

UNIVERSIDADE ESTADUAL PAULISTA "JÚLIO DE MESQUITA FILHO"

FACULDADE DE CIÊNCIAS - CAMPUS BAURU

DEPARTAMENTO DE COMPUTAÇÃO

BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

EDUARDO ALVARO MONTEIRO

**APRENDIZADO DE MÁQUINA APLICADO À QUANTIFICAÇÃO
DO ÍNDICE DE PROLIFERAÇÃO EM OSTEOSSARCOMAS
CANINOS**

BAURU

Novembro/2025

EDUARDO ALVARO MONTEIRO

**APRENDIZADO DE MÁQUINA APLICADO À QUANTIFICAÇÃO
DO ÍNDICE DE PROLIFERAÇÃO EM OSTEOSSARCOMAS
CANINOS**

Trabalho de Conclusão de Curso do Curso
de Bacharelado em Ciência da Computação
da Universidade Estadual Paulista “Júlio
de Mesquita Filho”, Faculdade de Ciências,
Campus Bauru.

Orientador: Prof. Dr. Clayton Reginaldo Pereira

BAURU

Novembro/2025

M775a	Monteiro, Eduardo Alvaro Aprendizado de Máquina aplicado à quantificação do Índice de Proliferação em Osteossarcomas Caninos / Eduardo Monteiro. -- , 2025 62 p. Trabalho de conclusão de curso (-) - Universidade Estadual Paulista (UNESP), Faculdade de Ciências, Bauru, Orientador: Clayton Reginaldo Pereira 1. Osteossarcoma canino. 2. Aprendizado de máquinas. 3. Índice de proliferação. 4. Ki-67. 5. Processamento de imagens. I. Título.
-------	---

Eduardo Alvaro Monteiro

APRENDIZADO DE MÁQUINA APLICADO À QUANTIFICAÇÃO DO ÍNDICE DE PROLIFERAÇÃO EM OSTEOSSARCOMAS CANINOS

Trabalho de Conclusão de Curso do Curso de
Bacharelado em Ciência da Computação da Uni-
versidade Estadual Paulista "Júlio de Mesquita
Filho", Faculdade de Ciências, Campus Bauru.

Banca Examinadora

Prof. Dr. Clayton Reginaldo Pereira

Orientador

Departamento de Computação

Faculdade de Ciências

Universidade Estadual Paulista "Júlio de
Mesquita Filho"

**Profa. Dra. Simone das Graças
Domingues Prado**

Departamento de Computação

Faculdade de Ciências

Universidade Estadual Paulista "Júlio de
Mesquita Filho"

Prof. Dr. Douglas Rodrigues

Departamento de Computação

Faculdade de Ciências

Universidade Estadual Paulista "Júlio de
Mesquita Filho"

Bauru, 14 de Novembro de 2025.

Agradecimentos

Com imensa gratidão, agradeço primeiramente à UNESP pela formação de excelência e por ter possibilitado o desenvolvimento deste trabalho. Expresso meu profundo reconhecimento ao meu orientador, Prof. Dr. Clayton Reginaldo Pereira, cuja orientação, paciência e conhecimento técnico foram cruciais para guiar e concluir esta pesquisa com sucesso. Estendo esses agradecimentos à Profa. Dra. Simone das Graças Domingues Prado e ao Prof. Dr. Douglas Rodrigues pela composição da banca.

Agradeço também aos nossos colegas veterinários da UNESP campus de Botucatu, principalmente à Layla Contessotto de Oliveira e ao Prof. Dr. Alexandre Hataka, pela troca interdisciplinar e auxílio na aquisição das imagens que viabilizaram esta pesquisa.

O apoio incondicional da minha família foi a força motriz e o alicerce emocional em toda a jornada. Agradeço aos meus pais, Denise e Fábio, e a meu irmão, Murilo, por serem meu porto seguro e exemplo de amor e valores. O carinho e o suporte dos meus avós, tios e primos demonstraram a força e a união que me permitiram superar cada desafio. Em especial, reconheço o Alexandre como minha maior inspiração na área, fonte de orientação e de conselhos valiosos.

Aos amigos, tanto de Bauru quanto de Sorocaba, sou grato por serem o alívio e a companhia nos momentos mais difíceis da graduação e da pesquisa, tornando a jornada mais leve e prazerosa. A amizade de vocês foi essencial.

Por fim, à minha namorada, Letícia, por ser o meu maior incentivo. Obrigado por ser minha companheira de vida.

"You can't connect the dots looking forward; you can only connect them looking backwards."
(STEVE JOBS, 2005)

Resumo

O osteossarcoma canino é uma das neoplasias ósseas mais agressivas na espécie, cujo prognóstico é frequentemente avaliado pelo índice de proliferação celular através do marcador imuno-histoquímico Ki-67. A obtenção deste índice depende da contagem manual de células em lâminas histopatológicas, um processo tradicionalmente laborioso, repetitivo e sujeito a elevada variabilidade inter e intraobservador. O advento da patologia digital permitiu o desenvolvimento de ferramentas computacionais para superar essas limitações. Neste contexto, o presente trabalho tem como objetivo propor e validar uma rotina computacional para a segmentação e contagem automatizada de células em proliferação e células saudáveis em imagens histológicas, visando a obtenção do índice de proliferação celular. A metodologia proposta explora duas abordagens distintas: o desenvolvimento de uma *pipeline* baseada em técnicas clássicas de processamento de imagens, e, para fins comparativos, o desenvolvimento de um modelo de aprendizado profundo baseado em Redes Neurais Convolucionais. Para viabilizar o estudo, foi criada uma base de dados de lâminas digitalizadas e anotadas por patologistas especialistas.

Palavras-chave: Osteossarcoma Canino, Aprendizado de Máquina, Índice de Proliferação, Ki-67, Processamento de Imagens, Patologia Digital, Segmentação.

Abstract

Canine osteosarcoma is one of the most aggressive bone neoplasms in the species, whose prognosis is frequently assessed by the cell proliferation index using the Ki-67 immunohistochemical marker. Obtaining this index depends on the manual counting of cells in histopathological slides, a process that is traditionally laborious, repetitive, and subject to high inter- and intra-observer variability. The advent of digital pathology has allowed the development of computational tools to overcome these limitations. In this context, this work proposes and validate a computational routine for the automated segmentation and counting of proliferating cells and healthy cells in histological images, aiming to obtain the cell proliferation index. The proposed methodology explores two distinct approaches: the development of a *pipeline* based on classic image processing techniques, and, for comparative purposes, the development of a deep learning model based on Convolutional Neural Networks. To enable the study, a database of digitized slides annotated by expert pathologists was created.

Keywords: Canine Osteosarcoma, Machine Learning, Proliferation Index, Ki-67, Image Processing, Digital Pathology, Segmentation.

Lista de figuras

Figura 1 – Regiões mais afetadas pelo OSA	15
Figura 2 – Lâmina do Controle Positivo	16
Figura 3 – Lâmina do Controle Negativo	17
Figura 4 – Arquitetura da CNN	23
Figura 5 – Arquitetura da YOLOv8-Seg	25
Figura 6 – Arquitetura da U-Net	26
Figura 7 – Imagens de lâminas com o "padrão ouro" de preparação	31
Figura 8 – Imagens de lâminas com o "padrão prata" de preparação	31
Figura 9 – Pipeline de processamento de imagens	34
Figura 10 – RGB (a) e HSV (b)	35
Figura 11 – Treinamento e validação do modelo U-Net	37
Figura 12 – Imagem de exemplo	43
Figura 13 – Contornos detectados pelo algoritmo Watershed	44
Figura 14 – Separação cromática (à esquerda) e balanceamento de branco (à direita)	44
Figura 15 – Imagem original e Canais de cor HSV separados	45
Figura 16 – Células em proliferação isoladas por limiarização	46
Figura 17 – Células saudáveis isoladas por limiarização	46
Figura 18 – Células em proliferação filtradas	47
Figura 19 – Células saudáveis filtradas	47
Figura 20 – Células em proliferação preenchidas	48
Figura 21 – Células saudáveis preenchidas	48
Figura 22 – Células em proliferação e saudáveis segmentadas	49
Figura 23 – Classificação das células em proliferação - <i>pipeline</i>	49
Figura 24 – Classificação das células saudáveis - <i>pipeline</i>	50
Figura 25 – Distribuição de Densidade dos valores de Precisão	51
Figura 26 – Distribuição de Densidade dos valores de <i>Recall</i>	52
Figura 27 – Perda do Treinamento vs. Validação	53
Figura 28 – Predição pura do modelo treinado	54
Figura 29 – Predição filtrada	54
Figura 30 – Células em proliferação segmentadas - U-Net	55
Figura 31 – Classificação das células em proliferação - U-Net	55

Lista de Quadros

Quadro 1 – Composição da Base de Dados	33
Quadro 2 – Resultados da <i>pipeline</i>	51
Quadro 3 – Resultados da U-Net	56
Quadro 4 – Comparativo células em proliferação <i>pipeline</i> e U-Net para figura 12 . . .	56

Lista de abreviaturas e siglas

BCE	Entropia Cruzada Binária
CNNs	Redes Neurais Convolucionais
DAB	Diaminobenzina
DL	Aprendizado Profundo
FN	Falso Negativo
FP	Falso Positivo
HE	Hematoxilina e Eosina
IA	Inteligência Artificial
ML	Aprendizado de Máquina
NMS	Supressão de não Máximos
OSA	Osteossarcoma Canino
TN	Verdadeiro Negativo
TP	Verdadeiro Positivo
WSI	Digitalização de Lâmina Inteira
YOLO	You Only Look Once

Sumário

1	INTRODUÇÃO	13
2	FUNDAMENTAÇÃO TEÓRICA	15
2.1	Osteossarcoma Canino	15
2.2	Visão computacional aplicada	17
2.3	Segmentação	18
2.3.1	Segmentação por Processamento de imagens	19
2.3.1.1	Limiarização	20
2.3.1.2	Técnicas baseadas em região	20
2.3.1.3	<i>Watershed</i>	21
2.3.1.4	Técnicas baseadas em borda	21
2.3.2	Segmentação por Aprendizado de máquina	22
2.3.2.1	Redes Neurais Convolucionais (CNNs)	22
2.3.2.2	You Only Look Once (YOLO)	24
2.3.2.3	U-Net	26
2.3.2.4	<i>Data Augmentation</i>	28
3	METODOLOGIA	30
3.1	Criação da base de dados	30
3.1.1	Coleta e Digitalização de Dados Brutos	30
3.1.2	Anotação das células	32
3.2	Processamento de imagens	33
3.2.1	Conversão do espaço de cor	34
3.2.2	Isolamento das cores	35
3.2.3	Aplicação de filtros	35
3.2.4	Contagem das células	36
3.3	Aprendizado de máquina	36
3.3.1	Arquitetura	37
3.3.2	Modelagem do problema	38
3.3.3	Geração de máscaras e Data Augmentation	38
3.3.4	Treinamento do modelo	39
3.3.5	Inferência	40
3.4	Validação dos processos	40
3.4.1	Geração de métricas	40
3.4.2	Análise Estatística	41

4	RESULTADOS	43
4.1	Experimentos	43
4.2	Classificação	52
4.3	Discussão	56
5	CONCLUSÃO	58
	REFERÊNCIAS	60

1 Introdução

A população de cães como animais de companhia tem crescido significativamente nos últimos anos, tornando a espécie parte importante do cotidiano de milhares de pessoas. Paralelo a esse cenário, observa-se o crescimento da prevalência de neoplasias - fenômeno atribuído tanto ao progresso das tecnologias de diagnóstico, quanto à longevidade crescente dos animais de estimação, já que cães em idade avançada possuem um risco maior de desenvolver essa doença em relação a animais jovens (FILHO et al., 2010).

Dentre as neoplasias caninas, 80% dos casos de tumores malignos ósseos correspondem ao osteossarcoma canino (OSA), um dos mais agressivos na espécie (BRODEY; RISER, 1969). Tradicionalmente, o diagnóstico do OSA é realizado através da constatação da proliferação de células mesenquimais malignas, produtoras de osteóide e de osso (MEUTEN, 2020). Como fator prognóstico, utiliza-se o índice de proliferação celular, obtido com uso da análise imuno-histoquímica, ou seja, a relação entre células doentes e células totais (saudáveis e doentes). Quanto maior o índice proliferativo, maior o risco de ocorrerem metástases e maior a agressividade do tumor (SCHOLZEN; GERDES, 2000).

Nesse contexto, o Ki-67, é uma proteína nuclear associada à proliferação celular e frequentemente utilizada como marcador para avaliar a taxa de crescimento celular em diferentes tecidos (SCHOLZEN; GERDES, 2000). Um nível elevado de expressão de Ki-67 geralmente está associado a um crescimento tumoral mais rápido e a um comportamento mais agressivo da neoplasia (HORIMOTO et al., 2015). Dessa forma, para determinar o grau de agressividade do tumor, e assim influenciar a conduta médica, é essencial calcular o índice de proliferação celular com o auxílio da imuno-histoquímica Ki-67.

A obtenção do índice proliferativo é um processo manual e repetitivo, que consiste na análise de partes do tecido celular, utilizando um microscópio óptico e um contador de células. Normalmente essa contagem é realizada por dois patologistas independentes, buscando uma maior assertividade no diagnóstico.

No entanto, esse processo em OSA é um desafio crítico na patologia veterinária. Os métodos convencionais, como a contagem manual de células proliferativas em imagens histopatológicas, são altamente dependentes da disponibilidade dos patologistas e estão sujeitos a variabilidade inter e intraobservador. Além disso, a subjetividade e o tempo demandado por esses métodos dificultam a padronização dos prognósticos, podendo impactar negativamente no tratamento adequado dos pacientes.

O advento da patologia digital, impulsionado pela Digitalização de Lâmina Inteira (*Whole Slide Imaging* - WSI), inaugurou uma nova era na oncologia diagnóstica, permitindo o desenvolvimento de ferramentas de análise de imagem computacional. Essas ferramentas ofere-

cem o potencial de superar as limitações da análise manual, proporcionando uma quantificação objetiva e rápida de biomarcadores morfológicos (GURCAN et al., 2009). O processamento de imagens digitais emergiu como a tecnologia central nesse campo, fornecendo um conjunto de algoritmos capazes de identificar, segmentar e quantificar estruturas de interesse, como núcleos celulares, de forma automatizada. As abordagens clássicas frequentemente envolvem um fluxo de trabalho que inclui etapas de pré-processamento da imagem, segmentação para isolar as células do fundo tecidual, extração de características (como forma, tamanho e textura) e, por fim, a classificação ou contagem dos objetos detectados (MADABHUSHI; LEE, 2016).

Outra possibilidade, com a evolução do Aprendizado de Máquina (*Machine Learning* - ML) na área da visão computacional e o surgimento do Aprendizado Profundo (*Deep Learning* - DL), introduzido por LeCun, Bengio e Hinton (2015), são as Redes Neurais Convolucionais (*Convolutional Neural Networks* - CNNs), que têm mostrado eficiência em tarefas de patologia digital. Esses modelos, a partir de um conjunto de dados previamente rotulado por especialistas, são capazes de aprender relações não lineares e realizar a quantificação do índice de proliferação de forma automática e mais reprodutível.

Neste contexto, o presente trabalho tem como objetivo propor e validar uma rotina computacional para a contagem de células em imagens histológicas de OSA. A metodologia proposta busca enfrentar os desafios de variabilidade de cor e complexidade morfológica por meio de uma sequência de etapas de processamento de imagem. Além disso, para efeito de comparação, um banco de dados será produzido com o intuito de viabilizar o treinamento de um modelo baseado em CNNs, que também realizará a segmentação e contagem das células.

A criação de um método automatizado e confiável visa fornecer uma ferramenta de apoio ao patologista, aumentando a objetividade e a eficiência na avaliação da proliferação celular, um fator crucial para o prognóstico desta neoplasia. Além dos benefícios na prática clínica veterinária, o desenvolvimento de um modelo robusto pode contribuir para a expansão do uso de técnicas de Inteligência Artificial (IA) na pesquisa oncológica, abrindo caminho para estudos comparativos com outras neoplasias e potencializando descobertas sobre a biologia dos tumores. Dessa forma, este trabalho pode ter impactos tanto na melhoria da precisão diagnóstica quanto na evolução da medicina veterinária baseada em evidências.

Este trabalho está estruturado da seguinte forma: no Capítulo 2 apresenta-se a Fundamentação Teórica; no Capítulo 3 descreve-se a Metodologia; no Capítulo 4 são discutidos os Resultados; e no Capítulo 5 são apresentadas as Conclusões e propostas de trabalhos futuros.

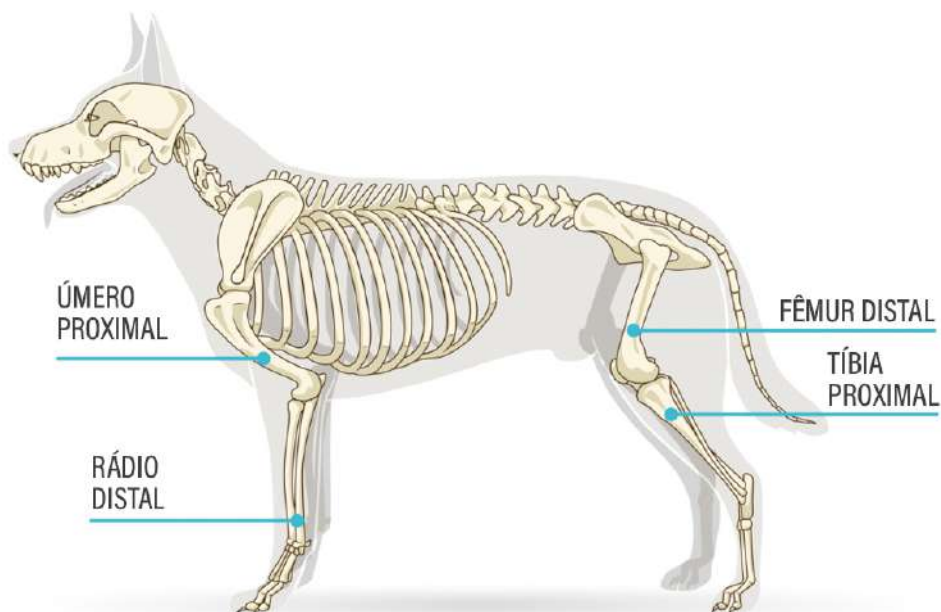
2 Fundamentação Teórica

Este capítulo apresenta a fundamentação teórica que serve como alicerce para o desenvolvimento deste trabalho. A abordagem inicia-se com uma revisão da doença, relacionando as propriedades visuais características de uma lâmina de OSA corada para análise de Ki-67 e a obtenção do índice de proliferação. Em seguida, são levantadas as abordagens de visão computacional mais comuns na patologia digital, tanto a partir de métodos tradicionais quanto de ML. Por fim, são discutidas as principais técnicas de segmentação disponíveis na literatura recente que formam a base para os capítulos subsequentes.

2.1 Osteossarcoma Canino

O OSA é a neoplasia óssea primária mais comum em cães, caracterizando-se por sua alta malignidade e comportamento agressivo. Trata-se de um tumor mesenquimal em que células neoplásicas produzem matriz osteoide ou óssea imatura. Geralmente, acomete cães de raças grandes a gigantes, de meia-idade a idosos, e possui uma elevada taxa de metástase, principalmente para os pulmões, o que torna o diagnóstico precoce um fator crucial para o manejo clínico e prognóstico do paciente (FILHO et al., 2010). A figura 1 evidencia os membros que são pela doença com mais frequência, os apêndices.

Figura 1 – Regiões mais afetadas pelo OSA

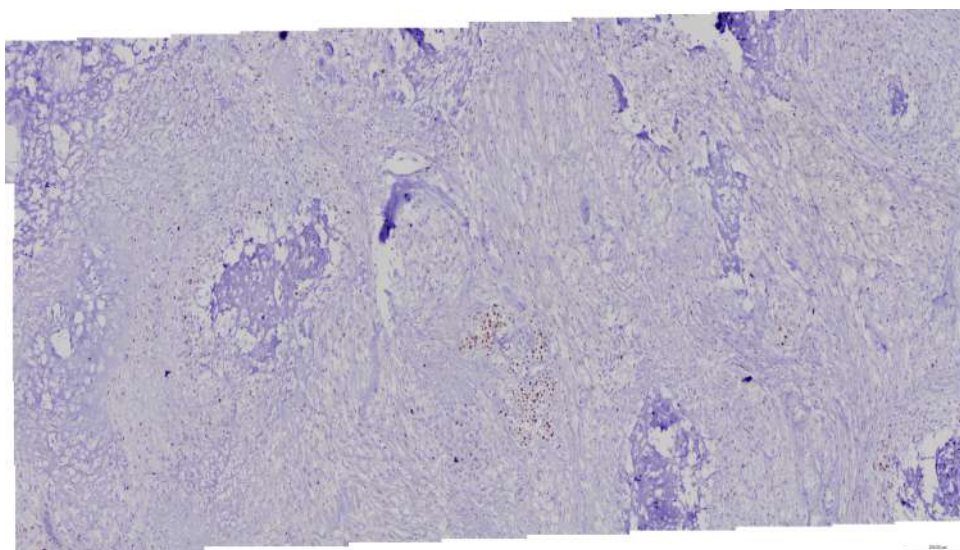


Ao microscópio, as amostras de OSA revelam uma notável heterogeneidade celular após a diluição do cromógeno diaminobenzina (DAB) e coração com Hematoxilina e Eosina (HE). As células tumorais exibem um acentuado pleomorfismo, com significativa variação no tamanho e na forma da célula como um todo. Frequentemente, os núcleos apresentam-se hipercromáticos, com a cromatina grosseiramente agregada, e nucléolos evidentes, indicativos de intensa atividade metabólica. A presença de mitose, muitas vezes atípicas, é um achado comum que reflete a rápida divisão celular. A matriz osteoide produzida por essas células aparece como um material eosinofílico e amorfo entremeadado às células (CAVALCANTI, 2007).

Para avaliar objetivamente a taxa de proliferação celular, utiliza-se a técnica de imunohistoquímica para detectar o antígeno nuclear Ki-67. Esta proteína está presente apenas em células que estão nas fases ativas do ciclo celular (G1, S, G2 e mitose), mas ausente na fase de repouso (G0). Nas amostras imunomarcadas, os núcleos das células em proliferação exibem uma coloração acastanhada, permitindo a quantificação da fração de crescimento do tumor. A expressão do Ki-67, portanto, funciona como um importante indicador prognóstico, onde um alto índice de marcação nuclear está diretamente correlacionado a um maior grau de malignidade e a um comportamento biológico mais agressivo da neoplasia (HORIMOTO et al., 2015).

A figura 2 evidencia um exemplo de lâmina do Controle positivo, ou seja, o tecido foi preparado especificamente para evidenciar a proteína Ki-67, e portanto a proliferação das células, na cor marrom. Essas são as lâminas que são utilizadas para a obtenção do índice de proliferação. Note que as faixas brancas nos cantos da imagem são falhas/limitações do scanner utilizado, o que não configura um problema, uma vez que apenas regiões específicas da lâmina para análise são selecionadas a partir de cortes e ampliações.

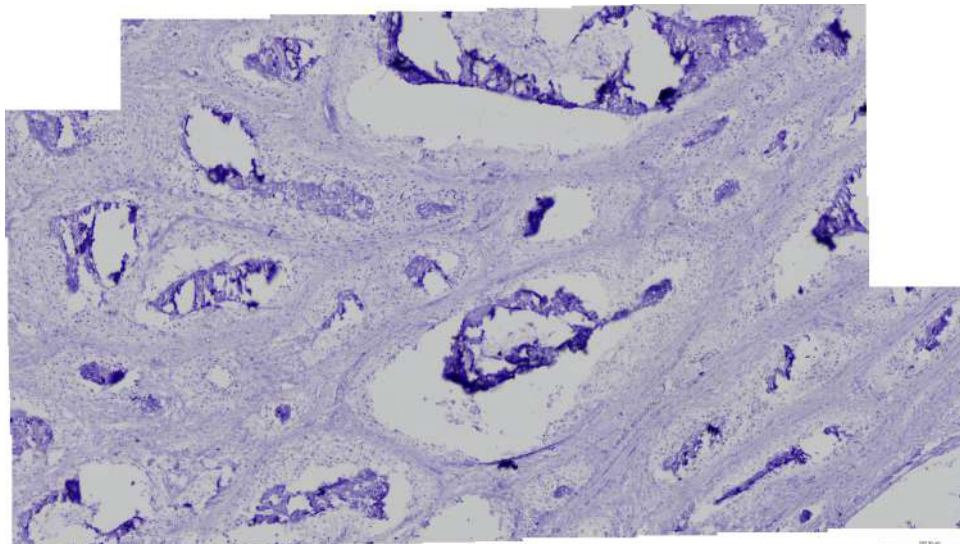
Figura 2 – Lâmina do Controle Positivo



Fonte: Elaborada pelo autor

Para efeito de comparação, a lâmina representada na figura 3 é um outro corte do mesmo animal, porém pertence ao Controle negativo, ou seja, não foi corada para esse propósito.

Figura 3 – Lâmina do Controle Negativo



Fonte: Elaborada pelo autor

A partir de uma lâmina do Controle positivo, um patologista é capaz de obter o índice de proliferação calculando uma razão simples entre a contagem das células em proliferação e células totais (em proliferação e saudáveis), exemplificada na equação 2.1:

$$\text{Índice de proliferação} = \frac{\text{Quantidade de células em proliferação}}{\text{Quantidade de células totais}} * 100 \quad (2.1)$$

No entanto, a não existência de um consenso entre os patologista a respeito da quantidade de células contadas necessárias para considerar aquele resultado válido, somado ao grande tempo exigido para a aplicação dessa técnica, inviabilizam a determinação do índice de proliferação na rotina do patologista.

2.2 Visão computacional aplicada

A patologia digital representa a evolução do diagnóstico anatomopatológico, consistindo na transição da análise de lâminas de vidro em microscópios ópticos para a avaliação de imagens digitais de alta resolução em monitores de computador. Esse processo é viabilizado por scanners especializados que realizam a WSI, convertendo o tecido físico em arquivos de imagem que podem ser armazenados, gerenciados e compartilhados eletronicamente.

Essa tecnologia rompe as barreiras físicas do laboratório, permitindo que patologistas acessem casos remotamente, facilitem interconsultas com especialistas em qualquer lugar do

mundo e criem arquivos digitais organizados para ensino e pesquisa. Além de otimizar o fluxo de trabalho e reduzir o risco de danos ou perda de lâminas físicas, a patologia digital abre caminho para a integração de ferramentas de análise de imagem e algoritmos de IA, que auxiliam na detecção de padrões, na quantificação de biomarcadores e na extração de dados objetivos, promovendo maior precisão, eficiência e reprodutibilidade ao diagnóstico.

A existência de um conjunto de características visuais típicas do OSA, aliado à digitalização de lâminas, viabiliza a aplicação de técnicas de visão computacional para auxiliar na análise histopatológica. As técnicas aplicadas devem objetivar a contagem de objetos, no contexto do desenvolvimento de rotinas computacionais capazes de quantificar elementos histológicos relevantes.

A literatura atual sobre OSA é predominantemente teórica/exploratória, concentrando-se na utilização do Ki-67 como marcador celular. Dessa forma, há uma lacuna crítica na aplicação prática desses conceitos na área da computação, o que impede a validação e o avanço da pesquisa. Embora trabalhos em OSA sejam limitados, pesquisas correlatas, com aplicações em câncer de pele (NAUTA et al., 2021), (BALAJI et al., 2023) e câncer de mama (JAGANATHAN et al., 2024), (BURRAI et al., 2023) demonstram a viabilidade e o potencial da visão computacional aplicada na patologia digital para gerar resultados tangíveis e inovadores.

Nesse contexto, a detecção de objetos, que normalmente resulta em caixas delimitadoras ou centróides é uma opção. No entanto, pela presença de regiões com semântica distinta, como tecido tumoral viável, áreas de necrose, estroma e a matriz osteoide característica do OSA, a segmentação se apresenta uma opção mais sofisticada e robusta.

2.3 Segmentação

A segmentação é uma técnica fundamental que visa particionar uma imagem em múltiplos segmentos ou conjuntos de pixels, com o objetivo de isolar e identificar objetos de interesse. A superioridade da segmentação como paradigma computacional para a análise de imagens médicas reside em sua capacidade de habilitar uma análise morfolétrica e espacial exaustiva. Ao delinear com precisão as fronteiras de estruturas como núcleos celulares ou regiões teciduais, a segmentação transforma a imagem em uma representação vetorial rica. Isso permite a extração de um vasto conjunto de características como área, perímetro, circularidade, excentricidade, e descritores de textura.

Adicionalmente, a informação espacial intrínseca preservada pelas máscaras de segmentação é crucial, pois não apenas refina a quantificação, mas seu maior nível de granularidade também potencializa a descoberta de novos biomarcadores de imagem. Por fim, a técnica se consolida como a mais apropriada para decodificar a heterogeneidade e a complexidade do OSA a partir de dados histopatológicos digitais, devido ao seu aprofundado nível de detalhe e

análise.

Essencialmente, a segmentação pode ser alcançada a partir de métodos tradicionais de processamento de imagem e abordagens baseadas em ML, especialmente o DL.

2.3.1 Segmentação por Processamento de imagens

Uma aplicação conhecida na patologia digital assistida por processamento de imagens é o *ImmunoRatio*, uma ferramenta de software baseada na web e código aberto, desenvolvida para auxiliar a análise de secções de tecido diversos. Dentre as inúmeras possibilidades de *plug-ins*, técnicas comumente utilizadas para análises quantitativas incluem: *Color Deconvolution*, que é crucial para isolar o corante DAB marrom (núcleos positivos) do azul de HE (contracolorante), Limiarização Adaptativa para separar os núcleos, distinguindo-os do fundo da imagem e acomodando variações de intensidade. E por fim, a segmentação de núcleos individuais possibilita a contagem do número total e dos núcleos positivamente marcados. Para garantir a máxima acurácia e mitigar a variabilidade, o software inclui uma etapa de calibração manual (TUOMINEN et al., 2010).

A avaliação e validação dessa ferramenta na obtenção do índice de Ki-67 foi realizada através de uma comparação, desenvolvida por Serras, Matos e Faísca (2020), entre a contagem manual e a análise automatizada. Os resultados, apesar de apresentarem boa correlação, foram marcadas por diversos fatores limitantes do software, como: Subestimação do índice de Ki-67 em magnificação de 200x, particularmente devido à dificuldade do software de contabilizar núcleos fracamente corados; diferenciar entre células tumorais e outros tipos celulares, como eosinófilos e mastócitos; distinguir tecido estromal de tecido tumoral; análise em magnificações mais baixas como em 50x.

Além das questões comuns quando se trata da análise de tecidos, a aplicação de algoritmos de processamento de imagem em cortes histológicos de OSA apresenta desafios intrínsecos que podem comprometer a robustez e a precisão dos sistemas automatizados. A própria natureza do OSA impõe dificuldades: o acentuado pleomorfismo celular, com grande variação no tamanho e na forma das células, e a frequente sobreposição de núcleos em áreas de alta densidade celular podem levar a erros de segmentação, onde múltiplos núcleos são interpretados como um único objeto (XING; YANG, 2016). As fronteiras celulares, muitas vezes tênues e mal definidas, também dificultam a separação precisa entre as células e o estroma circundante.

O levantamento das principais técnicas de processamento de imagens, com enfoque em segmentação, foi fundamental para a definição da metodologia para a contagem de células.

2.3.1.1 Limiarização

A limiarização, ou *thresholding*, visa particionar uma imagem em *foreground* e *background* por meio da criação de uma versão binária da imagem original em escala de cinza. A operação se baseia na comparação da intensidade de cada pixel com um valor de referência específico, denominado limiar T . A equação 2.2 exemplifica o *thresholding* binário invertido, no qual pixels (src) com intensidade superior ao limiar são atribuídos ao valor zero, enquanto aqueles com intensidade igual ou inferior são atribuídos ao valor máximo definido, geralmente o branco absoluto, representado pelo valor de pixel 1 ou 255.

$$dst(x, y) = \begin{cases} 0 & \text{se } src(x, y) > T(x, y) \\ \text{Valor Máximo} & \text{se não} \end{cases} \quad (2.2)$$

A abordagem mais direta é a limiarização global, na qual um único valor de T é aplicado a toda a imagem. A determinação desse valor é um passo crítico e frequentemente se baseia na análise do histograma de intensidades da imagem. Métodos automáticos se destacam, sendo o mais conhecido o Método de Otsu. Seu princípio é minimizar a variância intraclasse, que é a variância ponderada dentro das duas classes de pixels (*background* e *foreground*), ou, de forma equivalente, maximizar a variância interclasse (OTSU, 1975).

Apesar de sua eficácia, a limiarização global apresenta limitações significativas quando aplicada a imagens com condições de iluminação não uniformes, como a presença de gradientes de luz, sombras ou reflexos (SEZGIN; SANKUR, 2004). A limiarização adaptativa calcula um valor de limiar distinto para cada pixel ou para diferentes sub-regiões da imagem. O limiar é determinado com base nas características de intensidade da vizinhança local do pixel, utilizando medidas estatísticas.

2.3.1.2 Técnicas baseadas em região

As técnicas de segmentação baseadas em região operam sob o princípio da similaridade, agrupando pixels ou sub-regiões adjacentes que compartilham atributos comuns em relação a um conjunto de critérios predefinidos, como intensidade, cor ou textura. Duas estratégias clássicas e contrastantes dominam esta categoria: o *region growing* e o *split-and-merge*.

O *region growing* é uma abordagem do tipo *bottom-up*, que se inicia com a seleção de um ou mais pixels-semente, os quais servem como pontos de partida para cada região. De forma iterativa, os pixels vizinhos a essas sementes são avaliados e, se atenderem a um critério de homogeneidade são agregados a ela. Este processo de expansão continua até que nenhum pixel vizinho possa ser adicionado (ADAMS; BISCHOF, 1994). A qualidade da segmentação resultante é altamente dependente da escolha das sementes iniciais e da definição precisa do critério de homogeneidade.

Por outro lado, a técnica de *split-and-merge* adota uma abordagem hierárquica baseada em uma estrutura de dados conhecida como *quadtree*. O processo inicia-se com a imagem inteira como uma única região. Na fase de divisão *split*, qualquer região que não satisfaça o critério de homogeneidade é recursivamente dividida em quatro sub-regiões de igual tamanho, repetindo até que todas as sub-regiões resultantes sejam homogêneas. Subsequentemente, na fase de *merge*, duas ou mais regiões adjacentes que, quando combinadas, ainda satisfazem o predicado de homogeneidade, são unidas (HOROWITZ; PAVLIDIS, 1976). O resultado é uma partição da imagem mais robusta e menos sensível a ruídos do que a simples divisão.

2.3.1.3 Watershed

A segmentação por *Watershed* (ou divisor de águas) é uma abordagem que trata a imagem em escala de cinza como um relevo topográfico no qual a intensidade de cada pixel corresponde a uma elevação. O princípio do algoritmo pode ser visualizado como um processo de inundação simulada, onde bacias são progressivamente preenchidas a partir de seus mínimos locais. Quando a água proveniente de duas bacias distintas está prestes a se fundir, uma barreira, ou *watershed line*, é erguida para separá-las, formando contornos fechados (VINCENT; SOILLE, 1991).

Contudo, a aplicação direta do algoritmo de *Watershed* comumente resulta em um número excessivo de regiões pequenas e irrelevantes devido à presença de múltiplos mínimos locais causados por ruído ou variações de textura na imagem (MEYER; BEUCHER, 1990). Para mitigar este problema, a abordagem mais eficaz é o *Marker-Controlled Watershed*). Nesta variante, o processo de inundação é modificado para ser iniciado apenas a partir de um conjunto de marcadores previamente definidos, que podem ser pixels ou componentes conectados que identificam com segurança o interior dos objetos de interesse e o fundo.

2.3.1.4 Técnicas baseadas em borda

As técnicas de segmentação baseadas em borda buscam encontrar descontinuidades locais no sinal da imagem. Uma borda é definida como uma transição abrupta nos valores de intensidade dos pixels, sendo um forte indicativo dos limites entre objetos distintos ou entre um objeto e o fundo (SZELISKI, 2022). O processo de segmentação, nesta abordagem, consiste tipicamente em duas etapas sequenciais: primeiramente, a detecção das bordas e, em segundo lugar, a ligação desses pixels.

O detector de Canny otimiza a detecção de bordas ao satisfazer três critérios principais: alta taxa de detecção de bordas reais, boa localização (as bordas detectadas devem estar o mais próximo possível das bordas reais) e resposta única (uma única borda não deve gerar múltiplas respostas). Para alcançar isso, o método de Canny emprega um processo multifásico que inclui a suavização da imagem com um filtro gaussiano para reduzir ruídos, o cálculo da magnitude e da orientação do gradiente, a Supressão de não Máximos (*Non Max Supression*

- NMS) para afinar as bordas (NEUBECK; GOOL, 2006) e, por fim, um procedimento de limiarização por histerese, que utiliza dois limiares (alto e baixo) para conectar bordas fortes a contornos mais fracos, resultando em bordas mais contínuas e menos suscetíveis a ruídos (CANNY, 2009).

Adicionalmente, um método influente para a extração de estrutura topológica diretamente do processo de rastreamento de bordas foi proposto por Suzuki (1985). O algoritmo opera sobre uma varredura raster da imagem binária para detectar pontos de início de contornos, diferenciando entre bordas externas (de componentes de 1-pixels) e bordas de furos (de componentes de 0-pixels). A principal inovação da técnica reside em sua capacidade de determinar a relação de hierarquia e vizinhança entre os contornos detectados. Para cada nova borda encontrada, um número sequencial único é atribuído, e sua *parent border* é identificada com base na última borda encontrada durante a varredura na mesma linha. Após a identificação, o algoritmo percorre e marca os pixels pertencentes à borda, construindo uma representação estrutural da imagem.

2.3.2 Segmentação por Aprendizado de máquina

Modelos de ML, especialmente DL, tem demonstrado resultados promissores na patologia digital. Litjens et al. (2017) revisa coletivamente a aplicação extensiva das CNNs em várias modalidades de imagens médicas, incluindo raios-X de tórax, tomografias computadorizadas, patologia digital, imagens da mama, imagens cardíacas e ressonância magnética do cérebro. O avanço em tarefas como segmentação, classificação, detecção e diagnóstico de doenças, enfatizando a integração de modelos multi-escala, multi-visão e ensemble juntamente com técnicas tradicionais de processamento de imagens é destacado.

Perle (2019) explora a crescente integração de técnicas de ML na patologia veterinária. O artigo ressalta seu potencial para melhorar a precisão e a eficiência diagnósticas ao aproveitar grandes conjuntos de dados, como imagens digitais e prontuários eletrônicos de saúde. O artigo enfatiza que o ML serve como uma ferramenta complementar aos patologistas, e não como um substituto, destacando o papel crítico da rotulagem precisa dos dados e a importância da colaboração interdisciplinar.

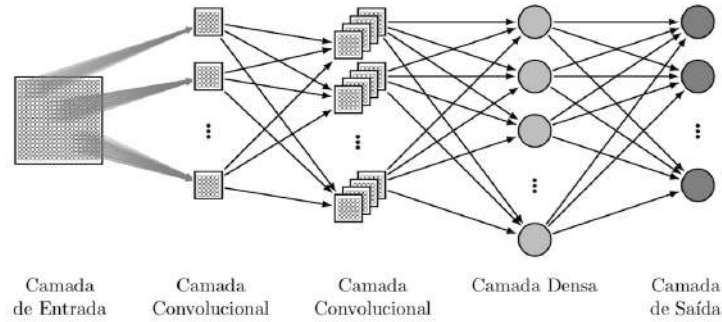
O levantamento das principais tecnologias de visão computacional, com enfoque em segmentação, no campo do ML foi fundamental para a escolha da arquitetura. As CNNs constituem o alicerce da visão computacional moderna e serve como o principal extrator de características para as arquiteturas de segmentação discutidas.

2.3.2.1 Redes Neurais Convolucionais (CNNs)

Sua origem remonta ao princípio biológico do córtex visual, concebida para processar dados com uma topologia de grade, como imagens. As CNNs funcionam matematicamente

através de uma sequência hierárquica de operações, como pode ser observado na figura 4.

Figura 4 – Arquitetura da CNN



Fonte: <https://www.sakurai.dev.br/cnn-mapreduce/> - Acessado dia 30/10/2025

Sua arquitetura pode ser dividida em três principais etapas, iniciando-se pela Camada Convolutiva, cuja operação central é a convolução, formalizada pela equação 2.3. Nesta expressão, o valor resultante na matriz de saída $Y(i, j)$ é obtido pela interação entre o filtro convolutivo $K(m, n)$ — um tensor de pesos treináveis — e a matriz de entrada X , indexada nas posições correspondentes. Fisicamente, esse conjunto de filtros desliza sobre a imagem de entrada realizando multiplicações elemento a elemento seguidas de uma soma, processo que gera os chamados Mapas de Ativação; dessa forma, cada filtro se especializa na detecção de uma *feature* visual específica, como bordas verticais ou horizontais, independentemente de sua posição na imagem, garantindo a propriedade de invariância translacional.

$$Y(i, j) = \sum_m \sum_n X(i - m, j - n) \cdot K(m, n) \quad (2.3)$$

A saída desta camada é então passada por uma função de ativação, tipicamente a ReLU, expressa pela equação 2.4. Nesta relação, a função $f(x)$ atua sobre o dado de entrada x selecionando o máximo entre zero e o próprio valor, o que efetivamente anula as entradas negativas enquanto preserva integralmente as positivas. Essa operação, além de ser computacionalmente mais eficiente que outras alternativas, desempenha o papel crucial de introduzir não-linearidade ao modelo, capacitando a rede neural a interpretar e modelar padrões de alta complexidade que operações puramente lineares não conseguiriam capturar.

$$f(x) = \max(0, x) \quad (2.4)$$

A segunda etapa constitui a Camada de *Pooling*, cuja função primordial é reduzir a dimensionalidade espacial do mapa de ativação, resultando em otimização do custo computacional e maior invariância a pequenas distorções na imagem. Para realizar essa compressão, utiliza-se frequentemente a técnica de *Max Pooling*, que opera aplicando uma janela retangular deslizante sobre os dados; o processo é regido pelo operador R_{max} , que seleciona exclusivamente a magnitude máxima dentro dessa região delimitada, descartando valores menores e preservando apenas a característica mais saliente daquela área específica.

$$Y(i, j) = (m, n) \in R_{max}X(m, n) \quad (2.5)$$

Por fim, após as camadas de extração de características, os dados são vetorizados em um formato unidimensional para alimentar as Camadas Totalmente Conectadas, onde a densa interconexão entre neurônios permite o aprendizado de correlações globais. A etapa de decisão é regida pela função *Softmax*, apresentada na equação 2.6, que converte o vetor de pontuações brutas (logits) z em uma saída interpretável. Matematicamente, a função calcula a exponencial da pontuação z_j referente a uma classe específica j e a normaliza pela somatória dos exponenciais de todas as K classes possíveis; esse processo assegura que o resultado $P(y = j|z)$ constitua uma distribuição de probabilidade válida, onde a soma de todas as saídas é igual a 1, facilitando a determinação da classe final.

$$P(y = j|z) = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}} \quad (2.6)$$

As características mais marcantes da CNN incluem o compartilhamento de pesos (o mesmo filtro é aplicado em toda a imagem), a conectividade esparsa (cada neurônio de saída é conectado apenas a uma pequena região da entrada) e a já mencionada invariância translacional, o que permite que a rede aprenda padrões espaciais de forma eficiente e com menos parâmetros do que as redes neurais tradicionais.

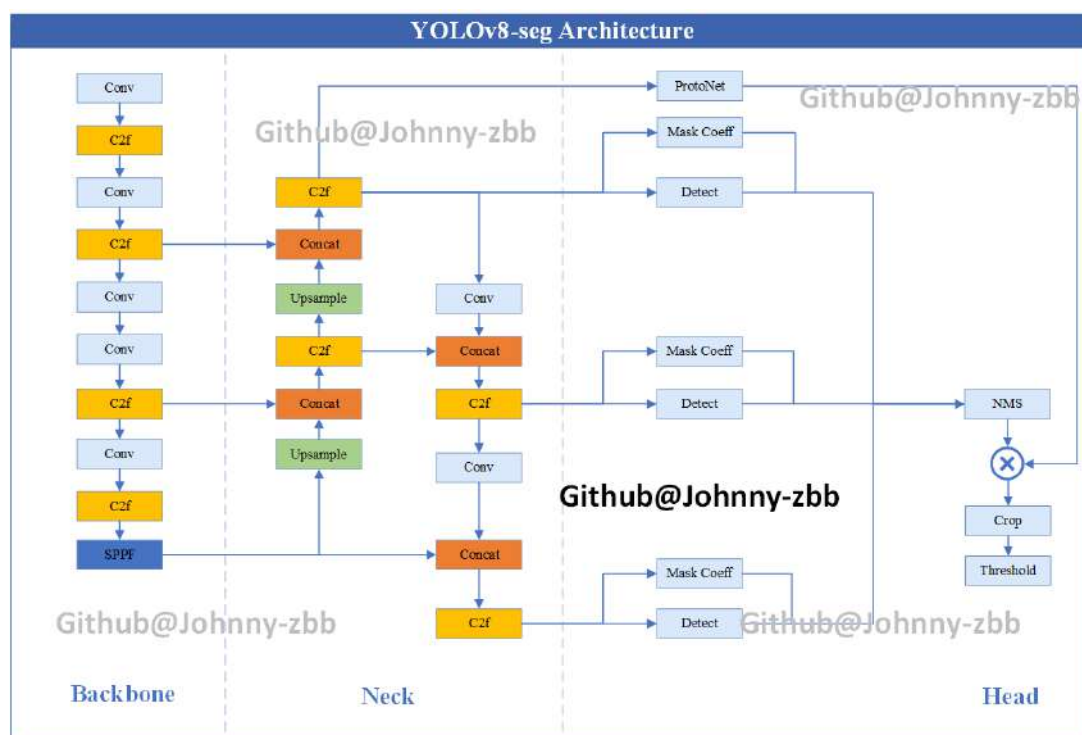
2.3.2.2 You Only Look Once (YOLO)

A arquitetura YOLO revolucionou a Detecção de Objetos ao reformulá-la como um problema de regressão espacial unificada, contrastando com os métodos anteriores de duas etapas, como R-CNN e o Faster R-CNN. Esse diferencial se dá pelo fato de que a YOLO processa a imagem em uma única avaliação, eliminando a necessidade de um passo intermediário de proposição de regiões (REDMON et al., 2016).

Sua fundamentação reside na utilização de CNNs profundas para extrair características hierárquicas diretamente do dado. Ela funciona como um *Backbone*, progressivamente reduzindo a resolução espacial da imagem de entrada enquanto enriquece a representação semântica através do aumento no número de canais de *feature*.

Embora mais conhecido e utilizado para a detecção de objetos, a YOLO tem se mostrado uma arquitetura robusta e eficiente também para a segmentação, introduzida a partir da YOLOv8. A YOLOv8-Seg representa a evolução do framework YOLOv8, estendendo suas capacidades da detecção de objetos para a segmentação de instâncias. Essa arquitetura pode ser visualizada na figura 5

Figura 5 – Arquitetura da YOLOv8-Seg



Fonte: <https://github.com/Johnny-zbb> - Acessado dia 30/10/2025

A estrutura do modelo é organizada em três componentes principais: o *Backbone*, já discutido anteriormente; o *Neck*, que processa as *features* brutas ao empregar mecanismos como as Redes de Agregação de Caminhos e Redes Piramidais de *features* para combiná-las em diferentes escalas; e a *Head*, a parte do modelo onde as predições finais são geradas, sendo dividida em duas saídas distintas: a *Detection Head* e a *Segmentation Head*. A primeira realiza a tarefa tradicional do YOLO, regredindo as coordenadas da caixa delimitadora e classificando a probabilidade de cada classe. A segunda, por sua vez, é a responsável pela segmentação de instâncias. Matematicamente, esta *Head* tipicamente envolve a predição de coeficientes que, quando combinados linearmente com um conjunto de máscaras protótipo de baixa resolução, geram a máscara final de alta resolução. Essa máscara é essencialmente uma classificação *pixel-wise* para determinar se cada pixel pertence ou não a uma instância detectada. Finalmente, o processo de pós-processamento refina as saídas, aplicando técnicas como a NMS.

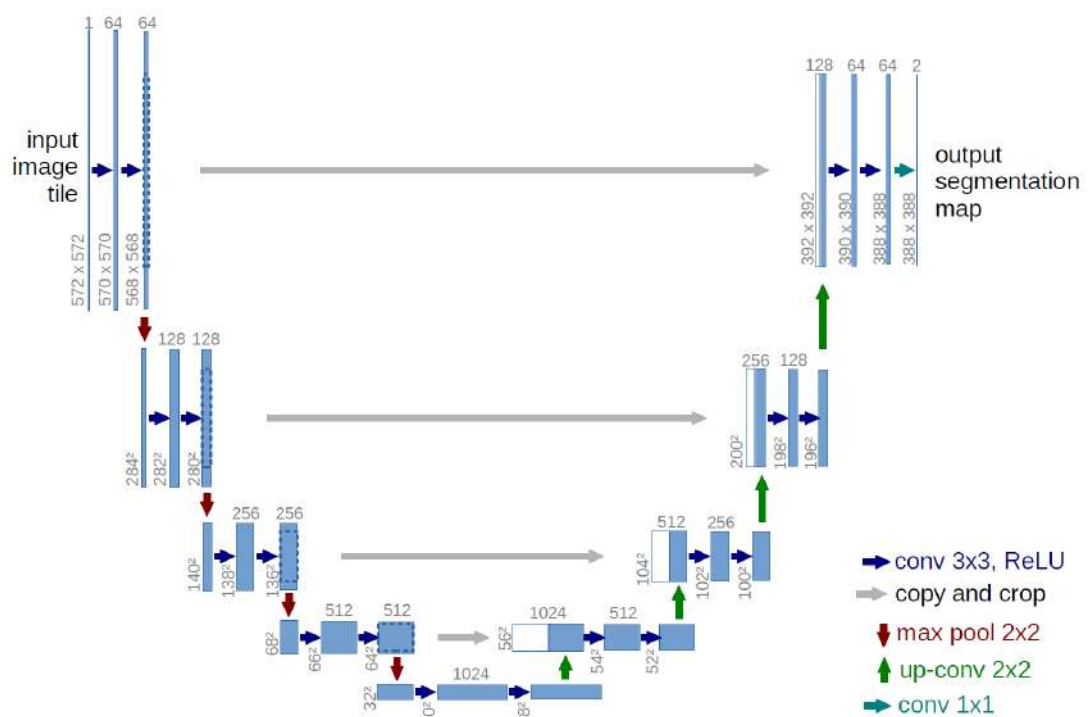
As características mais notáveis do YOLO são sua velocidade em tempo real e sua capacidade de utilizar o contexto global da imagem para fazer predições, o que reduz significa-

tivamente o número de falsos positivos em comparação com algoritmos baseados em regiões. Em essência, o YOLO atinge um equilíbrio notável entre precisão e velocidade, tornando-o um pilar para aplicações de visão computacional de baixa latência.

2.3.2.3 U-Net

A U-Net também é uma arquitetura baseada em CNNs. No entanto, a redução dimensional característica dessas redes, embora eficiente para a classificação de imagens, resulta na perda de informações de localização dos objetos. Para tarefas de segmentação, onde a saída deve ser um mapa de pixels que preserve o tamanho da imagem de entrada, essa perda de detalhes é extremamente prejudicial (RONNEBERGER; FISCHER; BROX, 2015). Para superar essa limitação, a U-Net foi projetada adotando uma arquitetura simétrica, em forma de 'U', como pode ser observado na figura 6:

Figura 6 – Arquitetura da U-Net



Fonte: U-Net: Convolutional Networks for Biomedical Image Segmentation

Na prática, a arquitetura pode ser dividida em duas partes principais:

- **Encoder (Caminho de Contração):** A primeira metade da rede, na esquerda do "U", funciona como um extrator de características, similar a uma CNN padrão. É composta por blocos de duas camadas convolucionais seguidas por uma camada de *Max Pooling*. À medida que a imagem avança por essa via, suas dimensões espaciais são reduzidas (subamostradas), enquanto o número de canais de características é dobrado. O objetivo

aqui é capturar o contexto e as características de alto nível da imagem, como a presença de objetos. Nesse caminho, a rede aprende o "O quê" da imagem; e

- **Decoder (Caminho de Expansão):** A segunda metade da rede, na direita do "U", é responsável por reconstruir a imagem segmentada. Ela utiliza camadas de *upsampling* (transposição de convolução ou convolução fracionária) para aumentar as dimensões espaciais da imagem de volta ao seu tamanho original. Cada passo de *upsampling* é seguido por camadas convolucionais que refinam a segmentação, utilizando as informações de contexto extraídas pelo *encoder*. Nesse caminho, a rede aprende o "Onde" da imagem.

Ainda, a principal revolução da arquitetura U-Net foi a introdução das *skip connections*. Essas conexões ligam diretamente as camadas correspondentes do caminho de contração com as do caminho de expansão, na mesma resolução. Elas fornecem ao *decoder* informações de alta resolução que foram perdidas durante a subamostragem no *encoder*. Sem elas, o *decoder* teria que confiar apenas nas informações de baixo nível e de alto contexto para reconstruir os detalhes finos da segmentação, o que levaria a resultados imprecisos e contornos borrados.

Outro fator importante é o desequilíbrio entre o pixels de *background* e as regiões de interesse em tarefas de segmentação. Nesse sentido, a escolha da função de perda é crucial, e o coeficiente de *Dice* desempenha um papel fundamental nesse contexto. Tradicionalmente, o desempenho da U-Net é otimizado pela minimização da *Dice Loss*, que é calculada a partir do Coeficiente de *Dice*, uma métrica de similaridade de conjuntos que avalia a sobreposição entre a segmentação predita pela rede e o *ground truth*, definida pela equação 2.7

$$DSC = \frac{2|X \cap Y|}{|X| + |Y|} \quad (2.7)$$

Nesta formulação, o termo $|X \cap Y|$ representa a interseção (pixels comuns) entre o conjunto de predições X e o conjunto de referência Y , enquanto o denominador normaliza o resultado pela soma das cardinalidades, ou seja, o total de pixels ativos em ambas as máscaras. O Coeficiente varia de 0 a 1 e penaliza discrepâncias em classes minoritárias de forma mais severa que a Entropia Cruzada Binária (BCE). A BCE, por sua vez, trata cada pixel de forma independente e busca maximizar a verossimilhança da classe correta, sendo expressa pela equação 2.8:

$$L_{BCE}(y, p) = -\frac{1}{N} \sum_{i=1}^N [y_i \log(p_i) + (1 - y_i) \log(1 - p_i)] \quad (2.8)$$

Nesta equação, o erro é computado como a média sobre o total de N pixels, onde y_i denota o rótulo verdadeiro (0 ou 1) do i -ésimo pixel e p_i corresponde à probabilidade estimada pelo modelo para aquele ponto. Devido a essas características complementares, frequentemente adota-se uma função de perda composta, somando a *Dice Loss* com a BCE. Essa estratégia

une o melhor de ambos os mundos: enquanto a *Dice Loss* foca na coerência global da forma e no desequilíbrio de classes, a BCE assegura que a rede aprenda as probabilidades de cada pixel de maneira robusta, refinando os pesos para produzir segmentações mais precisas nas regiões de interesse.

2.3.2.4 *Data Augmentation*

O Aumento de Dados, ou *Data augmentation*, é uma técnica de ML que expande artificialmente um conjunto de dados, criando versões modificadas dos dados já existentes. Esse processo ajuda a melhorar o desempenho do modelo, a evitar o *overfitting* ao introduzir variações e aprimora a capacidade de um modelo de generalizar para dados novos e não vistos. Dado a quantidade limitada de dados, característica de um problema da área médica, técnicas de *Data Augmentation* são amplamente utilizadas.

Para [Shorten e Khoshgoftaar \(2019\)](#), a implementação prática do *Data Augmentation* abrange um vasto espectro de técnicas. As mais fundamentais são as manipulações geométricas e fotométricas, que englobam transformações como espelhamento, escalonamento, rotação, translação, *cropping* e ajustes de brilho ou nitidez. Tais operações, notáveis pela eficiência computacional e facilidade de aplicação, demonstram ganhos de desempenho consistentes. Métodos mais sofisticados, como as técnicas baseadas em *erasing*, introduzem oclusões sintéticas ao apagar regiões da imagem. O propósito é forçar o modelo a focar em um contexto mais amplo, tornando-o mais robusto a oclusões parciais nos dados reais. Outra exemplo é a dos métodos baseados em *mixing*, como *Mixup*, que criam novos exemplos de treinamento pela combinação linear de duas ou mais imagens e seus respectivos rótulos. Embora os resultados possam ser visualmente não intuitivos, essas técnicas regularizam o comportamento da rede neural, promovendo uma transição mais suave entre as classes e gerando melhorias notáveis de desempenho.

Além das técnicas de aplicação manual, a pesquisa tem evoluído para estratégias que automatizam e otimizam o processo. O *AutoAugment*, por exemplo, visa descobrir sequências e combinações de operações adaptadas a um conjunto de dados e modelo específicos, frequentemente alcançando resultados superiores ao custo da sobrecarga computacional na fase de busca. O uso de modelos generativos profundos, como as Redes Adversárias Generativas, oferece outra via para sintetizar amostras de treinamento. Ainda, o *feature-level augmentation* atua não no espaço dos pixels, mas diretamente sobre as representações latentes extraídas pela rede, enriquecendo as características aprendidas ([SHORTEN; KHOSHGOFTAAR, 2019](#)).

Contudo, é imperativo reconhecer que a eficácia de uma estratégia de aumento de dados não é universal, sendo intrinsecamente dependente da tarefa e das características do conjunto de dados. No contexto das arquiteturas discutidas, a YOLO possui um módulo de *Data Augmentation* extremamente robusto, com diversas opções que podem ser especificadas como parâmetros durante o treinamento. Já a U-Net, destaca a importância do deslocamento e

rotação para reforçar as propriedades de invariância e robustez no aprendizado da rede, além de deformações elásticas aleatórias, que podem ser aplicadas de forma suave ([RONNEBERGER; FISCHER; BROX, 2015](#)).

3 Metodologia

O presente estudo desenvolveu uma metodologia de processamento de imagens digitais para a segmentação e contagem automatizada de células em lâminas histológicas de OSA. Uma metodologia para desenvolvimento de modelos baseados em CNNs para o mesmo fim foi aplicada para efeito de comparação.

Para dar suporte ao estudo, uma base de dados com peças cirúrgicas de tumores ósseos em membros apendiculares foi levantada por patologistas do Hospital Veterinário da Faculdade de Medicina Veterinária e Zootecnia/UNESP, campus de Botucatu, São Paulo.

3.1 Criação da base de dados

A base de dados utilizada foi construída a partir de um processo rigoroso dividido em duas fases principais: coleta e digitalização de dados brutos e anotação das células, com foco na geração de um conjunto de imagens de alta qualidade.

3.1.1 Coleta e Digitalização de Dados Brutos

Os dados brutos foram obtidos a partir de um estudo retrospectivo de tumores ósseos em cães. Informações clínicas foram resgatadas de um sistema eletrônico integrado e as amostras foram coletadas de arquivos físicos abrangendo o período de 2008 a 2025.

As amostras passaram por um rigoroso processo de preparação, que envolveu técnicas como inclusão de parafina, microtomia, desparafinagem, reidratação, bloqueio de peroxidase endógena, recuperação antigênica, entre outros. A detecção foi realizada com o kit Novolink, que utiliza a incubação do anticorpo secundário e do polímero. A revelação do sinal ocorreu com o cromógeno DAB, seguida pela contracoloração com HE e montagem final. Controles positivos e negativos foram empregados para validação. O resultado desse processo são lâminas de vidro devidamente preparadas para a análise das células.

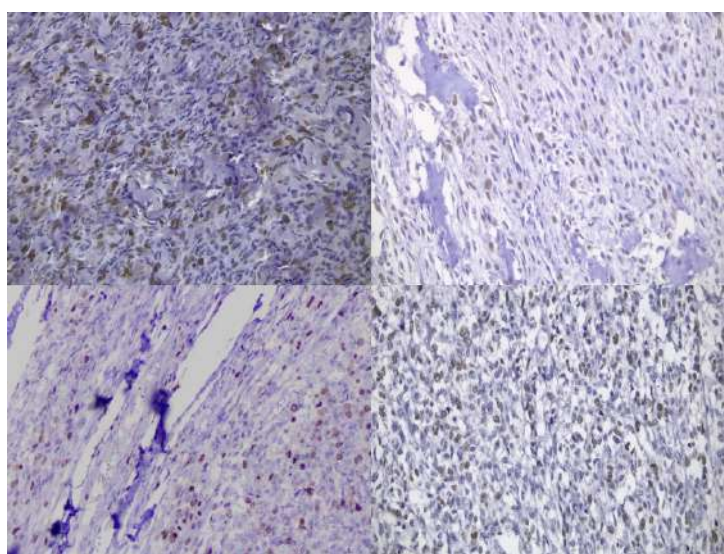
A digitalização contou com um scanner de lâminas com um aumento de 200x. Esta etapa é crucial para a conversão dos dados físicos em um formato digital. A captura de cada lâmina permitiu a geração de uma grande quantidade de detalhes microscópicos, essenciais para a análise histopatológica automatizada. As imagens resultantes, foram divididas em campos representativos. Para cada lâmina, foram selecionados aleatoriamente pelo menos 10 campos de visualização, totalizando 72 imagens de 7 animais distintos.

Um dos principais desafios desse processo como um todo é lidar com a heterogeneidade cromática, resultante da falta de padronização nos protocolos de preparação e coloração de

lâminas. Variações na concentração dos corantes, tempo de fixação e calibração dos scanners produzem imagens com aparências de cor drasticamente diferentes, dificultando a criação de um modelo de segmentação universalmente aplicável (VAHADANE et al., 2016).

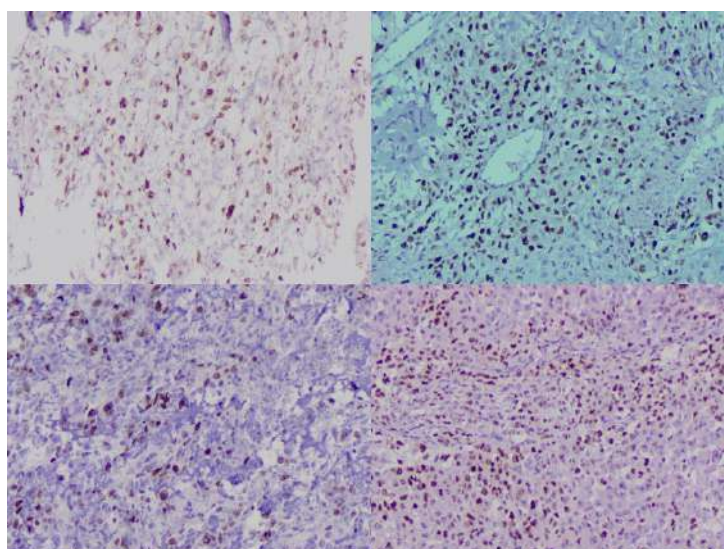
Diante dessa variabilidade, é útil classificar a qualidade das amostras. Definimos como "padrão ouro" as lâminas de alta qualidade, caracterizadas por uma nítida separação morfológica entre células adjacentes e entre células e estroma, além de alta consistência entre amostras (figura 7). Em contrapartida, o "padrão prata" descreve lâminas de qualidade inferior, onde essa separação é indistinta ou a coloração é irregular (figura 8).

Figura 7 – Imagens de lâminas com o "padrão ouro" de preparação



Fonte: Elaborada pelo autor

Figura 8 – Imagens de lâminas com o "padrão prata" de preparação



Fonte: Elaborada pelo autor

Para a construção da base de dados deste estudo, adotamos um critério de seleção focado na qualidade. Foram selecionadas, majoritariamente, lâminas classificadas como "padrão ouro". Essa decisão visa garantir que os alicerces do processo sejam construídos sob exemplos com maior clareza morfológica, facilitando a etapa de anotação e estabelecendo uma linha de base robusta.

3.1.2 Anotação das células

Após a virtualização, um patologista foi responsável por delimitar o contorno (segmentação) de cada célula em proliferação em um software de anotação de imagens. A plataforma escolhida foi o RoboFlow (<https://roboflow.com/> - Acessado dia 30/10/2025) pelos seguintes motivos:

- **Reconhecimento e Confiabilidade:** Sua reputação no mercado oferece garantia e uma base sólida para o desenvolvimento do projeto;
- **Conjunto de Ferramentas Abrangente:** Oferece uma gama de ferramentas de auxílio, como modelos pré-treinados e recursos de automação, que agilizam o processo de anotação;
- **Interface Intuitiva:** Mesmo para usuários sem experiência, como a equipe de veterinários e patologistas, a interface é intuitiva e de fácil uso; e
- **Gerenciamento em Nuvem e Colaboração:** O armazenamento em nuvem e a capacidade de compartilhar projetos permitiu a segurança e disponibilidade dos dados.

Para garantir a uniformidade da análise, foi estabelecido um critério de rotulagem com duas classes: células em proliferação, caracterizadas pela cor marrom, e células saudáveis, pela cor azul-escura. As anotações foram fundamentais para as duas abordagens do projeto: o processamento de imagens, na qual foram usadas para desenvolver algoritmos e gerar métricas de validação, e o DL, para o treinamento e teste dos modelos.

No entanto, o projeto enfrentou uma forte limitação devido à escassez de dados rotulados. Do universo de 72 imagens, apenas 18 foram anotadas por um patologista especialista e, destas, somente 1 continha a rotulagem completa de ambas as classes de células. A principal razão para essa disparidade foi o tempo exigido para o processo manual, especialmente o esforço extra para anotar as numerosas células saudáveis.

Essa restrição de dados impactou diretamente as duas frentes do trabalho. No processamento de imagens, a validação do algoritmo de contagem de células saudáveis ficou limitada a uma única amostra. De forma análoga, o modelo de ML pôde ser treinado exclusivamente com exemplos de células em proliferação, o que restringiu sua capacidade de predição apenas a essa classe.

Quadro 1 – Composição da Base de Dados

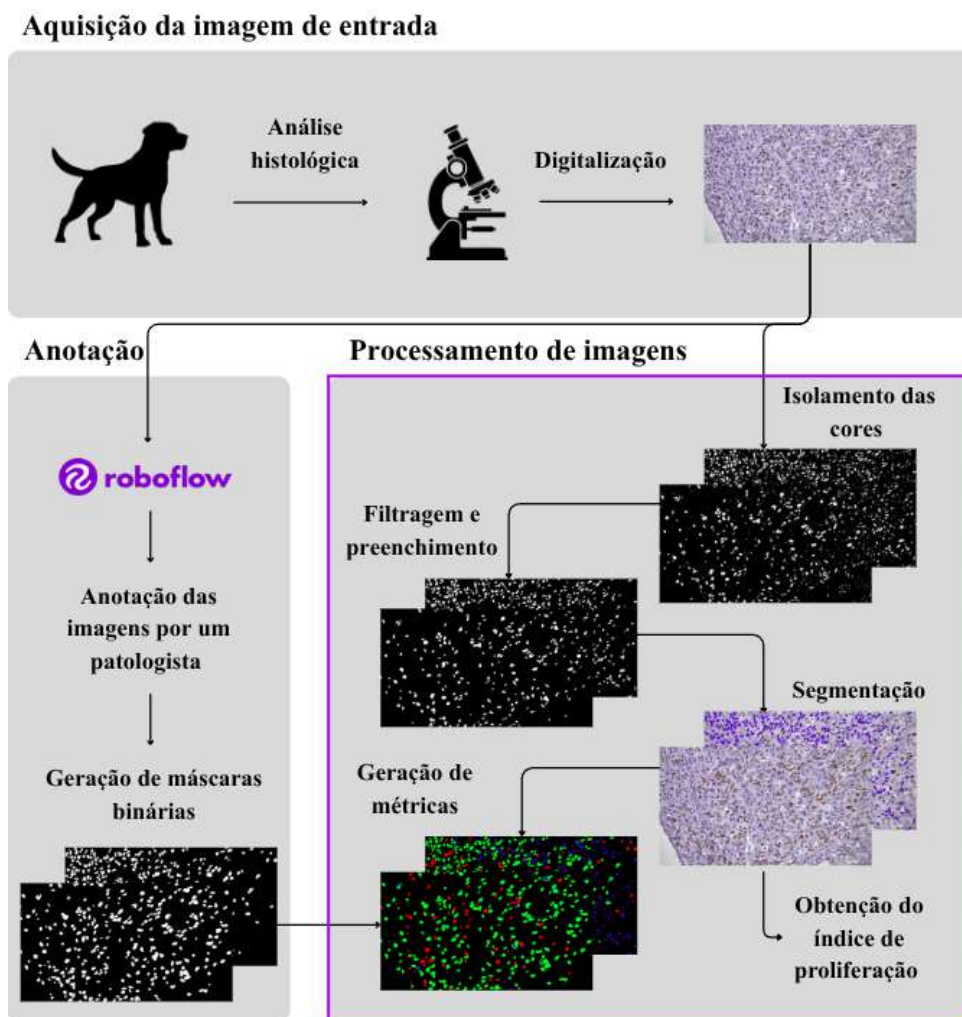
Tipo de dado	Quantidade
Animais	7
Lâminas	72
Imagens rotuladas parcialmente	18
Imagens rotuladas completamente	1

O quadro 1 resume a quantidade de Animais, Lâminas e Imagens rotuladas tanto parcial (apenas células em proliferação), quanto completamente. Observe que dessas 18 imagens efetivamente utilizadas ao longo do projeto, cerca de 70% são lâminas do conjunto definido como "padrão ouro" e os restantes 30% do "padrão prata".

3.2 Processamento de imagens

Esta seção descreve todo o processamento de imagens, desde do tratamento de uma imagem, até a segmentação e geração de métricas. O diagrama apresentado na figura 9 contempla as principais etapas empregadas em um passo a passo. Observe como cada técnica é exemplificada com duas imagens, sendo a mais à frente referente às células em proliferação e a mais ao fundo referente às células saudáveis.

Figura 9 – Pipeline de processamento de imagens



Fonte: Elaborada pelo autor

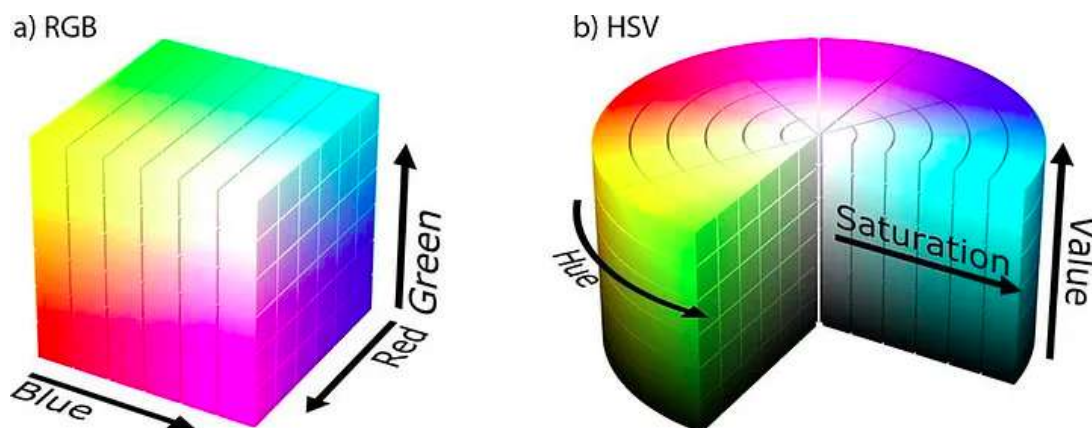
3.2.1 Conversão do espaço de cor

As imagens digitais, obtidas a partir dos cortes histológicos, são inicialmente representadas no modelo de cores RGB. Embora seja o esquema padrão para captura e exibição de imagens, o modelo apresenta limitações para a análise de imagens biológicas, uma vez que a informação de cor (crominância) está intrinsecamente acoplada à informação de brilho (luminância), tornando as operações sensíveis à variações de iluminação.

Para mitigar a heterogeneidade cromática intrínseca aos cortes histológicos, decorrente de variações no processo de coloração e na iluminação durante a captura, o espaço de cor HSV foi escolhido. Este modelo é mais alinhado à percepção humana, pois desacopla a matiz (*Hue* - H), que representa a cor pura, da saturação (*Saturation* - S), que indica a pureza da cor, e do valor (*Value* - V), que corresponde ao brilho (SZELISKI, 2022). A análise dos canais individualmente permite a identificação consistente das tonalidades de interesse, correspondentes às células em proliferação, na cor marrom e as demais células, na cor azul

escuro, independentemente das variações de intensidade luminosa presentes na amostra. A figura 10 representa visualmente os espaços de cor RGB e HSV, respectivamente:

Figura 10 – RGB (a) e HSV (b)



Fonte: <https://medium.com/neurosapiens/segmentation-and-classification-with-hsv-8f2406c62b39> - Acessado dia 30/10/2025

3.2.2 Isolamento das cores

A etapa de separação das cores foi realizada por meio de thresholding no espaço de cores HSV. Para cada tipo de célula de interesse, foram definidos intervalos de valores mínimos e máximos para os canais H, S e V. Com base nesses limiares, a imagem foi processada pixel a pixel para gerar uma máscara binária. O resultado deste processo é uma imagem em preto e branco que isola as áreas correspondentes às células-alvo.

Para encontrar os intervalos de valores mínimos e máximos foi desenvolvido um algoritmo capaz de selecionar um pixel com o mouse e retornar seu valor HSV. Com base nisso, foi realizada um levantamento dos pontos mais claros e mais escuros para as células saudáveis e em proliferação.

A partir de agora, convém trabalhar com as imagens em escala de cinza, uma vez que a informação que provém da cor já foi utilizada e as operações seguintes são mais facilmente aplicadas sobre a imagem com apenas um canal.

3.2.3 Aplicação de filtros

As células isoladas, contudo, frequentemente apresentam ruídos, como artefatos e pequenos agrupamentos de pixels incorretamente classificados, além de descontinuidades no interior dos objetos de interesse. Para refinar a máscara e prepará-la para a segmentação, foram aplicadas duas técnicas para pré-processamento: Análise de Componentes Conectados e operação morfológica de fechamento.

Primeiramente, utilizou-se a Análise de Componentes Conectados com critério de conectividade-8 para agrupar pixels adjacentes em objetos distintos (ROSENFELD; PFALTZ, 1966). Em seguida, um filtro de área foi aplicado para remover todos os componentes cuja área era inferior a um limiar pré-estabelecido, eliminando eficientemente os ruídos espúrios.

Para corrigir as falhas internas nas representações das células, a segunda técnica, operações de morfologia matemática, foi empregada. Especificamente, aplicou-se a operação de Fechamento, que consiste em uma dilatação seguida por uma erosão, utilizando um elemento estruturante. Primeiro, o processo de dilatação expande a área dos objetos, o que ajuda a conectar partes próximas e a fechar pequenas cavidades e lacunas. Em seguida, a erosão remove as bordas das áreas expandidas, restaurando os objetos ao seu tamanho original ou próximo a ele, mas mantendo a conectividade criada na etapa anterior. (HARALICK; STERNBERG; ZHUANG, 1987)

3.2.4 Contagem das células

Com a máscara devidamente refinada, procedeu-se à extração dos contornos individuais de cada célula. Subsequentemente, para a extração final dos contornos, empregou-se o algoritmo de rastreamento de bordas e análise topológica proposto por Suzuki (1985), que identifica e descreve a fronteira de cada objeto de forma hierárquica.

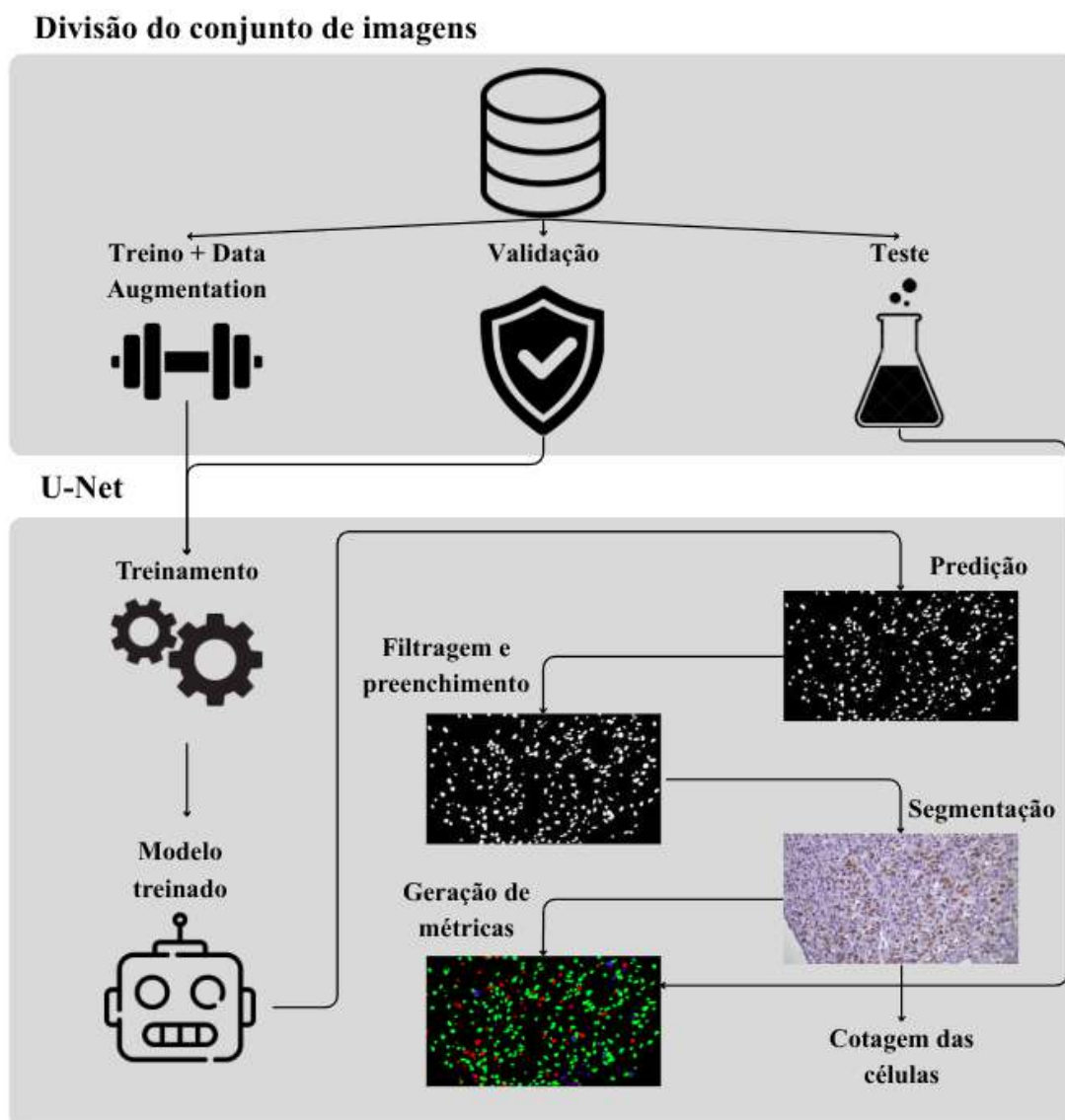
Apesar de sua eficiência, a principal característica desse algoritmo, o rastreo com base na vizinhança, não consegue resolver o principal problema da segmentação: Dois ou mais objetos conectados entre si. Esta pesquisa não empregou nenhuma técnica adicional para mitigar esse problema.

A partir das células segmentadas, é possível simplesmente verificar o número de contornos obtidos para encontrar a contagem das células. Além disso, como as células em proliferação e saudáveis já foram isoladas em etapas anteriores com base na cor, a segmentação pode ser realizada separadamente gerando duas máscaras finais e contagens diferentes.

3.3 Aprendizado de máquina

Essa seção descreve todo o processo baseado em DL, desde a preparação da base de dados para o treinamento do modelo e a geração de métricas. O diagrama representado na figura 11 destaca esse processo em duas fases principais, culminando na contagem das células. Observe que, diferentemente da rotina de processamento de imagens, o modelo foi treinado para prever apenas as células em proliferação, dado as limitações de dados, e por isso essa abordagem por si só não permite a obtenção do índice de proliferação.

Figura 11 – Treinamento e validação do modelo U-Net



Fonte: Elaborada pelo autor

3.3.1 Arquitetura

A escolha da arquitetura exigiu uma análise aprofundada das tecnologias disponíveis, considerando suas características, vantagens e limitações. Apesar das capacidades da YOLO, por ter um foco em velocidade e aplicações em tempo real, a qualidade das predições não pode ser comparada com modelos dedicados.

Diante da necessidade de alta precisão na segmentação de imagens, a U-Net emerge como uma escolha proeminente e amplamente validada na comunidade científica. Criada originalmente para imagens biomédicas, sua arquitetura se mostrou própria para o problema de segmentação de células. Além disso, o modelo é treinado sob máscaras de segmentação e essa também é sua principal saída, o que tornou o processo de transição da abordagem de

segmentação por processamento de imagens para a U-Net mais natural.

Quanto à aplicação da rede, foi escolhida a implementação customizada disponível no repositório: <https://github.com/milesial/Pytorch-UNet> (Acessado dia 30/10/2025). Seu reconhecimento na comunidade científica, disponibilidade em frameworks modernos, como o Pytorch, e validação em bases de dados conhecidas, como no desafio da empresa Carvana sediado pela plataforma Kaggle, evidenciaram sua robustez e confiabilidade.

3.3.2 Modelagem do problema

A modelagem de um problema de segmentação que possui apenas um objeto de interesse pode ser formalizada de duas maneiras distintas no que tange ao número de classes e o tratamento do *background* como uma classe ou não.

O primeiro método consiste em modelar o problema como segmentação binária, no qual se define a presença de duas classes. Neste cenário, a classe 0 representa o restante da imagem e a classe 1 corresponde ao objeto de interesse. O modelo, portanto, é treinado para distinguir pixel a pixel se o elemento pertence ao objeto alvo ou ao fundo da imagem.

A segunda abordagem, por sua vez, é a segmentação unária, onde se considera formalmente apenas uma classe, que é o próprio objeto de interesse. Neste modelo, o *background* é tratado implicitamente como "não-classe" ou como o resíduo do que não foi classificado. O foco do treinamento e da inferência é gerar uma máscara que demarque estritamente a localização e a forma do único objeto de interesse, sendo o que estiver fora dessa área automaticamente considerado *background*. Embora o resultado prático de uma máscara de segmentação possa parecer semelhante em ambos os casos (objeto versus não-objeto), a diferença reside na formulação da função de perda e na saída do modelo (por exemplo, duas probabilidades versus uma probabilidade ou logit), o que impacta o processo de treinamento e a interpretação das previsões.

Este trabalho optou por modelar o problema como uma segmentação binária, uma vez que muitos *frameworks* e funções de perda são otimizados e mais estáveis para problemas em que o objeto de interesse é pequeno. Ainda, essa estrutura oferece uma maior flexibilidade caso seja possível a obtenção de dados rotulados de células saudáveis, permitindo a expansão para mais objetos no futuro.

3.3.3 Geração de máscaras e Data Augmentation

A preparação dos dados para o formato exigido pelo modelo iniciou-se com a conversão dos contornos de célula das 18 imagens anotadas pelos patologistas em máscaras binárias. Foi desenvolvido um algoritmo para ler os polígonos do arquivo texto, desenhar seus contornos e preenche-los em branco (valor de pixel igual à 255), representando a classe 1, sob uma imagem

completamente preta (valor de pixel igual à 0), representando a classe 0 ou *background*. Essas máscaras formam um par com a imagem original para constituir uma amostra de treinamento.

A diversidade na origem das amostras, coletadas de múltiplos animais em momentos distintos, gerou um conjunto de máscaras de segmentação com dimensões não uniformes. Diante da exigência de uma entrada com altura e largura predefinidas pela primeira camada do modelo, a padronização das dimensões foi indispensável. Embora abordagens de redimensionamento como *upscaling* e *downscaling* sejam comuns para essa finalidade, sua aplicação é inviável em máscaras binárias. Tais métodos, fundamentados em interpolação, poderiam corromper os valores de pixel, que devem ser estritamente 0 ou 255, nesse caso. Consequentemente, optou-se pela aplicação de *padding*, adicionando-se uma moldura preta para adequar todas as imagens ao formato da maior imagem do conjunto, assegurando assim a integridade dos dados e a compatibilidade com a arquitetura da rede.

Para viabilizar o treinamento e a avaliação imparcial do modelo, o conjunto de dados foi estratificado em três subconjuntos: treino, validação e teste. Adotou-se a proporção de 70% para treino, 15% para validação e 15% para teste, e as imagens foram escolhidas de forma aleatória resultando em 12, 3 e 3 imagens para cada conjunto, respectivamente.

Subsequentemente, técnicas de *Data Augmentation* foram aplicadas sob o conjunto de treino para mitigar a variância e fragilidade do modelo treinado. No entanto, esse processo deve ser realizado com extrema cautela, uma vez que transformações inapropriadas podem inadvertidamente introduzir artefatos ou distorções que alteram a informação clínica essencial para o diagnóstico, levando o modelo a aprender padrões irrelevantes ou até incorretos. Dessa forma, optou-se por abordagens mais simples, especificamente por operações de espelhamento vertical, horizontal, e vertical e horizontal foram aplicadas sob as imagens e máscaras para transformar uma amostra em três diferentes, possibilitando a quadruplicação do conjunto de dados. Ao final desse processo, o conjunto de treino originalmente de tamanho 12 passou a ter 48 imagens.

3.3.4 Treinamento do modelo

O treinamento do modelo U-Net foi conduzido em uma estação de trabalho equipada com uma GPU NVIDIA GeForce RTX 1650 Ti com 4GB GDDR6. Utilizou-se o framework PyTorch, versão 1.9.1. Para a otimização dos pesos da rede, foi empregado o otimizador Adam, (KINGMA, 2014) com uma taxa de aprendizado inicial de 1×10^{-5} . A função de perda selecionada foi a combinação da *Dice Loss* com a BCE para equilibrar a segmentação em nível de pixel e de região. O modelo foi treinado com um tamanho de *batch* de 2 imagens, dado as limitações da máquina. A performance foi monitorada a cada época no conjunto de validação utilizando o Coeficiente de *Dice* juntamente com a técnica de *Early Stopping* para evitar o *overfitting* com um *patience* baixo de 10 épocas, uma vez que o conjunto de dados é pequeno.

3.3.5 Inferência

A inferência foi realizada com um *script* disponibilizado no mesmo repositório do treinamento. Após a correção de tamanho da imagem com um *padding*, o modelo treinado é capaz de prever uma máscara binária que representa as células em proliferação.

A partir dessa máscara inicial, pode-se afirmar que os processos subsequentes são aplicados de forma análoga à rotina de processamento de imagens. A iniciar-se pela Análise de Componentes e Operações de Fechamento, que foram aplicadas para filtrar pequenos ruídos e preencher lacunas nos objetos segmentados, respectivamente. Seguido pela segmentação, para a obtenção dos contornos e transformação da máscara binária em contornos que podem ser utilizados para a contagem das células

3.4 Validação dos processos

Esta seção descreve a validação das metodologias de processamento de imagens e aprendizado de máquina que foi realizada em duas etapas: Geração de Métricas com base nas máscaras preditas e Análise Estatística dos resultados.

3.4.1 Geração de métricas

A geração de métricas ocorreu pela comparação das máscaras preditas pelos algoritmos com o conjunto de anotações de referência (*ground truth*) presentes na base de dados levantada.

Primeiramente, foi feito o *download* dos objetos segmentados no formato YOLOv8. Esse formato foi escolhido pela sua simplicidade e facilidade de manipulação, uma vez que as anotações em uma imagem são representadas em um único arquivo texto. No arquivo, cada linha corresponde a um objeto, contendo a informação da classe seguido de um conjunto, de tamanho variável, de pontos cartesianos x e y que representam o polígono que delimita a célula.

Em sequência, foi desenvolvido um algoritmo para iterar sobre as células encontradas pela segmentação e comparar com os polígonos do arquivo texto. A performance do modelo foi quantificada por meio da classificação de cada detecção como:

- Verdadeiro Positivo (*True Positive* - TP) - Ocorre quando há pelo menos um ponto de interseção entre o contorno predito e um contorno de referência correspondente;
- Falso Positivo (*False Positive* - FP) - Ocorre quando o contorno predito não possui nenhum ponto de interseção com qualquer contorno de referência;
- Falso Negativo (*False Negative* - FN) - Ocorre quando um contorno de referência não possui nenhum ponto de interseção com qualquer contorno predito; e

- Verdadeiro Negativo (*True Negative* - TN) - Ocorre quando a classe negativa é corretamente identificada. Nesse caso, o *background* pode ser considerado, porém esse estudo não quantificou essa contagem.

A partir dessas contagens, podem ser calculadas as métricas de Precisão e *Recall* para avaliar objetivamente a acurácia e a robustez do método (FAWCETT, 2006). A Precisão mede a proporção de detecções corretas entre todas as detecções realizadas e é formalizada pela equação 3.1:

$$\text{Precisão} = \frac{TP}{TP + FP} \quad (3.1)$$

Já a *Recall* avalia a capacidade do algoritmo de identificar todas as células realmente presentes e é formalizada pela equação 3.2:

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3.2)$$

Outras métricas de desempenho consolidadas na literatura para a tarefa de segmentação, como o Coeficiente de *Dice* e a Interseção sobre União (IoU), baseiam-se na sobreposição dos objetos preditos com os dados de referência. Tais métricas, contudo, não foram consideradas para avaliar a performance do algoritmo. A justificativa é que o objetivo final da rotina é a contagem das células, um propósito para o qual a sobreposição de qualquer parte do contorno com o *ground truth* já é considerada suficiente.

3.4.2 Análise Estatística

Os valores das métricas são informativos por si só, mas, ao lidar com um conjunto de dados, as Análises Estatísticas tornam-se fundamentais. Elas permitem não apenas a quantificação da performance geral das abordagens, mas também a comparação da variabilidade dos resultados. Assim, fornecem uma base sólida para a discussão dos achados, bem como para a identificação de vieses ou áreas de melhoria. Para a realização destas análises neste trabalho, foram selecionadas a Média e o Desvio Padrão. A Média Aritmética (ou simplesmente média) é a medida de tendência central mais amplamente utilizada, representando o valor central ou "esperado" de uma distribuição. Ela é calculada através do somatório de todas as observações dividido pelo número total de elementos do conjunto.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (3.3)$$

Na equação 3.3, x_i representa cada valor individual de desempenho (como a precisão de cada teste) e n é o número total de testes realizados. A intuição por trás do uso da média neste contexto é obter um valor que sintetize a performance global do modelo, levando em

consideração a magnitude de cada resultado obtido. Diferente da mediana, que considera apenas a posição, a média é sensível a todos os valores do conjunto de dados, o que a torna a base matemática necessária para o cálculo de métricas de dispersão mais complexas, como a variância e o desvio padrão.

Consequentemente, o Desvio Padrão (σ ou S) torna-se a medida natural para acompanhar a média, pois ele quantifica a dispersão dos dados exatamente em relação a ela (μ ou $\bar{x}$). Uma baixa magnitude de desvio padrão indica que os resultados dos testes tendem a ser muito próximos da média calculada, sugerindo estabilidade no modelo. Por outro lado, uma alta magnitude sinaliza que os dados estão espalhados, refletindo instabilidade ou heterogeneidade na performance. A fórmula do desvio padrão é obtida através da raiz quadrada da variância; para uma amostra de n observações, ele é calculado de acordo com a equação 3.4:

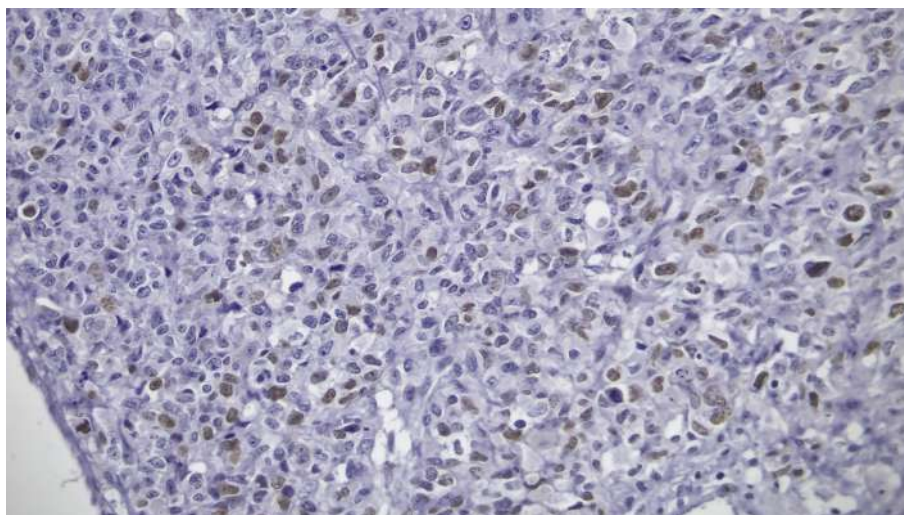
$$S = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}} \quad (3.4)$$

Na equação 3.4, observa-se a dependência direta do termo $\bar{x}$ (a média amostral definida anteriormente). O divisor $n - 1$ (graus de liberdade de Bessel) é empregado para fornecer um estimador não enviesado da variância populacional. Essas análises foram aplicadas aos valores de Precisão e *Recall* para melhorar a compreensão dos resultados e viabilizar a discussão dos mesmos.

4 Resultados

Os resultados deste trabalho são apresentados em três seções principais, alinhadas com a metodologia desenvolvida. Primeiramente, são detalhados os resultados do processamento de imagens e em seguida, do ML. Por fim, uma seção de discussão é desenvolvida com o intuito de comparar os resultados das duas abordagens, suas vantagens, desvantagens e limitações. Para todas as seções, a figura 12 será utilizada para demonstrar os resultados das técnicas aplicadas, uma vez que pertence ao conjunto de teste:

Figura 12 – Imagem de exemplo



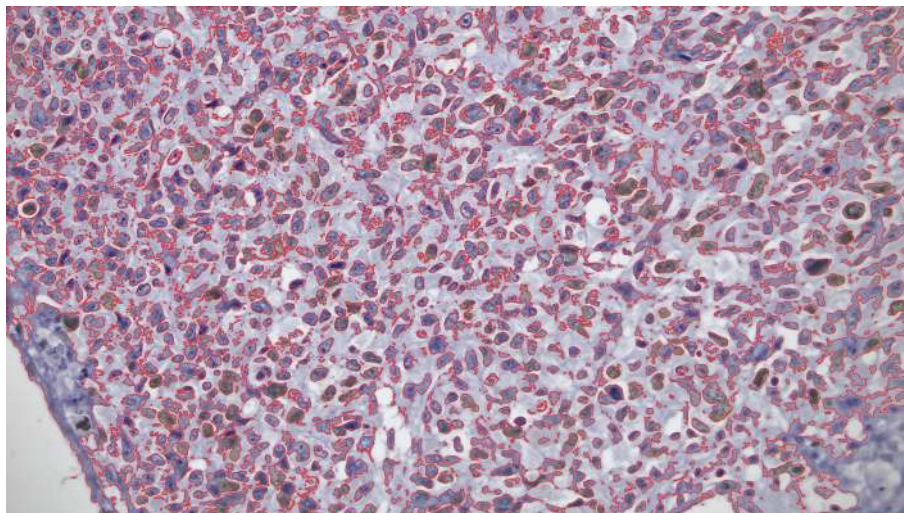
Fonte: Elaborada pelo autor

4.1 Experimentos

Os primeiros testes visaram a segmentação diretamente sobre a imagem original a partir da ideia de encontrar bordas e contornos com base na diferença entre pixels adjacentes. As principais técnicas utilizadas foram a limiarização global e o algoritmo *Watershed*.

A figura 13 exibe na cor vermelha os objetos encontrados pelo experimento. Observa-se uma grande dificuldade na separação das células do *background* e uma tendência prejudicial de aglomerar células saudáveis e em proliferação em um único contorno.

Figura 13 – Contornos detectados pelo algoritmo Watershed

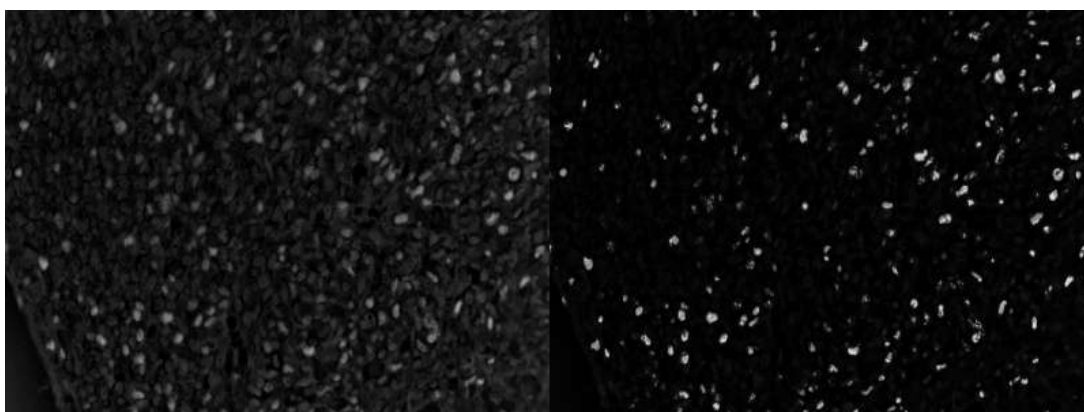


Fonte: Elaborada pelo autor

Os resultados iniciais obtidos pela segmentação pura evidenciaram a necessidade de uma mudança de abordagem. A partir disso, os próximos experimentos foram baseados completamente na informação de cor para realizar qualquer processamento inicial.

Em relação à cor, a natureza das imagens histopatológicas de sobreposição física dos corantes utilizados é um grande fator a ser considerado durante a manipulação de imagens desse tipo. A técnica de *Color Deconvolution* permite a separação quantitativa das contribuições individuais de cada agente de coloração utilizados para realçar estruturas celulares e teciduais, no caso do das lâminas de OSA para Ki-67, os corantes DAB e HE. No entanto, apesar dessa técnica mitigar as variações inerentes ao processo de preparação de lâminas, ela não se mostrou suficiente para a separação completa das células. Optou-se por utilizar a técnica de *White Balancing* combinada à saída anterior para ressaltar as células mais marcadas, porém isso não foi possível sem uma perda substancial de informação, como pode ser observado na figura 14:

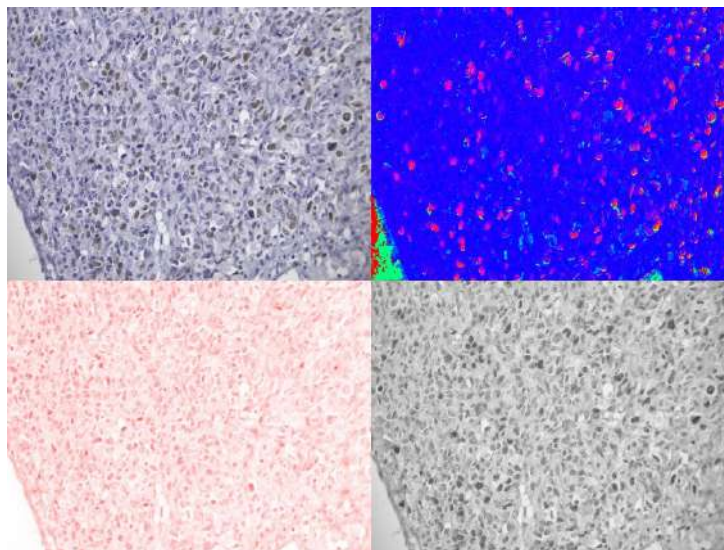
Figura 14 – Separação cromática (à esquerda) e balanceamento de branco (à direita)



Fonte: Elaborada pelo autor

Os próximos experimentos foram chaves na construção da rotina de processamento de imagens e evidenciaram as técnicas que fariam parte da *pipeline final*. A principal ideia foi tratar de forma mais direta o problema da separação de cor através da definição de dois intervalos: um para as cores marrom, evidenciando as células em proliferação, e outro para o azul escuro, evidenciando as células saudáveis. Para tornar o processo mais robusto, primeiro foi realizada a conversão da imagem original para o espaço de cor HSV.

Figura 15 – Imagem original e Canais de cor HSV separados

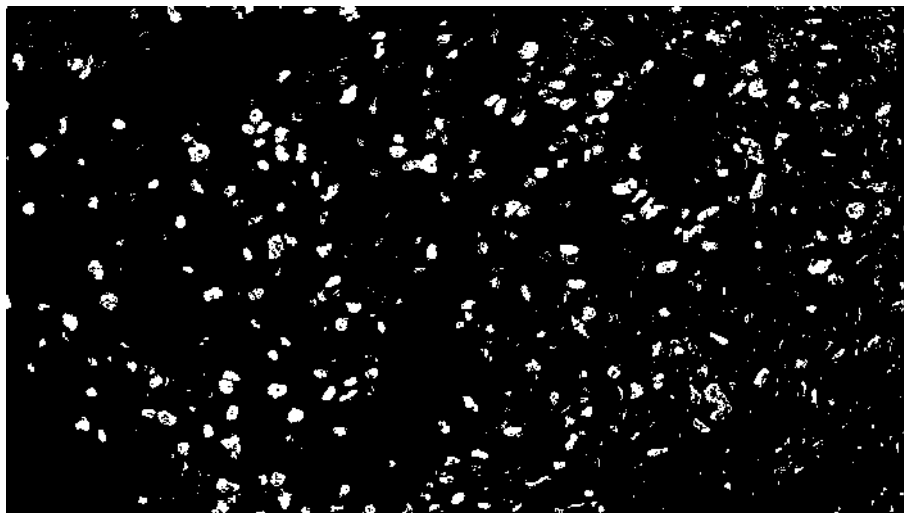


Fonte: Elaborada pelo autor

Com base nas imagens de matiz, saturação e valor, apresentadas respectivamente na figura 15, foi possível encontrar valores mínimos e máximos utilizadas para a Limirização. A correta definição desses valores é de extrema importância para o resultado final, sendo essa a etapa mais crítica do processo como um todo. Isso se dá pelo fato de que se o intervalo for muito abrangente, objetos indesejados e ruídos serão limiarizados junto com os objetos de interesse. Por outro lado, se o intervalo for pequeno e muito específico para uma única imagem, ele não generalizará para as demais e acarretará em uma grande perda de informação.

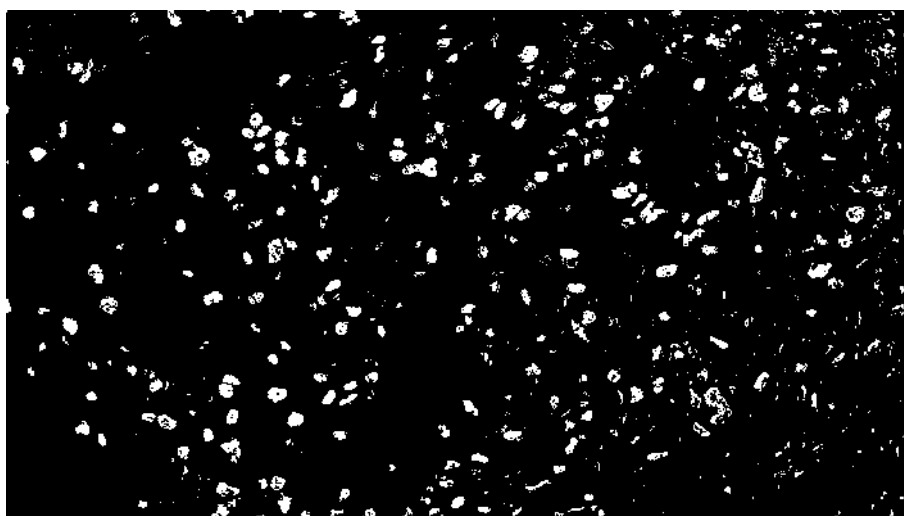
Os intervalos encontrados foram de $H=129$, $S=15$, $V=60$ (mínimo) e $H=190$, $S=90$, $V=170$ (máximo) para a cor marrom, e de $H=120$, $S=55$, $V=115$ (mínimo) e $H=130$, $S=180$, $V=180$ (máximo) para a cor azul escuro. A Limirização a partir desses valores pode ser observada na figura 16 para as células em proliferação e a figura 17 para as células saudáveis. A partir das máscaras binárias iniciais, as células saudáveis e em proliferação já foram separadas e, dessa forma, as técnicas subsequentes também foram aplicadas individualmente.

Figura 16 – Células em proliferação isoladas por limiarização



Fonte: Elaborada pelo autor

Figura 17 – Células saudáveis isoladas por limiarização



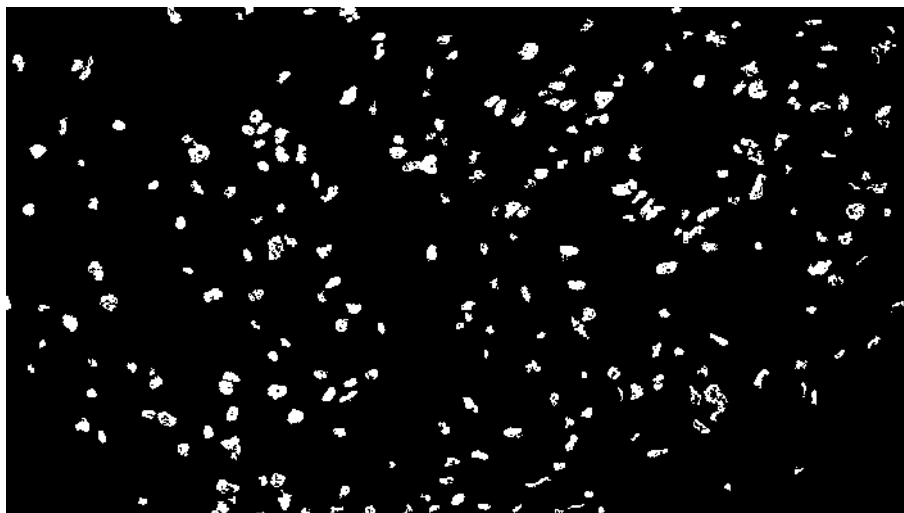
Fonte: Elaborada pelo autor

Por se tratarem de intervalos relativamente abrangentes na limiarização da imagem, é notável que inúmeros ruídos fazem parte da máscara inicial. Tais perturbações consistem em pequenas flutuações de intensidade de pixel que não correspondem à estrutura real de interesse. Para mitigar esse problema, foram aplicados filtros baseados em Análise de Componentes Conectados e operações morfológicas, especialmente a operação de fechamento, respectivamente.

Em primeiro lugar, após a aplicação da técnica estatística, foram realizados testes empíricos objetivando encontrar o valor ideal para o tamanho mínimo de área dos objetos identificados. Esse processo iniciou-se com um valor pequeno que foi gradativamente aumentado até que objetos de interesse comesçassem a ser descartados, chegando em um valor final de

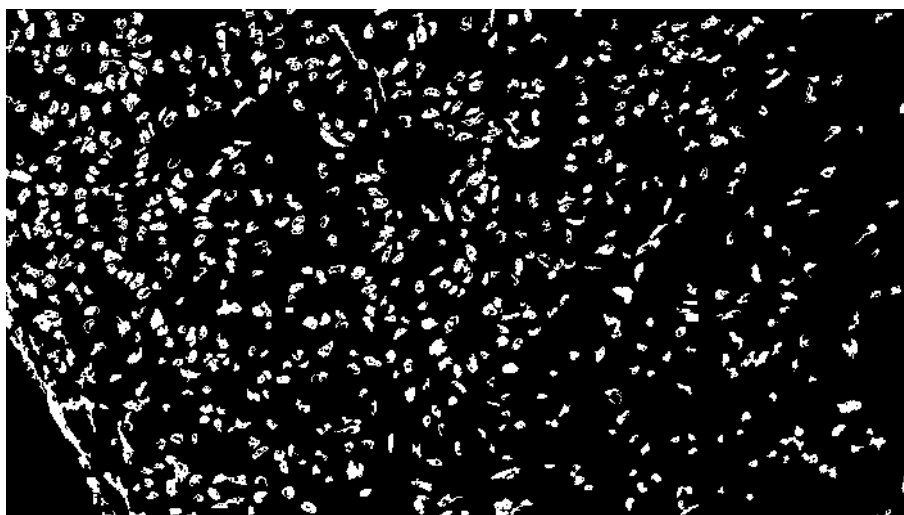
1000 pixels. A figura 18 evidencia as células em proliferação após o filtro de área e a figura 19 as células saudáveis.

Figura 18 – Células em proliferação filtradas



Fonte: Elaborada pelo autor

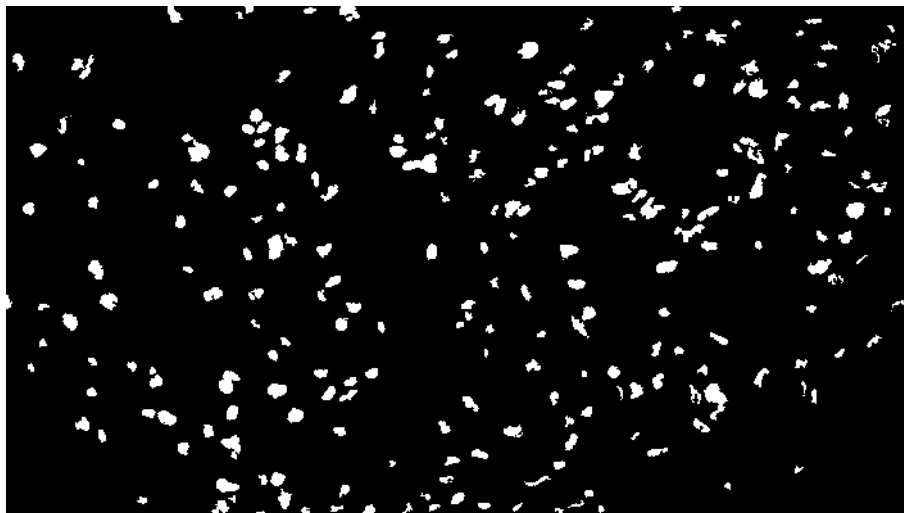
Figura 19 – Células saudáveis filtradas



Fonte: Elaborada pelo autor

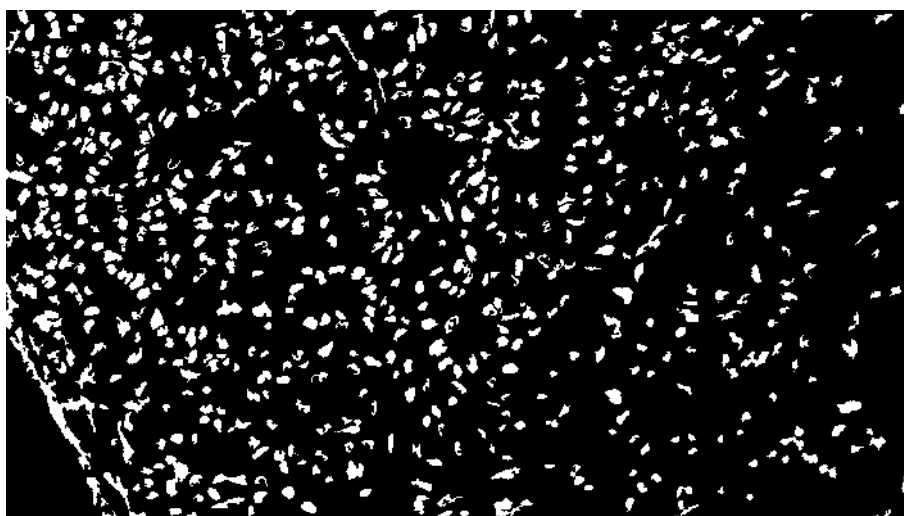
Em seguida, foram aplicadas operações de fechamento objetivando o preenchimento de pequenos buracos e a suavização dos contornos dos objetos sem alterar significativamente sua área original. Por se basear em um elemento estruturante (*kernel*), que pode assumir formas e tamanhos variáveis, foi necessário um processo empírico similar ao descrito anteriormente que resultou na escolha de um elemento com formato elíptico e de tamanho 9x9. A figura 20 evidencia as células em proliferação preenchidas e a figura 21 as células saudáveis.

Figura 20 – Células em proliferação preenchidas



Fonte: Elaborada pelo autor

Figura 21 – Células saudáveis preenchidas

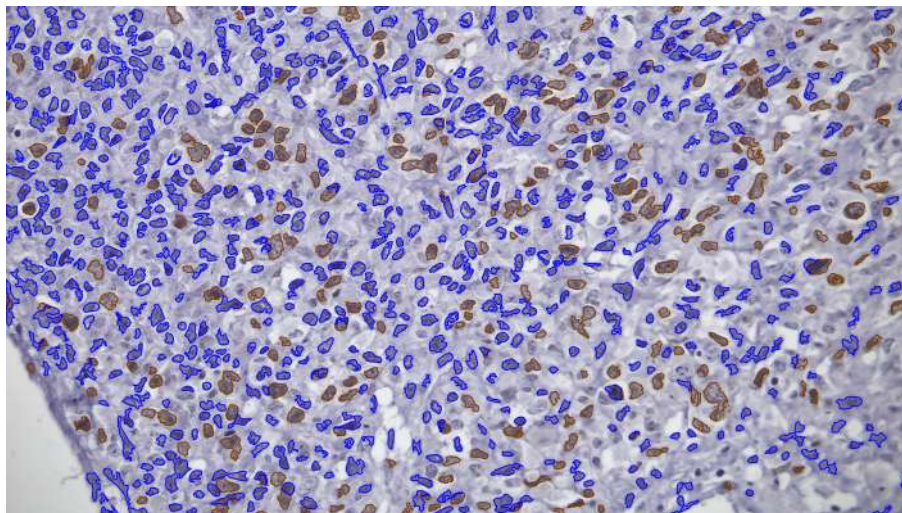


Fonte: Elaborada pelo autor

Finalmente, a segmentação foi realizada a partir do algoritmo de rastreamento de bordas e análise topológica. Uma vez que contornos podem ser expresso de múltiplas formas em uma imagem, por exemplo aninhados um dentro do outro, foi de extrema importância especificar para o algoritmo retornar apenas contornos externos, sem relações de parantesco.

O resultado dessa operação é uma lista de contornos externos das células. Cada contorno é representado por pontos x e y (em relação ao canto superior esquerdo da imagem), um para cada vértice do polígono delimitador. Na figura 22, as células encontradas pelo algoritmo podem ser visualizadas sobre a imagem original, sendo as células saudáveis representadas na cor azul e as células em proliferação na cor marrom.

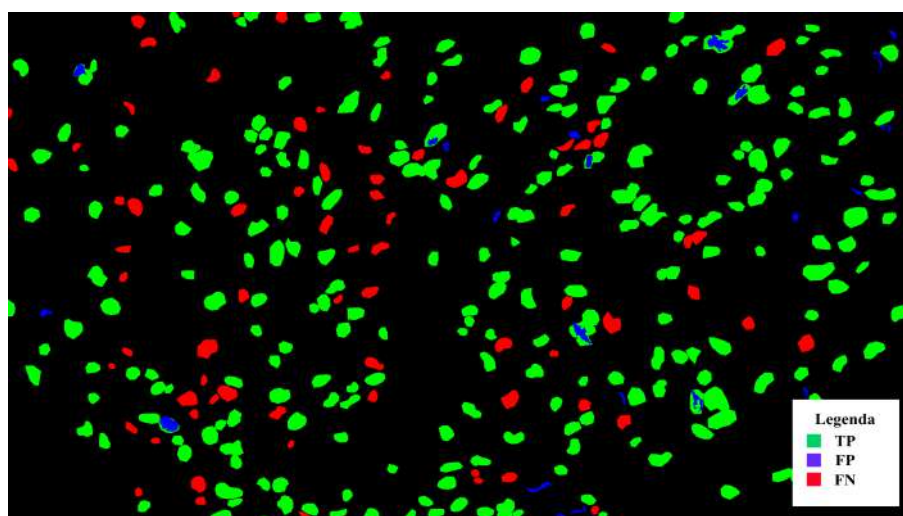
Figura 22 – Células em proliferação e saudáveis segmentadas



Fonte: Elaborada pelo autor

A contagem final das células de cada classe é realizada pela verificação da quantidade de contornos. Esta etapa encerra a rotina de processamento de imagens, visto que os dados de contagem permitem a obtenção do índice de proliferação do tecido. Todas as operações foram modularizadas em um *script* abrangente, que exige apenas uma imagem de entrada do usuário e este será referido como *pipeline* de processamento de imagens a partir de agora.

Para a validação dos resultados, a comparação entre os contornos encontrados com as células anotadas resulta na contagem dos TP, FP e FN , utilizados para o cálculo de métricas, e uma imagem para facilitar a visualização dos erros e acertos. A figura 23 evidencia a classificação das detecções de células em proliferação e a figura 24 das células saudáveis.

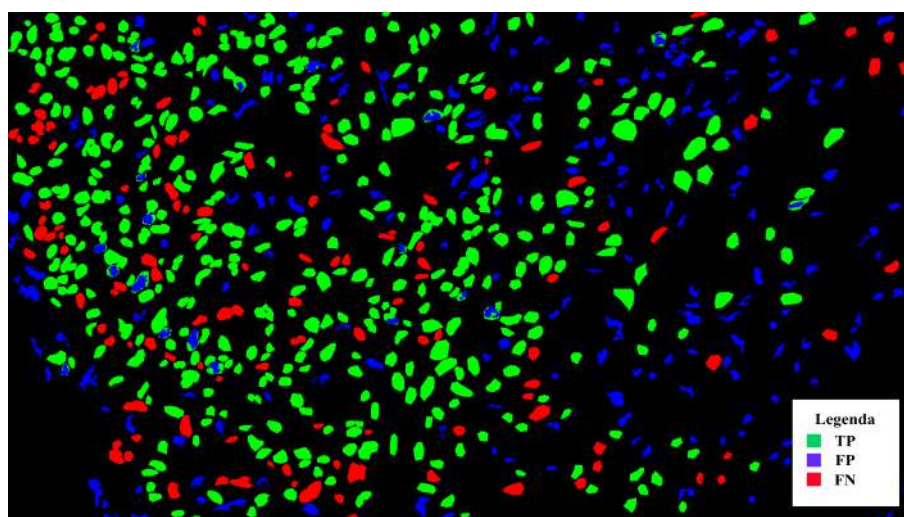
Figura 23 – Classificação das células em proliferação - *pipeline*

Fonte: Elaborada pelo autor

Observe que algumas células marcadas como FP (em azul) podem aparecer em uma intersecção ou completamente dentro de uma célula marcada como TP (em verde). Isso acontece quando mais de um contorno é encontrado para a mesma célula, seja por uma descontinuidade ou fraca coloração de um trecho do objeto, o que é corretamente penalizado pelo *script*. É também importante ressaltar que células TP aglomeradas, ou seja, duas ou mais unidas e/ou sem uma separação clara, foram consideradas apenas uma vez.

Para as células em proliferação nessa imagem, os patologistas contaram 296 células e a rotina de processamento de imagens encontrou 245, sendo 221 TP e 23 FP (75 FN). Dessa forma, a precisão foi de 90,57% e a revocação de 74,66%.

Figura 24 – Classificação das células saudáveis - *pipeline*



Fonte: Elaborada pelo autor

Para as células saudáveis, o patologista contou 492 células e a rotina de processamento de imagens encontrou 596, sendo 374 TP e 222 FP (118 FN). Dessa forma, a precisão foi de 62,75% e a revocação de 76,01%. Esse valor significativamente mais baixo de precisão é justificado pela concentração de células saudáveis no canto inferior direito da imagem, o qual apresentou borramento no momento de digitalização.

Para a aplicação da ferramenta na rotina do patologista, o resultado mais importante seria o cálculo do índice de proliferação, que para essa imagem foi previsto em 29,13%. Pode-se afirmar que houve uma subestimação do índice real, de 37,56%, configurando um erro absoluto de 8,43%. Essa diferença se justifica principalmente pelo baixo *Recall* de ambas as classes de células.

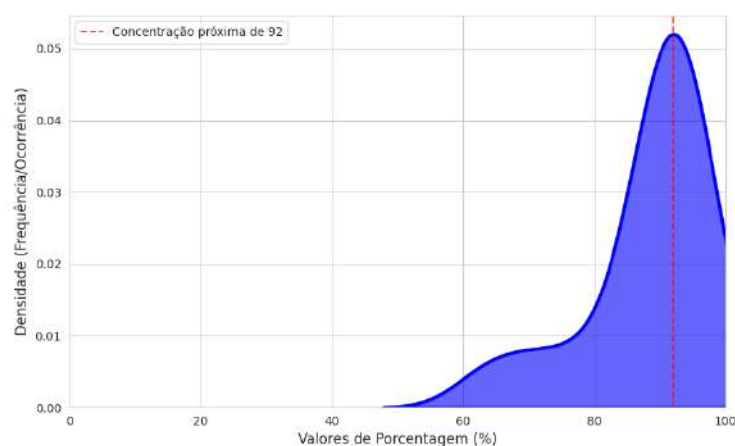
Para obter os resultados de forma mais geral, todas as 18 imagens da base de dados foram submetidas à contagem de células em proliferação através da *pipeline* desenvolvida. As máscaras preditas foram validadas e os valores de Média e Desvio Padrão para cada uma das métricas são apresentados no quadro 2.

Quadro 2 – Resultados da *pipeline*

	Precisão	<i>Recall</i>
Média	88,3	63,96
Desvio padrão	9,3	16,02

Em relação aos valores de precisão, a análise da distribuição de densidade das porcentagens na figura 25 demonstra um comportamento marcadamente unimodal, com os dados se aglomerando em torno de um único ponto central de alta frequência. Esta característica evidencia uma forte tendência de consistência nos resultados obtidos.

Figura 25 – Distribuição de Densidade dos valores de Precisão

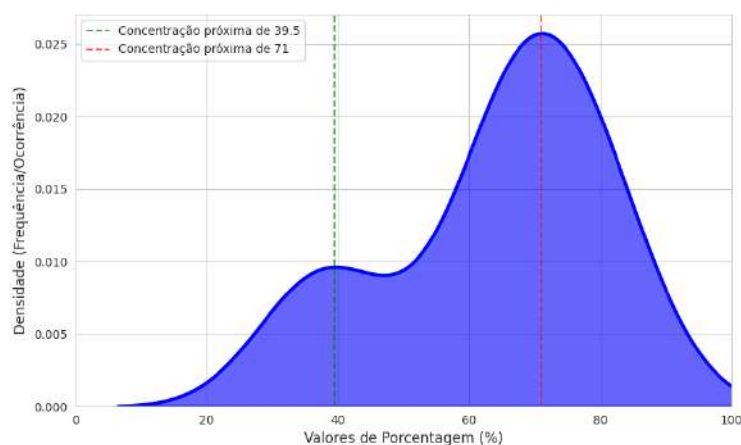


Fonte: Elaborada pelo autor

O pico da distribuição está localizado em aproximadamente 92%, indicando que este é o valor de precisão mais recorrente e representativo do conjunto de dados. A grande maioria das ocorrências concentra-se em uma faixa adjacente a este pico, o que reforça a previsibilidade e a estabilidade do desempenho medido. A alta precisão é justificada pela característica marcante de cor das células em proliferação, sendo essa forte assinatura cromática um diferenciador primário na classificação, o que torna FP mais raros.

Em relação aos valores de *Recall*, o alto valor de desvio padrão observado aliado a análise da distribuição de densidade dos valores na figura 26 revela um comportamento bimodal, caracterizado pela presença de dois picos distintos de concentração.

Figura 26 – Distribuição de Densidade dos valores de *Recall*



Fonte: Elaborada pelo autor

O primeiro centro de concentração, de menor proeminência, localiza-se em torno de 39,5%. Este agrupamento se deve às imagens em menor quantidade que não pertenciam ao "padrão ouro" de preparação das lâminas. Essa constatação indica que de fato uma menor clareza na separação entre células adjacentes e entre células e estroma foi extremamente prejudicial para o processo. Ainda mais, a volatilidade em termos de cor, que variaram tanto para tonalidades mais claras quanto para mais escuras, foi um fator crucial para queda da métrica de *Recall*, principalmente pela grande quantidade de pixels de células que se encontravam fora dos intervalos de cor definidos durante a etapa de separação.

O segundo e mais significativo pico de densidade ocorre em aproximadamente 71%. Esta região representa o modo dominante da distribuição, onde se encontra a maior frequência de ocorrências: lâminas do "padrão ouro", majoritárias na base de dados. A maior altura deste pico indica que para essas características, a rotina de processamento de imagens consistentemente retornou um número elevado de células em proliferação comparada com as anotações do patologista. Nesses casos, os poucos FN são justificados pelas células que foram coradas mais fracamente, ou então foram descartadas por filtros de área, principalmente. Essa é uma das métricas mais importantes, pois está diretamente relacionada à contagem das células.

4.2 Classificação

Esta seção dedica-se à apresentação e à análise dos resultados obtidos com o modelo U-Net desenvolvido para a tarefa de segmentação de células em proliferação. A avaliação inicial da eficácia do modelo concentra-se na validação do seu processo de treinamento e na sua capacidade de convergência. Para isso, o comportamento do modelo foi rigorosamente monitorado a cada época, comparando seu desempenho nos conjuntos de dados de treinamento

e de validação. A análise das curvas de aprendizado é fundamental para diagnosticar a estabilidade do treinamento e verificar se o modelo generaliza adequadamente para dados não vistos, evitando o *overfitting*. A figura 27 ilustra essa dinâmica, apresentando a evolução da função de perda ao longo das 50 épocas de treinamento, tanto para o conjunto de treinamento quanto para o de validação.

Figura 27 – Perda do Treinamento vs. Validação



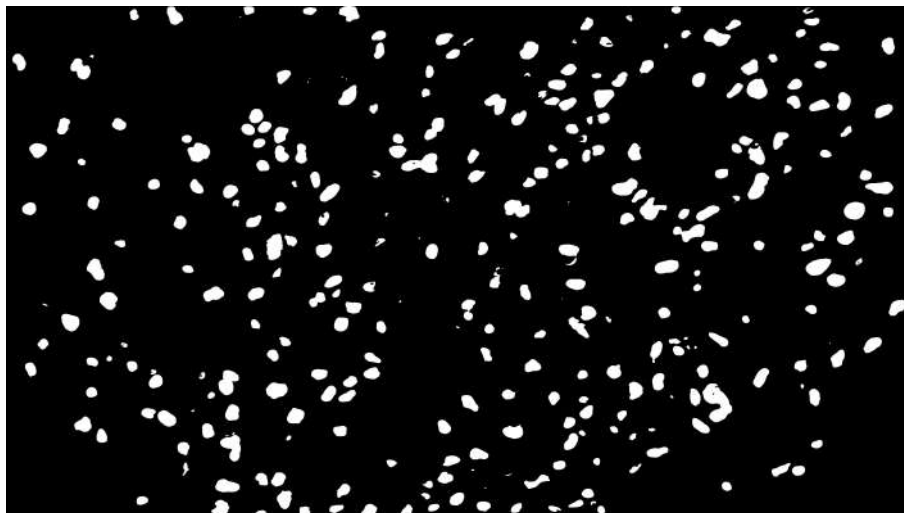
Fonte: Elaborada pelo autor

Observa-se uma rápida convergência inicial, onde tanto a perda de treinamento quanto a de validação decrescem abruptamente nas primeiras épocas, indicando um aprendizado efetivo. Após essa fase inicial, a curva de validação atinge um patamar baixo e demonstra notável estabilidade, mantendo-se consistentemente em um valor reduzido a partir, aproximadamente, da 15ª época. Este comportamento é um forte indicativo de que o modelo generaliza bem para dados não vistos e não está sofrendo de *overfitting*, já que a perda de validação não apresenta uma tendência de aumento. Em contraste, a perda de treinamento, embora também convirja para valores baixos, exibe uma volatilidade considerável, com picos e vales frequentes ao longo das épocas subsequentes. Essa instabilidade no treinamento, justaposta à estabilidade da validação, pode sugerir fatores como variância introduzida por um tamanho de lote *batch size* pequeno de tamanho 2, valor esse escolhido pelas limitações de *hardware*. Os melhores resultados foram obtidos na época 40, com um Coeficiente de *Dice* próximo de 0.8.

Após o treinamento do modelo, sua capacidade de gerar previsões de máscaras binárias foi avaliada qualitativamente para inspecionar o resultado do aprendizado. O processo de otimização permitiu ao modelo aprender a identificar os padrões e características complexas que definem as células em proliferação, distinguindo-as efetivamente do fundo da imagem e das outras células. Mais do que simplesmente convergir em termos de métricas, o modelo desenvolveu a habilidade prática de traduzir a imagem de entrada em um mapa de segmentação

coerente, delineando as áreas de interesse. Para ilustrar essa capacidade adquirida, a figura 28 exibe um exemplo visual retirado do conjunto de teste.

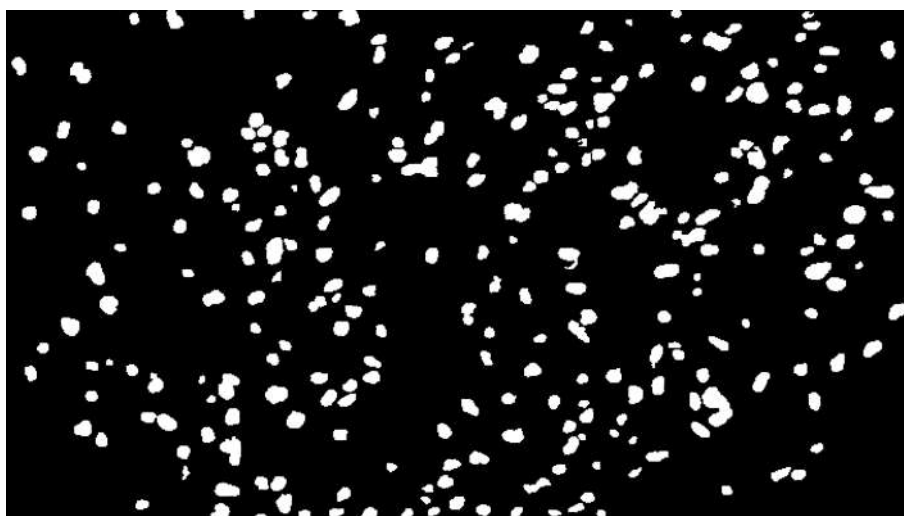
Figura 28 – Predição pura do modelo treinado



Fonte: Elaborada pelo autor

As predições brutas frequentemente continham pequenos pontos brancos indesejados, caracterizados como ruído de falso positivo e que seriam prejudiciais à contagem. Para refinar esses resultados e aumentar a robustez da segmentação, foi aplicada um filtro de área baseado na Análise de Componentes Conectados como pós-processamento, da mesma forma que na *pipeline* e uma iteração da operação de Fechamento para cobrir buracos. A figura 29 apresenta o resultado final deste processo, exibindo a máscara binária após a aplicação do filtro.

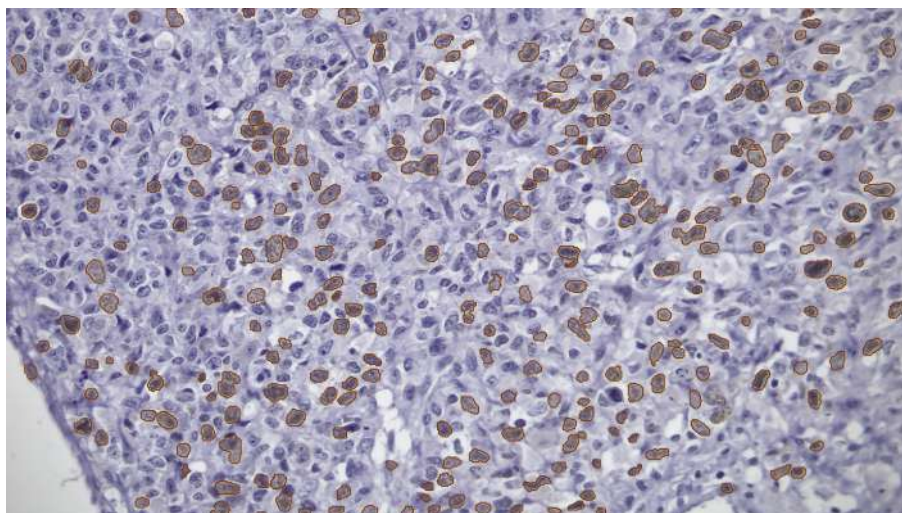
Figura 29 – Predição filtrada



Fonte: Elaborada pelo autor

Antes da validação final, a máscara binária final passou pelo processo de extração dos contornos seguindo o mesmo protocolo já discutido anteriormente. Os resultados da segmentação podem ser compreendidos na figura 30, com os contornos preditos sendo desenhados na cor marrom sobre a imagem original.

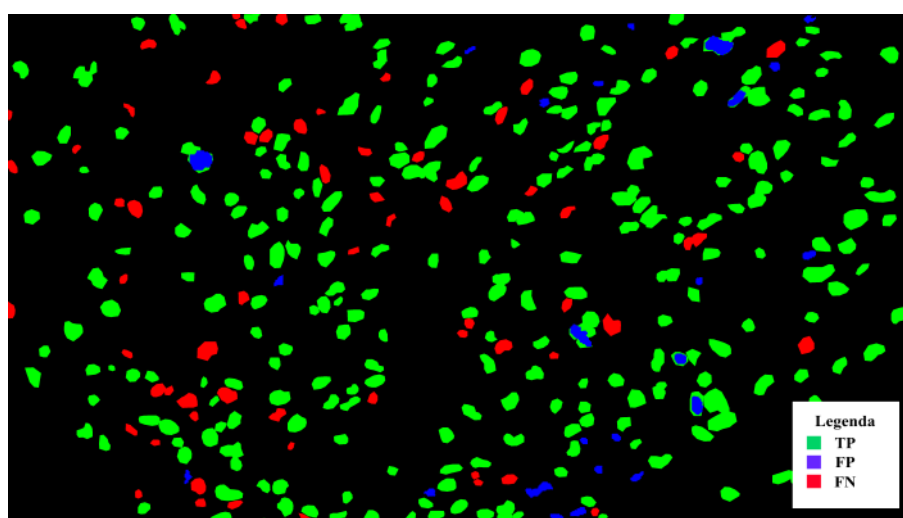
Figura 30 – Células em proliferação segmentadas - U-Net



Fonte: Elaborada pelo autor

Por fim, sua performance é avaliada sob a perspectiva quantitativa. De forma análoga à rotina de processamento de imagens, os contornos preditos foram submetidos ao *script* de comparação com os contornos das células anotadas. A figura 31 ilustra a classificação:

Figura 31 – Classificação das células em proliferação - U-Net



Fonte: Elaborada pelo autor

Observa-se que o modelo U-Net aprendeu uma característica prejudicial à rede, que é a tendência de juntar células adjacentes em um único objeto. A visualização da classificação na

figura 31 deixa claro que, embora o modelo delimite as áreas de interesse, ele nem sempre consegue separar individualmente células que estão próximas, por mais que exista um espaçamento visível.

A contagem de 255 células, sendo 232 TP e 25 FP (64 FN), configurando uma Precisão de 90% e um *Recall* de 78,3%, demonstram a capacidade de aprendizado de modelo a partir dos dados fornecidos durante o treinamento.

Para validar o aprendizado da rede, o modelo foi avaliado nas outras duas imagens de teste e o quadro 3 indica os resultados obtidos no conjunto de teste em sua totalidade.

Quadro 3 – Resultados da U-Net

	Precisão	<i>Recall</i>
Média	94,2	75,9
Desvio padrão	5,93	3,39

Dado sua natureza, é empírico que o modelo obtenha maior capacidade de generalização para imagens de entrada diferentes, o que é observado pelo baixo desvio padrão nas métricas. No entanto, qualquer afirmação sobre a performance do modelo não pode descartar o fato de que o teste foi realizado em apenas 3 imagens. Apesar da variedade de origem, animal e coloração das lâminas, a baixa quantidade desse tipo de dado não é capaz de representar todas as variações necessárias para consolidar o modelo na rotina do patologista.

4.3 Discussão

A análise comparativa dos resultados obtidos pela *pipeline* de processamento de imagens tradicional e pela abordagem baseada em DL com a arquitetura U-Net revela uma complexa permuta entre robustez, precisão e a natureza fundamental dos erros de cada método. Ambas as estratégias visaram a segmentação de células em proliferação, mas seus perfis de desempenho divergem significativamente, especialmente em termos de estabilidade e capacidade de generalização frente às variações intrínsecas dos dados histopatológicos.

Para fornecer uma comparação direta entre o resultado da *pipeline* e o resultado do modelo treinado, o quadro 4 apresenta as métricas das duas abordagens de forma resumida para a figura 12:

Quadro 4 – Comparativo células em proliferação *pipeline* e U-Net para figura 12

	Contagem	TP	FP	FN	Precisão	<i>Recall</i>
Pipeline	257	221	23	75	90,57%	74,66%
U-Net	245	232	25	64	90%	78,3%

Apesar da base de dados limitada e a dificuldade na separação de células, o modelo foi capaz de obter resultados equiparados à *pipeline*. Conforme demonstrado no quadro 4, para a imagem de exemplo, o modelo U-Net alcançou uma Precisão de 90% e um *Recall* de 78,3% e, valores muito próximos aos 90,57% de Precisão e 74,66% de *Recall* da *pipeline*.

Esta discrepância quantitativa é a chave para entender a principal vantagem da U-Net. A instabilidade da *pipeline* de processamento, evidenciada pela distribuição bimodal do *Recall* (figura 26), confirma sua natureza heurística e sua dependência crítica de parâmetros pré-definidos. A *pipeline* depende de intervalos de cor fixos no espaço HSV, que se mostraram eficazes apenas para as lâminas "padrão ouro", mas falharam substancialmente em amostras com variações de coloração ou preparação. Em contrapartida, a U-Net, ao aprender características hierárquicas diretamente dos dados, demonstrou uma capacidade de generalização intrinsecamente superior. A baixa variância em seus resultados no conjunto de teste sugere que o modelo aprendeu a identificar células em proliferação com base em características texturais e morfológicas mais complexas do que a simples assinatura cromática, tornando-o robusto a variações que seriam prejudiciais à abordagem por limiarização.

No entanto, uma análise puramente métrica oculta as diferenças qualitativas nos tipos de erro. O principal modo de falha da *pipeline* de processamento é a omissão (FN), resultante da rigidez dos seus limiares de cor e do filtro de área, que descartava células mais fracamente coradas ou de menor tamanho. O resultado, como visto na figura 23, é uma contagem subestimada (alto número de FNs), mas com boa separação individual das células que são detectadas. Por outro lado, a U-Net introduz um modo de falha distinto e problemático: a tendência de aglomerar células adjacentes em um único contorno mesmo que haja uma clara separação entre elas (figura 31). Embora o modelo identifique corretamente a região contendo células em proliferação, a fusão dos objetos impacta diretamente a contagem final, que é o objetivo primário para o cálculo do índice de proliferação.

Ademais, a comparação é limitada pelo escopo de cada experimento. A *pipeline* de processamento, apesar de suas falhas de *Recall* e da baixa precisão para células saudáveis (62,75% no exemplo, figura 24), foi concebida como uma solução completa, fornecendo a contagem de ambas as classes celulares necessárias para o índice de proliferação. A U-Net, na forma como foi treinada e avaliada, focou-se exclusivamente nas células em proliferação, não fornecendo, portanto, os dados para o cálculo do índice. O erro absoluto de 8,43% no índice calculado pela *pipeline* (para a imagem de exemplo) é uma consequência direta do baixo *Recall* em ambas as classes, um problema que a U-Net resolveu parcialmente para as células em proliferação.

5 Conclusão

A quantificação do índice de proliferação celular, um fator prognóstico essencial na avaliação da agressividade do osteossarcoma canino, permanece um desafio crítico na rotina da patologia veterinária. O processo manual, baseado na contagem de células imunomarcadas com Ki-67, é reconhecidamente laborioso, subjetivo e suscetível à variabilidade inter e intraobservador, o que pode impactar diretamente as decisões clínicas. Diante deste cenário, o presente trabalho teve como objetivo central propor e validar rotinas computacionais automatizadas para a segmentação e contagem de células em lâminas histológicas de OSA, visando fornecer uma ferramenta objetiva e eficiente de apoio ao diagnóstico, comparando uma abordagem clássica de processamento de imagens a um modelo de DL.

A primeira contribuição deste estudo foi o desenvolvimento de uma *pipeline* de processamento de imagens robusta, capaz de diferenciar e quantificar as duas populações celulares de interesse: as células em proliferação e as células saudáveis para a obtenção do índice de proliferação celular. A metodologia demonstrou que a conversão do espaço de cor RGB para HSV foi uma etapa fundamental, permitindo desacoplar a informação de matiz da luminância e, assim, mitigar parcialmente a heterogeneidade cromática. A aplicação de *thresholding* específicos no espaço HSV permitiu um isolamento inicial eficaz das duas classes de células. Contudo, os resultados evidenciaram que esta etapa por si só era insuficiente. A subsequente aplicação de filtros, notadamente a Análise de Componentes Conectados para remoção de ruídos por área e a operação morfológica de Fechamento para preenchimento de discontinuidades, mostrou-se indispensável para refinar as máscaras binárias e garantir a integridade dos objetos de interesse antes da contagem final por extração de contornos.

A validação da *pipeline*, realizada contra anotações de um patologista especialista nas 18 imagens disponíveis, revelou um desempenho promissor para as células em proliferação. A rotina alcançou um nível de Precisão consistentemente elevado, com a maioria das amostras concentrando-se em torno de 92%. Este resultado positivo indica que a assinatura cromática marrom, quando presente de forma clara, é um diferenciador muito confiável, tornando a ocorrência de FPs rara.

Em contrapartida, a métrica de *Recall* apresentou um comportamento bimodal, expondo a principal vulnerabilidade da abordagem: a sensibilidade à qualidade da preparação histológica. Amostras consideradas "padrão-ouro", com boa separação celular e coloração nítida, atingiram um *Recall* satisfatório (pico de 71%). No entanto, em lâminas com variações tonais, coloração mais fraca ou sobreposição de tecidos, o *Recall* caiu drasticamente (pico de 39.5%), pois muitas células reais não foram capturadas pelos intervalos de cor pré-definidos. Esta descoberta sublinha que, embora eficaz, a metodologia baseada em processamento clássico de imagens

possui uma generalização limitada diante da variabilidade inerente aos protocolos de coloração.

Paralelamente, este trabalho desenvolveu e avaliou um modelo de segmentação baseado em ML, utilizando a arquitetura U-Net. Apesar da escassez de dados de treinamento, o modelo demonstrou uma rápida convergência e notável estabilidade na validação, atingindo um Coeficiente de *Dice* próximo de 0.8. A análise comparativa direta para a segmentação de células em proliferação revelou que o modelo U-Net obteve um desempenho equiparado e, em métricas chave como o *Recall* (78.3% vs 74.66%), superior ao da *pipeline*. Sua capacidade de aprender relações não lineares e o fato de não ter a forte dependência da informação de cor, sugere uma maior capacidade de generalização frente a dados não vistos.

No entanto, o avanço de ambas as frentes metodológicas foi severamente restringido pela escassez de dados anotados. A base de dados, embora contendo 72 imagens, dispunha de poucas imagens com anotações de células em proliferação e, de forma crítica, somente uma única imagem com a anotação completa de ambas as classes. Esta limitação teve um impacto profundo: impossibilitou uma validação estatística robusta da contagem de células saudáveis pela *pipeline* (que apresentou baixa precisão, 62.75%, na única imagem testada) e, consequentemente, do cálculo final do índice de proliferação. De forma ainda mais impactante, restringiu o treinamento do modelo U-Net exclusivamente à classe de células em proliferação, impedindo que o modelo aprendesse a tarefa completa de segmentação multiclasse.

Como trabalhos futuros, a etapa mais urgente é a expansão da base de dados anotada. É imperativo obter anotações completas de ambas as classes de células, realizadas por patologistas especialistas, para todas as imagens disponíveis. Com um conjunto de dados robusto, será possível, primeiramente, validar de forma abrangente o cálculo do índice de proliferação. Em segundo lugar, será viável treinar o modelo U-Net para a tarefa de segmentação multiclasse (fundo, célula em proliferação, célula saudável), permitindo uma avaliação justa de seu desempenho e de sua capacidade de generalização frente às variações de coloração, que se mostrou promissora.

Sugere-se também o aprimoramento de ambas as rotinas pela incorporação de técnicas de separação de aglomerados, um desafio recorrente. Além disso, a investigação mais profunda da técnica *Color Deconvolution* pode ser uma estratégia eficaz para padronizar as imagens e aumentar a robustez de ambos os métodos contra a variabilidade.

Em suma, este trabalho logrou desenvolver e comparar duas rotinas computacionais funcionais. A *pipeline* de processamento de imagens demonstrou alto potencial de precisão na identificação de células em proliferação, embora sua eficácia seja dependente da qualidade da lâmina. A abordagem por DL, mesmo treinada em dados limitados, estabeleceu-se como uma alternativa superior em termos de generalização e métricas. A pesquisa confirma a viabilidade da automação na quantificação do Ki-67 em OSA canino, mas destaca a aquisição de dados anotados de alta qualidade como o principal gargalo a ser superado para a transição desta ferramenta da prova de conceito para a aplicação clínica robusta.

Referências

- ADAMS, R.; BISCHOF, L. Seeded region growing. *IEEE Transactions on pattern analysis and machine intelligence*, IEEE, v. 16, n. 6, p. 641–647, 1994.
- BALAJI, P.; HUNG, B. T.; CHAKRABARTI, P.; CHAKRABARTI, T.; ELNGAR, A. A.; ALUVALU, R. A novel artificial intelligence-based predictive analytics technique to detect skin cancer. *PeerJ Computer Science*, PeerJ Inc., v. 9, p. e1387, 2023.
- BRODEY, R. S.; RISER, W. H. 7 canine osteosarcoma: A clinicopathologic study of 194 cases. *Clinical Orthopaedics and Related Research (1976-2007)*, LWW, v. 62, p. 54–64, 1969.
- BURRAI, G. P.; GABRIELI, A.; POLINAS, M.; MURGIA, C.; BECCHERE, M. P.; DEMONTIS, P.; ANTUOFERMO, E. Canine mammary tumor histopathological image classification via computer-aided pathology: an available dataset for imaging analysis. *Animals*, MDPI, v. 13, n. 9, p. 1563, 2023.
- CANNY, J. A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence*, IEEE, n. 6, p. 679–698, 2009.
- CAVALCANTI, J. N. *Osteossarcoma canino: estudo histopatológico, imunoistoquímico e da atividade proliferativa*. Tese (Doutorado) — [sn], 2007.
- FAWCETT, T. An introduction to roc analysis. *Pattern recognition letters*, Elsevier, v. 27, n. 8, p. 861–874, 2006.
- FILHO, J. C. O.; KOMMERS, G. D.; MASUDA, E. K.; MARQUES, B. M.; FIGHERA, R. A.; IRIGOYEN, L. F.; BARROS, C. S. Estudo retrospectivo de 1.647 tumores mamários em cães. *Pesquisa Veterinária Brasileira*, SciELO Brasil, v. 30, p. 177–185, 2010.
- GURCAN, M. N.; BOUCHERON, L. E.; CAN, A.; MADABHUSHI, A.; RAJPOOT, N. M.; YENER, B. Histopathological image analysis: A review. *IEEE reviews in biomedical engineering*, IEEE, v. 2, p. 147–171, 2009.
- HARALICK, R. M.; STERNBERG, S. R.; ZHUANG, X. Image analysis using mathematical morphology. *IEEE transactions on pattern analysis and machine intelligence*, IEEE, n. 4, p. 532–550, 1987.
- HORIMOTO, Y.; ARAKAWA, A.; TANABE, M.; KURODA, K.; MATSUOKA, J.; IGARI, F.; HIMURO, T.; YOSHIDA, Y.; TOKUDA, E.; SHIMIZU, H. et al. Menstrual cycle could affect ki67 expression in estrogen receptor-positive breast cancer patients. *Journal of Clinical Pathology*, BMJ Publishing Group, v. 68, n. 10, p. 825–829, 2015.
- HOROWITZ, S. L.; PAVLIDIS, T. Picture segmentation by a tree traversal algorithm. *Journal of the ACM (JACM)*, ACM New York, NY, USA, v. 23, n. 2, p. 368–388, 1976.
- JAGANATHAN, D.; BALASUBRAMANIAM, S.; SURESHKUMAR, V.; DHANASEKARAN, S. Revolutionizing breast cancer diagnosis: a concatenated precision through transfer learning in histopathological data analysis. *Diagnostics*, MDPI, v. 14, n. 4, p. 422, 2024.

- KINGMA, D. P. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *nature*, Nature Publishing Group UK London, v. 521, n. 7553, p. 436–444, 2015.
- LITJENS, G.; KOOI, T.; BEJNORDI, B. E.; SETIO, A. A. A.; CIOMPI, F.; GHAFORIAN, M.; LAAK, J. A. V. D.; GINNEKEN, B. V.; SÁNCHEZ, C. I. A survey on deep learning in medical image analysis. *Medical image analysis*, Elsevier, v. 42, p. 60–88, 2017.
- MADABHUSHI, A.; LEE, G. Image analysis and machine learning in digital pathology: Challenges and opportunities. *Medical image analysis*, Elsevier, v. 33, p. 170–175, 2016.
- MEUTEN, D. J. *Tumors in domestic animals*. [S.l.]: John Wiley & Sons, 2020.
- MEYER, F.; BEUCHER, S. Morphological segmentation. *Journal of visual communication and image representation*, Elsevier, v. 1, n. 1, p. 21–46, 1990.
- NAUTA, M.; WALSH, R.; DUBOWSKI, A.; SEIFERT, C. Uncovering and correcting shortcut learning in machine learning models for skin cancer diagnosis. *Diagnostics*, MDPI, v. 12, n. 1, p. 40, 2021.
- NEUBECK, A.; GOOL, L. V. Efficient non-maximum suppression. In: IEEE. *18th international conference on pattern recognition (ICPR'06)*. [S.l.], 2006. v. 3, p. 850–855.
- OTSU. A threshold selection method from gray-level histograms. *Automatica*, v. 11, n. 285–296, p. 23–27, 1975.
- PERLE, K. M. L. *Machine learning and veterinary pathology: be not afraid!* [S.l.]: SAGE Publications Sage CA: Los Angeles, CA, 2019. 506–507 p.
- REDMON, J.; DIVVALA, S.; GIRSHICK, R.; FARHADI, A. You only look once: Unified, real-time object detection. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2016. p. 779–788.
- RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. In: SPRINGER. *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5–9, 2015, proceedings, part III 18*. [S.l.], 2015. p. 234–241.
- ROSENFELD, A.; PFALTZ, J. L. Sequential operations in digital picture processing. *Journal of the ACM (JACM)*, ACM New York, NY, USA, v. 13, n. 4, p. 471–494, 1966.
- SCHOLZEN, T.; GERDES, J. The ki-67 protein: from the known and the unknown. *Journal of cellular physiology*, Wiley Online Library, v. 182, n. 3, p. 311–322, 2000.
- SERRAS, M. J. C.; MATOS, R.; FAÍSCA, P. Comparação do índice de ki-67 obtido através de contagem manual e de análise automatizada imunoratió em amostras de tumores mamários caninos, carcinomas mamários felinos e mastocitomas. *Revista Lusófona de Ciência e Medicina Veterinária*, Edições Universitárias Lusófonas, 2020.
- SEZGIN, M.; SANKUR, B. I. Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic imaging*, Society of Photo-Optical Instrumentation Engineers, v. 13, n. 1, p. 146–168, 2004.

SHORTEN, C.; KHOSHGOFTAAR, T. M. A survey on image data augmentation for deep learning. *Journal of big data*, Springer, v. 6, n. 1, p. 1–48, 2019.

SUZUKI, S. Topological structural analysis of digitized binary images by border following. *Computer vision, graphics, and image processing*, Elsevier, v. 30, n. 1, p. 32–46, 1985.

SZELISKI, R. *Computer vision: algorithms and applications*. [S.l.]: Springer Nature, 2022.

TUOMINEN, V. J.; RUOTOISTENMÄKI, S.; VIITANEN, A.; JUMPPANEN, M.; ISOLA, J. Immunoratio: a publicly available web application for quantitative image analysis of estrogen receptor (er), progesterone receptor (pr), and ki-67. *Breast cancer research*, Springer, v. 12, n. 4, p. R56, 2010.

VAHADANE, A.; PENG, T.; SETHI, A.; ALBARQOUNI, S.; WANG, L.; BAUST, M.; STEIGER, K.; SCHLITTER, A. M.; ESPOSITO, I.; NAVAB, N. Structure-preserving color normalization and sparse stain separation for histological images. *IEEE transactions on medical imaging*, IEEE, v. 35, n. 8, p. 1962–1971, 2016.

VINCENT, L.; SOILLE, P. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, IEEE Computer Society, v. 13, n. 06, p. 583–598, 1991.

XING, F.; YANG, L. Robust nucleus/cell detection and segmentation in digital pathology and microscopy images: a comprehensive review. *IEEE reviews in biomedical engineering*, IEEE, v. 9, p. 234–263, 2016.