame e tornando o procedimento altamente suscetível a erros. Pesquisas emergentes sugerem que técnicas auxiliadas

por computador podem desempenhar um papel significativo na automatização do diagnóstico do câncer (TAN et al., 2024). Cada vez mais, procedimentos que são manuais, estão sendo substituídos por automatizados, isso é possível com uso de técnicas de *machine learning* (ML) e inteligência artificial (IA), onde um sistema computadorizado aprende a executar tais tarefas.

A IA é um campo multidisciplinar da ciência da computação que busca desenvolver sistemas capazes de realizar tarefas que, até então, eram exclusivas da inteligência humana, como aprendizado, raciocínio, resolução de problemas, reconhecimento de padrões, e tomada de decisão. IA refere-se a procedimentos computacionais que realizam funções que, se executadas por humanos, seriam consideradas inteligentes. Busca desenvolver sistemas com capacidades como aquisição de conhecimentos, permitindo aprender e armazenar informações; resolução de problemas complexos; reconhecimento de padrões em dados; e comunicação através de linguagens naturais. Essas habilidades são aplicadas em diversas áreas, como ML, automação, robótica e processamento de linguagem natural (LIMA; PINHEIRO; OLIVEIRA SANTOS, 2014).

O objetivo central de sistemas inteligentes é simular processos cognitivos humanos para criar ferramentas que sejam capazes de "aprender" e melhorar seu desempenho com base em experiências, sem a necessidade de serem programados para cada tarefa. Isso é alcançado principalmente por meio de abordagens como redes neurais artificiais, aprendizado supervisionado e não supervisionado, e processamento de linguagem natural, entre outras técnicas. A aprendizagem por transferência é referida como a capacidade de aplicar o conhecimento e a habilidade adquiridos em trabalhos anteriores a novas tarefas (TAN et al., 2024). A abordagem mais utilizada atualmente é a rede neural convolucional profunda (CNN). Ela é chamada "rede" por ser uma estrutura matemática complexa formada por subexpressões interligadas. O termo "convolucional" refere-se à repetição de um padrão fixo em toda a imagem de entrada. Já "profunda" indica que a rede possui várias camadas, o que a torna mais poderosa e eficaz (RUSSEL, 2021).

Nos últimos anos, a combinação de IA e ML tem se mostrado promissora no campo da saúde, especialmente em exames diagnósticos. Uma das áreas com

crescente interesse é a aplicação dessas tecnologias na citologia cervical. Com o intuito de aumentar a precisão, reduzir erros humanos e otimizar a eficiência nos processos de análise, pesquisadores têm explorado como tecnologias baseadas em IA podem complementar ou até substituir a interpretação manual de resultados citológicos. Com a evolução tecnológica, diversos sistemas baseados em algoritmos computacionais foram desenvolvidos para a análise automática de imagens celulares, visando aprimorar tanto a eficiência quanto a precisão na triagem (DINIZ et al., 2021).

O objetivo deste trabalho foi revisar estudos que exploraram o uso de IA na detecção automática de alterações celulares, especificamente em exames de citologia cervical. Além disso, foi realizada a análise das diversas abordagens, como CNNs e sistemas de triagem automatizados, na busca de avaliar o desempenho dessas tecnologias em comparação aos métodos manuais, considerando a precisão diagnóstica, sensibilidade, especificidade e a viabilidade clínica para auxiliar na detecção precoce de lesões pré-malignas e malignas.

METODOLOGIA

Trata-se de uma revisão narrativa da literatura, de natureza retrospectiva, descritiva e exploratória, com enfoque qualitativo. O objetivo desta abordagem é proporcionar uma compreensão sobre o uso de IA, no contexto da detecção de anomalias celulares, no exame de citologia cervical. A escolha da revisão narrativa permite maior flexibilidade na exploração do tema, facilitando a identificação de padrões e lacunas no conhecimento atual, assim como a síntese crítica de resultados obtidos em diferentes estudos.

A pesquisa foi realizada utilizando as bases de dados em saúde: PubMed (United States National Library of Medicine) e LILACS (Literatura Latino-Americana e do Caribe em Ciências da Saúde). Para a busca nas bases de dados, foram selecionados os seguintes termos relacionados ao foco do estudo: "*machine learning*", "artificial intelligence" e "papanicolaou". Estes termos foram combinados utilizando operadores booleanos (AND e OR) de forma a ampliar a abrangência dos resultados, sem perder o foco no escopo do trabalho. A estratégia de busca foi estruturada da seguinte maneira: “*machine learning*” OR “artificial intelligence” AND “papanicolaou”. Essa combinação permitiu captar artigos que discutem tanto o desenvolvimento e

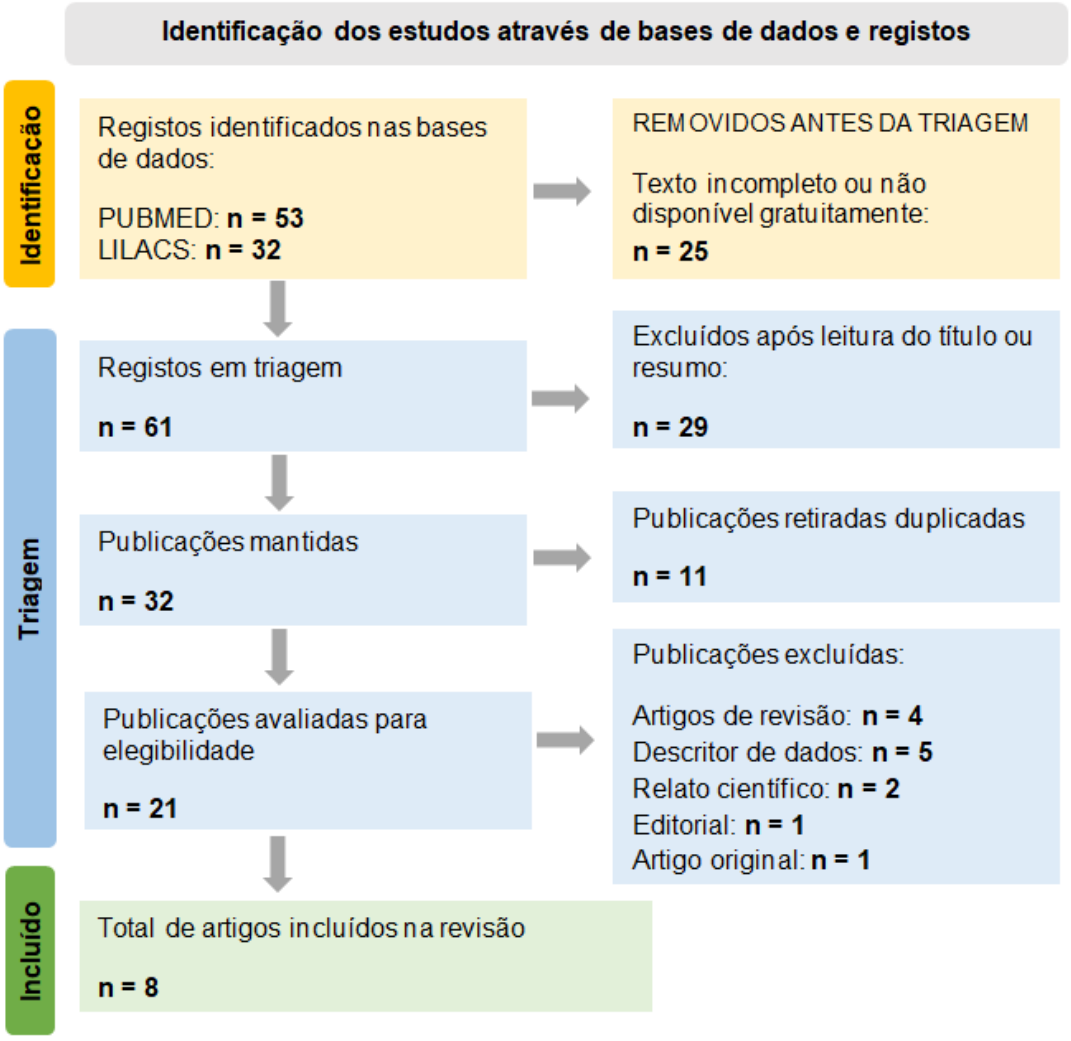
aplicação de sistemas baseados em IA quanto sua relevância no aprimoramento dos métodos de diagnóstico citológico. Os critérios de inclusão para a seleção dos artigos foram definidos com base em sua relevância temática e metodológica.

Para garantir transparência e rigor no processo de seleção dos artigos, foi elaborado um fluxograma utilizando a metodologia PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses). O PRISMA é amplamente reconhecido por fornecer diretrizes claras para condução de revisões sistemáticas, facilitando a replicabilidade e a verificação dos passos metodológicos adotados, o objetivo do PRISMA é ajudar os autores a melhorarem o relato de revisões sistemáticas e meta-análises (GALVÃO; PANSANI; HARRAD, 2015). O fluxograma delineia as etapas da busca, desde a identificação inicial dos artigos até a seleção final, oferecendo uma visão detalhada das exclusões em cada fase. Dessa forma, a metodologia deste estudo assegura uma abordagem criteriosa para a revisão da literatura, contribuindo para uma análise crítica e detalhada do uso de ML e IA no aprimoramento do diagnóstico citológico.

Foram considerados elegíveis os estudos originais, publicados entre 2019 e 2024, redigidos em inglês, disponíveis integralmente e de forma gratuita. Além disso, foram priorizados os artigos que apresentavam resultados robustos e discutiam em profundidade a aplicação de sistemas inteligentes na citologia cervical, de modo a garantir que a revisão abrangesse tanto a teoria quanto os avanços práticos. Estudos que se distanciavam do foco principal, como aqueles que abordavam outras áreas da medicina, foram excluídos.

A triagem inicial dos artigos foi realizada por meio da leitura dos títulos e resumos, visando identificar rapidamente a pertinência de cada estudo em relação ao tema investigado. Os artigos que não estavam diretamente relacionados ao uso de soluções computacionais inteligentes na citologia cervical foram eliminados nesta fase. Em seguida, os artigos pré-selecionados passaram por uma análise mais aprofundada de seu conteúdo integral, avaliando-se aspectos como metodologia, resultados e discussões. Estudos que se destacaram por sua relevância metodológica ou inovadora foram incluídos na análise final (Figura 3).

Figura 3. Fluxograma de seleção dos artigos de acordo com as recomendações da metodologia PRISMA



O quadro 1 apresentada a seguir, resume os tópicos analisados nos artigos selecionados para compor a revisão da literatura. O quadro 2 mostra as principais limitações encontradas pelos autores.

Quadro 1. Apresentação dos autores, ano, objetivo, metodologia e principais resultados dos artigos selecionados para a revisão.

Autores / Ano	Objetivo	Metodologia	Resultados
SHANTHI et al., 2019	Desenvolver um modelo de rede neural de convolução (CNN) profunda para detecção e classificação de malignidade em imagens microscópicas de células cervicais.	Utilizaram uma CNN para classificar imagens de células cervicais em diferentes classes. Avaliaram a precisão para 5 classes, de normal até carcinoma. Foram utilizadas 749 imagens do banco de dados Herlev, utilizou-se cinco algoritmos de pré-processamento para melhorar as imagens, foram segmentados apenas os núcleos das células e comparados com normais de referência. Finalmente o modelo de CNN foi usado para extrair de forma automática as características das células e classificá-las.	O modelo de CNN alcançou precisão de: 94,1% para problemas de 5 classes, 96,2% para 4 classes, 94,8% para 3 classes e 95,7% para problemas de 2 classes.
SANYAL et al., 2020	Desenvolver uma CNN capaz de identificar focos "anormais" em esfregaços cervicais convencionais para posterior avaliação adicional por um patologista.	Foram coletadas 1.838 microfotografias de esfregaços cervicais contendo focos "normais" e "anormais", esses dados foram divididos em conjuntos de treinamento, teste e validação. As microfotografias foram selecionadas e classificadas previamente por um grupo de patologistas. A CNN foi desenvolvida em Python, utilizando bibliotecas como OpenCv, TensorFlow e Keras.	A CNN alcançou uma precisão diagnóstica de 95,46%, sensibilidade de 94,28% e especificidade de 96,01%, valor preditivo positivo 91,66% e valor preditivo negativo 97,30%. A CNN reconheceu células normais sobrepostas, neutrófilos e detritos, classificando corretamente focos anormais.

Continuação quadro 1

Autores / Ano	Objetivo	Metodologia	Resultados
HOLMSTRÖM et al., 2021	Determinar se o diagnóstico de microscopia digital apoiado por inteligência artificial pode ser implementado em um ambiente com recursos limitados e usado para análise de testes de Papanicolaou.	Foram coletados esfregaços cervicais de 740 mulheres HIV-positivas em uma clínica rural no Quênia. As amostras foram digitalizadas com um scanner portátil de lâmina e carregadas para um servidor na nuvem. Foi treinado um sistema de aprendizagem profunda (DLS) e validado para detectar células cervicais atípicas, utilizando 350 imagens para treinamento e 361 para validação.	O DLS alcançou sensibilidade de 95,7% a 100% na detecção de atipias celulares cervicais. A especificidade variou entre 78,4% a 93,3%, sendo maior para lesões de alto grau. A área sob a curva de característica de operação do receptor (AUC) foi de 0,94 a 0,96, indicando alta precisão na classificação de atipias. A concordância entre avaliadores foi substancial (Kappa = 0,72) para lâminas digitais e razoável (Kappa 0,36) para lâminas de vidro.
BHATT et al., 2021	Propor uma metodologia inovadora para a classificação multiclasse de células cervicais a partir de imagens de lâminas inteiras (WSI) usando CNN com aprendizado por transferência e redimensionamento progressivo.	Foram utilizadas CNN (ConvNet) pré-treinadas em grandes conjuntos de dados e re-treinamento nos conjuntos de dados específicos de câncer cervical (Herlev e SIPaKMed). Treinamento interativo das ConvNets com imagens de tamanhos progressivamente maiores, começando com 224x224 pixels até 2048x2048 pixels. Uso de modelos EfficientNet-B3, B4 e B5 que equilibram profundidade, largura e resolução da rede. Geração de mapas de calor GradCam para visualizar e interpretar as características aprendidas pelo modelo, destacando lesões pré-malignas e malignas.	A metodologia proposta alcançou precisão de 99,7% na classificação multiclasse de WSI do conjunto de dados SIPaKMed. As técnicas de aprendizado por transferência e redimensionamento progressivo resultaram em melhorias significativas no desempenho dos modelos. A utilização do GradCam melhorou a interpretação dos resultados, destacando visualmente as lesões pré-malignas e malignas nas imagens. Os resultados foram validados com sucesso por patologistas, confirmando a eficácia do sistema proposto.

Continuação quadro 1

Autores / Ano	Objetivo	Metodologia	Resultados
TAO et al., 2022	Avaliar o desempenho de um sistema de triagem baseado na inteligência artificial (IA) para prever a citologia de células escamosas atípicas de significado indeterminado (ASC-US) para lesões de NIC2.	Foram utilizadas mais de 60.000 imagens para treinar o sistema de triagem ASC-US baseado em aprendizagem profunda. O estudo incluiu 1.967 exames de Papanicolaou com interpretação de ASC-US entre 2017 e 2019. O desempenho da triagem entre o sistema de IA e o teste de HPV de alto risco (hrHPV) foi comparado em 622 pacientes com teste simultâneos.	O sistema de triagem baseado em IA apresentou sensibilidade de 92,9% e especificidade de 49,7% comparável ao teste hrHPV. Mostrou desempenho superior ao teste hrHPV sem a necessidade de exames adicionais. A maior especificidade do sistema reduziu a ansiedade dos pacientes e evitou biópsias e tratamentos excessivos. O sistema pode ser usado de forma independente ou em conjunto com o teste hrHPV para triagem de ASC-US.
AKCA; ERKILIC, 2023	Investigar a utilidade diagnóstica do ThinPrep Imaging System (TIS) na detecção de lesões glandulares pré-malignas e malignas em esfregaços cervicais. Comparar a sensibilidade dos resultados do TIS com a triagem manual para determinar a eficácia do sistema em identificar células glandulares atípicas (AGC).	Foram incluídas 126 amostras de esfregaços cervicais entre 2010 e 2022, com confirmação histológica de adenocarcinoma endometrial ou adenocarcinoma *in situ*. As amostras foram reavaliadas manualmente e pelo TIS por dois patologistas experientes, comparando os resultados em termos de sensibilidade. Foram excluídas do estudo amostras que apresentaram problemas técnicos ou insatisfatórias para avaliação.	O TIS detectou AGC em 66 dos 70 casos (94,3%), enquanto a triagem manual detectou apenas 48 casos (68,5%). O TIS encontrou 30 casos positivos a mais em comparação com os diagnósticos iniciais feitos por diferentes patologistas. O uso do TIS economiza tempo e ajuda a garantir a padronização, sendo especialmente útil em laboratórios com alta carga de trabalho, é um método eficaz e com alta sensibilidade para detectar AGC, comparável à detecção de anomalias de células escamosas.

Continuação quadro 1

Autores / Ano	Objetivo	Metodologia	Resultados
TAN et al., 2024	Desenvolver modelos de aprendizagem profunda para detecção automatizada de câncer cervical a partir de imagens de exames de Papanicolaou. Criar modelos que não dependam de métodos de segmentação ou recursos customizados, utilizando aprendizagem por transferência com modelos de CNN pré-treinados. Melhorar a precisão e o desempenho na classificação de células cervicais, mesmo com a disponibilidade limitada de dados.	Utilização de modelos de CNN pré-treinadas para operar diretamente em imagens de esfregaço de Papanicolaou. Uso do conjunto de dados Herlev, para a tarefa de classificação de sete classes. Avaliação e comparação de 13 modelos de CNN profundos pré-treinados utilizando o pacote Keras no Google Collaboratory. Identificação do DenseNet-201 como o modelo com melhor desempenho em termos de precisão e eficiência computacional.	O modelo DenseNet-201 apresentou maior sensibilidade e especificidade, tendo maior desempenho, com uma precisão de 87,02%. Todos os 13 modelos de CNN pré-treinados foram treinados em menos de 6 minutos. DenseNet-201, DenseNet-169 e Inception tiveram precisão significativa para as classes 2 e 5.
BAI et al., 2024	Avaliar a eficácia e precisão do sistema auxiliar de IA AICyte na triagem citológica cervical. Investigar a utilidade do sistema para auxiliar no fluxo de trabalho e na capacidade de diagnóstico, comparando-o com métodos tradicionais de triagem, Determinar se o AICyte pode reduzir a carga de trabalho, mantendo a precisão diagnóstica.	O estudo utilizou o scanner Ruiqian WSI-2400, que digitaliza lâminas de citologia cervical e fornece imagens de dados diagnósticos em 2 minutos. O sistema AICyte usa modelos de aprendizado profundo (CNN Yolo5 e ResNet50) para segmentar e classificar células. Dois patologistas revisaram 32.451 casos de forma independente, utilizando o sistema AICyte para identificar objetos de interesse e fazer o diagnóstico. A análise estatística Kappa foi utilizada para medir concordância.	O tempo médio de leitura dos patologistas usando o AICyte foi 22,23 segundos /caso, comparado a cerca de 180 segundos para leitura manual. O AICyte apresentou sensibilidade: 97,89% e aumento significativo na especificidade (16,19% vs 6,77%) para a detecção NIC2 e leões acima. A concordância Kappa entre os patologistas usando AICyte e a interpretação manual foi aceitável (0,65 e 0,61 respectivamente). O AICyte tem potencial para reduzir quase metade da carga de trabalho na triagem primária da citologia cervical, mantendo alta precisão e eficiência.

Quadro 2. Principais limitações encontradas de acordo com cada autor.

Autores	Limitações
SHANTHI et al., 2019	Necessidade de grandes conjuntos de dados para garantir alta precisão; dificuldades com segmentação de núcleos celulares complexos.
SANYAL et al., 2020	Problemas com sobreposição de células e detritos nas imagens, afetando a precisão; aplicabilidade limitada a esfregaços convencionais.
HOLMSTRÖM et al., 2021	Comparação limitada com poucos especialistas; natureza subjetiva dos diagnósticos citológicos.
BHATT et al., 2021	Dependência da localização precisa dos núcleos celulares; dificuldades com a extração de características em células sobrepostas.
TAO et al., 2022	Generalização limitada devido ao uso de uma plataforma de preparação específica; insuficiência de poder de cálculo para digitalização.
AKCA et al., 2023	Problemas com amostras mascaradas por infecções ou sangue; necessidade de especialistas para interpretação precisa.
TAN et al., 2024	Pequeno tamanho do conjunto de dados e desequilíbrio de classes; necessidade de ajustes e testes adicionais para generalização.
BAI et al., 2024	Limitações relacionadas à natureza retrospectiva do estudo e ao número limitado de objetos de interesse analisados.

DISCUSSÃO

A aplicação da IA é vasta e impacta inúmeros setores, desde a saúde, com sistemas de diagnóstico, até a indústria, onde automação inteligente vem sendo utilizada para otimizar a produção. Sistemas inteligentes estão revolucionando a medicina com diagnósticos mais precisos, tratamentos personalizados e análise de grandes volumes de dados médicos. Ferramentas de IA podem ajudar a detectar doenças em estágios iniciais e prever surtos e epidemias (SICHMAN, 2021).

Atualmente existem alguns conjuntos de dados gratuitos com imagens de citologia cervical, que podem ser usados em pesquisas e desenvolvimento de novas ferramentas, sendo os mais populares SIPakMed e Herlev, onde são incorporadas imagens de células individuais (BHATT et al., 2021).

Neste estudo, ao comparar os resultados sobre o uso de sistemas de raciocínio automático no diagnóstico citológico e na detecção de alterações cervicais, foram identificadas variações significativas tanto nos resultados quanto nas limitações enfrentadas. O desempenho dos modelos depende das técnicas utilizadas, dos conjuntos de dados empregados e dos desafios inerentes à classificação de imagens

citológicas complexas. Os estudos de BHATT et al. (2021), BAI et al. (2024), e AKCA et al. (2023) destacam-se pelo desempenho robusto.

BHATT et al. (2021) alcançaram precisão de 99,7% na classificação multiclasse de imagens de lâminas inteiras, utilizando o conjunto de dados SIPaKMed. O uso de aprendizado por transferência e técnicas como o GradCam permitiu uma visualização mais clara das lesões pré-malignas e malignas, tornando o sistema altamente eficaz, apesar da dependência da localização dos núcleos celulares e da dificuldade de extrair características em imagens com sobreposição de células. BAI et al. (2024) obtiveram sensibilidade de 97,89% e destacou-se pela eficiência na redução do tempo de leitura dos patologistas, embora o estudo tenha enfrentado limitações como a falta de exames histológicos para correlação dos casos e ao número limitado de objetos de interesse por lâmina. Já AKCA et al. (2023) focaram na detecção de células glandulares atípicas, alcançando 94,3% de sensibilidade com o *ThinPrep Imaging System*, superando a triagem manual, mas enfrentando desafios na detecção manual inicial e na interpretação de amostras com sobreposição celular.

Outros estudos apresentaram bons resultados, embora com limitações que restringem seu uso em cenários clínicos. SANYAL et al. (2020) observaram precisão diagnóstica de 95,46%, com boa sensibilidade e especificidade, mas tiveram problemas de sobreposição de células e detritos nas imagens, além de limitações na aplicabilidade a outros tipos de esfregaços. SHANTHI et al. (2019) demostraram precisão entre 94,1% e 96,2%, destacando-se no uso de CNNs, mas a necessidade de grandes conjuntos de dados e a queda de desempenho em cenários complexos limitaram a generalização do modelo.

Por outro lado, alguns estudos apresentaram resultados mais modestos. TAO et al. (2022) relataram sensibilidade de 92,9%, mas com especificidade de apenas 49,7%, indicando a necessidade de melhorias na diversificação do banco de dados. TAN et al. (2024) utilizaram o modelo de rede neural DenseNet-201, alcançando 87,02% de precisão, mas o pequeno tamanho do conjunto de dados Herlev e o desequilíbrio de classes prejudicaram o desempenho.

Entre as limitações comuns aos estudos, destacam-se o tamanho e a qualidade dos conjuntos de dados, como fatores que precisam ser melhorados para aumentar a generalização dos modelos (SHANTI et al., 2019; TAN et al., 2024). A sobreposição de células e artefatos também foi uma dificuldade mencionada, resultando em falsos positivos e negativos (SANYAL et al., 2020; AKCA et al., 2023). A dependência de uma segmentação eficaz, essencial para a alta acurácia, é outro desafio enfrentado por muitos modelos, como os de SHANTI et al. (2019) e BHATT et al. (2021), especialmente em imagens complexas. Por fim, a generalização dos resultados foi apontada como um problema devido ao uso de plataformas específicas ou ao foco em subconjuntos de células, o que pode limitar a aplicação dos modelos a diferentes contextos clínicos (TAO et al, 2022; BAI et al 2024).

O avanço de mecanismos de simulação cognitiva também levanta questões éticas e sociais profundas. A automação do trabalho, o viés nos algoritmos, a privacidade dos dados e os riscos de uma IA fora de controle são temas amplamente discutidos pela academia, indústria e sociedade civil. Algoritmos podem apresentar falibilidade, gerando conclusões incorretas ao se basearem em correlações inválidas entre dados, o que pode resultar em decisões injustas. Outro desafio importante é o viés associado à discriminação, uma vez que os algoritmos podem refletir preconceitos humanos, levando a decisões tendenciosas e discriminatórias. A autonomia dessas tecnologias também levanta questões éticas sobre responsabilidade, especialmente em casos de danos causados por suas ações (ROSSETTI; ANGELUCI, 2021). Assim, o desenvolvimento de IA não apenas requer avanços tecnológicos, mas também uma consideração cuidadosa dos impactos sobre a sociedade.

CONCLUSÃO

A aplicação de IA na detecção de anomalias cervicais, apresenta resultados promissores, com destaque para o uso de redes neurais CNNs e ML. Estudos recentes demonstraram a capacidade dessas técnicas de aumentar a precisão e reduzir o tempo de diagnóstico, contribuindo significativamente para a detecção precoce de lesões pré-malignas e malignas. No entanto, desafios importantes ainda

precisam ser superados, como a necessidade de conjuntos de dados maiores e mais diversos, melhorias nos algoritmos de segmentação de imagens e maior generalização dos modelos para diferentes contextos clínicos.

Apesar das limitações, o potencial da IA em complementar e aprimorar o diagnóstico citológico é inegável, apontando para um futuro em que esses sistemas automatizados desempenharão um papel importante na triagem e prevenção do câncer cervical. Assim, esforços contínuos de pesquisa e desenvolvimento serão primordiais para tornar essas tecnologias mais acessíveis e eficazes em larga escala.

REFERÊNCIAS

AKCA Y, ERKILIC S. Diagnostic utility of ThinPrep imaging System® for detecting atypical glandular cells in cervical smear samples. **Diagnostic Cytopathology**. 2023;51(2):135-139. doi:10.1002/dc.25066.

AZEVEDO E SILVA, G. et al. Exame de Papanicolaou no Brasil: análise da Pesquisa Nacional de Saúde em 2013 e 2019. **Revista de Saúde Pública**, v. 57, n. 1, 2023.

BAI, X. et al. Assessment of Efficacy and Accuracy of Cervical Cytology Screening With Artificial Intelligence Assistive System. **Modern Pathology***, v. 37*, n. 6, 1 jun. 2024.

BHATT, A. R. et al. Cervical cancer detection in pap smear whole slide images using convNet with transfer learning and progressive resizing. **PeerJ Computer Science**, v. 7, p. 1–18, 2021.

BRASIL**.** Ministério da Saúde. Comissão Nacional de Incorporação de Tecnologias no Sistema Único de Saúde (Conitec). **Relatório de Recomendação: Testagem molecular para detecção de HPV**, nº 65/2023. 2023. Disponível em: https://www.gov.br/conitec/ptbr/midias/consultas/relatorios/2023/relatorio_cp_testagem-molecular-deteccao-hpv_65_2023.pdf. Acesso em: 07 out. 2024.

DINIZ, D. N. et al. A deep learning ensemble method to assist cytopathologists in pap test image classification. **Journal of Imaging**, v. 7, n. 7, 1 jul. 2021.

GALVÃO, T. F.; PANSANI, T. DE S. A.; HARRAD, D. Principais itens para relatar Revisões sistemáticas e Meta-análises: A recomendação PRISMA. **Epidemiologia e Serviços de Saúde**, v. 24, n. 2, p. 335–342, abr. 2015.

HOFF, P. M. G. **Tratado de oncologia**. 4 ed. São Paulo: Atheneu, 2013. 2829p.

HOLMSTRÖM, O. et al. Point-of-Care Digital Cytology with Artificial Intelligence for Cervical Cancer Screening in a Resource-Limited Setting. **JAMA Network Open**, v. 4, n. 3, 17 mar. 2021.

INSTITUTO NACIONAL DO CÂNCER. **SUS passa a ter teste inovador para detecção do HPV em mulheres**. Disponível em: www.gov.br/inca. Acesso em: 20 set. 2024.

LIMA, D. E. de O. B. et all. Conhecimento de Mulheres acerca do Exame Papanicolaou. **Revista Brasileira de Cancerologia**, *[S. l.]*, v. 70, n. 1, p. e–054393, 2024. DOI: 10.32635/2176-9745.RBC.2024v70n1.4393.

LIMA, I.; PINHEIRO, C. A. M.; OLIVEIRA SANTOS, F. A. **Inteligência Artificial**. 1. ed. Rio de Janeiro: Elsevier, 2014.

NAYAR, R.; WILBUR, D. C. **O Sistema Bethesda para Relatórios Citologia Cervical**. 3. ed. São Paulo: Livraria Livromed, 2018.

OPAS, OMS: **HPV e Câncer do colo do útero**. Disponível em: httwww.paho.org/pt/topicos/hpv-e-cancer-do-colo-do-utero. Acesso em: 18 set. 2024.

ROSSETTI, R.; ANGELUCI, A. Ética algorítmica: questões e desafios éticos do avanço tecnológico da sociedade da informação. **Galáxia**, n. 46, 2021.

RUSSELL, S. **Inteligencia artificial a nosso favor**. como manter o controle sobre a tecnologia. São Paulo: Editora Schwarcz, 2021.

SANYAL, P.; GANGULI, P.; BARUI, S. Performance characteristics of an artificial intelligence based on convolutional neural network for screening conventional Papanicolaou-stained cervical smears. **Medical Journal Armed Forces India**, v. 76, n. 4, p. 418–424, 1 out. 2020.

SHANTHI, P. B. et al. Deep Convolution Neural Network for malignancy detection and classification in microscopic uterine cervix cell images. **Asian Pacific Journal of Cancer Prevention**, v. 20, n. 11, p. 3447–3456, 2019.

SICHMAN, J. S. Inteligência Artificial e sociedade: avanços e riscos. **Estudos Avançados**, v. 35, n. 101, p. 37–49, 1 jan. 2021.

TAN, S. L. et al. Cervical Cancer Classification From Pap Smear Images Using Deep Convolutional Neural Network Models. **Interdisciplinary Sciences – Computational Life Sciences**, v. 16, n. 1, p. 16–38, 1 mar. 2024.

TAO, X. et al. Scrutinizing high-risk patients from ASC-US cytology via a deep learning model. **Cancer Cytopathology**, v. 130, n. 6, p. 407–414, 1 jun. 2022.