Journal of Health Informatics

J Health Inform. 2026;18:e1519
DOI: https://doi.org/10.59681/2175-4411.v18.2026.1519

ARTIGO DE REVISÃO

Revisão sistemática sobre técnicas de visão Computacional aplicadas à classificação de leucócitos

Systematic review on computer vision techniques applied to leukocyte classification

Revisión sistemática sobre técnicas de visión por computadora aplicadas a la clasificación de leucocitos

João Kasprowicz[1], Halan Germano Bacca[1], Alexandre Gonçalves Silva[1]

ORCID: 0009-0002-1703-8630 · 0000-0001-7694-1390 · 0000-0002-0189-8540

[1] Universidade Federal de Santa Catarina.

FICHA DO ARTIGO

TIPO
Artigo de Revisão

IDENTIFICADOR
e1519

DOI
10.59681/2175-4411.v18.2026.1519

RECEBIDO
15/07/2025

APROVADO
06/01/2026

PUBLICADO
21/01/2026

PROCESSO DE AVALIAÇÃO
Double Blind Review

EDITOR RESPONSÁVEL
Juliano de Souza Gaspar

ACESSO
Aberto, CC BY 4.0

AUTOR CORRESPONDENTE
João Kasprowicz
joao.kasprowicz@univali.br

LICENÇA

Este é um artigo publicado em acesso aberto sob a licença Creative Commons CC BY 4.0, que permite uso, distribuição e reprodução em qualquer meio, desde que o trabalho original seja corretamente citado.

CC BY 4.0

RESUMO

Introdução: A classificação automatizada de leucócitos por visão computacional é uma alternativa promissora à análise manual, que é demorada e subjetiva. **Objetivo:** Esta revisão sistemática visa analisar e sintetizar as evidências sobre os algoritmos de aprendizado de máquina aplicados à classificação de leucócitos em imagens de esfregaço sanguíneo, avaliando seu desempenho e desafios. **Métodos:** Foi realizada uma revisão sistemática (PRISMA) nas bases MEDLINE/PubMed, Embase e Scopus, para estudos publicados entre 2020 e 2025. **Resultados:** A partir de 300 registros iniciais, 28 estudos foram incluídos. Os resultados indicam um predomínio de modelos de aprendizado profundo (CNNs, YOLO) com acurácia frequentemente superior a 95% na classificação de leucócitos maduros. **Conclusão:** Embora tecnicamente madura para células normais, a área enfrenta desafios como a heterogeneidade metodológica entre os estudos e uma lacuna na classificação de células imaturas e atípicas, que possui maior relevância clínica.

Descritores: Inteligência Artificial; Contagem de Leucócitos; Processamento de Imagem Assistida por Computador.

ABSTRACT

Introduction: Automated leukocyte classification using computer vision is a promising alternative to manual analysis, which is time-consuming and subjective. **Objective:** This systematic review aims to analyze and synthesize evidence on machine learning algorithms applied to leukocyte classification in peripheral blood smear images, evaluating their performance and challenges. **Methods:** A systematic review (PRISMA) was conducted in MEDLINE/PubMed, Embase, and Scopus for studies published between 2020 and 2025. **Results:** Out of 300 initial records, 28 studies were included. The findings indicate a predominance of deep learning models (CNNs, YOLO) with accuracies often exceeding 95% in the classification of mature leukocytes. **Conclusion:** Although technically mature for normal cell classification, the field faces challenges such as methodological heterogeneity across studies and a gap in the classification of immature and atypical cells, which have greater clinical relevance.

Keywords: Artificial Intelligence; Leukocytes; Image Interpretation, Computer-Assisted.

RESUMEN

Introducción: La clasificación automatizada de leucocitos mediante visión por computadora es una alternativa prometedora al análisis manual, que resulta laborioso y subjetivo. **Objetivo:** Esta revisión sistemática tiene como objetivo analizar y sintetizar la evidencia sobre los algoritmos de aprendizaje automático aplicados a la clasificación de leucocitos en imágenes de frotis de sangre periférica, evaluando su desempeño y los desafíos asociados. **Métodos:** Se realizó una revisión sistemática (PRISMA) en las bases de datos MEDLINE/PubMed, Embase y Scopus para estudios publicados entre 2020 y 2025. **Resultados:** De 300 registros iniciales, se incluyeron 28 estudios. Los resultados indican un predominio de modelos de aprendizaje profundo (CNN, YOLO), con precisiones que superan frecuentemente el 95% en la clasificación de leucocitos maduros. **Conclusión:** Aunque técnicamente madura para la clasificación de células normales, el área enfrenta desafíos como la heterogeneidad metodológica entre los estudios y una brecha en la clasificación de células inmaduras y atípicas, que poseen mayor relevancia clínica.

Descriptores: Inteligencia Artificial; Leucocitos; Interpretación de Imágenes Asistida por Computadora.

COMO CITAR ESTE ARTIGO: Kasprowicz J, Bacca HG, Silva AG. Revisão sistemática sobre técnicas de visão Computacional aplicadas à classificação de leucócitos. J Health Inform 2026;18:e1519. doi: https://doi.org/10.59681/2175-4411.v18.2026.1519

1 INTRODUÇÃO

As células brancas do sangue (leucócitos) são componentes vitais do sistema imune, e sua classificação em subtipos é fundamental para o diagnóstico de diversas condições, como infecções, inflamações e neoplasias hematológicas. A contagem diferencial de leucócitos é, portanto, um dos pilares da análise hematológica clínica [1].

Tradicionalmente, a análise de esfregaços de sangue periférico é realizada manualmente ao microscópio. No entanto, este método é laborioso, demorado e suscetível à variabilidade do observador, o que representa um desafio significativo para a eficiência e padronização do diagnóstico em laboratórios com alta demanda [2].

Para superar essas limitações, técnicas de aprendizado profundo (*deep learning*) têm sido aplicadas com resultados promissores na classificação automatizada de leucócitos. Apesar do sucesso, a área ainda enfrenta desafios importantes, incluindo a variabilidade morfológica das células, a falta de padronização nos métodos de aquisição de imagem e a escassez de grandes conjuntos de dados rotulados [1].

Diante da crescente produção científica, uma análise sistemática da literatura torna-se essencial para consolidar o conhecimento atual. Portanto, o objetivo desta revisão é investigar e analisar criticamente os algoritmos de visão computacional aplicados à classificação de leucócitos, identificando as principais arquiteturas, avaliando seu desempenho diagnóstico e discutindo os desafios e oportunidades para futuras pesquisas.

2 MÉTODOS

Esta revisão sistemática foi conduzida de acordo com as diretrizes PRISMA 2020 [31] e seu protocolo foi previamente registrado nas plataformas PROSPERO (CRD420251060493) e Open Science Framework (DOI: 10.17605/OSF.IO/ND8W5). A pesquisa buscou responder à seguinte pergunta, estruturada pela metodologia PCC: "Como os modelos de aprendizado de máquina e visão computacional têm sido aplicados na classificação automática de leucócitos em amostras de sangue periférico?".

2.1 Estratégia de busca

A busca foi realizada em junho de 2025 nas seguintes bases de dados eletrônicas: MEDLINE (via PubMed), Embase e Scopus. A estratégia de busca foi desenvolvida utilizando descritores controlados (MeSH e DeCS) e termos livres, combinados com operadores booleanos "AND" e "OR". Embora os artigos pudessem ser em inglês, português ou espanhol, os termos de busca foram mantidos em inglês para maximizar a sensibilidade da busca nas bases de dados internacionais. Foram excluídos artigos de revisão, estudos não humanos ou de medula óssea, e publicações sem dados originais. O Quadro 1 detalha as estratégias utilizadas.

Quadro 1. Estratégia de busca.

Base de dados	Estratégia de busca
PubMed/ MEDLINE	(("Artificial Intelligence"[MeSH Terms] OR "Machine Learning"[MeSH Terms] OR "Deep Learning"[MeSH Terms] OR "neural networks, computer"[MeSH Terms] OR "Artificial Intelligence"[All Fields] OR "Machine Learning"[All Fields] OR "Deep Learning"[All Fields] OR "neural networks"[All Fields]) AND ("image interpretation, computer assisted"[MeSH Terms] OR "computer vision"[All Fields] OR "image analysis"[All Fields] OR "image processing"[All Fields]) AND ("leukocytes"[MeSH Terms] OR "white blood cells"[All Fields] OR ("leukocyte count"[MeSH Terms] OR ("leukocyte"[All Fields] AND "count"[All Fields]) OR "leukocyte count"[All Fields] OR "wbc"[All Fields]) OR "leukocyte classification"[All Fields]) AND "blood"[MeSH Terms]) AND (y_5[Filter])
Embase (Elsevier)	('Artificial Intelligence'/exp OR 'Machine Learning'/exp OR 'Deep Learning'/exp OR 'Neural Networks, Computer'/exp OR 'artificial intelligence' OR 'machine learning' OR 'deep learning' OR 'neural networks') AND ('Image Interpretation, Computer-Assisted'/exp OR 'computer vision' OR 'image analysis' OR 'image processing') AND (Leukocytes/exp OR 'white blood cells' OR WBC OR 'leukocyte classification') AND ('Peripheral Blood'/exp OR 'blood smear' OR 'peripheral blood') AND ('automated classification' OR 'automatic detection' OR 'cell classification')
Scopus Elsevier	("Artificial Intelligence" OR "Machine Learning" OR "Deep Learning" OR "Neural Networks, Computer" OR "artificial intelligence" OR "machine learning" OR "deep learning" OR "neural networks") AND ("Image Interpretation, Computer-Assisted" OR "computer vision" OR "image analysis" OR "image processing") AND (Leukocytes OR "white blood cells" OR WBC OR "leukocyte classification") AND ("Peripheral Blood" OR "blood smear" OR "peripheral blood") AND ("automated classification" OR "automatic detection" OR "cell classification")

PubMed/MEDLINE = *United States National Library of Medicine/Medical Literature Analysis and Retrieval System Online*

Fonte: Elaborado pelos Autores (2025).

O processo de seleção dos estudos foi realizado por dois revisores independentes, com um terceiro para resolver divergências, e está detalhado no fluxograma PRISMA (Figura 1). A extração de dados focou nas características dos estudos, algoritmos, conjuntos de dados e métricas de desempenho. A qualidade metodológica e o risco de viés foram avaliados com a ferramenta QUADAS-2, e a certeza da evidência foi classificada utilizando a abordagem GRADE.

3 RESULTADOS

A busca inicial nas bases de dados identificou 300 registros. Após a remoção de 70 duplicatas, 230 artigos foram triados por título e resumo, resultando na seleção de 49 estudos para a análise do texto completo. Nesta etapa de elegibilidade, 21 artigos foram excluídos

por não atenderem aos critérios de inclusão. Os principais motivos para exclusão foram a não abordagem da classificação leucocitária (n=5), o uso de material biológico inadequado (n=4) e baixo rigor metodológico (n=4). O processo detalhado de seleção está apresentado no fluxograma PRISMA (Figura 1), e os 28 estudos incluídos estão caracterizados no Quadro 2.

A qualidade metodológica e o risco de viés dos 28 estudos foram avaliados com a ferramenta QUADAS-2. Os resultados consolidados são apresentados visualmente nos gráficos de risco de viés (Figura 2). A análise demonstrou um baixo risco de viés na maioria dos domínios, especialmente no Teste Índice (D2) e no Fluxo e Tempo (D4), indicando que os algoritmos foram, em geral, bem descritos e aplicados de forma robusta. O principal ponto de preocupação foi identificado no Domínio 3 (Padrão de Referência), onde uma parcela dos estudos foi classificada como de alto risco por utilizar dados sintéticos como padrão-ouro para validação, o que limita a aplicabilidade clínica dos seus achados.

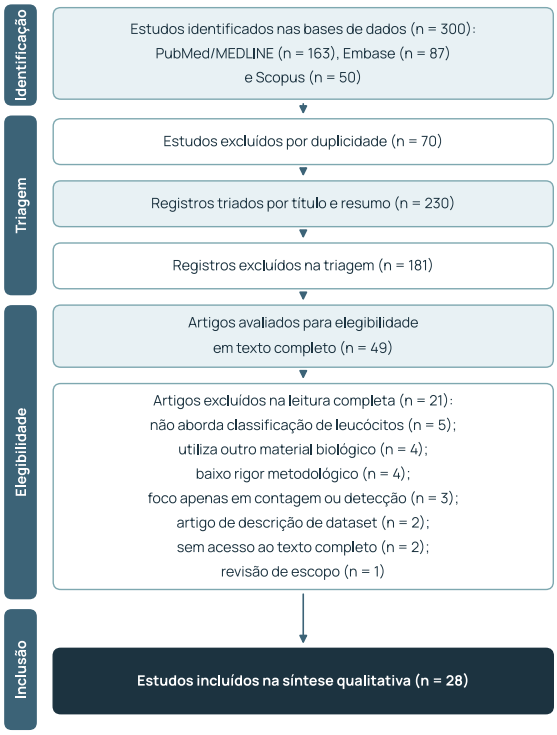
A síntese da evidência foi realizada utilizando a metodologia GRADE, e os resultados estão detalhados na Tabela de Resumo dos Achados (Tabela 1). A análise consolidada revela que, apesar das altas acurácias reportadas individualmente, a certeza geral na evidência é classificada como "Muito Baixa".

A classificação de "Muito Baixa" certeza foi determinada por fragilidades metodológicas e de relato significativas presentes no corpo da literatura. As principais razões para o rebaixamento da evidência incluíram o risco de viés pelo uso de dados sintéticos como padrão de referência, a extrema variedade de modelos e protocolos experimentais (inconsistência), a aplicabilidade clínica limitada de alguns estudos (evidência indireta) e, de forma crítica, a falha quase unânime em reportar intervalos de confiança para as métricas de desempenho (imprecisão).

4 DISCUSSÃO

Esta revisão sistemática de 28 artigos demonstra a notável evolução e diversidade das abordagens de inteligência artificial para a classificação leucocitária. Uma observação central é o predomínio de modelos de aprendizado profundo que vão desde arquiteturas clássicas de Redes Neurais Convolucionais (CNN) até sistemas de detecção em tempo real e modelos híbridos complexos, e que consistentemente atingem alto desempenho, com a maioria dos estudos relatando acurácias superiores a 95%, evidenciando a maturidade

Figura 1. Fluxograma PRISMA do processo de busca e seleção dos estudos.



Fonte: Elaborado pelos Autores (2025).

da área para aplicações de auxílio diagnóstico. Até o momento, não foram identificadas na literatura outras revisões sistemáticas que consolidassem essas abordagens, o que reforça a originalidade e a principal contribuição deste trabalho: a síntese e a análise crítica das metodologias, algoritmos e resultados que definem o estado da arte no campo.

Os primeiros trabalhos no campo frequentemente utilizaram modelos seminais como a AlexNet e a VGG16 como base. Essas redes, embora mais simples que as atuais, foram cruciais para estabelecer a viabilidade das CNNs na extração de características morfológicas de imagens de células sanguíneas [6,7].

Posteriormente, observou-se a consolidação de arquiteturas mais profundas e robustas como ResNet e DenseNet. Esses modelos tornaram-se um padrão na área, sendo amplamente utilizados tanto para a extração direta de características, que alimentam outros classificadores, quanto como a espinha dorsal (*backbone*) para modelos mais complexos e customizados [18-20].

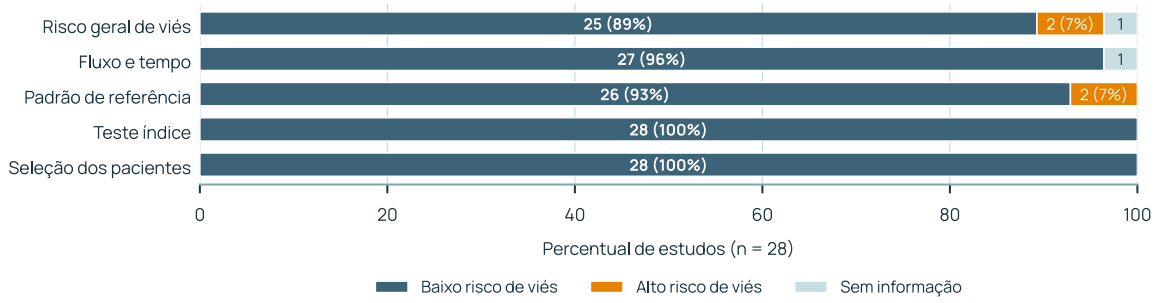
Uma tendência mais recente e notável é a adoção de modelos da família YOLO (*You Only Look Once*), que evoluíram para versões como YOLOv7, v8, v10 e v11. Essa mudança indica um foco crescente não apenas na acurácia da classificação final, mas também na eficiência e na aplicação prática em sistemas de dois estágios, onde a detecção ou segmentação em tempo real precede a classificação detalhada. Paralelamente, a necessidade de implementação em dispositivos com

Quadro 2. Caracterização dos estudos incluídos na revisão sistemática.

Autoria / Periódico	Conjunto de dados	Algoritmo(s)	Resultados
Zhai Q, et al. (2022) [20] – *Journal of Medical and Biological Engineering*	8.915 sub-imagens de esfregaços de sangue (próprio)	Segmentação: Thresholding Classificação: Deeply Aggregated CNN	Acurácia média 98,43%; F1 98,47%
Manthouri M, et al. (2022) [14] – *Computational and Mathematical Methods in Medicine*	LISC e WBCis	Segmentação: Gram-Schmidt Classificação: SIFT + CNN + WTPTSSR	Segmentação: 96,49–97,33%; Classificação: 97,14%
Houssein EH, et al. (2023) [19] – *Frontiers in Oncology*	BCCD (GitHub)	Segmentação: BO-FRS e Hybrid Ellipse Fitting Classificação: DenseNet-161 + CLR	Acurácia 99,8%
Zhu Z, Wang SH, Zhang YD (2023) [18] – *Technology in Cancer Research & Treatment*	Kaggle, 3 tipos de células	Extração: ResNet50 Classificação: Ensemble de 3 redes (SNN, ELM, dRVFL) – ReRNet	Acurácia média 99,97%; F1 99,97%
Sazak H, Kotan M (2025) [17] – *Diagnostics*	BCCD	YOLOv10 / YOLOv11	mAP geral 93,8%; acurácia WBC 99,0%, plaquetas 91,8%
Wang X, et al. (2025) [15] – *Scientific Reports*	BCCD, LDWBC, LISC, Raabin.	Detecção: YOLOv7 aprimorado Classificação: EfficientNetv2 aprimorado	Detecção BCCD: 94,7%; Classificação: precisão 95,12%, recall 97%
Kalnoor G, Kadrolli V (2025) [16] – *Scientific Reports*	ALL-IDB, Cell Morphological Dataset, WBC-3K	Segmentação: DBO + Levy flight Classificação: ADCU-Net	Acurácia 98,4%; Dice 0,94
Jung C, et al. (2022) [4] – *BMC Medical Imaging*	Dataset privado + LISC	Classificação: W-Net Geração de imagens: DCGAN	Acurácia 97% (privado), 96% (LISC)
Chen X, et al. (2023) [3] – *Journal of Biophotonics*	PBC e BCISCIK	Auto-supervisionado com CNN	PBC: 88,88%; BCISCIK: 71,82%
Fang T, et al. (2024) [6] – *Cytometry Part A*	CellaVision, BCCD, ALL-IDB2, PBC-HCB.	• Segmentação: K-means, U-Net, U-Net+ResNet. • Classificação: AlexNet, VGG16, ResNet18.	ResNet18: 100% CellaVision, 98,49% PBC-HCB
Khan S, et al. (2024) [11] – *Computers in Biology and Medicine*	PBC, LISC, Raabin-WBC.	CNN MobileNetV2 + Rede de Atenção Dupla (DA) + DCGAN	Acurácia 99,83% (PBC), 99,35% (LISC), 99,60% (Raabin-WBC).
Tavakoli S, et al. (2021) [10] – *Scientific Reports*	Raabin-WBC, LISC, BCCD.	Extração de cor e forma + SVM	Acurácia 92,21–94,65%
Anand V, et al. (2024) [5] – *BMC Medical Imaging*	Kaggle, 12.444 imagens (4 classes)	Segmentação: processamento de imagem Classificação: CNN própria	Acurácia 97,98%; precisão 97,97%
Özcan ŞN, Uyar T, Karayeğen G (2024) [9] – *Cytometry Part A*	Dataset misto (CellaVision, Raabin-WBC, Kaggle)	CNN própria + UNet + SegNet	Acurácia balanceada: 97,27–98,03%
Ametefe DS, et al. (2024) [8] – *International Journal of Laboratory Hematology*	C_NMC_2019, ALL IDB	Classificação: EfficientNetB3 (ALLDet) Segmentação: Chan-Vese	Acurácia geral 97,13% (LLA)
Zhang C, et al. (2020) [24] – *Medical Physics*	Dataset privado (Hospital Mianyang, China)	Rede híbrida adversária-discriminativa (HADN) baseada em ResNet	Acurácia média 96,53 ±1,14%
Reena MR, Ameer PM (2020) [7] – *Computers in Biology and Medicine*	LISC	Segmentação: DeepLabv3+ Classificação: AlexNet + transferência	mAP 98,42% (segmentação); acurácia 98,87 ±1%
Patil AM, et al. (2021) [23] – *IRBM*	BCCD	CNN + RNN (LSTM) + CCA	Acurácia Xception-LSTM: 95,89%
Girdhar A, et al. (2022) [22] – *Biomedical Signal Processing and Control*	BCCD	CNN própria	Acurácia 98,55%; F1 97,11%
Hosseini M, et al. (2022) [29] – *Expert Systems with Applications*	BCCD	CNN otimizado por GWO-GA	Acurácia 97–99%; sensibilidade 94%; precisão 95%
Bayat N, et al. (2022) [28] – *Big Data and Cognitive Computing*	PBC (Kaggle)	CNN (EfficientNet) + atenção	Acurácia 99,69%; robustez mantida mesmo com redução de dados (<1%)
Cheuque C, et al. (2022) [21] – *Diagnostics*	BCD, CBC, WBC, KBC, LISC.	Detecção: Faster R-CNN Classificação: CNN MobileNet paralelas	Métricas de performance ~98,4%
Ahmad R, et al. (2023) [27] – *Diagnostics*	Dataset privado (Hospital Militar Combinado, Paquistão)	DenseNet201/Darknet53 + ECMPA + Fine KNN	Acurácia 99,58%
Abou Ali M, et al. (2023) [26] – *Algorithms*	BCCD, PBC	Comparativo: CNN própria, VGG16, ViT	Melhor resultado ViT: 97,1%
Khan S, et al. (2025) [13] – *Computers in Biology and Medicine*	Raabin-WBC, PBC, LISC	Segmentação: YOLOv8m-seg Classificação: MobileNetV3	Acurácias 98,49–99,63%
Shaik A, et al. (2024) [25] – *International Journal of Imaging Systems and Technology*	BCCD	YOLOv5 customizado	Precisão 99,20%; F1 99%; sensibilidade 96,9%
Al-qudah R, Suen CY (2021) [30] – *Computers in Biology and Medicine*	ALL-IDB	Pipeline 1: ResNet50 + processamento de imagem Pipeline 2: VGG19/VGG16 para WBCs e plaquetas	Plaquetas 98,6%; leucócitos 95%; combinada 96,7%, superando referência
Almezhghwi K, Serte S (2020) [12] – *Computational Intelligence and Neuroscience*	LISC	VGG-16, ResNet, DenseNet com GAN	DenseNet-169 acurácia 98,8%

Fonte: Elaborado pelos Autores (2025).

Figura 2. Gráfico de Risco de Viés.



Fonte: Elaborado pelos Autores (2025).

Tabela 1. Resumo dos achados segundo a abordagem GRADE

Desfecho	Nº de Estudos	Resumo dos Efeitos (Acurácia Reportada)	Qualidade/Certeza da Evidência (GRADE)	Comentários e Justificativa para a Classificação
Acurácia na Classificação de Leucócitos	28 estudos (Nº de imagens por estudo varia de ~242 a >12.000)	A maioria dos estudos reporta acurácias muito altas, geralmente na faixa de 95% a 99% para a classificação de 4 a 5 subtipos de leucócitos.	⨁◯◯◯ MUITO BAIXA	A certeza da evidência foi rebaixada devido a sérias limitações em múltiplos domínios: Risco de Viés, Inconsistência, Evidência Indireta e Imprecisão.

Fonte: Elaborado pelos Autores (2025).

menor capacidade computacional, como analisadores hematológicos automatizados ou plataformas de telediagnóstico, impulsionou o uso de modelos leves como a MobileNet e a EfficientNet. Essas arquiteturas buscam um equilíbrio otimizado entre alta acurácia e baixo custo computacional [11,17,21,25].

Finalmente, a exploração de novos paradigmas para além das CNNs começa a emergir com a introdução dos *Vision Transformers* (ViT), que representam uma nova e promissora frente de pesquisa, com estudos comparativos já demonstrando um desempenho superior para esta tarefa específica [26].

A partir da análise dos artigos, emergem duas estratégias metodológicas principais para a classificação automatizada de leucócitos: sistemas de estágio único e de dois estágios. A abordagem de estágio único utiliza um modelo de ponta a ponta, geralmente uma CNN como ResNet50 ou DenseNet-161, que recebe a imagem de uma célula e realiza diretamente a sua classificação em uma única pipeline integrada [19].

Em contrapartida, a abordagem de dois estágios, cada vez mais comum na literatura recente, divide o problema em duas tarefas sequenciais. Primeiramente, um algoritmo é dedicado à detecção e segmentação precisa dos leucócitos, isolando-os de outras células e do fundo da imagem, utilizando técnicas que vão desde a limiarização (*thresholding*) até modelos de detecção de objetos de alta performance como a família YOLO. A principal vantagem desta abordagem é que a segunda etapa, a classificação, recebe como entrada uma imagem "limpa" e focada, contendo apenas a região de interesse (ROI) do leucócito. Isso pode aumentar a precisão, pois o classificador (como MobileNetV3 ou EfficientNetv2) não precisa lidar com o ruído e a complexidade do fundo, concentrando seu poder de aprendizado exclusivamente nas características morfológicas relevantes da célula. Essa separação de tarefas, embora potencialmente mais complexa em sua implementação, permite otimizar cada etapa de forma independente, visando um resultado final mais acurado [13-15].

Além da escolha da arquitetura principal, a análise dos artigos revela um esforço consistente dos pesquisadores em empregar técnicas avançadas para extrair o máximo de desempenho dos modelos, refinando a capacidade de classificação em tarefas de alta complexidade. Esses recursos adicionais podem ser agrupados em quatro categorias principais:

- Mecanismos de Atenção: Uma técnica proeminente é a incorporação de mecanismos de atenção, que permitem ao modelo focar dinamicamente nas características morfológicas mais discriminativas da célula, como a textura do núcleo ou a granulação do citoplasma. Ao ponderar a importância de diferentes regiões da imagem, esses mecanismos melhoram a distinção entre classes com alta similaridade visual, como demonstrado em arquiteturas como a ADCU-Net e em modelos baseados em YOLOv7 com atenção multi-cabeça [16].
- Aumento e Geração de Dados: Para contornar a limitação de dados, especialmente para classes de leucócitos mais raras que podem levar a um treinamento desbalanceado, diversos estudos empregam estratégias de aumento e geração de

dados. Notavelmente, o uso de Redes Adversárias Geradoras (GANs), especificamente a DCGAN (*Deep Convolutional Generative Adversarial Network*), foi relatado como uma abordagem para criar imagens sintéticas realistas. Essa técnica enriquece o conjunto de treinamento, melhora a robustez do classificador e ajuda a mitigar o sobreajuste (*overfitting*) [4-11].

- Modelos em *Ensemble*: Outra abordagem para aumentar a precisão e a confiabilidade é o uso de modelos em *ensemble*. Em vez de confiar em um único classificador, essa técnica combina as predições de múltiplos modelos — sejam eles da mesma arquitetura ou de diferentes — e agrega os resultados, frequentemente por meio de um sistema de votação majoritária. Essa estratégia tende a mitigar erros individuais de cada modelo e a produzir um resultado final mais estável e acurado [18].
- Otimização de Hiperparâmetros: Finalmente, a otimização de hiperparâmetros por meio de métodos sofisticados também foi observada como um fator chave para o desempenho. Técnicas que utilizam algoritmos genéticos e de otimização de enxames (como o GWO-GA) ou estratégias de treinamento como a Taxa de Aprendizagem Cíclica (CLR) são aplicadas para realizar um ajuste fino do processo de treinamento. O objetivo é encontrar a combinação ideal de parâmetros (como taxa de aprendizado e momentum) que garanta uma convergência mais rápida e um desempenho superior do modelo [29].

A base de evidências para a classificação de leucócitos é amplamente construída sobre *datasets* públicos e bem conhecidos, como o BCCD, LISC, PBC e Raabin-WBC. A utilização recorrente desses conjuntos de dados oferece a vantagem superficial de permitir um certo nível de comparabilidade entre os estudos. Contudo, essa comparabilidade é frequentemente ilusória. Reconhecendo os limites de treinar e validar modelos em um único *dataset*, o que invariavelmente leva a um sobreajuste (*overfitting*) e a uma baixa capacidade de generalização para novos dados, vários estudos buscaram provar a robustez de seus algoritmos. Abordagens notáveis incluem o uso combinado de múltiplos *datasets* ou, de forma ainda mais rigorosa, a validação de um modelo contra um conjunto de teste completamente independente [6-26]. Esse desafio de validação externa é central, pois expõe a fragilidade de muitos modelos quando confrontados com as variações do mundo real, como diferentes protocolos de coloração, scanners de lâminas e perfis de pacientes, que não estavam presentes nos dados de treinamento.

A principal força desta revisão sistemática reside na sua abrangência, consolidando as diversas arquiteturas e estratégias que definem o estado da arte. No entanto, é na análise crítica de suas limitações que emergem as barreiras mais significativas para a translação clínica. A significativa heterogeneidade metodológica é a limitação mais proeminente, com implicações clínicas diretas. Mesmo quando utilizam o mesmo *dataset* público, os autores aplicam diferentes técnicas de pré-processamento, divisão de dados e aumento sintético, o que impede uma comparação direta de desempenho e torna impossível para um laboratório clínico determinar, com base na literatura, qual o algoritmo mais robusto para sua implementação. Essa falta de padronização leva a um cenário onde métricas impressionantes, como acurácias acima de 99%, perdem o significado prático, pois não são reprodutíveis em diferentes contextos.

Além disso, a inconsistência no reporte de métricas representa um risco de viés com consequências clínicas graves. Muitos artigos relatam apenas a acurácia geral, uma medida notoriamente enganosa em *datasets* com desbalanceamento de classes, como os de hematologia. Um modelo pode, por exemplo, atingir 99% de acurácia ao identificar corretamente milhares de neutrófilos, mas falhar em reconhecer os poucos blastos presentes, que são o achado mais crítico para o diagnóstico de uma leucemia aguda. A ausência de um padrão para o reporte de matrizes de confusão detalhadas e métricas por classe (precisão, sensibilidade) oculta essas fraquezas e impede uma avaliação real da segurança do modelo para uso diagnóstico. Essa variabilidade metodológica, portanto, não apenas impede a realização de uma metanálise quantitativa, mas também representa a principal barreira de confiança para a adoção clínica.

Apesar desses desafios, a aplicabilidade prática das tecnologias revisadas é inegável, especialmente quando posicionadas como ferramentas de apoio, e não como substitutos do analista. Em laboratórios clínicos, os modelos podem ser integrados a sistemas de patologia digital para funcionar como uma triagem inteligente, analisando milhares de células em segundos para separar os esfregaços normais (que poderiam ser liberados com mínima supervisão) daqueles que exigem atenção detalhada de um especialista. Outra aplicação é como uma ferramenta de aumento de produtividade, onde o software pré-localiza e pré-classifica os leucócitos em uma lâmina digital, permitindo que o analista revise e confirme os achados de forma muito mais rápida e ergonômica, além de ajudar a padronizar a classificação entre diferentes profissionais. Em plataformas digitais de saúde, essas ferramentas são a espinha dorsal da telehematologia, permitindo que um laboratório em uma área remota digitalize uma lâmina e receba uma análise preliminar precisa, que pode ser validada à distância por um hematologista, democratizando o acesso a diagnósticos de alta qualidade.

No entanto, a segunda limitação crucial identificada é o escopo da classificação. A grande maioria dos estudos

foca na diferenciação de um número limitado de leucócitos maduros. Embora os modelos demonstrem alta performance nessa tarefa, o verdadeiro desafio e a maior necessidade clínica na hematologia residem na identificação de células imaturas e atípicas, cuja morfologia é sutil e a classificação, complexa. Esta lacuna indica que, apesar do enorme progresso na automação da contagem diferencial de rotina, os modelos atuais ainda não abordam o problema de maior impacto diagnóstico. O avanço futuro do campo depende, portanto, da criação de *datasets* robustos e anotados por especialistas, que incluam toda a gama de anormalidades morfológicas, para que a inteligência artificial possa evoluir de uma ferramenta de automação para um verdadeiro parceiro no diagnóstico hematológico de precisão.

5 CONCLUSÃO

Esta revisão sistemática consolida as evidências sobre a aplicação do aprendizado profundo na classificação de leucócitos, evidenciando que diferentes arquiteturas, especialmente CNNs e modelos da família YOLO, atingem alta acurácia para a diferenciação de leucócitos maduros. Entretanto, a análise crítica mostra que a heterogeneidade metodológica entre os estudos, a dependência de bases de dados públicas limitadas, a ausência de validações externas consistentes e a escassez de métricas detalhadas ainda limitam a confiança e a generalização dos achados.

A principal lacuna clínica permanece na detecção de células imaturas e atípicas, de maior relevância diagnóstica. Nesse cenário, embora os resultados sejam encorajadores, o campo exige validação mais robusta em dados representativos do mundo real, além da padronização de protocolos experimentais. O fortalecimento dessas evidências será fundamental para que tais avanços tecnológicos possam evoluir de forma segura e contribuir efetivamente no apoio ao diagnóstico hematológico.

Contribuição dos autores. Os autores contribuíram igualmente em todas as etapas de construção do artigo: concepção e delineamento do estudo; coleta, análise e interpretação dos dados; redação e revisão crítica do manuscrito; e aprovação da versão final.

Financiamento. Os autores não declararam fonte de financiamento.

Disponibilidade de dados e materiais. Os autores não declararam a disponibilidade dos dados e materiais da pesquisa.

Conflitos de interesse. Os autores não declararam conflitos de interesse.

Declaração de uso de IA. Os autores não declararam uso de ferramentas de Inteligência Artificial.

REFERÊNCIAS

1. Asghar R, Kumar S, Shaukat A, Hynds P. Classification of white blood cells (leucocytes) from blood smear imagery using machine and deep learning models: a global scoping review. PLoS One. 2024;19(6):e0292026. doi: https://doi.org/10.1371/journal.pone.0292026
2. Kratz A, Lee SH, Zini G, Riedl JA, Hur M, Machin S, et al. Digital morphology analyzers in hematology: ICSH review and recommendations. Int J Lab Hematol. 2019;41(4):437-47. doi: https://doi.org/10.1111/ijlh.13042
3. Chen X, Zheng G, Zhou L, Li Z, Fan H. Deep self-supervised transformation learning for leukocyte classification. J Biophotonics. 2023;16(3):e202200244. doi: https://doi.org/10.1002/jbio.202200244
4. Jung C, Abuhamad M, Mohaisen D, Han K, Nyang D. WBC image classification and generative models based on convolutional neural network. BMC Med Imaging. 2022;22(1):94. doi: https://doi.org/10.1186/s12880-022-00818-1
5. Anand V, Gupta S, Koundal D, Alghamdi WY, Alsharbi BM. Deep learning-based image annotation for leukocyte segmentation and classification of blood cell morphology. BMC Med Imaging. 2024;24(1):83. doi: https://doi.org/10.1186/s12880-024-01254-z
6. Fang T, Huang X, Chen X, Chen D, Wang J, Chen J. Segmentation, feature extraction and classification of leukocytes leveraging neural networks, a comparative study. Cytometry A. 2024;105(7):536-46. doi: https://doi.org/10.1002/cyto.a.24832
7. Reena MR, Ameer PM. Localization and recognition of leukocytes in peripheral blood: a deep learning approach. Comput Biol Med. 2020;126:104034. doi: https://doi.org/10.1016/j.compbiomed.2020.104034
8. Ametefe DS, Sarnin SS, Ali DM, Ametefe GD, John D, Aliu AA, et al. Automatic classification and segmentation of blast cells using deep transfer learning and active contours. Int J Lab Hematol. 2024;46(5):837-49. doi: https://doi.org/10.1111/ijlh.14305
9. Özcan ŞN, Uyar T, Karayeğen G. Comprehensive data analysis of white blood cells with classification and segmentation by using deep learning approaches. Cytometry A. 2024;105(7):501-20. doi: https://doi.org/10.1002/cyto.a.24839
10. Tavakoli S, Ghaffari A, Kouzehkanan ZM, Hosseini R. New segmentation and feature extraction algorithm for classification of white blood cells in peripheral smear images. Sci Rep. 2021;11(1):19428. doi: https://doi.org/10.1038/s41598-021-98599-0
11. Khan S, Sajjad M, Abbas N, Escorcia-Gutierrez J, Gamarra M, Muhammad K. Efficient leukocytes detection and classification in microscopic blood images using convolutional neural network coupled with a dual attention network. Comput Biol Med. 2024;174:108146. doi: https://doi.org/10.1016/j.compbiomed.2024.108146
12. Almezhghwi K, Serte S. Improved classification of white blood cells with the generative adversarial network and deep convolutional neural network. Comput Intell

Neurosci. 2020;2020:6490479. doi: https://doi.org/10.1155/2020/6490479

13. Khan S, Sajjad M, Escorcia-Gutierrez J, Dhahbi S, Hijji M, Muhammad K. Two-stage CNN-based framework for leukocytes classification. Comput Biol Med. 2025;187:109616. doi: https://doi.org/10.1016/j.compbiomed.2024.109616
14. Manthouri M, Aghajari Z, Safary S. Computational intelligence method for detection of white blood cells using hybrid of convolutional deep learning and SIFT. Comput Math Methods Med. 2022;2022:9934144. doi: https://doi.org/10.1155/2022/9934144
15. Wang X, Pan G, Hu Z, Ge A. A two stage blood cell detection and classification algorithm based on improved YOLOv7 and EfficientNetv2. Sci Rep. 2025;15(1):8427. doi: https://doi.org/10.1038/s41598-025-91720-7
16. Kalnoor G, Kadrolli V. Advanced leukocyte classification using attention mechanisms and dual channel U-Net architecture. Sci Rep. 2025;15(1):13825. doi: https://doi.org/10.1038/s41598-025-96918-3
17. Sazak H, Kotan M. Automated blood cell detection and classification in microscopic images using YOLOv11 and optimized weights. Diagnostics (Basel). 2025;15(1):22. doi: https://doi.org/10.3390/diagnostics15010022
18. Zhu Z, Wang SH, Zhang YD. ReRNet: a deep learning network for classifying blood cells. Technol Cancer Res Treat. 2023;22:15330338231165856. doi: https://doi.org/10.1177/15330338231165856
19. Houssein EH, Mohamed O, Abdel Samee N, Mahmoud NF, Talaat R, Al-Hejri AM, et al. Using deep DenseNet with cyclical learning rate to classify leukocytes for leukemia identification. Front Oncol. 2023;13:1230434. doi: https://doi.org/10.3389/fonc.2023.1230434
20. Zhai Q, Fan B, Zhang B, Li JH, Liu JZ. Automatic white blood cell classification based on whole-slide images with a deeply aggregated neural network. J Med Biol Eng. 2022;42(1):126-37. doi: https://doi.org/10.1007/s40846-022-00683-x
21. Cheuque C, Querales M, León R, Salas R, Torres R. An efficient multi-level convolutional neural network approach for white blood cells classification. Diagnostics (Basel). 2022;12(2):248. doi: https://doi.org/10.3390/diagnostics12020248
22. Girdhar A, Kapur H, Kumar V. Classification of white blood cell using convolution neural network. Biomed Signal Process Control. 2022;71:103156. doi: https://doi.org/10.1016/j.bspc.2021.103156
23. Patil AM, Patil MD, Birajdar GK. White blood cells image classification using deep learning with canonical correlation analysis. IRBM. 2021;42(5):378-89. doi: https://doi.org/10.1016/j.irbm.2020.08.005
24. Zhang C, Wu S, Lu Z, Shen Y, Wang J, Huang P, et al. Hybrid adversarial-discriminative network for leukocyte classification in leukemia. Med Phys. 2020;47(8):3732-44. doi: https://doi.org/10.1002/mp.14144
25. Shaik A, Tiwari A, Neelapu B, Jain PK, Banoth E. Advancing leukocyte classification: a cutting-edge deep learning approach for AI-driven clinical diagnosis. Int J Imaging Syst Technol. 2024;34(6):e23204. doi: https://doi.org/10.1002/ima.23204
26. Abou Ali M, Dornaika F, Arganda-Carreras I. White blood cell classification: convolutional neural network (CNN) and vision transformer (ViT) under medical microscope. Algorithms. 2023;16(11):525. doi: https://doi.org/10.3390/a16110525
27. Ahmad R, Awais M, Kausar N, Akram T. White blood cells classification using entropy-controlled deep features optimization. Diagnostics (Basel). 2023;13(3):352. doi: https://doi.org/10.3390/diagnostics13030352
28. Bayat N, Davey DD, Coathup M, Park JH. White blood cell classification using multi-attention data augmentation and regularization. Big Data Cogn Comput. 2022;6(4):122. doi: https://doi.org/10.3390/bdcc6040122
29. Hosseini M, Bani-Hani D, Lam SS. Leukocytes image classification using optimized convolutional neural networks. Expert Syst Appl. 2022;205:117672. doi: https://doi.org/10.1016/j.eswa.2022.117672
30. Al-qudah R, Suen CY. Improving blood cells classification in peripheral blood smears using enhanced incremental training. Comput Biol Med. 2021;131:104265. doi: https://doi.org/10.1016/j.compbiomed.2021.104265
31. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. doi: https://doi.org/10.1136/bmj.n71

NOTAS EDITORIAIS

O Journal of Health Informatics é uma publicação científica mantida pela Sociedade Brasileira de Informática em Saúde, associação científica sem fins lucrativos dedicada ao desenvolvimento da Informática em Saúde e da Saúde Digital no Brasil. O processo editorial é conduzido com independência científica, sem interferência institucional, comercial ou de patrocinadores, e segue as boas práticas internacionais de ética em publicação científica.

Este artigo foi avaliado por pares, em sistema duplo-cego, por especialistas da área. A equipe editorial usou ferramentas de inteligência artificial, com supervisão humana, na triagem inicial, na conferência do texto, das referências e dos números e, quando necessário, no redesenho de figuras, quadros e tabelas, aprovado pelos autores na prova. As ferramentas e as salvaguardas estão na Política de uso de inteligência artificial da revista. As decisões editoriais foram tomadas por pessoas.

SBIS Sociedade Brasileira de Informática em Saúde