

João Virgilio Bezerra Martins
Orientador: Leonardo Pereira de Sousa

Revisão Sistemática dos Métodos de Aprendizagem Profundo para Detecção e Classificação da Leucemia Mieloide Aguda

Picos - PI
1 de fevereiro de 2024

João Virgilio Bezerra Martins
Orientador: Leonardo Pereira de Sousa

Revisão Sistemática dos Métodos de Aprendizagem Profundo para Detecção e Classificação da Leucemia Mieloide Aguda

Monografia submetida ao Curso de Bacharelado em Sistemas de Informação como requisito parcial para obtenção de grau de Bacharel em Sistemas de Informação.

Universidade Federal do Piauí
Campus Senador Heuvídio Nunes de Barros
Bacharelado em Sistemas de Informação

Picos - PI
1 de fevereiro de 2024

FICHA CATALOGRÁFICA
Serviço de Processamento Técnico da Universidade Federal do Piauí
Biblioteca José Albano de Macêdo

M386r Martins, João Virgílio Bezerra.

Revisão sistemática dos métodos de aprendizagem profundo para detecção e classificação da Leucemia Mieloide Aguda./ João Virgílio Bezerra Martins. – 2024.

43 f.

1 Arquivo em PDF

Indexado no catálogo *online* da biblioteca José Albano de Macêdo-CSHNB
Aberto a pesquisadores, com restrições da Biblioteca

Trabalho de Conclusão de Curso (Graduação) – Universidade Federal do Piauí, Curso de Bacharelado em Sistemas de Informação, Picos, 2024.

“Orientador: Prof. Leonardo Pereira de Sousa”

1. Leucemia Mieloide. 2. Revisão sistemática. 3. Aprendizagem-métodos.

4. Bioacústica. I. Sousa, Leonardo Pereira de. II. Título.

CDD 616.155

Elaborado por Sérvulo Fernandes da Silva Neto CRB 15/603

REVISÃO SISTEMÁTICA DOS MÉTODOS DE APRENDIZAGEM PROFUNDO PARA
DETECÇÃO E CLASSIFICAÇÃO DA LEUCEMIA MIELOIDE AGUDA

JOÃO VIRGÍLIO BEZERRA MARTINS

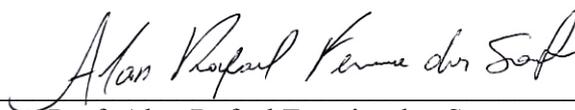
Monografia aprovada como exigência parcial para obtenção do grau de Bacharel em Sistemas
de Informação.

Data de Aprovação

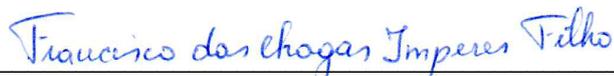
Picos – PI, 01 de fevereiro de 2024



Prof. Leonardo Pereira de Sousa



Prof. Alan Rafael Ferreira dos Santos



Prof. Francisco das Chagas Imperes Filho

Agradecimentos

Primeiramente quero expressar gratidão aos meus pais, Conceição da Silva Bezerra Martins e Joaquim Rocha Martins. Sem o apoio deles, não teria alcançado essa conquista em minha jornada acadêmica. Seus esforços e orientações foram importantes para meu crescimento. Também não posso deixar de agradecer minhas duas irmãs Lorena Bezerra Martins e Isabela Bezerra Martins, elas sempre me motivam. Agradeço de coração por tudo que fazem por mim.

Durante a minha jornada acadêmica, conheci pessoas incríveis que me acolheram. Agradeço ao Anderson Lourenço, meu irmão, e à minha segunda mãe, Antônia que infelizmente não poderá comemorar esta conquista comigo. Vocês são como uma segunda família, e sem o apoio de vocês, nada disso seria possível.

*Coisas incríveis no mundo dos negócios nunca são feitas por uma única pessoa, mas sim
por uma equipe.*

Steve Jobs

Resumo

A Leucemia Mieloide Aguda (LMA) é um tipo de câncer que começa na medula óssea e é caracterizado por um crescimento anormal dos glóbulos brancos. É uma doença grave que faz parte do grupo de neoplasias malignas que afeta pessoas do mundo todo. O hematologista é um médico especializado em hematologia, que é o estudo dos elementos formados no sangue e dos órgãos relacionados, como a medula óssea. Esses profissionais diagnosticam e tratam distúrbios do sangue. O diagnóstico da doença é desafiador devido à ampla gama de condições, que vão de anemias simples até leucemias complexas. A exigência de conhecimento extenso e atualizado significa que a falta de experiência ou o cansaço do hematologista podem levar a diagnósticos menos precisos. Portanto, a integração de tecnologia para a detecção e classificação da Leucemia Mieloide Aguda (LMA) é uma aliada valiosa, mantendo a qualidade diagnóstica mesmo diante de desafios. Esta revisão sistemática realiza uma análise abrangente do método de Deep Learning aplicado à classificação e detecção de Leucemia Mieloide Aguda (LMA). Utilizando as bases bibliográficas IEEE Xplore, Web of Science, Science Direct e Scopus, a pesquisa identificou artigos relevantes por meio de uma busca por palavras-chave específicas. A seleção de publicações seguiu critérios rigorosos de inclusão e exclusão, abordando seis questões de pesquisa. Os resultados destacam os desafios enfrentados pelos pesquisadores, bem como as diversas etapas, incluindo pré-processamento e aumento de dados, necessárias antes da classificação final. Os resultados desta análise são baseados na revisão de 12 artigos sobre a identificação da LMA, destacaram a principal deficiência no acesso a bases de dados específicas para classificação. Além disso, observou-se uma lacuna na abordagem dos trabalhos em relação à identificação dos tipos e subtipos da LMA. Diante disso, esta pesquisa sugere a necessidade de novas abordagens, especialmente focadas no aumento de dados, visando obter novos tipos de informações e aprimorar as técnicas de detecção.

palavras-chave: Revisão Sistemática, Leucemia Mieloide, Deep Learning, Classificação, Identificação

Abstract

Acute Myeloid Leukemia (AML) is a type of cancer that begins in the bone marrow and is characterized by an abnormal growth of white globes. It is a serious disease that is part of the group of malignant neoplasms that affect people around the world. A hematologist is a doctor specializing in hematology, which is the study of elements formed in the blood and related organs, such as the bone marrow. These professionals diagnose and treat blood disorders. Diagnosis of the disease is required for a wide range of conditions, ranging from simple anemias to complex leukemias. The requirement for extensive and up-to-date knowledge means that a hematologist's lack of experience or fatigue can lead to less accurate diagnoses. Therefore, the integration of technology for the detection and classification of Acute Myeloid Leukemia is an excellent ally, maintaining diagnostic quality even in the face of challenges. This systematic review performs a comprehensive analysis of the Deep Learning method applied to LMA classification and detection. Using IEEE Xplore, Web of Science, Science Direct and Scopus as bibliographic bases, the search for relevant articles through a search for specific keywords. The selection of publications with strict inclusion and exclusion criteria, addressing six research questions. The results highlight the challenges faced by researchers, as well as the various steps including pre-processing and data augmentation, permission before final classification. The results of this analysis are based on the review of 12 articles on the identification of AML, highlighting the main deficiency in access to specific databases for classification. Furthermore, we comment on a gap in the approach of the work in relation to the identification of types and subtypes of AML. Given this, this research suggests the need for new approaches, especially focused on data augmentation, to obtain new types of information and improve detection techniques.

keywords: Systematic Review, Myeloid Leukemia, Deep Learning, Classification, Identification

Lista de ilustrações

Figura 1 – Tipos de Leucemia	16
Figura 2 – Comparação entre Aprendizado de Máquina e Aprendizado Profundo .	21
Figura 3 – Fluxograma da Metodologia	23
Figura 4 – Resultado da Busca	26
Figura 5 – Seleção dos Artigos	27
Figura 6 – Base de Imagens	36
Figura 7 – Comparativo entre as Abordagens	37

Lista de tabelas

Tabela 1 – Classificação da LMA segundo a FAB	18
Tabela 2 – Critérios de Inclusão e Exclusão	25
Tabela 3 – <i>Strings</i> de Buscas	26
Tabela 4 – Extração de Dados	28
Tabela 5 – Bases de Dados	31
Tabela 6 – Técnicas de Pré-processamento e Aumento de Dados	33
Tabela 7 – Abordagens Utilizadas	35

Lista de abreviaturas e siglas

AA	Anemia Aplástica
AIIMS	All India Institute of Medical
AML-IDB	<i>Acute Myeloid Leukemia Image Database</i>
AOI	<i>Area of Interest</i>
ASH	<i>American Society of Hematology</i>
BCCD	<i>Blood Cell Counting Dataset</i>
CNN	<i>Convolutional Neural Network</i>
DL	<i>Deep Learning</i>
DNN	<i>Deep Neural Network</i>
FAB	Britânico-Americano-Francês
HBS	<i>Healthy Blood Slides</i>
LLA	Leucemia Linfoblástica Aguda
LM	Leucemia
LMA	Leucemia Mieloide Aguda
MO	Medula óssea
OMS	Organização Mundial de Saúde.
RS	Revisão Sistemática
SMD	Síndromes Mielodisplásicas
StArt	<i>State of the Art through Systematic Review</i>
SXMU	<i>Shanxi Medical University</i>
ZCA	<i>Zero-phase Component Analysis</i>

Lista de símbolos

inv	Uma inversão em um cromossomo.
p	O braço curto de um cromossomo (a metade superior)
q	O braço longo de um cromossomo (a metade inferior)
t	translocação entre os cromossomos.

Sumário

1	Introdução	14
1.1	Objetivo Geral	15
1.2	Objetivo Específicos	15
1.3	Organização do Trabalho	15
2	Referencial Teórico	16
2.1	Leucemia	16
2.2	A Leucemia Mieloide Aguda (LMA)	17
2.2.1	Classificação da LMA	17
2.2.2	Sistema de classificação FAB	17
2.2.3	Sistema de classificação OMS	18
2.2.4	Diagnóstico	19
2.3	Aprendizado de Máquina	20
2.3.1	Aprendizado Profundo	20
2.3.2	Transferência de Aprendizagem	21
3	Metodologia	23
3.1	Planejamento	23
3.1.1	Questões de pesquisa	24
3.1.2	Termos de Pesquisa	24
3.1.3	Bases de busca bibliográficas	24
3.1.4	Crítérios de Inclusão e Exclusão	25
3.2	Execução	25
3.3	Seleção	26
3.4	Extração	27
4	Análise dos resultados	29
4.1	Aquisição dos Dados	29
4.2	Pré-processamento	31
4.3	Classificação e Validação	33
4.4	Resultados	36
5	Conclusão e Trabalhos Futuros	38
5.1	Aquisição de Imagens	38
5.2	Capacidades de Generalização de Preditores	38
5.3	Gasto Computacional	39

5.4	Classificação dos subtipos da LMA	39
5.5	Conclusão	39
Referências	41

1 Introdução

A Leucemia Mieloide Aguda (LMA) integra o conjunto de doenças malignas e é uma forma de neoplasia maligna originada de células mieloides imaturas, derivadas da Medula Óssea (MO), apresentando variabilidade em sua diferenciação. Este grupo de doenças biologicamente diverso afeta predominantemente a MO e o sangue periférico, mas pode também manifestar-se em tecidos extramedulares. A LMA exibe um curso clínico agressivo, com uma taxa de sobrevida global em cinco anos situada em torno de aproximadamente 25% (NARAYANAN; WEINBERG, 2020).

A Leucemia (LM) é classificada em quatro tipos: agudas e crônicas, que se subdividem em linfóides ou mieloides. Dentro dessa categorização, destaca-se a leucemia mieloide (LM). São classificadas de acordo com o tipo celular envolvido e o estado de maturidade das células leucêmicas. Na fase aguda, ocorre uma proliferação de células da linhagem mieloide, conhecidas como mieloblastos. Isso resulta na produção insuficiente de células sanguíneas maduras e normais na medula óssea e no sangue periférico, levando à substituição do tecido normal (SANTOS et al., 2019).

Existem diversos fatores de predisposição à doença conhecidos, tais como tabagismo, exposição a produtos químicos, uso de medicamentos quimioterápicos, exposição à radiação, presença de doenças sanguíneas (pacientes com condições mieloproliferativas crônicas, como policitemia vera, trombocitopenia essencial e mielofibrose idiopática, apresentam risco aumentado de desenvolver leucemia mieloide aguda), síndromes hereditárias, histórico familiar, idade e gênero. Os pacientes com LMA frequentemente manifestam vários sintomas não específicos, como perda de peso, febre, sudorese noturna, fadiga e perda de apetite (TRESSO, 2021).

Diante dos desafios impostos pela alta incidência e mortalidade da leucemia, especialmente da Leucemia Mieloide Aguda, conforme destacado pelas projeções da SOCIETY (2023), a incorporação de tecnologias inovadoras torna-se crucial. Nesse cenário, o aprendizado profundo emerge como uma ferramenta promissora. Sua aplicação na análise de imagens médicas, conforme descrito por (KUMAR et al., 2023), pode revolucionar o diagnóstico e tratamento da leucemia, aumentando a precisão e eficiência, reduzindo erros e perdas. Assim, a integração do aprendizado profundo na pesquisa e prática médica pode ser um passo significativo na luta contra essa doença devastadora, oferecendo esperança através de diagnósticos mais precoces e tratamentos mais eficazes.

Considerando esse contexto, a pesquisa avançada voltada para a compreensão dos métodos de aprendizagem profunda aplicados à classificação e detecção da LMA suscita uma promissora perspectiva de aprimorar as taxas de sobrevivência e a qualidade de vida dos pacientes afetados por essa condição desafiadora. A aplicação de tecnologias inovadoras nesse campo pode abrir caminho para diagnósticos mais precisos, tratamentos

mais eficientes e, em última instância, melhores prognósticos para aqueles que enfrentam os desafios dessa enfermidade.

1.1 Objetivo Geral

Essa Revisão Sistemática tem objetivo realizar uma análise das técnicas de aprendizado profundo disponíveis e sua aplicação em imagens médicas, com foco principal na classificação e detecção da LMA em imagens microscópicas de sangue.

1.2 Objetivo Específicos

Considerando o objetivo geral da pesquisa, foram formulados objetivos específicos, visando complementar o estudo realizado:

- Buscar as bases de dados utilizadas;
- Identificar métodos de pré-processamento aplicados nas bases de dados;
- Apresentar os tipos de testes conduzidos;
- Destacar qual abordagem mostrou resultados mais favoráveis.

1.3 Organização do Trabalho

Este trabalho está estruturado em seis capítulos. O Capítulo 2 oferece um referencial teórico com conceitos fundamentais para compreender a classificação e identificação da LMA. O Capítulo 3 detalha o processo conduzido para realizar esta Revisão Sistemática. No Capítulo 4, é apresentada a análise dos resultados, que compõe uma discussão sobre os resultados obtidos. Por fim, o Capítulo 5 aborda os desafios enfrentados pelos autores, escopo para trabalhos futuros e a conclusão do estudo realizado.

2 Referencial Teórico

Para uma compreensão mais aprofundada dos termos abordados e da metodologia proposta, este capítulo oferece uma breve revisão literária relacionada à identificação e classificação da LM utilizando técnicas de aprendizado profundo.

2.1 Leucemia

A Leucemia (LM) é uma doença maligna dos glóbulos brancos, caracterizada pelo acúmulo de células doentes na medula óssea, substituindo as células sanguíneas saudáveis. Existem mais de 12 tipos de leucemia, sendo os quatro principais: leucemia mieloide aguda (LMA), leucemia mieloide crônica (LMC), leucemia linfocítica aguda (LLA) e leucemia linfocítica crônica (LLC). Cada tipo possui características distintas e requer abordagens específicas de diagnóstico e tratamento (JUNIOR et al., 2022). A Figura 1 resume a divisão dos tipos de Leucemia mais comuns.

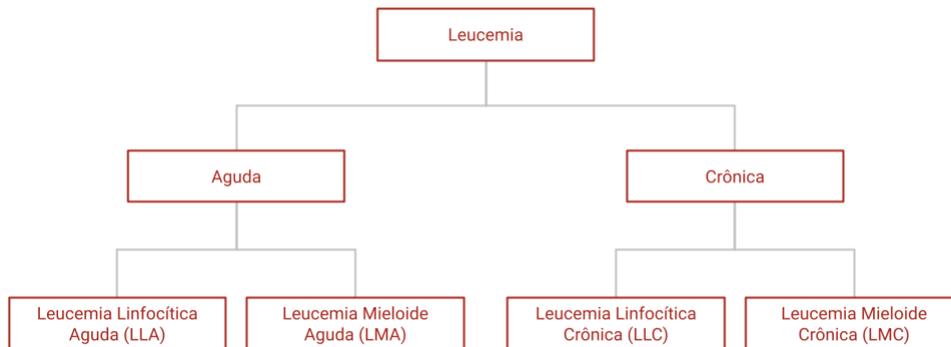


Figura 1 – Tipos de Leucemia

Fonte: elaborado pelo autor (2024)

Os leucócitos, também conhecidos como glóbulos brancos, desempenham um papel fundamental no sistema imunológico, protegendo o corpo contra vírus, infecções, bactérias e outros invasores. Originados na medula óssea, essas células nucleadas circulam pelo sangue e estão presentes no sistema linfático. Diferenciando-se de outras células sanguíneas, como glóbulos vermelhos e plaquetas, os leucócitos são classificados principalmente em duas categorias com base na hereditariedade celular, conhecidas como linfóides e mielóides, e na estrutura celular, classificadas como agranulócitos e granulócitos. Essas categorias são então subdivididas em cinco subcategorias: neutrófilos, basófilos, eosinófilos, linfócitos e monócitos. Cada subcategoria desempenha papéis específicos na resposta imunológica do corpo (SHAH et al., 2021).

As classes ou tipos de LM dependem do tipo de leucócitos afetados. Se os leucócitos imaturos forem linfócitos, a leucemia é classificada como leucemia linfocítica, e se os leucócitos imaturos forem monócitos e granulócitos, então é considerada leucemia mielóide. Em termos gerais, a leucemia pode ser categorizada como aguda, quando os leucócitos afetados são incapazes de funcionar como leucócitos normais, ou crônica, quando os leucócitos afetados podem funcionar como células normais, tornando a leucemia crônica mais prolongada. (SHAH et al., 2021)

2.2 A Leucemia Mieloide Aguda (LMA)

A LMA é uma condição maligna caracterizada pelo desenvolvimento clonal anormal de células-tronco hematopoiéticas. Essa doença envolve uma variedade de anormalidades cromossômicas, mutações genéticas frequentes, alterações epigenéticas que impactam a estrutura da cromatina e desregulações nos microRNAs. A presença de uma ampla heterogeneidade genômica, as diferenças individuais entre os pacientes e as mutações genéticas representam desafios significativos que influenciam a eficácia do tratamento para aqueles diagnosticados com LMA (YU et al., 2020).

2.2.1 Classificação da LMA

Na grande maioria dos cânceres a avaliação do estágio (extensão) é um elemento crucial. Esse estágio é determinado pelo tamanho do tumor principal e pela disseminação do câncer. Por outro lado, a LMA não segue o padrão comum de formação de tumores. Em vez disso, ela geralmente se dissemina por toda a medula óssea e ocasionalmente atinge outros órgãos, como fígado e baço. Como resultado, a LMA não é classificada da mesma maneira que a maioria dos outros tipos de câncer. A avaliação da perspectiva de uma pessoa com LMA baseia-se em outros fatores, como o subtipo de LMA (determinado por exames laboratoriais), a idade do paciente e outros resultados de exames laboratoriais. Para isso, os principais sistemas utilizados para categorizar os subtipos da LMA são a classificação Franco-Americana-Britânica (FAB) e a classificação mais recente da Organização Mundial da Saúde (OMS) SOCIETY (2018).

2.2.2 Sistema de classificação FAB

Em 1976, o Dr. John M. Bennett e seus colegas do grupo cooperativo Franco-Americano-Britânico (FAB) lançaram um artigo seminal, apresentando a primeira classificação da LMA. Existem dois tipos de blastócitos de acordo com a classificação FAB: Tipo I e Tipo II. Os blastos do Tipo I apresentam menos grânulos citoplasmáticos, enquanto os do Tipo II têm um número reduzido de grânulos principais. Ambos os tipos possuem nucléolos

proeminentes (normalmente de 3 a 5) e uma cromatina delicada e não condensada em seus núcleos centrais. (MITTELMAN, 2023)

Como mostrado na Tabela 1 o grupo dividiu a leucemia mieloide aguda em subtipos, M0 a M7, com base no tipo de célula em que a leucemia se desenvolve e o grau de maturidade das células.

Tabela 1 – Classificação da LMA segundo a FAB

Subtipo	Nome
M0	Leucemia mieloblástica aguda indiferenciada
M1	Leucemia mieloblástica aguda com maturação mínima
M2	Leucemia mieloblástica aguda com maturação
M3	Leucemia promielocítica aguda
M4	Leucemia mielomonocítica aguda
M4 eos	Leucemia mielomonocítica aguda com eosinofilia
M5	Leucemia monocítica aguda
M6	Leucemia eritroide aguda
M7	Leucemia megacarioblástica aguda

Os subtipos M0 a M5 da LMA se origina em formas imaturas dos glóbulos brancos. O subtipo M6 tem início em formas extremamente imaturas dos glóbulos vermelhos, enquanto o M7 se origina em formas imaturas das células responsáveis pela produção de plaquetas SOCIETY (2018).

2.2.3 Sistema de classificação OMS

Embora o sistema de classificação FAB possa ser útil, ele não incorpora diversos fatores agora reconhecidos como influentes no prognóstico. O sistema da Organização Mundial da Saúde (OMS), atualizado em 2016, aborda alguns desses fatores, visando uma classificação mais precisa da Leucemia Mieloide Aguda (LMA). (SOCIETY, 2018)

A OMS revisa sua classificação a cada oito anos, levando em consideração anomalias cromossômicas e mutações genéticas. Em contraste, o grupo FAB não considera esses fatores, baseando-se exclusivamente na morfologia e citoquímica para identificar os mieloblastos, categorizando-os conforme a conformação celular. O mieloblasto tipo I não apresenta maturação e granulação; o tipo II tem pouca diferenciação no citoplasma com poucos grânulos (inferior a 20%), enquanto o tipo III possui citoplasma granular (superior a 20%) com núcleo central e evidência da zona de Golgi. A classificação categoriza a LMA em grupos de acordo com descobertas mais recentes sobre suas características citogenéticas e clínicas (ZERBINI et al., 2011). Divisão de grupos da LMA, segundo a OMS (SOCIETY, 2018):

- LMA com certas anomalias genéticas (alterações genéticas ou cromossômicas)
 - LMA com translocação entre os cromossomos 8 e 21 [t(8;21)]

- LMA com translocação ou inversão no cromossomo 16 [t(16;16) ou inv(16)]
 - APL com o gene de fusão PML-RARA
 - LMA com translocação entre os cromossomos 9 e 11 [t(9;11)]
 - LMA com translocação entre os cromossomos 6 e 9 [t(6;9)]
 - LMA com translocação ou inversão no cromossomo 3 [t(3;3) ou inv(3)]
 - LMA (megacarioblástica) com translocação entre os cromossomos 1 e 22 [t(1;22)]
 - LMA com o gene de fusão BCR-ABL1 (BCR-ABL)
 - LMA com gene NPM1 mutado
 - LMA com mutações bialélicas do gene CEBPA (ou seja, mutações em ambas as cópias do gene)
 - LMA com gene RUNX1 mutado
- LMA com alterações relacionadas à mielodisplasia
 - LMA relacionada a quimioterapia ou radiação anterior
 - LMA sem outra especificação (isso inclui casos de LMA que não se enquadram em nenhum dos grupos acima e é semelhante à classificação FAB.)
 - LMA com diferenciação mínima (FAB M0)
 - LMA sem maturação (FAB M1)
 - LMA com maturação (FAB M2)
 - Leucemia mielomonocítica aguda (FAB M4)
 - Leucemia monoblástica/monocítica aguda (FAB M5)
 - Leucemia eritróide pura (FAB M6)
 - Leucemia megacarioblástica aguda (FAB M7)
 - Leucemia basofílica aguda
 - Panmielose aguda com fibrose
 - Sarcoma mieloide (também conhecido como sarcoma granulocítico ou cloroma)
 - Proliferações mieloides relacionadas à síndrome de Down
 - As leucemias agudas indiferenciadas e bifenotípicas não são estritamente LMA, mas são leucemias que apresentam características linfocíticas e mieloides. Às vezes são chamadas de leucemias agudas de fenótipo misto (MPALs).

2.2.4 Diagnóstico

Os patologistas e hematologistas costumam diagnosticar a Leucemia ao examinar manualmente o esfregaço de sangue periférico do paciente através de um microscópio. Este processo permite uma avaliação detalhada das células sanguíneas para identificar possíveis anomalias e confirmar a presença de leucemia. É utilizada uma variedade de células distintas e suas propriedades físicas para identificar e classificar a leucemia. As características específicas dessas células permitem a detecção, evidenciada por um aumento anormal no número de glóbulos brancos (leucócitos) e uma diminuição no número de outras células

sanguíneas. Essa análise minuciosa ajuda a confirmar o diagnóstico e a compreender a natureza específica da leucemia presente. (DHARANI; HARIPRASATH, 2018)

A classificação FAB estabelece que o diagnóstico de LMA requer a presença de mais de 30% de blastos na medula óssea. Em contraste, segundo a classificação da OMS, o diagnóstico pode ser feito com mais de 20% de blastos na medula óssea. No entanto, técnicas adicionais de diagnóstico são essenciais para uma análise mais detalhada, ajudando a determinar o tipo celular presente na leucemogênese, orientar a escolha da terapêutica apropriada e prognosticar a doença. (SOCIETY, 2018)

2.3 Aprendizado de Máquina

O aprendizado de máquina tem experimentado um crescimento acelerado nos últimos anos no campo da análise de dados e computação. Oferece aos sistemas a capacidade de aprender e melhorar automaticamente com base na experiência. Os algoritmos são geralmente divididos em quatro categorias principais: Aprendizado Supervisionado, nesse tipo de aprendizado, o algoritmo aprende a função que mapeia uma entrada para uma saída com base em pares de entrada-saída rotulados em amostras; Aprendizado Não Supervisionado, este método envolve a análise de conjuntos de dados não rotulados, dispensando a necessidade de intervenção humana. É um processo baseado em dados; Aprendizado Semisupervisionado, é uma combinação de métodos supervisionados e não supervisionados, onde o modelo é treinado em um conjunto de dados parcialmente rotulado; Aprendizado por Reforço, os agentes de software e máquinas avaliam automaticamente o comportamento ideal em um contexto ou ambiente específico para melhorar sua eficiência ao longo do tempo (SARKER, 2021).

2.3.1 Aprendizado Profundo

O aprendizado profundo ou *Deep Learning (DL)* é uma subcategoria do aprendizado de máquina que emprega diversas redes neurais para conduzir processos de aprendizado supervisionado e não supervisionado. Essas redes neurais, modeladas com base na funcionalidade do cérebro, geralmente consistem em três camadas: a camada de entrada, a camada oculta e a camada de saída. Cada camada contém um conjunto de nós, com a camada de entrada recebendo as entradas, as camadas ocultas processando os dados e a camada de saída gerando resultados, que podem ser utilizados para classificação, regressão, entre outras tarefas (KUMAR et al., 2023). A Figura 2 ilustra esse processo.

O DL tem se destacado pela sua capacidade de processar e aprender com grandes volumes de dados. Dentro desse campo, as Deep Neural Network (DNN) desempenham um papel crucial, sendo estruturas complexas que permitem a aprendizagem de representações abstratas e hierárquicas. As DNNs consistem em múltiplas camadas de neurônios interconectados, permitindo a extração de características complexas dos dados. Um tipo

de DNN é a Convolutional Neural Network (CNN), particularmente eficaz em tarefas de visão computacional. O diferencial das CNNs reside na sua capacidade de eliminar a necessidade de extração manual de características (KRISHNA; KALLURI, 2019).

A Figura 2 mostra a diferença entre aprendizado profundo e aprendizado de máquina, que tem como vantagem sobre abordagens tradicionais a não utilização de algoritmos como SIFT ou LBP para extrair características de uma imagem, as CNNs podem aprender, durante o treinamento, representações hierárquicas diretamente dos dados brutos. Essa característica adaptativa permite que as CNNs identifiquem padrões específicos do conjunto de dados, resultando em modelos mais precisos em tarefas de visão computacional, como detecção de objetos, classificação e reconhecimento (KRISHNA; KALLURI, 2019).

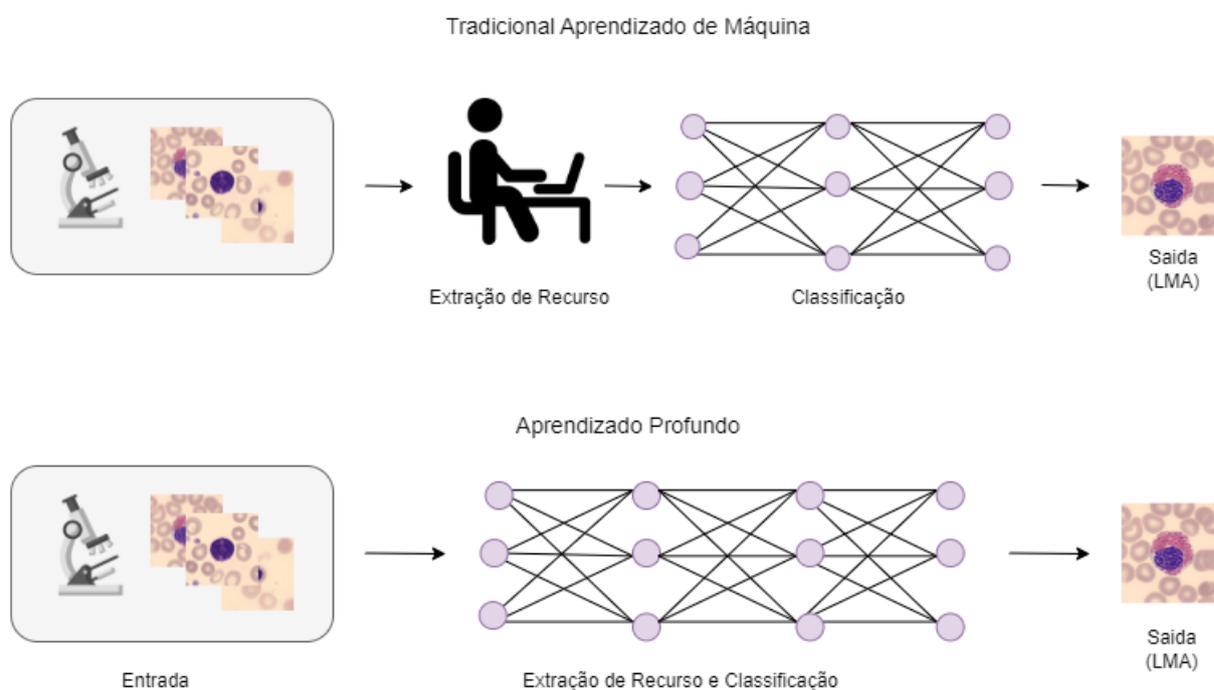


Figura 2 – Comparação entre Aprendizado de Máquina e Aprendizado Profundo

Fonte: elaborado pelo autor (2024)

2.3.2 Transferência de Aprendizagem

A aprendizagem por transferência visa utilizar o conhecimento adquirido em um domínio de origem para aprimorar a aprendizagem em um domínio de destino, estabelecendo um paradigma crucial no campo de aprendizado de máquina. Esse processo difere do aprendizado de máquina tradicional, onde, para uma nova tarefa, o aprendizado tradicional parte do zero para cada tarefa. Em contraste, a aprendizagem por transferência utiliza o conhecimento previamente adquirido em uma tarefa anterior, permitindo que

uma máquina explore esse conhecimento para melhorar a generalização em uma nova tarefa (LIANG; FU; YI, 2019).

3 Metodologia

Revisão Sistemática (SR) é definida por (GALVÃO; RICARTE, 2019) como uma modalidade de pesquisa, que segue protocolos específicos, e que busca entender e dar alguma logicidade a um grande corpus documental, especialmente, verificando o que funciona e o que não funciona num dado contexto. Com base nessa declaração, o presente trabalho utilizou a ferramenta *StArt* (*State of the Art through Systematic Review*)¹ para conduzir uma Revisão Sistemática da literatura.



Figura 3 – Fluxograma da Metodologia
Fonte: elaborado pelo autor (2024)

Conforme ilustrado na Figura 3, o processo é composto por cinco etapas: Planejamento, Execução, Seleção, Extração e Sumarização. No Planejamento, foi definido o protocolo que envolve a elaboração dos objetivos, questões de pesquisas, definição da palavra-chave, escolha das bases bibliográficas e critérios de inclusão e exclusão. Em seguida, na etapa de Execução, ocorreu a busca nas bases bibliográficas escolhidas, possibilitando a Seleção por meio da leitura das partes principais dos trabalhos, atendendo aos critérios definidos. A Extração envolveu o refinamento por meio da leitura completa dos trabalhos selecionados, respondendo às questões de pesquisa. Por fim, na etapa de Sumarização, foram criadas tabelas com o resumo da revisão e a discussão dos resultados. As subseções seguintes detalharão a execução de cada etapa desse processo.

3.1 Planejamento

Essa etapa consiste na criação do protocolo que envolve a definição dos objetivos, delimitação das questões de pesquisa, escolha das bases de busca bibliográfica e a escolha dos critérios de inclusão e exclusão. As próximas subseções fornecerão detalhes sobre a execução de cada aspecto dessa etapa

¹ <https://www.lapes.ufscar.br/resources/tools-1/start-1>

3.1.1 Questões de pesquisa

Considerando o objetivo inicial da pesquisa, foram formuladas questões secundárias para abordar aspectos adicionais. Essas questões foram aplicadas a todos os materiais utilizados, proporcionando uma abordagem abrangente para complementar o problema inicial:

- Q1: Quais bases de dados foram utilizadas?
- Q2 :Quais métodos de pré-processamento foram aplicados nas bases de dados?
- Q3: Quais tipos de testes foram conduzidos?
- Q4: Quais métricas e metodologias de comparação foram empregadas?
- Q5: Houve comparação entre as abordagens?
- Q6: Qual abordagem demonstrou os resultados mais favoráveis?

Tendo em vista essas questões, o objetivo é identificar as abordagens, buscando os melhores resultados, compreendendo seus "ambientes de atuação" e examinando seus pontos positivos e negativos. Com base nisso, este trabalho proporciona uma análise que pode servir como referência para futuras pesquisas que busquem explorar novas abordagens para a identificação e classificação da Leucemia (LMA).

3.1.2 Termos de Pesquisa

Uma parte essencial de uma revisão sistemática é a condução de uma pesquisa abrangente da literatura para identificar todos os estudos relevantes sobre o tema. Portanto, com intuito de obter resultados mais precisos na busca de trabalhos relevantes, foram definidas palavras-chaves que foram incorporados à estratégia de filtragem. Esses termos foram verificados em partes relevantes dos artigos, como título e resumo. As palavras-chave incluíram: *Leukemia*, *Myeloid*, *Acute*, *Classification*, *Detection* e *Identification*.

3.1.3 Bases de busca bibliográficas

Na seleção das bases bibliográficas, diversos critérios foram considerados, incluindo o tamanho do acervo de obras, a relevância para a área de ciências da computação e a capacidade de exportar referências dos trabalhos selecionados no formato *bibtex*. Levando em conta esses requisitos, as bases *IEEE Xplore*², *Web of Science*³, *Science Direct*⁴ e *Scopus*⁵ foram escolhidas. As plataformas possui um grande acervo de obras voltados para

² <https://ieeexplore.ieee.org/search/advanced>

³ <https://www.webofknowledge.com>

⁴ <https://www.sciencedirect.com>

⁵ <https://www.scopus.com/home.uri>

área da tecnologia, é possível acessar sua maioria pelo Portal de Periódicos da CAPES (PP-CAPES).

3.1.4 Critérios de Inclusão e Exclusão

Certos parâmetros foram adotados para determinar a inclusão ou exclusão dos artigos retornados na etapa de Seleção 3.3, de acordo com os critérios estipulados na proposta do projeto. Considerando esses critérios, os trabalhos são primeiramente avaliados com base no título e no resumo para determinar se atendem aos parâmetros estabelecidos. Em alguns casos, foi necessário realizar a leitura de trechos específicos para verificar se não se enquadravam nos termos de exclusão. Esses critérios estão detalhados na Tabela 2.

Tabela 2 – Critérios de Inclusão e Exclusão

Critérios de Inclusão:
I.1 Idioma de escrita deve ser em inglês
I.2 Quando as palavras-chave predefinidas estão presentes no conjunto completo ou, pelo menos, no título, ou na seção do resumo do documento.
I.3 Disponibilizado gratuitamente ou pelo Periódicos CAPES
I.4 Voltado para utilização de Deep Learning para a classificação e detecção da LMA
I.5 O artigo deve ser um estudo primário e apresentar resultados
Critérios de Exclusão
E.1 Não ser referente ao tema de pesquisa
E.2 O artigo não estar disponível de forma completa
E.3 Artigos duplicados nos documentos de pesquisa
E.4 Artigos publicados antes de 2018

A análise dos critérios mencionados na Tabela 2, envolveu duas fases de refinamento. Na primeira fase, foram considerados os critérios de inclusão (I) I.1 e I.2. Durante as pesquisas, os artigos foram inicialmente limitados ao critério I.1, visto que todos os trabalhos obtidos estavam em inglês. Na segunda fase, os critérios I.4 e I.5 foram analisados por meio de uma leitura mais detalhada. Vale ressaltar que apenas foram considerados artigos gratuitos ou disponibilizados pelo Portal de Periódicos CAPES ⁶. Para garantir a obtenção de resultados mais recentes, os artigos foram restritos a um período específico, não sendo aceitos aqueles publicados antes de 2018.

3.2 Execução

Nesta etapa, foram realizadas as buscas por artigos nas bases bibliográficas escolhidas. Os métodos de busca foram refinados por meio da criação de strings que incorporam os termos mencionados na subseção 3.1.2. Para otimizar os resultados, o termo "*deep learning*" foi incluído nas *strings* utilizadas nos Repositório Virtuais *Web of Science, Science*

⁶ <https://www.periodicos.capes.gov.br/>

Direct e *Scopus*. Além disso, as buscas foram realizadas restringidos os trabalhos publicados antes de 2018, e o idioma foi limitado ao inglês. Dessa forma, cada base bibliográfica de busca possui *strings* distintas, conforme detalhado na tabela 3.

Tabela 3 – *Strings* de Buscas

Repositório Virtual	String de Busca
<i>IEEE Xplore</i>	("Document Title":Acute AND "Document Title":Myeloid AND "Document Title":Leukemia) AND("Abstract":identification OR "Abstract":classification OR "Abstract":detection)
<i>Web of Science</i>	((((TI=(Acute AND Myeloid AND Leukemia)) AND AB=(classification OR identification OR detection))) AND TS=(deep learning))
<i>Science Direct</i>	(Acute AND Myeloid AND Leukemia) AND (classification OR identification OR detection) AND deep learning
<i>Scopus</i>	(TITLE (acute AND myeloid AND leukemia) AND ABS (identification OR classification OR detection) AND TITLE-ABS-KEY (deep AND learning)) AND PUBYEAR >2017

As buscas utilizando as *strings* mencionadas na Tabela resultou em um total de 51 trabalhos. Destes, 19 foram provenientes da *Science Direct*, a base *Scopus* retornou 14, *IEEE Xplore* retornou 09 e a *Web of Science* apresentou 09 trabalhos. A Figura 4 demonstra o resultado do processo de busca.

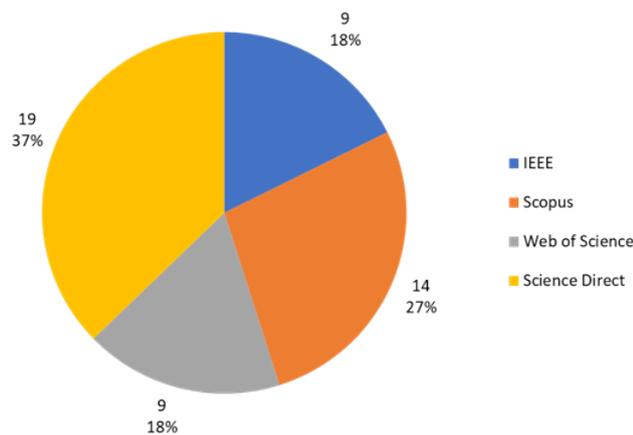


Figura 4 – Resultado da Busca

Fonte: elaborado pelo autor (2024)

3.3 Seleção

Nesta etapa, houve um refinamento dos artigos obtidos na busca inicial, ou seja, foi feita uma seleção de acordo com os critérios de inclusão e exclusão mencionados na Subseção 3.1.4. Para um trabalho ser considerado aceito e prosseguir para a próxima etapa, ele

não pode atender a nenhum critério de exclusão e precisa atender a todos os critérios de inclusão. Além disso, durante esse processo, foram identificados e removidos os trabalhos duplicados. Após a classificação, todos os artigos aceitos foram revisados individualmente. A Figura 5 mostra o processo de seleção.

Os artigos foram classificados como duplicados, aceitos e rejeitados, a busca inicial obteve um total de 51 artigos. Em seguida, foram identificadas duplicatas reduzido o número de artigos para 43, diminuindo os artigos a serem verificados. Após essas etapas, foi realizada a filtragem, atendendo aos Critérios de Inclusão e Exclusão descritos na Subseção 3.1.4, resultando em um total de 15 trabalhos selecionados. Por fim, após analisar todos os critérios foi possível obter 12 artigos como resultado final. A Figura 5 resume todos os passos seguidos para adquirir os trabalhos utilizados nesse estudo.

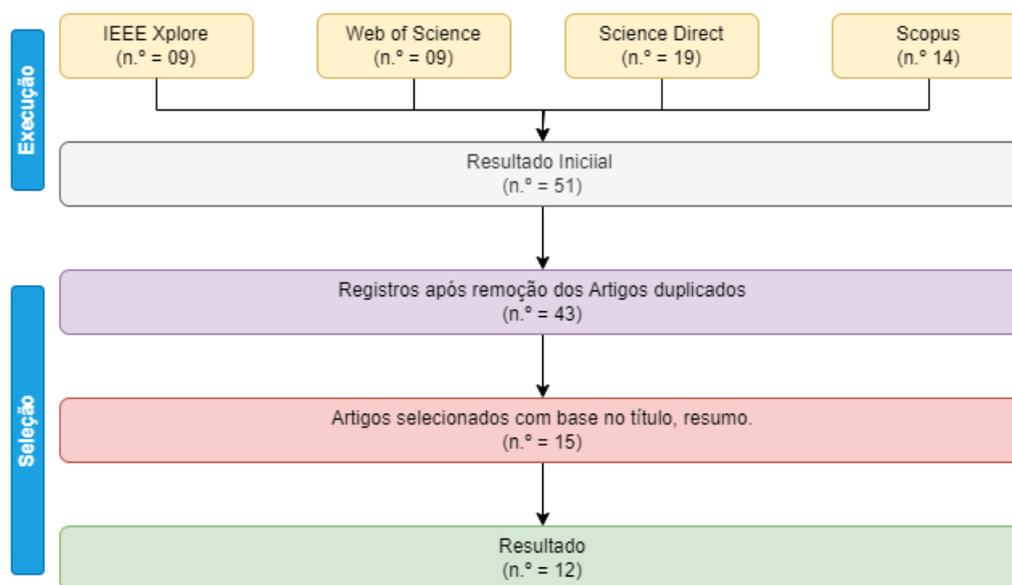


Figura 5 – Seleção dos Artigos

Fonte: elaborado pelo autor (2024)

3.4 Extração

Com o objetivo de extrair informações relevantes de cada trabalho, foi criado um formulário de extração, elaborado com base nas questões de pesquisas citados na subseção 3.1.1 O formulário está estruturado em quatro grupos: Aquisição dos Dados, Pré-processamento, Classificação e Validação.

No Grupo 1, as perguntas abordam a aquisição de dados, incluindo detalhes sobre o equipamento utilizado para a coleta, a existência de melhorias e se a coleta foi conduzida pelos próprios autores. O objetivo é fornecer informações sobre o processo de obtenção das imagens.

No Grupo 2, as perguntas são direcionadas ao pré-processamento das imagens coletadas, visando compreender como características relevantes para a identificação da LMA

foram extraídas. Além disso, buscar informações sobre a aplicação de técnicas de normalização e aumento de dados.

Os Grupos 3 e 4 contêm perguntas relacionadas aos classificadores e métricas utilizadas. No Grupo 3, o foco está nos classificadores, buscando compreender a abordagem utilizada. No Grupo 4, as perguntas se referem às métricas utilizadas, incluindo detalhes sobre a divisão entre conjuntos de treinamento e teste, e como essa abordagem foi validada. A Tabela 4 detalha cada as questões definidas em cada grupo

Tabela 4 – Extração de Dados

Grupo	Questões
Aquisição dos Dados	Qual conjunto de dados é usado? Se o conjunto de dados é público ou privado? Quantas amostras tem a base? Qual tipo de extensão das imagens (JPG, PNG..)? Qual resolução das imagens? Quais classes são usadas? Qual a distribuição por classe?
Pré-processamento	Houve alguma normalização aos dados? Se sim, qual? Houve alguma técnica de aumento de dados? Se sim, qual?
Classificação	Qual abordagem de aprendizado profundo? Como foi definido o conjunto de treinamento/teste? Validação cruzada, se sim, quantas pastas ou dobras? Qual é a divisão de treinamento/teste adotada? Houve otimização dos parâmetros de classificação? Se sim, qual técnica?
Validação	Quais métricas foram usadas? Quais os valores obtidos?

4 Análise dos resultados

Este capítulo apresenta a análise dos resultados, que envolve a destilação dos principais pontos e descobertas durante o estudo. Nele, estão descritas as informações coletadas dos 12 artigos selecionados que atenderam a todos os requisitos impostos nesta Revisão Sistemática, conforme apresentado na Subseção 3.1.4. Os resultados foram extraídos com base no descrito na Seção 3.4. Para um melhor entendimento, este capítulo foi dividido em quatro seções: Aquisição dos Dados, Pré-processamento, Classificação e Validação. Cada seção é detalhada por meio de um resumo das informações e técnicas comumente utilizadas nos trabalhos selecionados. Por fim, a Seção Resultados compila todas as informações detalhadas.

4.1 Aquisição dos Dados

Cinco dos artigos analisados, (MATEK et al., 2019b), (RASTOGI; KHANNA; SINGH, 2022), (ROY; AMEER, 2022), (ELHASSAN et al., 2022) e (RAHMAN; AHMAD, 2023) fizeram uso do mesmo conjunto de dados. Eles utilizaram o conjunto de imagens de morfologia AML de Munique (*AML-Cytomorphology_LMU*), que contém 18.365 imagens unicelulares rotuladas por especialistas. Essas imagens foram retiradas de esfregaços de sangue periférico de 100 pacientes diagnosticados com leucemia mieloide aguda no Hospital Universitário de Munique entre 2014 e 2017, além de 100 pacientes sem sinais de malignidade hematológica. A aquisição de imagens foi feita utilizando um microscópio/scanner digital M8 (Precipoint GmbH, Freising, Alemanha) com ampliação óptica de 100 vezes e imersão em óleo. (MATEK et al., 2019a).

Em (NAGIUB et al., 2020), o conjunto de dados utilizado foi o AML Image Database (AML-IDB), composto por imagens de esfregaços de sangue periférico e medula óssea coletadas entre 2017 e 2019 na Unidade de Hematologia de imagens do Departamento de Patologia Clínica dos Hospitais Universitários de Assiut, Egito. Este conjunto inclui 412 amostras de imagens digitais, divididas em duas classes: LMA e controle de sangue normal. São 206 imagens de pacientes com LMA e 206 imagens de controles saudáveis normais.

No estudo conduzido por (WANG et al., 2022), os dados foram obtidos de duas fontes diferentes: a *American Society of Hematology* (ASH) e o Hospital da *Shanxi Medical University* (SXMU) no período de julho de 2016 a dezembro de 2020. Do ASH, um total de 115 esfregaços de medula óssea foi analisado, dos quais 32 correspondiam SMD, 26 a AA e 57 a LMA. Do SXMU, foram utilizadas 432 imagens, (SMD: 214, AA: 115 e LMA: 103). As imagens do *ASH Image Bank* foram empregadas para o desenvolvimento e validação interna do modelo, enquanto os dados hospitalares foram utilizados para a va-

lidação externa. O *ASH Image Bank* é uma biblioteca de imagens acessível publicamente, oferecendo uma coleção abrangente relacionada a diversos tópicos hematológicos.

Com o objetivo de aumentar o número de amostras, além do repositório de imagens *AML-Cytomorphology_LMU*, os autores (RASTOGI; KHANNA; SINGH, 2022) utilizaram as imagens de outro conjunto de dados: PBC-Barcelona. Este conjunto de dados contém um total de 17.092 imagens de células normais individuais, adquiridas usando o analisador *CellaVision DM96* no Laboratório Central do Hospital Clínico de Barcelona. O conjunto de dados está organizado em oito grupos, incluindo neutrófilos, eosinófilos, basófilos, linfócitos, monócitos, granulócitos imaturos (promielócitos, mielócitos e metamielócitos), eritroblastos e plaquetas ou trombócitos. O tamanho das imagens é de 360x363 *pixels*, no formato JPG (ACEVEDO, 2020).

No estudo realizado pelos autores (CLARO et al., 2022), foram reunidos 18 conjuntos de dados públicos, possibilitando a obtenção de 3.536 imagens. Dentre o total de imagens, 1.434 pertenciam à classe HBS (*Healthy Blood Slides*) (40,55% do total), 881 pertenciam à classe Leucemia Linfoblástica Aguda (LLA) (24,92%), 978 pertenciam à classe AML (27,66%) e 243 pertenciam à classe "outras modalidades" (6,87%).

O trabalho realizado pelos autores (ABHISHEK et al., 2022) utilizou uma base de dados contendo 500 imagens que foram obtidas em parceria com o All India Institute of Medical Sciences (AIIMS) onde coletaram dados que incluíram imagens referentes as doenças ALL, LMA e casos normais. O conjunto contém 125 amostras para LLA, 125 LMA e 250 são normais

No trabalho realizado por (YU et al., 2023), foi criado um banco de dados próprio contendo um total de 8.245 imagens de esfregaço de medula óssea de 651 pacientes com diagnóstico de diferentes subtipos da LMA. Todos os pacientes foram diagnosticados com base nas categorias morfológicas padrão do sistema de classificação FAB, o conjunto de dados foi derivado do Primeiro Hospital Afiliado da Universidade de Zhejiang de 2010 a 2020, e o conjunto de dados de teste foi derivado do Primeiro Hospital Afiliado da Universidade de Zhejiang e do Primeiro Hospital Popular Afiliado de Hangzhou entre 2020 e 2021.

A base de dados utilizada pelos autores (BAIRABOINA; BATTULA, 2023) foi o conjunto de dados *Blood Cell Counting Dataset* (BCCD), disponível publicamente no Kaggle. Ele contém designações de tipos de células em um arquivo CSV e 12.453 imagens no formato JPEG de leucócitos, neutrófilos, monócitos e eosinófilos, sendo esses quatro tipos de células incluídos nesta coleção.

A Tabela 5 apresenta uma lista detalhada dos artigos, suas respectivas bases de dados e o número original de imagens associadas a cada um. É importante observar que os trabalhos (RASTOGI; KHANNA; SINGH, 2022) e (WANG et al., 2022) mencionados utilizaram mais de uma base de imagens para compor seus conjuntos de teste, evidenciando uma abordagem abrangente na coleta de dados. Essa prática sugere uma busca

por diversidade e representatividade nos conjuntos de teste, o que pode ser crucial para a validade e generalização dos resultados obtidos.

Tabela 5 – Bases de Dados

Artigo	Base de Dados	Nº Original de Imagens
(MATEK et al., 2019b)	AML-Cytomorphology_LMU	18.365
(NAGIUB et al., 2020)	AML-IDB	412
(SARKER, 2021)	PBC – Barcelona	17.092
(WANG et al., 2022)	ASH Image Bank	115
	SXMU	432
(RASTOGI; KHANNA; SINGH, 2022)	AML-Cytomorphology_LMU	18.365
	PBC – Barcelona	17.092
(ROY; AMEER, 2022)	AML-Cytomorphology_LMU	18.365
(CLARO et al., 2022)	Própria	3.536
(ABHISHEK et al., 2022)	AIIMS	500
(ELHASSAN et al., 2022)	AML-Cytomorphology_LMU	18.365
(YU et al., 2023)	Própria	8.245
(BAIRABOINA; BATTULA, 2023)	BCCD	12.453
(RAHMAN; AHMAD, 2023)	AML-Cytomorphology_LMU	18.365

4.2 Pré-processamento

Os artigos realizados por (MATEK et al., 2019b) e (NAGIUB et al., 2020) destacaram o uso de técnicas específicas de microscopia para a captura das amostras. Em (NAGIUB et al., 2020), foi empregado o corante de Leishman, proporcionando uma coloração azul-púrpura escura. Já em (MATEK et al., 2019b), a imersão em óleo com uma ampliação de 100 vezes foi a técnica utilizada. Essas abordagens não se enquadram diretamente nos tópicos de pré-processamento ou aumento de dados.

Em (MATEK et al., 2019b), foram geradas imagens adicionais aplicando transformações rotacionais aleatórias de 0 a 359 graus, além de inversões horizontais e verticais. As células fagocitárias e os promielócitos, com e sem bastonetes de Auer, foram agrupados em uma classe comum de promielócitos, resultando em 15 classes para treinamento e avaliação. Classes morfológicas com menos de 10 imagens foram combinadas com classes vizinhas na taxonomia em uma classe abrangente de nível superior. Mais especificamente, os mieloblastos, com e sem bastonetes de Auer, foram combinados em uma classe comum de mieloblastos, e as imagens unicelulares em nosso conjunto de dados. Com essas operações, o conjunto de dados foi expandido de forma que cada classe continha aproximadamente 10.000 imagens para treinamento.

O artigo conduzido por (WANG et al., 2022), fez o uso de oito métodos para transformações de imagens para obter diferentes versões das imagens originais, são eles: Rotação, a imagem foi girada em 35 graus em direção aleatória (esquerda ou direita); Branqueamento ZCA (Zero-phase Component Analysis), usado para reduzir informações redundantes de imagens e preservar informações importantes; Deslocamento de Largura, a imagem foi

deslocada aleatoriamente para a esquerda ou direita em 40%; Deslocamento de Altura, obtido pelo deslocamento aleatório da imagem para cima ou para baixo com 35%; Cisalhamento; realizado deslocando a imagem em 35 graus no sentido anti-horário; Zoom, a imagem foi ampliada em 35% para aproximar a aparência dos objetos na imagem; Inversão horizontal, obtida invertendo a imagem para cima e para baixo. Por fim, inversão vertical, obtida invertendo a imagem para a esquerda e para a direita. Cada imagem original foi aumentada para três amostras usando cada método de aprimoramento de dados. Após esse processo, o número de amostras no conjunto de dados do *ASH Image Bank* foi expandido em 24 vezes, indo de 115 para 2.760.

No artigo realizado pelos autores ([RASTOGI; KHANNA; SINGH, 2022](#)), devido ao desequilíbrio do conjunto de dados, foi empregada a técnica de amostragem aleatória (*up-sampling*) para balanceamento. O conjunto de dados remontado tinha 10.000 amostras em cada classe. As imagens foram redimensionadas para $(200 \times 200 \times 3)$ pixels e normalizadas para uma escala de 0 a 1. Por fim, as imagens do conjunto de treinamento foram aumentadas usando técnicas como inversão aleatória, rotações aleatórias e translações.

Para lidar com desequilíbrio entre as classes após a compilação das 18 bases de imagens, os autores ([CLARO et al., 2022](#)) optaram por empregar técnicas comuns de aumento de dados baseadas em transformações geométricas, tais como rotação, translação, inversão, escala e cisalhamento.

O pré-processamento realizado por ([BAIRABOINA; BATTULA, 2023](#)), envolveu o redirecionamento de cada uma das imagens para 256×256 , em seguida foi realizado o aprimoramento de contraste com base em Equalização adaptativa limitada de histograma (CLAHE) para aumentar o contraste e criar corpos celulares mais visível. A técnica de aumento de dados Redes Generativas Adversariais Convolucionais (DCGAN) foi aplicada para gerar dados adicionais e, assim, aumentar o tamanho do conjunto de dados.

Em ([RAHMAN; AHMAD, 2023](#)), durante a etapa de pré-processamento, as imagens são inicialmente divididas em três conjuntos: treinamento, validação e teste, cada um com seus rótulos correspondentes. A divisão fixa do conjunto de dados é utilizada para avaliar o desempenho dos modelos. as imagens originais são redimensionadas de $400 \times 400 \times 1$ pixels para $224 \times 224 \times 1$ pixels. Em seguida, cada imagem é convertida em um tensor. Finalmente, são aplicadas inversões horizontais e verticais para aumentar apenas as imagens no conjunto de treinamento.

A Tabela 6 ilustra as técnicas de pré-processamento e aumento de dados aplicadas em cada artigo. O pré-processamento muitas vezes envolve a normalização ou branqueamento, enquanto o aumento de dados inclui técnicas comuns que são baseadas em transformação geométricas, operações como rotação, inversão, deslocamento, zoom, entre outras, para enriquecer a diversidade do conjunto de dados utilizado nos estudos.

Tabela 6 – Técnicas de Pré-processamento e Aumento de Dados

Artigo	Pré-processamento	Aumento de Dados
(MATEK et al., 2019b)	Redirecionamento, segmentação da região	Rotação aleatórias, inversão horizontal e inversão vertical
(NAGIUB et al., 2020)	Segmentação da região	-
(SARKER, 2021)	Redimensionamento	-
(WANG et al., 2022)	Branqueamento ZCA	Rotação, deslocamento de largura, deslocamento de altura, cisalhamento, zoom, inversão horizontal e inversão vertical
(RASTOGI; KHANNA; SINGH, 2022)	Normalização, redirecionamento	Inversão aleatória, rotações aleatórias e translações
(ROY; AMEER, 2022)	Redirecionamento	Rotação e translação
(CLARO et al., 2022)	-	Rotação, translação, inversão, escala e cisalhamento
(ABHISHEK et al., 2022)	-	-
(ELHASSAN et al., 2022)	Segmentação da região	-
(YU et al., 2023)	Mapa de Gradiente	-
(BAIRABOINA; BATTULA, 2023)	Equalização adaptativa limitada de histograma (CLAHE)	DCGAN
(RAHMAN; AHMAD, 2023)	Redirecionamento	Inversões horizontais e verticais

4.3 Classificação e Validação

No trabalho feito por (MATEK et al., 2019b), o treinamento foi realizado utilizando a ResNeXt CNN, para a tarefa de classificação, as imagens em cada classe do conjunto de dados foram divididas aleatoriamente em conjuntos de teste e treinamento. O conjunto de treinamento continha aproximadamente 80% das imagens, enquanto o conjunto de teste continha os 20% restantes. Para realizar uma validação cruzada quádrupla, foi feita uma divisão estratificada e aleatória das imagens em cinco dobras. Cada dobra continha aproximadamente 20% das imagens de cada classe, e uma imagem específica estava presente em apenas uma dobra. Para avaliar quantitativamente a qualidade da previsão em classe, foram utilizadas as métricas precisão, especificidade e sensibilidade.

No artigo realizado por (NAGIUB et al., 2020), foi realizada uma comparação entre os modelos AlexNet, VGG16, GoogleNet, ResNet101 e Inception-v3. Para identificar imagens de esfregaços de sangue normais ou com LMA, as imagens foram divididas em dois grupos: 80% das imagens foram definidas como conjunto de treinamento, e 20% das imagens restantes foram definidas como conjunto de teste. Para validar o trabalho realizado. O modelo Inception-v3 teve a maior precisão, atingindo 99%, podendo detectar todos os

casos leucêmicos de LMA.

(SARKER, 2021) propôs um modelo de classificação baseado em AlexNet para o diagnóstico de da LMA e, em seguida, os resultados foram comparados com um modelo baseado em LeNet-5. As investigações foram baseadas em amostras de esfregaços de sangue em um banco de dados. AlexNet identificou com precisão 88,9% das imagens com 87,4% de precisão e 98,58% de exatidão, enquanto o LeNet-5 identificou corretamente 85,3% das imagens com 83,6% de precisão e 96,25% de exatidão.

Em (WANG et al., 2022), o desempenho do modelo ResNet50 foi avaliado para identificar pacientes com Síndromes Mielodisplásicas (SMD), Anemia Aplástica (AA) e Leucemia Mieloide Aguda (LMA). O diagnóstico de LMA foi baseado nos critérios da OMS, que exigem a identificação de anormalidades citogenéticas clonais e reprodutíveis, como $t(8;21)(q22;q22.1)$, $inv(16)(p13.1;q22)$, $t(16;16)(p13.1;q22)$ e $t(15;17)(q22;q12)$. Mesmo que a porcentagem de blastos na medula óssea seja inferior a 20%, o diagnóstico de LMA é estabelecido. Os dados do ASH Image Bank foram divididos aleatoriamente em conjunto de treinamento e conjunto de teste com uma proporção de 7:3. O modelo tinha duas camadas de saída: uma para determinar se o paciente tinha SMD (duas classificações) e outra para identificar se o paciente tinha AA, SMD ou LMA (três classificações). O modelo alcançou uma precisão de 92,90%.

No artigo conduzido pelos autores (RASTOGI; KHANNA; SINGH, 2022), foi proposta uma nova metodologia em duas etapas para a classificação robusta para o diagnóstico de Leucemia Mieloide (LM). Isso foi alcançado construindo um modelo extrator de recursos ajustado e adaptado à arquitetura VGG16, denominado "LeuFeatx". A divisão do conjunto de dados foi realizada em conjuntos de treinamento e teste, utilizando a divisão estratificada com base no número de classes no conjunto de dados. Os autores conduziram as classificações utilizando vários classificadores baseados em aprendizado de máquina, como SVM, floresta aleatória e Extreme Gradient Boosting (XGB).

O estudo conduzido por (CLARO et al., 2022) busca demonstrar que o aumento de dados pode contribuir para a melhoria do desempenho de algumas redes pré-treinadas, incluindo abordagens como DenseNet121, DenseNet121-DA, Xception-DA, ResNet50, entre outras. A abordagem que alcançou a melhor precisão foi VGG16-DA, atingindo 99,6% na tarefa de diferenciar a LMA de células sanguíneas normais.

No contexto do trabalho proposto por (ROY; AMEER, 2022), a abordagem proposta envolve a combinação das previsões de redes pré-treinadas, como ResNet-50, VGG-19 e GoogLeNet. As saídas dessas três redes são somadas e fornecidas à camada de saída de classificação por meio de uma camada softmax. Essa estratégia de aprendizado conjunto busca explorar as vantagens de diferentes modelos para obter um desempenho geral aprimorado.

Em (ABHISHEK et al., 2022) foram utilizados SVM, LR, K-NN como classificadores e CNNs pré-treinadas DenseNet121, InceptionV3, MobileNet, ResNet50, VGG16 e VGG19.

Para o estudo são utilizadas 300 imagens para treinamento e 125 imagens são usadas para testar os modelos treinados. Durante os teste a rede , ResNer50 mais classificado SVM obterão uma melhor precisão de 95%

Em (ELHASSAN et al., 2022) O método de extração de recursos proposto é uma combinação híbrida de localização do momento CMYK e fusão de recursos CNN. Neste estudo, o método de localização do momento CMYK foi empregado para extrair a Região de Interesse (ROI) utilizando uma combinação do método de transformação de cores (CMYK) e momentos da imagem. A utilização do ROI visa reduzir a quantidade de informações irrelevantes para a extração de recursos de *White Blood Cell* (WBC). O método alcanço precisões de classificação de 97,57%

O estudo realizado por (YU et al., 2023) envolveu o desenvolvimento de um pipeline denominado AMLnet, que teve como objetivo identificar os subtipos da Leucemia Mieloide Aguda (LMA) entre os pacientes. A proposta alcançou uma AUC de 0,885 no nível da imagem e 0,921 no nível do paciente na distinção de nove subtipos de LMA no conjunto de dados de teste. O sistema demonstrou um desempenho superior em certos subtipos, tais como M2b, M3, M4Eo, M6 e M7.

O artigo realizado pelos autores (BAIRABOINA; BATTULA, 2023), propôs uma extração e classificação eficaz de recursos com base em GhostNet e ResNeXt otimizada. Essa abordagem foi conduzida com o auxílio da segmentação baseada em W-net para classificar leucócitos maduros e imaturos. O modelo alcançou taxas de precisão de 99,24% para os conjuntos de dados BCCD.

O trabalho proposto pelos autores (RAHMAN; AHMAD, 2023), realizou os testes utilizando AlexNet, ResNet50, DenseNet161 e VGG-16 modificados foi usada para detectar leucócitos imaturos. Após o treinamento e validação do modelo utilizando diferentes parâmetros, os modelos com os melhores parâmetros foram aplicados ao conjunto de teste. Entre outros modelos, o AlexNet modificado precisão, 94,94%

A Tabela 7 resume as informações sobre os artigos, suas áreas de aplicação, arquiteturas de modelos utilizadas e a precisão alcançada nos resultados de classificação.

Tabela 7 – Abordagens Utilizadas

Artigo	Classifica	Abordagem	Precisão(%)
(MATEK et al., 2019b)	LMA	ResNeXt	90%
(NAGIUB et al., 2020)	LMA	Inception-v3	99%
(SARKER, 2021)	LMA	AlexNet	88,90%
(WANG et al., 2022)	AA, LMA e SMD	ResNet50	92,90%
(RASTOGI; KHANNA; SINGH, 2022)	LMA	Própria	96,15%
(ROY; AMEER, 2022)	LMA	Própria	91%
(ABHISHEK et al., 2022)	LMA, LLA, Normal	ResNer50	95%
(ELHASSAN et al., 2022)	LMA	Própria	97,57%
(CLARO et al., 2022)	LMA, LLA, Normal	VGG16-DA	99,60%
(YU et al., 2023)	LMA	Própria	90%
(BAIRABOINA; BATTULA, 2023)	LMA e LLA	GhostNet e ResNeXt	99,24%
(RAHMAN; AHMAD, 2023)	LMA	AlexNet	94,94%

4.4 Resultados

A escolha da base de dados desempenha um papel crucial na validade e generalização dos resultados de estudos relacionados à morfologia citológica em LMA. A diversidade e o tamanho dos conjuntos de dados refletem a complexidade da área de estudo, fornecendo uma base robusta para avanços significativos na compreensão e diagnóstico dessas condições clínicas.

Destacam-se algumas bases de dados recorrentes, como AML-Cytomorphology_LMU, AML-IDB, e PBC – Barcelona, que foram utilizadas em múltiplos estudos, consolidando-se como fontes significativas para pesquisas nessa área. Além disso, a criação de bases de dados próprias, como nos pelos estudos de (CLARO et al., 2022) e (YU et al., 2023), destaca a importância de adaptar os conjuntos de dados às necessidades específicas de cada pesquisa.

O tamanho variado dos conjuntos de dados, também é uma consideração relevante. Conjuntos de dados maiores, podem aumentar a capacidade o desempalho da abordagem, o conjunto com mais amostras é AML-Cytomorphology_LMU. A Figura 6 mostra um comparativo entre a quantidade de amostras de cada base de dados.

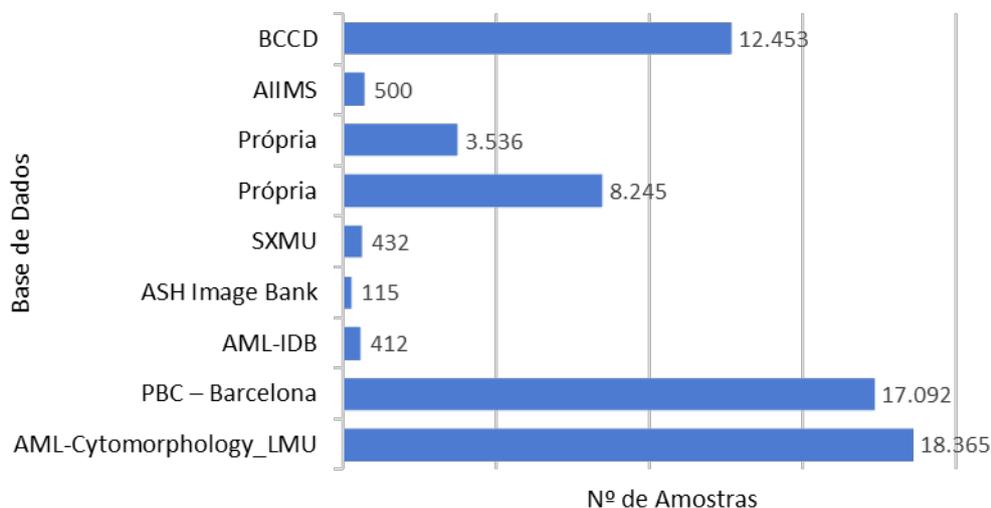


Figura 6 – Base de Imagens

Fonte: elaborado pelo autor (2024)

Dado que as bases de dados utilizadas forneciam poucas amostras e classes desequilibradas, os trabalhos analisados optaram por usar o aumento de dados. O intuito foi empregar uma estratégia para expandir e diversificar conjuntos de dados, com o objetivo de melhorar o desempenho e a generalização dos modelos utilizados. A maioria dos trabalhos utilizou transformações geométricas nas imagens, como rotação aleatória, inversão horizontal e vertical, deslocamento de largura e altura. O trabalho realizado por (BAI-

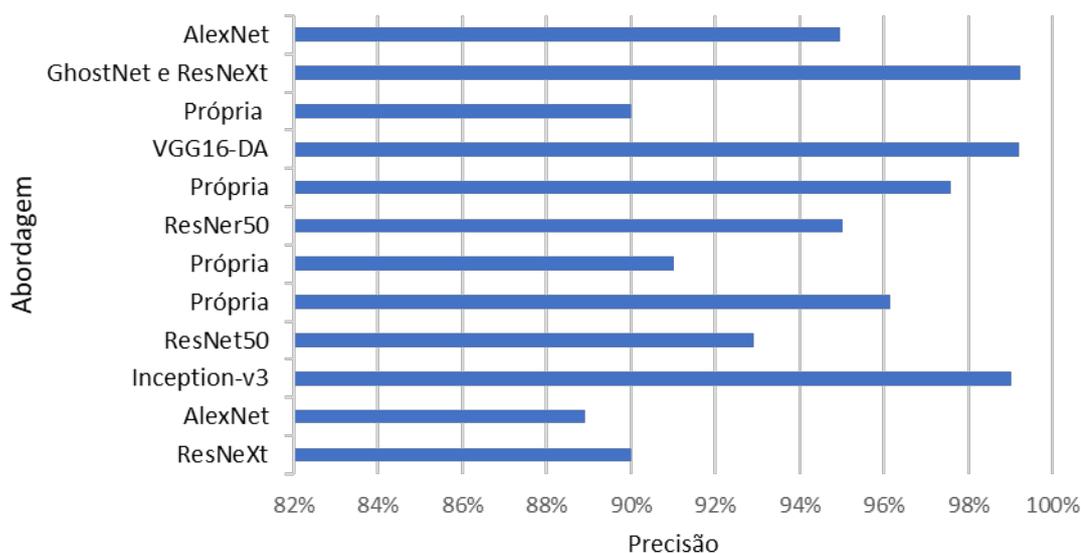


Figura 7 – Comparativo entre as Abordagens

Fonte: elaborado pelo autor (2024)

[RABOINA; BATTULA, 2023](#)) se destacou por utilizar Redes Generativas Adversariais para gerar novas amostras.

A análise dos artigos revela uma variedade de abordagens na aplicação de modelos de classificação médica em diferentes condições. As arquiteturas dos modelos variam de redes consolidadas, como ResNeXt, Inception-v3, ResNet50, VGG16-DA, GhostNet, a modelos desenvolvidos internamente pelos pesquisadores. A precisão obtida nos resultados de classificação oscila entre 88.90% e 99.60%. Foi possível notar a alta precisão alcançada por alguns modelos proprietários, como na abordagem de ([RASTOGI; KHANNA; SINGH, 2022](#)) com 96.15%, ([ROY; AMEER, 2022](#)) com 91%, ([ELHASSAN et al., 2022](#)) com 97.57%, e ([CLARO et al., 2022](#)) com a maior precisão alcançada de 99.60%. A Figura 7 ilustra o comparativo entre as abordagens utilizadas em cada um dos artigos.

5 Conclusão e Trabalhos Futuros

A detecção e classificação da LMA por meio do processamento de imagens de esfregaço de sangue é uma área promissora. No entanto, esta revisão identificou alguns problemas significativos. A falta de conjuntos de dados grandes e de alta qualidade disponíveis publicamente, a escassez de profissionais médicos especializados e a ausência de abordagens leves estão entre esses problemas que a comunidade científica precisa enfrentar ou superar. Neste capítulo, serão abordados os principais desafios encontrados durante a análise dos artigos selecionados.

5.1 Aquisição de Imagens

A falta de conjuntos de dados disponíveis ao público é uma questão séria. Para solucionar esses problemas, os pesquisadores devem incentivar os centros de saúde a disponibilizar publicamente os dados originais. Além disso, é crucial promover atividades que incentivem a disponibilidade de dados abertos por parte de diversas organizações de saúde em todo o mundo. Pois, quando um grande volume de dados está disponível, o problema se torna mais atraente para a pesquisa, especialmente para aplicações de abordagens de DL, onde a aquisição de conjuntos de dados grandes e de alta qualidade com rotulagem verdadeira é fundamental.

5.2 Capacidades de Generalização de Preditores

Apesar dos avanços significativos em técnicas de DL, foi possível identificar vários desafios ao treinar modelos e garantir que eles generalizem bem para novos dados. Alguns dos desafios comuns incluíram: *Overfitting*, isso pode prejudicar a capacidade do modelo de generalizar para novos dados, levando a previsões imprecisas. *Underfitting*, ocorre quando um modelo é muito simples para capturar a complexidade dos dados de treinamento. Desequilíbrio de Classes, durante os estudos alguns das bases de dados utilizadas apresentaram o desequilíbrio, onde uma das classes tem mais imagens do que a outra, Dados Ruins ou Insuficientes, um dos grandes desafios destacados pelos autores foi a melhoria das imagens utilizadas. Portanto, abordar esses desafios requer uma combinação de escolha cuidadosa de algoritmos, técnicas de pré-processamento de dados, ajuste fino de hiperparâmetros e considerações éticas durante o desenvolvimento e implementação de modelos de aprendizado de máquina.

5.3 Gasto Computacional

O desafio de tempo gasto é um aspecto identificado nesta análise, especialmente ao lidar com conjuntos de dados extensos. O processo consome considerável tempo e recursos de memória, especialmente quando a complexidade da técnica está vinculada ao número de camadas utilizadas nas Redes Neurais Convolucionais. Por exemplo, o uso da validação cruzada, que é crucial para avaliar a capacidade de generalização de um modelo, aumenta o gasto computacional, pois envolve treinar e avaliar o modelo várias vezes em diferentes divisões dos dados. Esta complexidade computacional destaca a necessidade de estratégias eficientes para enfrentar os desafios temporais e demanda por recursos na implementação de abordagens inovadoras.

5.4 Classificação dos subtipos da LMA

Além de identificar a presença da LMA, é essencial classificar os subtipos para administrar o tratamento adequado aos pacientes. Embora alguns trabalhos não tenham focado nessa classificação devido à sua complexidade, os estudos bem-sucedidos na identificação dos subtipos frequentemente utilizaram o sistema mais antigo FAB. Esse sistema baseia-se principalmente nas características da célula leucêmica. Embora a classificação FAB tenha sido uma ferramenta valiosa para os hematologistas durante muitos anos, abordagens mais recentes, como a classificação da OMS é mais abrangente e reflete melhor a heterogeneidade da LMA, considerando as características clínicas e moleculares.

5.5 Conclusão

Esta monografia compila a literatura de 12 artigos, com foco na detecção e classificação LMA em imagens de esfregaço de sangue usando DL. O trabalho resume a visão geral dos diversos métodos utilizados pelos autores para classificar os subtipos, destacando as etapas utilizadas. São abordados os vários desafios e obstáculos enfrentados pelos autores. Embora a maioria das técnicas tenha proporcionado excelentes resultados, muitos estudos focaram apenas na diferenciação da LMA, deixando de abordar suas subdivisões, aspecto crucial para um tratamento adequado. Além disso, durante a investigação, observou-se uma carência de bases de dados públicas e um desbalanceamento entre as classes.

A aquisição de imagens é um processo delicado e pode ser considerado uma das fases mais importantes. Por esse motivo, esta Revisão Sistemática buscou informações valiosas sobre as principais bases de dados utilizadas e os métodos para adquiri-las. As imagens precisam ser validadas por algum centro médico, e as bases existentes mostraram-se ineficientes devido à pequena quantidade e ao desbalanceamento das classes. Portanto, foi possível identificar que o processo de aumento de dados é extremamente importante.

A escassez de conjuntos de dados específicos para a LMA é uma questão frequentemente mencionada. Portanto, a aplicação de técnicas de aumento de dados, juntamente com o uso de modelos de aprendizado de transferência, é altamente recomendada para aprimorar o tamanho dos conjuntos de dados destinados ao treinamento e teste. Transformações de imagem convencionais, como rotação, zoom e métodos gráficos de ajuste de intensidade.

Com base nas informações apresentadas neste trabalho, a utilização de técnicas de aumento de dados se torna indispensável devido à escassez de base de imagens. Além dos métodos comuns de aumento de dados, o Aumento Baseado em Aprendizado Generativo pode ser testado. A segmentação da região de interesse é algo crucial para identificar os leucócitos imaturos. Além disso, novos métodos de diagnóstico assistidos por computador para a classificação de diferentes tipos e subtipos de LMA, baseados no esquema de classificação revisto da OMS, podem ser desenvolvidos para trabalhos futuros.

Referências

- ABHISHEK, A.; SANTHANAM, N.; JHA, R. K.; SINHA, R.; JHA, K. Multi class classification of acute leukemia using transfer learning. In: *2022 International Conference for Advancement in Technology (ICONAT)*. IEEE, 2022. Disponível em: <<http://dx.doi.org/10.1109/ICONAT53423.2022.9726083>><http://dx.doi.org/10.1109/ICONAT53423.2022.9726083>. Citado 5 vezes nas páginas 30, 31, 33, 34 e 35.
- ACEVEDO, A. *A dataset for microscopic peripheral blood cell images for development of automatic recognition systems*. Mendeley, 2020. Disponível em: <<https://data.mendeley.com/datasets/snkd93bnjr/1>>. Citado na página 30.
- BAIRABOINA, S. S. R.; BATTULA, S. R. Ghost-resnext: An effective deep learning based on mature and immature wbc classification. *Applied Sciences*, MDPI AG, v. 13, n. 6, p. 4054, mar. 2023. ISSN 2076-3417. Disponível em: <<http://dx.doi.org/10.3390/app13064054>><http://dx.doi.org/10.3390/app13064054>. Citado 6 vezes nas páginas 30, 31, 32, 33, 35 e 37.
- CLARO, M. L.; VERAS, R. d. M.; SANTANA, A. M.; VOGADO, L. H. S.; JUNIOR, G. B.; MEDEIROS, F. N. d.; TAVARES, J. M. R. Assessing the impact of data augmentation and a combination of cnns on leukemia classification. *Information Sciences*, Elsevier BV, v. 609, p. 1010–1029, set. 2022. ISSN 0020-0255. Disponível em: <<http://dx.doi.org/10.1016/j.ins.2022.07.059>><http://dx.doi.org/10.1016/j.ins.2022.07.059>. Citado 8 vezes nas páginas 30, 31, 32, 33, 34, 35, 36 e 37.
- DHARANI, T.; HARIPRASATH, S. Diagnosis of leukemia and its types using digital image processing techniques. In: *2018 3rd International Conference on Communication and Electronics Systems (ICCES)*. [S.l.: s.n.], 2018. p. 275–279. Citado na página 20.
- ELHASSAN, T. A. M.; RAHIM, M. S. M.; SWEE, T. T.; HASHIM, S. Z. M.; ALJURF, M. Feature extraction of white blood cells using cmyk-moment localization and deep learning in acute myeloid leukemia blood smear microscopic images. *IEEE Access*, Institute of Electrical and Electronics Engineers (IEEE), v. 10, p. 16577–16591, 2022. ISSN 2169-3536. Disponível em: <<http://dx.doi.org/10.1109/ACCESS.2022.3149637>><http://dx.doi.org/10.1109/ACCESS.2022.3149637>. Citado 5 vezes nas páginas 29, 31, 33, 35 e 37.
- GALVÃO, M. C. B.; RICARTE, I. L. M. Revisão sistemática da literatura: conceituação, produção e publicação. *Logeion: Filosofia da informação*, v. 6, n. 1, p. 57–73, 2019. Citado na página 23.
- JUNIOR, A. A.; MARINHO, M.; MEDEIROS, A. C. F. de; FILHO, C. J. A. B. Comparação de modelos de segmentação utilizados na identificação de linfócitos maduros e imaturos em lâminas sanguíneas. *Revista de Engenharia e Pesquisa Aplicada*, v. 7, n. 2, p. 12–22, 2022. Citado na página 16.
- KRISHNA, S. T.; KALLURI, H. K. Deep learning and transfer learning approaches for image classification. *International Journal of Recent Technology and Engineering (IJRTE)*, v. 7, n. 5S4, p. 427–432, 2019. Citado na página 21.

KUMAR, R.; KUMBHARKAR, P.; VANAM, S.; SHARMA, S. Medical images classification using deep learning: a survey. *Multimedia Tools and Applications*, Springer, p. 1–46, 2023. Citado 2 vezes nas páginas 14 e 20.

LIANG, H.; FU, W.; YI, F. A survey of recent advances in transfer learning. In: *2019 IEEE 19th International Conference on Communication Technology (ICCT)*. IEEE, 2019. Disponível em: <<http://dx.doi.org/10.1109/ICCT46805.2019.8947072>><http://dx.doi.org/10.1109/ICCT46805.2019.8947072>. Citado na página 22.

MATEK, C.; SCHWARZ, S.; MARR, C.; SPIEKERMANN, K. *A Single-cell Morphological Dataset of Leukocytes from AML Patients and Non-malignant Controls*. The Cancer Imaging Archive, 2019. Disponível em: <https://www.cancerimagingarchive.net/collection/aml-cytomorphology_lmu/>. Citado na página 29.

MATEK, C.; SCHWARZ, S.; SPIEKERMANN, K.; MARR, C. Human-level recognition of blast cells in acute myeloid leukemia with convolutional neural networks. Cold Spring Harbor Laboratory, fev. 2019. Disponível em: <<http://dx.doi.org/10.1101/564039>><http://dx.doi.org/10.1101/564039>. Citado 4 vezes nas páginas 29, 31, 33 e 35.

MITTELMAN, M. The french-american-british classification of myelodysplastic syndromes. *Haematologica*, Ferrata Storti Foundation, v. 108, n. 10, p. 2559, 2023. Citado na página 18.

NAGIUB, E. M.; HUSSAIN, K. F.; OMAR, N. M.; AL-RASHEDI, Q. Acute myeloid leukemia diagnosis using deep learning. *EGYPTIAN JOURNAL OF HAEMATOLOGY*, v. 45, n. 4, p. 167–174, OCT-DEC 2020. ISSN 1110-1067. Citado 4 vezes nas páginas 29, 31, 33 e 35.

NARAYANAN, D.; WEINBERG, O. K. How i investigate acute myeloid leukemia. *International journal of laboratory hematology*, Wiley Online Library, v. 42, n. 1, p. 3–15, 2020. Citado na página 14.

RAHMAN, J. F.; AHMAD, M. Detection of acute myeloid leukemia from peripheral blood smear images using transfer learning in modified cnn architectures. In: _____. *Studies in Autonomic, Data-driven and Industrial Computing*. Springer Nature Singapore, 2023. p. 447–459. ISBN 9789811975288. Disponível em: <http://dx.doi.org/10.1007/978-981-19-7528-8_35>http://dx.doi.org/10.1007/978-981-19-7528-8_35. Citado 5 vezes nas páginas 29, 31, 32, 33 e 35.

RASTOGI, P.; KHANNA, K.; SINGH, V. Leufeatx: Deep learning–based feature extractor for the diagnosis of acute leukemia from microscopic images of peripheral blood smear. *Computers in Biology and Medicine*, Elsevier BV, v. 142, p. 105236, mar. 2022. ISSN 0010-4825. Disponível em: <<http://dx.doi.org/10.1016/j.combiomed.2022.105236>><http://dx.doi.org/10.1016/j.combiomed.2022.105236>. Citado 8 vezes nas páginas 29, 30, 31, 32, 33, 34, 35 e 37.

ROY, R. M.; AMEER, P. M. Identification of white blood cells for the diagnosis of acute myeloid leukemia. *International Journal of Imaging Systems and Technology*, Wiley, v. 32, n. 4, p. 1307–1317, jan. 2022. ISSN 1098-1098. Disponível em: <<http://dx.doi.org/10.1002/ima.22702>><http://dx.doi.org/10.1002/ima.22702>. Citado 6 vezes nas páginas 29, 31, 33, 34, 35 e 37.

SANTOS, M. M. F. dos; JESUS, G. P. de; FERREIRA, L. P.; FERREIRA, R. Leucemia mielóide, aguda e crônica: diagnósticos e possíveis tratamentos. 2019. Citado na página 14.

SARKER, I. H. Machine learning: Algorithms, real-world applications and research directions. *SN Computer Science*, Springer Science and Business Media LLC, v. 2, n. 3, mar. 2021. ISSN 2661-8907. Disponível em: <<http://dx.doi.org/10.1007/s42979-021-00592-x>><http://dx.doi.org/10.1007/s42979-021-00592-x>. Citado 5 vezes nas páginas 20, 31, 33, 34 e 35.

SHAH, A.; NAQVI, S. S.; NAVEED, K.; SALEM, N.; KHAN, M. A. U.; ALIMGEER, K. S. Automated diagnosis of leukemia: A comprehensive review. *IEEE Access*, Institute of Electrical and Electronics Engineers (IEEE), v. 9, p. 132097–132124, 2021. ISSN 2169-3536. Disponível em: <<http://dx.doi.org/10.1109/ACCESS.2021.3114059>><http://dx.doi.org/10.1109/ACCESS.2021.3114059>. Citado 2 vezes nas páginas 16 e 17.

SOCIETY, A. C. *Acute Myeloid Leukemia (AML) Subtypes and Prognostic Factors*. 2018. Disponível em: <<https://www.cancer.org/cancer/types/acute-myeloid-leukemia/detection-diagnosis-staging/how-classified.html>>. Citado 3 vezes nas páginas 17, 18 e 20.

SOCIETY, A. C. *Key Statistics for Acute Myeloid Leukemia (AML)*. 2023. Disponível em: <<https://www.cancer.org/cancer/types/acute-myeloid-leukemia/about/key-statistics.html>>. Citado na página 14.

TRESSO, M. Métodos diagnósticos da leucemia mieloide aguda. *Ciência News*. Disponível em: http://www.ciencianews.com.br/arquivos/ACET/IMAGENS/biblioteca-digital/hematologia/serie_branca/leucemias_linfomas_mieloma/leucemias/24-Metodos-diagnostico-da-leucemia-mieloide-aguda.pdf. Acesso em, v. 29, 2021. Citado na página 14.

WANG, M.; DONG, C.; GAO, Y.; LI, J.; HAN, M.; WANG, L. A deep learning model for the automatic recognition of aplastic anemia, myelodysplastic syndromes, and acute myeloid leukemia based on bone marrow smear. *Frontiers in Oncology*, v. 12, 2022. Cited by: 2; All Open Access, Gold Open Access, Green Open Access. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85148372714doi=10.3389%2ffonc.2022.844978partnerID=40md5=845b2ffd5f100131bd4e51afe167c905>>. Citado 6 vezes nas páginas 29, 30, 31, 33, 34 e 35.

YU, J.; JIANG, P. Y.; SUN, H.; ZHANG, X.; JIANG, Z.; LI, Y.; SONG, Y. Advances in targeted therapy for acute myeloid leukemia. *Biomarker research*, BioMed Central, v. 8, n. 1, p. 1–11, 2020. Citado na página 17.

YU, Z.; LI, J.; WEN, X.; HAN, Y.; JIANG, P.; ZHU, M.; WANG, M.; GAO, X.; SHEN, D.; ZHANG, T.; ZHAO, S.; ZHU, Y.; TONG, J.; YUAN, S.; ZHU, H.; HUANG, H.; QIAN, P. Amlnet, a deep-learning pipeline for the differential diagnosis of acute myeloid leukemia from bone marrow smears. *Journal of Hematology amp; Oncology*, Final, v. 16, n. 1, mar. 2023. ISSN 1756-8722. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85150905647doi=10.1186%2fs13045-023-01419-3partnerID=40md5=da6130204ddd093a882332396880365f>>. Citado 5 vezes nas páginas 30, 31, 33, 35 e 36.

ZERBINI, M. C. N.; SOARES, F. A.; VELLOSO, E. D. R. P.; CHAUFAILLE, M. d. L. L.; PAES, R. P. Classificação da organização mundial da saúde para os tumores dos tecidos hematopoético e linfóide, 4^a edição, 2008: principais modificações introduzidas em relação à 3^a edição, 2001. *Revista da Associação Médica Brasileira*, SciELO Brasil, v. 57, p. 6–73, 2011. Citado na página 18.



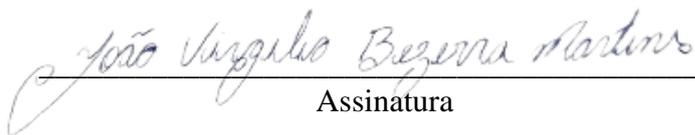
**TERMO DE AUTORIZAÇÃO PARA PUBLICAÇÃO DIGITAL NA BIBLIOTECA
“JOSÉ ALBANO DE MACEDO”**

Identificação do Tipo de Documento

- () Tese
- () Dissertação
- (X) Monografia
- () Artigo

Eu, **João Virgílio Bezerra Martins**, autorizo com base na Lei Federal nº 9.610 de 19 de Fevereiro de 1998 e na Lei nº 10.973 de 02 de dezembro de 2004, a biblioteca da Universidade Federal do Piauí a divulgar, gratuitamente, sem ressarcimento de direitos autorais, o texto integral da publicação **Revisão Sistemática dos Métodos de Aprendizagem Profundo para Detecção e Classificação da Leucemia Mieloide Aguda** de minha autoria, em formato PDF, para fins de leitura e/ou impressão, pela internet a título de divulgação da produção científica gerada pela Universidade.

Picos-PI 01 de fevereiro de 2024


Assinatura