

UNIVERSIDADE FEDERAL DE JUIZ DE FORA
INSTITUTO DE CIÊNCIAS EXATAS
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

Juliana Hansen Basdão Fam

**Desenvolvimento de Modelo de Aprendizado de Máquina para Segmentação
de Eosinófilos em Imagens Histológicas**

Juiz de Fora

2026

Juliana Hansen Basdão Fam

**Desenvolvimento de Modelo de Aprendizado de Máquina para Segmentação
de Eosinófilos em Imagens Histológicas**

Trabalho de conclusão de curso apresentado
ao Instituto de Ciências Exatas da Univer-
sidade Federal de Juiz de Fora como requisito
parcial à obtenção do grau de bacharel em
Ciência da Computação.

Orientadora: Profa Dra. Bárbara de Melo Quintela

Coorientadores:

M.Sc. Ana Clara Verly

M.Sc. Thiago Esterci Fernandes

Juiz de Fora

2026

Hansen Basdão Fam, Juliana.

Desenvolvimento de Modelo de Aprendizado de Máquina para Segmentação de Eosinófilos em Imagens Histológicas / Juliana Hansen Basdão Fam. – 2026.

44 f. : il.

Orientadora: Profa Dra. Bárbara de Melo Quintela

Coorientadores:

M.Sc. Ana Clara Verly

M.Sc. Thiago Esterci Fernandes

Trabalho de Conclusão de Curso (graduação) – Universidade Federal de Juiz de Fora, Instituto de Ciências Exatas. Bacharelado em Ciência da Computação, 2026.

1. Eosinófilos. 2. Imagens Histológicas. 3. Processamento de Imagens. 4. Aprendizado de Máquina. 5. Redes Neurais Convolucionais. I. de Melo Quintela, Bárbara, orient. II. Título

Juliana Hansen Basdão Fam

**Desenvolvimento de Modelo de Aprendizado de Máquina para Segmentação
de Eosinófilos em Imagens Histológicas**

Trabalho de conclusão de curso apresentado
ao Instituto de Ciências Exatas da Universidade Federal de Juiz de Fora como requisito
parcial à obtenção do grau de bacharel em
Ciência da Computação.

Aprovada em 10 de Julho de 2026.

BANCA EXAMINADORA

Profa Dra. Bárbara de Melo Quintela - Orientadora
Universidade Federal de Juiz de Fora

Profa Dra Luciana Conceição Dias Campos
Universidade Federal de Juiz de Fora

Prof Dr Luiz Maurilio da Silva Maciel
Universidade Federal de Juiz de Fora

RESUMO

A contagem de eosinófilos, células do sistema imunológico, em imagens histológicas é fundamental para pesquisas biomédicas que estudam respostas inflamatórias do organismo. Porém, a demarcação dessas células nas imagens para contagem é, em grande parte, realizada manualmente, demandando muitas horas de trabalho dos pesquisadores - tempo que poderia ser direcionado a atividades intelectuais, analíticas e interpretativas. Diante desse cenário, torna-se relevante o desenvolvimento de soluções automatizadas que auxiliem na identificação e quantificação dessas células. Assim, com o objetivo de automatizar a segmentação e contagem de eosinófilos e aumentar a produtividade das pesquisas biomédicas, este trabalho propõe o desenvolvimento de um modelo computacional baseado em técnicas de aprendizado de máquina com redes neurais convolucionais para a detecção de eosinófilos em imagens microscópicas. Para isso, foi treinado um modelo YOLO11 em sua variante nano, que foi avaliado qualitativamente por Biólogos e quantitativamente por métricas utilizadas na visão computacional. Assim, este trabalho contribui para a redução do esforço manual e maior eficiência na análise de imagens histológicas, reduzindo o tempo de contagem de células de horas para segundos.

Palavras-chave: eosinófilos; patologia digital; processamento de imagens; aprendizado de máquina; redes neurais convolucionais.

ABSTRACT

The counting of eosinophils, immune system cells, in histological images is essential for biomedical research that investigates inflammatory responses of the organism. However, the annotation of these cells for counting purposes is largely performed manually, requiring many hours of work from researchers - time that could otherwise be devoted to intellectual, analytical, and interpretative activities. In this context, the development of automated solutions to assist in the identification of these cells becomes highly relevant. Therefore, in order to automate the segmentation and counting of eosinophils and to increase the productivity of biomedical research, this work proposes the development of a computational model based on machine learning techniques for the detection and counting of eosinophils in microscopic images. To achieve this, a YOLO11 model in its “nano” variant was trained and evaluated qualitatively by biologists and quantitatively using metrics employed in computer vision. Thus, this project contributes to reducing manual effort and increasing efficiency in the analysis of histological images, cutting cell counting time from hours to seconds.

Keywords: eosinophils; histological images; image processing; machine learning; convolutional neural networks.

LISTA DE ILUSTRAÇÕES

Figura 1 – Representação de eosinófilos em granuloma.	13
Figura 2 – Representações esquemáticas: (a) rede neural artificial com múltiplas camadas ocultas; (b) neurônio artificial, incluindo entradas, pesos, soma ponderada, função de ativação e saída.	18
Figura 3 – Relação hierárquica entre Inteligência Artificial, Aprendizado de Máquina, Redes Neurais, Aprendizado Profundo e Redes Neurais Convolucionais.	20
Figura 4 – Representação de exemplo de convolução.	20
Figura 5– Arquitetura do YOLO11.	22
Figura 6 – Matriz de confusão para geração das métricas.	23
Figura 7 – Fluxograma do <i>pipeline</i> metodológico proposto.	28
Figura 8– Divisão de Grupos para Validação Cruzada.	33
Figura 9 – Marcações dos eosinófilos.	37

LISTA DE TABELAS

Tabela 1 – Resumo estatístico final consolidado (Média $\pm$ Desvio Padrão).	36
--	----

SUMÁRIO

1	INTRODUÇÃO	10
1.1	JUSTIFICATIVA	10
1.2	OBJETIVOS	11
1.3	ORGANIZAÇÃO DO TRABALHO	12
2	FUNDAMENTAÇÃO TEÓRICA	13
2.1	CONTEXTO BIOLÓGICO	13
2.1.1	Eosinófilos	13
2.1.2	Granulomas	13
2.1.3	Processos Inflamatórios e o Papel dos Eosinófilos	14
2.2	HISTOLOGIA	14
2.3	MÉTODOS COMPUTACIONAIS PARA ANÁLISE DE IMAGENS HISTOLÓGICAS	15
2.3.1	Segmentação de Imagens	16
2.3.2	Extração de Características	16
2.3.3	Aprendizado de Máquina	17
2.3.4	Redes Neurais Artificiais	17
2.3.5	Aprendizado Profundo (<i>Deep Learning</i>)	18
2.3.6	Redes Neurais Convolucionais (CNNs)	19
2.3.7	Dados para Treinamento de Modelos	21
2.3.8	YOLO11	21
2.4	MÉTRICAS DE AVALIAÇÃO	23
2.4.1	Precisão (<i>Precision</i>)	24
2.4.2	Sensibilidade ou Revocação (<i>Recall</i>)	24
2.4.3	Acurácia (<i>Accuracy</i>)	24
2.4.4	<i>F1-Score</i>	24
2.4.5	Média da Precisão Média (mAP - <i>Mean Average Precision</i>)	25
2.5	DESAFIOS DA ANÁLISE COMPUTACIONAL DE IMAGENS HISTOLÓGICAS	26
3	MÉTODOS	28
3.1	ACESSO AOS DADOS	29
3.2	PRÉ-PROCESSAMENTO DAS IMAGENS	29
3.2.1	Criação dos Gabaritos	29
3.2.2	Divisão em Patches	29
3.2.3	<i>Data Augmentation</i>	31
3.3	TREINAMENTO DO MODELO	31
3.3.1	Validação Cruzada	31
3.3.2	Desenvolvimento e Aplicação do <i>Pipeline</i>	33

3.4	AVALIAÇÃO DO MODELO	35
4	RESULTADOS E DISCUSSÃO	36
4.1	DESEMPENHO DO MODELO NO <i>PIPELINE</i>	36
4.2	VALIDAÇÃO POR ESPECIALISTAS	38
5	CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS . .	39
	REFERÊNCIAS	40

1 INTRODUÇÃO

Imagens histológicas são aquelas obtidas com o uso do microscópio que mostram a estrutura microscópica de células e tecidos biológicos. As análises destas imagens desempenham um papel essencial no diagnóstico de doenças e em pesquisas que buscam entender as respostas imunológicas do organismo a certas patologias e tratamentos destas (GUPTA et al. 2020).

A interseção entre a Ciência da Computação e a Patologia Digital tem avançado através da tecnologia de *Whole Slide Imaging* (WSI), que permite a digitalização de seções histológicas inteiras em alta resolução (MELO et al. 2020), fornecendo imagens detalhadas para que modelos computacionais sejam treinados e consigam avaliar essas imagens com alta precisão.

Cortes histológicos digitalizados em alta resolução podem conter milhares de células distribuídas em tecidos complexos. A análise dessas imagens e a demarcação de células e de estruturas de interesse demandam tempo excessivo de trabalho humano quando realizadas manualmente. Por conta disso, estudar esses cortes é um desafio que demanda esforço humano para realizar essas demarcações (DIMITRIOU et al. 2019).

No contexto deste trabalho, as células de interesse são eosinófilos, glóbulos brancos que atuam na defesa do sistema imunológico, presentes em granulomas, área de inflamação com um aglomerado de células imunes que tentam isolar fisicamente corpos estranhos. Essas células são objeto de estudo dos pesquisadores do Laboratório de Biologia Celular da Universidade Federal de Juiz de Fora, que realizam as marcações manualmente e forneceram as imagens e marcações geradas por eles para que seja treinado um modelo de aprendizado de máquina para identificar, classificar, demarcar e contabilizar essas células, criando uma ferramenta de automação do reconhecimento dessas estruturas.

1.1 JUSTIFICATIVA

Os avanços em automação e inteligência artificial têm transformado a forma de lidar com tarefas de identificação, marcação e classificação de células em imagens histológicas. Essa interdisciplinaridade com as ciências biológicas está alinhada às recomendações das National Academies of Sciences, Engineering, and Medicine, que defendem a convergência entre as ciências da vida e a computação para a resolução de problemas complexos NATIONAL RESEARCH COUNCIL 2014. Ao substituir etapas repetitivas da análise por um *pipeline* automatizado, este trabalho busca reduzir o tempo gasto com a contagem e a marcação manuais de células, permitindo que os pesquisadores se concentrem na interpretação dos resultados.

Entretanto, para que uma arquitetura de inteligência artificial seja capaz de identificar corretamente os eosinófilos, é necessário treiná-la com imagens histológicas repre-

sentativas do problema em investigação. Embora modelos de detecção e segmentação apresentem arquiteturas previamente definidas, não possuem, inicialmente, conhecimento específico sobre a aparência dos eosinófilos nas lâminas utilizadas neste trabalho. Durante o treinamento, são apresentadas ao modelo imagens previamente anotadas por especialistas, nas quais a localização e a classe das células de interesse são indicadas. A partir desses exemplos, o modelo ajusta seus parâmetros internos e aprende características visuais relacionadas à forma, à dimensão, à textura e à coloração dos eosinófilos, diferenciando-os das demais estruturas presentes no tecido.

Esse processo é particularmente importante em imagens histológicas, uma vez que a aparência das células pode variar em decorrência da técnica de coloração, das condições de preparação da lâmina, do equipamento utilizado na digitalização e das próprias características morfológicas do tecido. Dessa forma, um modelo treinado com imagens obtidas sob determinada coloração ou em condições distintas pode não apresentar desempenho satisfatório quando aplicado diretamente a outro conjunto de dados. O treinamento com imagens específicas do contexto analisado permite adaptar o modelo a essas particularidades, reduzindo erros de identificação, como falsos positivos e falsos negativos, e aumentando a confiabilidade das contagens realizadas.

Muitos estudos relacionados à utilização de inteligência artificial na análise de imagens histológicas ainda não exploram abordagens modernas e arquiteturas de detecção em tempo real de última geração, como o YOLO11, que possibilitam a detecção e a segmentação de objetos. Além disso, grande parte dos conjuntos de dados, das anotações e dos códigos desenvolvidos não está disponível no domínio público BAMWENDA et al. 2025. A indisponibilidade desses recursos dificulta a reprodução dos estudos e impede a utilização direta de modelos previamente treinados. Por esse motivo, torna-se necessário construir um conjunto de dados próprio, anotar as células de interesse e treinar uma ferramenta adequada às imagens utilizadas pelos pesquisadores da área de Ciências Biológicas da Universidade Federal de Juiz de Fora.

Assim, este trabalho contribui para a ciência ao desenvolver uma ferramenta capaz de substituir horas de atividades repetitivas por um processo automatizado de identificação e quantificação celulares. Essa automatização pode ampliar a produtividade e a reprodutibilidade das pesquisas biomédicas, permitindo que um maior tempo seja dedicado à análise dos dados coletados e à investigação das respostas dos organismos a diferentes patologias.

1.2 OBJETIVOS

O objetivo deste trabalho é desenvolver um *pipeline* computacional baseado em aprendizado de máquina capaz de identificar, classificar e contar eosinófilos em imagens histológicas, contornando as dificuldades da marcação e contagem manual, através dos

seguintes objetivos específicos:

- Implementar técnicas de pré-processamento nas imagens .tif, formato de imagem digital de alta qualidade que se destaca por não perder nenhuma informação ou cor original ao ser salvo;
- Treinar e ajustar modelos de Redes Neurais Convolucionais utilizando a base de dados rotulada pelo Laboratório de Biologia Celular;
- Avaliar o desempenho do modelo através de métricas;
- Propor melhorias e trabalhos futuros de desenvolvimento de um *software* funcional que forneça uma interface para que o usuário envie uma imagem e receba as células demarcadas e a contagem total.

1.3 ORGANIZAÇÃO DO TRABALHO

Este trabalho está estruturado em cinco capítulos.

- **Capítulo 1 - Introdução:** Contextualiza o problema abordado, sua justificativa, os objetivos do trabalho e a organização do documento.
- **Capítulo 2 - Fundamentação Teórica:** Apresenta os conceitos necessários para a compreensão do trabalho. Apresenta fundamentos da biologia de histologia, processamento de imagens biomédicas e conceitos da computação de aprendizado de máquina, aprendizado profundo, redes neurais convolucionais, a arquitetura YOLO11 e as métricas utilizadas para a avaliação.
- **Capítulo 3 - Métodos:** Descreve a metodologia empregada no desenvolvimento do trabalho, como a obtenção e preparação dos dados, o pré-processamento das imagens, o processo de treinamento e a estratégia adotada para avaliar o modelo.
- **Capítulo 4 - Resultados e Discussão:** Apresenta os resultados obtidos por meio de métricas e validação qualitativa realizada por especialistas.
- **Capítulo 5 - Considerações Finais e Trabalhos Futuros:** Apresenta as considerações finais, destacando as contribuições do trabalho, suas limitações e possíveis direções para desenvolvimentos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

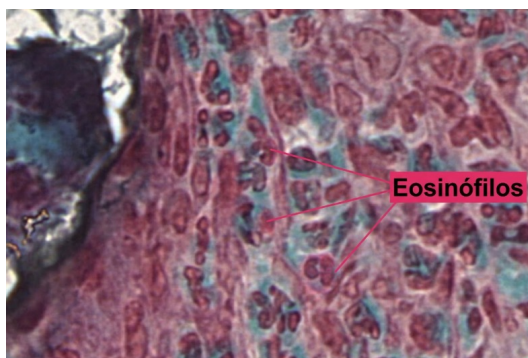
2.1 CONTEXTO BIOLÓGICO

Alguns conceitos de citologia e imunologia, fundamentais para a compreensão desse trabalho, são apresentados nesta seção.

2.1.1 Eosinófilos

Eosinófilos são glóbulos brancos, células do sistema imunológico que defendem o organismo, por exemplo, contra infecções por parasitas, alergias e doenças inflamatórias. São facilmente identificados nas lâminas por possuírem núcleo segmentado (geralmente bilobado, ou seja, dividido em dois segmentos) e grande afinidade com certos corantes, que destacam suas cores (BARATA et al. 2025), como apresentado na Figura 1. Esse é um passo necessário para a etapa de treinamento

Figura 1 - Representação de eosinófilos em granuloma.



Fonte: Laboratório de Biologia Celular da Universidade Federal de Juiz de Fora.

2.1.2 Granulomas

Granulomas são estruturas formadas por células do sistema imunológico que atuam como mecanismo de defesa, criando uma barreira física que isola um agente invasor, como ovos de parasitas, minimizando os danos causados por corpos estranhos (MALTA et al. 2022).

O granuloma apresenta uma organização espacial previsível, dividindo-se em três zonas: periovular (ao redor do ovo), medial e periférica (BARATA et al. (2025)).

2.1.3 Processos Inflamatórios e o Papel dos Eosinófilos

A inflamação é uma resposta do sistema imune desencadeada por lesões teciduais, infecções microbianas, reações alérgicas ou estímulos nocivos. Esse processo tem como objetivo principal localizar e eliminar o agente agressor, remover as células danificadas e ativar os mecanismos de reparo do tecido afetado. A resposta imune ocorre por meio de uma cascata de eventos vasculares e celulares, caracterizada pela liberação de mediadores químicos que promovem o recrutamento de leucócitos do compartimento sanguíneo para o foco da lesão (MEDZHITOV 2008).

Os eosinófilos são leucócitos granulócitos especializados que desempenham funções cruciais em processos inflamatórios específicos, com destaque para as reações alérgicas crônicas (como a asma brônquica), infecções por parasitas helmintos e na formação de granulomas (MCBRIEN et al. 2017).

Diante desse cenário, a contagem de eosinófilos em cortes histológicos (fatias extremamente finas de tecidos biológicos que permitem a análise detalhada de células) tem grande relevância diagnóstica e prognóstica. A quantificação exata dessas células em biópsias permite aos patologistas avaliar a gravidade da resposta inflamatória, monitorar a progressão de doenças de perfil eosinofílico e mensurar a eficácia de tratamentos.

2.2 HISTOLOGIA

A histologia é a ciência dedicada ao estudo dos tecidos biológicos, estruturas formadas por conjuntos organizados de células que atuam de maneira integrada no desempenho de funções específicas SOUZA et al. 2010. A análise histológica permite investigar a organização e as alterações morfológicas dos tecidos, sendo amplamente utilizada em pesquisas biológicas e na prática clínica. Para que essa análise seja realizada adequadamente, diferentes etapas devem ser consideradas, desde a preparação e coloração dos tecidos até a digitalização e interpretação das lâminas histológicas.

A escolha da técnica de coloração é fundamental para proporcionar o contraste necessário à identificação das estruturas celulares, tanto em análises manuais quanto em abordagens automatizadas. Para a identificação de eosinófilos em cortes histológicos, por exemplo, pode-se utilizar a coloração por Hematoxilina e Eosina (HE), técnica padrão na patologia clínica. Nessa técnica, os núcleos celulares são corados em azul ou púrpura, enquanto o citoplasma e os grânulos dos eosinófilos apresentam tonalidades rosa-intensas ou laranja-avermelhadas, devido à sua natureza acidofílica. Nas imagens utilizadas neste trabalho, foi empregada a técnica *Fast Green-Neutral Red*, na qual os eosinófilos apresentam coloração verde, proporcionando elevado contraste em relação às demais estruturas do tecido KIERNAN 2015.

Após a preparação e a coloração, as lâminas podem ser convertidas em formato

digital por meio da tecnologia de *Whole Slide Imaging* (WSI). Esse processo consiste na digitalização completa da lâmina histológica utilizando *scanners* de alta resolução, que capturam milhares de campos individuais e os combinam para formar uma única imagem digital de dimensões da ordem de *gigapixels* FARAHANI et al. 2015. Essas imagens podem ser navegadas em um computador de maneira semelhante à observação realizada em um microscópio convencional, permitindo a visualização do tecido em diferentes níveis de ampliação, como $20\times$ e $40\times$, com elevado detalhamento morfológico. A preservação desses detalhes é especialmente importante para a aplicação de algoritmos de inteligência artificial em patologia, pois esses métodos aprendem a reconhecer células e padrões teciduais com base nas características presentes nas imagens.

A digitalização também possibilita o desenvolvimento de métodos computacionais voltados à redução das limitações associadas à análise manual das lâminas. A avaliação realizada exclusivamente por especialistas é geralmente trabalhosa e demorada, além de estar sujeita à fadiga e a falhas humanas que podem resultar em erros de identificação ou de contagem de células. Outro desafio relevante é a variabilidade interobservador e intraobservador, pois diferentes especialistas, ou até mesmo um único especialista em momentos distintos, podem apresentar resultados divergentes ao analisar a mesma amostra SANTOS et al. 2023. Nesse contexto, a análise automatizada de imagens histológicas pode contribuir para tornar o processo mais rápido, reprodutível e objetivo, desde que sejam consideradas as particularidades das técnicas de coloração, da aquisição de imagens e das estruturas celulares investigadas.

2.3 MÉTODOS COMPUTACIONAIS PARA ANÁLISE DE IMAGENS HISTOLÓGICAS

Uma imagem digital é composta por uma grade de elementos fundamentais chamados *pixels*. Cada *pixel* possui um valor numérico que representa a intensidade ou cor naquela posição específica da imagem (CHADHA et al. 2020). Em patologia digital, as imagens de lâmina inteira (WSI) são capturadas em resoluções altíssimas, permitindo que o modelo aprenda representações complexas e abstratas diretamente a partir de valores tipicamente entre $0,25\ \mu\text{m}$ e $0,5\ \mu\text{m}$ por *pixel* (MELO et al. 2020). Uma única WSI pode conter bilhões de *pixels*.

O processamento de imagens biomédicas compreende o conjunto de técnicas computacionais utilizadas para manipular, extrair e destacar informações relevantes. No contexto da microscopia, essas técnicas preparam os dados brutos para que algoritmos de alta complexidade consigam isolar estruturas celulares com precisão (BANKMAN 2008).

Uma imagem digital pode ser representada matematicamente por uma matriz bidimensional $f(x, y)$, na qual cada par de coordenadas (x, y) corresponde a um elemento discreto denominado *pixel*. O valor de $f(x, y)$ representa a intensidade luminosa ou o nível

de cinza associado a esse ponto da imagem (GONZALEZ et al. 2018).

A resolução da imagem define a densidade de *pixels* por unidade de área espacial, limitando o nível de detalhes morfológicos visíveis. Nas imagens histológicas, as dimensões são frequentemente ampliadas para preservar estruturas subcelulares (PANTANOWITZ et al. 2010).

O espaço de cor mais comum na computação é o RGB (*Red, Green, Blue*), composto por três canais de cor independentes. Cada canal armazena uma matriz de intensidade correspondente ao seu comprimento de onda, sendo cada *pixel* a intensidade daquele comprimento de onda na formação da cor do *pixel*. (GONZALEZ et al. 2018). Para a análise de tecidos corados, os canais de cor atuam como base para diferenciar os tons que caracterizam os eosinófilos e aprimorar a visualização (BARATA et al. 2025).

O pré-processamento visa elevar a qualidade visual das imagens e padronizar o comportamento dos dados antes de alimentarem os modelos preditivos (JANOWCZYK et al. 2016). Entre as técnicas principais, destacam-se a segmentação de imagens e a extração de características.

2.3.1 Segmentação de Imagens

A segmentação consiste no processo de agrupar subconjuntos de *pixels* que compartilham características visuais semelhantes, isolando os objetos de interesse do plano de fundo (*background*). Em contextos biomédicos, a segmentação pode ocorrer de forma semântica (classificando cada *pixel* da imagem em uma categoria, sem distinguir objetos individuais) ou por instância (onde cada célula identificada recebe um rótulo exclusivo, permitindo a contagem individualizada) (XING et al. 2016).

2.3.2 Extração de Características

Após o isolamento das regiões de interesse, aplica-se a extração de características para transformar os *pixels* brutos em vetores de propriedades quantificáveis. Essas propriedades dividem-se em descritores de forma (área, perímetro, circularidade e excentricidade), descritores de textura (padrões de rugosidade e distribuição de granulações no citoplasma) e descritores cromáticos (médias e variâncias de intensidade nos canais de cor) (GURCAN et al. 2009). Esses vetores servem como a assinatura matemática que permite aos classificadores diferenciar uma célula real de um artefato de tecido.

A análise moderna de imagens histológicas baseia-se em algoritmos capazes de aprender, de forma autônoma, a reconhecer padrões biológicos complexos (SANTOS et al. 2022), unindo as tarefas de segmentação, que localiza os limites de cada estrutura de interesse no fragmento de tecido, e de classificação, que determina o tipo de cada estrutura encontrada. Detectores de objetos realizam a classificação e desenham caixas

delimitadoras (*bounding boxes*) ao redor dos alvos simultaneamente, otimizando o tempo de processamento em pipelines de contagem (LITJENS et al. 2017).

2.3.3 Aprendizado de Máquina

O Aprendizado de Máquina (*Machine Learning*) refere-se ao subcampo da Ciência da Computação que desenvolve algoritmos capazes de extrair regras de associação, identificar padrões e tomar decisões com base em dados. Em sistemas supervisionados, o algoritmo recebe um conjunto de dados de entrada associado a suas respectivas respostas corretas (*labels*) e busca ajustar uma função matemática que minimize o erro entre suas previsões realizadas e as respostas corretas conhecidas da entrada (MITCHELL 1997).

2.3.4 Redes Neurais Artificiais

Redes Neurais Artificiais são algoritmos de Aprendizado de Máquina inspirados no funcionamento dos neurônios biológicos (HAYKIN 2001). Como mostrado na Figura 2(a), possuem uma camada de entrada dos dados (representados por x_1, x_2, x_3), camadas ocultas (representadas por $L_1, L_2, \dots, L_{n-1}$), que extraem e abstraem as características intrínsecas dos dados, e a camada de saída que exhibe o resultado final da predição do modelo (y_1, y_2) (FACELI et al. 2021).

Cada neurônio de uma camada é conectado aos neurônios da camada seguinte por meio de valores numéricos chamados pesos sinápticos (representados por w_{ij}^k). O número superior (k) indica em qual camada a conexão ocorre, enquanto os números inferiores (i e j) mostram exatamente de qual neurônio a informação sai e em qual ela será representada. Esse conjunto de pesos define como os dados trafegam pela rede neural (HAYKIN 2001).

A unidade fundamental que compõe as camadas ocultas e de saída é o neurônio artificial, representado na Figura 2(b). O fluxo de processamento de um neurônio segue as seguintes etapas:

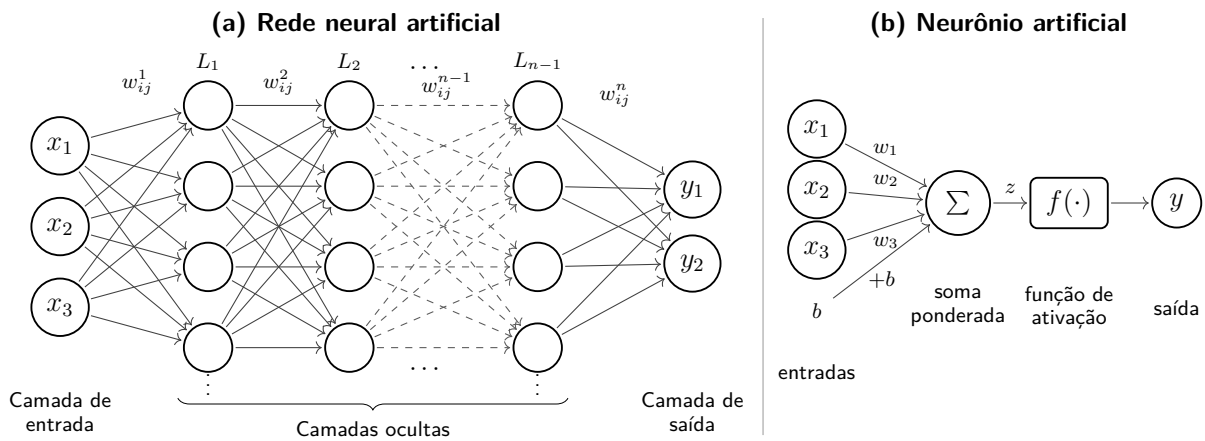
1. **Soma Ponderada:** O neurônio recebe o vetor de m entradas ($x_1, x_2, x_3, \dots, x_m$) multiplicado por seus respectivos pesos sinápticos ($w_1, w_2, w_3, \dots, w_m$). Depois, o neurônio soma todos esses resultados e adiciona um valor de ajuste, chamado *bias* (b), que desloca a curva da função de ativação para permitir uma melhor convergência (FACELI et al. 2021). O resultado gera um potencial de ativação z , expresso por:

$$z = \left(\sum_{i=1}^m x_i \cdot w_i \right) + b \quad (2.1)$$

2. **Função de Ativação:** O valor linear consolidado z é submetido a uma função de ativação não-linear, representada genericamente por $f(\cdot)$, que atua como um "filtro" e injeta não-linearidade no sistema (GOODFELLOW et al. 2016). Isso é o que permite à rede neural aprender padrões complexos que equações lineares simples nunca conseguiriam resolver, como a segmentação de imagens de biologia celular.

3. **Saída:** Após passar pelo filtro da função de ativação, o neurônio gera o valor de saída final $y = f(z)$. Nas camadas ocultas do modelo, essa saída serve imediatamente como entrada para os neurônios da próxima camada, até atingir a camada de saída final da rede neural (HAYKIN 2001).

Figura 2 - Representações esquemáticas: (a) rede neural artificial com múltiplas camadas ocultas; (b) neurônio artificial, incluindo entradas, pesos, soma ponderada, função de ativação e saída.



Fonte: Adaptado de FU et al. 2023.

2.3.5 Aprendizado Profundo (*Deep Learning*)

Aprendizado Profundo (*Deep Learning*) é uma evolução do conceito de Redes Neurais Artificiais, cuja principal diferença reside na quantidade de camadas ocultas empilhadas entre a entrada e a saída do modelo. Enquanto as redes tradicionais utilizam apenas uma ou duas camadas intermediárias, as redes profundas são compostas por dezenas ou até centenas delas. Esse processo de propagação cria uma hierarquia de aprendizado automatizada, onde cada nível de camadas especializa-se em extrair um tipo de informação (LECUN et al. 2015).

O *Deep Learning* pode ser aplicado em diferentes contextos. Tratando-se de imagens, como neste trabalho, pode-se considerar os seguintes níveis de camadas:

- **Camadas Iniciais:** Aprendem características visuais de baixo nível, como bordas geométricas, transições de contraste, linhas e texturas isoladas.
- **Camadas Intermediárias:** Combinam os padrões de linhas e bordas identificados anteriormente para reconhecer formas geométricas mais complexas e contornos específicos.

- **Camadas Profundas:** Agrupam as formas abstratas para consolidar a compreensão de estruturas completas, tornando o modelo capaz de distinguir objetos complexos com alta precisão.

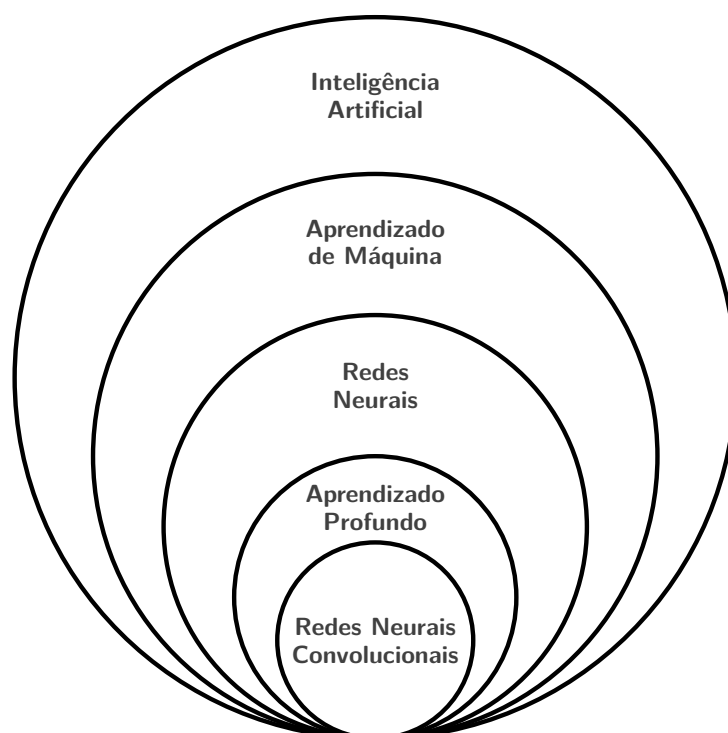
A vantagem do *Deep Learning* é o aprendizado autônomo de características, eliminando a necessidade de engenharia de características manual (*feature engineering*). Assim, não é preciso programar algoritmos matemáticos específicos para extrair propriedades texturais ou formatos celulares, o que, com *Deep Learning*, ocorre automaticamente durante o treinamento da rede (GOODFELLOW et al. 2016).

2.3.6 Redes Neurais Convolucionais (CNNs)

Essa capacidade de autoextração de características abstratas tornou o *Deep Learning* a abordagem ideal para lidar com a alta complexidade de imagens médicas e biológicas. Porém, o empilhamento massivo de camadas densas (*fully connected*) apresenta um desafio computacional: o crescimento exponencial do número de parâmetros (pesos) a serem treinados. Para contornar essa limitação no processamento de matrizes de pixels, surgiram arquiteturas especializadas que aplicam operações matemáticas de filtragem local espacial, conhecidas como Redes Neurais Convolucionais (como representado na hierarquia na Figura 3). Além disso, as CNNs permitem explorar a vizinhança bidimensional em imagens, o que não é possível apenas com camadas densas. (ALOYSIUS et al. 2017).

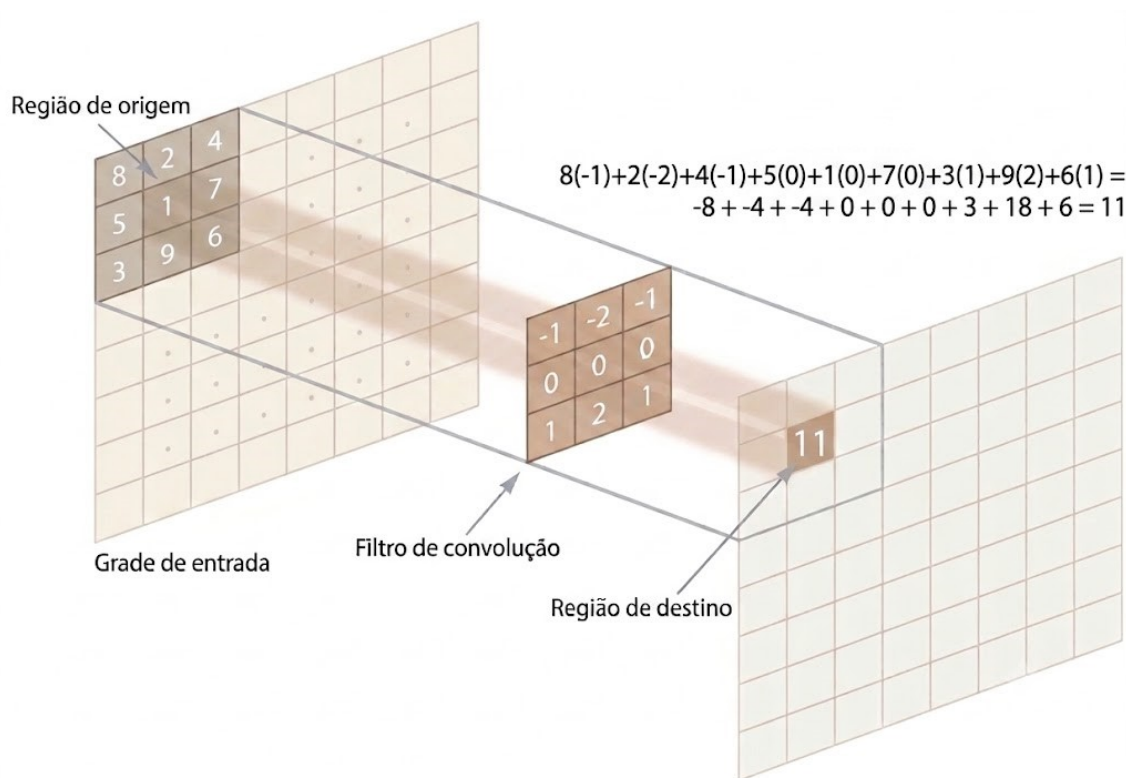
Diferente das redes neurais tradicionais, as Redes Neurais Convolucionais (*Convolutional Neural Networks* - CNNs) utilizam a operação matemática de convolução, na qual pequenos filtros de pesos (matrizes de tamanho reduzido, como 3×3) deslizam por toda a extensão da imagem, como exemplificado na Figura 4). Isso permite que uma célula seja reconhecida independentemente de sua posição na foto e reduz o número de parâmetros necessários para o treinamento, reutilizando os mesmos filtros em toda a extensão da matriz (SIMONYAN et al. 2014).

Figura 3 – Relação hierárquica entre Inteligência Artificial, Aprendizado de Máquina, Redes Neurais, Aprendizado Profundo e Redes Neurais Convolucionais.



Fonte: Elaborado pela autora (2026).

Figura 4 – Representação de exemplo de convolução.



Fonte: Adaptado de KALEJAHİ et al. 2019.

2.3.7 Dados para Treinamento de Modelos

O ciclo de desenvolvimento de um modelo preditivo apoia-se em três pilares fundamentais de dados (GOODFELLOW et al. 2016):

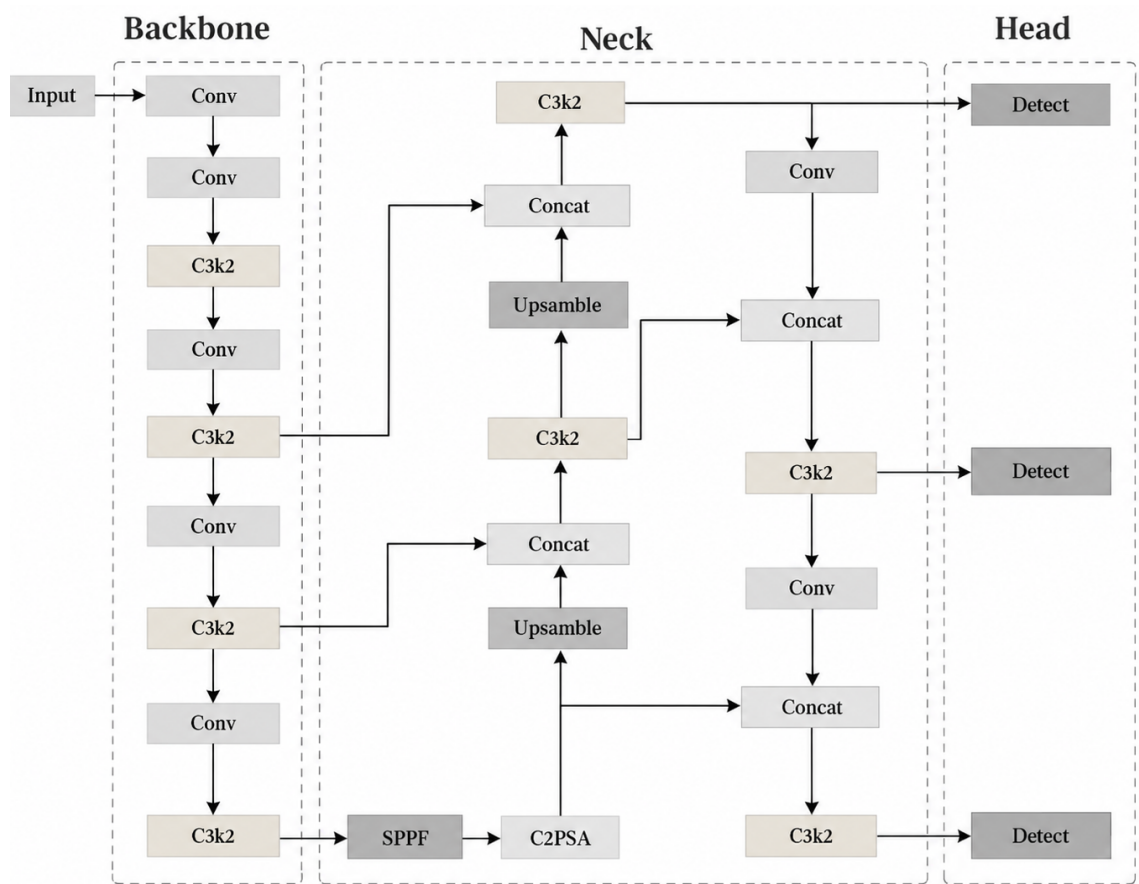
- **Base de dados:** O conjunto de dados bruto coletado. Em contextos médicos, a qualidade e representatividade da base de dados (incluindo variações de pacientes e processos de geração das imagens) influenciam na capacidade do modelo.
- **Anotação (*Ground Truth*):** É o mapeamento manual realizado por especialistas, indicando onde cada célula está localizada. Serve como a resposta alvo utilizada no cálculo de erro fazendo a diferença com a saída da rede.
- **Aumento de Dados (*Data Augmentation*):** Conjunto de transformações geométricas (rotações, espelhamentos) e cromáticas (ajustes de brilho e contraste) aplicadas de forma estocástica durante o treinamento. Essa técnica multiplica artificialmente a variedade de contextos visualizados pela rede, atuando como um potente regularizador contra o *overfitting* (quando o modelo passa a "decorar" os dados do treinamento em vez de aprender padrões gerais, resultando em um desempenho ruim com dados novos).

2.3.8 YOLO11

A arquitetura YOLO (*You Only Look Once*), em sua décima primeira versão (YOLO11), representa o estado da arte em modelos de detecção de objetos em tempo real. Diferentemente de modelos históricos em dois estágios, que primeiro geravam propostas de regiões para depois classificá-las, o YOLO aborda a detecção como um problema de regressão único (REDMON et al. 2016).

A imagem de entrada é processada de ponta a ponta em uma única passagem pela rede, prevendo simultaneamente as coordenadas de todas as caixas delimitadoras e suas respectivas probabilidades de classe. Em sua variante nano (YOLO11n), a arquitetura representada na Figura 5 incorpora blocos otimizados de atenção espacial (C2PSA) e extratores densos de características (C3K2), permitindo que o modelo aprenda detalhes morfológicos finos de objetos pequenos (como leucócitos), sendo computacionalmente extremamente leve, ideal para ambientes de hardware restritos. As etapas da arquitetura (JOCHER et al. 2024) estão descritas a seguir.

Figura 5– Arquitetura do YOLO11.



Fonte: Adaptado de CHAMAN et al. 2025.

O **Backbone** extrai as características principais convolucionais. Os blocos **Conv** (Convolucionais) aplicam filtros matemáticos locais para reduzir a resolução espacial da imagem enquanto aumentam a profundidade de canais, capturando traços básicos, como linhas, bordas e texturas celulares. Já o bloco **C3k2** divide o fluxo em dois caminhos, processa um deles com convoluções mais leves e guarda as informações originais da imagem naquele momento no outro. Isso permite que a rede se torne mais profunda e capture padrões morfológicos complexos com menos custo computacional.

O bloco **C2PSA** injeta um mecanismo de atenção espacial e de canal que força a rede a “prestar mais atenção” nas regiões geometricamente relevantes da imagem e ignorar o ruído de fundo. O **SPPF** (*Spatial Pyramid Pooling - Fast*) recebe o mapa de características e aplica operações de *MaxPooling* (que extraem apenas os valores máximos de cada região filtrada, retendo os recursos visuais mais marcantes e reduzindo o tamanho da matriz) em múltiplas escalas, de forma sequencial, consolidando os recursos visuais em um vetor de tamanho fixo, o que garante que o modelo consiga detectar objetos independentemente do tamanho que ocupam na imagem.

Os blocos **Upsample** redimensionam os mapas de características abstratas para que possam ser combinados com mapas de camadas anteriores na etapa de **Concat**, fundindo as formas originais preservadas por meio das conexões de salto (*skip connections*) com os detalhes refinados capturados.

Na etapa **Head**, esses mapas de características que foram misturados e enriquecidos pelo **Neck** são convertidos em previsões concretas, gerando as coordenadas das caixas delimitadoras (*bounding boxes*) e as classes dos objetos. Essas previsões são realizadas a partir de blocos com resoluções diferentes, permitindo identificar desde objetos menores na imagem com maior resolução, até os objetos maiores na imagem com menor resolução.

2.4 MÉTRICAS DE AVALIAÇÃO

A avaliação do desempenho de modelos de aprendizado profundo aplicados à detecção de objetos baseia-se na quantificação das previsões em relação a um gabarito anotado por especialistas (*ground truth*). No contexto da identificação de eosinófilos em imagens histológicas, as previsões geradas pela arquitetura YOLO11n são categorizadas a partir de uma matriz de confusão adaptada para a tarefa de localização espacial, gerando quatro componentes fundamentais: Verdadeiro Positivo, Falso Positivo, Verdadeiro Negativo e Falso Negativo, como apresentado na Figura 6.

Figura 6 – Matriz de confusão para geração das métricas.

		Predição	
		Positivo	Negativo
Real	Positivo	Verdadeiro Positivo (TP)	Falso Negativo (FN)
	Negativo	Falso Positivo (FP)	Verdadeiro Negativo (TN)

Fonte: Elaborado pela autora (2026).

- **Verdadeiro Positivo (VP):** O modelo detecta a presença de um eosinófilo e desenha a caixa delimitadora (*bounding box*) na localização correta, sobrepondo-se ao gabarito de acordo com um limiar preestabelecido.
- **Verdadeiro Negativo (VN):** O modelo classifica corretamente as demais células detectadas como "não eosinófilos".
- **Falso Positivo (FP):** O modelo marca um objeto onde não há eosinófilo, segundo o gabarito (alarme falso), ou erra severamente na localização geométrica da célula.

- **Falso Negativo (FN):** O modelo omite um eosinófilo legítimo presente na lâmina histológica, falhando em gerar a marcação correspondente (omissão).

A partir desses componentes estruturais, derivam-se as métricas estatísticas utilizadas para a validação do *pipeline* desenvolvido, conforme detalhado a seguir.

2.4.1 Precisão (*Precision*)

A Precisão quantifica a confiabilidade das predições positivas do modelo. Ela determina a proporção de eosinófilos corretamente identificados em relação ao total de marcações que o algoritmo classificou como daquela classe. A expressão matemática que a define é dada pela equação:

$$\text{Precisão} = \frac{VP}{VP + FP}. \quad (2.2)$$

O valor da Precisão varia estritamente no intervalo $[0, 1]$. Quanto maior o seu valor (mais próximo de 1,0), menor é a taxa de falsos alarmes gerados.

2.4.2 Sensibilidade ou Revocação (*Recall*)

A sensibilidade, frequentemente denominada *recall* na literatura de aprendizado de máquina, mede a capacidade do modelo de detectar todos os alvos reais contidos na imagem. Ela expressa a proporção de eosinófilos identificados pelo sistema em relação à quantidade total de células que de fato existiam no gabarito manual, conforme a equação:

$$\text{Recall} = \frac{VP}{VP + FN}. \quad (2.3)$$

Assim como a precisão, o *recall* assume valores entre 0 e 1. Quanto maior o valor obtido, menor é a quantidade de células que o modelo deixou passar despercebidas.

2.4.3 Acurácia (*Accuracy*)

A acurácia representa a fração global de acertos do classificador, ponderando a soma de todas as predições corretas sobre o universo total de amostras avaliadas, conforme expresso na equação:

$$\text{Acurácia} = \frac{VP + VN}{VP + VN + FP + FN}. \quad (2.4)$$

2.4.4 *F1-Score*

Em problemas práticos de visão computacional, há um dilema de compensação (*trade-off*) entre a precisão e o *recall*: algoritmos muito rigorosos tendem a exibir alta

precisão, mas omitir muitas células (*recall* baixo), ao passo que algoritmos permissivos capturam todas as células (*recall* alto), mas geram muitos alarmes falsos (precisão baixa). O *F1-Score* resolve esse impasse ao calcular a média ponderada entre ambas as métricas, penalizando grandes assimetrias, conforme denotado na equação:

$$\text{F1-Score} = 2 \times \frac{\text{Precisão} \times \text{Recall}}{\text{Precisão} + \text{Recall}}. \quad (2.5)$$

O *F1-Score* é limitado entre $[0, 1]$. Valores elevados indicam um bom equilíbrio entre a capacidade de assertividade e a sensibilidade de detecção da rede.

2.4.5 Média da Precisão Média (mAP - *Mean Average Precision*)

A métrica padrão para a validação de algoritmos da família YOLO é o mAP. Para compreendê-la, é necessário introduzir o conceito de Interseção sobre União (*IoU*), que mede o alinhamento geométrico entre a caixa prevista pelo modelo (B_{pred}) e a caixa do gabarito (B_{gt}):

$$\text{IoU} = \frac{\text{Área}(B_{\text{pred}} \cap B_{\text{gt}})}{\text{Área}(B_{\text{pred}} \cup B_{\text{gt}})}. \quad (2.6)$$

Ao alterar o limiar de confiança do modelo, obtêm-se diferentes pares de valores de precisão (P) e *recall* (R), que definem a chamada Curva Precisão-Recall. A Precisão Média (AP) é calculada computando-se a área sob essa curva integrada, conforme a equação:

$$\text{AP} = \int_0^1 P(R) dR. \quad (2.7)$$

Como o mAP representa a média aritmética do AP calculado para todas as classes do *dataset*, e o presente trabalho foca na classe única que avalia a capacidade do modelo de localizar eosinófilos, o valor de AP equivale diretamente ao mAP.

A notação mAP@0,5 estabelece que uma predição é considerada um Verdadeiro Positivo se o cálculo do IoU for maior ou igual a 0,5 (50%). Trata-se de uma métrica de validação tolerante quanto à exatidão do contorno da caixa, que avalia prioritariamente a capacidade do modelo de localizar a localização macro da célula na lâmina.

Considerada a métrica mais rigorosa e realista para a detecção de objetos da visão computacional moderna, o mAP@0,5–0,95 calcula a média do mAP obtido em 10 limiares sucessivos e incrementais de IoU (variando de 0,50 a 0,95 com passos de 0,05) (LIN et al. 2014), como demonstrado na equação:

$$\text{mAP@0,5-0,95} = \frac{1}{10} \sum_{\text{limiar}=0,50}^{0,95} \text{mAP}_{\text{limiar}}. \quad (2.8)$$

Para pontuar satisfatoriamente nos limiares mais altos (e.g., $\text{IoU} = 0,95$), o modelo precisa desenhar uma caixa limitadora praticamente idêntica à definida manualmente pelo especialista humano. Valores acima de 0,50 (50%) na literatura de detecção de objetos são historicamente classificados como desempenhos excelentes, dadas as dificuldades de ajustar contornos perfeitamente quadrados ao redor de estruturas celulares arredondadas e dinâmicas (BOCHKOVSKIY et al. 2020).

2.5 DESAFIOS DA ANÁLISE COMPUTACIONAL DE IMAGENS HISTOLÓGICAS

O processamento de tecidos biológicos é impactado por alguns desafios principais:

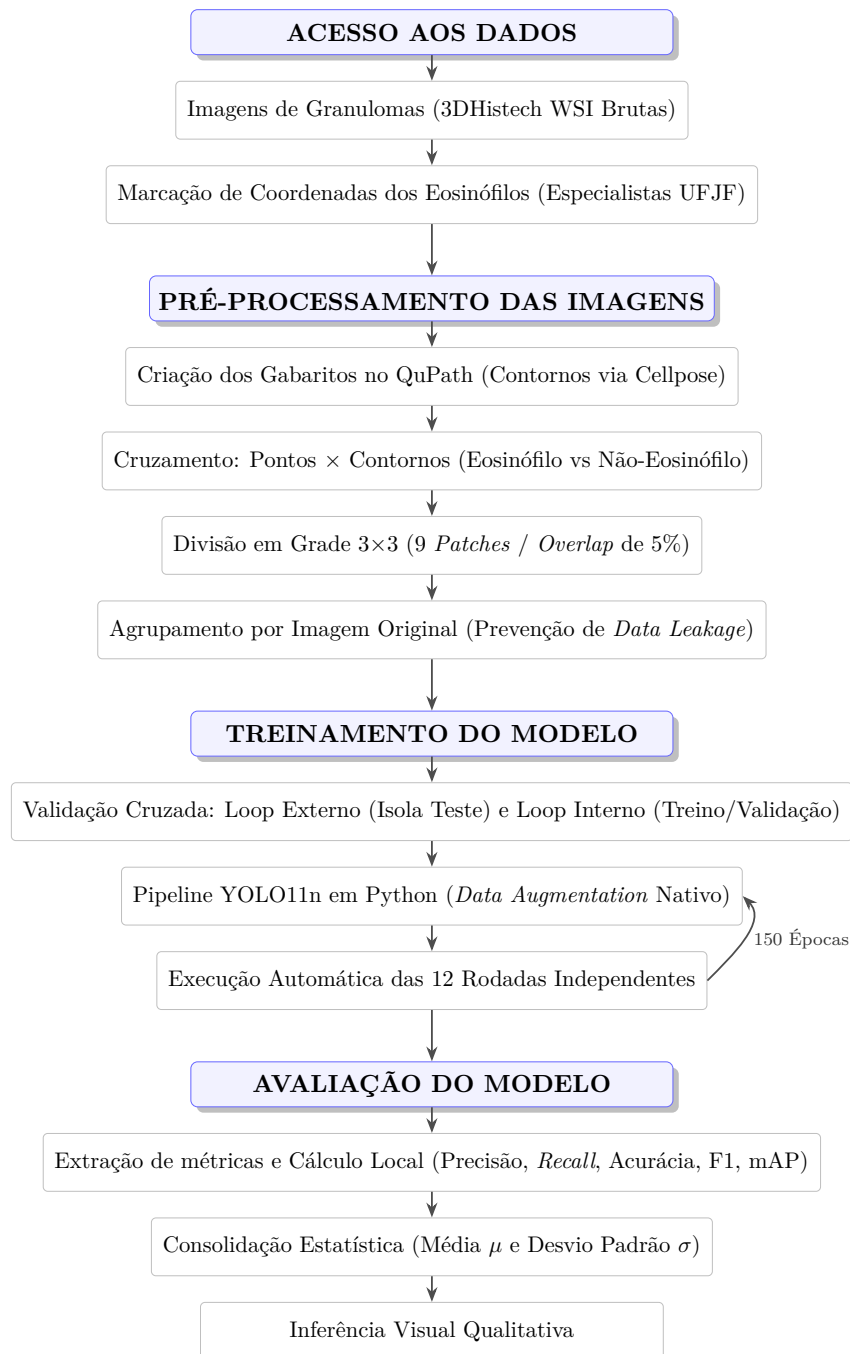
- **Variação de Coloração:** O processo manual ou automatizado de coloração das lâminas introduz variações de tonalidade entre laboratórios ou equipamentos. Dois tecidos idênticos podem apresentar variações de tons, o que exige que o modelo seja flexível para não ignorar células devido a essas variações (MELO et al. 2020).
- **Aglomeração Celular:** Em focos inflamatórios intensos, os eosinófilos costumam ficar muito próximos ou sobrepostos. Para o algoritmo, separar e desenhar caixas individuais ao redor de células coladas é complexo, pois os limites visuais das bordas se perdem no agrupamento (TURNER et al. 2022).
- **Complexidade do Fundo e *Overfitting*:** O fundo das imagens de microscopia é ruidoso, contendo fragmentos de colágeno, hemácias e artefatos que mimetizam a textura das células alvo. Como a área de fundo é imensamente maior do que o número de eosinófilos, há um desbalanceamento severo dos dados (ADORNO et al. 2021). Isso pode levar a um cenário em que o modelo decore o cenário de fundo ou ruídos específicos do *dataset* de treino, em vez de aprender as características gerais da célula (XIE et al. 2015), falhando ao ver novas imagens e aumentando o risco de *overfitting*.
- **Dimensões Elevadas e Resolução:** Imagens de biópsias possuem resoluções elevadas e podem conter bilhões de *pixels*. Reduzi-las diretamente ao tamanho padrão de entrada das redes degradaria os eosinófilos, transformando-os em pequenos pontos sem detalhes morfológicos (KOMURA et al. 2018). Isso obriga o uso de técnicas de corte em partes menores (*patches*), limitadas pela memória de vídeo (VRAM) disponível (LITJENS et al. 2017).
- **Subjetividade e Fadiga Humana:** A marcação do gabarito (*ground truth*) depende do trabalho visual exaustivo de especialistas. A fadiga humana causa omissões involuntárias de células no gabarito original. Consequentemente, a mesma imagem analisada por diferentes especialistas apresenta marcações divergentes. Houve um estudo no qual vários radiologistas anotaram os mesmos dados, e a discordância

entre eles foi três vezes maior do que a concordância plena (LITJENS et al. 2017). Essas falhas humanas confundem o treinamento do modelo e penalizam injustamente as métricas matemáticas, pois o algoritmo pode detectar uma célula corretamente e ser pontuado como erro por ela não constar na anotação original.

3 MÉTODOS

O trabalho está estruturado em quatro etapas principais: a coleta e curadoria dos dados histológicos, o pré-processamento das imagens, o treinamento de modelos e a avaliação de desempenho baseada em métricas de visão computacional, conforme sintetizado na Figura 7.

Figura 7 – Fluxograma do *pipeline* metodológico proposto.



Fonte: Elaborado pela autora (2026).

3.1 ACESSO AOS DADOS

Os dados de treinamento e validação do modelo foram fornecidos pelo Laboratório de Biologia Celular da Universidade Federal de Juiz de Fora, grupo que realiza pesquisas sobre os mecanismos celulares relacionados à inflamação. Foi disponibilizada uma base de imagens de granulomas extraídos de imagens de lâminas inteiras (WSI) digitalizadas usando um *scanner* de lâminas digitais 3DHistech Panoramic Scan.

A demarcação e rotulação dos eosinófilos foram realizadas por pesquisadores do laboratório, biólogos, o que garantiu dados confiáveis e maior precisão no treinamento e validação do modelo. As imagens foram analisadas pelos pesquisadores, que marcaram um ponto em cada eosinófilo identificado, fornecendo para o trabalho as coordenadas dos pontos em que existiam eosinófilos em cada imagem.

3.2 PRÉ-PROCESSAMENTO DAS IMAGENS

3.2.1 Criação dos Gabaritos

Para o treinamento do modelo, é importante que as demarcações do “gabarito” incluam o contorno de cada célula. Porém, nas imagens de granulomas disponibilizadas, os eosinófilos estavam marcados apenas com pontos. Por isso, foi necessária uma etapa inicial para identificar e delimitar o contorno dessas células. Para isso, foi utilizado o QuPath, software voltado para análises de bioimagem e que aceita extensões e *scripts*. Utilizando uma extensão do *cellpose* no QuPath, um modelo pré-treinado de *deep learning*, foi realizada a marcação dos contornos de todas as células presentes nas imagens. Cruzando as marcações das células com os pontos que identificavam os eosinófilos, foi possível classificar as células como “eosinófilo” ou “não eosinófilo”.

3.2.2 Divisão em Patches

Inicialmente, foi avaliada a necessidade de dividir a imagem em partes menores (*patches*), o que é muito comum em trabalhos que envolvem WSIs devido às suas grandes dimensões, que podem conter bilhões de *pixels*, exigindo um grande poder computacional para o processamento dessas imagens (KOMURA et al. 2018). Como as imagens utilizadas neste trabalho não são de lâminas inteiras, mas apenas das estruturas de granuloma retiradas desses WSIs, as dimensões permitem trabalhar com o granuloma inteiro, sem que a imagem seja dividida. Apesar disso, observou-se que ainda haveria ganhos com a divisão da imagem em *patches*, o que levou à segmentação de cada matriz em uma grade de 3×3 , com sobreposição (*overlapping*) de 5% entre as bordas dos recortes adjacentes, para evitar a perda de objetos que seriam recortados na divisão, garantindo que células localizadas exatamente na fronteira entre dois *patches* não sejam perdidas. Essa estratégia trouxe impactos positivos, como:

- **Preservação dos Detalhes Celulares:** O redimensionamento direto da imagem do granuloma inteiro para o padrão de entrada nativo da arquitetura YOLO11 (640×640 *pixels*) resultaria na redução da matriz original. Esse processo reduziria os eosinófilos, que já são células pequenas, eliminando suas características morfológicas citoplasmáticas. O fracionamento em *patches* permitiu que cada quadrante mantivesse sua dimensão original, preservando a textura e os contornos necessários.
- **Otimização da Memória de Vídeo (VRAM):** Ao dividir o granuloma em subimagens menores, reduziu-se o consumo de VRAM por operação. Essa estratégia permitiu a adoção de um tamanho de lote (*batch size*) igual a 4, que mostrou-se o equilíbrio ideal para este trabalho. Ele garantiu que o treinamento permanecesse leve e estável dentro dos limites físicos de memória da GPU utilizada, sem comprometer a convergência do modelo e sem erros de estouro de memória (*out of memory*).
- **Prevenção de *Overfitting* e Variedade dos Lotes (*Mini-batches*):** Durante o treinamento, os parâmetros do modelo são atualizados em pequenas parcelas de dados a cada lote. Ao segmentar a lâmina em múltiplos *patches*, o algoritmo é exposto, separadamente, a diferentes cenários do granuloma, o que ajuda a regular a rede. Em vez de o modelo memorizar o arranjo global e o fundo de uma única imagem inteira, o que poderia causar *overfitting*, ele é forçado a aprender características visuais específicas e locais dos eosinófilos, aumentando sua capacidade de funcionar em novas imagens.

Como as imagens utilizadas no treinamento não possuíam grande variação de dimensões, optou-se por dividir cada imagem em 9 *patches* ao invés de dividir em *patches* de tamanho padrão pois quando uma imagem era poucos *pixels* maior que o tamanho padrão eram criados 4 *patches* praticamente iguais da imagem inteira (com poucos *pixels* de diferença entre cada *patch*, de forma que todas as divisões pareciam iguais). Dividir a imagem em 9 partes, independentemente de suas dimensões, garante que cada corte seja único. Apenas foi possível utilizar essa técnica pois as imagens estavam dentro de dimensões bem controladas. Caso fossem utilizadas WSIs completas, a divisão em 9 não seria suficiente para evitar problemas de processamento e seria necessário aplicar um tamanho de *patch* padrão em uma janela deslizante.

Para evitar *data leakage*, todos os 9 *patches* de uma mesma imagem original sempre ficam juntos no mesmo conjunto (treino, validação ou teste), nunca são divididos. Se *patches* da mesma imagem estivessem em conjuntos diferentes, o modelo seria exposto a cortes que compartilham exatamente as mesmas condições de iluminação, ruído de fundo, espessura do corte histológico e padrão de coloração da lâmina original. Durante a fase de otimização, a rede convolucional poderia facilmente memorizar essas características globais do plano de fundo em vez de aprender as propriedades morfológicas intrínsecas dos

eosinófilos. Como consequência, ao avaliar o desempenho nos conjuntos de validação ou de teste, o modelo apresentaria métricas de acurácia e precisão artificialmente elevadas, decorrentes da capacidade de reconhecer o cenário de fundo previamente visto durante o treinamento.

3.2.3 *Data Augmentation*

A aplicação de técnicas de aumento de dados (*Data Augmentation*) desempenha um papel crítico na mitigação do *overfitting* em cenários com escassez de dados (SHORTEN et al. 2019). No escopo deste trabalho, cujo conjunto amostral limita-se a 12 imagens originais e, após o fracionamento, a um conjunto de 108 *patches* de análise, a ampliação da diversidade visual foi indispensável para assegurar a convergência estável dos parâmetros dos blocos extratores de características C3K2 e C2PSA.

Não foi necessária a implementação manual de rotinas de transformação geométrica ou cromática no *pipeline* de pré-processamento. A biblioteca *Ultralytics*, que gerencia o ecossistema do YOLO11, executa, de forma nativa, um conjunto robusto e automatizado de técnicas de aumento de dados em memória volátil a cada época de treinamento. Durante o fluxo de otimização, o modelo foi exposto dinamicamente a algoritmos avançados, como o *Mosaic* (que funde quatro subimagens aleatórias em uma composição contextual única) e o *Mixup* (que sobrepõe imagens com pesos probabilísticos), além de rotações, translações, espelhamentos horizontais e verticais e distorções aleatórias no espaço de cor HSV (matiz, saturação e brilho). Essas transformações simularam as variações de iluminação e de calibração de cor típicas de diferentes equipamentos laboratoriais, enriquecendo os lotes (*mini-batches*) com exposição a diferentes condições, mesmo com o número reduzido de imagens iniciais.

3.3 TREINAMENTO DO MODELO

3.3.1 Validação Cruzada

Utilizou-se validação cruzada para avaliar as métricas obtidas em diferentes cenários de divisão dos dados nos grupos de treinamento, validação e teste. Assim, foi possível avaliar a estabilidade do modelo sem viés de amostragem, já que uma distribuição específica dos dados entre esses conjuntos, seja por manipulação ou sorte, poderia gerar um resultado muito positivo para o caso testado que não corresponderia à realidade quando o modelo fosse utilizado para novas imagens.

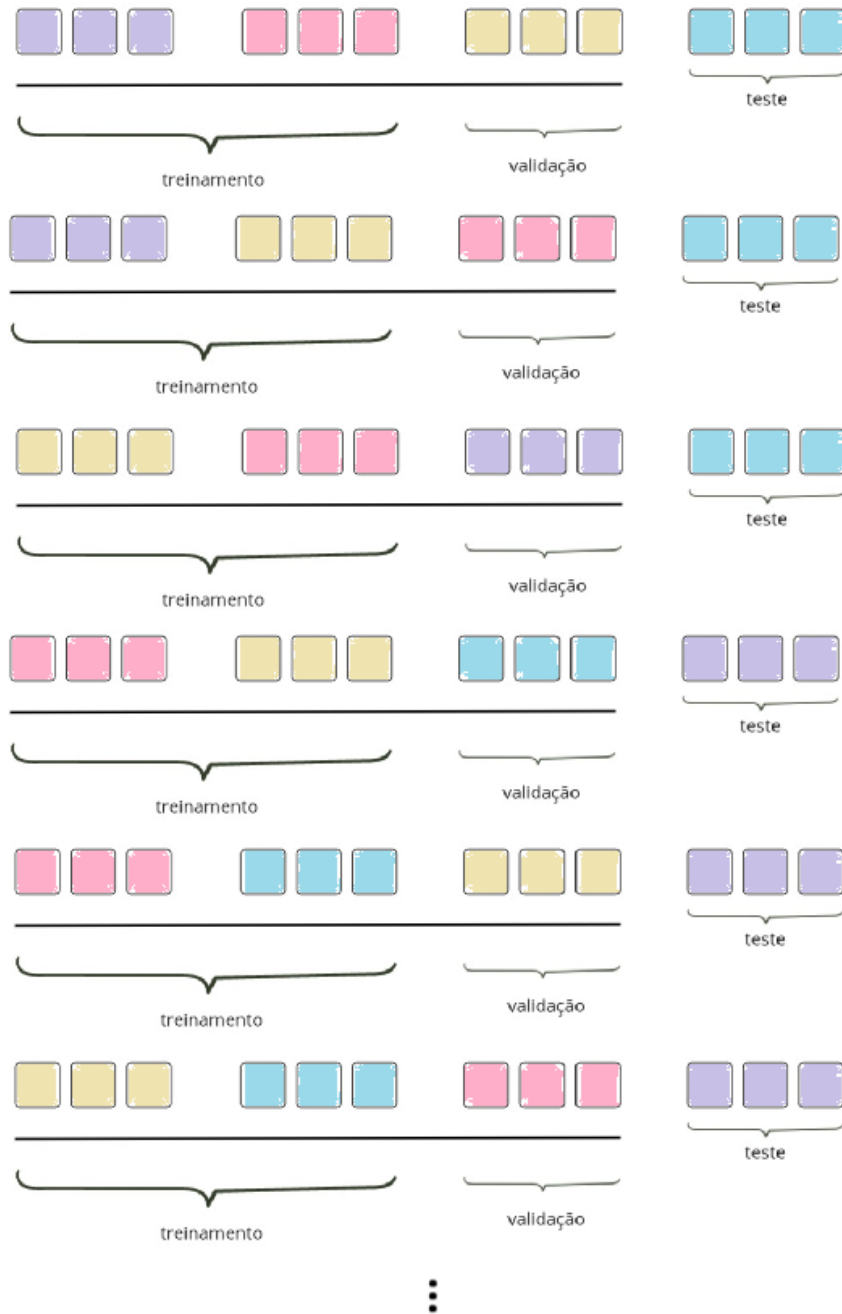
Por isso, as 12 imagens foram divididas em quatro grupos de 3 imagens (divididas em seus respectivos *patches*) e combinadas de forma a garantir que todas as imagens sirvam como dados de treino, validação e teste em momentos distintos sem que haja sobreposição entre os três conjuntos de um mesmo treinamento, como apresentado na Figura 8, em que

cada cor representa um dos quatro grupos. Em cada rodada, um grupo foi mantido como conjunto de teste separado, um grupo atuou como conjunto de validação e os dois grupos restantes foram utilizados no treinamento.

A separação de grupos foi feita entre as imagens inteiras, e não entre os *patches*, para evitar o vazamento de dados. Por exemplo, se *patches*) da mesma imagem estivessem tanto no grupo de treinamento quanto no de teste, o modelo já conheceria características daquela imagem, como cor, contraste e iluminação. Assim, mesmo sendo partes diferente da imagem, ainda teriam características semelhantes, que poderiam gerar um resultado de teste enviesado. Por isso, todos os *patches*) de uma imagem estão sempre dentro do mesmo conjunto.

Cada um dos 12 treinamentos foi realizado separadamente, cada um resultando em suas próprias métricas de avaliação. Ao comparar a média dessas métricas e o desvio padrão ao longo das 12 execuções é possível avaliar a estabilidade do modelo utilizado e se os resultados são coerentes nas diferentes combinações dos grupos, descartando a possibilidade de um resultado enviesado por uma combinação específica.

Figura 8– Divisão de Grupos para Validação Cruzada.



Fonte: Elaborado pela autora (2026).

3.3.2 Desenvolvimento e Aplicação do *Pipeline*

Para realizar a detecção e a contagem automáticas de eosinófilos, foi desenvolvido um *pipeline* em Python, utilizando a biblioteca *Ultralytics*. O fluxo principal do código foi concentrado no script `src/yolo/train.py`, responsável por gerenciar, de forma automati-

zada, o treinamento, a validação e a extração de métricas ao longo de todo o processo de validação cruzada.

O funcionamento do sistema inicia-se com a função `_get_device`, que verifica se há uma placa de vídeo (GPU) disponível no computador. Caso uma interface CUDA seja encontrada, o sistema direciona o processamento para a GPU e registra as especificações do *hardware* nos arquivos de log; caso contrário, o treinamento é executado pela CPU.

O treinamento é controlado por um *script* que funciona como um laço de repetição principal que gerencia as 12 combinações independentes criadas pelas divisões dos dados. Para cada uma das 12 rodadas, o *pipeline* executa os seguintes passos:

1. **Carga do Modelo:** Inicializa a arquitetura YOLO11 na versão nano (`yolo11n.pt`). Esse modelo mais leve foi escolhido por oferecer processamento rápido e eficiente, ideal para conjuntos de dados menores e placas de vídeo com menos memória disponível (VRAM).
2. **Treinamento:** Executa a função `model.train()` utilizando o arquivo de configuração `data_fold_0_I.yaml` da rodada atual. O treinamento foi configurado com imagens de entrada no tamanho 640×640 *pixels*, lotes (*batch size*) de 4 *patches* e um limite de 150 épocas por combinação.
3. **Extração de Métricas:** Após o fim das épocas, a função `model.val()` avalia o melhor modelo gerado no conjunto de validação daquela rodada. Os resultados alimentam a rotina `compute_metrics_from_val`, que isola a classe de eosinófilos diretamente a partir da matriz de confusão do YOLO. Os Verdadeiros Positivos (VP) são retirados da diagonal principal da matriz, enquanto os Falsos Positivos (FP) e os Falsos Negativos (FN) são calculados a partir das somas das linhas e das colunas. Com base nesses valores, o script calcula a Precisão, o *Recall*, a Acurácia e o *F1-Score*. As métricas de $mAP@0,5$ e $mAP@0,5-0,95$ são obtidas diretamente do relatório do YOLO.
4. **Salvamento dos Resultados:** A função `save_eosinofilo_metrics` exporta os dados de desempenho de cada rodada em formato JSON para análises futuras confiáveis e também em formato de texto TXT para leitura humana.

Após o término de todas as 12 execuções, a rotina `save_consolidated_metrics` combina os resultados do trabalho, calculando a média aritmética e o desvio padrão de cada métrica. Esse processo garante uma avaliação estatística confiável, eliminando distorções que poderiam ocorrer caso as imagens fossem divididas aleatoriamente apenas uma vez.

especificar parametros do YOLO que foram alterados, epocas etc Por fim, o módulo de inferência visual (`src/yolo/predict.py`) realiza uma validação qualitativa do modelo.

Diferente do treinamento, a predição é aplicada sobre novas imagens histológicas originais e inteiras (sem divisão em *patches*) do grupo de teste, utilizando uma estratégia de janela deslizante (*sliding-window*) combinada com um algoritmo de Supressão Não-Máxima (NMS) global para eliminar marcações duplicadas. O script desenha caixas delimitadoras vermelhas ao redor dos eosinófilos encontrados pelo YOLO11n e plota círculos amarelos no centro das células do gabarito original (*ground truth*), lidos da pasta `data/masks-yolo/`. O resultado visual é salvo no diretório `predict_<run_name>`, permitindo que os especialistas comparem diretamente o acerto do modelo às anotações manuais.

3.4 AVALIAÇÃO DO MODELO

A extração das métricas foi realizada pelo script `src/yolo/train.py` por meio da função `compute_metrics_from_val`. O fluxo de processamento dos dados de validação seguiu os seguintes critérios de engenharia:

1. **Cálculo dos Parâmetros Locais:** A partir da matriz isolada, o algoritmo extraiu os Verdadeiros Positivos (VP), Falsos Positivos (FP) e Falsos Negativos (FN). Com base nesses parâmetros, o *script* calculou localmente os índices de Precisão, *Recall* (Sensibilidade), Acurácia global e o *F1-Score*. Esse procedimento permitiu avaliar o equilíbrio do modelo, minimizando a omissão de células legítimas e a geração de alarmes falsos em tecidos histológicos complexos.
2. **Recuperação dos Índices de Precisão Média:** Os valores de $mAP@0,5$ e $mAP@0,5-0,95$ foram recuperados diretamente a partir do objeto de saída de validação do YOLO11n. Essas métricas foram adotadas como critérios de avaliação geométrica, medindo a exatidão com que o modelo ajustou as caixas delimitadoras ao redor das células em diferentes níveis de exigência de sobreposição (*IoU*).

Para mitigar a variabilidade estatística introduzida por qualquer divisão arbitrária ou afortunada dos dados, os resultados das execuções foram processados pela rotina `save_consolidated_metrics`. O algoritmo agregou as tabelas JSON individuais e calculou a média aritmética (μ) e o desvio padrão (σ) para cada uma das métricas avaliadas.

Essa abordagem de consolidação estatística cumpre um papel metodológico duplo: a média fornece uma estimativa realista e robusta do desempenho esperado do modelo em ambiente de produção biomédica, enquanto o desvio padrão quantifica a estabilidade e a sensibilidade da arquitetura frente a diferentes fatias de dados inéditos. O *pipeline* foi considerado validado apenas após a verificação de que a dispersão dos resultados se mantinha dentro de limites estatisticamente estreitos, assegurando a confiabilidade do sistema para auditoria clínica.

4 RESULTADOS E DISCUSSÃO

O desempenho do modelo foi avaliado por meio de métricas estatísticas robustas e consolidadas por validação cruzada, mitigando vieses de partição de dados e garantindo a capacidade de generalização do algoritmo.

A Figura 9 apresenta um exemplo de resultado das predições do modelo treinado na imagem de um granuloma, em que o centro é um agente invasor isolado pelas células em sua volta, incluindo os eosinófilos. Os quadrados vermelhos representam as predições realizadas, enquanto os pontos amarelos correspondem aos gabaritos fornecidos pelos pesquisadores da Universidade Federal de Juiz de Fora (UFJF).

Foi observado pelos pesquisadores do Laboratório de Biologia Celular que em alguns casos, onde houve predição e não há gabarito, na verdade a predição foi legítima apesar de não ter sido marcada inicialmente. Em outros casos, os falsos positivos podem ocorrer pela presença de outras células com morfologia semelhante e núcleos próximos que podem ser interpretados pelo extrator de características como um núcleo bilobado. Também é possível observar regiões em que havia marcação de gabarito e que não foram detectadas pelo modelo (falsos negativos). Alguns casos apresentam o núcleo cortado (como na borda do granuloma), possuem sobreposição de outros núcleos próximos ou não apresentam o formato padrão de seu núcleo bilobado (a morte celular faz com que a membrana do eosinófilo se desintegre e perca o formato clássico de seu núcleo). Portanto, mesmo apresentando resultados satisfatórios, ainda é possível melhorar as predições ao usar um conjunto de dados mais diversos e em diferentes momentos do ciclo de vida do eosinófilo.

4.1 DESEMPENHO DO MODELO NO *PIPELINE*

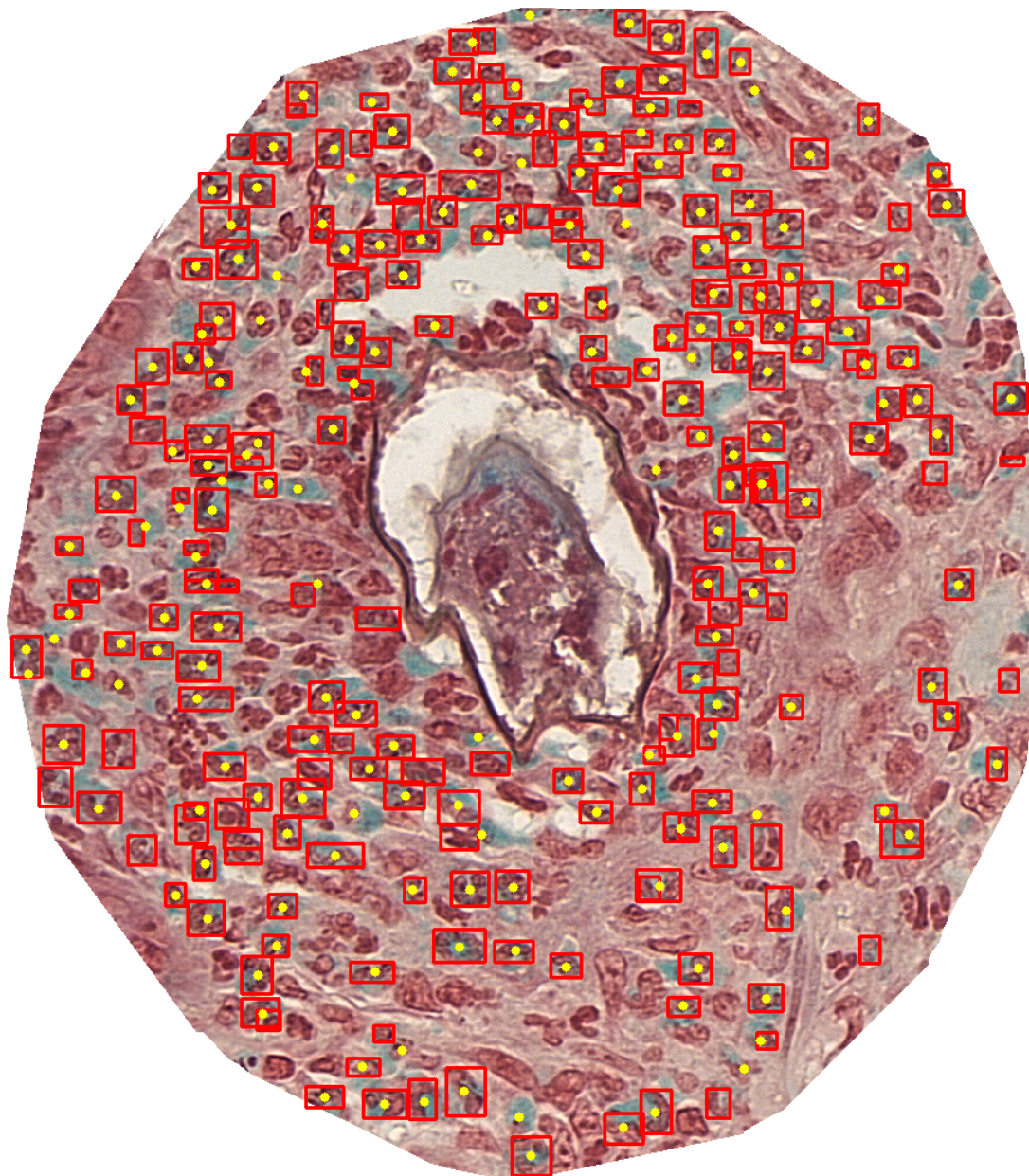
O modelo foi submetido a um arranjo de validação cruzada aninhada com 4 partições no loop externo (outer folds) e 3 no interno (inner folds), totalizando 12 rodadas independentes de treinamento e teste. A Tabela 1 consolida as métricas obtidas para a classe de eosinófilos em cada iteração.

Tabela 1 – Resumo estatístico final consolidado (Média $\pm$ Desvio Padrão).

Métrica	Valor Médio $\pm$ Desvio Padrão
Precisão	0,7839 $\pm$ 0,0236
<i>Recall</i> (Sensibilidade)	0,7598 $\pm$ 0,0157
Acurácia	0,7761 $\pm$ 0,0239
<i>F1-Score</i>	0,7715 $\pm$ 0,0154
mAP@0,5	0,8393 $\pm$ 0,0134
mAP@0,5–0,95	0,5678 $\pm$ 0,0219

Fonte: Elaborado pela autora a partir dos resultados dos testes do modelo (2026).

Figura 9 – Marcações dos eosinófilos.



Fonte: Elaborado pela autora (2026).

O modelo demonstrou alta estabilidade e reprodutibilidade, evidenciadas pelos baixos desvios-padrão em todas as métricas analíticas. O comportamento do $mAP@0,5$ ($0,8393 \pm 0,0134$) indica que a rede obteve excelente desempenho em delimitar a presença das estruturas celulares em um limiar de sobreposição de interseção sobre união acima de 0,5, patamar este considerado robusto para aplicações biológicas onde a morfologia celular pode sofrer variações decorrentes do processo de fixação e coloração das lâminas. O valor médio alcançado pelo $F1-Score$ ($0,7715$) comprova o equilíbrio dinâmico entre a capacidade do modelo de classificar corretamente os eosinófilos sem gerar falsos positivos

em tecidos complexos (caracterizado pela Precisão de 0,7839) e a sua sensibilidade em mapear a maioria dos alvos reais presentes (caracterizado pelo *Recall* de 0,7598).

4.2 VALIDAÇÃO POR ESPECIALISTAS

Ao apresentar os resultados das predições do modelo para validação por pesquisadores especialistas da Universidade Federal de Juiz de Fora (UFJF), verificou-se que parte dos elementos contabilizados estatisticamente como Falsos Positivos (FP) pelo *script* automatizado correspondia, na realidade, a eosinófilos legítimos que haviam sido omitidos durante o processo de anotação manual dos gabaritos.

Este cenário evidencia dois pontos críticos:

1. **Subestimação das Métricas Reais:** O valor nominal de Precisão (78,39%) é artificialmente penalizado pela presença de marcações corretas do modelo que não constavam no gabarito falho. Em um cenário de validação clínica perfeitamente revisada, a taxa de falsos positivos reais decresceria, elevando o valor de precisão analítica.
2. **Capacidade Auxiliar de Diagnóstico:** Os blocos de atenção espacial e de canal do YOLO11 (C2PSA) mostraram-se eficazes em abstrair padrões cromáticos (afinidade por corantes eosinofílicos) e texturas granulares citoplasmáticas consistentes. O modelo conseguiu manter a consistência preditiva mesmo diante da fadiga visual humana ou de variações na atenção do anotador clínico. A validação consensual por parte dos pesquisadores da UFJF confirma que o sistema desenvolvido ultrapassa a função de mero reprodutor de anotações, atuando como uma ferramenta preditiva robusta, capaz de atuar em auditorias de exames histopatológicos e em sistemas automatizados de suporte à decisão clínica.

5 CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS

Este trabalho apresentou o desenvolvimento, a implementação e a validação de um *pipeline* automatizado de visão computacional voltado à detecção de eosinófilos em imagens histológicas de granulomas. Centrado na arquitetura YOLO11, em sua variante nano (YOLO11n), o sistema foi projetado para alto desempenho preditivo, viabilizando a extração de características morfológicas complexas mesmo sob restrições severas de *hardware* e escopo amostral.

Diante de um cenário de alta complexidade analítica e dados limitados, composto por somente 12 imagens de granulomas, a adoção da validação cruzada garantiu que ao longo das 12 execuções cada lâmina histológica atuasse de forma estritamente isolada como conjunto de teste. A consolidação dos resultados, por meio da média e do desvio padrão das métricas de mAP@0.5, Precisão, *Recall* e *F1-Score*, mostrou, matematicamente, que o modelo é estável e funciona bem em imagens inéditas, evitando distorções.

A validação das marcações geradas, realizada por pesquisadores especialistas da UFJF, revelou que o modelo foi capaz de identificar eosinófilos legítimos que haviam sido omitidos no gabarito manual devido à fadiga humana. Esse fenômeno mostra que as predições do modelo podem alcançar uma consistência analítica superior à da anotação humana isolada, consolidando a ferramenta como um mecanismo viável de auditoria.

Adicionalmente, embora este trabalho tenha focado especificamente na contagem de eosinófilos, o modelo demonstrou capacidade de mapear e marcar outros tipos celulares presentes nas lâminas, o que servirá de base para etapas futuras do trabalho.

A continuidade desta pesquisa e as próximas fases do trabalho devem focar na expansão da capacidade preditiva do modelo e na consolidação da ferramenta no ambiente prático de patologia a partir dos seguintes passos:

- **Expansão da Base de Dados:** Incorporar um volume maior de imagens histológicas provenientes de diferentes seres, coletadas e anotadas por pesquisadores distintos. Essa variedade é essencial para testar os limites do modelo e garantir que ele funcione bem diante de novos padrões de lâminas.
- **Exploração do Treinamento Multiclasse:** Utilizar as marcações de outras células já realizadas pelo modelo para expandir o sistema. O objetivo é fazer com que a rede deixe de contar apenas os eosinófilos e passe a classificar e quantificar simultaneamente outros tipos de células presentes nos granulomas.
- **Aprimoramento da Interface de Usuário:** Desenvolver uma interface gráfica intuitiva para o executável gerado pelo PyInstaller. O foco é criar um painel interativo simples para que os pesquisadores da UFJF possam aplicar o modelo treinado a novas imagens e visualizar os resultados no QuPath, sem interações com o terminal.

REFERÊNCIAS

- ADORNO, WILLIAM, ALEXIS CATALANO, LUBAINA EHSAN, HANS VITZHUM VON ECKSTAEDT, BARRETT BARNES, EMILY MCGOWAN, SANA SYED e DONALD E. BROWN (fev. de 2021). “Advancing Eosinophilic Esophagitis Diagnosis and Phenotype Assessment with Deep Learning Computer Vision”. Em: *Biomedical engineering systems and technologies, international joint conference, BIOSTEC ... revised selected papers. BIOSTEC (Conference) 2021.2 Bioimaging*, pp. 44–55. DOI: 10.5220/0010241900002865. URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC8144887/> (acesso em 17/03/2026).
- ALOYSIUS, N e M GEETHA (2017). “A review on deep convolutional neural networks”. Em: *2017 International Conference on Communication and Signal Processing (ICCSP)*. IEEE, pp. 0588–0592.
- BAMWENDA, JULIUS et al. (2025). “A Hybrid Deep Learning Framework for Accurate Cell Segmentation in Whole Slide Images Using YOLOv11, StarDist, and SAM2”. Em: *Bioengineering* 12.6, p. 674.
- BANKMAN, ISAAC (2008). *Handbook of Medical Image Processing and Analysis*. 2nd. Academic Press.
- BARATA, LUCCAS M., KÁSSIA K. MALTA, VITOR H. NEVES, CINTHIA PALAZZI, ELIANE G. OLIVEIRA-BARROS, YASMIN AGUIAR, FELIPE KNEIP, BRUNO A. MARCELINO, LÍVIA A. S. CARMO, JOÃO FELIPE AUDI-GAZETA, PEDRO H. S. SANTOS, MARIA KAROLYNNA B. MILANI, MICHELLE C. A. PAULA, ELIANA C. B. TOSCANO, ROSANA GENTILE, FELIPE F. DIAS, THIAGO P. SILVA e ROSSANA C. N. MELO (jun. de 2025). “A major ecological niche of eosinophils in evolving *Schistosoma* granulomas challenges the eosinophil view as “helminth killer” cells”. Em: *Science Advances* 11.24, eadt2779. DOI: 10.1126/sciadv.adt2779. URL: <https://www.science.org/doi/10.1126/sciadv.adt2779> (acesso em 18/03/2026).
- BOCHKOVSKIY, ALEX, CHIEN-YAO WANG e HONG-YUAN MARK LIAO (2020). “YOLOv4: Optimal Speed and Accuracy of Object Detection”. Em: *arXiv preprint arXiv:2004.10934*.
- CHADHA, GULPREET KAUR, AAKARSH SRIVASTAVA, ABHILASHA SINGH, RITU GUPTA e DEEPANSHI SINGLA (jan. de 2020). “An Automated Method for Counting Red Blood Cells using Image Processing”. Em: *Procedia Computer Science*. International Conference on Computational Intelligence and Data Science 167, pp. 769–778. ISSN: 1877-0509. DOI: 10.1016/j.procs.2020.03.408. URL: <https://www.sciencedirect.com/science/article/pii/S1877050920308747> (acesso em 25/01/2026).

- CHAMAN, MOHAMMED, ANAS EL MALIKI, HAMAD DAHOU e ABDELKADER HADJOUDJA (2025). “Benchmarking YOLO-based deep learning models for real-time object detection in hybrid ADAS and intelligent transportation systems”. Em: *Journal of Real-Time Image Processing* 22.1, pp. 102–115.
- DIMITRIOU, NEOFYTOS, OGNJEN ARANDJELOVIĆ e PETER D. CAIE (2019). “Deep learning for whole slide image analysis: an overview”. Em: *Frontiers in Medicine* 6, p. 264.
- FACELI, KATTI, ANA CAROLINA LORENA, JOÃO GAMA e ANDRÉ C. P. L. F. CARVALHO (2021). *Inteligência Artificial: Uma Abordagem de Aprendizado de Máquina*. 2^a ed. Rio de Janeiro: LTC.
- FARAHANI, NAVID, ANIL V PARWANI e LIRON PANTANOWITZ (2015). “Whole slide imaging in pathology: advantages, limitations, and emerging perspectives”. Em: *Pathology and Laboratory Medicine International* 7, pp. 23–33.
- FU, TINGZHAO, JIANFA ZHANG, RUN SUN, YUYAO HUANG, WEI XU, SIGANG YANG, ZHIHONG ZHU e HONGWEI CHEN (2023). “Optical neural networks: progress and challenges”. Em: *Optical Review* 30.4, pp. 375–392.
- GONZALEZ, RAFAEL C. e RICHARD E. WOODS (2018). *Digital Image Processing*. 4th. New York: Pearson.
- GOODFELLOW, IAN IMPLICITLY, YOSHUA BENGIO e AARON COURVILLE (2016). *Deep Learning*. Disponível em: <http://www.deeplearningbook.org>. Cambridge: MIT Press.
- GUPTA, RAJARSI, TAHSIN KURC, ASHISH SHARMA, JONAS S. ALMEIDA e JOEL H. SALTZ (2020). “Characterizing immune responses in whole slide images of cancer with digital pathology and pathomics”. Em: *Current Pathobiology Reports* 8.4, pp. 133–148.
- GURCAN, METIN N., LAURA E. BOUCHERON, ALI CAN, ANANT MADABHUSHI, NASIR M. RAJPOOT e BÜLENT YENER (2009). “Histopathological image analysis: A review”. Em: *IEEE Reviews in Biomedical Engineering* 2, pp. 147–171.
- HAYKIN, SIMON (2001). *Redes Neurais: Princípios e Prática*. Vol. 2. Porto Alegre: Bookman Editora.
- JANOWCZYK, ANDREW e ANANT MADABHUSHI (jul. de 2016). “Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases”. Em: *Journal of Pathology Informatics* 7, p. 29. ISSN: 2229-5089. DOI: 10.4103/2153-3539.186902. URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC4977982/> (acesso em 25/01/2026).
- JOCHER, GLENN e JING QIU (2024). *Ultralytics YOLO11*. Versão 11.0.0. URL: <https://github.com/ultralytics/ultralytics>.

- KALEJAHÍ, BEHNAM KIANI, SAEED MESHGINI, SABALAN DANESHVAR e SHIVA ASADZADEH (2019). “Diagnosis of liver disease using computer-assisted imaging techniques: A Review”. Em: *arXiv preprint arXiv:1912.09572*. License: CC0.
- KIERNAN, JOHN A. (2015). “Histological and histochemical methods: theory and practice”. Em: *Microscopy and Microanalysis* 21.4, pp. 1023–1025.
- KOMURA, DAISUKE e SHUMPEI ISHIKAWA (jan. de 2018). “Machine Learning Methods for Histopathological Image Analysis”. English. Em: *Computational and Structural Biotechnology Journal* 16, pp. 34–42. ISSN: 2001-0370. DOI: 10.1016/j.csbj.2018.01.001. URL: [https://www.csbj.org/article/S2001-0370\(17\)30086-7/fulltext](https://www.csbj.org/article/S2001-0370(17)30086-7/fulltext) (acesso em 18/03/2026).
- LECUN, YANN, YOSHUA BENGIO e GEOFFREY HINTON (2015). “Deep learning”. Em: *Nature* 521.7553, pp. 436–444.
- LIN, TSUNG-YI, MICHAEL MAIRE, SERGE BELONGIE, JAMES HAYS, PIETRO PERONA, DEVA RAMANAN, PIOTR DOLLÁR e C LAWRENCE ZITNICK (2014). “Microsoft COCO: Common objects in context”. Em: *Computer Vision—ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13*. Springer, pp. 740–755.
- LITJENS, GEERT, THIJS KOOI, BABAK EHTESHAMI BEJNORDI, ARNAUD ARINDRA ADIYOSO SETIO, FRANCESCO CIOMPI, MOHSEN GHAFOORIAN, JEROEN A. W. M. VAN DER LAAK, BRAM VAN GINNEKEN e CLARA I. SÁNCHEZ (2017). “A survey on deep learning in medical image analysis”. Em: *Medical Image Analysis* 42, pp. 60–88.
- MALTA, KÁSSIA K., CINTHIA PALAZZI, VITOR H. NEVES, YASMIN AGUIAR, THIAGO P. SILVA e ROSSANA C. N. MELO (out. de 2022). “Schistosomiasis Mansonii-Recruited Eosinophils: An Overview in the Granuloma Context”. en. Em: *Microorganisms* 10.10, p. 2022. ISSN: 2076-2607. DOI: 10.3390/microorganisms10102022. URL: <https://www.mdpi.com/2076-2607/10/10/2022> (acesso em 18/03/2026).
- MCBRIEN, CLAIRE N. e ANDREW MENZIES-GOW (2017). “The biology of eosinophils and their role in asthma”. Em: *Frontiers in Medicine* 4, p. 93.
- MEDZHITOV, RUSLAN (2008). “Origin and physiological roles of inflammation”. Em: *Nature* 454.7203, pp. 428–435.
- MELO, ROSSANA C. N., MAXIMILIAN W. D. RAAS, CINTHIA PALAZZI, VITOR H. NEVES, KÁSSIA K. MALTA e THIAGO P. SILVA (jan. de 2020). “Whole Slide Imaging and Its Applications to Histopathological Studies of Liver Disorders”. English. Em: *Frontiers in Medicine* 6. ISSN: 2296-858X. DOI: 10.3389/fmed.2019.00310. URL: <https://www.frontiersin.org/journals/medicine/articles/10.3389/fmed.2019.00310/full> (acesso em 17/03/2026).
- MITCHELL, TOM M. (1997). *Machine Learning*. New York: McGraw-Hill.

- NATIONAL RESEARCH COUNCIL (jun. de 2014). *Convergence: Facilitating Transdisciplinary Integration of Life Sciences, Physical Sciences, Engineering, and Beyond*. en. Pages: 18722. Washington, D.C.: National Academies Press. ISBN: 978-0-309-30151-0. DOI: 10.17226/18722. URL: <https://www.nationalacademies.org/publications/18722> (acesso em 19/03/2026).
- PANTANOWITZ, LIRON, PAUL N. VALENSTEIN, ANDREW J. EVANS, KEITH J. KAPLAN, JOHN D. PFEIFER, DAVID C. WILBUR, LINDA C. COLLINS e THOMAS J. COLGAN (2010). “Review of digital pathology”. Em: *Journal of Pathology Informatics* 1.1, p. 15.
- REDMON, JOSEPH, SANTOSH DIVVALA, ROSS GIRSHICK e ALI FARHADI (2016). “You only look once: Unified, real-time object detection”. Em: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779–788.
- SANTOS, DALÍ F. D. DOS, PAULO R. DE FARIA, BRUNO A. N. TRAVENÇOLO e MARCELO Z. DO NASCIMENTO (ago. de 2023). “Influence of Data Augmentation Strategies on the Segmentation of Oral Histological Images Using Fully Convolutional Neural Networks”. Em: *Journal of Digital Imaging* 36.4, pp. 1608–1623. ISSN: 0897-1889. DOI: 10.1007/s10278-023-00814-z. URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC10406800/> (acesso em 09/04/2026).
- SANTOS, MARIANA DOURADO X. S., WILLIAM LAUS BERTEMES, IAAN MESQUITA DE SOUZA, MATEUS HENRIQUE B. ANDRADES, DAVID ANTONIO T. M. BARROS e VINICIUS SEBBA PATTO (out. de 2022). “Algoritmos de Aprendizado de Máquina para Classificação de Células Nucleadas do Sangue Periférico - Uma Experiência do Projeto Hemovision”. pt. Em: *Escola Regional de Informática de Goiás (ERI-GO)*. SBC, pp. 130–140. DOI: 10.5753/erigo.2022.227698. URL: <https://sol.sbc.org.br/index.php/erigo/article/view/22543> (acesso em 25/01/2026).
- SHORTEN, CONNOR e TAGHI M. KHOSHGOFTAAR (2019). “A survey on Image Data Augmentation for Deep Learning”. Em: *Journal of Big Data* 6.1, pp. 1–48. DOI: 10.1186/s40537-019-0197-0.
- SIMONYAN, KAREN e ANDREW ZISSERMAN (2014). “Very deep convolutional networks for large-scale image recognition”. Em: *arXiv preprint arXiv:1409.1556*.
- SOUZA, DANIEL SANTOS, LEANDRO MEDRADO e LYCIA de BRITO GITIRANA (2010). “Histologia”. Em: *Conceitos e métodos para a formação de profissionais em laboratórios de saúde*. Ed. por ETELCIA MORAES MOLINARO, LUZIA FÁTIMA GONÇALVES CAPUTO e MARIA REGINA REIS AMENDOEIRA. Vol. 2. Rio de Janeiro: EPSJV, pp. 43–88. URL: https://www.epsjv.fiocruz.br/sites/default/files/capitulo_2_vol2.pdf.

- TURNER, KEVIN O., MARGARET H. COLLINS, MARJORIE M. WALKER e ROBERT M. GENTA (abr. de 2022). “Quantification of Mucosal Eosinophils for the Histopathologic Diagnosis of Eosinophilic Gastritis and Duodenitis”. Em: *The American Journal of Surgical Pathology* 46.4, pp. 557–566. ISSN: 0147-5185. DOI: 10.1097/PAS.0000000000001843. URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC8923360/> (acesso em 25/01/2026).
- XIE, WEIDI, J ALISON NOBLE e ANDREW ZISSERMAN (2015). “Microscopy Cell Counting with Fully Convolutional Regression Networks”. en. Em: *MICCAI 1st Workshop on Deep Learning in Medical Image Analysis*.
- XING, FUYONG e LIN YANG (2016). “Robust Nucleus/Cell Detection and Segmentation in Digital Pathology and Microscopy Images: A Comprehensive Review”. Em: *IEEE reviews in biomedical engineering* 9, pp. 234–263. ISSN: 1937-3333. DOI: 10.1109/RBME.2016.2515127. URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC5233461/> (acesso em 25/01/2026).