

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO CEARÁ
PRÓ-REITORIA DE PESQUISA, PÓS-GRADUAÇÃO E INOVAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

José Jerovane da Costa Nascimento

**Nova abordagem por meio de *deep learning* e *fine-tuning* para
detecção e segmentação de melanoma em imagens dermatoscópicas**

Fortaleza – CE

20 de Fevereiro de 2024

Dados Internacionais de Catalogação na Publicação
Instituto Federal do Ceará - IFCE
Sistema de Bibliotecas - SIBI
Ficha catalográfica elaborada pelo SIBI/IFCE, com os dados fornecidos pelo(a) autor(a)

NASCIMENTO, JOSÉ JEROVANE DA COSTA.

Nova abordagem por meio de deep learning e fine-tuning para detecção e segmentação de melanoma em imagens dermatoscópicas / JOSÉ JEROVANE DA COSTA NASCIMENTO. - 2025.
84 f. : il. color.

Dissertação (Mestrado) - Instituto Federal do Ceará, Mestrado em Ciência da Computação, Campus Fortaleza, 2025.

Orientação: Prof. Dr. Carlos Mauricio Jaborandy de Mattos Dourado Junior.

1. Detecção e Segmentação de Melanoma. 2. Aprendizado Profundo. 3. Fine-Tuning. 4. Imagens Médicas. 5. Dermatoscopia. I. Título.

CDD 004

José Jerovane da Costa Nascimento

**Nova abordagem por meio de *deep learning* e *fine-tuning*
para detecção e segmentação de melanoma em imagens
dermatoscópicas**

Dissertação submetida ao Programa de Pós-Graduação em Ciência da Computação do Instituto Federal de Educação, Ciência e Tecnologia do Ceará, como requisito parcial para obtenção do grau de Mestre em Ciência da Computação.

Linha de Pesquisa: Processamento Digitais de Sinais e Imagens

Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE)

Pró-Reitoria de Pesquisa, Pós-Graduação e Inovação (PRPI)

Programa de Pós-Graduação em Ciência da Computação (PPGCC)

Fortaleza – CE

20 de Fevereiro de 2024

Agradecimentos

Em primeiro lugar, expresso minha profunda gratidão ao Pai Celestial pelo dom da vida, pela força, saúde e determinação que me proporcionou para alcançar este objetivo. Agradeço imensamente pelo dom da graça concedido por meu Salvador Jesus Cristo, completando aquilo que eu, por mim só, não conseguiria. Sou igualmente grato pelo Espírito Santo, cujas inspirações, orientações e conforto foram luz nos momentos desafiadores.

À minha mãe, Maria da Conceição da Costa Nascimento, vai meu sincero agradecimento pelo amor incondicional, carinho e dedicação. Um exemplo de mulher guerreira, que dedicou toda sua vida à família.

Ao meu pai, Raimundo Nonato do Nascimento, expresso gratidão pelo exemplo de dedicação ao trabalho, honestidade e integridade. Muito obrigado por todo o cuidado dispensado.

À minha linda esposa, Ana Paula Lira Nascimento, dedico palavras de profundo agradecimento pelo seu amor, dedicação, apoio, amizade e companheirismo. Agradeço pelos momentos em que estive ao meu lado, tanto nos desafios quanto nas alegrias. Sou grato por seu apoio constante, por torcer por mim e por nunca desistir. Você é meu exemplo. Obrigado por cuidar de mim e de nossos filhos com dedicação, sempre com um sorriso no rosto e muito amor. Que possamos continuar a viver cada dia nossa vida, nossa história, nossos sonhos e felicidades. Que seja eterno o que começamos juntos aqui nesta terra. Muito obrigado, minha amada.

Não posso deixar de agradecer aos meus filhos: Jenniffer Lira Nascimento, Giovana Letícia Lira Nascimento, Ismael Lira Nascimento e Emanuel Lira Nascimento. Vocês são minhas maiores inspirações e motivações. Sou grato por ter filhos talentosos, dedicados e tementes a Deus, sempre prontos a servir ao próximo. Um dos meus objetivos em continuar estudando é deixar um legado para vocês. Agradeço por ver cada um progredindo em todos os aspectos: fisicamente, intelectualmente, emocionalmente e espiritualmente.

Aos meus irmãos Maria Aparecida, Edvan Nascimento e Maria da Paz, expresso agradecimentos pela amizade, companheirismo e torcida constante. Agradeço imensamente pelo respeito sempre demonstrado. Meus agradecimentos também se estendem a todos os meus familiares e amigos pelo apoio constante.

Ao meu Orientador, Prof. Dr. Carlos Maurício Jaborandy de Mattos Dourado Júnior, agradeço pela paciência e generosidade durante o período acadêmico. Muito obrigado pela parceria e orientações.

Ao meu Coorientador, Prof. Dr. Luís Fabrício de Freitas Souza, expresso agrade-

cimentos pela paciência, orientações, motivações e oportunidades concedidas. Obrigado por ensinar o caminho das pedras, por inspirar e motivar. Que o Pai Celestial continue iluminando, protegendo e abençoando seu caminho e de sua família. Muito obrigado por seu mais do que um Coorientador.

Não posso deixar de expressar minha profunda gratidão aos amigos e parceiros de pesquisa, Luís Fabrício, Adriel Gomes, Guilherme Brilhante e Yasmim Rodrigues. Vão além de colegas, são verdadeiros companheiros, sempre prontos para a batalha, incansáveis e determinados. Agradeço não apenas pela colaboração nas pesquisas, mas sobretudo pela amizade e pelo estímulo diário. Minhas preces diárias incluem pedidos para que o Senhor os guie, concedendo a vocês a realização de todos os nobres objetivos em suas vidas.

“E como nem todos têm fé, buscai diligentemente e ensinai-vos uns aos outros palavras de sabedoria; sim, nos melhores livros buscai palavras de sabedoria; procurai conhecimento, sim, pelo estudo e também pela fé.”

Doutrina e Convênios 88:118

Resumo

O câncer de pele é uma patologia que vem se desenvolvendo de forma alarmante no mundo. De acordo com a Organização Mundial da Saúde (OMS) o melanoma é um tipo de câncer que afeta a população mundial em diferentes partes do corpo humano, trazendo o óbito de milhares de pessoas cada ano no mundo. Ferramentas de diagnóstico inteligente por meio de detecção automática em imagens médicas são extremamente eficazes para auxílio ao diagnóstico médico. Sistemas baseados em *Computer Aided Diagnostics* (CADs) são extremamente importantes para o pré-diagnóstico por imagens, o uso de ferramentas baseadas em inteligência artificial para monitoramento, detecção e segmentação da região patológica são cada vez mais utilizadas em soluções inteligentes integradas a sistemas de suporte à tomada de decisões. Este estudo tem como proposta uma nova abordagem para segmentação de melanoma em imagens dermatoscópicas com uso de CNN para detecção e *Fine-Tuning* baseado em processamento digital de imagens. O método apresenta uma abordagem *deep learning* com base nas diferentes versões do *framework YOLOv8* para detecção de melanomas em imagens dermatoscópicas, na qual a abordagem propõe novas técnicas de *Fine-Tuning* para segmentação da região do melanoma. A nova abordagem obteve resultados satisfatórios de 98.67% de acurácia para detecção e de 98.68% de acurácia para segmentação do câncer de pele, superando diferentes trabalhos encontrados no estado da arte em diferentes métodos clássicos e atuais com abordagens, tais como métodos manuais, semi-automáticos e automáticos.

Palavras-chave: Detecção e Segmentação de Melanoma, Aprendizado Profundo, *Fine-Tuning*, Imagens Médicas, Dermatoscopia.

Abstract

Skin cancer is a pathology that has been alarmingly developing worldwide. According to the World Health Organization (WHO), melanoma is a type of cancer that affects the global population in different parts of the human body, causing the death of thousands of people each year. Intelligent diagnostic tools using automatic detection in medical images are extremely effective for assisting in medical diagnosis. Computer-Aided Diagnosis (CAD) systems are crucial for pre-diagnosis through image analysis, and the use of artificial intelligence-based tools for monitoring, detection, and segmentation of pathological regions is increasingly integrated into intelligent solutions supporting decision-making systems. This study proposes a new approach for melanoma segmentation in dermatoscopic images using CNN for detection and Fine-Tuning based on digital image processing. The method presents a deep learning approach based on different versions of the YOLOv8 framework for melanoma detection in dermatoscopic images, introducing new Fine-Tuning techniques for melanoma region segmentation. The new approach achieved satisfactory results with 98.67% accuracy for detection and 98.68% accuracy for skin cancer segmentation, surpassing various state-of-the-art works employing different classical and current methods, including manual, semi-automatic, and automatic approaches.

Keywords: Melanoma Detection and Segmentation, Deep Learning, Fine-Tuning, Medical Images, Dermoscopy.

Lista de ilustrações

Figura 1 – Exame de Dermatoscopia: método diagnóstico não invasivo, que auxilia na avaliação de lesões pigmentadas da pele. A imagem ilustrativa mostra um médico especialista utilizando o dermatoscópio para uma visualização precisa da região de interesse, identificando potencialmente um melanoma. Fonte: Autor.	22
Figura 2 – Diferentes tipos e formas de melanomas em imagens dermatoscópicas. .	23
Figura 3 – (A)- Ilustração da ligação (conexão) entre as camadas de uma rede CNN. (B)- Representação de 3 neurônios da camada oculta interligados ao mesmo Mapa de Características. Fonte: O próprio autor.	34
Figura 4 – Ilustração do processo de camadas convolucionais. Fonte: Adaptada de < https://repositorio.ufc.br/handle/riufc/72342 >.	35
Figura 5 – Ilustração de diferentes camadas encarregadas em um processo de convolucional. Fonte: o próprio autor.	36
Figura 6 – Ilustra a arquitetura do YOLOv8. Este modelo utiliza uma espinha dorsal chamada C2f, substituindo a CSPLayer. O C2f destaca-se por integrar características de alto nível com informações contextuais, aprimorando a precisão na detecção de objetos. Adotando um modelo sem âncoras e uma abordagem de cabeça desvinculada, o YOLOv8 permite que cada ramo se concentre em tarefas específicas, resultando em um aumento geral na precisão. A camada de saída usa ativação sigmoid para pontuação de objetividade e softmax para probabilidades de classe. As funções de perda CIoU e DFL são implementadas para melhoria na detecção de objetos, especialmente em cenários com objetos de menor escala. Fonte: diagrama baseado em (TERVEN; CórDOVA-ESPARZA; ROMERO-GONZÁLEZ, 2023) e (CONTRIBUTORS, 2023).	39
Figura 7 – Representação do modelo fine-tuning do Janelamento de Parzen; 1 - inicialização do método; 2 - Geração do Hiperplano; 3 - Mapa das densidades de probabilidades; 4 - Resultado das estimativas de diferentes densidades do mapa; 5 - Máscara baseada na densidade do mapa; 6 - Resultado da Segmentação.	46
Figura 8 – Representação do modelo Fine-tuning do Janelamento de Parzen com Clusterização; 1 - inicialização do método; 2 - Clusterização; 3 - Geração do Hiperplano; 4 - Mapa das densidades de probabilidades; 5 - Resultado das estimativas de diferentes densidades do mapa; 6 - Máscara baseada na densidade do mapa; 7 - Resultado da Segmentação.	47

Figura 9 – Representação do modelo fine-tuning do Janelamento de Parzen com Crescimento de Região: 1 - inicialização; 2 - Cálculo do centroide; 3 - Crescimento de Região; 4 - Geração do Hiperplano; 5 - Mapa das densidades de probabilidades; 6 - Resultado das estimativas de diferentes densidades do mapa; 7 - Máscara baseada na densidade do mapa; 8 - Resultado da Segmentação.	48
Figura 10 – Ilustra uma visualização da metodologia proposta. Na Etapa (a), destaca-se o treinamento dos modelos YOLO, enquanto a Etapa (b) demonstra a detecção da região do melanoma por meio de aprendizagem profunda. A máscara binária de detecção é ilustrada na Etapa (c), revelando a delimitação da região de interesse. A Etapa (d) mostra o processo de <i>fine-tuning</i> dos modelos, e na Etapa (e) são apresentados os resultados obtidos. Por fim, a Etapa (f) destaca o melhor modelo. Fonte: o próprio autor.	54
Figura 11 – imagem da etapa de treinamento. Fonte: o próprio autor.	56
Figura 12 – Funcionamento do Estágio 1 da metodologia proposta com os estágios a1 - a5 sendo ilustrados.	56
Figura 13 – Imagem da evolução do janelamento de parzen. Fonte: o próprio autor.	59
Figura 14 – Modelos <i>Fine-Tuning</i> desenvolvidos a partir dos métodos de processamento digital de imagens. Fonte: o próprio autor.	61
Figura 15 – Resultado da detecção de diferentes melanomas em imagens dermatoscópicas com as diversas versões do YOLOv8, com destaque para a YOLOv8x, na qual apresentou melhores resultados para detecção.	64
Figura 16 – Detecções por caixa delimitadora para o modelo YOLOv8x.	65
Figura 17 – Comparação gráficas da Tabela 3. Gráfico de ilustração da Comparação das métricas de segmentação realizadas pelo modelos <i>Fine-tuning</i> tabela.	68
Figura 18 – Comparação gráficas da Tabela 3. Gráfico de ilustração da Comparação das métricas de segmentação realizadas pelo modelos <i>Fine-tuning</i> tabela.	70
Figura 19 – Resultado de Segmentação de casos distintos para o melhor modelo proposto, YOLOv8x + Parzen_Region_G.	71
Figura 20 – Gráfico de comparação do tempo de segmentação para diferentes combinações de detecção e ajuste fino (<i>fine-tuning</i>).	73
Figura 21 – Comparação entre os resultados de segmentação entre os diferentes modelos propostos por este estudo YOLOv8x.	74
Figura 22 – Gráfico obtido a partir da tabela de comparação com o estado da arte.	75

Lista de tabelas

Tabela 1	– Resultados métricos da detecção do melanoma em imagens dermatoscópicas para as diferentes versões do YOLOv8 abordadas neste estudo.	63
Tabela 2	– Comparação de Tempo de detecção para cada um dos modelos.	63
Tabela 3	– Resultados de métricas para segmentação nos modelos desenvolvidos neste estudo. Os modelos são compostos pela combinação de uma versão do YOLOv8 e com <i>Fine Tuning</i> que segmentam a região das bordas do melanoma. A tabela em questão cobre as métricas de ACC, DICE, SEN e SPE.	67
Tabela 4	– Resultados de métricas para segmentação nos modelos desenvolvidos neste estudo. Os modelos são compostos pela combinação de uma versão do YOLOv8 e com <i>Fine Tuning</i> . A tabela em questão abrange as métricas de PPV, NPV, MCC e JAC.	69
Tabela 5	– Representação dos resultados do teste estatístico de Friedman realizado entre os diferentes resultados de segmentação por <i>Fine-tuning</i> , <i>YOLOv8x</i> + <i>Parzen_Region_G</i>	72
Tabela 6	– Comparação de tempo de segmentação entre os métodos de <i>Fine-Tuning</i> proposto por este estudo.	72
Tabela 7	– Comparação do melhor modelo proposto por este estudo (<i>YOLOv8x</i> + <i>Parzen_Region_Growth</i>), comparados a trabalhos encontrados no estado da arte.	75

Lista de abreviaturas e siglas

OMS	Organização Mundial da Saúde
FINE-TUNING	Ajuste Fino
CAD	Diagnóstico Auxiliado por Computador (<i>Computer-Aided Diagnosis</i>)
PDI	Processamento Digital de Imagens
PARZEM	Janelamento de Parzem
CNN	Rede Neural Convolutacional (<i>Convolutional Neural Network</i>)
RNAs	Redes Neurais Artificiais
R-CNN	Redes Neurais Convolucionais baseadas em Regiões
ROI	Região de Interesse
ACC	Acurácia (<i>Accuracy</i>)
DICE	Coefficiente de Dice (<i>Dice Coefficient</i>)
SEN	Sensibilidade (<i>Sensitivity</i>)
SPE	Especificidade (<i>Specificity</i>)
PPV	Valor Preditivo Positivo (<i>Positive Predictive Value</i>)
NPV	Valor Preditivo Negativo (<i>Negative Predictive Value</i>)
MCC	Coefficiente de Correlação de Matthews (<i>Matthews Correlation Coefficient</i>)
JAC	Índice de Jaccard (<i>Jaccard Index</i>)
TP	Verdadeiro Positivo
FP	Falso Positivo
TN	Verdadeiro Negativo
FN	Falso Negativo
YOLOv8	<i>You Only Look One Level</i>
YOLOv8l	<i>YOLOv8 Large</i>

YOLOv8m	<i>YOLOv8 Medium</i>
YOLOv8n	<i>YOLOv8 Nano</i>
YOLOv8s	<i>YOLOv8 Small</i>
YOLOv8x	<i>YOLOv8 Extra Large</i>
UFCA	Universidade Federal do Cariri
UFC	Universidade Federal do Ceará
IFCE	Instituto Federal de Educação, Ciência e Tecnologia do Ceará

Lista de símbolos

ω	Letra grega Ômega minúsculo, indica a operação de máximo local no detector de fronteira difuso.
Ω	Letra grega Ômega maiúsculo, indica a operação de mínimo local no detector de fronteira difuso.
Π	Letra grega Pi Maiúscula, denominada Produto ou Produtório, no conceito matemático, refere-se à multiplicação.
Φ	Letra grega Fi, equivale ao conjunto de parâmetros definidos pela i-ésima componente da mistura.
Θ	Letra grega Teta, representa a medida de um determinado vetor, sendo um conjunto de parâmetros da mistura.
θ	Letra grega teta, referência ao ângulo.
Σ	Notação Sigma, somatório ou somatória, significa a soma de termos.
π	Símbolo Pi, representa uma proporção numérica definida pela relação entre o perímetro de uma circunferência e seu diâmetro.
ϕ	Letra grega Fi (variante), correspondente a uma função de densidade Gaussiana.
α	Letra grega Alpha, conotação angular de início ou inicialização.
β	Letra grega Beta, conotação ponderação ao ângulo.
$\int$	Símbolo Integral, representa um sinal que indica a integração e o resultado de integrar uma diferencial ou derivada.
μ	Letra grega Mi (Pronúncia: Miú), refere-se ao peso específico de um coeficiente.
∂	Símbolo "del", usado para distinguir derivadas parciais das derivadas ordinárias de uma variável.
$\cap$	Símbolo Intersecção, corresponde à intersecção entre conjuntos.
ω	Letra grega Ômega minúsculo, indica o mínimo local.
Ω	Letra grega Ômega maiúsculo, indica a composição do mínimo com o máximo local.

ν	Letra grega Nu, indica o limiar da força balão.
∇	Símbolo grego Nabla, denomina o operador diferencial.
$\in$	Símbolo que indica a relação de pertinência.

Sumário

1	INTRODUÇÃO	18
1.1	Motivação para a Dissertação	20
1.2	Descrição do problema	22
1.3	Objetivo Geral e Específicos	24
1.3.1	Objetivo Geral	24
1.3.2	Objetivos Específicos	24
1.4	Produção científica	24
1.5	Estrutura da Dissertação	26
2	TRABALHOS RELACIONADOS	27
3	FUNDAMENTAÇÃO TEÓRICA	31
3.1	Conjunto de Dados	31
3.2	Aprendizagem de Máquina	31
3.3	Aprendizado Profundo	32
3.4	Redes Neurais Convolucionais - CNN	33
3.5	Frameworks de Detecção	36
3.5.1	Framework YOLO	37
3.6	Segmentação - Ajuste Fino (<i>Fine-tuning</i>) e Otimização dos Modelos	40
3.6.1	Crescimento de Região	40
3.6.2	Operações Morfológicas	41
3.6.3	Agrupamento <i>K-means</i>	42
3.6.4	Janelamento de Parzen	43
3.6.5	<i>Fine-tuning</i> no processo de segmentação	44
3.7	Teste Estatístico	49
3.8	Métricas de Avaliação	50
3.8.1	Acurácia	51
3.8.2	Coeficiente dos dados - (<i>Dice's coefficient</i>)	51
3.8.3	Sensitividade	51
3.8.4	Especificidade	52
3.8.5	Valor Preditivo Positivo	52
3.8.6	Valor Preditivo Negativo	53
3.8.7	Coeficiente de Correlação de Matthews	53
3.8.8	JAC (Coeficiente de Jaccard)	53
4	MÉTODO PROPOSTO	54

4.1	Metodologia	55
4.1.1	Etapa 1 - Treinamento de modelos e Detecção do Melanoma	55
4.1.2	Etapa 2 - Segmentação do Melanoma com Modelos <i>Fine-Tuning</i>	58
5	RESULTADOS E DISCUSSÃO	62
5.1	Resultados e Discussão	62
5.1.1	Resultados da Primeira Fase de Experimentos - Detecção	62
5.1.2	Resultados da Segunda Fase de Experimento - Segmentação (<i>Fine-tuning</i>) .	66
5.1.3	Resultados do Terceiro Experimento - Comparação com o Estado da Arte .	75
6	CONCLUSÃO	77
	REFERÊNCIAS	78

1 Introdução

O melanoma é uma variação de câncer de pele originada a partir das células produtoras de melanina, os melanócitos. Esse tipo de câncer ocorre quando há multiplicação descontrolada dessas células em determinadas regiões, formando tumores ([LONG et al., 2023](#)). Ele pode se desenvolver em qualquer parte do corpo onde há presença significativa de melanócitos; no entanto, os casos mais comuns ocorrem na pele, devido à exposição a altas doses de radiação ultravioleta (UV) ([LONG et al., 2023](#)). Vale ressaltar que esse tipo de patologia é mais comum em pessoas de pele mais clara, devido à menor taxa de melanina na pele e maior sensibilidade à mutação cancerígena das células mediante altas doses de raios UV ([LONG et al., 2023](#)).

No que diz respeito aos índices deste tipo de câncer, os dados são alarmantes. Relatórios indicam aproximadamente 325.000 novos casos em 2020, resultando em cerca de 57.000 óbitos. As projeções futuras são ainda mais preocupantes, estimando-se 510.000 novos casos dessa patologia até 2040, com aproximadamente 96.000 casos fatais ([ARNOLD et al., 2022](#)). Nos Estados Unidos, Reino Unido e Austrália, a média anual de óbitos chega a 15.500 casos ([GORDON et al., 2020](#)).

Esses dados destacam a urgência da prevenção, tratamento e acompanhamento do melanoma, ressaltando a importância crucial da detecção precoce para a sobrevivência do paciente. Conforme os estudos de([SANDRU et al., 2014](#)), a taxa de sobrevivência para pacientes com melanoma maligno ao longo de 5 anos de progressão varia entre 5% e 19%. Essa variação é influenciada pela localização e quantidade de metástases da patologia, que pode se espalhar para outros órgãos do corpo.

Além dos altos índices de incidência de novos casos e das baixas taxas de sobrevivência para casos mais avançados, o melanoma também acarreta em elevados custos médicos para todos os países do mundo. Apenas nos Estados Unidos, em 2011, foram gastos aproximadamente US\$4.8 bilhões em custos relacionados à prevenção e tratamento dessa patologia ([GORDON et al., 2020](#)). O custo médio por paciente para o tratamento do melanoma em países europeus, como a Suécia, é de cerca de 17.408 € para casos avançados, 923 € para tratamento ambulatorial, hospitalização e consulta, e 3.511 € para custos relacionados à mortalidade. Essa última taxa atinge 20.418 € em países como a Inglaterra ([KRENSEL; SCHÄFER; AUGUSTIN, 2019](#)).

Dentre os exames utilizados no diagnóstico do melanoma, esses podem variar desde exames clínicos, como autoexames e exames dermatológicos com a utilização de imagens dermatoscópicas de alta resolução; biópsia, considerada o exame definitivo para confirmação da cancerosidade de uma lesão, embora seja invasivo; mapeamento corporal, que documenta

todas as pintas e lesões na pele para acompanhamento ao longo do tempo; até tomografia computadorizada e ressonância, amplamente aplicados para detectar metástase do câncer. Inclusive, em determinadas situações, exames de sangue podem ser aplicados com o intuito de detectar marcadores tumorais no organismo do paciente (ROSSI et al., 2019; JONES et al., 2019).

Dentre esses exames, a dermatoscopia é uma valiosa e não invasiva técnica de diagnóstico e monitoramento de diversas lesões de pele, incluindo o melanoma. Esse método utiliza o dermatoscópio, um instrumento que possibilita a visualização precisa de potenciais anormalidades cutâneas, garantindo, assim, um aumento significativo na precisão do diagnóstico de lesões suspeitas de câncer de pele, tanto melanomas quanto carcinomas, mesmo em estágios extremamente iniciais.

A dermatoscopia revela-se especialmente eficaz ao permitir a detecção precoce da doença em estágios iniciais, por meio de imagens de alta resolução. Além disso, contribui para elevar a taxa de sensibilidade e especificidade na detecção, permitindo ao médico especialista diferenciar entre melanomas benignos e malignos por meio dos detalhes visuais fornecidos. Esse método também facilita um acompanhamento eficaz ao longo do tempo, possibilitando a comparação visual das imagens de alta resolução durante o intervalo entre os exames (CELEBI; CODELLA; HALPERN, 2019). Antes da adoção da dermatoscopia, o exame de sinais cutâneos era conduzido a olho nu, sem auxílio, o que impossibilitava uma avaliação detalhada das lesões cutâneas e, em algumas situações, resultava em atrasos no diagnóstico de cânceres já presentes.

Outro fator importante da dermatoscopia é a facilidade de integração com sistemas de detecção automáticos que processam as imagens dermatoscópicas de alta resolução, gerando análises médicas úteis para o diagnóstico e acompanhamento mais eficazes por parte de profissionais da área. Esses sistemas são denominados sistemas de auxílio ao diagnóstico médico (CAD) e são constituídos por diversas tecnologias de inteligência artificial, como *Machine Learning*, *Deep Learning* e *CNNs* - (*Convolutional Neural Networks*). Essas tecnologias possibilitam não apenas a detecção com alta eficácia, mas também o ajuste fino da segmentação do melanoma, proporcionando uma maior precisão diagnóstica (GAJERA; NAYAK; ZAVERI, 2023).

Os sistemas CAD são desenvolvidos para realizar a detecção do melanoma na imagem do exame, classificar o tipo de melanoma entre benigno e maligno e segmentar a superfície da lesão, utilizando métodos cada vez mais performáticos encontrados no estado da arte. Esses métodos são capazes de detectar e segmentar a região afetada de maneira bastante próxima à exatidão do médico especialista, com métricas crescentes e uma taxa de precisão superior a 99%. Essas ferramentas CAD estão cada vez mais em alta, fornecendo análises e sugestões que resultam em decisões de tratamento adequadas em tempo hábil à detecção precoce do exame. Isso auxilia na obtenção de resultados clínicos mais favoráveis

e eficazes, além de reduzir os custos clínicos de acompanhamento provenientes de soluções mais dispendiosas e invasivas de diagnóstico. Tais aplicações automáticas são incluídas em estudos das mais renomadas instituições de pesquisa (ACETO; PERSICO; PESCAPÉ, 2020).

Considerando a urgência do diagnóstico do melanoma, especialmente devido às elevadas taxas de casos e mortalidade em pacientes com estágios mais avançados, e diante dos desafios associados ao desenvolvimento de sistemas automáticos de detecção e segmentação do melanoma em imagens dermatoscópicas, buscando maior agilidade e eficiência na execução da atividade principal, este estudo propõe um novo método baseado em inteligência artificial para detecção e segmentação de melanoma em imagens dermatoscópicas. Essa abordagem utiliza modelos de *deep learning* e *fine-tuning*, adotando o conceito de CNN (*Convolutional Neural Network*).

Este trabalho aborda diversas temáticas com base nos seguintes eixos:

- Detecção de melanoma em imagens dermatoscópicas usando aprendizado profundo.
- Segmentação do câncer de pele usando ajuste fino aplicado a modelos computacionais.
- Sistema de diagnóstico auxiliado por computador para imagens médicas por meio de processamento de imagens usando inteligência artificial e técnicas de visão computacional.
- Auxílio ao diagnóstico por processamento em imagens médicas.

1.1 Motivação para a Dissertação

No vasto campo da pesquisa médica, surgem inúmeras estratégias para abordar a complexa tarefa de segmentar regiões em imagens dermatoscópicas, especialmente ao identificar e delimitar melanomas em imagens médicas. Tradicionalmente, muitos desses métodos empregam técnicas clássicas que dependem de modelos manuais ou semi-automatizados, o que torna o processo de segmentação excessivamente vinculado à intervenção humana, como ilustrado na Figura 1, que apresenta o fluxo do exame de dermatoscopia.

No entanto, nota-se uma tendência em crescimento na atualidade, focada no uso de Redes Neurais Convolucionais (CNNs), impulsionadas por sua notável capacidade de processamento de dados e pela habilidade de gerar resultados de alta qualidade. Contudo, ao aplicar essas inovações ao âmbito das imagens médicas, surgem desafios significativos. Muitas aplicações demandam volumes substanciais de dados, resultando em custos elevados, tanto em termos de recursos computacionais quanto de investimentos financeiros. Isso abrange desde a aquisição das imagens e a necessária marcação por especialistas até a anotação dos dados e sua posterior armazenagem, formando uma cadeia de processamento

complexa e dispendiosa.

Além disso, a diversidade de problemas que surgem no campo da medicina, sem uma abordagem refinada e específica, limita a aplicação de modelos genéricos a situações particulares, reduzindo a adaptabilidade na criação de modelos com aplicações em desafios médicos variados. Por outro lado, o avanço contínuo das CNNs se beneficia da crescente capacidade de processamento, da automatização dos modelos e do desenvolvimento de soluções que geram resultados altamente satisfatórios. Técnicas inspiradas em (*fine-tuning*) e na aprendizagem por transferência permitem a criação de modelos mais resistentes e adaptados para resolver problemas específicos. Quando essas técnicas são combinadas com outros métodos de aprendizado de máquina, é possível ajustar os modelos com maior precisão, tornando-os adequados para enfrentar desafios médicos específicos. Isso, por sua vez, oferece estratégias mais eficazes para melhorar os resultados.

No que diz respeito aos métodos de auxílio ao diagnóstico aplicados nas imagens dermatoscópicas, a utilização de CNNs capazes de processar imagens e detectar objetos possibilita a criação de modelos totalmente automáticos. Esses métodos podem ser desenvolvidos ao combinar os modelos de detecção automática com técnicas de *fine-tuning*, proporcionando assim uma possibilidade de aprimorar a eficiência na segmentação da região lesionada na imagem.

Esta dissertação tem como foco explorar e aprimorar essas abordagens, direcionando-as para a detecção e segmentação de melanoma em imagens dermatoscópicas. Com o objetivo de otimizar o uso de recursos, reduzir os custos associados ao processamento de imagens médicas e fornecer soluções eficazes e adaptáveis para superar os desafios únicos encontrados na área de diagnóstico médico em imagens de câncer de pele.



Figura 1 – Exame de Dermatoscopia: método diagnóstico não invasivo, que auxilia na avaliação de lesões pigmentadas da pele. A imagem ilustrativa mostra um médico especialista utilizando o dermatoscópio para uma visualização precisa da região de interesse, identificando potencialmente um melanoma. Fonte: Autor.

1.2 Descrição do problema

A detecção e segmentação de melanomas em imagens dermatoscópicas são tarefas altamente desafiadoras e de grande importância na área da dermatologia. A complexidade decorre de vários fatores intrínsecos relacionados à natureza dos melanomas e às características das imagens dermatoscópicas.

Os melanomas podem apresentar uma ampla variedade de tamanhos, formas e localizações na pele. Alguns podem ser muito pequenos e difíceis de detectar, enquanto outros podem ser extensos, abrangendo diferentes áreas da pele rapidamente. A variação nas características morfológicas dos melanomas, como bordas irregulares e texturas distintas, torna desafiadora a delimitação precisa de sua extensão na imagem dermatoscópica. Essa diversidade morfológica destaca a complexidade na identificação dos melanomas.

Além disso, as imagens dermatoscópicas podem conter artefatos, como ruídos e variações de intensidade, que interferem na identificação correta dos melanomas. Essas interferências podem ser causadas por vários fatores, como movimentos do paciente durante a aquisição da imagem, qualidade do equipamento dermatoscópico ou outros elementos externos, que podem obscurecer a presença dos melanomas, conforme ilustrado na Figura 2. Até mesmo a presença de pelos ou sinais na pele do paciente pode influenciar diretamente na complexidade da detecção do melanoma.

Criar novos modelos, diante desses desafios, requer combinações de métodos e técnicas para gerar modelos inovadores. Nessa temática, este trabalho apresenta como inovação uma solução para a segmentação de melanoma em imagens dermatoscópicas,



Figura 2 – Diferentes tipos e formas de melanomas em imagens dermatoscópicas.

reunindo conceitos atuais inspirados no modelo baseado em CNN combinado a métodos de *fine-tuning* baseados em processamento digital de imagens.

1.3 Objetivo Geral e Específicos

1.3.1 Objetivo Geral

Esta dissertação tem como objetivo geral propor uma nova abordagem para detectar e segmentar regiões de melanoma em imagens dermatoscópicas, utilizando CNNs combinadas com métodos baseados em *fine-tuning* de forma automática. Espera-se que os modelos desenvolvidos por essas combinações sejam capazes de realizar a detecção com altas taxas de confiabilidade e a segmentação com elevados índices de similaridade e sobreposição, conforme indicado pelas métricas de avaliação adotadas.

1.3.2 Objetivos Específicos

Além do objetivo geral, este trabalho contempla outros objetivos específicos, tais como:

1. Desenvolver uma nova abordagem para o processamento de imagens dermatoscópicas.
2. Propor diferentes modelos para detectar a região de melanomas utilizando métodos de aprendizagem profunda.
3. Apresentar uma nova abordagem para a segmentação da região do melanoma.
4. Realizar a comparação do melhor modelo proposto com métodos encontrados na literatura que foram testados na mesma base de dados.
5. Proporcionar uma discussão quantitativa e qualitativa, elencando os pontos-chave dos métodos abordados e como a sua combinação com o dataset em questão pode ser refletida nas métricas.

1.4 Produção científica

Durante este projeto de mestrado, os seguintes trabalhos científicos foram aceitos e publicados, a saber:

- (*International Conference on Intelligent Systems Design and Applications* publicado no ISDA 2023) - **SDA-Detection Melanoma: Deep Approach System for Detection and Segmentation in Melanoma Images Using Fine-tuning**;
- (CBA 2022) - An Approach to Location Estimation and Navigation Support Aided by Depth Estimation Map;

- (CBA 2022) - Uma Nova Abordagem para Localização de Robô Móvel em Ambiente Virtual utilizando Mapas Topológicos e Classificação com Opção de Rejeição a partir de Matriz de Co-ocorrência Estrutural em Imagens Omnidirecionais;
- (International Joint Conference on Neural Networks - IJCNN 2022) - Automatic Segmentation of Hemorrhagic Stroke on Brain CT Images Using Convolutional Neural Networks Through Fine-Tuning;
- (International Joint Conference on Neural Networks - IJCNN 2023) - New Health of Things Approach to Classification and Detection of Brain Tumors Using Transfer Learning for Segmentation in MR Images;
- (International Joint Conference on Neural Networks - IJCNN 2023) - Divisible Cell-Segmentation: A New Approach for Stroke Detection and Segmentation in CT Scans Using Deep Learning and Fine-Tuning;
- (International Conference on Intelligent Systems Design and Applications” publicado no ISDA 2022) - New Approach in LPR Systems Using Deep Learning to Classify Mercosur License Plates with Perspective Adjustment;
- (International Conference on Intelligent Systems Design and Applications” publicado no ISDA 2023) - ALPR System Perspective Adjustment: New Automatic License Plate Recognition Approach for Brazilian Mercosur Model Vehicle Plates;
- (International Conference on Systems, Signals and Image Processing, IWSSIP 2021) - New Approach Detection of COVID-19 by Cough Spectrogram Using IoT and Learning Transfer Technique for Classification;
- (IET Image Processing, v. 2022) - New fully automatic approach for tissue identification in histopathological examinations using transfer learning;
- (Learning and Nonlinear Models - Journal of the Brazilian Society on Computational Intelligence (SBIC), Vol. 19, Iss. 1, pp. 45-54, 2021) - Detection and Segmentation of Lungs Regions Using CNN Combined with Levelset;
- (Learning and Nonlinear Models - Journal of the Brazilian Society on Computational Intelligence (SBIC), Vol. 19, Iss. 2, pp. 4-14, 2021) - Detection and Segmentation of Damaged Photovoltaic Panels Using Deep Learning and Fine-tuning in Images Captured by Drone.

1.5 Estrutura da Dissertação

O capítulo 2 deste estudo apresenta uma revisão literária de trabalhos relacionados a métodos de detecção e segmentação melanoma, abrangendo desde métodos tradicionais até abordagens mais recentes baseadas em *deep learning* e *fine-tuning*. Em seguida, são apresentados métodos que serviram como base de análise e comparação de resultados para a base de dados em estudo. No capítulo 3, são detalhados os métodos e modelos de CNN aplicados na metodologia e utilizados para a criação e desenvolvimento do método proposto.

No Capítulo 4, é apresentada a nova abordagem para a segmentação de melanoma baseada em *deep learning*, aplicando técnicas de *fine-tuning*. Serão abordados desde métodos de Processamento Digital de Imagens até técnicas de *deep learning* e CNN. Finalizando, a discussão dos resultados e comparação com diferentes trabalhos da literatura será apresentada no Capítulo de Resultados e Discussão 5, seguida pela Conclusão e Trabalhos Futuros no Capítulo 6.

2 Trabalhos relacionados

Nesta seção, discutimos trabalhos relacionados ao tema desta pesquisa, abordando o uso de técnicas de classificação, detecção e segmentação em diferentes tipos de imagens médicas.

Em um cenário amplo de imagens médicas, as técnicas de classificação desempenham um papel crucial no diagnóstico de doenças, destacando-se as Redes Neurais Convolucionais (CNNs). Um exemplo é o estudo de (TRIPATHI; KOPPARAPU, 2020), que utilizou CNN para classificar sinais e identificar níveis de gravidade em pacientes com a doença de Parkinson, evidenciando a eficácia dessas redes na interpretação de padrões complexos em imagens médicas.

Outro avanço relevante é encontrado no trabalho de (RAJ et al., 2020), que empregou Redes Neurais Artificiais para classificação de imagens de câncer de pulmão, imagens cerebrais e doença de Alzheimer. Apesar da alta acurácia média de 95,22%, a discussão sobre a aplicação de diferentes algoritmos de otimização pode enriquecer ainda mais a compreensão dos modelos utilizados, proporcionando maior robustez e generalização.

No contexto da detecção de imagens médicas, métodos avançados têm obtido sucesso. (GAUR et al., 2021) utilizaram CNNs, como *EfficientNetB0*, *VGG16* e *InceptionV3*, aliadas a *transfer learning*, para detectar Covid-19 em tomografias computadorizadas, alcançando uma acurácia de 92,93%, evidenciando a aplicação promissora dessas técnicas em cenários clínicos complexos, embora desafios relacionados à disponibilidade de dados também sejam reconhecidos.

(HAN et al., 2019) apresentaram resultados positivos na detecção de metástases cerebrais utilizando YOLOv3, associado à técnica de aumento de dados por *Generative Adversarial Network* (GAN). A eficácia dessa abordagem, apesar dos riscos associados ao aumento de falsos positivos, ressalta a importância de estratégias inovadoras na detecção de condições médicas específicas.

A segmentação de tumores cerebrais em imagens de ressonância magnética também tem sido alvo de estudos, como demonstrado por (TRIPATHI; VERMA; SHARMA, 2021), que adaptaram a rede *Seg-Net* com melhorias na normalização cruzada de canal. Enquanto (BUDA; SAHA; MAZUROWSKI, 2019) empregaram a arquitetura U-Net para a mesma finalidade. Ambas as abordagens enfatizam a necessidade de considerar características específicas das imagens e a aplicação de métodos robustos de segmentação para lidar com desafios clínicos.

Explorando o cenário do AVC, (HAAN et al., 2015) propuseram um método de

segmentação semiautomático para AVC isquêmico em imagens de Ressonância Magnética (RM) e TC, alcançando índices Dice e Jaccard médios satisfatórios. A dependência da intervenção humana para melhores resultados destaca a importância de abordagens automatizadas mais avançadas.

Em AVC, (KAMNITSAS et al., 2017) propuseram um modelo de CNN para segmentação de AVC hemorrágico nas imagens de RM, superando limitações de propostas anteriores. A necessidade de ajustes adicionais nos dados e a aplicação de técnicas de pré-processamento indicam a constante busca por refinamento desses modelos.

A incorporação da rede *Mask R-CNN* em estudos de segmentação em imagens médicas tem ganhado destaque. (XIE et al., 2018) alcançaram resultados impressionantes na segmentação de núcleos celulares em imagens de microscopia, enquanto (ANANTHA-RAMAN; VELAZQUEZ; LEE, 2018) utilizaram a mesma abordagem para segmentar doenças orais. (HU et al., 2020) e (HAN et al., 2020) expandiram o uso da *Mask R-CNN* para a segmentação do pulmão em exames de TC, evidenciando a eficácia dessas técnicas em diversas aplicações médicas.

Avançando para a detecção e segmentação de melanomas em imagens dermatoscópicas, estudos recentes, como o de (RODRIGUES et al., 2023), propõem abordagens inovadoras, combinando redes de aprendizado profundo com técnicas de janelamento de Parzen e clusterização. Essas técnicas mostraram resultados promissores, superando métodos existentes no estado da arte. No entanto, a aplicabilidade em diferentes tipos de pele, variações de melanoma e a integração desses modelos em ambientes clínicos reais são aspectos que merecem uma análise mais aprofundada.

Focando na segmentação de lesões cutâneas, (AL-MASNI et al., 2018) propuseram o uso de redes convolucionais de resolução total, a FRCN, alcançando um índice Dice médio de 91,5% em conjuntos de dados desafiadores como PH2 e ISIC 2017. Esse avanço demonstra o potencial das CNNs para aprender características individuais de pixels sem a necessidade de pré e pós-processamento. Mesmo assim, seu Dice médio sugere uma análise mais ampla com o uso de métricas adicionais de avaliação.

Concentrando-se na dimensão colorimétrica da análise de imagens, (PATIÑO; AVENDAÑO; BRANCH, 2018) desenvolveram um método de segmentação de superpixels para lesões cutâneas em imagens dermatoscópicas, aplicando o algoritmo *Simple Linear Iterative Clustering (SLIC)*. Este método, baseado na média das cores RGB para fundir super pixels, alcançou uma precisão de 90% em bases de dados como PH2 e ISIC 2017. No entanto, encontrou desafios ao segmentar lesões com uma média de cor RGB heterogênea.

Aprofundando-se nas profundezas das redes neurais, (BAGHERSALIMI et al., 2019) introduziram a Dermonet, uma inovadora rede neural baseada em blocos convolucionais densamente conectados, que demonstrou a habilidade de extrair recursos de alto nível com

eficácia. Essa rede conseguiu uma especificidade notável de 95% no conjunto de dados PH2. A pesquisa sugere que a realização de ajustes finos e a inclusão de outras métricas de avaliação poderiam potencializar ainda mais a eficiência da Dermonet.

Paralelamente, ([BADRINARAYANAN; KENDALL; CIPOLLA, 2017](#)) desbravaram novos caminhos com a SegNet, uma arquitetura de rede neural totalmente convolucional voltada para a segmentação semântica pixel a pixel. Equipada com um codificador composto por 13 camadas convolucionais, a *SegNet* comprovou sua robustez ao alcançar 93% de precisão no conjunto de dados PH2, reafirmando o potencial das redes convolucionais na segmentação de imagens médicas.

Avançando ainda na abordagem de segmentação, ([PENNISI et al., 2016](#)) desenvolveram um método eficiente usando a Triangulação de Delaunay, que mostrou ser particularmente benéfico para conjuntos de dados limitados de imagens. Este método alcançou uma precisão de 89% no conjunto de dados de melanoma, embora sua eficácia seja moderada pela sensibilidade às distorções de imagem, dada a ênfase em características colorimétricas e geométricas.

Em uma via similar, ([BI et al., 2019](#)) reconheceram as limitações das redes convolucionais e propuseram a DCL-PSI, uma metodologia de aprendizado profundo que melhora a segmentação através de etapas baseadas em probabilidade. A abordagem resultou em uma precisão de 95,30%, indicando que a adição de um espectro mais amplo de ferramentas probabilísticas poderia oferecer melhorias substanciais.

Por sua vez, ([VASCONCELOS et al., 2019](#)) empregaram a morfologia matemática com o método MGAC, que se destacou pelo baixo custo computacional e alcançou 94,59% de precisão no conjunto de dados PH2. No entanto, a técnica mostrou-se sensível a variações em áreas de lesões mais claras, evidenciando sua dependência da modelagem precisa de uma equação diferencial.

Seguindo nesta linha, ([REBOUÇAS et al., 2021](#)) propuseram um método de segmentação avançado baseado em contornos ativos, que se aproveita das técnicas de janelamento de Parzen e clusterização pelo método *Floor of Log*. Este método alcançou uma precisão de 94,25% no conjunto de dados PH2, sinalizando que uma investigação mais aprofundada sobre os parâmetros ótimos para estas técnicas poderia resultar em melhorias significativas no desempenho da segmentação.

Em resumo, a pesquisa proposta, que se concentra na detecção e segmentação de melanoma em imagens dermatoscópicas, através da combinação de *deep learning* e *fine-tuning*, visa preencher lacunas identificadas na literatura. A integração do aprendizado profundo com ajustes finos tem como objetivo otimizar o processo de diagnóstico, oferecendo soluções mais eficazes e adaptáveis. Alinhada às tendências de inovação e eficiência na análise de imagens médicas, essa abordagem busca contribuir de maneira significativa para

o avanço do campo e o aprimoramento das práticas clínicas.

3 Fundamentação Teórica

Nesta seção, são apresentados os diferentes métodos empregados no modelo proposto, assim como os materiais utilizados no estudo.

3.1 Conjunto de Dados

O banco de dados PH2, composto por 200 imagens dermatoscópicas de alta qualidade, foi derivado de exames microscópicos de pele, os quais foram ampliados em uma escala de vinte vezes seu tamanho original. Essas imagens, provenientes do renomado Serviço de Dermatologia do Hospital Pedro Hispano, localizado em Portugal, são caracterizadas por uma resolução de 8 *bits* por *pixel*. Cada imagem no conjunto de dados possui dimensões específicas de 768x560 *pixels*, proporcionando uma visão detalhada para análises dermatológicas. Este banco de dados é particularmente valioso para pesquisas e estudos na área de dermatologia, notadamente para o treinamento e aprimoramento de algoritmos de inteligência artificial voltados para o diagnóstico e identificação de condições cutâneas.

3.2 Aprendizagem de Máquina

Ao revisitar os cenários delineados nos capítulos 1 e 2, destaca-se de forma expressiva o notável avanço que permeia as pesquisas e implementações no âmbito da aprendizagem de máquina na última década. Em termos gerais, a aprendizagem de máquina figura como uma disciplina específica inserida no vasto campo da inteligência artificial. Seu cerne reside na concepção, desenvolvimento e aprimoramento de técnicas computacionais, por meio de modelos de aprendizado automatizado. Esses modelos fundamentam-se no reconhecimento de padrões e na capacidade de tomada de decisões ([PROVOST; KOHAVI, 1998](#)).

A capacidade de aprendizado automático confere aos computadores a habilidade de aprender de maneira não programática e explícita ([SAMUEL, 1967](#)). Esse procedimento ocorre através de métodos probabilísticos que se apoiam em seus próprios erros, gerando previsões com base nos dados de entrada fornecidos aos modelos. O objetivo é identificar padrões e estabelecer conexões entre eles, a fim de adquirir a capacidade de realizar tarefas específicas sem a necessidade de intervenção humana direta ou indireta. Isso culmina na formalização do conceito de inteligência computacional. Esses modelos se valem de análises estatísticas com a finalidade de prever resultados, incrementando a precisão e reduzindo as chances de erro no processamento dos dados.

Dentro desse contexto, a literatura engloba uma ampla diversidade de modelos

que traçam suas raízes na aprendizagem de máquina, abrangendo desde abordagens clássicas e eficazes ([MOHRI; ROSTAMIZADEH; TALWALKAR, 2018](#)), até técnicas mais contemporâneas ancoradas na aprendizagem profunda ([LECUN; BENGIO; HINTON, 2015](#)).

3.3 Aprendizado Profundo

O cérebro humano, com sua intrincada rede de neurônios, possui uma capacidade notável de processar informações, desempenhando um vasto conjunto de funções, desde as mais simples até as mais complexas. Essa extraordinária máquina biológica se manifesta por meio de atividades que envolvem percepção, reconhecimento de padrões, tomada de decisões e inúmeras outras tarefas desenvolvidas pelo corpo humano ([KOCH; TSUCHIYA, 2007](#)).

Na tentativa de compreender e replicar esse processo tão sofisticado, surgiu o campo das Redes Neurais Artificiais (RNAs). Essas redes são máquinas projetadas especificamente para imitar as funcionalidades do cérebro, possibilitando a tomada de decisões ou realização de atividades específicas com base nos dados fornecidos ([HAYKIN, 2007](#)). Ao simular o processamento de informações semelhante ao do cérebro, as RNAs conseguem oferecer uma variedade de soluções para problemas que antes eram quase insuperáveis para métodos computacionais tradicionais.

Junto a essa abordagem inspirada na biologia, a computação tem desempenhado um papel vital. Utilizando conceitos matemáticos, estatísticos e linguagens de programação, os cientistas foram capazes de construir modelos que não apenas simulassem, mas em certos casos, otimizassem as atividades de inteligência realizadas pelo cérebro humano. Estes modelos, fundamentados em lógica computacional, probabilidade e estatística, têm conduzido a inovações significativas no campo da Inteligência Artificial.

O contínuo avanço tecnológico trouxe consigo um aumento notável na capacidade de processamento. Isso possibilitou a evolução das RNAs para o que hoje é conhecido como Aprendizado Profundo. Este subcampo da aprendizagem de máquinas tem revolucionado diversas áreas, desde visão computacional a processamento de linguagem natural. As redes neurais profundas têm a habilidade de modelar e processar informações com uma complexidade sem precedentes, e como resultado, estamos presenciando avanços surpreendentes em diversas aplicações práticas.

A revolução trazida pelas RNAs e, posteriormente, pelo Aprendizado Profundo, é refletida não apenas em avanços acadêmicos, mas também em aplicações práticas. Setores como saúde, finanças, transporte e entretenimento foram drasticamente transformados. Na área médica, por exemplo, sistemas de diagnóstico assistido por computador estão sendo implementados para identificar doenças com precisão comparável ou, em alguns casos,

superior à dos especialistas. No setor financeiro, algoritmos de Aprendizado Profundo são empregados para detectar fraudes e automatizar decisões de investimentos, tornando os processos mais rápidos e precisos.

No entanto, a implementação e generalização do Aprendizado Profundo e das RNAs não estão isentas de desafios. A necessidade de grandes conjuntos de dados para treinamento, a explicabilidade dos modelos e a otimização de sua arquitetura são questões em aberto na pesquisa. Assim, enquanto celebramos os progressos e benefícios destas tecnologias, é imperativo abordar e resolver os desafios inerentes para garantir seu uso eficaz.

3.4 Redes Neurais Convolucionais - CNN

Ao longo de mais de duas décadas, as Redes Neurais Convolucionais (CNNs) têm sido objeto de exploração em diversas vertentes de pesquisa, especialmente para a segmentação, detecção e reconhecimento de objetos e regiões em grandes conjuntos de imagens, caracterizados por dados rotulados em larga escala. Contudo, um marco significativo para as CNNs surgiu a partir de 2012, quando essas redes exibiram um desempenho notável durante a competição *ImageNet*. Esse evento teve um impacto profundo no desenvolvimento de aplicações e estudos nesse domínio específico. Na ocasião, mais de um milhão de imagens distribuídas em mil categorias distintas foram empregadas, e os resultados obtidos foram notáveis, solidificando a posição das CNNs (LECUN; BENGIO; HINTON, 2015). Esse avanço representou, efetivamente, uma grande revolução na área de visão computacional, impulsionada tanto pelo expressivo progresso na capacidade de processamento ao longo das últimas décadas quanto pelo desenvolvimento de novas tecnologias de hardware específicas.

A técnica utilizada pelas Redes Neurais Convolucionais (CNNs) envolve a aplicação de vários filtros distintos sobre dados visuais, como imagens, preservando a conexão espacial entre *pixels* adjacentes ao longo do processo de convolução. Isso significa que as entradas para os neurônios ocultos em uma dada camada, denominada x , são derivadas de um subset específico da camada precedente, referida como camada $x-1$. A Figura 3 (A) ilustra essa interconexão entre as camadas, evidenciando que, no método das CNNs, cada filtro f_i é aplicado de forma repetida em toda a extensão da imagem. Dessa maneira, as unidades que são replicadas compartilham os mesmos parâmetros, incluindo vetores, pesos e polarizações, criando um mapa de características, como demonstrado na Figura 3 (B).

Portanto, entende-se que o mapa de características apresentado na Figura 3 (B) é composto por três neurônios da camada oculta. Assim, fica claro que cada uma das camadas ocultas é constituída por uma multiplicidade de mapas de características. O gradiente do peso compartilhado é obtido pela agregação dos gradientes dos parâmetros

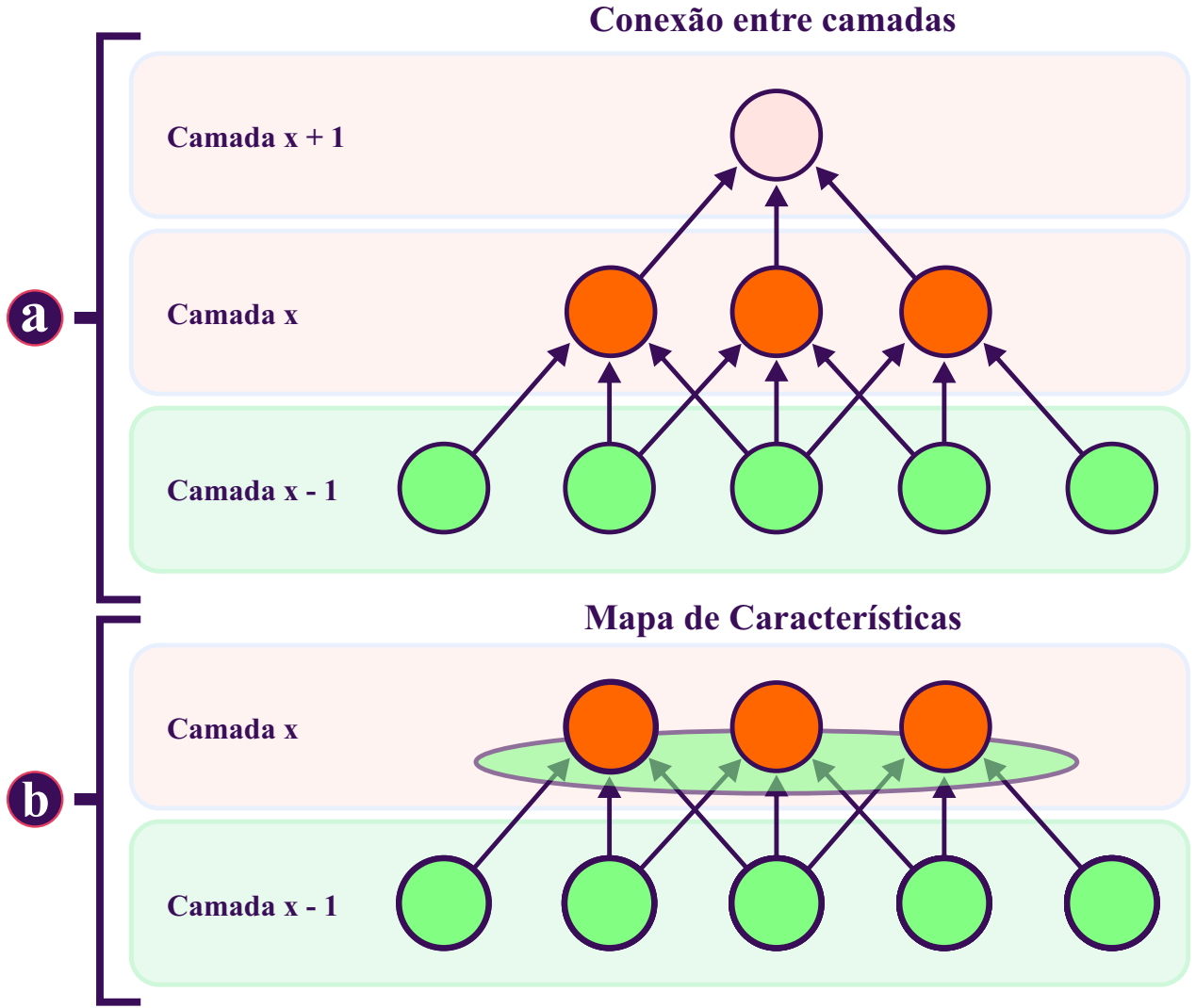


Figura 3 – (A)- Ilustração da ligação (conexão) entre as camadas de uma rede CNN. (B)- Representação de 3 neurônios da camada oculta interligados ao mesmo Mapa de Características. Fonte: O próprio autor.

que são compartilhados, facilitando o cálculo do k -ésimo mapa de características f , com os filtros determinados pelos pesos w^k e pelo viés b_k , conforme descrito na equação 3.1.

$$f_{ij}^k = \tanh((w^k * x)_{ij} + b_k) \quad (3.1)$$

Neste contexto, a duplicação das unidades facilita a detecção de características, e o compartilhamento de pesos melhora a eficiência do aprendizado ao mesmo tempo que diminui o total de parâmetros que precisam ser aprendidos. Os pesos das camadas ocultas são representados por um tensor de quatro dimensões (4D), com elementos correspondentes a cada par de mapas de características, incluindo informações sobre o destino, a origem, e suas posições (vertical e horizontal). Quanto às polarizações b , elas são descritas por um vetor dedicado, com um elemento para cada mapa de características alvo.

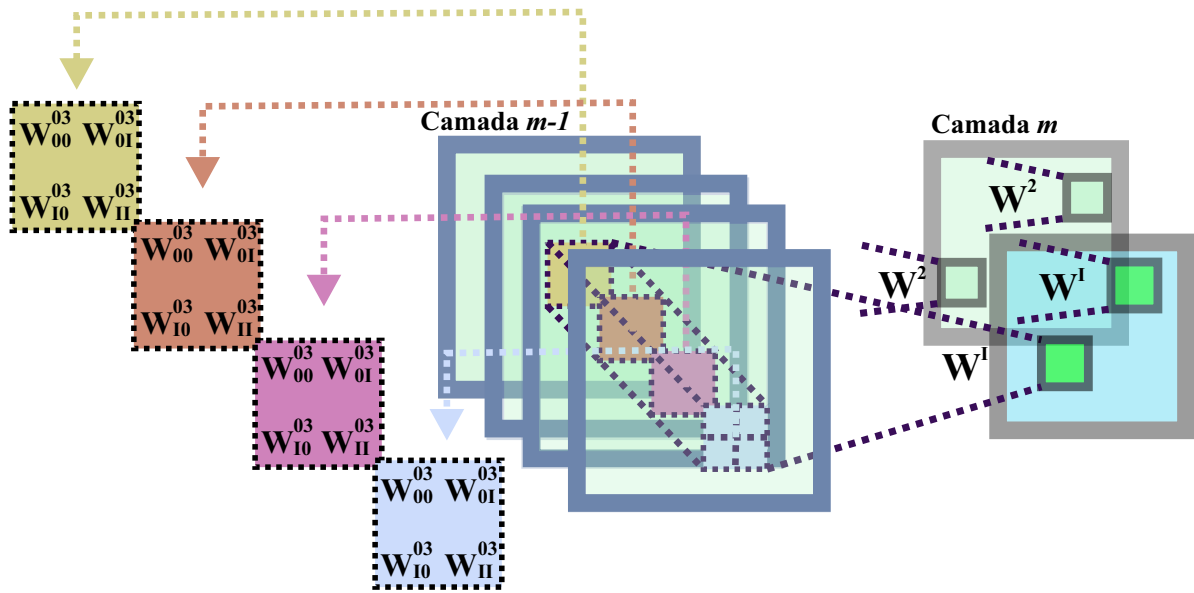


Figura 4 – Ilustração do processo de camadas convolucionais. Fonte: Adaptada de <<https://repositorio.ufc.br/handle/riufc/72342>>.

A Figura 4 detalha os componentes específicos das camadas envolvidas no processo convolucional. Observa-se que a camada $m - 1$ é composta por quatro mapas de características, enquanto a camada oculta m processa dois mapas de características identificados como $(f^1$ e $f^2)$. A extração dos valores de pixel para os neurônios $(f^1$ e $f^2)$ na camada oculta m , simbolizados pelos quadrados azuis e vermelhos, é feita a partir dos pixels presentes na camada $m - 1$, conforme demonstrado na Figura 4. Posteriormente, esses valores são encaminhados para as camadas convolucionais.

As camadas convolucionais incorporam os quatro mapas de características que constituem suas entradas. Portanto, percebe-se que os pesos w^0 e w^1 para os neurônios f^1 e f^2 são expressos através de tensores 3D. A dimensão principal desses tensores organiza-se indexando os mapas de características de entrada, enquanto as outras duas dimensões apontam para as localizações (coordenadas) dos pixels. Isso implica que o peso liga cada pixel do k -ésimo mapa de características na camada m ao pixel situado nas coordenadas (i, j) do l -ésimo mapa de características na camada $m - 1$.

É crucial destacar o papel do *Max-Pooling* nas Redes Neurais Convolucionais. O *Max-Pooling* opera como um mecanismo de agrupamento máximo, partindo a imagem de entrada em conjuntos retangulares não sobrepostos e, para cada um desses conjuntos, seleciona-se o valor máximo. Esta característica do *max-pooling* minimiza a carga computacional nas camadas subsequentes, criando uma conexão em cascata eficaz entre a camada de *max-pooling* e as camadas de convolução.

A ilustração na Figura 5 demonstra o processo convolucional numa Rede Neural Convolucional. A área destacada na imagem de entrada representa a primeira camada, que

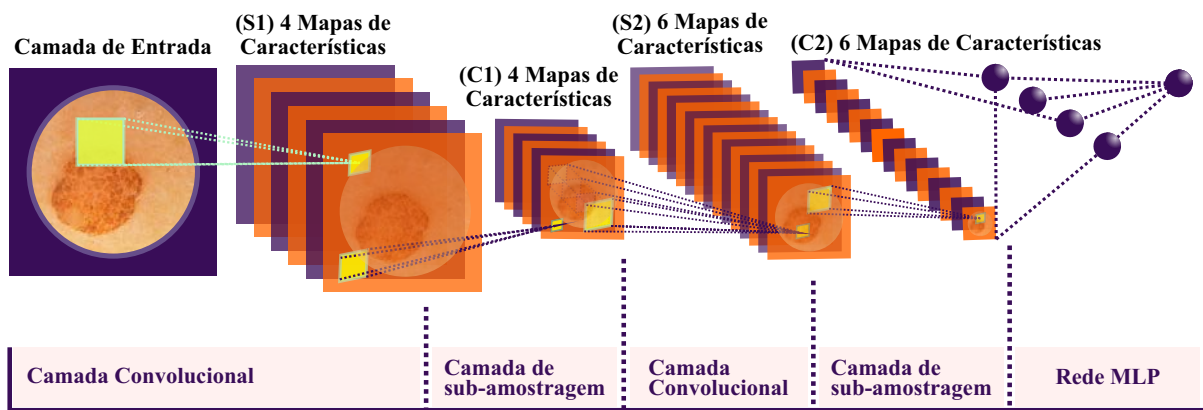


Figura 5 – Ilustração de diferentes camadas encarregadas em um processo de convolucional.
Fonte: o próprio autor.

é processada através de camadas convolucionais alternadas com camadas de *max-pooling*, seguindo a metodologia descrita pela equação 3.1 para a identificação de características significativas.

Nas etapas subsequentes, as camadas são totalmente conectadas, compreendendo uma camada oculta que antecede uma regressão logística. É importante ressaltar que a entrada para a primeira camada totalmente conectada é composta pelo agrupamento de todos os mapas de características provenientes da camada $x - 1$, tal como é evidenciado na Figura 3 (B).

3.5 Frameworks de Detecção

Os frameworks de detecção são ferramentas cada vez mais utilizadas na literatura, sobretudo para experimentações de casos, com variações de parâmetros e condições de imagem. Frameworks como a Detectron2, desenvolvida pelo Facebook AI Research (FAIR) para a implementação de algoritmos de visão computacional com foco particular na detecção de objetos e segmentação de instâncias, vale ressaltar que o framework também tem suporte com bibliotecas nativas de inteligência artificial, como a biblioteca PyTorch (YAGÜE et al., 2022).

Outro framework bem difundido na literatura é o TensorFlow, Desenvolvido pelo Google Brain, o TensorFlow é um framework de aprendizado de máquina de código aberto para treinamento, construção e implantação de modelos de inteligência artificial. Embora seja utilizado para uma ampla gama de tarefas de IA, TensorFlow também inclui suporte para visão computacional (SANCHEZ; ROMERO; MORALES, 2020).

Dentre vários frameworks conhecido na literatura, este estudo optou pelo YOLO, devido à praticidade de aplicação bem como a necessidade de um baixo conjunto de dados

para obtenção de bons resultados de detecção.

3.5.1 Framework YOLO

O *framework* YOLO (*You Only Look Once*) ([REDMON et al., 2016](#)) destaca-se pela sua eficiência e precisão na detecção de objetos em imagens, executando a identificação em um único processo por meio de uma rede neural. Desde sua concepção inicial, a família YOLO passou por diversas iterações, cada uma construída sobre as versões anteriores para superar limitações e aprimorar o desempenho. As iterações do YOLO abrangem desde o YOLOv1 em 2015 até a mais recente versão lançada em 2023, denominada YOLOv8 ([TERVEN; C6RDOVA-ESPARZA; ROMERO-GONZÁLEZ, 2023](#)).

No início de 2023, a Ultralytics apresentou ao mundo a mais recente iteração do YOLO, batizada de YOLOv8. A robustez e eficiência do *framework* YOLOv8 mais uma vez se destacaram, consolidando sua posição de destaque na área de visão computacional ([JOCHER; CHAURASIA; QIU, 2023a](#)). Existem um total de 5 variantes da YOLOv8, todas oscilando em parâmetros de precisão média, velocidade de CPU e operações de ponto flutuante:

YOLOv8n (nano): Com um mAP (Mean Average Precision) de 37.3 na escala de 50-95, esta variante oferece a maior velocidade em CPUs (80.4 ms) e no A100 TensorRT (0.99 ms), com o menor número de parâmetros (3.2 M) e operações de ponto flutuante (FLOPs) (8.7 B). Isso a torna ideal para ambientes com recursos limitados onde a velocidade é crucial.

YOLOv8s (small): Apresenta um equilíbrio entre velocidade e precisão com um mAP de 44.9. Sua velocidade é moderada em CPUs (128.4 ms) e A100 TensorRT (1.20 ms), tendo um número maior de parâmetros (11.2 M) e FLOPs (28.6 B) em comparação ao YOLOv8n, o que a torna adequada para uma ampla gama de aplicações.

YOLOv8m (medium): Com um mAP de 50.2, esta variante é projetada para oferecer um alto desempenho sem comprometer significativamente a velocidade (234.7 ms em CPUs e 1.83 ms em A100 TensorRT). Com 25.9 M de parâmetros e 78.9 B de FLOPs, é ideal para tarefas que necessitam de uma detecção de objetos mais precisa.

YOLOv8l (large): Oferece alta precisão (mAP de 52.9) à custa de uma menor velocidade (375.2 ms em CPUs e 2.39 ms em A100 TensorRT) e maior complexidade (43.7 M de parâmetros e 165.2 B de FLOPs). É indicada para cenários onde a precisão é mais valorizada que a velocidade.

YOLOv8x (extra-large): A variante mais avançada, com o maior mAP (53.9), também é a mais lenta (479.1 ms em CPUs e 3.53 ms em A100 TensorRT) e a mais complexa, possuindo 68.2 M de parâmetros e 257.8 B de FLOPs. É adequada para aplicações que demandam a mais alta precisão em detecção de objetos, independentemente

das limitações de recursos.

Essa diversidade de variantes oferece aos desenvolvedores uma flexibilidade sem precedentes para escolher a configuração mais adequada às suas necessidades específicas, garantindo que o YOLOv8 permaneça na vanguarda da inovação em visão computacional (JOCHER; CHAURASIA; QIU, 2023a).

A versatilidade do YOLOv8 se manifesta na sua capacidade de enfrentar desafios variados, incluindo detecção de objetos, segmentação, estimativa de pose, rastreamento e classificação. Ao incorporar avançadas técnicas de detecção e processamento, o YOLOv8 não apenas supera as limitações de suas versões anteriores, mas também estabelece novos padrões de desempenho e precisão.

A arquitetura da YOLOv8 é apresentada na Figura 6. O YOLOv8, derivando inspiração do YOLOv5 (JOCHER, 2020), emprega uma espinha dorsal ajustada, denominada C2f, em substituição à CSPLayer. Este C2f (*bottleneck* parcial entre estágios com duas convoluções) se destaca por integrar características de alto nível com informações contextuais, aprimorando assim a acuidade na detecção de objetos. (TERVEN; CórDOVA-ESPARZA; ROMERO-GONZÁLEZ, 2023)

No paradigma do YOLOv8, adota-se um modelo sem âncoras, com uma abordagem de cabeça desvinculada para processar distintamente as tarefas de objetividade, classificação e regressão. Essa arquitetura inovadora permite que cada ramo focalize exclusivamente em sua função específica, resultando em um aumento geral na precisão do modelo. Na camada de saída do YOLOv8, a função sigmoid é utilizada como ativação para a pontuação de objetividade, indicando a probabilidade de a caixa delimitadora conter um objeto, enquanto a função softmax modela as probabilidades de classe, refletindo as probabilidades associadas a cada classe possível (TERVEN; CórDOVA-ESPARZA; ROMERO-GONZÁLEZ, 2023).

As funções de perda CIoU (ZHENG et al., 2020) e DFL (LI et al., 2020) são implementadas no YOLOv8 para a perda de caixa delimitadora, e a entropia cruzada binária é empregada para a perda de classificação. Essas abordagens de perda demonstram melhorias significativas no desempenho de detecção de objetos, especialmente em cenários envolvendo objetos de menor escala. (JOCHER; CHAURASIA; QIU, 2023b)

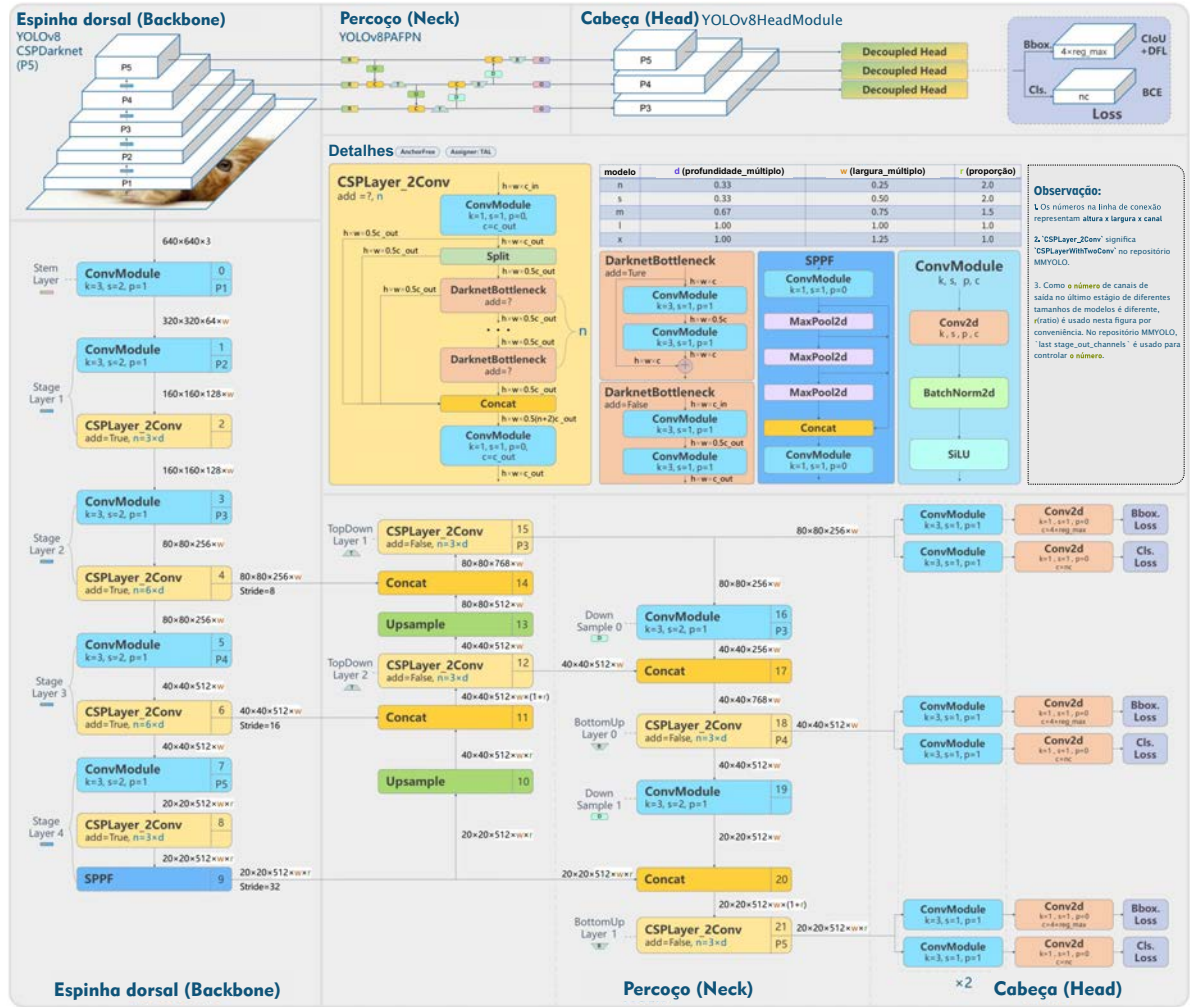


Figura 6 – Ilustra a arquitetura do YOLOv8. Este modelo utiliza uma espinha dorsal chamada C2f, substituindo a CSPLayer. O C2f destaca-se por integrar características de alto nível com informações contextuais, aprimorando a precisão na detecção de objetos. Adotando um modelo sem âncoras e uma abordagem de cabeça desvinculada, o YOLOv8 permite que cada ramo se concentre em tarefas específicas, resultando em um aumento geral na precisão. A camada de saída usa ativação sigmoid para pontuação de objetividade e softmax para probabilidades de classe. As funções de perda CIoU e DFL são implementadas para melhoria na detecção de objetos, especialmente em cenários com objetos de menor escala. Fonte: diagrama baseado em (TERVEN; C6RDOVA-ESPARZA; ROMERO-GONZÁLEZ, 2023) e (CONTRIBUTORS, 2023).

O modelo YOLOv8n destaca-se por sua velocidade, atingindo um tempo de inferência de 80.4 ms em CPUs, mas com uma precisão mAP de 37.3. Esta versão é particularmente útil em contextos que demandam rapidez, embora, para tarefas que requerem maior precisão, como na segmentação de lesões cutâneas, melhorias futuras na arquitetura da rede poderiam elevar a precisão sem impactar drasticamente o desempenho. (JOCHER; CHAURASIA; QIU, 2023c)

YOLOv8s proporciona uma harmonia entre velocidade e precisão, alcançando um

mAP de 44.9. Com um tempo de inferência de 128.4 ms em CPUs, posiciona-se como uma opção intermediária. Acredita-se que otimizações futuras no ajuste fino do modelo poderiam aprimorar sua precisão, tornando-o mais apropriado para identificação de lesões cutâneas complexas.([JOCHER; CHAURASIA; QIU, 2023c](#))

YOLOv8m, com um mAP de 50.2, é projetado para aplicações mais amplas, apresentando um tempo de inferência de 234.7 ms. A incorporação de conjuntos de dados específicos de lesões cutâneas no treinamento pode potencializar a capacidade do modelo de detectar peculiaridades do melanoma com maior sensibilidade.([JOCHER; CHAURASIA; QIU, 2023c](#))

O modelo YOLOv8l, que alcança um mAP de 52.9, é indicado para aplicações que exigem elevada precisão. No entanto, seu tempo de inferência de 375.2 ms pode restringir seu uso em aplicações de tempo real. Ajustes na estrutura da rede, visando a redução da complexidade sem perder precisão, são caminhos possíveis para otimização da velocidade.([JOCHER; CHAURASIA; QIU, 2023c](#))

Por fim, YOLOv8x é o modelo de maior precisão, com um mAP de 53.9, embora seja o mais lento, apresentando um tempo de inferência de 479.1 ms. Este modelo é ótimo para análises detalhadas onde a precisão é crucial. Técnicas de otimização, como quantização e poda da rede, podem ser exploradas para melhorar a eficiência de processamento.([JOCHER; CHAURASIA; QIU, 2023c](#))

As diversas versões do YOLOv8 marcam um progresso importante na detecção de objetos, oferecendo um leque de opções para satisfazer as variadas exigências de aplicação. A otimização contínua desses modelos, particularmente para tarefas especializadas como a detecção de lesões de melanoma em imagens dermatoscópicas, constitui uma área de pesquisa ativa e promissora.

3.6 Segmentação - Ajuste Fino (*Fine-tuning*) e Otimização dos Modelos

A segmentação de imagens dermatoscópicas, especialmente em aplicações médicas para a identificação de lesões cutâneas, requer uma abordagem metódica e detalhada. Esta seção combina diferentes técnicas e modelos utilizados em nosso estudo, focando no ajuste fino e otimização dos processos de segmentação.

3.6.1 Crescimento de Região

O Crescimento de Região representa uma técnica de segmentação de imagens que se baseia na ideia de começar a segmentação a partir de pontos específicos, chamados pontos de semente, e expandi-los por agregação de pixels vizinhos que compartilham certas

características similares, como intensidade de cor ou textura. Este método é particularmente valorizado por sua capacidade de adaptar-se às formas específicas das regiões de interesse na imagem, oferecendo uma abordagem intuitiva e eficaz para a segmentação.

O algoritmo de crescimento de região inicia com a seleção de pontos de semente, seguido pela agregação de pixels vizinhos baseada em critérios de similaridade predefinidos. O processo é iterativo e continua até que não haja mais pixels que satisfaçam os critérios de similaridade para serem agregados. A expansão da região pode ser expressa pela seguinte relação:

$$R_{n+1} = R_n \cup \{p \in N | S(p, R_n) \leq T\} \quad (3.2)$$

onde R_n representa a região após a n -ésima iteração, p denota um pixel candidato pertencente ao conjunto de vizinhos N de R_n , S é uma função de similaridade que mede a afinidade de p para com a região R_n , e T é um limiar de similaridade.

O critério de similaridade é fundamental para determinar como a região cresce. Por exemplo, a similaridade baseada na intensidade pode ser definida como a diferença absoluta entre a intensidade do pixel e a média da região, tal que:

$$S(p, R) = |I(p) - \mu_R| \quad (3.3)$$

onde $I(p)$ é a intensidade do pixel p e μ_R é a média de intensidade dos pixels na região R . A região se expande incluindo pixels cuja diferença de intensidade está dentro de um limiar T .

3.6.2 Operações Morfológicas

A aplicação da morfologia matemática, fundamentada na teoria da interseção de áreas, emerge como uma ferramenta crucial no processamento de imagens, especialmente na detecção e segmentação de regiões pela homogeneidade dos pixels. As técnicas morfológicas, como Erosão e Dilatação, baseiam-se na manipulação de conjuntos de pixels a partir da convolução por elementos estruturantes para manipular os níveis de cinza do mesmo, de acordo com sua intensidade e posição na matriz bidimensional, no caso a imagem, promovendo ajustes na representação visual da região de interesse, no contexto da aplicação, as operações morfológicas são de grande ajuda no ajuste da máscara binária de segmentação.

A Erosão, uma das operações morfológicas mais aplicadas, é definida pela equação 3.4:

$$A \varepsilon B = \{z \in E | B_z \subseteq A\} \quad (3.4)$$

onde A representa a imagem original, B é o elemento estruturante e E é o espaço discreto ou contínuo. Essa operação, no geral, é aplicada na imagem com o propósito de ampliar os

níveis de cinza em tom escuro na imagem. Em imagens binárias, o efeito é de redução nas bordas dos objetos com nível de cinza claro, a redução se dá de acordo com o elemento estruturante aplicado, sendo o mais comum, o elemento quadrado de ordem 3x3. cujos valores internos são 1. No contexto dermatoscópico, a Erosão pode ser interpretada como uma ferramenta eficaz para simplificar detalhes indesejados, facilitando assim a identificação de características relevantes (COMER; III, 1999a).

A operação de dilatação na morfologia matemática é definida conforme a equação 3.5:

$$A \delta B = \{z \in E | B_z \cap A \neq \emptyset\} \quad (3.5)$$

onde A representa a imagem original, B é o elemento estruturante e E é o espaço discreto ou contínuo. A dilatação é aplicada na imagem com o intuito de se aumentar a proporção dos pixels de nível de cinza mais claro, aumentando certos atributos na imagem, que dizem respeito à intensidade clara. Nesta operação, cada pixel tem seu valor modificado de acordo com o cálculo de pesos do elemento estruturante. Para uma imagem binária por exemplo a dilatação irá aumentar a massa de pixels brancos de uma região. No contexto dermatoscópico, esta operação é aplicada com o intuito de ampliar a borda da região externa, de acordo com o elemento estruturante (COMER; III, 1999a).

Existem uma outra gama de aplicações envolvendo operações morfológicas, na qual a aplicação de diferentes elementos estruturantes, em diferentes sequências, ocasiona em efeitos distintos, como esqueletonização de imagens binárias, abertura, fechamento e reconstrução (HANBURY; SERRA, 2001).

Além disso, as operações morfológicas têm uma base teórica sólida, sendo fundamentadas na teoria dos conjuntos. Elas são aplicadas a uma imagem representada como um conjunto X em um espaço discreto ou contínuo, onde X representa os pontos no primeiro plano da imagem e X^c representa o complemento de X . A transformação da imagem é realizada por meio de outro conjunto conhecido como elemento estruturante, cuja forma e tamanho influenciam diretamente a imagem resultante (COMER; III, 1999b).

3.6.3 Agrupamento *K-means*

O K-means é um algoritmo de agrupamento que divide um conjunto de dados em *clusters*, sendo cada ponto de dado, atribuído a um cluster próximo, com base na distância euclidiana dos centroides desses objetos, centroide este calculado a partir dos atributos de cada dado. Durante o processo de ajuste para obtenção dos valores ótimos dos centroides, a posição dos mesmos é recalculada iterativamente de acordo com os pontos próximos. A quantidade de *clusters* é selecionada como parâmetros (PHAM; DIMOV; NGUYEN, 2005). A equação 3.6 ilustra a atualização dos centroides, recalculados a a partir da média dos pontos atribuídos à cada *cluster*.

$$\mu_i^{(t+1)} = \frac{1}{|S_i^{(t)}|} \sum_{x_j \in S_i^{(t)}} x_j \quad (3.6)$$

Esta equação representa a atualização dos centroides μ_i na iteração $t + 1$, onde $S_i^{(t)}$ é o conjunto de pontos atribuídos ao *cluster* i na iteração t . O operador $|\cdot|$ denota o tamanho do conjunto, ou seja, o número de elementos em $S_i^{(t)}$.

A clusterização pelo método K-means realiza a homogeneização do nível de cinza dos grupos de pixels presentes na imagem de acordo com a quantidade de K *clusters* selecionados. Dado o valor de K , o algoritmo realiza iterações não supervisionadas para encontrar valores estáveis de limiares de nível de cinza para cada K . Após a descoberta destes valores de limiares, os níveis de cinza da imagem são agrupados em um dos K existentes e tem seu valor substituído pelo valor limiar. Em imagens de cinza, o efeito notado é a homogeneização das cores na imagem de acordo com o valor K selecionado. Este método pode vir a facilitar a atuação de métodos de segmentação baseado em regiões (LIKAS; VLASSIS; VERBEEK, 2003).

3.6.4 Janelamento de Parzen

O método probabilístico de Janelamento de Parzen é utilizado para estimar a densidade de probabilidade. Ao empregar uma função de kernel, geralmente Gaussiana, esse método desempenha um papel fundamental na suavização da segmentação e na adaptação das fronteiras das lesões, considerando a proximidade dos pontos de dados ao ponto de interesse (REBOUÇAS et al., 2021).

O janelamento de Parzen é um método não paramétrico de estimação de funções de densidade de probabilidade, ele estima a probabilidade do píxel pertencer ou não a um determinado domínio, através da convolução da matriz de imagem por um elemento estruturante, se utilizando de cálculos probabilísticos como mostra a equação 3.7

$$p(z) = \frac{1}{n} \sum_{i=1}^n \delta_n(z - z_i) \quad (3.7)$$

Na qual $p(z)$ representa a probabilidade do píxel pertencer ou não à região, z representa o píxel, δ é a função kernel usada para limitar a vizinhança do píxel, h é o tamanho da aresta da região, e n é o número total de píxels da região (YEUNG; CHOW, 2002).

A função kernel, ou elemento estruturante selecionado foi a Gaussiana por dois motivos. Primeiro, a função Gaussiana é suavizada e, portanto, a função de densidade estimada $p(z)$ também varia suavemente. O Segundo motivo, se assumida uma forma especial da família Gaussiana na qual a função é radialmente simétrica, a função pode ser

especificada completamente apenas por um parâmetro de variação. Assim, $p(z)$ pode ser expresso como uma mistura de núcleos gaussianos radialmente simétricos com variância comum δ :(VINCENT; BENGIO, 2003), Determinada pela equação 3.8:

$$\delta(z) = \frac{1}{\sqrt{(2\pi)^n |C|}} e^{-\frac{1}{2}(z_i - z)' C^{-1} (z_i - z)} \quad (3.8)$$

Em que C é a matriz de covariância $d \times d$ onde d representa a dimensão, $|C|$ é o determinante de C , z e z_i são os pixels.

Em síntese, o Janelamento de Parzen atua sobre a imagem, criando um plano sobreposto, na qual a máscara binária terá suas bordas e formatos ajustados por uma força probabilística advinda dos cálculos fornecidos pela função de densidade de probabilidade, que irão acusar se os pixels localizados na fronteira da borda deverão ser ou não englobados como parte da região de interesse. No contexto da aplicação, os cálculos são realizados sobre os níveis de cinza da região do melanoma. O elemento estruturante gaussiano suavizando esta transição, de maneira com a qual o ajuste não descaraterize a região segmentada.

3.6.5 *Fine-tuning* no processo de segmentação

O ajuste fino, ou *fine-tuning*, desempenha um papel crucial na otimização dos processos de segmentação de imagens, especialmente em contextos críticos, cuja melhora da performance, mesmo que mínima, acarretará melhorias no sistema, como em modelos de auxílio ao diagnóstico diagnóstico médico. Nesta subseção, exploramos como o *fine-tuning* pode ser estrategicamente aplicado para elevar a precisão e a eficácia da segmentação em aplicações médicas.

Inicialmente, a segmentação de imagens pode ser realizada utilizando métodos convencionais, como o Janelamento de Parzen, Crescimento de Região, dentre outros ou o algoritmos de inteligência artificial, como algumas CNNs voltadas para a segmentação, como a *Mask R-CNN* (HE et al., 2017) e a *Detectron2* (PHAM; PHAM; DANG, 2020). Em suma, o processo de *fine-tuning* aplicado á modelos clássicos de segmentação começa com uma análise cuidadosa dos resultados iniciais, que serão utilizados como entrada do mesmo, servindo de semente. Em seguida, podem haver estudos á cerca dos ajustes meticulosos dos parâmetros desses métodos, para que se alinhem melhor com as especificidades das imagens em análise, Algoritmos de *Grid Search* ou alguma outra sub rotina pode ser implementada com a finalidade de se encontrar esses parâmetros ótimos. Por exemplo, a estrutura do *kernel* no Janelamento de Parzen, que é gaussiano para suavizar, mas poderia ser de outras formas, pode ser também ajustados os parâmetros de força de ajuste de bordas, e a quantidade de iterações possíveis que o Janelamento irá realizar, oque acarretará em ajustes mais ou menos abruptos por parte do mesmo. No Crescimento de Região, poderão ser aplicados diferentes limiares de nível de cinza, e *kerneis* de diferentes

dimensões, o que acarretará em um crescimento mais ou menos sensível à variação de nível de cinza de região.

A segunda forma encontrada de aplicar o *fine-tuning* em um método, seria transformando-o em mais uma etapa metodológica. Normalmente, o *fine-tuning* aplicado desta maneira será considerado a etapa final de melhora da segmentação do modelo, e normalmente consiste em um conjunto de técnicas, que se relacionam entre si, sendo aplicadas em conjunto sobre a máscara binária de segmentação obtida pelas etapas anteriores do método, como uma semente para o *fine-tuning* (MARQUES et al., 2022).

O *fine-tuning* aplicado no método proposto por este estudo se dá na vertente de aplicação como sendo o último passo do método, sendo constituído por um conjunto de modelos criados a partir da combinação em cadeia de métodos de Processamento Digital de Imagens, sendo esses métodos: Crescimento de Região, Agrupamento K-means e Janelamento de Parzen. Os modelos foram gerados pelas seguintes combinações:

- **Janelamento de Parzen:** Este modelo *fine-tuning* é composto pela aplicação do Janelamento de Parzen sobre a máscara binária obtida da região de interesse. O processo de ajuste de borda probabilístico do Janelamento ocorre com a máscara binária sendo remodelada em um plano sobreposto ao plano do exame da imagem, na qual os níveis de cinza são usados para realização do cálculo da probabilidade do pixel ser ou não pertencente à região de interesse.
- **Parzen + Crescimento de Região:** O *fine-tuning* composto pela junção destas duas técnicas funciona com a re-segmentação da região de interesse com o crescimento de região, cuja semente inicial foi o centroide da máscara binária de segmentação da etapa anterior. Após a re-segmentação, o Janelamento de Parzen atua ajustando as novas bordas, com a finalidade de melhorar a performance da re-segmentação.
- **Parzen + Clusterização K-means:** O modelo composto pela junção do Parzen e Clusterização, aplica Janelamento de parzen na máscara binária de segmentação, sobreposto ao plano da imagem do exame cujos pixels foram modificados pela clusterização, de maneira a homogeneizar as regiões de interesse na imagem.

As operações morfológicas são utilizadas como base para a aplicação do janelamento de parzen.

A seguir, cada um dos modelos de Fine-tuning desenvolvidos para esta abordagem será analisado, destacando suas características únicas e seu papel na otimização do processo de segmentação.

A Figura 7 ilustra o funcionamento do modelo Fine-tuning composto pela aplicação do Janelamento de Parzen sobre o exame dermatoscópico. Na primeira etapa, a inicialização

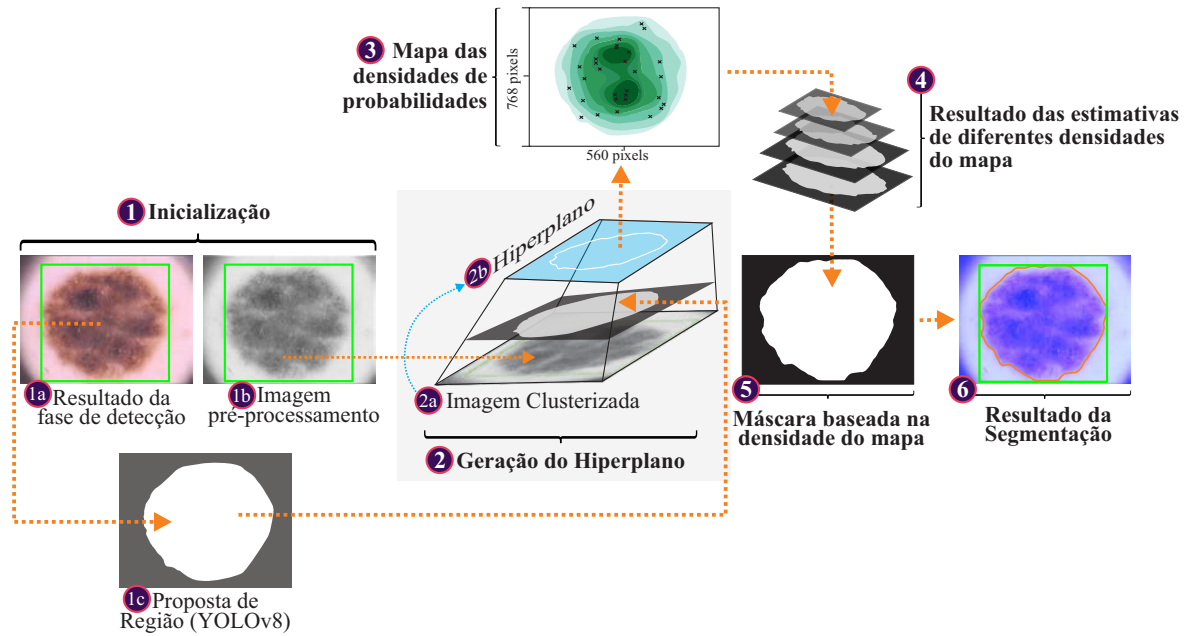


Figura 7 – Representação do modelo fine-tuning do Janelamento de Parzen; 1 - inicialização do método; 2 - Geração do Hiperplano; 3 - Mapa das densidades de probabilidades; 4 - Resultado das estimativas de diferentes densidades do mapa; 5 - Máscara baseada na densidade do mapa; 6 - Resultado da Segmentação.

do método (Passo 1) envolve o recebimento da imagem do exame dermatoscópico e da máscara binária de detecção gerada pelo YOLOv8.

Em seguida, no Passo 2, ocorre a geração do hiperplano. Este é criado pela sobreposição da imagem em tons de cinza do exame pela máscara binária. O hiperplano preserva a posição dos pixels na borda da máscara binária, delineando assim o contorno preciso da região de interesse.

No Passo 3, o hiperplano é utilizado para criar o mapa de densidades de probabilidade. Este mapa consiste em uma coleção de regiões probabilísticas que representam os possíveis níveis de contorno dos pixels pertencentes à região de interesse.

A etapa 4 é responsável pelo cálculo da estimativa de ajuste de borda da região de interesse com base no mapa de densidades de probabilidade. Os parâmetros adotados, como energia de expansão e quantidade máxima de iterações do método influenciam diretamente no comportamento do Janelamento de Parzen, resultando em um ajuste mais suave com valores baixos de parâmetros e um ajuste mais robusto com valores altos.

Na etapa 5, ocorre a criação da máscara baseada na densidade do mapa, representando o resultado da estimativa do ajuste de borda da região de interesse.

Por fim, no Passo 6, observamos o resultado final da segmentação. A máscara de segmentação produzida pelo ajuste fino do Janelamento de Parzen é então sobreposta à imagem original do melanoma, proporcionando uma visualização clara da região

segmentada.

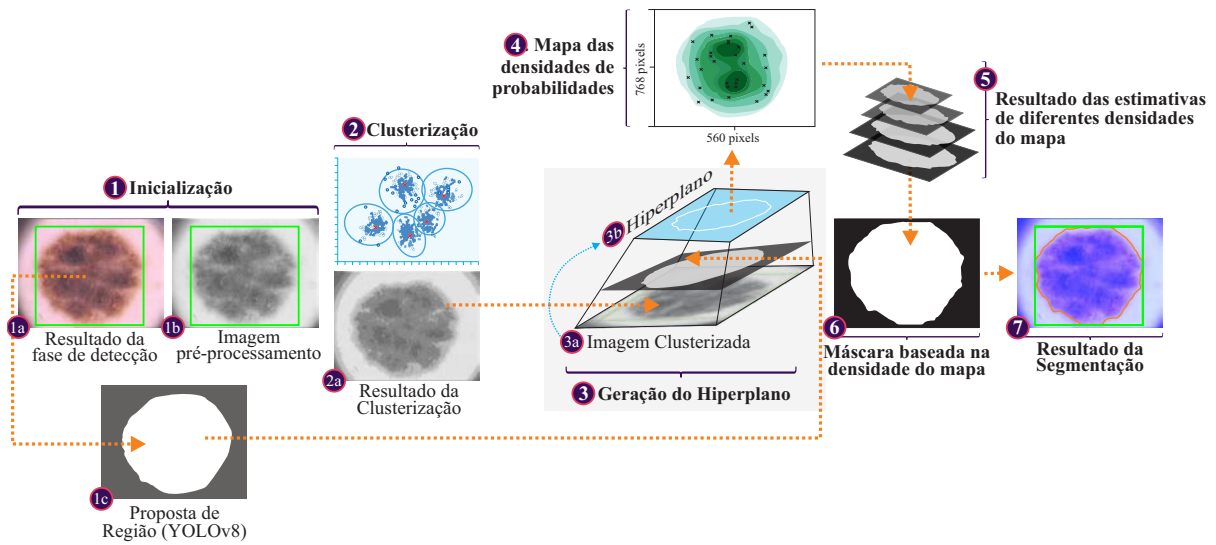


Figura 8 – Representação do modelo Fine-tuning do Janelamento de Parzen com Clusterização; 1 - inicialização do método; 2 - Clusterização; 3 - Geração do Hiperplano; 4 - Mapa das densidades de probabilidades; 5 - Resultado das estimativas de diferentes densidades do mapa; 6 - Máscara baseada na densidade do mapa; 7 - Resultado da Segmentação.

A Figura 8 ilustra o funcionamento do modelo Fine-tuning composto pela combinação do Janelamento de Parzen com a Clusterização pelo método Kmeans, aplicados sobre a saída da detecção. Esse modelo guarda semelhanças significativas com o modelo anterior composto pelo Janelamento de Parzen.

A distinção desse modelo reside no fato de que, ao construir o hiperplano para criar o mapa de densidades de probabilidades, não é utilizada a imagem do exame dermatoscópico, mas sim a clusterização desse mesmo exame, conforme apresentado em 2. A perspectiva é que a aplicação da clusterização torne a região do melanoma mais homogênea, simplificando os cálculos probabilísticos para o ajuste de bordas.

Em termos de segmentação de melanoma, a homogeneidade indicaria que os elementos dentro das áreas do melanoma apresentam características semelhantes, facilitando a identificação precisa das bordas. Em um contexto mais amplo, a homogeneidade na segmentação implica na redução de variações substanciais entre os pixels da região, contribuindo para uma segmentação mais eficiente.

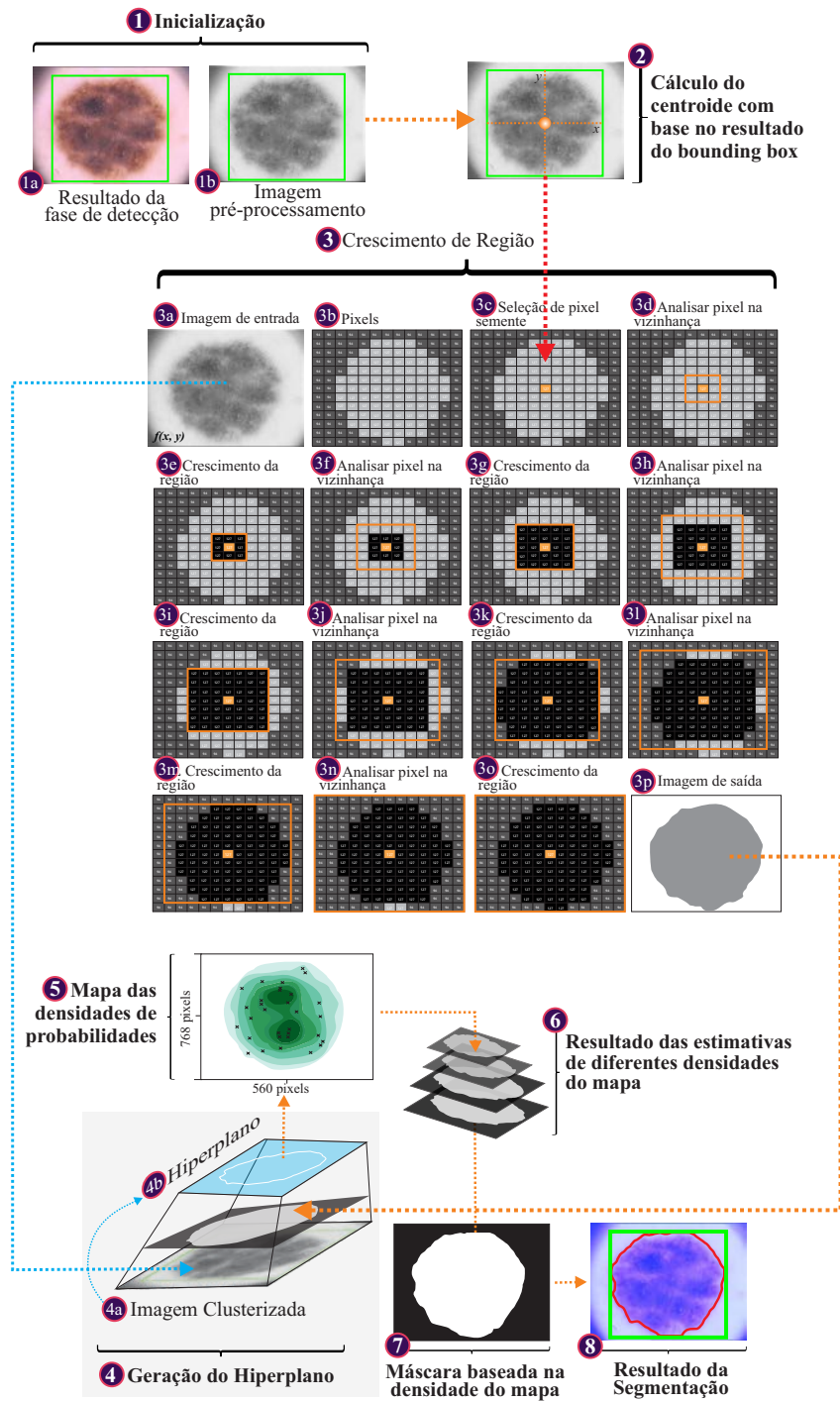


Figura 9 – Representação do modelo fine-tuning do Janelamento de Parzen com Crescimento de Região: 1 - inicialização; 2 - Cálculo do centroide; 3 - Crescimento de Região; 4 - Geração do Hiperplano; 5 - Mapa das densidades de probabilidades; 6 - Resultado das estimativas de diferentes densidades do mapa; 7 - Máscara baseada na densidade do mapa; 8 - Resultado da Segmentação.

A Figura 9 ilustra o funcionamento do modelo Fine-tuning composto pela combinação do Janelamento de Parzen com o Crescimento de Região, aplicados sobre a saída da detecção. Esse modelo guarda semelhanças significativas com o modelo anterior composto apenas pelo Janelamento de Parzen.

Este modelo de ajuste fino remove o centroide da caixa delimitadora e o utiliza como semente no método de crescimento de região, conforme observado no passo 2 e 3. Este método realiza a re-segmentação da área, utilizando um limiar definido (com 40 níveis de cinza acima e abaixo) para otimizar o processo.

Em síntese, o crescimento de região agrega ao pixel semente, os demais pixels vizinhos com base no critério de adesão do limiar, a cada iteração, os vizinhos próximos são aderidos ou não se o seu valor de cinza estiver dentro do limite de 40 níveis de cinza acima ou abaixo do nível de cinza do pixel semente.

Como observado na imagem, em comparação com os outros modelos de ajuste fino, o método de Janelamento de Parzen com Crescimento de Região não requer a máscara binária para inicialização. Isso acontece porque o crescimento de região precisa apenas do pixel semente, o que o diferencia dos outros dois métodos, uma vez que esta abordagem não está diretamente ligado à performance de detecção do YOLO.

Após a etapa de segmentação com crescimento, aplica-se o Janelamento de Parzen para obter uma máscara binária definitiva.

3.7 Teste Estatístico

O teste estatístico de *Kolmogorov-Smirnov* é uma ferramenta fundamental e versátil na análise de dados para comparar distribuições de amostras contínuas. Diferentemente dos testes paramétricos que requerem suposições sobre a distribuição dos dados, o teste K-S é não paramétrico e adequado para situações em que a distribuição subjacente é desconhecida ou não segue um padrão específico, como a distribuição normal ([ASLAM, 2019](#)).

Ao realizar o teste de *Kolmogorov-Smirnov*, formula-se duas hipóteses: a hipótese nula (H_0), que afirma que as duas amostras têm a mesma distribuição, e a hipótese alternativa (H_1), que afirma que as amostras têm distribuições diferentes. Em seguida, coleta-se duas amostras de dados, representando diferentes grupos ou populações que se deseja comparar. Essas amostras devem ser contínuas e ter o mesmo tamanho ([BAUMGARTNER; KOLASSA, 2023](#)).

Para calcular a estatística de teste, ordenam-se os valores de cada amostra em ordem crescente e calcula-se a função de distribuição empírica (ECDF) para ambas as amostras. A ECDF é uma estimativa da função de distribuição cumulativa real com base nos dados coletados e atribui a probabilidade acumulada de encontrar um valor menor ou igual a cada valor da amostra ([OTSU; TANIGUCHI, 2020](#)).

A estatística de teste de *Kolmogorov-Smirnov* é a maior discrepância absoluta entre as duas ECDFs, representando a maior diferença vertical entre elas. Em seguida, calcula-se

o P-valor, que é a probabilidade de obter uma estatística de teste tão extrema quanto a observada, sob a hipótese nula ser verdadeira (WARE et al., 2019).

A interpretação do P-valor é fundamental na tomada de decisão. Se o P-valor for maior que o coeficiente de confiabilidade α escolhido (como 5% ou 3%), não há evidências suficientes para rejeitar a hipótese nula, e conclui-se que as duas amostras têm a mesma distribuição (igualdade estatística). No entanto, se o P-valor for menor que α , rejeita-se a hipótese nula em favor da hipótese alternativa e conclui-se que as duas distribuições são estatisticamente diferentes, como mostra a expressão 3.9, indicando que as amostras têm distribuições diferentes (diferença estatística) (WARE et al., 2019).

$$\text{Teste de } K \quad \begin{cases} P > \alpha & = & H_0 & - & \text{Equivalencia Estatistica} \\ P < \alpha & = & H_1 & - & \text{Diferenca Estatistica} \end{cases} \quad (3.9)$$

Na equação apresentada, P simboliza o valor-p expresso pela estatística D , enquanto α representa o nível de significância adotado para o teste, estabelecido em 1% neste contexto, denotando assim um elevado nível de confiabilidade para o teste realizado. As hipóteses do teste, ou seja, a hipótese de igualdade estatística entre as amostras (H_0) e a hipótese de equivalência estatística entre as amostras (H_1), são fundamentais nesse contexto. A comparação entre o valor-p obtido e a significância adotada determinará se a hipótese de equivalência estatística entre as amostras será confirmada ou refutada. Esse resultado contribui para fortalecer as comparações métricas, evidenciando se há ou não equivalência estatística entre as amostras. Mesmo que apresentem médias e desvios padrões distintos, a análise estatística pode indicar que essa diferença é mínima em termos significativos.

3.8 Métricas de Avaliação

Foram utilizadas algumas métricas de segmentação para a avaliação dos experimentos, além de validar o método proposto mediante comparação com trabalhos relacionados. As métricas específicas utilizadas são delineadas nesta subseção.

A fim de avaliar os diferentes resultados gerados por este estudo, são utilizadas às métricas de sobreposição, tais como Acurácia (ACC), Precisão (PRE) e Sensibilidade (SEN). No tocante às métricas de sobreposição, são utilizadas: Especificidade (SPE), Coeficiente de Correlação de Matthews (MCC), Valor Preditivo Negativo (NPV) e Valor Preditivo Positivo (PPV) (TAHA; HANBURY, 2015).

De modo geral, as métricas abordadas neste estudo envolvem cálculos baseados em Verdadeiro Positivo (TP), Falso Positivo (FP), Verdadeiro Negativo (TN) e Falso Negativo (FN). Essas grandezas, por sua vez, são calculadas em uma matriz de confusão

após a comparação entre os resultados de detecção, classificação ou segmentação do método desenvolvido e o padrão ouro de referência, ou seja, o valor de gabarito.

3.8.1 Acurácia

A Acurácia é uma métrica amplamente utilizada em diversos tipos de problema, como classificação, detecção e segmentação, sendo para este último caso, a acurácia se relaciona diretamente ao número de *pixels* que foram segmentados corretamente sobre o número total de *pixels* segmentados correspondente à região de interesse em uma imagem (ALJABAR et al., 2009). sua fórmula é dada como:

$$ACC = \frac{VP + VP}{VP + VP + FP + FN} \quad (3.10)$$

A acurácia é uma métrica essencial para avaliar a performance geral de algoritmos de segmentação, pois fornece uma visão geral da capacidade do modelo em segmentar corretamente os *pixels* de interesse. No entanto, é importante considerar que, em alguns casos, quando há um desequilíbrio significativo entre as classes (por exemplo, uma classe de interesse com muito mais pixels do que a outra), a acurácia pode não ser suficiente para avaliar completamente a eficácia do modelo. Nesses casos, outras métricas, como a F1-Score ou a sensibilidade, podem ser utilizadas para fornecer uma avaliação mais abrangente do desempenho da segmentação (ALJABAR et al., 2009).

3.8.2 Coeficiente dos dados - (*Dice's coefficient*)

O coeficiente dos dados (DC) é usado para medir a similaridade entre as regiões segmentadas automática e manualmente, e é definido como. Neste métrica a segmentação proveniente do algoritmo e a segmentação do especialista são comparadas por sobreposição.

$$DC = 2 \times \frac{|S_m \cap S_t|}{|A_k| + |A_l|}, \quad (3.11)$$

Conforme mostrado na Equação 3.11, S_m significa a região segmentada automaticamente e S_t o padrão de referência. Quando o valor de sobreposição dos intervalos é igual a 0, não há sobreposição entre as duas regiões, enquanto que quando é igual a 1, a sobreposição é perfeita (DICE, 1945). O valor DC aumenta juntamente com a semelhança entre a região segmentada e a verdade básica (QIAN et al., 2013).

3.8.3 Sensitividade

A sensibilidade é uma métrica de Classificação responsável por avaliar o índice de pixels corretamente segmentados como pertencentes à região de interesse dentre todos os pixels realmente pertencentes à tal grupo (BHADAURIA; DEWAL, 2014). Em síntese, a

sensitividade é uma medida que enfatiza a capacidade do modelo em detectar corretamente os pixels da região de interesse, evitando os falsos negativos, que seriam pixels relevantes erroneamente classificados como não pertencentes à região de interesse (TAHA; HANBURY, 2015). Sua fórmula é dada como:

$$Sen = \frac{VP}{VP + FN} \quad (3.12)$$

3.8.4 Especificidade

A Especificidade é uma métrica de Classificação responsável por indicar o índice de pixels corretamente segmentados como pertencentes à região de fundo dentre todos os pixels realmente pertencentes ao fundo, isto é, os pixels que não pertencem à região de interesse (BHADARIA; DEWAL, 2014). Sua fórmula é dada como:

$$Spe = \frac{VN}{VN + FP} \quad (3.13)$$

A Especificidade é um indicador crucial para avaliar o desempenho geral do algoritmo de segmentação, pois proporciona uma visão complementar à sensibilidade. Enquanto a sensibilidade foca na detecção correta dos *pixels* da região de interesse, a Especificidade enfatiza a precisão na identificação dos *pixels* que não fazem parte dessa região. Juntas, essas métricas fornecem uma avaliação abrangente da qualidade da segmentação, permitindo ajustes e melhorias no modelo para diversas aplicações (TAHA; HANBURY, 2015).

3.8.5 Valor Preditivo Positivo

O PPV é uma métrica crucial para avaliar a capacidade de um modelo em identificar corretamente os pixels que fazem parte da região de interesse em uma imagem, evitando assim falsos positivos. Os falsos positivos referem-se aos pixels erroneamente classificados como pertencentes à região de interesse. Essa métrica é especialmente relevante em aplicações médicas, como a detecção de melanoma em imagens dermatoscópicas, onde é vital minimizar diagnósticos incorretos (BHADARIA; DEWAL, 2014; TAHA; HANBURY, 2015).

A fórmula para calcular o PPV, ou Precisão, é dada por:

$$PPV = \frac{VP}{VP + FP} \quad (3.14)$$

Onde VP representa os verdadeiros positivos (*pixels* corretamente classificados como positivos), e FP representa os falsos positivos (*pixels* erroneamente classificados como positivos). A precisão é, portanto, uma medida essencial para garantir a confiabilidade

do modelo de segmentação, especialmente em situações em que a minimização de falsos positivos é crucial.

3.8.6 Valor Preditivo Negativo

O NPV é uma medida que avalia a proporção de pixels corretamente identificados como pertencentes ao fundo (região que não é de interesse) em relação ao número total de pixels classificados como fundo pelo modelo de segmentação (BHADARIA; DEWAL, 2014). Sua fórmula é dada como:

$$NPV = \frac{VP}{VP + FN} \quad (3.15)$$

3.8.7 Coeficiente de Correlação de Matthews

O Coeficiente de Correlação de Matthews (MCC) é uma métrica altamente informativa que leva em consideração não apenas os verdadeiros positivos e verdadeiros negativos, mas também os falsos positivos e falsos negativos na avaliação de algoritmos de segmentação de imagens, especialmente em situações em que as classes de pixels estão desequilibradas ou quando há interesse específico em avaliar o desempenho do modelo em relação a todas as quatro possíveis classificações do pixel na segmentação de imagens (BHADARIA; DEWAL, 2014). Sua fórmula é dada como:

$$MCC = \frac{VP \times VP - FP \times FN}{\sqrt{(VP + FP)(VP + FN)(VP + FP)(VP + FN)}} \quad (3.16)$$

3.8.8 JAC (Coeficiente de Jaccard)

O Coeficiente de Jaccard é uma métrica utilizada para a avaliação da sobreposição entre a máscara segmentada e a máscara de referência (TAHA; HANBURY, 2015), sua fórmula é dada como:

$$JAC = \frac{TP}{TP + FP + FN} \quad (3.17)$$

Esta métrica é de suma importância na avaliação de proporções e bordas da segmentação, uma vez que, por ser uma métrica calculada a partir da sobreposição das regiões, tende a mensurar de maneira mais rígida a proporção de pixels verdadeiros positivos e falsos positivos, mas além disso, trás um peso adequado aos falsos negativos e falsos positivos em seu cálculo, apontando métricas com médias normalmente mais baixas que as demais, porém com uma perspectiva de avaliação e discussão bastante rica (TAHA; HANBURY, 2015).

4 Método Proposto

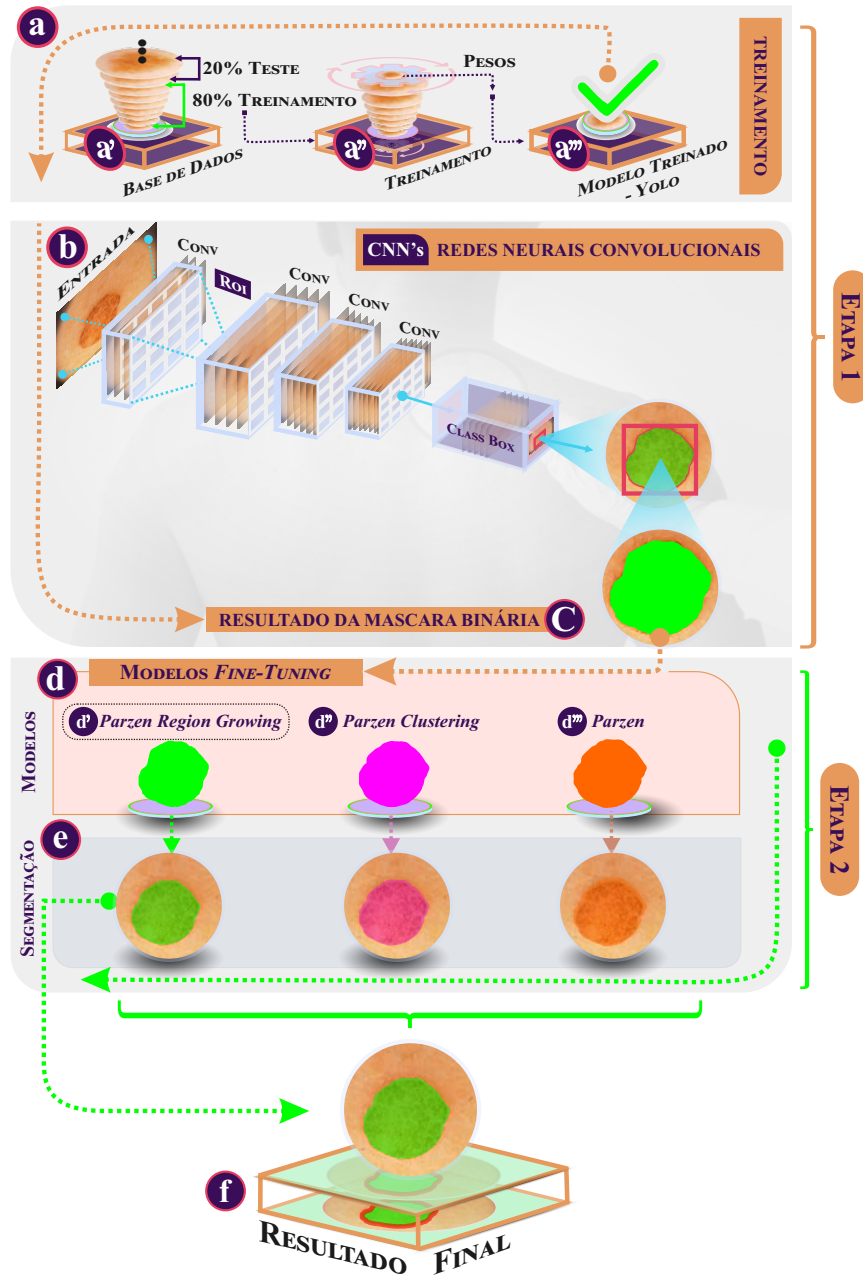


Figura 10 – Ilustra uma visualização da metodologia proposta. Na Etapa (a), destaca-se o treinamento dos modelos YOLO, enquanto a Etapa (b) demonstra a detecção da região do melanoma por meio de aprendizagem profunda. A máscara binária de detecção é ilustrada na Etapa (c), revelando a delimitação da região de interesse. A Etapa (d) mostra o processo de *fine-tuning* dos modelos, e na Etapa (e) são apresentados os resultados obtidos. Por fim, a Etapa (f) destaca o melhor modelo. Fonte: o próprio autor.

Esta seção detalha a metodologia da abordagem proposta para a detecção e segmentação de melanoma em imagens médicas dermatoscópicas. A estratégia desenvolvida fundamenta-se em *deep learning* para identificar a região do melanoma, seguida por uma etapa de segmentação totalmente automática da área previamente detectada. A principal contribuição desta nova abordagem reside na automatização do processo de detecção por meio de aprendizagem profunda, com a utilização de *fine-tuning* para ajustes precisos ao redor da borda da região do melanoma. Essa etapa é crucial para a identificação da região patológica e auxílio ao diagnóstico médico.

4.1 Metodologia

A metodologia proposta para a detecção e segmentação do melanoma é dividida em duas etapas distintas. Na primeira etapa, realiza-se o treinamento dos modelos YOLO, seguido pela detecção da região do melanoma na imagem médica por meio de aprendizagem profunda. Na segunda etapa, procede-se à segmentação da região do melanoma, ou seja, à identificação das bordas do melanoma. Esse processo é conduzido por meio de uma combinação de técnicas computacionais, aliadas a técnica de *fine-tuning* após a detecção da região de interesse.

A Figura 10 proporciona uma visualização clara das duas etapas mencionadas, identificadas pelos números 1 e 2, e seus respectivos estágios, representados pelas letras (a, b, c, d, e, f), bem como suas variações em a' , a'' , a''' , d' , d'' , e d''' . Essa representação gráfica facilita a compreensão do fluxo do processo, destacando cada fase distintiva no desenvolvimento da detecção e segmentação do melanoma, proporcionando uma visão abrangente e detalhada do método proposto.

4.1.1 Etapa 1 - Treinamento de modelos e Detecção do Melanoma

Estágio **a** - A base de dados, composta por 200 imagens dermatoscópicas, é subdividida em conjuntos de treino e teste, seguindo uma proporção de 80/20. Esses conjuntos são, então, submetidos aos diferentes modelos YOLOv8, dando início ao processo de treinamento das amostras para a detecção por caixa delimitadora e máscara binária. É importante salientar que os modelos adotam como referência a demarcação realizada por um médico especialista, estabelecendo assim o padrão-ouro. Durante o treinamento, cada modelo YOLOv8 executa um número específico de épocas, o qual é influenciado pelas dimensões da arquitetura e pelos parâmetros internos da própria rede. Após essa fase, os pesos gerados durante o treinamento são salvos como arquivos de texto. Os estágios relacionados à separação da base de dados, treinamento e geração dos modelos são abordados nos estágios a' , que representa a divisão da base de dados, a'' e a''' que representam, respectivamente, o treinamento e o modelo treinado. Essas etapas são

visualmente apresentadas na Figura 10.

A Figura 11 representa a divisão da base de dados entre treino e teste, sendo as proporções utilizadas para a obtenção dos pesos de detecção do modelo. As imagens de teste são empregadas para a obtenção das métricas de segmentação, discutidas de maneira quantitativa e qualitativa com o objetivo de compreender o comportamento dos modelos em relação à base de dados abordada.

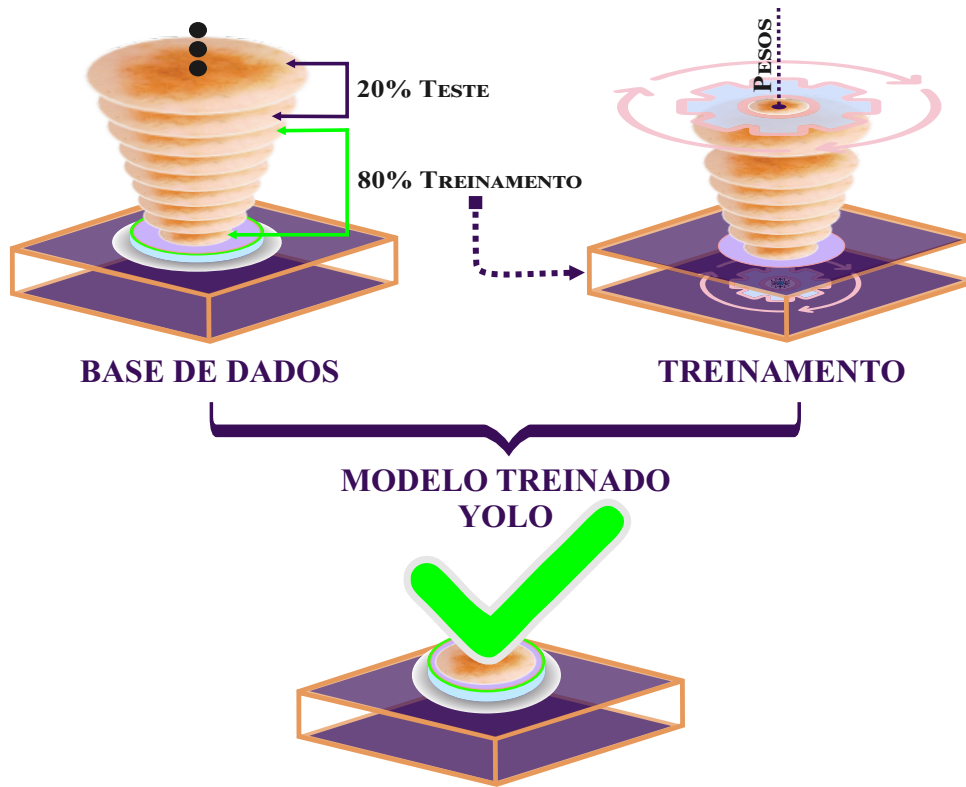


Figura 11 – imagem da etapa de treinamento. Fonte: o próprio autor.

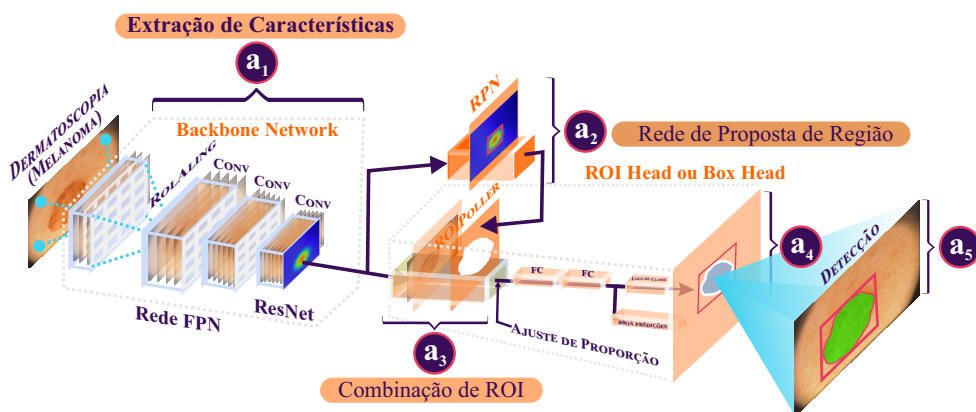


Figura 12 – Funcionamento do Estágio 1 da metodologia proposta com os estágios a1 - a5 sendo ilustrados.

A Figura 12 ilustra de maneira resumida o Estágio 1 da metodologia proposta, com a entrada da imagem médica, o processamento da CNN, ocasionando na detecção do

melanoma via *bounding box* e máscara binária.

Na primeira etapa (A1), ocorre o processo de extração de características, no qual as várias camadas convolucionais da arquitetura neural realizam a extração de atributos profundos. Esses atributos são posteriormente usados para a geração de propostas de região.

Na segunda etapa (A2), através da rede de proposta de região, a arquitetura gera propostas de região utilizando os atributos profundos fornecidos, as propostas buscam delimitar a região e interesse de acordo com os atributos, várias propostas são geradas com o objetivo de se obter a proposta ótima.

Na terceira etapa (A3), todas as regiões de interesse propostas são combinadas e refinadas em uma única. O ROI Poller, ou o mecanismo de pooling de região de interesse, desempenha um papel crucial nesta etapa, pois concatena as ROIs geradas de maneira a abranger a região de interesse com trechos ótimos gerados, em síntese, os pontos de acerto das propostas anteriores são conectados, formando assim a detecção da região pela rede, que inicialmente se dá pela demarcação com caixa delimitadora.

Em seguida, na etapa A4, após a detecção por Caixa delimitadora, a rede neural constrói a máscara binária de detecção. Isso é feito classificando os pixels internos da caixa delimitadora como pertencentes ou não à região de interesse, de acordo com os atributos profundos da mesma, neste caso, a região de interesse na imagem é a afetada pelo melanoma. Os pixels identificados como pertencentes à região do melanoma são então utilizados para criar a máscara binária de detecção, que destaca de forma mais precisa a área suspeita de melanoma na imagem.

Finalmente, na última etapa (A5), é possível visualizar o resultado da detecção realizada através do YOLOv8. Este resultado inclui a máscara binária gerada na etapa anterior, juntamente com a caixa delimitadora que define a região de interesse. Esses elementos são utilizados como inicializadores de quaisquer outras técnicas aplicadas sobre a região do melanoma, como modelos fine tuning para melhoria de detecção e segmentação.

Estágio **b** - Após o treinamento adequado, com a geração de pesos para a detecção da *bounding box* e máscara binária, a *Convolutional Neural Network* (CNN) utiliza esses pesos para realizar a detecção automática do melanoma em imagens dermatoscópicas de teste. Em resumo, a arquitetura da rede convolucional processa a imagem de entrada, realizando transformações espaciais que destacam a classe dos *pixels* na imagem com base nos atributos extraídos.

Dentre as principais camadas da CNN, a camada de RPN é responsável por gerar as propostas de regiões de interesse, onde a presença de melanoma pode ser positiva. A camada de *ROI Pooling* normaliza as dimensões das propostas de regiões de interesse para que sejam processadas pelas camadas de classificação e regressão; regiões de áreas comuns

são unificadas, reduzindo drasticamente as regiões propostas. Vale ressaltar que, nesta etapa, são construídas as caixas delimitadoras sobre as regiões de interesse detectadas. A camada de classificação atribui a classe da região de interesse na demarcação por caixa, enquanto que a camada de regressão realiza ajustes nas caixas, como correções de coordenadas e dimensões.

Embora seja um *framework*, o YOLOv8 possui a mesma lógica de funcionamento das CNNs, com as funcionalidades das camadas principais, entretanto, sua arquitetura é totalmente otimizada para aplicações mais robustas e com baixo tempo de resposta. As diferentes versões do YOLOv8 se distinguem na quantidade de neurônios e nas ligações entre eles, variando de uma arquitetura voltada para respostas rápidas, como o YOLOv8n, até uma arquitetura focada em desempenho e eficácia de detecção, como é o caso do YOLOv8X. O Estágio de Detecção pode ser observado no Estágio b da Figura 10. Os resultados gerados pelos diferentes modelos são avaliados entre si para determinar a variação mais eficaz do YOLOv8 para a base de dados abordada. Os resultados obtidos de cada modelo serão utilizados para a próxima etapa do método proposto.

Estágio c - Após a detecção da região do melanoma mediante a caixa delimitadora e a máscara binária, este segundo objeto é enviado para a Etapa 2 do método proposto. Nesta etapa, diferentes modelos de *fine-tuning* são aplicados e comparados para determinar a melhor combinação do YOLOv8 com *fine-tuning*.

4.1.2 Etapa 2 - Segmentação do Melanoma com Modelos *Fine-Tuning*

Estágio d - Nesta etapa, vários modelos computacionais são desenvolvidos por meio da combinação de três técnicas de processamento digital de imagens: Janelamento de Parzen, clusterização e Crescimento de Região. Os modelos de *fine-tuning* são empregados para delinear com precisão a região de interesse, tendo como ponto de partida a imagem de saída do Estágio c, que contém a detecção da região com o tumor por meio da máscara binária. Este estágio é ilustrado pelo Estágio d da Figura 10.

A Figura 13 ilustra o processo de crescimento do Janelamento de Parzen a partir de uma semente inicial sobreposta à região do melanoma. Mesmo não sendo tão eficaz quando aplicado de maneira isolada e sem uma semente inicial bem estabelecida, o Janelamento de Parzen é capaz de segmentar com precisão satisfatória a região de interesse. Na imagem, é possível perceber que o modelo realiza variações nas bordas de acordo com a densidade de probabilidade estimada para os *pixels* ao redor da máscara por iteração. Dessa forma, o crescimento da máscara binária formada ocorre por etapas, e a quantidade de iterações e a força de crescimento selecionadas devem ser escolhidas de maneira precisa, mesmo que de forma empírica.

O Janelamento de Parzen foi aplicado com uma janela de dimensão 4x4, ou seja,

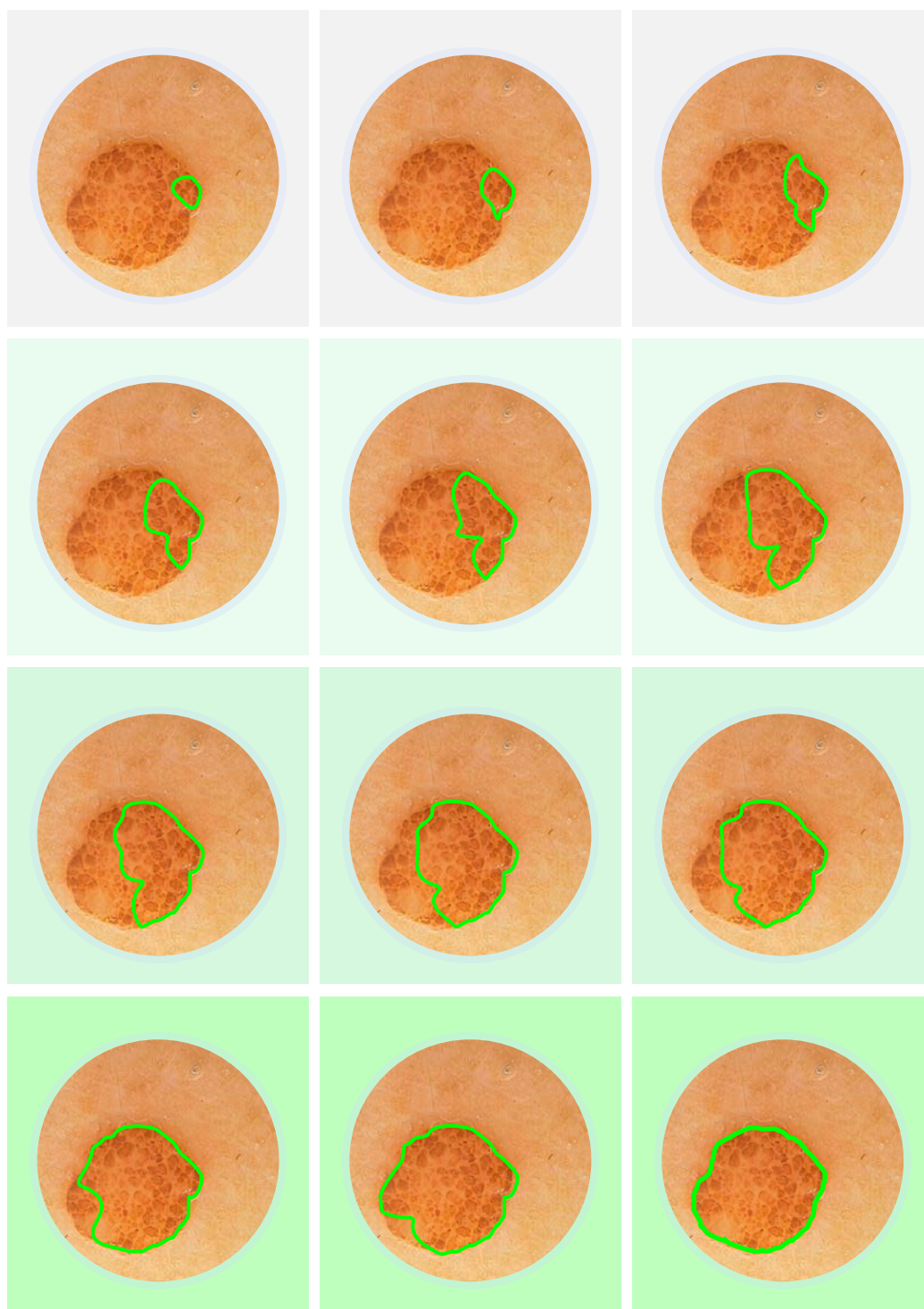


Figura 13 – Imagem da evolução do janelamento de parzen. Fonte: o próprio autor.

uma janela quadrada. O *kernel* utilizado foi um Kernel gaussiano, a fim de suavizar a transição dos *pixels*. O método, por si só, enfrenta dificuldades em algumas imagens dermatoscópicas devido ao baixo contraste entre o final da região da pele não lesionada e o início do melanoma. No entanto, esses casos são mais comuns em aplicações com maior energia de janelamento, quando se deseja abranger grandes áreas com o modelo.

Nesta aplicação, o Janelamento de Parzen foi ajustado com baixa força cinética, de maneira a realizar apenas pequenos ajustes na máscara binária de saída. Dessa forma, o modelo é capaz de expandir a máscara binária sem grandes perdas nas bordas, uma vez que a detecção realizada através do YOLOv8 possui alta performance.

A clusterização pelo método *K-means* foi aplicada com $K=80$, visando agrupar os níveis de cinza da imagem sem introduzir grandes modificações nos tons de cinza. A ideia é agrupar as regiões da lesão, que possuem relativa homogeneidade em relação aos tons do melanoma. Portanto, o valor 80 foi adotado sem consumir abruptamente os *pixels* mais sutis do melanoma na imagem. Assim, a clusterização *K-means* foi utilizada como uma ferramenta de suporte para o Janelamento de Parzen e o Crescimento de Região, suavizando os grupos de níveis de cinza da imagem, nos quais o Janelamento e a regra de adesão do crescimento serão aplicados para os cálculos probabilísticos de adesão de *pixels* à máscara binária de segmentação.

O Crescimento de Região é um método de segmentação amplamente utilizado em abordagens que visam segmentar regiões homogêneas. Quando parametrizado adequadamente, esse método apresenta alta sensibilidade à variação de níveis de cinza em uma região. Para o método proposto, o Crescimento de Região foi aplicado com uma regra de adesão que considera 25 níveis de cinza a mais e a menos em relação ao nível de cinza do *pixel* central da semente utilizada. Quando combinado com a clusterização *K-means*, o Crescimento de Região encontra facilidade em abranger certas regiões, uma vez que estas estão clusterizadas. Além disso, quando aplicado com o Janelamento de Parzen, o Crescimento de Região remodela a área da segmentação para que o Janelamento ajuste as bordas posteriormente.

Estágio e - Os modelos de *Fine Tuning* são aplicados às máscaras binárias de detecção das imagens de teste geradas pelas diferentes versões do YOLOv8. Dessa forma, são gerados 15 modelos no total, onde cada um consiste em uma das 5 versões do YOLOv8 (n, m, l, s, x) combinada com um modelo de *Fine Tuning* (*Parzen*, *Parzen_Clustering* e *Parzen_Clustering_Region_G*). Este estágio é representado pela Etapa (e) na Figura 10.

Estágio f - Após a geração de resultados para as diferentes combinações do YOLOv8 + *Fine Tuning*, as saídas são comparadas para determinar a melhor combinação de segmentação para a base de dados dermatoscópica abordada neste estudo. A Figura 14 ilustra a Etapa 2 da metodologia do estudo proposto.



Figura 14 – Modelos *Fine-Tuning* desenvolvidos a partir dos métodos de processamento digital de imagens. Fonte: o próprio autor.

5 Resultados e Discussão

Esta seção está dividida em três subseções: Experimentos de Detecção do Melanoma (*Deep Learning*), Experimentos de Segmentação do Melanoma (*Fine-tuning*), seguida pela subseção de Comparação com o Estado da Arte.

5.1 Resultados e Discussão

O uso de métricas específicas para a avaliação quantitativa da segmentação em modelos de detecção e segmentação de melanoma é fundamental para a validade e confiabilidade dos resultados obtidos. Neste estudo, as métricas escolhidas incluem Acurácia, Índice Jaccard, Coeficiente Dice, Sensibilidade e Especificidade, todas amplamente reconhecidas e consagradas na literatura científica, conforme destacado por (TAHA; HANBURY, 2015).

A escolha dessas métricas não é aleatória, mas baseia-se em sua reputação e frequente uso em trabalhos renomados no campo da detecção e segmentação de melanoma. Essas métricas são consideradas padrões de ouro em muitos estudos devido à sua capacidade de fornecer uma avaliação abrangente e detalhada do desempenho dos modelos.

5.1.1 Resultados da Primeira Fase de Experimentos - Detecção

Nesta subseção, são apresentados os resultados dos experimentos para a detecção de melanoma em imagens de exames com o uso de dermatoscópio. Foram conduzidos experimentos com diferentes modelos computacionais do *framework* YOLOv8: YOLOv8l, YOLOv8m, YOLOv8n, YOLOv8s e YOLOv8x.

O treinamento foi realizado com diversas versões do YOLOv8, todas apresentando resultados relevantes e satisfatórios para a detecção da região de interesse (melanoma). A Tabela 1 apresenta os valores obtidos por diferentes métricas relacionadas aos modelos treinados com o conjunto de dados abordado.

Na Tabela 1, destaca-se a versão YOLOv8x, que obteve os melhores resultados para a detecção do melanoma. O modelo alcançou 98,67% de acurácia (ACC), com o menor desvio padrão entre os modelos treinados com a mesma base de dados. É importante ressaltar que o YOLOv8X apresentou, em todas as métricas, o menor desvio padrão entre os modelos treinados, indicando um bom desempenho e uniformidade nos resultados de detecção, proporcionando uma detecção mais precisa e consistente da região do melanoma.

A métrica de coeficiente Dice apresentou uma similaridade de 94,77% entre as imagens localizadas na região do melanoma e a região marcada pelos especialistas. A

sobreposição entre o resultado da detecção pelo modelo YOLOv8x e a região real do melanoma teve uma diferença inferior a 1,63% no pior caso (94.77 contra 93.14), o que é considerado um valor relevante em análises de imagens médicas. Cada *pixel* na região de interesse pode determinar a progressão ou regressão da doença, tornando essa precisão crucial. O modelo YOLOv8s obteve o melhor resultado na métrica de Sensibilidade (SEN), com 98,63%, demonstrando sua capacidade de identificar com precisão a proporção real de *pixels* pertencentes à região do câncer de pele na imagem.

Tabela 1 – Resultados métricos da detecção do melanoma em imagens dermatoscópicas para as diferentes versões do YOLOv8 abordadas neste estudo.

Modelos	ACC(%)	DICE(%)	SEN(%)	SPE(%)	PPV(%)	NPV(%)	MCC(%)	JAC(%)
YOLOv8l	98.44±1.06	94.03±3.54	96.11±3.62	98.65±1.69	92.54±7.36	99.43±0.55	93.25±3.71	88.94±6.19
YOLOv8m	98.38±1.50	93.97±3.89	95.03±4.84	98.68±2.17	93.63±7.99	99.32±0.71	93.21±4.24	88.87±6.70
YOLOv8n	98.54±1.13	94.61±3.21	97.46±2.61	98.49±1.89	92.32±6.72	99.67±0.34	93.87±3.35	89.95±5.54
YOLOv8s	98.08±1.62	93.14±4.80	98.63±1.29	97.77±2.46	88.73±8.73	99.81±0.17	92.28±4.96	87.51±7.99
YOLOv8x	98.67±0.77	94.77±2.99	95.55±4.09	99.02±1.14	94.24±4.76	99.34±0.62	94.02±3.20	90.20±5.17

É relevante destacar que o processo de detecção automática proporcionou ganhos significativos no tempo de resposta dos modelos. Todos os modelos alcançaram uma taxa de detecção inferior a 1%, tornando-os atraentes para diversas aplicações. Essa eficiência na detecção sem intervenção humana representa uma evolução significativa, garantindo precisão na localização da patologia.

O YOLOv8x demonstrou um desempenho superior em relação às outras versões do experimento. Apresentou resultados notáveis em métricas como SPE, que superou 99%, e PPV, com uma métrica de 94,24%, seguida por MCC e JAC. Esses resultados consolidam o YOLOv8x como o modelo mais eficiente na detecção do melanoma, destacando-se em comparação com diferentes abordagens presentes na literatura, tanto locais quanto automáticas.

Tabela 2 – Comparação de Tempo de detecção para cada um dos modelos.

Modelos	Tempo(s)
YOLOv8l	0.85 ± 1.90
YOLOv8m	0.44 ± 0.72
YOLOv8n	0.24 ± 0.63
YOLOv8s	0.30 ± 0.68
YOLOv8x	0.74 ± 0.69

A Figura 15 exibe os resultados dos modelos YOLOv8 na detecção da região cancerígena processados pelo método proposto. Nesta representação visual, são apresentados diversos casos de melanoma, cada um acompanhado da porcentagem de detecção associada à Acurácia (ACC) obtida para a respectiva imagem. Essa visualização oferece uma análise comparativa dos resultados de detecção para diferentes casos de melanoma processados pelo método proposto.

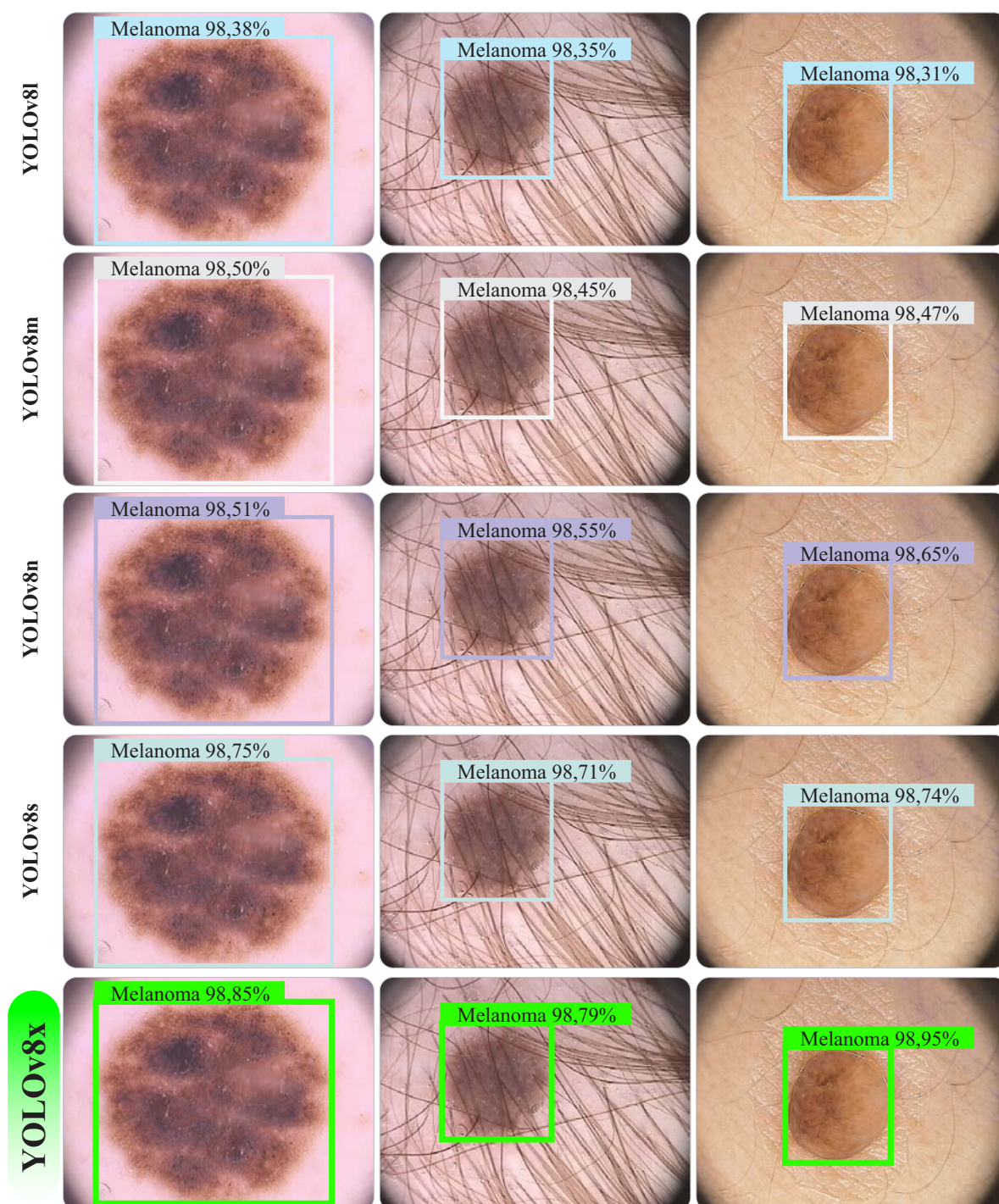


Figura 15 – Resultado da detecção de diferentes melanomas em imagens dermatoscópicas com as diversas versões do YOLOv8, com destaque para a YOLOv8x, na qual apresentou melhores resultados para detecção.

Ainda assim, é possível observar visualmente que o modelo YOLOv8x foi capaz de delinear com precisão a região de interesse por meio da caixa delimitadora. Este feito é desafiador, pois exige extrema precisão ao definir o escopo da área em que o crescimento da borda do câncer pode ser estimado. Dessa maneira, pode-se concluir que a detecção eficaz oferece um suporte valioso à camada de segmentação por meio do processo de fine-tuning.

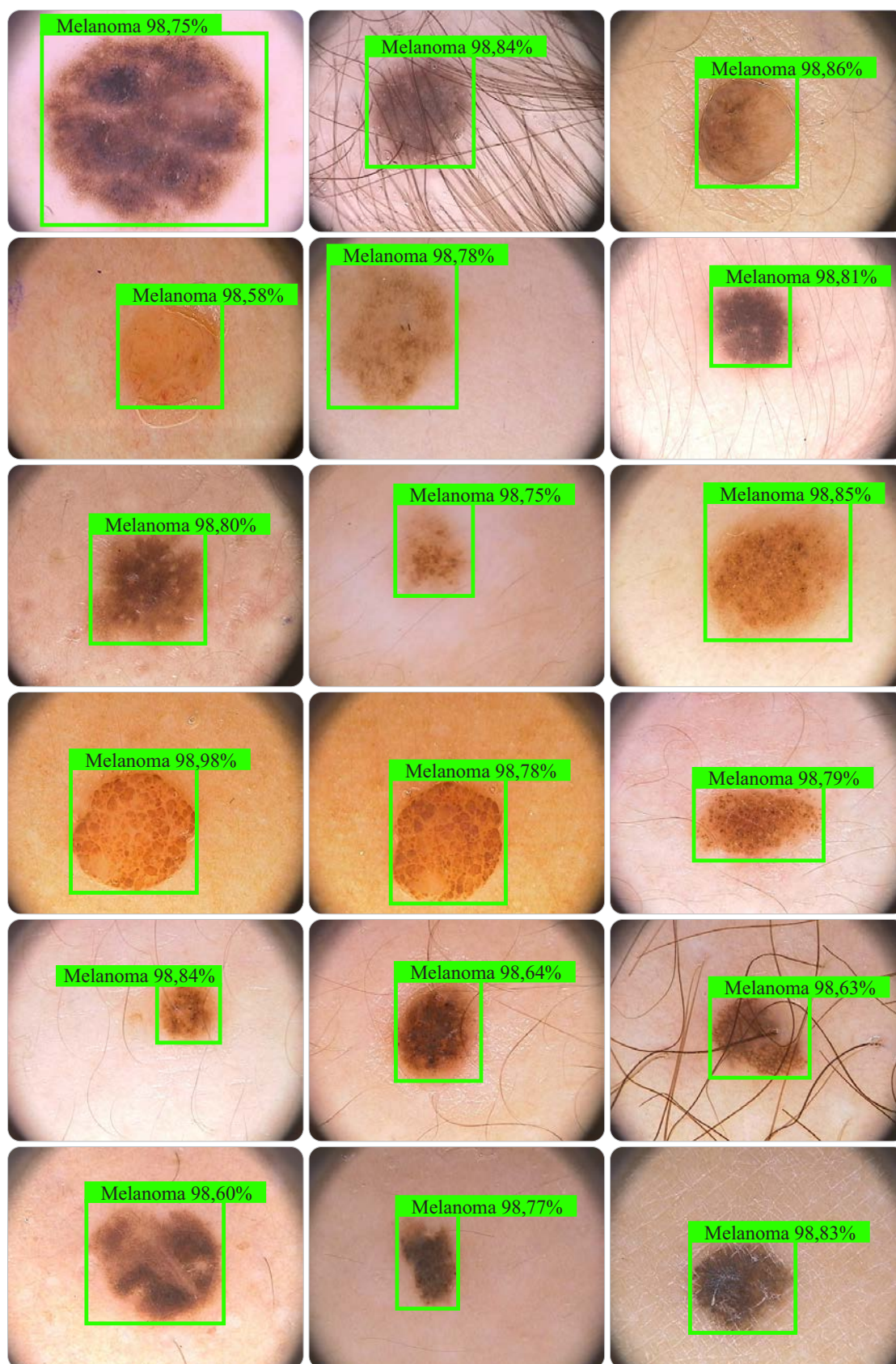


Figura 16 – Detecções por caixa delimitadora para o modelo YOLOv8x.

A Figura 16 ilustra resultados de alguns casos de detecção por *bounding box* para o YOLOv8x, realizados pelo estudo proposto.

5.1.2 Resultados da Segunda Fase de Experimento - Segmentação (*Fine-tuning*)

Nesta seção, serão abordados os métodos de segmentação desenvolvidos no estudo proposto, utilizando a detecção com as diferentes versões do YOLO, seguida da segmentação realizada com *fine-tuning* baseado em métodos de visão computacional com processamento digital de imagens. A avaliação da performance e eficácia da segmentação para os modelos propostos neste estudo ocorre por meio da comparação qualitativa e quantitativa das métricas ACC, DICE, SEN, SPE, PPV, NPV, MCC e JAC.

A Tabela 3 apresenta os resultados métricos de ACC, DICE, SEN e SPE para os modelos de segmentação desenvolvidos neste estudo. Ao todo, foram desenvolvidos 15 modelos de detecção e segmentação automática de melanoma em imagens dermatoscópicas. Cada modelo é constituído pela combinação de uma versão do YOLO para a detecção e um modelo *fine-tuning* existente. A Tabela 4 ilustra as médias e desvios padrão de cada modelo desenvolvido neste estudo para as métricas de PPV, NPV, MCC e JAC.

Dentre os modelos desenvolvidos para a detecção e segmentação de melanoma em imagens dermatoscópicas, a melhor combinação de detecção e segmentação foi a YOLOv8x com o *fine-tuning* baseado em Janelamento de Parzen com Crescimento de Região. Conforme observado na Tabela 3, o modelo obteve uma acurácia média de 98.68%, indicando uma boa proporção de *pixels* corretamente segmentados em relação ao total de *pixels* na imagem, uma característica valorizada para modelos CAD de segmentação.

O coeficiente Dice de 94.66% indica uma alta concordância entre a segmentação realizada pelo modelo e o padrão ouro para o dataset abordado. No entanto, o desvio padrão de 3.15% indica a existência de alguns casos nos quais o modelo teve sua eficácia reduzida, mas ainda dentro do padrão de variação dos demais modelos para esta métrica.

A especificidade de 99.16% indica uma boa capacidade de identificação dos *pixels* verdadeiros negativos. Por fim, a Sensitividade do melhor modelo proposto não foi a mais alta em termos de média, registrando um valor de 94.71%, o que é 4.88% menor que a maior Sensitividade observada na tabela (94.71 contra 99.59). Essa pequena margem de diferença sugere uma capacidade satisfatória do modelo na detecção de *pixels* verdadeiros positivos na imagem.

Por fim, o gráfico 17 ilustra a comparação entre as métricas para os modelos de segmentação desenvolvidos neste estudo. É importante notar a escala menor para a métrica de Dice em todos os modelos, indicando uma média relativamente menor em comparação com as demais métricas.

Pela Tabela 4, observa-se que o melhor modelo proposto possui o maior PPV da tabela, com uma média de 94.90%, indicando uma boa proporção de *pixels* segmentados corretamente nas imagens dermatoscópicas aplicadas. Quanto à métrica NPV, o melhor

Tabela 3 – Resultados de métricas para segmentação nos modelos desenvolvidos neste estudo. Os modelos são compostos pela combinação de uma versão do YOLOv8 e com *Fine Tuning* que segmentam a região das bordas do melanoma. A tabela em questão cobre as métricas de ACC, DICE, SEN e SPE.

	<i>Fine-Tuning</i>	ACC(%)	DICE(%)	SEN(%)	SPE(%)
YOLOv8l	Parzen	98.32±1.22	93.60±4.12	97.16±3.01	98.34±1.87
	Parzen_Clustering	98.20±1.37	93.26±4.66	97.70±2.64	98.13±1.99
	Parzen_Region_G	98.44±1.07	94.05±3.61	96.24±3.58	98.63±1.70
YOLOv8m	Parzen	98.31±1.67	93.92±4.30	96.47±3.91	98.40±2.37
	Parzen_Clustering	98.20±1.79	93.61±4.80	97.30±3.31	98.16±2.48
	Parzen_Region_G	98.38±1.51	94.01±3.99	95.29±4.79	98.66±2.17
YOLOv8n	Parzen	98.36±1.29	94.05±3.79	98.38±1.91	98.15±2.06
	Parzen_Clustering	98.20±1.43	93.61±4.02	98.95±1.42	97.87±2.23
	Parzen_Region_G	98.54±1.15	94.63±3.31	97.64±2.46	98.47±1.90
YOLOv8s	Parzen	97.83±1.77	92.32±5.31	99.26±0.77	97.38±2.63
	Parzen_Clustering	97.61±1.87	91.63±5.43	99.59±0.48	97.08±2.76
	Parzen_Region_G	98.08±1.63	93.14±4.83	98.73±1.26	97.76±2.46
YOLOv8x	Parzen	98.56±0.85	94.38±3.25	96.66±3.69	98.72±1.28
	Parzen_Clustering	98.56±0.85	94.42±3.22	96.67±3.70	98.72±1.28
	Parzen_Region_G	98.68±0.75	94.66±3.15	94.71±4.30	99.16±1.00

método proposto alcançou uma média de 99.22%, sendo ligeiramente menor (0.72%) que o maior NPV (99.22 contra 99.94), mas mantendo uma boa performance no que diz respeito à não segmentação dos *pixels* de fundo, ou seja, a pele não lesionada do paciente.

Em relação à métrica MCC, o método proposto obteve o maior resultado, com 93.94%, um valor satisfatório para esta métrica, pois ela indica uma medida global de desempenho da segmentação, considerando a proporção entre os *pixels* verdadeiros positivos, falsos positivos, verdadeiros negativos e falsos negativos. Quanto à métrica JAC, o melhor modelo também obteve o maior valor, com 90.02%. Essa métrica é especialmente relevante para medir a similaridade das bordas da segmentação realizada pelo modelo em comparação com o especialista médico.

A Figura 18 ilustra a comparação visual entre as métricas discutidas na Tabela 4. É perceptível que as escalas das métricas PPV e JAC são menores em comparação às demais, sendo consistente com as análises quantitativas realizadas.

Vale ressaltar que, ao obter as maiores métricas de ACC, DICE, SPE, PPV, MCC e JAC para o dataset abordado, a combinação do YOLOv8X com o *fine-tuning* de Janelamento de Parzem com Crescimento de Região mostrou-se o método mais performático em comparação aos demais para a segmentação de melanoma em imagens dermatoscópicas.

O YOLOv8X é a versão mais robusta do YOLOv8, apresentando uma arquitetura mais robusta, um tempo de treinamento mais longo e geralmente uma quantidade maior de épocas de treinamento. Esse modelo realiza a detecção por máscara binária com alta

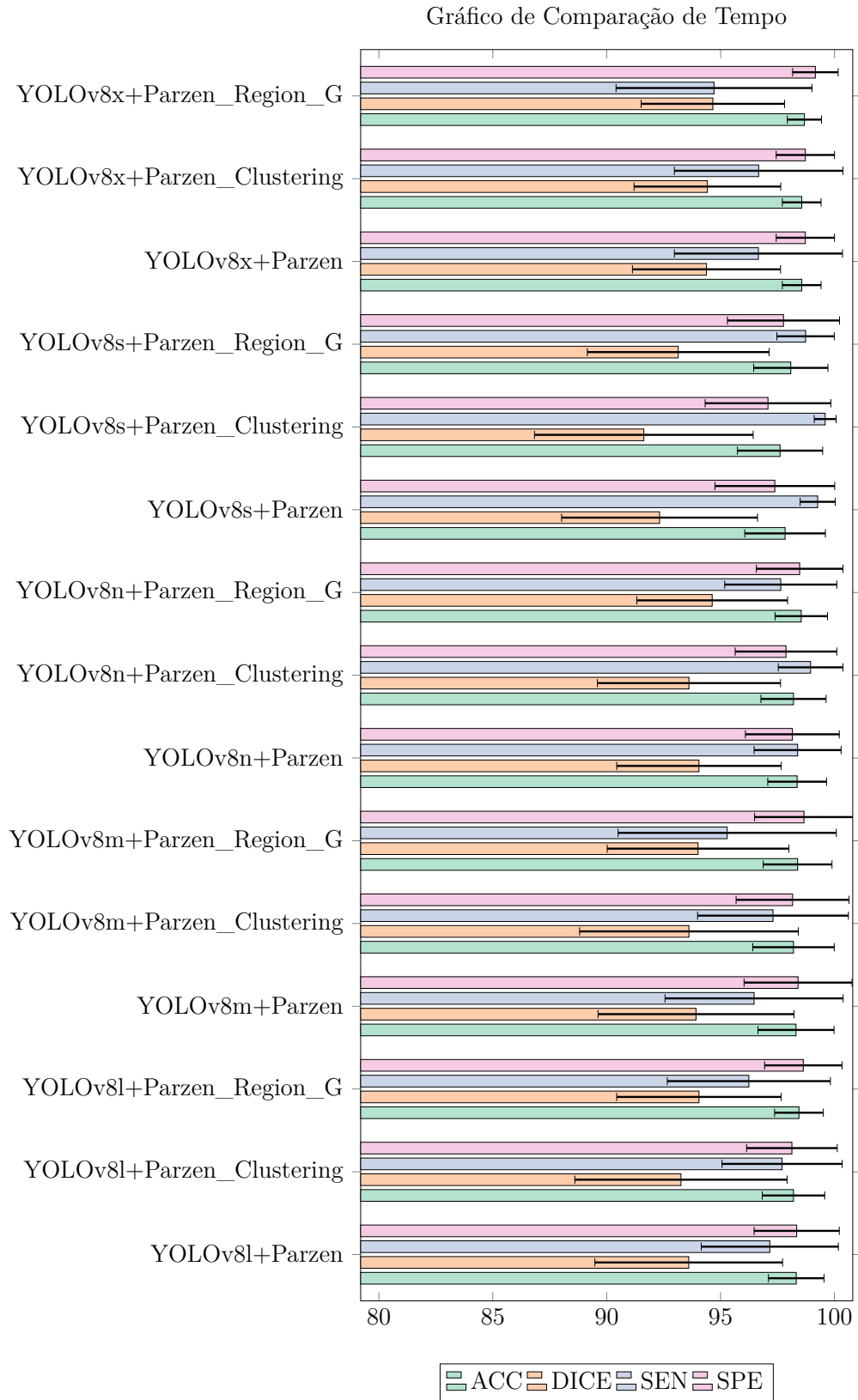


Figura 17 – Comparação gráfica da Tabela 3. Gráfico de ilustração da Comparação das métricas de segmentação realizadas pelo modelos *Fine-tuning* tabela.

resolução, exibindo bordas bastante semelhantes ao padrão-ouro e uma taxa de acerto de detecção elevada. Ao combinar esse alto desempenho de detecção com uma boa máscara

Tabela 4 – Resultados de métricas para segmentação nos modelos desenvolvidos neste estudo. Os modelos são compostos pela combinação de uma versão do YOLOv8 e com *Fine Tuning*. A tabela em questão abrange as métricas de PPV, NPV, MCC e JAC.

	Fine-Tuning	PPV(%)	NPV(%)	MCC(%)	JAC(%)
YOLOv8l	Parzen	90.81±8.02	99.57±0.45	92.80±4.27	88.25±7.06
	Parzen_Clustering	89.75±8.69	99.66±0.37	92.45±4.81	87.71±7.88
	Parzen_Region_G	92.46±7.44	99.44±0.55	93.28±3.79	88.98±6.31
YOLOv8m	Parzen	92.19±8.62	99.50±0.59	93.14±4.66	88.84±7.32
	Parzen_Clustering	90.89±9.29	99.61±0.50	92.82±5.13	88.35±8.06
	Parzen_Region_G	93.48±8.13	99.34±0.71	93.26±4.34	88.96±6.89
YOLOv8n	Parzen	90.49±7.32	99.79±0.24	93.26±3.91	89.00±6.42
	Parzen_Clustering	89.19±7.43	99.86±0.18	92.77±4.18	88.24±6.77
	Parzen_Region_G	92.18±6.84	99.68±0.33	93.88±3.46	89.98±5.72
YOLOv8s	Parzen	86.80±9.09	99.90±0.10	91.42±5.42	86.17±8.67
	Parzen_Clustering	85.34±9.05	99.94±0.07	90.68±5.53	85.00±8.80
	Parzen_Region_G	88.64±8.73	99.82±0.17	92.28±5.00	87.52±8.04
YOLOv8x	Parzen	92.47±5.46	99.49±0.54	93.60±3.45	89.53±5.58
	Parzen_Clustering	92.51±5.37	99.49±0.55	93.64±3.43	89.59±5.54
	Parzen_Region_G	94.90±5.19	99.22±0.67	93.94±3.35	90.02±5.45

binária de entrada, o *fine-tuning* baseado em Janelamento de Parzen com Crescimento de Região, quando devidamente parametrizado para a aplicação, é capaz de aprimorar as bordas da máscara binária de maneira eficaz. O crescimento de região utiliza uma regra de adesão com uma variação de nível de cinza de mais ou menos 10, expandindo a máscara de forma sutil e criteriosa. Por outro lado, o janelamento de Parzen foi configurado para realizar ajustes uniformes nas bordas com poucas iterações e baixa energia. Dessa forma, o *fine-tuning* tinha como objetivo melhorar, a nível de *pixel*, a qualidade das bordas da segmentação para torná-las mais semelhantes ao padrão-ouro.

A Figura 19 ilustra o resultado da segmentação para diferentes imagens dermatoscópicas do dataset abordado.

$$Equivalence \begin{cases} \oplus = Yes \\ x = No \end{cases} \quad (5.1)$$

A Tabela 5 apresenta os resultados do teste estatístico de Friedman realizado entre o melhor modelo, YOLOv8x + Parzen_Region_G (YOLOv8x com Janelamento de Parzen e Crescimento de Região), e os demais métodos encontrados no estado da arte. A equação 5.1 fornece o significado dos símbolos na tabela de teste estatístico, onde " $\oplus$ " indica que as médias e desvios daquela métrica para aquele modelo foram construídos sob vetores de métricas estatisticamente equivalentes aos do melhor modelo proposto. Por outro lado, "x" indica o oposto.

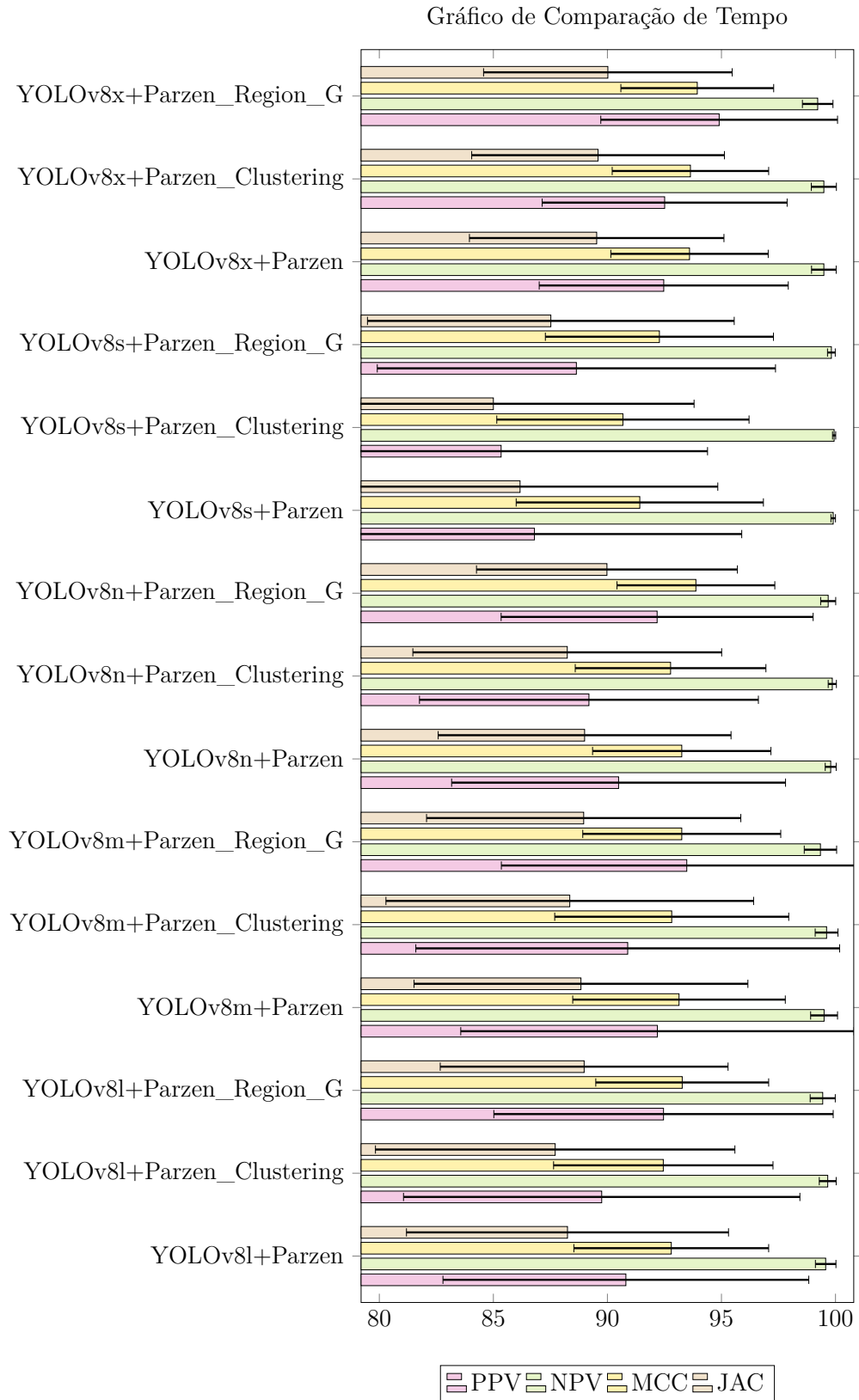


Figura 18 – Comparação gráfica da Tabela 3. Gráfico de ilustração da Comparação das métricas de segmentação realizadas pelo modelos *Fine-tuning* tabela.

O teste indica que praticamente todos os vetores de ACC, DICE, MCC e JAC são estatisticamente equivalentes aos vetores do melhor modelo proposto, sugerindo uma

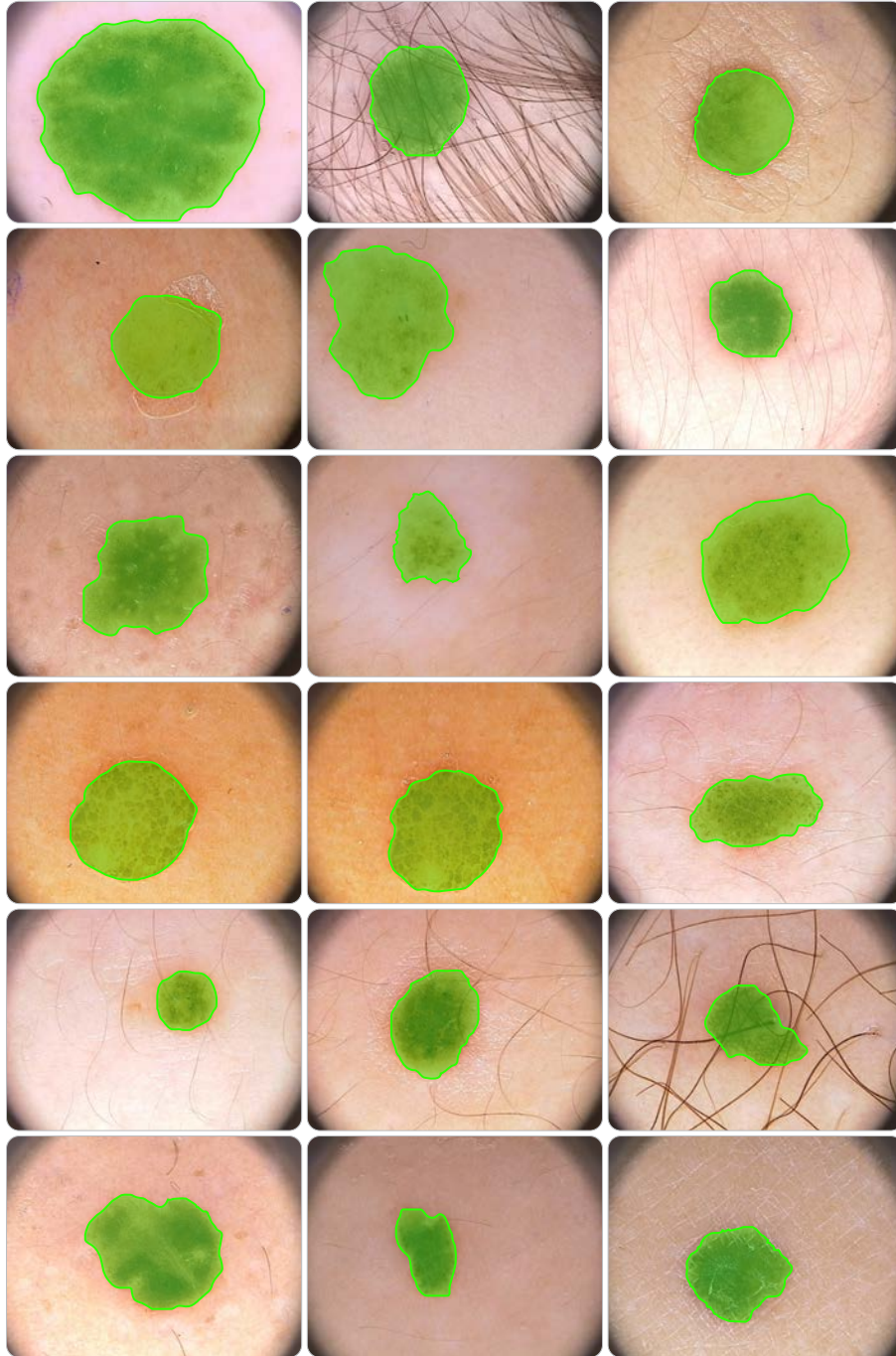


Figura 19 – Resultado de Segmentação de casos distintos para o melhor modelo proposto, YOLOv8x + Parzen_Region_G.

relativa semelhança nas métricas, com os *outliers* explicando as variações. Por outro lado, os vetores métricos de SEN, SPE, PPC e NPV, em geral, são estatisticamente diferentes dos vetores métricos do melhor modelo. Isso indica que o melhor modelo proposto foi, em geral, distinto na maioria dos casos em relação a essas métricas quando comparado aos outros modelos.

Por fim, ao observar a tabela 6 para a comparação de tempo, o modelo de segmentação mais rápido foi o YOLOv8n com o *fine-tuning* baseado em Janelamento de Parzem,

Tabela 5 – Representação dos resultados do teste estatístico de Friedman realizado entre os diferentes resultados de segmentação por *Fine-tuning*, *YOLOv8x + Parzen_Region_G*.

Detecção	Fine-Tuning	ACC	DICE	SEN	SPE	PPV	NPV	MCC	JAC
YOLOv8l	Parzen	⊕	⊕	x	x	x	x	⊕	⊕
	Parzen_Clustering	⊕	⊕	x	x	x	x	⊕	⊕
	Parzen_Region_G	⊕	⊕	x	x	x	x	⊕	⊕
YOLOv8m	Parzen	⊕	⊕	x	x	x	x	⊕	⊕
	Parzen_Clustering	⊕	⊕	x	x	x	x	⊕	⊕
	Parzen_Region_G	⊕	⊕	⊕	⊕	⊕	⊕	⊕	⊕
YOLOv8n	Parzen	⊕	⊕	x	x	x	x	⊕	⊕
	Parzen_Clustering	⊕	⊕	x	x	x	x	⊕	⊕
	Parzen_Region_G	⊕	⊕	x	x	x	x	⊕	⊕
YOLOv8s	Parzen	⊕	⊕	x	x	x	x	⊕	⊕
	Parzen_Clustering	x	⊕	x	x	x	x	⊕	⊕
	Parzen_Region_G	⊕	⊕	x	x	x	x	⊕	⊕
YOLOv8x	Parzen	⊕	⊕	x	x	x	x	⊕	⊕
	Parzen_Clustering	⊕	⊕	x	x	x	x	⊕	⊕

Tabela 6 – Comparação de tempo de segmentação entre os métodos de *Fine-Tuning* proposto por este estudo.

Detecção	Fine-Tuning	Time(s)
YOLOv8l	Parzen	6.13 ± 1.95
	Parzen_Clustering	6.96 ± 2.08
	Parzen_Region_G	6.34 ± 2.07
YOLOv8m	Parzen	5.78 ± 0.88
	Parzen_Clustering	6.29 ± 0.83
	Parzen_Region_G	5.94 ± 0.92
YOLOv8n	Parzen	5.59 ± 0.79
	Parzen_Clustering	6.10 ± 0.77
	Parzen_Region_G	5.70 ± 0.84
YOLOv8s	Parzen	5.64 ± 0.82
	Parzen_Clustering	6.15 ± 0.82
	Parzen_Region_G	5.75 ± 0.89
YOLOv8x	Parzen	6.07 ± 0.82
	Parzen_Clustering	6.31 ± 0.96
	Parzen_Region_G	6.23 ± 0.96

com 5,59 segundos em média, sendo mais rápido em 0,64 segundos que o melhor modelo em termos de desempenho, o YOLOv8x + Parzen_Region_Growing, que levou em média 6,23 segundos. A YOLOv8n é uma versão nano da YOLOv8, com uma arquitetura menor e focada na rapidez, perdendo em processamento bruto quando comparada às suas versões mais robustas, enquanto o janelamento de Parzen, sem crescimento de região, mostrou-se eficiente no ajuste de bordas. A Figura 20 ilustra graficamente a comparação de tempo da

tabela.

A Figura 21 ilustra a comparação gráfica da segmentação para os modelos propostos, é perceptível que o modelo que se utiliza da YOLOv8x possui as bordas de segmentação mais expressivas e fidedignas com o Padrão Ouro.

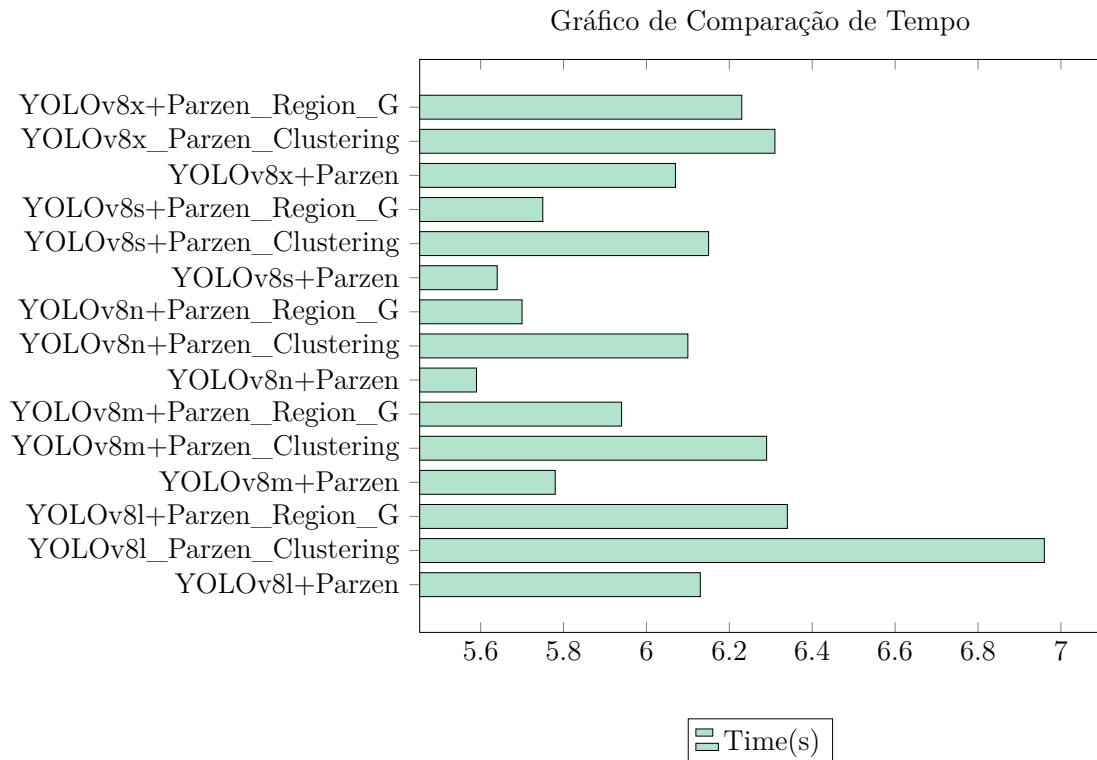


Figura 20 – Gráfico de comparação do tempo de segmentação para diferentes combinações de detecção e ajuste fino (*fine-tuning*).

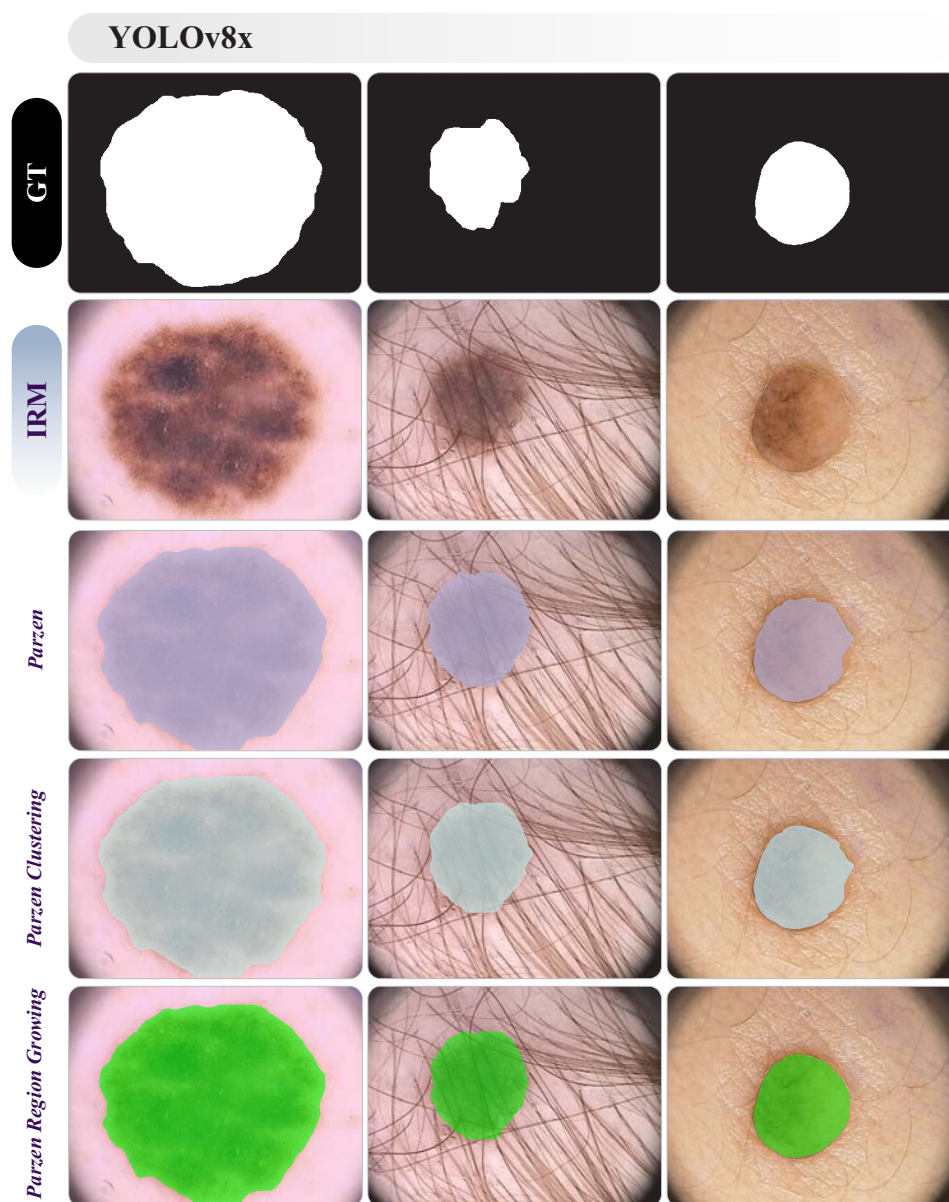


Figura 21 – Comparação entre os resultados de segmentação entre os diferentes modelos propostos por este estudo YOLOv8x.

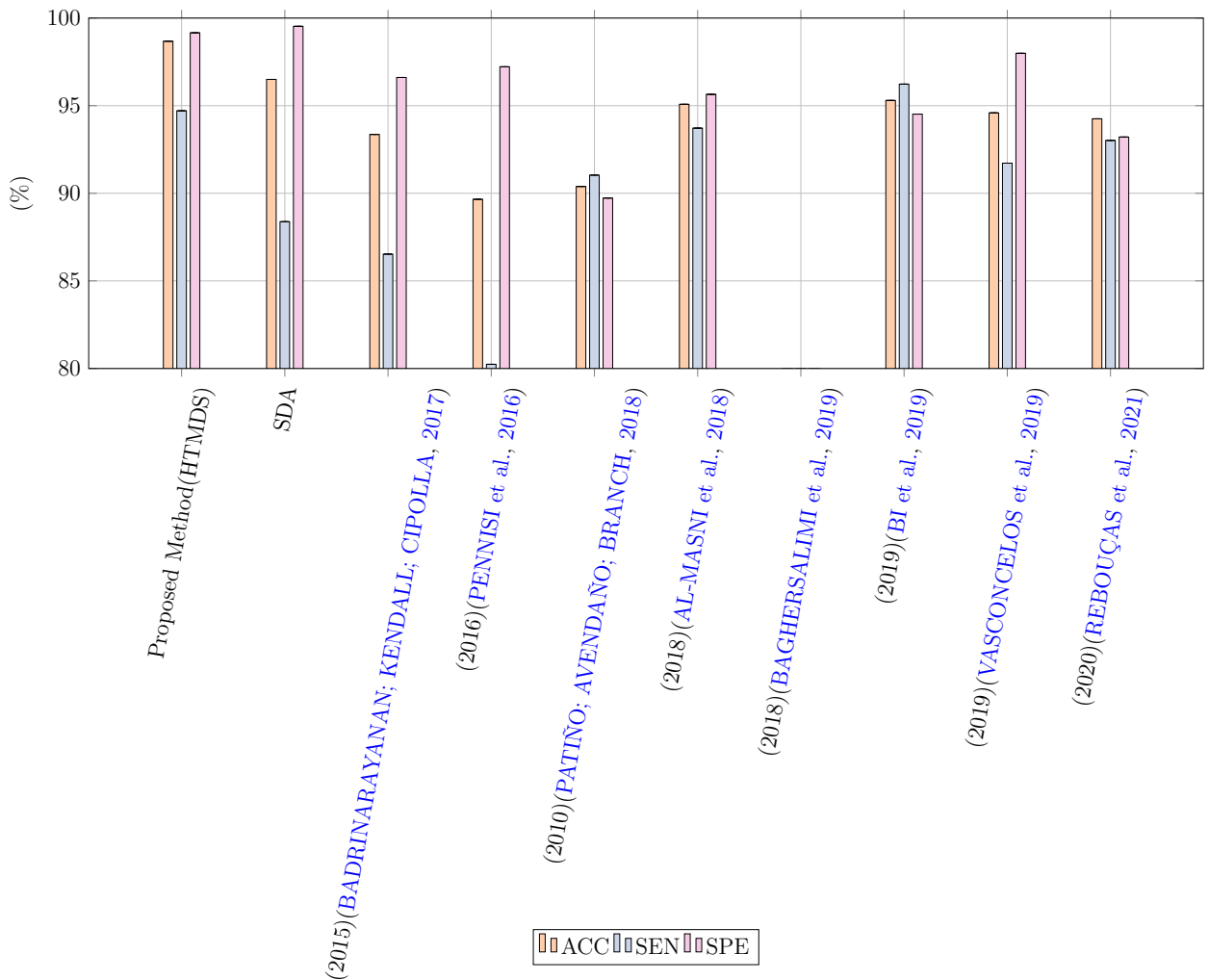
5.1.3 Resultados do Terceiro Experimento - Comparação com o Estado da Arte

Esta subseção apresenta os resultados de métodos renomados da literatura para detecção e segmentação de melanoma em imagens dermatoscópicas. A tabela 7 destaca os diferentes métodos do estado da arte para a base de dados abordada.

Tabela 7 – Comparação do melhor modelo proposto por este estudo (*YOLOv8x + Parzen_Region_Growth*), comparados a trabalhos encontrados no estado da arte.

Métodos	ACC(%)	JAC(%)	DICE(%)	SEN(%)	SPE(%)
Proposed Method (HTMDS) - Detection and segmentation of melanoma	98.68	90.02	94.66	94.71	99.16
SDA-Detection Melanoma (2023) (NASCIMENTO et al., 2023)	96.50	87.22	92.60	88.38	99.53
Deep convolutional network pixel wise segmentation (2015) (BADRINARAYANAN; KENDALL; CIPOLLA, 2017)	93.36	80.77	89.36	86.53	96.61
Delaunay triangulation (2016) (PENNISI et al., 2016)	89.66	-	-	80.24	97.22
Clustering-based (2018) (PATIÑO; AVENDAÑO; BRANCH, 2018)	90.39	-	-	91.04	89.73
Full resolution convolution networks (2018) (AL-MASNI et al., 2018)	95.08	84.79	91.77	93.72	95.65
FCN encoder-decoder (2018) (BAGHERSALIMI et al., 2019)	-	85.30	91.50	-	-
Deep class-specific learning features segmentation (2019) (BI et al., 2019)	95.30	85.90	92.10	96.23	94.52
Geodesic active contour (2019) (VASCONCELOS et al., 2019)	94.59	86.16	92.17	91.72	97.99
FLog Parzen level set (2020) (REBOUÇAS et al., 2021)	94.25	85.09	92.49	93.02	93.21

Figura 22 – Gráfico obtido a partir da tabela de comparação com o estado da arte.



O método desenvolvido neste estudo demonstrou resultados superiores em diversas métricas em comparação com uma variedade de abordagens encontradas na literatura. O

modelo proposto obteve um aumento significativo de mais de 4% em relação aos métodos com menor desempenho, alcançando uma acurácia de 98.68% em comparação com 94.25% de (REBOUÇAS et al., 2021). Mesmo considerando a diversidade de modelos utilizados na comparação, incluindo métodos semi-automáticos e detecção manual, a proposta de Detecção e Segmentação de Melanoma apresentou uma vantagem expressiva de mais de 9% em comparação com modelos reconhecidos, como evidenciado no estudo de (PENNISI et al., 2016), que registrou uma acurácia de 89.66%.

O modelo proposto superou métodos equivalentes com uso de *deep learning*, como os apresentados por (BADRINARAYANAN; KENDALL; CIPOLLA, 2017), que abordaram a mesma problemática, assim como o estudo de (REBOUÇAS et al., 2021). Alguns trabalhos, como o de (AL-MASNI et al., 2018), não forneceram métricas comparáveis, dificultando uma validação mais precisa da segmentação do melanoma.

Considerando os experimentos realizados e a comparação com a mesma base de dados utilizada em diferentes estudos, conclui-se que o método proposto se destacou em várias métricas, como ACC, JAC e DICE, tornando o modelo robusto para detecção precisa e segmentação totalmente automática do câncer de pele. O gráfico representado pela Figura 22 ilustra visualmente os diferentes valores obtidos por meio das métricas compartilhadas pelos diversos métodos do estado da arte, permitindo uma análise eficaz por meio das barras verticais. Repare que o trabalho de (BAGHERSALIMI et al., 2019) não tem dados no gráfico, devido ao seu trabalho não se utilizar das métricas de Acurácia, Sensitividade e Especificidade.

6 Conclusão

A pesquisa proposta fundamenta-se em uma nova abordagem, empregando *deep learning* e *fine-tuning* para a detecção e segmentação de melanoma em imagens dermatoscópicas. A metodologia adotada possibilitou a identificação e segmentação precisa da região afetada pelo melanoma, utilizando *deep learning* com o framework YOLOv8 para a detecção do câncer de pele. O *fine-tuning*, combinado a diversas técnicas, como crescimento de região, janelamento de Parzen e clusterização, desempenhou o papel de camada de refinamento, aprimorando a precisão do processo na segmentação da região de interesse.

Os resultados obtidos no estudo demonstram uma detecção e segmentação automáticas de melanoma utilizando *deep learning*, superando os padrões do estado da arte. O método YOLOv8x obteve valores métricas acima dos métodos da literatura, atingindo 98.67% de acurácia na detecção eficaz do melanoma em imagens dermatoscópicas. Posteriormente, o *fine-tuning*, especialmente com o modelo YOLOv8x *Parzen_Region_Growth*, elevou a acurácia de segmentação para 98.68%, superando diversas abordagens encontradas no estado da arte. Além disso, observou-se uma melhoria no desempenho temporal, com um tempo de detecção e segmentação de 5.59 segundos no cenário mais otimizado, destacando a eficiência do modelo em comparação com métodos manuais, semi-automáticos e automáticos encontrados na literatura.

Conclui-se que a abordagem proposta revelou-se eficaz tanto para a detecção quanto para a segmentação em imagens dermatoscópicas de melanoma, demonstrando robustez e precisão. Destaca-se sua aplicabilidade como ferramenta de auxílio ao diagnóstico médico, promovendo resultados satisfatórios em comparação com abordagens convencionais e inovadoras.

Quanto a trabalhos futuros, sugere-se explorar a classificação de diferentes tipos de melanomas, demandando aprofundamento e bases de dados rotuladas específicas. A aplicação da metodologia proposta em variados conjuntos de dados médicos, incluindo câncer de pele e hematomas na epiderme, representa uma extensão promissora. Ademais, a adaptação do modelo para treinamento em diferentes imagens médicas, como detecção de tumores cerebrais em ressonância magnética e tomografias computadorizadas, abre perspectivas para uma abordagem mais abrangente e versátil na área de diagnóstico médico.

Outra abordagem especialmente interessante é a aplicação do método desenvolvido em sistemas embarcados ou em nuvem, de maneira a dinamizar o fluxo da aplicação, sendo possível ao médico especialista receber o um pré-diagnóstico completo de detecção das bordas da lesão com um baixo tempo de resposta.

Referências

ACETO, G.; PERSICO, V.; PESCAPÉ, A. Industry 4.0 and health: Internet of things, big data, and cloud computing for healthcare 4.0. *Journal of Industrial Information Integration*, Elsevier, v. 18, p. 100129, 2020. Citado na página 20.

AL-MASNI, M. A. et al. Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks. *Computer methods and programs in biomedicine*, Elsevier, v. 162, p. 221–231, 2018. Citado 3 vezes nas páginas 28, 75 e 76.

ALJABAR, P. et al. Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. *Neuroimage*, Elsevier, v. 46, n. 3, p. 726–738, 2009. Citado na página 51.

ANANTHARAMAN, R.; VELAZQUEZ, M.; LEE, Y. Utilizing mask r-cnn for detection and segmentation of oral diseases. In: IEEE. *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. Madrid, Spain, 2018. p. 2197–2204. Citado na página 28.

ARNOLD, M. et al. Global burden of cutaneous melanoma in 2020 and projections to 2040. *JAMA dermatology*, American Medical Association, v. 158, n. 5, p. 495–503, 2022. Citado na página 18.

ASLAM, M. Introducing kolmogorov–smirnov tests under uncertainty: an application to radioactive data. *ACS omega*, ACS Publications, v. 5, n. 1, p. 914–917, 2019. Citado na página 49.

BADRINARAYANAN, V.; KENDALL, A.; CIPOLLA, R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, IEEE, v. 39, n. 12, p. 2481–2495, 2017. Citado 3 vezes nas páginas 29, 75 e 76.

BAGHERSALIMI, S. et al. Dermonet: densely linked convolutional neural network for efficient skin lesion segmentation. *EURASIP Journal on Image and Video Processing*, SpringerOpen, v. 2019, n. 1, p. 1–10, 2019. Citado 3 vezes nas páginas 28, 75 e 76.

BAUMGARTNER, D.; KOLASSA, J. Power considerations for kolmogorov–smirnov and anderson–darling two-sample tests. *Communications in Statistics-Simulation and Computation*, Taylor & Francis, v. 52, n. 7, p. 3137–3145, 2023. Citado na página 49.

BHADAURIA, H.; DEWAL, M. Intracranial hemorrhage detection using spatial fuzzy c-mean and region-based active contour on brain ct imaging. *Signal, Image and Video Processing*, Springer, v. 8, n. 2, p. 357–364, 2014. Citado 3 vezes nas páginas 51, 52 e 53.

BI, L. et al. Step-wise integration of deep class-specific learning for dermoscopic image segmentation. *Pattern recognition*, Elsevier, v. 85, p. 78–89, 2019. Citado 2 vezes nas páginas 29 e 75.

BUDA, M.; SAHA, A.; MAZUROWSKI, M. A. Association of genomic subtypes of lower-grade gliomas with shape features automatically extracted by a deep learning

- algorithm. *Computers in biology and medicine*, Elsevier, v. 109, p. 218–225, 2019. Citado na página 27.
- CELEBI, M. E.; CODELLA, N.; HALPERN, A. Dermoscopy image analysis: overview and future directions. *IEEE journal of biomedical and health informatics*, IEEE, v. 23, n. 2, p. 474–478, 2019. Citado na página 19.
- COMER, M. L.; III, E. J. D. Morphological operations for color image processing. *Journal of electronic imaging*, SPIE, v. 8, n. 3, p. 279–289, 1999. Citado na página 42.
- COMER, M. L.; III, E. J. D. Morphological operations for color image processing. *Journal of Electronic Imaging*, SPIE, v. 8, n. 3, p. 279 – 289, 1999. Disponível em: <<https://doi.org/10.1117/1.482677>>. Citado na página 42.
- CONTRIBUTORS, M. *YOLOv8 by MMYOLO*. 2023. <<https://github.com/open-mmlab/mmyolo/tree/main/configs/yolov8>>. Accessed on 13 May 2023. Citado 2 vezes nas páginas 9 e 39.
- DICE, L. R. Measures of the amount of ecologic association between species. *Ecology*, Wiley Online Library, v. 26, n. 3, p. 297–302, 1945. Citado na página 51.
- GAJERA, H. K.; NAYAK, D. R.; ZAVERI, M. A. A comprehensive analysis of dermoscopy images for melanoma detection via deep cnn features. *Biomedical Signal Processing and Control*, Elsevier, v. 79, p. 104186, 2023. Citado na página 19.
- GAUR, L. et al. Medical image-based detection of covid-19 using deep convolution neural networks. *Multimedia systems*, Springer, p. 1–10, 2021. Citado na página 27.
- GORDON, L. et al. Prevention versus early detection for long-term control of melanoma and keratinocyte carcinomas: a cost-effectiveness modelling study. *BMJ open*, British Medical Journal Publishing Group, v. 10, n. 2, p. e034388, 2020. Citado na página 18.
- HAAN, B. de et al. Fast semi-automated lesion demarcation in stroke. *NeuroImage: Clinical*, Elsevier, v. 9, p. 69–74, 2015. Citado na página 27.
- HAN, C. et al. Learning more with less: Conditional ppgan-based data augmentation for brain metastases detection using highly-rough annotation on mr images. In: *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*. New York, NY, USA: Association for Computing Machinery, 2019. (CIKM '19), p. 119–127. ISBN 9781450369763. Disponível em: <<https://doi.org/10.1145/3357384.3357890>>. Citado na página 27.
- HAN, T. et al. Internet of medical things—based on deep learning techniques for segmentation of lung and stroke regions in ct scans. *IEEE Access*, IEEE, v. 8, p. 71117–71135, 2020. Citado na página 28.
- HANBURY, A. G.; SERRA, J. Morphological operators on the unit circle. *IEEE transactions on image processing*, IEEE, v. 10, n. 12, p. 1842–1850, 2001. Citado na página 42.
- HAYKIN, S. *Redes neurais: princípios e prática*. São Paulo, Higienópolis: Bookman Editora, 2007. Citado na página 32.

- HE, K. et al. Mask r-cnn. In: *Proceedings of the IEEE international conference on computer vision*. [S.l.: s.n.], 2017. p. 2961–2969. Citado na página 44.
- HU, Q. et al. An effective approach for ct lung segmentation using mask region-based convolutional neural networks. *Artificial Intelligence in Medicine*, Elsevier, p. 101792, 2020. Citado na página 28.
- JOCHER, G. *Ultralytics YOLOv5*. 2020. Disponível em: <<https://github.com/ultralytics/yolov5>>. Citado na página 38.
- JOCHER, G.; CHAURASIA, A.; QIU, J. *Ultralytics YOLOv8*. 2023. Disponível em: <<https://github.com/ultralytics/ultralytics>>. Citado 2 vezes nas páginas 37 e 38.
- JOCHER, G.; CHAURASIA, A.; QIU, J. *YOLO by Ultralytics*. 2023. <<https://github.com/ultralytics/ultralytics>>. Accessed on 28 February 2023. Citado na página 38.
- JOCHER, G.; CHAURASIA, A.; QIU, J. *YOLO by Ultralytics*. 2023. <<https://github.com/ultralytics/ultralytics>>. (Accessed on 28 February 2023). Citado 2 vezes nas páginas 39 e 40.
- JONES, O. et al. Dermoscopy for melanoma detection and triage in primary care: a systematic review. *BMJ open*, British Medical Journal Publishing Group, v. 9, n. 8, p. e027529, 2019. Citado na página 19.
- KAMNITSAS, K. et al. Efficient multi-scale 3d cnn with fully connected crf for accurate brain lesion segmentation. *Medical image analysis*, Elsevier, v. 36, p. 61–78, 2017. Citado na página 28.
- KOCH, C.; TSUCHIYA, N. Attention and consciousness: two distinct brain processes. *Trends in cognitive sciences*, Elsevier, v. 11, n. 1, p. 16–22, 2007. Citado na página 32.
- KRENSEL, M.; SCHÄFER, I.; AUGUSTIN, M. Cost-of-illness of melanoma in europe—a systematic review of the published literature. *Journal of the European Academy of Dermatology and Venereology*, Wiley Online Library, v. 33, n. 3, p. 504–510, 2019. Citado na página 18.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *nature*, Nature Publishing Group UK London, v. 521, n. 7553, p. 436–444, 2015. Citado 2 vezes nas páginas 32 e 33.
- LI, X. et al. Generalized focal loss: Learning qualified and distributed bounding boxes for dense object detection. *Advances in Neural Information Processing Systems*, v. 33, p. 21002–21012, 2020. Citado na página 38.
- LIKAS, A.; VLASSIS, N.; VERBEEK, J. J. The global k-means clustering algorithm. *Pattern recognition*, Elsevier, v. 36, n. 2, p. 451–461, 2003. Citado na página 43.
- LONG, G. V. et al. Cutaneous melanoma. *The Lancet*, Elsevier, v. 402, n. 10400, p. 485–502, 2023. Citado na página 18.
- MARQUES, A. G. et al. Automatic segmentation of hemorrhagic stroke on brain ct images using convolutional neural networks through fine-tuning. In: IEEE. *2022 International Joint Conference on Neural Networks (IJCNN)*. Padua, Italy, 2022. p. 01–08. Citado na página 45.

- MOHRI, M.; ROSTAMIZADEH, A.; TALWALKAR, A. *Foundations of machine learning*. Massachusetts, Cambridge: MIT Press, 2018. Citado na página 32.
- NASCIMENTO, J. J. D. C. et al. New health of things approach to classification and detection of brain tumors using transfer learning for segmentation in imr images. In: IEEE. *2023 International Joint Conference on Neural Networks (IJCNN)*. Gold Coast, Australia, 2023. p. 1–8. Citado na página 75.
- OTSU, T.; TANIGUCHI, G. Kolmogorov–smirnov type test for generated variables. *Economics Letters*, Elsevier, v. 195, p. 109401, 2020. Citado na página 49.
- PATÍÑO, D.; AVENDAÑO, J.; BRANCH, J. W. Automatic skin lesion segmentation on dermoscopic images by the means of superpixel merging. In: SPRINGER. *International conference on medical image computing and computer-assisted intervention*. Granada, Spain, 2018. p. 728–736. Citado 2 vezes nas páginas 28 e 75.
- PENNISI, A. et al. Skin lesion image segmentation using delaunay triangulation for melanoma detection. *Computerized Medical Imaging and Graphics*, Elsevier, v. 52, p. 89–103, 2016. Citado 3 vezes nas páginas 29, 75 e 76.
- PHAM, D. T.; DIMOV, S. S.; NGUYEN, C. D. Selection of k in k-means clustering. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, SAGE Publications Sage UK: London, England, v. 219, n. 1, p. 103–119, 2005. Citado na página 42.
- PHAM, V.; PHAM, C.; DANG, T. Road damage detection and classification with detectron2 and faster r-cnn. In: IEEE. *2020 IEEE International Conference on Big Data (Big Data)*. [S.l.], 2020. p. 5592–5601. Citado na página 44.
- PROVOST, F.; KOHAVI, R. Glossary of terms. *Journal of Machine Learning*, v. 30, n. 2-3, p. 271–274, 1998. Citado na página 31.
- QIAN, W. et al. Mapping and candidate-gene screening of the novel turnip mosaic virus resistance gene retri02 in chinese cabbage (*brassica rapa* l.). *Theoretical and applied genetics*, Springer, v. 126, p. 179–188, 2013. Citado na página 51.
- RAJ, R. J. S. et al. Optimal feature selection-based medical image classification using deep learning model in internet of medical things. *IEEE Access*, IEEE, v. 8, p. 58006–58017, 2020. Citado na página 27.
- REBOUÇAS, E. de S. et al. Level set approach based on parzen window and floor of log for edge computing object segmentation in digital images. *Applied Soft Computing*, Elsevier, v. 105, p. 107273, 2021. Citado 4 vezes nas páginas 29, 43, 75 e 76.
- REDMON, J. et al. You only look once: Unified, real-time object detection. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. Las Vegas, Nevada: RSIP Vision, 2016. p. 779–788. Citado na página 37.
- RODRIGUES, Y. O. A. et al. Sda-detection melanoma: Deep approach system for detection and segmentation in melanoma images using fine-tuning. *Computerized Medical Imaging and Graphics*, 2023. Citado na página 28.

- ROSSI, M. et al. Familial melanoma: diagnostic and management implications. *Dermatology practical & conceptual*, Mattioli 1885, v. 9, n. 1, p. 10, 2019. Citado na página 19.
- SAMUEL, A. L. Some studies in machine learning using the game of checkers. ii—recent progress. *IBM Journal of research and development*, IBM, v. 11, n. 6, p. 601–617, 1967. Citado na página 31.
- SANCHEZ, S.; ROMERO, H.; MORALES, A. A review: Comparison of performance metrics of pretrained models for object detection using the tensorflow framework. In: IOP PUBLISHING. *IOP Conference Series: Materials Science and Engineering*. Bristol, England, 2020. v. 844, n. 1, p. 012024. Citado na página 36.
- SANDRU, A. et al. Survival rates of patients with metastatic malignant melanoma. *Journal of medicine and life*, Carol Davila-University Press, v. 7, n. 4, p. 572, 2014. Citado na página 18.
- TAHA, A. A.; HANBURY, A. Metrics for evaluating 3d medical image segmentation: analysis, selection, and tool. *BMC medical imaging*, BioMed Central, v. 15, n. 1, p. 1–28, 2015. Citado 4 vezes nas páginas 50, 52, 53 e 62.
- TERVEN, J.; Córdova-ESPARZA, D.-M.; ROMERO-GONZÁLEZ, J.-A. A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas. *Machine Learning and Knowledge Extraction*, v. 5, n. 4, p. 1680–1716, 2023. ISSN 2504-4990. Disponível em: <<https://www.mdpi.com/2504-4990/5/4/83>>. Citado 4 vezes nas páginas 9, 37, 38 e 39.
- TRIPATHI, A.; KOPPARAPU, S. K. Cnn based parkinson’s disease assessment using empirical mode decomposition. In: . Mumbai, India: CIKM (Workshops), 2020. Citado na página 27.
- TRIPATHI, S.; VERMA, A.; SHARMA, N. Automatic segmentation of brain tumour in mr images using an enhanced deep learning approach. Taylor & Francis, v. 9, n. 2, p. 121–130, 2021. Citado na página 27.
- VASCONCELOS, F. F. X. et al. Automatic skin lesions segmentation based on a new morphological approach via geodesic active contour. *Cognitive Systems Research*, Elsevier, v. 55, p. 44–59, 2019. Citado 2 vezes nas páginas 29 e 75.
- VINCENT, P.; BENGIO, Y. Manifold parzen windows. In: *Advances in neural information processing systems*. San Diego, CA: NeurIPS, 2003. p. 849–856. Citado na página 44.
- WARE, J. H. et al. P values. In: *Medical uses of statistics*. New England: CRC Press, 2019. p. 181–200. Citado na página 50.
- XIE, X. et al. Robust segmentation of nucleus in histopathology images via mask r-cnn. In: SPRINGER. *International MICCAI Brainlesion Workshop*. Cham, 2018. p. 428–436. Citado na página 28.
- YAGÜE, F. J. et al. Defect detection and segmentation in x-ray images of magnesium alloy castings using the detectron2 framework. *arXiv preprint arXiv:2202.13945*, 2022. Citado na página 36.

YEUNG, D.-Y.; CHOW, C. Parzen-window network intrusion detectors. In: IEEE. *Object recognition supported by user interaction for service robots*. Quebec City, QC, Canada, 2002. v. 4, p. 385–388. Citado na página [43](#).

ZHENG, Z. et al. Distance-iou loss: Faster and better learning for bounding box regression. In: *Proceedings of the AAAI conference on artificial intelligence*. Palo Alto, California USArk, USA: AAAI Press, 2020. v. 34, n. 07, p. 12993–13000. Citado na página [38](#).