
Abordagem Computacional para Auxiliar a Análise e Classificação de Câncer de Próstata por Meio de Imagens

Maxwell Gomes da Silva



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Uberlândia
2025

Maxwell Gomes da Silva

**Abordagem Computacional para Auxiliar a
Análise e Classificação de Câncer de Próstata
por Meio de Imagens**

Tese de doutorado apresentada ao Programa de Pós-graduação da Faculdade de Computação da Universidade Federal de Uberlândia como parte dos requisitos para a obtenção do título de Doutor em Ciência da Computação.

Área de concentração: Ciência da Computação

Orientador: André Ricardo Backes

Coorientador: Bruno Augusto Nassif Travençolo

Uberlândia

2025

Ficha Catalográfica Online do Sistema de Bibliotecas da UFU
com dados informados pelo(a) próprio(a) autor(a).

S586 Silva, Maxwell Gomes da, 1984-
2025 Abordagem Computacional para Auxiliar a Análise e
Classificação de Câncer de Próstata por Meio de Imagens [recurso
eletrônico] / Maxwell Gomes da Silva. - 2025.

Orientador: André Ricardo Backes.

Coorientador: Bruno Augusto Nassif Travençolo.

Tese (Doutorado) - Universidade Federal de Uberlândia, Pós-
graduação em Ciência da Computação.

Modo de acesso: Internet.

DOI <http://doi.org/10.14393/ufu.te.2025.659>

Inclui bibliografia.

1. Computação. I. Backes, André Ricardo, 1981-, (Orient.). II.
Travençolo, Bruno Augusto Nassif, 1981-, (Coorient.). III.
Universidade Federal de Uberlândia. Pós-graduação em Ciência da
Computação. IV. Título.

CDU: 681.3

Bibliotecários responsáveis pela estrutura de acordo com o AACR2:

Gizele Cristine Nunes do Couto - CRB6/2091

Nelson Marcos Ferreira - CRB6/3074



ATA DE DEFESA - PÓS-GRADUAÇÃO

Programa de Pós-Graduação em:	Ciência da Computação				
Defesa de:	Tese, 42/2025, PPGCO				
Data:	13 de Novembro de 2025	Hora de início:	13:00	Hora de encerramento:	17:00
Matrícula do Discente:	12113CCP004				
Nome do Discente:	Maxwell Gomes da Silva				
Título do Trabalho:	Abordagem Computacional para Auxiliar a Análise e Classificação de Câncer de Próstata por Meio de Imagens.				
Área de concentração:	Ciência da Computação				
Linha de pesquisa:	Ciência de Dados				
Projeto de Pesquisa de vinculação:	-----				

Reuniu-se por videoconferência, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Ciência da Computação, assim composta: Professores Doutores: Bruno Augusto Nassif Travençolo (Coorientador) - FACOM/UFU, Marcelo Zanchetta do Nascimento - FACOM/UFU, Henrique Coelho Fernandes - Cranfield University, Paulo Sergio Silva Rodrigues - UFABC, Alexandre Luís Magalhães Levada - DC/UFSCar e André Ricardo Backes - DC/UFSCar, orientador do(a) candidato(a).

Os examinadores participaram desde as seguintes localidades: André Ricardo Backes - São Carlos/SP, Henrique Coelho Fernandes - Milton Keynes/Inglaterra, Paulo Sergio Silva Rodrigues - São Bernardo do Campo/SP, Alexandre Luís Magalhães Levada - São Carlos/Sp. Os outros membros da banca e o aluno(a) participaram da cidade de Uberlândia.

Iniciando os trabalhos o presidente da mesa, Prof. Dr. André Ricardo Backes, apresentou a Comissão Examinadora e o(á) candidato(a), agradeceu a presença do público, e concedeu ao(á) Discente a palavra para a exposição do seu trabalho. A duração da apresentação do(a) Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir o senhor presidente concedeu a palavra, pela ordem sucessivamente, aos examinadores, que passaram a arguir o(á) candidato(a). Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o candidato(a):

Aprovado

Esta defesa faz parte dos requisitos necessários à obtenção do título de Doutor.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.



Documento assinado eletronicamente por **Alexandre Luis Magalhães Levada, Usuário Externo**, em 14/11/2025, às 15:14, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Bruno Augusto Nassif Travençolo, Professor(a) do Magistério Superior**, em 14/11/2025, às 16:21, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Henrique Coelho Fernandes, Professor(a) do Magistério Superior**, em 14/11/2025, às 18:41, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **André Ricardo Backes, Usuário Externo**, em 15/11/2025, às 09:41, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Paulo Sergio Silva Rodrigues, Usuário Externo**, em 17/11/2025, às 23:28, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Marcelo Zanchetta do Nascimento, Professor(a) do Magistério Superior**, em 18/11/2025, às 10:15, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **6845594** e o código CRC **5CBDAB37**.

*“Este trabalho é dedicado à memória do meu pai,
cuja batalha contra o câncer marcou minha vida
e despertou em mim o desejo de transformar dor em propósito”.*

Agradecimentos

Dedico esta seção para expressar minha profunda gratidão às pessoas e instituições que foram pilares essenciais na conclusão desta trajetória acadêmica.

Primeiramente, agradeço a Deus pelo dom da sabedoria e pela força que me guiaram ao longo deste caminho. Sua orientação foi o alicerce para que este trabalho fosse alcançado.

Ao meu esposo, Marcos Rivânio Gomes, toda a minha gratidão. Seu incentivo constante, paciência e otimismo foram fundamentais para que eu enxergasse o melhor em mim e superasse os desafios desta jornada.

Agradeço imensamente a Romualdo Filho, cuja generosidade em disponibilizar o servidor para o processamento de imagens foi uma contribuição instrumental para a execução da fase prática deste projeto.

Em memória do meu querido pai, Mário Antônio da Silva, que, ao lado de minha mãe, Eleusa Maria Gomes da Silva, me concedeu os maiores presentes: a vida e o apoio incondicional em todos os meus passos.

Aos meus orientadores, Professores André Backes e Bruno Travençolo, meu sincero reconhecimento. Sua orientação técnica, rigor acadêmico e encorajamento contínuo foram indispensáveis para a realização deste marco em minha formação.

À Universidade Federal de Uberlândia (UFU) e à Faculdade de Computação (FACOM), agradeço pela excelência do ambiente acadêmico e pela infraestrutura oferecida.

Por fim, agradeço a todos que, direta ou indiretamente, contribuíram para esta conquista. Cada gesto de apoio, conselho e incentivo foi profundamente valorizado e essencial para a conclusão deste trabalho.

A todos, minha mais sincera gratidão.

*“Quando o desconhecido se torna um desafio, é nele que encontramos a força para superar
nossos limites e descobrir o melhor de nós mesmos.”
(Marxwell Gomes da Silva)*

Resumo

A análise de imagens histológicas de lâminas inteiras (WSI) no diagnóstico do câncer de próstata impõe desafios computacionais, como o processamento de imagens em gigapixels e a variabilidade morfológica dos tecidos. Este trabalho propõe uma metodologia que integra um pré-processamento baseado em operações morfológicas a uma arquitetura de rede neural convolucional do tipo Mask R-CNN para segmentação de instâncias. Aplicada ao conjunto de dados PANDA, a abordagem incluiu a construção de um banco de dados curado, com geração de *patches* e balanceamento de classes. O modelo, implementado no *framework* Detectron2, obteve acurácia de 97,87% na classificação dos padrões de Gleason, confirmando que o pré-processamento direcionado favorece o aprendizado da rede em relação às abordagens *end-to-end* e foi possível realizar também a validação por meio do subconjunto Karolinska. Os resultados indicam o potencial da metodologia como ferramenta para patologia digital, com capacidade de generalização, estabelecendo um novo referencial para aplicação relacionados aos desafios das imagens médicas.

Palavras-chave: Processamento de Imagem, Câncer de Próstata, Segmentação de Imagem, Inteligência Artificial, Redes Neurais Convolucionais, Detectron2.

Abstract

The analysis of whole-slide histological images (WSI) in prostate cancer diagnosis presents significant computational challenges, such as gigapixel image processing and tissue morphological variability. This work proposes a methodology that integrates a preprocessing stage based on morphological operations with a convolutional neural network architecture of the Mask R-CNN type for instance segmentation. Applied to the PANDA dataset, the approach included the construction of a curated database with patch generation and class balancing. The model, implemented using the Detectron2 framework, achieved an accuracy of 97.87% in classifying Gleason patterns, confirming that targeted preprocessing enhances network learning compared to end-to-end approaches. Validation was also performed using the Karolinska subset. The results indicate the potential of the proposed methodology as a tool for digital pathology, demonstrating its generalization capability and establishing a new reference for applications addressing challenges in medical imaging.

Keywords: Image Processing, Prostate Cancer, Image Segmentation, Artificial Intelligence, Convolutional Neural Networks, Detectron2.

Lista de ilustrações

Figura 1 –	Escore de Gleason (BRAZIL, 2022).	24
Figura 2 –	Sistema genital masculino, destacando a próstata e exibindo região com câncer (SOCIETY, 2023).	27
Figura 3 –	Exemplo de imagens histológicas e principais estruturas das células da próstata, obtida em (KAGGLE, 2023; MESCHER, 2021). Legenda: A - lúmen, B - células secretoras, C - Citoplasma, D - Estroma, E - Núcleos celulares, F - Glândulas.	29
Figura 4 –	Visão geral do desafio PANDA e do desenho do estudo, conforme apresentado em (BULTEN et al., 2022). A fase de desenvolvimento (esquerda) envolveu uma competição internacional onde as equipes treinaram seus modelos. A fase de validação (direita) realizou uma avaliação cega e independente dos algoritmos selecionados em múltiplos <i>datasets</i>	32
Figura 5 –	Arquitetura MIL proposta em (MAI et al., 2024). A WSI é dividida em instâncias (<i>patches</i>). Um fluxo processa as características espaciais, enquanto o outro (SFRM-WT) reconstrói características no domínio da frequência. Um módulo de atenção cruzada funde as informações para gerar a representação final da lâmina (saco).	34
Figura 6 –	Arquitetura do <i>Hierarchical Vision Transformer (HIPT)</i> (CHEN et al., 2023). O modelo processa a WSI em escalas crescentes, desde características celulares (16×16) até fenótipos de tecido (4096×4096), para formar uma representação final da lâmina inteira.	34
Figura 7 –	Protocolo de criação do <i>dataset</i> CrowdGleason (LÓPEZ-PÉREZ et al., 2024). O fluxo ilustra a aquisição de WSIs, a extração de <i>patches</i> , o processo de rotulação distribuída (<i>crowd labeling</i>) para gerar o conjunto de treinamento, e a curadoria com especialistas para criar um conjunto de teste confiável.	35

Figura 8 – Diagrama de blocos do algoritmo de segmentação baseado em GAN proposto por (VATS et al., 2024). O modelo gerador recebe a imagem de entrada e produz uma máscara de segmentação, que é avaliada pelo discriminador em comparação com a máscara real (<i>ground-truth</i>).	36
Figura 9 – Exemplo de estrutura de Redes Neurais Convolucionais - <i>Convolutional Neural Networks</i> (CNN). A imagem de entrada passa por camadas de convolução e subamostragem para extração de mapas de características, que são então utilizados por uma camada totalmente conectada para a classificação final.	40
Figura 10 – Estrutura geral do Mask R-CNN. A imagem passa por um <i>backbone</i> (CNN) para extrair características. A RPN propõe regiões (ROIs), que são alinhadas pela camada ROIAlign. Em seguida, as cabeças da rede realizam a classificação, o refinamento da caixa delimitadora e a geração da máscara de segmentação para cada instância. Figura adaptada de (CHIAO et al., 2019).	43
Figura 11 – Passos da metodologia proposta.	46
Figura 12 – Exemplo do processo de graduação de Gleason para uma biópsia contendo câncer de próstata. Os padrões de crescimento canceroso mais comum (contorno azul, padrão Gleason 3) e o segundo mais comum (contorno vermelho, padrão Gleason 4) presentes na biópsia determinam a pontuação Gleason (3+4 para esta biópsia), que por sua vez é convertida em um grau <i>International Society of Urological Pathology</i> (ISUP) (2 para esta biópsia) seguindo as diretrizes da Sociedade Internacional de Patologia Urológica. Biópsias que não contêm câncer são representadas por um grau ISUP de 0 no conjunto de dados apresentado (KAGGLE, 2023).	47
Figura 13 – A imagem apresenta uma seção histológica de tecido em que foram realizadas operações morfológicas. As áreas indicadas pela borda de cor azul claro (A) representam as regiões em que as operações morfológicas foram aplicadas, resultando na remoção de determinadas estruturas. As áreas contornadas em preto (B) representam as regiões que foram mantidas e consideradas para o cálculo do escore de Gleason.	51
Figura 14 – Exemplo de pré-processamento: A) Imagem original; B) Máscara com anotações do <i>dataset</i> ; C) Máscara com as regiões delimitadoras para segmentar em <i>patches</i> ; e D) Imagem segmentada em <i>patches</i> por regiões determinadas por C.	53
Figura 15 – Exemplo de imagem carregada para realizar análise na GUI desenvolvida para este propósito.	56
Figura 16 – Exemplo de imagem processada com segmentação do escore de Gleason.	57

Figura 17 – Precisão ao longo das épocas.	64
Figura 18 – <i>F1-Score</i> ao longo das épocas.	64
Figura 19 – Precisão classificação ao longo das épocas.	64
Figura 20 – Perda segmentação ao longo das épocas.	64
Figura 21 – Comparação entre o resultado da pesquisa e imagens com anotações da base Karolinska. Legenda: Na primeira linha o escore identificado é de Grau 4 e na Segunda linha Grau 3.	66
Figura 22 – Imagem histológica com segmentação e classificação automáticas. A região destacada em vermelho B foi classificada como Escore de Gleason 4, enquanto a região em amarelo A foi classificada como Escore de Gleason 3. A máscara destacada em azul C foi utilizada como referência de anotação do <i>dataset</i> Karolinska, Grau 4.	67
Figura 23 – Métricas de avaliação dos resultados obtidos do treinamento no <i>dataset</i> PANDA.	69
Figura 24 – Kappa estimado em comparação com trabalhos da literatura.	69

Lista de tabelas

Tabela 1	– Tabela comparativa dos principais trabalhos relacionados.	38
Tabela 2	– Distribuição dos escores de Gleason por Instituto.	48
Tabela 3	– Resultados por Época no Conjunto PANDA (Intervalo de 4 Épocas). .	65
Tabela 4	– Resultados por Época no Conjunto SICAPV2 (Intervalo de 12 Épocas)	65
Tabela 5	– Equivalência qualitativa entre F1-Score, IoU e k_q estimado, com base em referências consolidadas.	70
Tabela 6	– Análise comparativa com trabalhos selecionados.	71

Lista de siglas

CNN Redes Neurais Convolucionais - *Convolutional Neural Networks*

COCO *Common Objects in Context*

H&E Hematoxilina-Eosina

ISUP *International Society of Urological Pathology*

PANDA *Prostate Cancer Grade Assessment*

TIFF *Tagged Image File Format*

WSI Lâminas Histológicas Inteiras - *Whole Slide Images*

Sumário

1	INTRODUÇÃO	23
1.1	Motivação	25
1.2	Hipótese	25
1.3	Objetivos e Desafios da Pesquisa	26
1.4	Organização desta Tese	26
2	CONTEXTUALIZAÇÃO	27
2.1	A próstata	27
2.2	Análise de Imagens Histológicas para Câncer de Próstata . . .	28
2.3	Trabalhos Relacionados	30
2.3.1	Síntese e Lacunas da Literatura	36
2.4	Processamento Digital de Imagens em Patologia	39
2.5	Redes Neurais Convolucionais	40
2.6	Segmentação de Instâncias com <i>Mask R-CNN</i> e Arquiteturas Avançadas	42
3	METODOLOGIA DE PESQUISA	45
3.1	Construção do Banco de Imagens	46
3.2	Pré-processamento	49
3.3	Treinamento da Rede Neural	52
3.3.1	Aplicação para Análise de Imagens	55
4	RESULTADOS E DISCUSSÃO	59
4.1	Treinamento do Modelo	59
4.2	Resultados e Experimentos para Autodetecção	63
4.3	Análise Comparativa com Estudos Anteriores	68
5	CONCLUSÃO	73

REFERÊNCIAS 77

Introdução

O câncer de próstata é a quarta principal causa de morte por câncer no Brasil, correspondendo a 6% de todas as mortes nesta categoria. A taxa de mortalidade total tem mostrado uma forte tendência de aumento, com 16.301 mortes e cerca de 71.730 novos casos registrados em 2022, correspondendo a 29,2% dos tumores incidentes em homens (BRAZIL, 2022). Nos Estados Unidos, os registros indicam que há cerca de 288,3 mil novos casos e cerca de 34,7 mil mortes causadas por câncer de próstata em 2023. Além disso, as estatísticas mostram que um em cada oito homens será diagnosticado com esse tipo de neoplasia (SOCIETY, 2023).

Observa-se que tanto no Brasil quanto nos Estados Unidos, devido ao aumento de casos de tumor e da taxa de mortalidade causada pela doença, há ações em nível nacional e políticas para prevenção e controle do câncer. O objetivo é reduzir a incidência e a mortalidade por meio de ações contínuas, como conscientização dos fatores de risco do câncer, promoção da detecção precoce de cânceres rastreáveis e fornecimento de tratamento de qualidade em todo o país (SOCIETY, 2023; BRAZIL, 2022).

Assim como em outros cânceres, a idade é um marcador de risco e tem significado especial no câncer de próstata, uma vez que a morbidade e a mortalidade aumentam exponencialmente após os 50 anos no Brasil; nos Estados Unidos, o aumento ocorre em homens com 65 anos ou mais. A justificativa para orientar a detecção precoce do câncer de próstata é a possibilidade de reduzir os custos decorrentes do tratamento do câncer em estágios avançados ou doença metastática, além de aumentar as taxas de cura dos pacientes (SOCIETY, 2023; BRAZIL, 2022).

O diagnóstico do câncer de próstata pode ser feito por estudo histopatológico do tecido obtido por biópsia de próstata, o que deve ser considerado sempre que houver anormalidades no exame digital do reto ou na dosagem do Antígeno Específico da Próstata ou PSA (LOEB et al., 2014). Utilizando um laudo médico ou anatomopatológico, avalia-se o grau histológico do câncer, conhecido como Escore de Gleason (Figura 1). O objetivo é informar sobre a provável taxa de crescimento do tumor e sua tendência à disseminação, além de auxiliar no melhor tratamento para o paciente (BRAZIL, 2002).

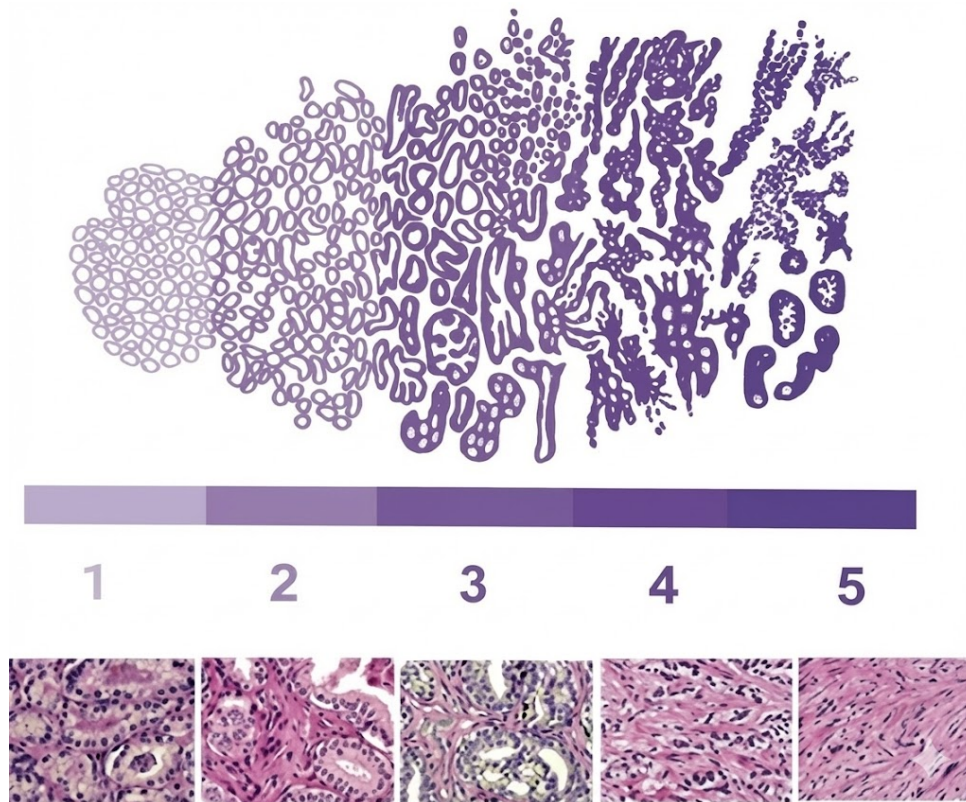


Figura 1 – Escore de Gleason (BRAZIL, 2022).

De acordo com o Instituto Nacional do Câncer, a escala de classificação do câncer de próstata varia de 1 a 5, como mostrado na Figura 1. São eles (BRAZIL, 2022):

- ❑ Grau 1 - As células são geralmente uniformes e pequenas, formando glândulas regulares com pouca variação de tamanho e forma, com margens bem definidas, densamente agrupadas e distribuídas uniformemente, com pouco estroma entre elas.
- ❑ Grau 2 - As células variam mais em tamanho e forma, as glândulas permanecem uniformes, com nódulos soltos com bordas irregulares.
- ❑ Grau 3 - As células são mais variáveis em tamanho e forma, constituindo glândulas muito pequenas, uniformes, angulares ou alongadas, individualmente e irregularmente distribuídas ao longo do estroma. Elas também podem formar nódulos fusiformes ou papilares com bordas lisas.
- ❑ Grau 4 - Muitas células se fundem em grandes massas amorfas ou formam glândulas irregulares, de distribuição desigual, mostrando infiltração e invasão irregulares de tecidos adjacentes.
- ❑ Grau 5 - Tumor anaplásico. A maioria das células se agrega em grandes aglomerados e invade órgãos e tecidos adjacentes. A massa celular pode apresentar necrose central, com padrão comedocarcinoma. A diferenciação glandular frequentemente pode estar ausente: padrões de crescimento celular cordão-like ou infiltração de células soltas.

1.1 Motivação

O diagnóstico do câncer de próstata, embora bem estabelecido clinicamente através do Escore de Gleason, apresenta desafios que motivam a busca por soluções computacionais. A análise manual de lâminas histológicas é um trabalho sujeito a viés e fadiga, dependendo da disponibilidade e da experiência do patologista, o que pode levar a variabilidades no diagnóstico. A aplicação de técnicas de inteligência artificial surge como uma oportunidade para aumentar a eficiência, a reprodutibilidade e a precisão do processo diagnóstico.

Do ponto de vista da Ciência da Computação, o problema impõe desafios complexos. As lâminas histológicas inteiras (WSIs) são imagens em escala de gigapixels, cujo processamento demanda alto poder computacional e algoritmos especializados (CHEN et al., 2023). Além disso, a variabilidade na coloração H&E e a complexidade morfológica dos tecidos exigem modelos com alta capacidade de generalização. Embora abordagens baseadas em aprendizado profundo tenham alcançado resultados promissores, como a acurácia de 77% relatada por (RODRIGUEZ et al., 2020) no *dataset* SICAPv2, muitos trabalhos aplicam redes neurais diretamente sobre as imagens, tratando o modelo como uma “caixa-preta”. Esta abordagem, embora funcional, negligencia a oportunidade de otimizar a extração de características e carece de interpretabilidade sobre quais regiões da imagem contribuíram para o diagnóstico.

A hipótese central desta tese é que um pré-processamento direcionado, que utiliza conhecimento de domínio para realçar estruturas glandulares através de operações morfológicas, permite que uma rede neural aprenda características mais discriminativas. Argumenta-se que, ao invés de uma abordagem *end-to-end*, a criação de um *pipeline* que limpa e prepara os dados melhora a capacidade de generalização do modelo para novos dados. Nesse contexto, a contribuição principal é um *pipeline* computacional agnóstico que demonstra como técnicas clássicas (morfologia) podem estabilizar o treinamento de redes profundas (*Deep Learning*) em cenários de alta variabilidade visual. Portanto, esta tese propõe e valida uma metodologia computacional que não apenas aplica o aprendizado profundo, mas o otimiza através de uma etapa de pré-processamento fundamentada, visando estabelecer um novo padrão de precisão para a análise do *dataset* PANDA.

1.2 Hipótese

A partir dos pressupostos e da análise do estado da arte no diagnóstico do câncer de próstata, este trabalho formula a hipótese de que a aplicação de uma arquitetura de segmentação de instâncias, enriquecida por um pré-processamento específico para a morfologia de tecidos prostáticos, pode proporcionar uma detecção e classificação mais precisa, reprodutível e clinicamente interpretável dos graus de câncer de próstata, em comparação com abordagens que tratam a tarefa como um problema de classificação de

imagem em caixa-preta.

1.3 Objetivos e Desafios da Pesquisa

Desenvolver e validar uma metodologia computacional, baseada em visão computacional e aprendizado profundo, para a segmentação e classificação automática de padrões de Escore de Gleason em imagens histológicas de lâminas inteiras de câncer de próstata.

Os objetivos específicos incluem:

- ❑ Construir um banco de dados curado e balanceado a partir do *dataset* PANDA, otimizado para o treinamento de redes neurais convolucionais.
- ❑ Implementar um *pipeline* de pré-processamento de imagens para realçar estruturas glandulares e remover artefatos.
- ❑ Avaliar um modelo de segmentação de instâncias (Mask R-CNN) quanto à sua capacidade de identificar e classificar os diferentes graus de Gleason.
- ❑ Comparar o desempenho da metodologia proposta com o estado da arte, analisando métricas de acurácia e segmentação.

1.4 Organização desta Tese

Este trabalho está estruturado em cinco capítulos para facilitar o entendimento e a progressão lógica da pesquisa. O **Capítulo 1** introduz o problema sob a ótica da saúde e da computação, delineando a motivação, os objetivos e a hipótese. O **Capítulo 2** apresenta a contextualização teórica, revisa o estado da arte computacional e detalha os fundamentos de processamento de imagem e *deep learning* aplicados à patologia digital. O **Capítulo 3** descreve a metodologia de pesquisa, desde a construção do banco de dados até o treinamento da rede neural. O **Capítulo 4** discute os experimentos e analisa os resultados obtidos. Por fim, o **Capítulo 5** apresenta a conclusão, as contribuições do trabalho e aponta direções para pesquisas futuras.

Contextualização

Neste capítulo, apresentam-se as teorias que fundamentam esta pesquisa, acompanhadas por uma revisão da literatura. O objetivo é proporcionar uma compreensão dos referenciais teóricos que sustentam este estudo e examinar o conjunto de trabalhos relacionados ao tema desta pesquisa.

2.1 A próstata

A próstata, glândula do sistema reprodutivo masculino, localiza-se anteriormente ao reto e inferiormente à bexiga urinária. Seu tamanho varia conforme a idade do indivíduo; o órgão compara-se a uma noz em homens jovens, mas pode aumentar em homens com idade avançada (ONCOGUIA, 2014). Sua função primária consiste na produção de um fluido essencial para a proteção e nutrição dos espermatozoides no sêmen, conferindo-lhe maior fluidez. Nas proximidades da próstata, situam-se as vesículas seminais, responsáveis pela produção da maior parte do fluido seminal. A uretra, canal que permite a passagem de urina e sêmen para fora do corpo através do pênis, atravessa o centro da próstata (ONCOGUIA, 2014).

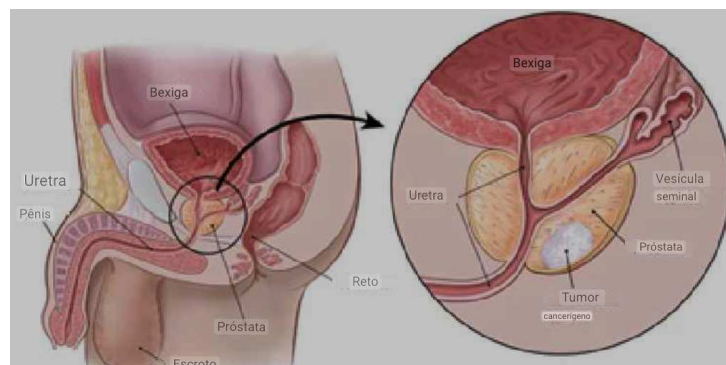


Figura 2 – Sistema genital masculino, destacando a próstata e exibindo região com câncer (SOCIETY, 2023).

A manutenção da saúde da próstata é um fator importante para a qualidade de vida, dada a prevalência de condições como a Hiperplasia Prostática Benigna (HPB) e o câncer de próstata (Adenocarcinoma). A HPB refere-se ao aumento não canceroso da próstata, comum em homens mais velhos, que pode causar dificuldades urinárias devido à compressão da uretra.

Embora o termo “câncer de próstata” seja utilizado de forma ampla, este trabalho foca especificamente no adenocarcinoma acinar, o subtipo histológico mais prevalente e aquele para o qual o sistema de graduação de Gleason foi originalmente desenvolvido e é primariamente aplicado (EPSTEIN et al., 2016).

Por outro lado, o câncer de próstata figura entre os tipos mais comuns de neoplasia entre os homens, desenvolvendo-se frequentemente de forma lenta e sem sintomas evidentes nos estágios iniciais. A detecção precoce do câncer de próstata eleva a probabilidade de um tratamento bem-sucedido. Exames regulares constituem os métodos de diagnóstico (ONCOGUIA, 2014):

- ❑ O exame digital retal, um procedimento físico, avalia a próstata por meio da inserção de um dedo lubrificado no reto para palpação.
- ❑ O teste de antígeno prostático específico ou *Prostate-specific antigen* (PSA) é um exame de sangue que detecta um antígeno produzido pela próstata, podendo sinalizar neoplasia ou outras condições.
- ❑ A biópsia, um procedimento cirúrgico, coleta amostras de tecido prostático para análise microscópica, fornecendo dados morfológicos detalhados para o diagnóstico de doenças da próstata.

A análise histológica de amostras de tecido, utilizando imagens, possibilita a identificação de células cancerosas e a avaliação da extensão da doença. As imagens resultantes fornecem detalhes sobre a morfologia celular, a estrutura do tecido e a gravidade das lesões encontradas.

2.2 Análise de Imagens Histológicas para Câncer de Próstata

Estudos diversos corroboram a eficácia da análise histológica na detecção do câncer de próstata. Uma pesquisa conduzida por (TRABULSI et al., 2020) demonstrou sensibilidade superior a 90% para este método na detecção do câncer prostático. Além disso, técnicas como a microscopia digital têm apresentado aumento na precisão diagnóstica e na identificação de lesões de alto risco. A histologia da próstata apresenta elementos com formas e disposições de relevância para o estudo do desenvolvimento de doenças. No contexto do câncer de próstata, os núcleos – organelas membranosas localizadas no

centro das células, responsáveis pelo controle da expressão gênica – são cruciais para a análise (MESCHER, 2021). Esses núcleos fornecem informações sobre o estado do tumor, empregadas no diagnóstico. Na Figura 3, observa-se a disposição nuclear em um epitélio glandular saudável. A compreensão detalhada dessa estrutura, em conjunto com a análise dos núcleos, desempenha um papel fundamental na avaliação da condição neoplásica da próstata.

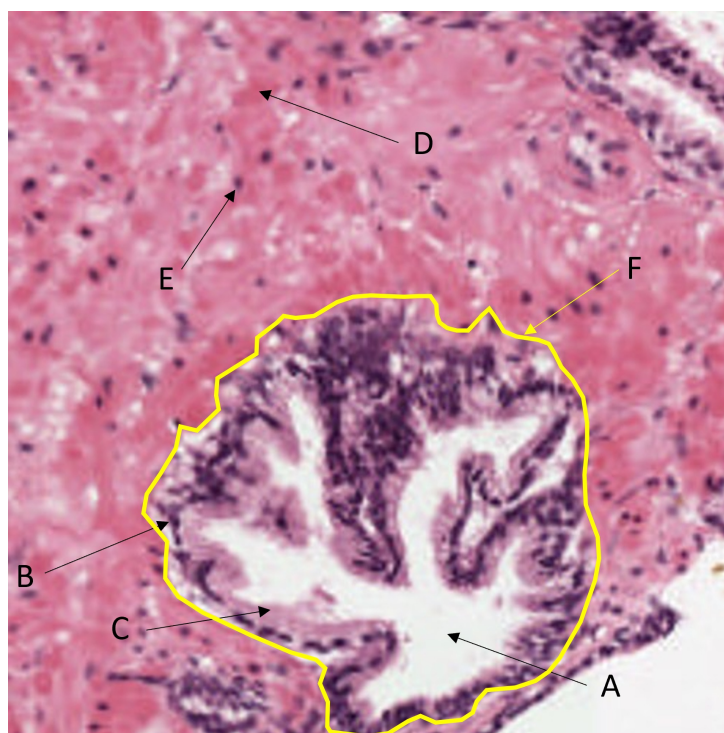


Figura 3 – Exemplo de imagens histológicas e principais estruturas das células da próstata, obtida em (KAGGLE, 2023; MESCHER, 2021). Legenda: A - lúmen, B - células secretoras, C - Citoplasma, D - Estroma, E - Núcleos celulares, F - Glândulas.

Na Figura 3, destacam-se as seguintes estruturas (MESCHER, 2021): A) O lúmen, que representa o espaço central dentro de uma estrutura tubular; B) Células secretoras, especializadas na produção de substâncias; C) Citoplasma, a região celular entre a membrana e o núcleo; D) Estroma, o tecido conjuntivo de suporte; E) Núcleos, que abrigam o material genético; F) Glândulas, estruturas secretoras.

Por meio da coloração Hematoxilina-Eosina (H&E), a distinção entre esses tipos celulares pode ser complexa. Na Figura 3, as células secretoras estão dispostas em colunas logo após o lúmen, seguidas pelas células basais (MESCHER, 2021). As células fibromusculares alongadas são identificáveis na imagem, distribuídas por todo o estroma (SHAH; ZHOU, 2012).

Por meio dessas imagens de biópsia é possível classificar o câncer de próstata, um passo crucial na determinação do prognóstico do paciente e na orientação das decisões terapêuticas. As imagens resultantes da análise histológica permitem a classificação do

câncer segundo o sistema de Escore de Gleason, amplamente utilizado na prática clínica (SHAH; ZHOU, 2012). O Escore de Gleason classifica as células cancerosas com base em sua aparência e padrão de crescimento. Esse sistema, que avalia a diferenciação celular e a arquitetura tumoral, categoriza o câncer de próstata em uma escala de 1 a 5, considerando os padrões predominante e secundário (SOCIETY, 2023). Se o câncer de próstata for identificado em uma biópsia, atribui-se uma classificação baseada na avaliação microscópica da aparência anômala do câncer. Cânceres de grau mais alto exibem características mais anômalas e apresentam maior propensão a crescer e se disseminar. Há duas formas principais de descrever essa classificação (SOCIETY, 2013):

- ❑ Se o câncer apresentar aparência semelhante ao tecido normal, pode receber uma pontuação de 1 ou 2, embora estas pontuações raramente sejam utilizadas em diagnósticos modernos de câncer.
- ❑ Cânceres com características anormais recebem uma pontuação de 5.
- ❑ As pontuações de 3 e 4 correspondem a características intermediárias entre os padrões normais e anormais.
- ❑ Em geral, cânceres clinicamente relevantes são classificados com pontuações de 3 a 5.

Como o câncer de próstata pode possuir áreas distintas com graus variados, uma classificação é atribuída às duas áreas que compõem a maioria das regiões do câncer. Esses dois escores são somados para produzir o Escore de Gleason. O primeiro número atribuído corresponde ao grau mais prevalente no tumor. Por exemplo, um Escore de Gleason de $3+4=7$ indica que a maior parte do tumor é de grau 3, com uma proporção menor de grau 4 (SOCIETY, 2013). Teoricamente, o Escore de Gleason pode variar de 2 a 10, mas escores abaixo de 6 são raramente utilizados. Com base no Escore de Gleason, o câncer de próstata é frequentemente categorizado em três grupos (SOCIETY, 2013):

- ❑ Cânceres com um Escore de Gleason de 6 ou menos são classificados como bem diferenciados ou de baixo grau.
- ❑ Cânceres com um Escore de Gleason de 7 são categorizados como moderadamente diferenciados ou de grau intermediário.
- ❑ Cânceres com escores de Gleason de 8 a 10 são descritos como pouco diferenciados ou de alto grau.

Não obstante, este método pode enfrentar limitações devido à subjetividade interpessoal e à variabilidade na interpretação da amostra. A classificação de Gleason, combinada com outras informações clínicas, auxilia na definição do tratamento adequado para cada paciente.

2.3 Trabalhos Relacionados

A revisão da literatura foi conduzida utilizando as bases de dados Scopus, IEEE Xplore e Google Scholar. A estratégia de busca empregou os seguintes operadores booleanos: ('Prostate Cancer' OR 'Gleason Score') AND ('Deep Learning' OR 'Mask R-CNN' OR 'Instance Segmentation'). Os critérios de inclusão compreenderam artigos publicados entre 2018–2024, com foco em histopatologia digital e no uso de CNNs. Foram excluídos trabalhos restritos a dados clínicos (não-imagem) ou baseados em ressonância magnética. A seguir, apresenta-se uma análise dos trabalhos selecionados.

O trabalho de (BULTEN et al., 2022) representa um marco na validação de algoritmos de IA em patologia, ao sumarizar os resultados do desafio PANDA (Prostate cANcer graDe Assessment). A metodologia do desafio, ilustrada na Figura 4, foi concebida para testar rigorosamente a generalização dos modelos. Estruturado em duas fases, o desafio iniciou com uma competição aberta que disponibilizou um vasto conjunto de dados de 10.616 biópsias para o treinamento dos algoritmos por 1.290 desenvolvedores. A segunda fase consistiu em uma validação cega e independente, na qual os algoritmos de melhor desempenho foram avaliados em coortes externas de múltiplos centros dos EUA e da Europa, com o padrão-ouro estabelecido por consensos de painéis de uropatologistas. Esta abordagem metodológica foi fundamental para expor os algoritmos a variações do mundo real, como diferentes protocolos de preparação de lâminas, *scanners* de digitalização e perfis de pacientes.

O principal resultado do desafio foi a demonstração empírica de que múltiplos algoritmos, desenvolvidos com abordagens distintas (incluindo redes convolucionais, métodos baseados em atenção e diferentes estratégias de pré-processamento), foram capazes de atingir desempenho comparável ao de patologistas humanos. A métrica de avaliação primária, o coeficiente Kappa quadrático ponderado, alcançou valores de 0,862 (IC 95%, 0,840-0,884) e 0,868 (IC 95%, 0,835-0,900) nas coortes de validação dos EUA e da Europa, respectivamente. Tais valores, próximos da unidade, indicam uma concordância elevada entre as predições da IA e o padrão-ouro, validando a robustez e o potencial de aplicabilidade clínica dos modelos para a graduação de Gleason.

A contribuição de (BULTEN et al., 2022) reside na criação de um parâmetro de referência de alta complexidade e na prova de conceito de que a IA pode atingir desempenho de nível especialista de forma generalizável. Para esta tese, o conjunto de dados PANDA, validado por este estudo, serve não apenas como a fonte de dados primária, mas também como o padrão de referência que define a complexidade do problema. Contudo, uma limitação fundamental da abordagem do desafio é o foco na classificação em nível de lâmina inteira (*slide-level*), que resulta em um único escore ISUP para toda a biópsia. Isso evidencia uma lacuna na localização espacial e na quantificação de múltiplos padrões de Gleason coexistentes no mesmo tecido, que é precisamente o problema de segmentação de instâncias que esta tese se propõe a resolver.

Para lidar com a natureza fracamente rotulada das WSIs, (MAI et al., 2024) explora o paradigma de *Multiple Instance Learning* (MIL). O MIL trata a WSI como um conjunto de instâncias (*patches*) e atribui um rótulo único a todo o conjunto, dispensando anotações em nível de *pixel*. A arquitetura proposta, FRCM-MIL, é um modelo de duplo fluxo (Figura 5): um ramo utiliza uma CNN para extrair características convencionais do domínio espacial, enquanto o segundo ramo (SFRM-WT) emprega a transformada *wavelet* para decompor os sinais da imagem em componentes de baixa e alta frequência, capturando informações de textura e bordas que poderiam ser perdidas no processamento espacial. Um Módulo de Atenção Cruzada (CAM) é subsequentemente utilizado para fundir as representações de ambos os domínios de forma ponderada.

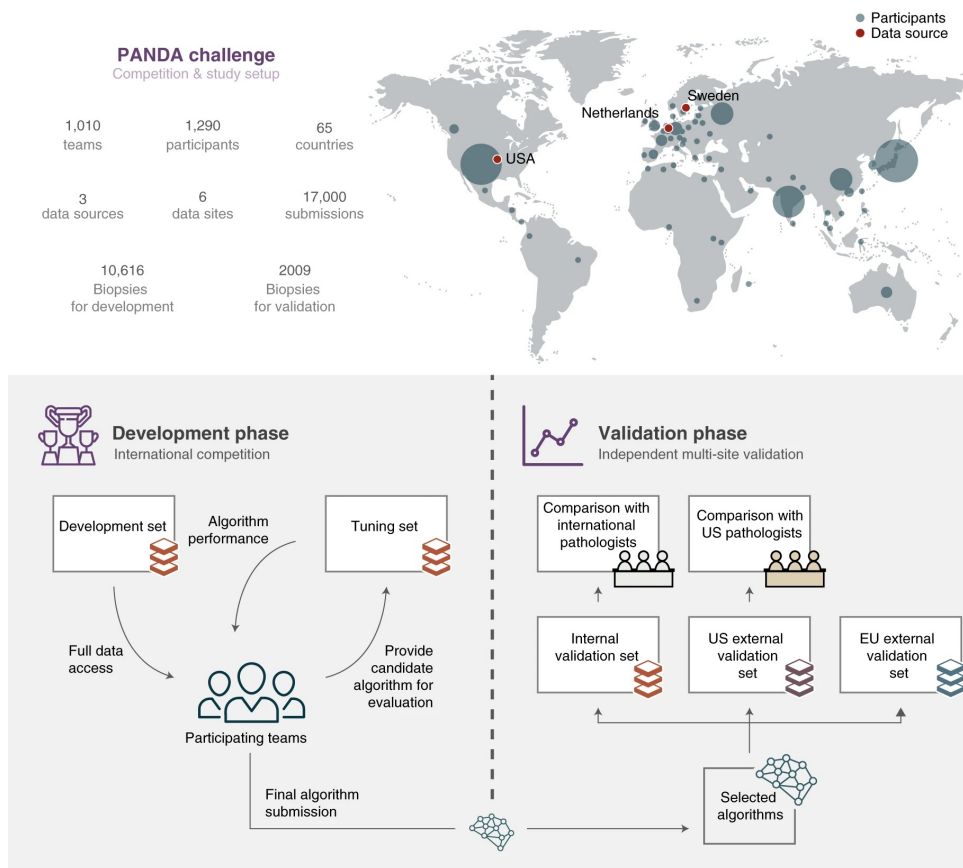


Figura 4 – Visão geral do desafio PANDA e do desenho do estudo, conforme apresentado em (BULTEN et al., 2022). A fase de desenvolvimento (esquerda) envolveu uma competição internacional onde as equipes treinaram seus modelos. A fase de validação (direita) realizou uma avaliação cega e independente dos algoritmos selecionados em múltiplos *datasets*.

Em termos de desempenho no conjunto de dados PANDA, a abordagem FRCM-MIL alcançou uma AUC (Área Sob a Curva ROC) de 91,69%, demonstrando boa capacidade de discriminar entre os diferentes graus de Gleason em nível de lâmina. A acurácia de 67,24% reflete a dificuldade inerente à classificação multiclasse em um conjunto de dados com desbalanceamento e alta variabilidade. A metodologia de agregação de características

vizinhas através do módulo CQAM, que utiliza o algoritmo K-vizinhos mais próximos (KNN), também contribuiu para a robustez do modelo ao gerar representações de consulta mais estáveis para o mecanismo de atenção.

A principal contribuição deste trabalho é apresentar uma arquitetura MIL que integra explicitamente características do domínio da frequência para melhorar a classificação de WSIs. No entanto, a limitação intrínseca do MIL persiste: a perda de interpretabilidade espacial. O modelo fornece uma predição para a lâmina inteira, mas não gera um mapa de segmentação que localize os diferentes padrões tumorais. Esta abordagem contrasta com o objetivo desta tese, que é empregar a segmentação de instâncias para fornecer não apenas o diagnóstico, mas também a localização e a extensão exatas de cada padrão, informação de elevado valor para o planejamento terapêutico e a avaliação da heterogeneidade tumoral.

Outra frente de pesquisa foca em arquiteturas capazes de lidar com a escala de gigapixels das WSIs de forma hierárquica. (CHEN et al., 2023) propõe o HIPT (*Hierarchical Image Pyramid Transformer*), um modelo baseado em *Vision Transformers* (ViTs) e treinado com aprendizado auto-supervisionado. A metodologia do HIPT, detalhada na Figura 6, processa a imagem em múltiplas escalas de forma ascendente (*bottom-up*). O processo inicia com um ViT operando em *patches* de 256×256 *pixels*, que aprende representações de “organização celular” ao tratar blocos de 16×16 *pixels* como “tokens”. As representações vetoriais resultantes são então utilizadas como uma nova sequência de “tokens” para um segundo nível de ViT, que analisa regiões maiores de 4096×4096 *pixels* para capturar “fenótipos de tecido”, como a interface tumor-estroma.

O modelo HIPT demonstrou desempenho de ponta em tarefas de subtipagem de câncer e predição de sobrevida em 33 tipos de câncer do *The Cancer Genome Atlas (TCGA)*. A análise dos mapas de atenção do modelo revelou que as camadas hierárquicas aprendem a focar em características histologicamente relevantes em suas respectivas escalas, desde células individuais até a organização tecidual mais ampla. Isso valida a hipótese de que a análise multiescala é fundamental para a extração de informações prognósticas de WSIs.

Apesar de sua *performance*, a abordagem HIPT possui limitações no contexto desta tese. Primeiro, sua complexidade computacional é alta, exigindo recursos de *hardware* para o treinamento em duas etapas. Segundo, o modelo foi projetado para gerar uma única representação vetorial para a WSI inteira, sendo otimizado para tarefas de classificação e prognóstico em nível de lâmina. A arquitetura não foi desenvolvida para a segmentação granular de múltiplos padrões de Gleason, que exige a delimitação precisa de fronteiras entre diferentes regiões tumorais, tarefa para a qual a segmentação de instâncias é mais adequada.

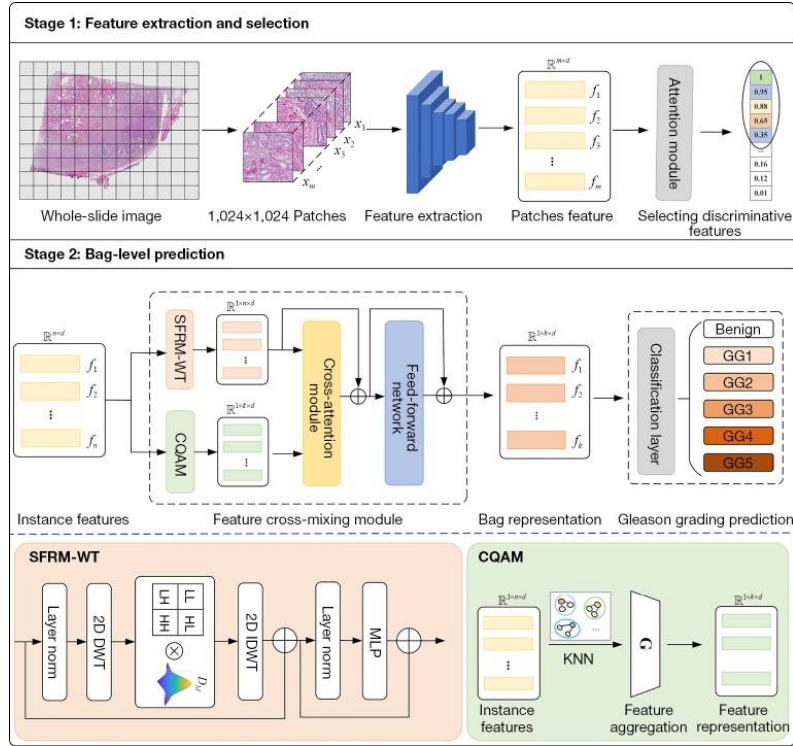


Figura 5 – Arquitetura MIL proposta em (MAI et al., 2024). A WSI é dividida em instâncias (*patches*). Um fluxo processa as características espaciais, enquanto o outro (SFRM-WT) reconstrói características no domínio da frequência. Um módulo de atenção cruzada funde as informações para gerar a representação final da lâmina (saco).

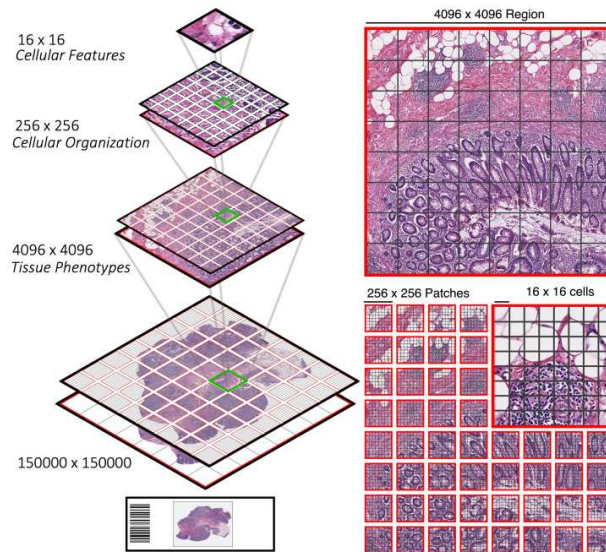


Figura 6 – Arquitetura do *Hierarchical Vision Transformer (HIPT)* (CHEN et al., 2023). O modelo processa a WSI em escalas crescentes, desde características celulares (16×16) até fenótipos de tecido (4096×4096), para formar uma representação final da lâmina inteira.

O custo da anotação por especialistas é um desafio persistente na área. O trabalho de (LÓPEZ-PÉREZ et al., 2024) aborda este problema com a criação e validação do conjunto de dados CrowdGleason. A metodologia, descrita na Figura 7, envolveu um protocolo de *crowdsourcing* onde sete patologistas em treinamento anotaram 19.077 *patches* de 1.045 WSIs. A inovação reside na forma de lidar com a variabilidade e o ruído esperados em anotações de não especialistas. Para isso, os autores empregaram processos Gaussianos, um método estatístico Bayesiano que aprende uma matriz de confusão para cada anotador, modelando sua *expertise* e permitindo uma agregação ponderada dos rótulos.

Os resultados quantitativos validaram a abordagem. Um modelo treinado apenas com os dados de *crowdsourcing* (SVGPCR) alcançou um Kappa de 0,7048. Mais relevante, um modelo híbrido (SVGPMIX), que combinou o grande volume de dados do CrowdGleason com um pequeno conjunto de dados de alta qualidade anotado por especialistas (SICAPv2), obteve a melhor *performance*, com um Kappa de 0,7814. Este resultado sugere que anotações em massa, mesmo que com ruído, podem ser utilizadas de forma eficaz para treinar modelos robustos, especialmente quando guiadas por uma pequena quantidade de dados de especialistas.

A contribuição deste trabalho é a apresentação de uma estratégia viável para escalar a criação de conjuntos de dados anotados, o que é relevante para um dos objetivos desta tese: “Construir um banco de dados curado”. Embora esta tese utilize o conjunto de dados PANDA, os achados de (LÓPEZ-PÉREZ et al., 2024) reforçam a complexidade da etapa de curadoria de dados e demonstram a utilidade de métodos probabilísticos para lidar com a incerteza dos rótulos, uma questão presente até entre especialistas.

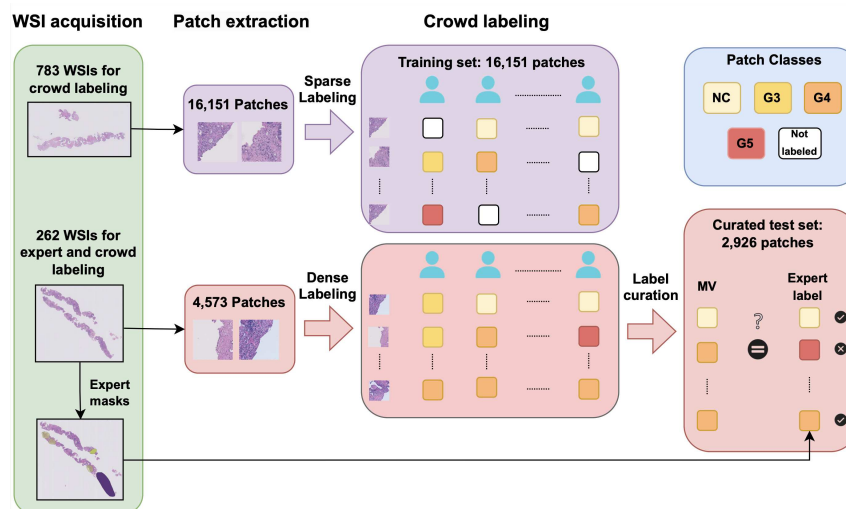


Figura 7 – Protocolo de criação do *dataset* CrowdGleason (LÓPEZ-PÉREZ et al., 2024). O fluxo ilustra a aquisição de WSIs, a extração de *patches*, o processo de rotulação distribuída (*crowd labeling*) para gerar o conjunto de treinamento, e a curadoria com especialistas para criar um conjunto de teste confiável.

Por fim, abordagens generativas também têm sido exploradas para a tarefa de seg-

mentação. (VATS et al., 2024) propõe um método baseado em *Generative Adversarial Networks* (GANs). Conforme ilustrado na Figura 8, a arquitetura GAN consiste em dois modelos que competem entre si: um Gerador, com arquitetura U-Net, que aprende a criar máscaras de segmentação a partir de uma imagem de entrada; e um Discriminador, treinado para distinguir entre as máscaras geradas e as máscaras reais do conjunto de dados (o *ground-truth*). O treinamento é um processo adversarial no qual o Gerador é otimizado para produzir resultados que “iludam” o Discriminador, forçando-o a gerar segmentações progressivamente mais realistas e precisas.

Esta abordagem generativa demonstrou eficácia para a tarefa de segmentação semântica no conjunto de dados do desafio MICCAI 2019, alcançando um *F1-score* de 0,872. Este resultado superou outras arquiteturas de segmentação, como a U-Net (0,843) e a DeepLabV3 (0,860), indicando que o aprendizado adversarial pode refinar os contornos e a qualidade das máscaras de segmentação.

Este trabalho é relevante para a tese, pois apresenta uma alternativa de vanguarda para a tarefa de segmentação. Contudo, há duas distinções pertinentes. Primeiro, a abordagem com GAN foi otimizada para a segmentação *semântica*, onde todos os *pixels* de uma mesma classe (exemplo: Gleason 4) recebem o mesmo rótulo. Isso difere da segmentação de *instâncias*, objetivo desta tese, que visa delinear cada glândula ou agrupamento glandular como um objeto distinto. Segundo, a validação em um conjunto de dados diferente (MICCAI) impede uma comparação direta de desempenho com os resultados obtidos no PANDA.

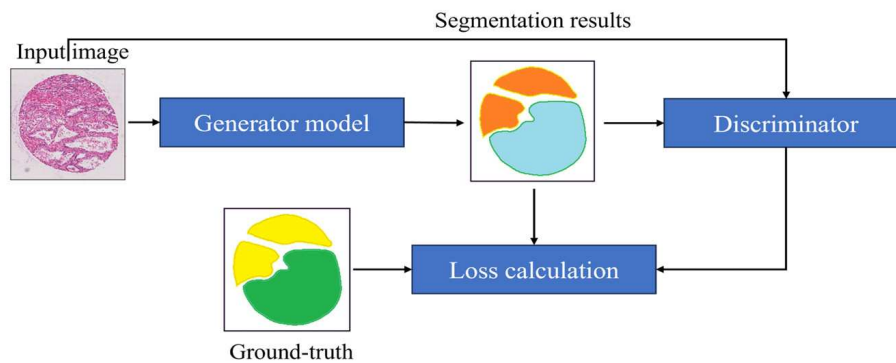


Figura 8 – Diagrama de blocos do algoritmo de segmentação baseado em GAN proposto por (VATS et al., 2024). O modelo gerador recebe a imagem de entrada e produz uma máscara de segmentação, que é avaliada pelo discriminador em comparação com a máscara real (*ground-truth*).

2.3.1 Síntese e Lacunas da Literatura

A literatura demonstra uma tendência em direção a modelos que analisam a WSI inteira. Não obstante, persistem lacunas que esta tese objetiva endereçar. A Tabela 1 sintetiza a análise dos trabalhos relacionados.

- ❑ **Classificação de *Slide* versus Segmentação de Instância:** A maioria dos trabalhos de alta *performance* concentra-se na classificação do *slide*, sem fornecer a localização e a extensão exata de cada padrão de Gleason, informação que a segmentação de instâncias pode oferecer.
- ❑ **Impacto do Pré-processamento:** Poucos trabalhos investigam o impacto de técnicas de pré-processamento especificamente desenvolvidas. Muitas abordagens tratam a rede como uma solução integral (*end-to-end*), evidenciando uma lacuna sobre como a otimização dos dados de entrada pode melhorar o desempenho.
- ❑ **Interpretabilidade:** Arquiteturas como MIL e ViTs, embora potentes, podem ser menos interpretáveis. Existe espaço para explorar como arquiteturas como o Mask R-CNN podem ser otimizadas para um balanço entre *performance* e interpretabilidade clínica.

Tabela 1 – Tabela comparativa dos principais trabalhos relacionados.

Autor(es)	Método	Base de Dados	Métrica/Resultado	Vantagens	Limitações
(ARVANITI et al., 2018)	MobileNet para classificação automática de padrões de Gleason.	641 pacientes (<i>base de teste</i> : 245)	Kappa quadrático: 75% e 71%	Alta concordância com patologistas especialistas.	Foco em microarranjos de tecido, não em WSIs completas.
(RODRIGUEZ et al., 2020)	<i>Deep learning</i> para previsão de Escore de Gleason em nível de <i>pixel</i> .	SICAPv2 (182 imagens)	Precisão: 77%	Fornecer mapas de calor com localização dos padrões.	<i>Dataset</i> pequeno, limitando a generalização do modelo.
(BULTEN et al., 2022)	IA para diagnóstico e gradação de Gleason, validada em múltiplas coortes (desafio PANDA).	PANDA e coortes externas	Kappa quadrático: 0,86	Alta generalização e desempenho de nível clínico em classificação de <i>slide</i> .	Foco na classificação do <i>slide</i> , sem segmentação detalhada dos padrões.
(MAI et al., 2024)	<i>Multiple Instance Learning</i> (MIL) com reconstrução de características espaciais e de frequência.	<i>Dataset</i> privado e PANDA	AUC: 91,69% (PANDA)	Eficiente para classificação de WSI sem anotações em nível de <i>pixel</i> .	Menor interpretabilidade; não localiza precisamente as regiões tumorais.
(CHEN et al., 2023)	<i>Vision Transformer</i> (ViT) hierárquico (HIPT) com aprendizado auto-supervisionado em múltiplas escalas.	TCGA (33 tipos de câncer)	Alta performance em subtipagem de câncer e previsão de sobrevida.	Lida com a escala de gigapixels das WSIs e aprende características em múltiplos níveis.	Requer grande poder computacional e não foi focado primariamente em segmentação.
(LÓPEZ-PÉREZ et al., 2024)	Aprendizado a partir de anotações de <i>crowdsourcing</i> com Processos Gaussianos para modelar ruído.	CrowdGleason (novo <i>dataset</i>) e SICAPv2	Demonstra que anotações “com ruído” podem treinar modelos robustos (Kappa: 0.78).	Reduz a dependência da necessidade de anotações de patologistas especialistas.	A qualidade do modelo final depende de algoritmos de agregação de rótulos.
(VATS et al., 2024)	Segmentação de padrões de Gleason utilizando uma Rede Generativa Adversarial (GAN) com gerador U-Net.	MICCAI 2019	F1-score: 0,872	Produz máscaras de segmentação precisas, superando outras arquiteturas.	Avaliado em um <i>dataset</i> diferente do PANDA, o que dificulta a comparação direta.

Fonte: O autor (2025).

2.4 Processamento Digital de Imagens em Patologia

O processamento de imagens digitais é uma área da ciência da computação com vastas aplicações, incluindo o campo da medicina. No contexto da patologia, aplica-se para extrair informações quantitativas e objetivas de imagens histológicas, visando a superação das limitações da avaliação qualitativa inerente ao diagnóstico humano (GONZALEZ; WOORDS, 2007; LITJENS et al., 2017). Esta seção detalha os conceitos de processamento que são relevantes para a análise de *Whole Slide Images* (WSIs) de câncer de próstata.

A análise de Lâminas Histológicas Inteiras - *Whole Slide Images* (WSI) impõe desafios computacionais significativos. Devido à sua resolução na ordem de gigapixels, estas imagens não podem ser processadas em uma única iteração, o que exige estratégias para a gestão de memória e tempo de processamento (MADABHUSHI; RANGANATHAN, 2019). Adicionalmente, a variabilidade na morfologia do tecido, a presença de artefatos e as inconsistências de coloração introduzidas durante a preparação das lâminas constituem fatores que comprometem a generalização dos modelos (TOTH et al., 2014; HOFMAN et al., 2021). As técnicas empregadas neste contexto incluem:

- *Geração de Patches (Tiling)*: Para mitigar o custo computacional imposto pela escala gigapixel, a WSI é particionada em milhares de *patches* (ladrilhos) menores (tipicamente 256×256 ou 512×512 pixels) (CAMPANELLA et al., 2019; CHEN et al., 2023). Esta abordagem torna a análise computacionalmente viável, permitindo que cada *patch* sirva como a unidade de entrada para a Rede Neural Convolucional (CNN). A eficiência do treinamento é otimizada pela amostragem guiada de *patches*, que filtra regiões de fundo e artefatos, e pela aplicação de técnicas como o *hard negative mining*, onde *patches* incorretamente classificados são reintroduzidos no treinamento para refinar as fronteiras de decisão do modelo (MADABHUSHI; RANGANATHAN, 2019).
- *Normalização de Coloração*: A coloração com Hematoxilina e Eosina (H&E) é suscetível a variações de intensidade e tonalidade entre diferentes laboratórios e lotes de reagentes (ARVANITI et al., 2018). Essa variação pode introduzir um viés no modelo de *deep learning*, levando-o a correlacionar o padrão cromático em detrimento das características morfológicas do tecido. A normalização de coloração é um passo de pré-processamento que ajusta o perfil de cores de uma imagem para corresponder a um padrão de referência. Metodologicamente, a normalização pode ser realizada através de técnicas de correspondência estatística de cores ou de métodos de separação de corantes no espaço de densidade óptica (OD). Recentemente, métodos baseados em *deep learning* têm sido investigados por sua capacidade de alcançar a normalização sem a dependência explícita de uma imagem de referência, podendo apresentar maior robustez e adaptabilidade (CHEN et al., 2024).

❑ *Operações Morfológicas e Segmentação de Artefatos:* Técnicas clássicas de processamento de imagem, como operações morfológicas, são aplicadas para refinar as máscaras de segmentação utilizadas no treinamento (GONZALEZ; WOORDS, 2007). Operações de abertura (erosão seguida de dilatação) removem pequenos objetos e ruídos (artefatos de digitalização). Operações de fechamento (dilatação seguida de erosão) preenchem pequenos orifícios dentro de regiões de interesse. A aplicação dessas operações resulta em anotações mais limpas. Adicionalmente, a segmentação do tecido em relação ao fundo da lâmina e a remoção de artefatos são realizadas por meio de limiarização (e.g., método de Otsu) ou por algoritmos de segmentação baseados em *deep learning* dedicados ao controle de qualidade (HOFMAN et al., 2021).

Essas etapas de pré-processamento, embora baseadas em conceitos clássicos, são aplicadas de forma direcionada para resolver os desafios de escala e variabilidade das imagens de patologia digital, constituindo a base sobre a qual os modelos de aprendizado profundo podem atuar com maior eficácia.

2.5 Redes Neurais Convolucionais

Uma *Rede Neural Convolutiva* (CNN) é uma arquitetura de *deep learning* inspirada na organização hierárquica do córtex visual humano, que é a arquitetura mais frequentemente empregada para a maioria das tarefas de visão computacional (LECUN; KAVUKCUOGLU; FARABET, 2010). Sua arquitetura é projetada para aprender representações hierárquicas de características a partir de dados brutos de *pixels*, o que a torna adequada para a análise de imagens histológicas.

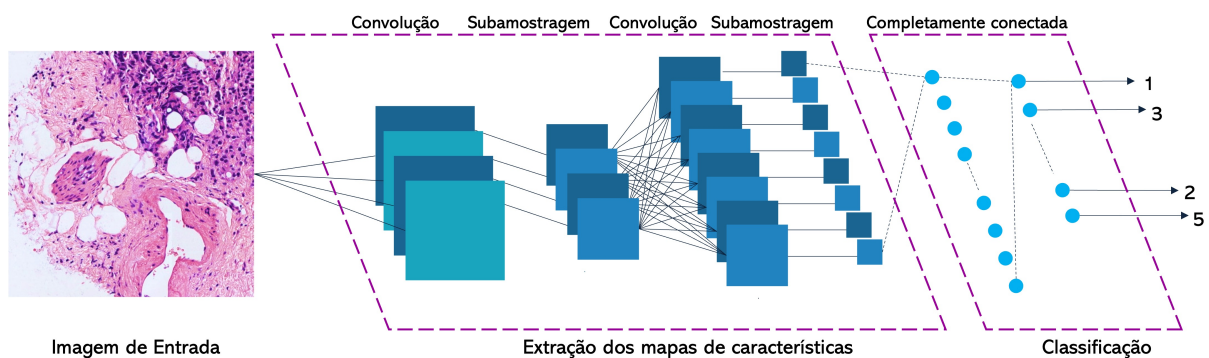


Figura 9 – Exemplo de estrutura de CNN. A imagem de entrada passa por camadas de convolução e subamostragem para extração de mapas de características, que são então utilizados por uma camada totalmente conectada para a classificação final.

A estrutura de uma CNN, ilustrada na Figura 9, é composta por:

1. *Camada de Convolução*: Onde filtros (*kernels*) aprendem a detectar padrões espaciais, como bordas, texturas e formas. Nas camadas iniciais, os filtros detectam características simples. Em camadas mais profundas, combinam essas características para identificar estruturas complexas, como a morfologia de glândulas cancerosas (KRIZHEVSKY; SUTSKEVER; HINTON, 2012). As *Deep Convolutional Neural Networks* (DCNNs) são utilizadas na identificação de detalhes finos na análise de lesões prostáticas.
2. *Camada de Subamostragem (Pooling)*: Reduz a dimensão espacial dos mapas de características, o que diminui a carga computacional e confere à representação maior invariância a pequenas translações na imagem de entrada (BOUREAU; PONCE; LECUN, 2010).
3. *Camada Totalmente Conectada*: Após a extração de características, estas são transformadas em um vetor e processadas por camadas densas para realizar a tarefa final, como a classificação do padrão de Gleason.

A arquitetura *ResNet (Residual Neural Network)* (HE et al., 2016), utilizada como *backbone* nesta pesquisa, possibilitou o treinamento de redes mais profundas. Introduziu as “conexões residuais” (*skip connections*), que permitem que o gradiente flua com maior facilidade através da rede durante o treinamento, o que contribui para mitigar o problema do desaparecimento do gradiente. Outras arquiteturas (*VGG*, *Inception* e *DenseNet*) são igualmente utilizadas, frequentemente empregando a técnica de *Transfer Learning* para otimizar o desempenho em *datasets* médicos limitados (COURTY et al., 2023). Adicionalmente, pesquisas exploram arquiteturas especializadas como as *Graph Convolutional Networks* (GCN) e modelos de fusão de características para refinar a análise estrutural do câncer de próstata (ALI et al., 2021; ZHU et al., 2022).

O treinamento de uma Rede Neural Convolucional envolve a minimização de uma função de perda (erro) através do ajuste iterativo dos pesos sinápticos. O algoritmo para este processo é o *Gradiente Descendente*. Este método de otimização calcula o gradiente da função de perda em relação aos parâmetros da rede e atualiza os pesos na direção oposta ao gradiente, buscando o mínimo global ou local da função de erro (GOODFELLOW; BENGIO; COURVILLE, 2016). Em arquiteturas modernas, utilizam-se variações estocásticas deste algoritmo, como o *AdamW* utilizado neste trabalho, que adaptam a taxa de aprendizado para acelerar a convergência.

Para viabilizar o cálculo desses gradientes em redes profundas com múltiplas camadas, emprega-se o algoritmo de Retropropagação (*Backpropagation*). Baseado na regra da cadeia do cálculo diferencial, o *Backpropagation* propaga o erro da camada de saída de volta para as camadas iniciais, permitindo calcular a contribuição exata (derivada parcial) de cada peso individual para o erro total da rede (BISHOP, 2006). É essa mecânica que

permite à rede “aprender” quais características da imagem (como bordas ou texturas de glândulas) são relevantes para a classificação correta do Escore de Gleason.

2.6 Segmentação de Instâncias com *Mask R-CNN* e Arquiteturas Avançadas

No contexto da graduação de Gleason, uma única biópsia pode conter múltiplas regiões com diferentes padrões (exemplo: áreas de Gleason 3 e Gleason 4) (EPSTEIN et al., 2016). Tarefas de classificação ou segmentação semântica são insuficientes para distinguir objetos individuais. A segmentação de instâncias é a tarefa computacional utilizada neste contexto, pois identifica e delinea cada objeto (ou instância) de forma independente (HE et al., 2017; ZHOU et al., 2019).

O *Mask R-CNN* (HE et al., 2017) é uma arquitetura de referência para esta tarefa, sendo uma extensão do *Faster R-CNN* (REN et al., 2015) com um ramo adicional para a geração de máscara. Sua abordagem de duas etapas é aplicada para problemas complexos:

1. *Geração de Propostas de Região*: Uma sub-rede denominada *Region Proposal Network* (RPN) analisa o mapa de características (frequentemente aprimorado com *Feature Pyramidal Network - FPN* para melhor multiescala (TIWARI et al., 2024)) e propõe regiões candidatas (*Regions of Interest - ROIs*) que provavelmente contêm um objeto.
2. *Classificação, Refinamento e Geração de Máscara*: Para cada ROI, três “cabeças” da rede atuam em paralelo: uma classifica o objeto (exemplo: “Gleason 3”), outra refina as coordenadas da caixa delimitadora, e uma terceira gera uma máscara de segmentação em nível de *pixel* para a instância.

Uma inovação técnica do Mask R-CNN é a camada *ROIAlign* (HE et al., 2017), que resolve um problema de desalinhamento de *pixels* (quantização) presente em arquiteturas anteriores, utilizando interpolação bilinear. Este mecanismo de alinhamento é importante para que a máscara de segmentação seja gerada com alta fidelidade espacial (HE et al., 2017). O *Detectron2*, o *framework* utilizado nesta tese, é uma implementação eficiente e modular desta arquitetura (WU et al., 2019b). O Mask R-CNN é reportado por apresentar um desempenho adequado na segmentação de instâncias que se sobrepõem, como células e glândulas densamente agrupadas (RETTENBERGER; SAGER; LELLMANN, 2023). Pesquisas atuais investigam otimizações do Mask R-CNN, incluindo o uso de *backbones* mais avançados (e.g., *Res2Net*) e modificações arquiteturais para aumentar a precisão nas bordas das instâncias (HUANG et al., 2023).

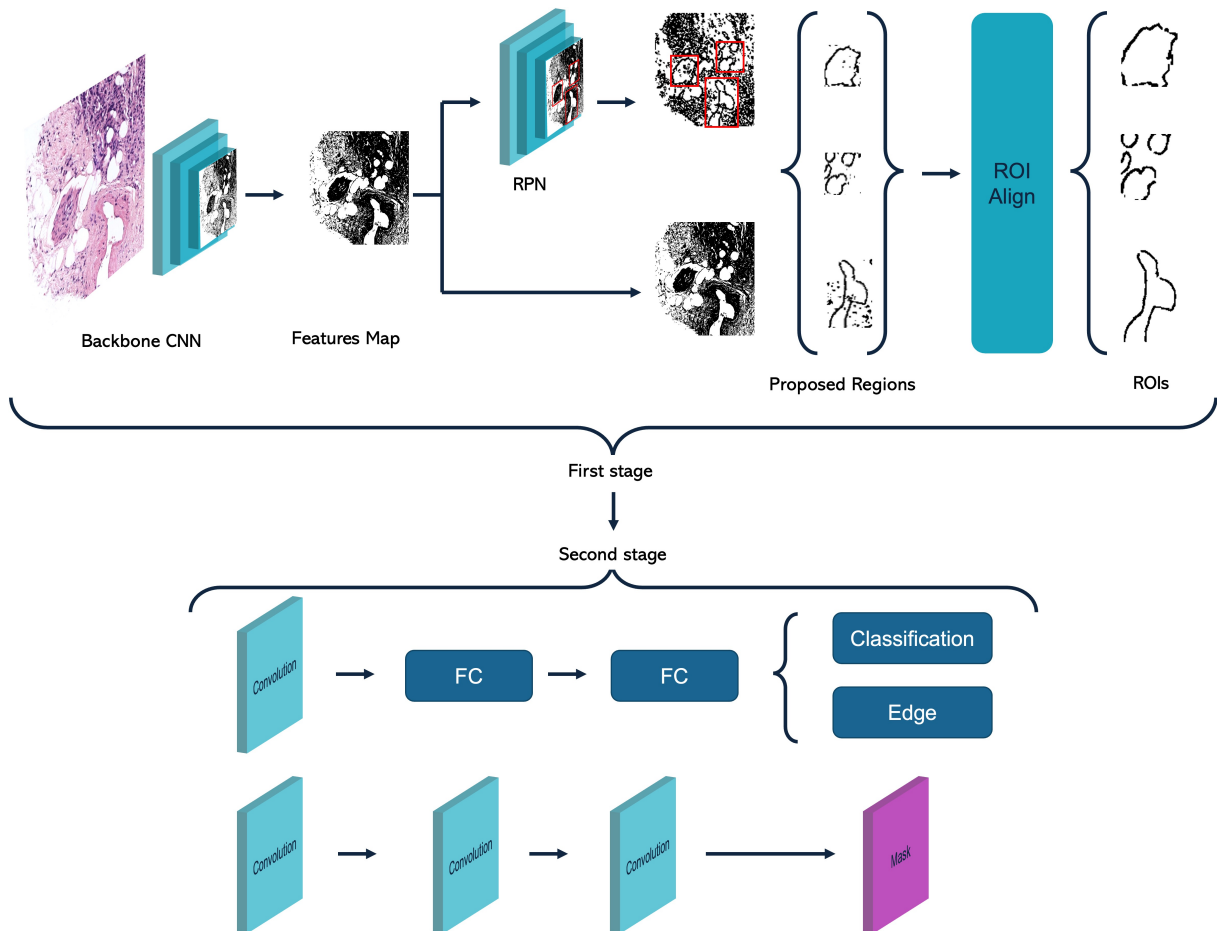


Figura 10 – Estrutura geral do Mask R-CNN. A imagem passa por um *backbone* (CNN) para extrair características. A RPN propõe regiões (ROIs), que são alinhadas pela camada ROIAlign. Em seguida, as cabeças da rede realizam a classificação, o refinamento da caixa delimitadora e a geração da máscara de segmentação para cada instância. Figura adaptada de (CHIAO et al., 2019).

Metodologia de Pesquisa

A análise computacional de imagens médicas tem se consolidado como uma área de pesquisa relevante, com a aplicação de algoritmos de Inteligência Artificial (IA) para auxiliar em tarefas de diagnóstico e prognóstico (GIANSANTI, 2025; RIBEIRO et al., 2025). Dentro deste campo, abordagens baseadas em *deep learning* têm sido amplamente investigadas para a classificação e segmentação de padrões em imagens patológicas, demonstrando capacidade de aprender hierarquias espaciais de características diretamente dos dados (HE; LUAN; HU, 2025). O presente estudo se insere neste contexto, propondo uma metodologia para a análise de imagens histológicas de câncer de próstata.

O trabalho foi desenvolvido utilizando o conjunto de dados *Prostate Cancer Grade Assessment* (PANDA). A escolha deste acervo justifica-se por ser um recurso público de grande escala, amplamente utilizado como referência (*benchmark*) na área. Ele é composto por imagens de lâminas inteiras (*Whole-Slide Images* - WSI) em formato *Tagged Image File Format* (TIFF) de alta resolução, acompanhadas de anotações do escore de Gleason, o que permite o treinamento supervisionado e a avaliação comparativa dos modelos propostos.

Dada a alta dimensionalidade das WSIs, uma etapa de pré-processamento dos dados é mandatória. Inicialmente, aplicam-se operações de morfologia matemática com o objetivo de atenuar ruídos e artefatos de coloração que poderiam impactar negativamente o desempenho do modelo (SILVA; NASCIMENTO; MEDEIROS, 2022). Em seguida, são geradas máscaras de tecido para isolar a área de interesse do fundo da lâmina, otimizando o processamento computacional.

Posteriormente, as regiões de tecido são divididas em múltiplos *patches* (recortes) de dimensões uniformes. Esta abordagem, conhecida como *tiling*, é um procedimento padrão no processamento de WSIs, pois permite que as imagens sejam tratadas por unidades de processamento gráfico (GPUs) com memória limitada e, simultaneamente, aumenta a quantidade de amostras disponíveis para o treinamento da rede neural.

Para a tarefa de identificação e delineamento das estruturas histológicas, foi selecionada a arquitetura *Mask R-CNN*, que utiliza a ResNet-50 como rede de extração de características (*backbone*). A ResNet-50 é uma arquitetura convolucional consolidada, reconhecida por

sua capacidade de aprender representações hierárquicas robustas de dados visuais. A implementação foi realizada utilizando o *framework* Detectron2, uma biblioteca de software que oferece componentes modulares e eficientes para tarefas de detecção e segmentação de objetos.

A escolha do *Mask R-CNN* fundamenta-se na natureza do problema, que exige a segmentação de instância. Esta é uma tarefa de visão computacional que vai além da detecção de objetos (que fornece apenas caixas delimitadoras) e da segmentação semântica (que classifica pixels, mas não distingue instâncias adjacentes da mesma classe) (HE et al., 2017). A segmentação de instância fornece uma máscara de segmentação distinta para cada objeto individual detectado.

A aplicação desta técnica em histopatologia é justificada pela necessidade de analisar individualmente as glândulas prostáticas. A morfologia e a disposição dessas estruturas são critérios essenciais no sistema de classificação de Gleason. A capacidade de isolar cada glândula permite a extração de características quantitativas específicas por instância, como área, perímetro e fatores de forma. A literatura recente corrobora a adequação de abordagens de *deep learning* para tarefas de segmentação em nível de objeto em imagens histopatológicas, como na segmentação de núcleos, apoiando a metodologia aqui adotada (BASU et al., 2024).

Por fim, foi desenvolvida uma interface gráfica de usuário (GUI) como ferramenta de apoio. Sua finalidade é permitir a execução do modelo treinado sobre novas imagens (inferência) e facilitar a inspeção visual qualitativa dos resultados da segmentação, servindo como um meio para a verificação e interpretação dos resultados gerados pelo modelo. Os passos sequenciais desta metodologia estão representados na Figura 11.

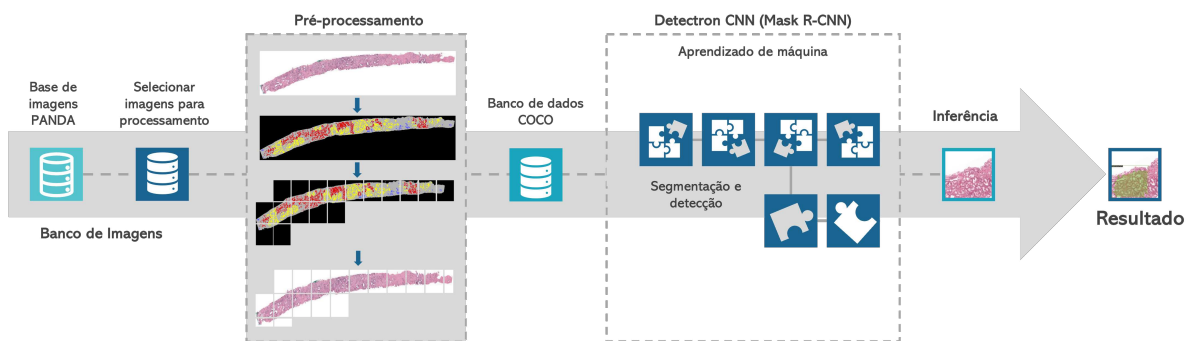


Figura 11 – Passos da metodologia proposta.

3.1 Construção do Banco de Imagens

O treinamento de redes neurais convolucionais para tarefas de supervisão requer um conjunto de dados extenso e com anotações de alta qualidade. Este processo metodológico

inicial envolve a seleção de um banco de dados de imagens apropriado para o estudo e a curadoria dos dados que servirão de base para o treinamento e validação do modelo (WANG et al., 2020).

Para este trabalho, foi selecionado o conjunto de dados PANDA, que consiste em imagens de biópsias de câncer de próstata. O conjunto de treinamento compreende aproximadamente 11.000 imagens digitalizadas de lâminas inteiras (*Whole-Slide Images* - WSI) coradas com hematoxilina e eosina (H&E), originárias de dois centros de referência: o *Radboud University Medical Center* e o *Karolinska Institute*. As imagens variam em dimensões, com médias de 8.192 pixels de largura por 22.528 pixels de altura, profundidade de cor de 24 bits e formato TIFF, totalizando 411,9 gigabytes de armazenamento (KAGGLE, 2023).

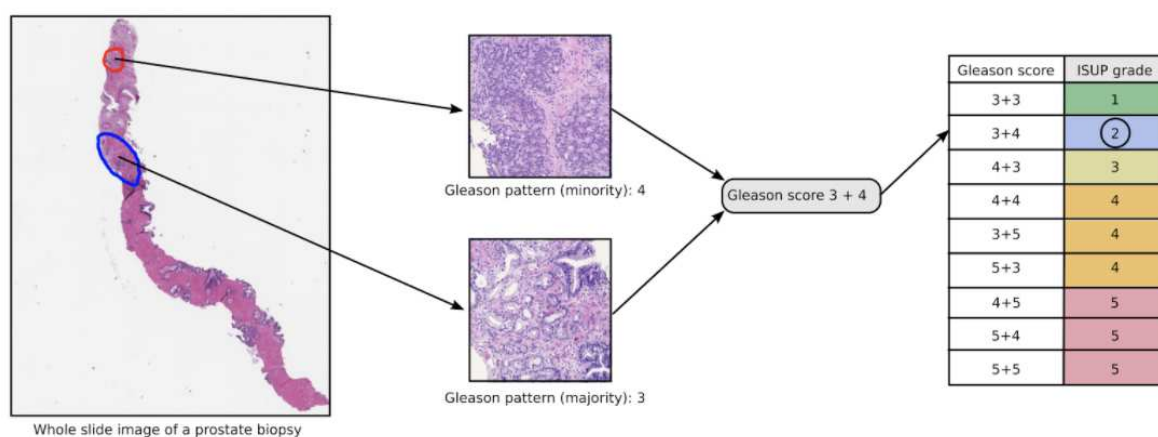


Figura 12 – Exemplo do processo de graduação de Gleason para uma biópsia contendo câncer de próstata. Os padrões de crescimento canceroso mais comum (contorno azul, padrão Gleason 3) e o segundo mais comum (contorno vermelho, padrão Gleason 4) presentes na biópsia determinam a pontuação Gleason (3+4 para esta biópsia), que por sua vez é convertida em um grau ISUP (2 para esta biópsia) seguindo as diretrizes da Sociedade Internacional de Patologia Urológica. Biópsias que não contêm câncer são representadas por um grau ISUP de 0 no conjunto de dados apresentado (KAGGLE, 2023).

A Figura 12 ilustra o processo de graduação de Gleason para uma biópsia com câncer de próstata, destacando os padrões de crescimento mais comuns que determinam a pontuação Gleason e, consequentemente, o grau ISUP. O *Radboud University Medical Center* e o *Karolinska Institute*, em colaboração com a *Tampere University*, coordenaram a criação deste conjunto de imagens. O grupo de pesquisa do Radboud University Medical Center foca no desenvolvimento de algoritmos computacionais para suporte clínico, enquanto o departamento do Karolinska Institute mantém um grupo interdisciplinar dedicado ao aprimoramento do diagnóstico e tratamento do câncer de próstata (KAGGLE, 2023).

Este conjunto de dados difere de outros bancos por incluir biópsias completas obtidas de testes multicêntricos e analisadas por uropatologistas especialistas (KAGGLE, 2023).

Essa característica permite uma representação mais abrangente das variações histológicas observadas na prática clínica, o que é fundamental para o treinamento de modelos que possam generalizar para dados do mundo real.

O acervo é dividido em dois subconjuntos com estratégias de anotação distintas. O primeiro, denominado Radboud, contém imagens histológicas de glândulas prostáticas rotuladas com anotações individuais e em nível de pixel do escore de Gleason. A classificação detalhada inclui: 1 para estroma; 2 para epitélio benigno; 3 para epitélio canceroso (Gleason 3); 4 para epitélio canceroso (Gleason 4); e 5 para epitélio canceroso (Gleason 5). A presença de múltiplas anotações em uma única imagem permite a identificação precisa de áreas com diferentes graus de agressividade (BULTEN et al., 2022).

O segundo subconjunto, Karolinska, adota uma abordagem de rotulação mais ampla, categorizando as regiões da imagem como: 0 para fundo (não tecido) ou desconhecido; 1 para tecido benigno (estroma e epitélio combinados); e 2 para tecido canceroso. Diferentemente do subconjunto Radboud, as imagens do Karolinska não apresentam máscaras de segmentação detalhadas com base nos escores de Gleason. Os escores são associados às amostras como um todo, indicando a malignidade geral da lâmina, mas sem especificar a localização exata dos diferentes padrões tumorais (BULTEN et al., 2022).

Os dados foram organizados na Tabela 2, que apresenta a distribuição desses escores para as duas instituições, representando o número de casos observados para cada combinação de escore de Gleason.

Tabela 2 – Distribuição dos escores de Gleason por Instituto.

Escore de Gleason	Karolinska	Radboud
0+0	1925	0
3+3	1814	852
3+4	667	675
3+5	13	67
4+3	318	925
4+4	466	660
4+5	208	641
5+3	2	41
5+4	27	221
5+5	16	111
Negativo	0	967

A análise dos dados revela tendências distintas. No Instituto Karolinska, predominam casos com escores de Gleason baixos (0+0 e 3+3), indicando uma proporção maior de cânceres de próstata menos agressivos. No Centro Médico Radboud, a distribuição é mais uniforme, com uma proporção significativa de casos com escores altos (4+3, 4+4, 4+5 e 5+4), sugerindo uma incidência maior de cânceres agressivos ou de estágio avançado. A presença de casos negativos em ambas as instituições (Negativo no Radboud e 0+0 no Karolinska) indica biópsias sem evidência de câncer.

A escolha das imagens do subconjunto Radboud para este estudo decorre, portanto, da disponibilidade de máscaras de segmentação correspondentes a cada imagem, contendo a classificação detalhada do escore de Gleason e a delimitação das regiões indicativas de câncer. Esse processo corresponde às etapas iniciais da metodologia adotada, representada na Figura 11.

A construção de bancos de dados anotados constitui um componente central na análise de imagens médicas, fornecendo rótulos precisos para o treinamento supervisionado de modelos de *deep learning*. A literatura recente indica que *datasets* anotados aprimoram a capacidade de generalização desses modelos, reduzindo vieses e elevando a precisão em tarefas como segmentação e classificação.

A relevância de *datasets* como o PANDA reside na sua capacidade de representar as variações encontradas na prática clínica, por serem provenientes de múltiplos centros. A literatura enfatiza que bancos de dados multicêntricos reduzem o viés institucional e melhoram a transferibilidade dos modelos para cenários clínicos diversificados.

Adicionalmente, a integração de tais *datasets* com técnicas de pré-processamento, como operações morfológicas, prepara os dados para um treinamento eficiente. Embora o foco desta seção seja na construção do banco, a literatura sobre morfologia matemática em imagens médicas, como em (SILVA; NASCIMENTO; MEDEIROS, 2022), destaca como anotações precisas combinadas com refinamentos morfológicos podem melhorar a segmentação de estruturas, uma lógica aplicável de forma análoga aos tecidos prostáticos.

A construção do *dataset* final não se limitou à extração de dados, mas envolveu um processo de curadoria manual e verificação de integridade das máscaras que perdurou por aproximadamente seis meses. Esta etapa foi crítica para remover anotações ruidosas inerentes a bases públicas e garantir que apenas patches com concordância clara de rotulação fossem apresentados à rede, para mitigar a propagação de erros durante o treinamento.

3.2 Pré-processamento

A etapa subsequente envolve o pré-processamento do conjunto de dados, que prepara as imagens para análise e segmentação. Este processo inclui operações automáticas de redimensionamento, normalização e filtragem. O redimensionamento ajusta as dimensões das imagens para compatibilidade com a rede neural convolucional, enquanto a normalização padroniza os valores dos pixels em uma faixa específica, facilitando o treinamento do modelo (HE et al., 2017). A filtragem remove ruídos ou realça características relevantes para a detecção de objetos, além de abordar o equilíbrio do conjunto de dados.

Estudos indicam que o desequilíbrio entre classes representa um desafio em tarefas de classificação, especialmente quando a classe minoritária possui poucas instâncias (HE; GARCIA, 2009; JOHNSON; KHOSHGOFTAAR, 2023; JAFARIGOLA; TRAFALISA,

2023). Nesse contexto, o classificador tende a priorizar a classe majoritária, resultando em viés nos resultados (HE; GARCIA, 2009). Pesquisas corroboram que o desequilíbrio afeta a classificação da classe minoritária sem necessariamente alterar a acurácia global, tornando-a um indicador potencialmente enganoso; por isso recomenda-se o uso de métricas mais adequadas como F1-score, IoU e correlatas (CHICCO; JURMAN, 2020; POWERS, 2011). A magnitude desse desafio varia conforme o tipo de dados, o tamanho do conjunto e a distribuição entre classes (JOHNSON; KHOSHGOFTAAR, 2023; HE; GARCIA, 2009; JAFARIGOLA; TRAFALISA, 2023).

O equilíbrio do banco de dados de imagens é adotado para o treinamento do modelo. A utilização de quantidades iguais de imagens por categoria busca reduzir o viés em direção às classes mais representadas. Em tarefas de detecção de objetos, diferenças na quantidade de exemplos por classe podem induzir viés para classes frequentes, afetando o desempenho em classes menos representadas.

Após a seleção do subconjunto de imagens, as imagens do banco de dados PANDA foram submetidas a pré-processamento para fornecer dados adequados ao modelo. Esta fase compreende as seguintes etapas:

1. **Aplicação de Operações Morfológicas:** Operações de abertura morfológica removem ruídos e objetos com área inferior a 100 pixels, para isso foi utilizado um kernel estruturante elíptico de tamanho 3×3 , eguida pela remoção de objetos conectados com área $A < 100$ pixels. O fechamento morfológico utilizou o mesmo kernel para preencher lacunas intra-glandulares, preservando a morfologia do Grau 5.
2. **Geração de Máscaras:** Máscaras são criadas com segmentação das anotações do *dataset*, atribuindo cores específicas: fundo em preto (0, 0, 0); estroma em cinza (180, 180, 180); epitélio saudável em azul (0, 0, 255); epitélio canceroso (Gleason 3) em amarelo (255, 255, 0); epitélio canceroso (Gleason 4) em vermelho (223, 0, 0); e epitélio canceroso (Gleason 5) em verde (0, 255, 0).
3. **Identificação de Regiões:** As máscaras são utilizadas para localizar regiões de interesse nas imagens.
4. **Divisão em *Patches*:** Imagens e máscaras são divididas em *patches* de 512×512 pixels, retendo apenas aqueles com objetos relevantes.
5. **Balanceamento do Conjunto:** O conjunto é balanceado por subamostragem das classes majoritárias, utilizando a contagem da classe minoritária (Grau 5, 4.794 ocorrências) como referência, resultando em 23.970 anotações para treinamento.
6. **Aumento de dados:** Técnicas de aumento de dados são aplicadas em tempo de execução durante o treinamento, incluindo rotações aleatórias (0-90 graus),

espelhamento horizontal e vertical, e variações de brilho e contraste para aprimorar a generalização.

Ressalta-se que as técnicas de *data augmentation* descritas (rotação, espelhamento, variação de brilho) foram aplicadas apenas no subconjunto de treinamento. O conjunto de validação e teste permaneceu inalterado (sem aumento artificial) para garantir uma avaliação fidedigna da capacidade de generalização do modelo em dados reais e evitar o vazamento de dados (*data leakage*).

Antes da divisão e padronização das imagens, operações morfológicas foram aplicadas para remover ruídos e imperfeições, conforme a etapa 1. A Figura 13 ilustra este processo, com as áreas afetadas destacadas. As regiões removidas aparecem em rosa, indicando objetos excluídos do cálculo do escore de Gleason. As demais categorias seguem as cores da etapa 2.

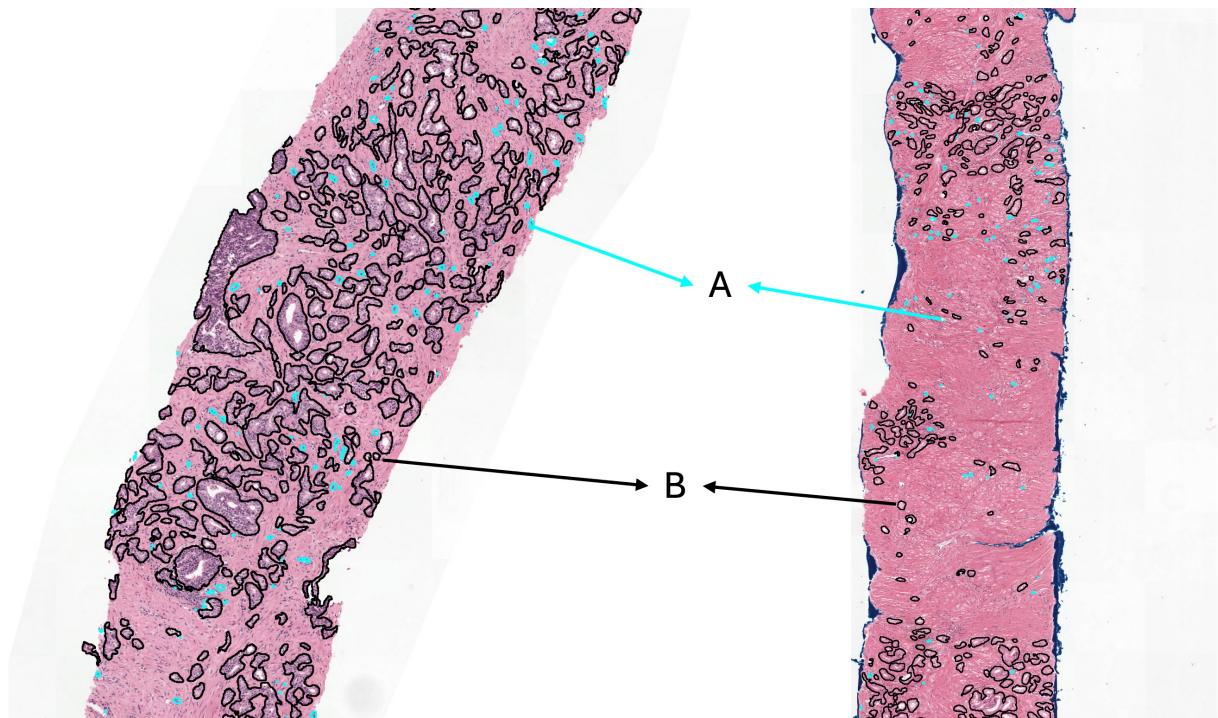


Figura 13 – A imagem apresenta uma seção histológica de tecido em que foram realizadas operações morfológicas. As áreas indicadas pela borda de cor azul claro (A) representam as regiões em que as operações morfológicas foram aplicadas, resultando na remoção de determinadas estruturas. As áreas contornadas em preto (B) representam as regiões que foram mantidas e consideradas para o cálculo do escore de Gleason.

As operações morfológicas foram utilizadas para refinar as máscaras de segmentação, removendo pequenos objetos e definindo contornos. A remoção de objetos com área inferior a um limiar mínimo utiliza filtros para eliminar componentes irrelevantes, aprimorando a precisão ao remover artefatos (STENNING et al., 2013). Após a remoção, a máscara é

reprocessada para identificar as regiões conectadas restantes. Em seguida, os contornos são extraídos e desenhados sobre a imagem original para permitir a visualização das regiões segmentadas, como visto na Figura 13.

Os processos como abertura e fechamento, refinam as máscaras, removendo ruídos e preenchendo lacunas. O trabalho de (SILVA; NASCIMENTO; MEDEIROS, 2022) demonstra a aplicação consolidada desta técnica na segmentação de imagens médicas para identificar estruturas como o miocárdio com CNNs, preservando a integridade morfológica. Indo além, estudos recentes propõem a integração de operadores morfológicos como camadas aprendíveis (*learnable*) em arquiteturas de *deep learning*, adaptando-os aos dados e melhorando a precisão em tarefas como a delimitação de estruturas cardíacas.

O *Data augmentation*, aplicado em tempo real, aumenta a variabilidade do conjunto de dados. Conforme discutido em teses sobre análise de imagens médicas com escassez de dados, essa técnica é utilizada para mitigar o *overfitting* e melhorar a capacidade de generalização do modelo (FAN, 2025). Em *frameworks* para detecção de anomalias, etapas de pré-processamento como normalização e balanceamento também são consideradas indispensáveis para garantir a robustez do sistema.

A Figura 14 apresenta a saída do pré-processamento: A) imagem original; B) máscara com anotações do *dataset*; C) máscara segmentada em regiões com objetos; D) imagem original segmentada de acordo com as regiões da máscara.

Após o processamento, os conjuntos de dados derivados do PANDA foram estruturados no formato *Common Objects in Context* (COCO), que é amplamente utilizado no treinamento de redes neurais para segmentação e detecção de objetos por fornecer uma estrutura padronizada para imagens, objetos e anotações (LIN et al., 2014; LIN, 2015). Conforme a Figura 11, a construção do banco de dados no formato COCO segue os critérios estabelecidos no pré-processamento, criando categorias para o escore de Gleason e definindo as regiões e máscaras que serão utilizadas no treinamento da CNN.

3.3 Treinamento da Rede Neural

A arquitetura Mask R-CNN foi empregada para a segmentação de instâncias em biópsias de câncer de próstata. A Mask R-CNN constitui uma extensão da Faster R-CNN, com a adição de uma ramificação dedicada à segmentação de máscaras de objetos. Essa arquitetura realiza a detecção e a segmentação de instâncias em uma única etapa (HE et al., 2017). A segmentação de instância envolve a aplicação de redes neurais convolucionais, como a Mask R-CNN, para o treinamento, identificação e segmentação de objetos (HE et al., 2017). O modelo foi treinado com o banco de dados COCO, derivado do conjunto de dados PANDA.

O treinamento ocorreu por meio do *framework* Detectron2 (WU et al., 2019a). A avaliação do modelo utilizou métricas padrão de detecção de objetos: a Intersecção sobre

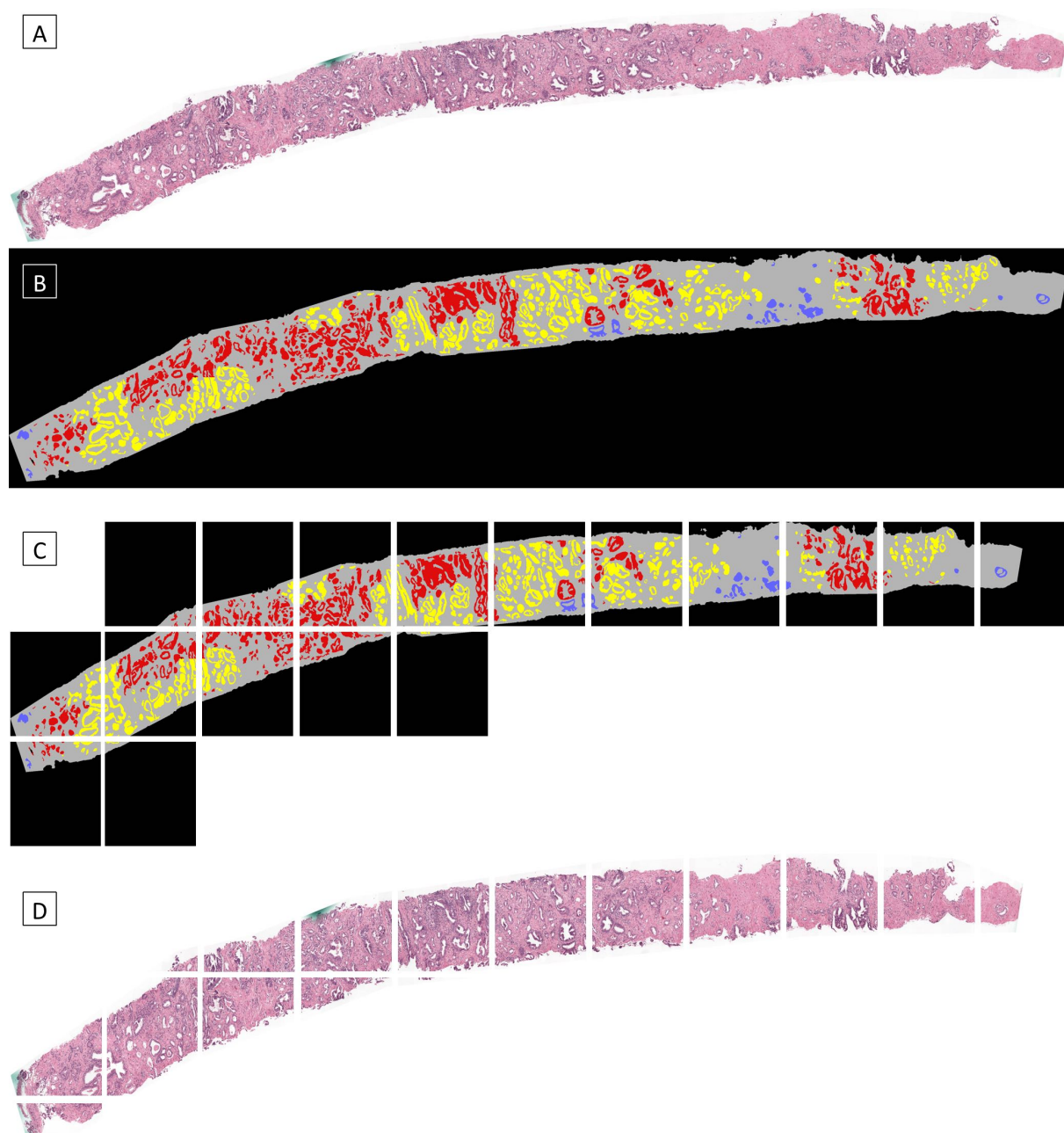


Figura 14 – Exemplo de pré-processamento: A) Imagem original; B) Máscara com anotações do *dataset*; C) Máscara com as regiões delimitadoras para segmentar em *patches*; e D) Imagem segmentada em *patches* por regiões determinadas por C.

União (IoU), que mede a precisão da localização geométrica das caixas delimitadoras, e a Precisão Média (AP), que resume o desempenho geral do classificador em termos de precisão e *recall*. Nesse contexto, o *framework* Detectron2 foi selecionado, com o algoritmo Mask R-CNN. A rede neural recebe imagens pré-processadas para aprender a identificar e segmentar estruturas histológicas.

A definição e otimização de hiperparâmetros são etapas que impactam diretamente a qualidade e a convergência do modelo treinado. A busca por uma combinação ótima, envolvendo parâmetros como taxa de aprendizado, tamanho do lote (*batch size*) e a própria arquitetura da rede, é um desafio complexo, dado o vasto espaço de busca e a independência *a priori* entre esses parâmetros (SILVA; CAMATA, 2024). A literatura apresenta diversas abordagens para essa otimização, desde métodos de busca sistemática até estratégias mais avançadas, como algoritmos genéticos (URBINATE, 2025) e otimização Bayesiana (SILVA; CAMATA, 2024).

No contexto do treinamento, define-se *Época* como um ciclo completo onde todo o conjunto de dados de treinamento passa pela rede neural. Já a *Iteração* refere-se ao processamento de um único lote (*batch*) de imagens e a atualização dos pesos (SILVA; CAMATA, 2024). O ajuste dos pesos foi realizado através do algoritmo de retropropagação (*backpropagation*). Este método calcula o gradiente da função de perda em relação a cada peso da rede, propagando o erro da saída para a entrada (*chain rule*). Associado ao otimizador AdamW, o *backpropagation* permite que o modelo ajuste seus parâmetros iterativamente na direção oposta ao gradiente do erro (gradiente descendente), minimizando a discrepância entre a máscara predita e o *ground-truth* (SILVA; CAMATA, 2024).

Neste trabalho, os hiperparâmetros foram definidos com base em práticas estabelecidas na literatura de segmentação de imagens médicas, visando um equilíbrio entre desempenho e eficiência computacional. Os parâmetros definidos para o treinamento incluem os seguintes:

1. **Arquitetura da Rede: ResNet-50.** Essa rede neural compreende 50 camadas e facilita a detecção e classificação de características em imagens. O ResNet-50 incorpora blocos residuais para mitigar a degradação do gradiente em redes profundas.
2. **Tamanho do lote: 12.** O tamanho do lote indica o número de imagens processadas simultaneamente antes da atualização dos parâmetros da rede. Esse valor de 12 foi adotado com base em considerações de uso de memória e estabilidade do gradiente.
3. **Taxa de aprendizado de 0,0001.** Esse parâmetro determina a magnitude dos ajustes nos parâmetros do modelo durante o treinamento. Uma taxa de 0,0001 permite ajustes graduais.
4. **Iterações de aquecimento: 100.** Nas primeiras 1000 iterações, a taxa de aprendizado é incrementada progressivamente para uma adaptação inicial.

5. **Número máximo de iterações de treinamento: 70.000.** Esse número estabelece o total de iterações para o treinamento do modelo.
6. **Redução de pontos e fator de redução da taxa de aprendizado:** A taxa de aprendizado é reduzida em 70% e 90% das iterações, com cada redução correspondendo a um fator de 0,5.
7. **Tamanho do lote para Regiões de Interesse (ROIs): 24.** Para as ROIs, o tamanho do lote de 24 define o número de regiões processadas por vez.
8. **Frequência de registro de métricas:** Métricas são registradas a cada iteração para monitorar o desempenho.

A otimização de hiperparâmetros no treinamento da rede neural seguiu estratégias documentadas na literatura. Por exemplo, abordagens adaptativas para taxas de aprendizado em segmentação de imagens médicas foram exploradas, como o otimizador *Cyclic Learning/Momentum Rate* (CLMR) proposto por Mortazi et al. (2023), que apoia a seleção de *batch sizes* e taxas de aprendizado para melhorar o desempenho (MORTAZI et al., 2023).

O processo de treinamento resultou em arquivos com pesos e categorias, utilizados para inferência. Esses arquivos permitem a detecção e segmentação de objetos em novas imagens.

3.3.1 Aplicação para Análise de Imagens

Após o treinamento, a CNN pode ser aplicada para inferir rótulos de objetos em novas imagens. A inferência consiste em fornecer uma imagem à rede treinada e interpretar as saídas, que incluem caixas delimitadoras e máscaras de segmentação. Esse processo facilita a aplicação do modelo em contextos práticos (WU et al., 2019b).

A biblioteca Tinker foi empregada para desenvolver uma GUI e realizar a inferência. Essa interface permite a seleção de imagens, ajustes de exibição e visualização das detecções (Figura 15). A interface carrega os pesos treinados pelo *framework* Detectron2 e apresenta as seguintes características:

- ❑ **Canvas:** Área principal para exibição de imagens.
- ❑ **Legenda:** Representação das cores das máscaras associadas ao score de Gleason.
- ❑ **Controles de exibição:** Permitem ajustes nas informações exibidas, como a seleção de máscaras específicas para escores de Gleason.
- ❑ **Botões de controle:** Incluem funções de Zoom, Rotacionar, Salvar, Analisar e Abrir Imagem.

- ❑ **Indicador de duração do processamento:** Mostra o tempo de processamento da imagem.

O software (Figura 15) possibilita o carregamento de imagens sem fragmentação prévia. A funcionalidade principal envolve a identificação e segmentação de regiões com câncer de próstata, permitindo a determinação do escore de Gleason a partir do slide.

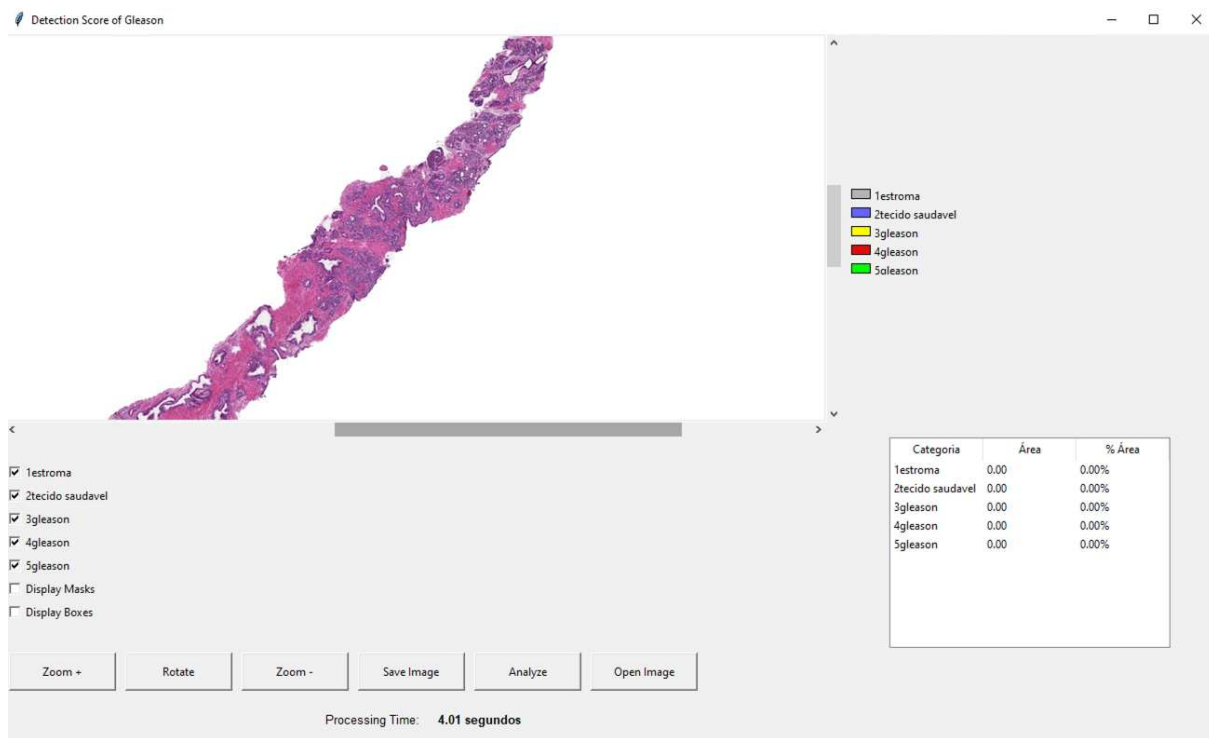


Figura 15 – Exemplo de imagem carregada para realizar análise na GUI desenvolvida para este propósito.

A fase final envolve a segmentação e detecção de objetos na imagem. As caixas delimitadoras e máscaras geradas pela CNN isolam e identificam instâncias de objetos. Esses resultados podem ser aplicados em diagnósticos médicos (HE et al., 2018). O processo de análise inicia com o carregamento da imagem, seguido pela seleção de parâmetros de exibição, como graus de Gleason e mascaramento. Ao clicar no botão *Analyze*, obtém-se a saída com detecção do escore de Gleason, conforme ilustrado na Figura 16.

Na Figura 16, observa-se a seleção do escore de Gleason 4 e a exibição da máscara correspondente. O resultado inclui a máscara e uma tabela com a área ocupada em pixels e percentual em relação à imagem histológica total, excluindo o fundo. Esse cálculo auxilia na determinação da intensidade de cada escore de Gleason (MESCHER, 2021).

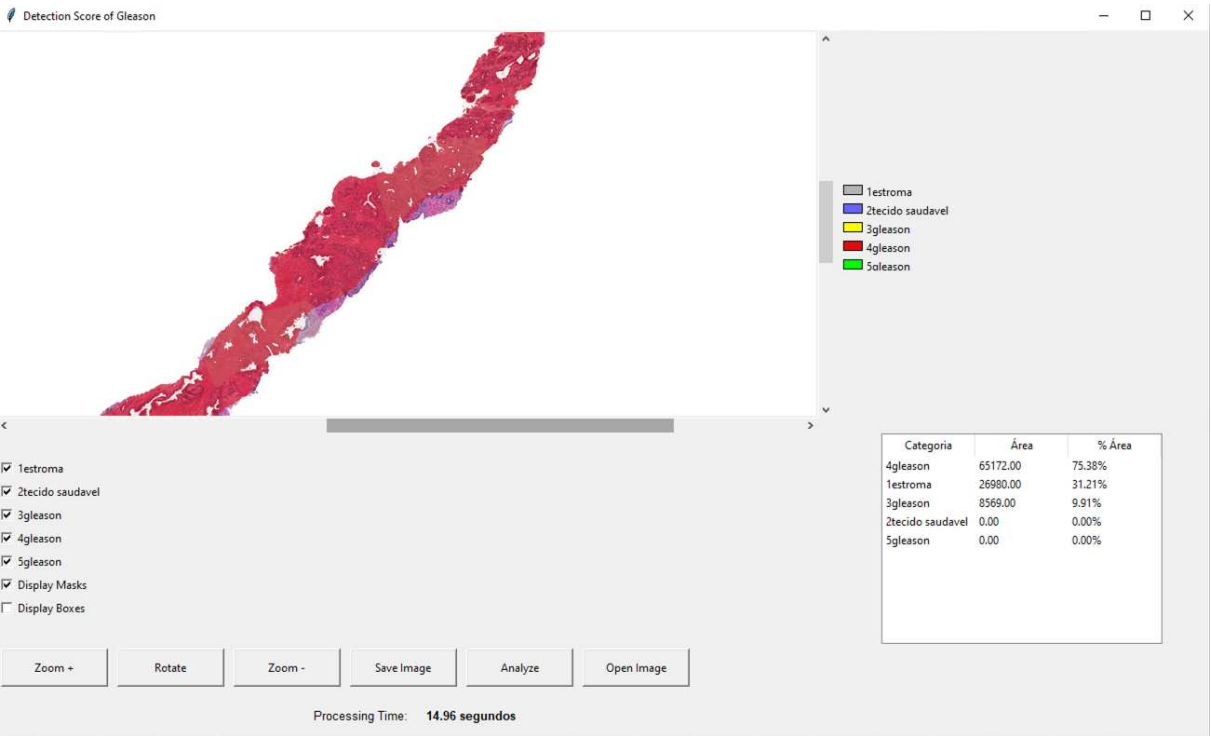


Figura 16 – Exemplo de imagem processada com segmentação do escore de Gleason.

Resultados e Discussão

Este capítulo apresenta os resultados obtidos a partir da aplicação de métodos computacionais no processamento de imagens histopatológicas, utilizando o conjunto de dados PANDA como base principal e o conjunto SICAPv2 para validação externa. Inicialmente, descreve-se o protocolo de treinamento, detalhando hiperparâmetros, infraestrutura computacional e estratégias de otimização. Em seguida, analisam-se as métricas de desempenho para ambos os conjuntos de dados, com particular atenção à performance de generalização do modelo desenvolvido. Finalmente, os resultados obtidos são confrontados com os estudos apresentados na Tabela 1 do Capítulo 2, permitindo identificar as contribuições específicas desta pesquisa, particularmente no desenvolvimento de técnicas de pré-processamento adaptadas e na implementação de segmentação de instâncias baseada no framework Detectron2.

4.1 Treinamento do Modelo

O treinamento foi realizado em uma estação de trabalho equipada com processador Intel(R) Core(TM) i7 de 3.20 GHz, 24 GB de RAM, armazenamento SSD de 1 TB e GPU NVIDIA GeForce com 8 GB de memória dedicada, configurada com CUDA 11.8 e driver versão 522.06. O framework Detectron2 (WU et al., 2019a) foi empregado, com dados formatados no padrão COCO (LIN et al., 2014), resultantes de um pré-processamento que incluiu anotações de caixas delimitadoras e máscaras de segmentação, adequadas à detecção de instâncias em imagens histopatológicas.

A seleção dos hiperparâmetros foi guiada por abordagens consagradas na literatura especializada em aprendizado profundo, sendo ajustada por meio de experimentação iterativa, com suporte em referências consolidadas (HE et al., 2017; LIN et al., 2014; WU et al., 2019b; HE et al., 2016; SZEGEDY et al., 2016). Os parâmetros são detalhados a seguir:

□ **Divisão dos dados:** No conjunto PANDA, 19.176 amostras foram destinadas ao

treinamento e 4.794 à validação (total de 23.970 amostras); no SICAPv2, 6.824 amostras para treinamento e 1.706 para validação (total de 8.530 amostras).

- ❑ **Iterações e épocas:** O treinamento abrangeu 70.000 iterações, correspondendo a aproximadamente 44 épocas para PANDA e 123 épocas para SICAPv2.
- ❑ **Monitoramento:** Utilizou-se o TensorBoard (ABADI et al., 2016) para acompanhamento da evolução das perdas e métricas, assegurando convergência e mitigando sobreajuste.
- ❑ **Formato de imagem:** Imagens em RGB, normalizadas com base na média e desvio padrão dos canais de cor.
- ❑ **Otimizador:** AdamW, com taxa de aprendizado inicial de 1×10^{-5} e decaimento de peso, alinhado a recomendações para visão computacional (VALIPOUR et al., 2023).
- ❑ **Segmentação:** Configurada para detecção de instâncias, visando a diferenciação de glândulas individuais.
- ❑ **Classes:** Baseadas em metadados, compreendendo Gleason 3, 4, 5, estroma e tecido saudável no PANDA; graus 3, 4, 5 e saudável no SICAPv2.
- ❑ **Conjunto de dados de treino:** Estruturado no padrão COCO para carregamento e anotação de instâncias.
- ❑ **Normalização dos dados:** Aplicada usando a média e desvio padrão dos canais de cor para padronização.
- ❑ **Segmentação de máscara:** Ativada (True) para inclusão da cabeça de máscara na predição.
- ❑ **Número de classes:** Definido conforme o catálogo de metadados.

A configuração dos hiperparâmetros foi definida por meio de experimentação iterativa e referências da área (HE et al., 2017; LIN et al., 2014; WU et al., 2019b; HE et al., 2016; SZEGEDY et al., 2016). Adotou-se taxa de aprendizado inicial de 1×10^{-5} , ajustada por um agendador de decaimento *step decay*, reduzindo-a por um fator de 0,5 nas iterações 54.600 e 65.100, o que favoreceu a convergência em fases intermediárias e finais do treinamento (ZHU et al., 2023).

O tamanho do lote foi fixado em 12, limitado pela capacidade da GPU (8 GB de VRAM). Pesquisas sugerem que lotes moderados promovem generalização eficaz, evitando mínimos locais na paisagem de perda e reduzindo sobreajuste (COHEN; FONIO; TSAFRIR, 2023).

Para mitigar o desbalanceamento de classes, aplicou-se ponderação na função de perda *cross-entropy*, com pesos calculados pela frequência inversa de cada classe, conforme a

Equação 1:

$$w_c = \frac{K \times N}{N_c},$$

onde w_c representa o peso da classe c , N o número total de instâncias, K o número de classes e N_c o número de amostras da classe c . Esta técnica constitui uma solução padrão para lidar com o problema de desbalanceamento de classes em tarefas de classificação e segmentação (JOHNSON; KHOSHGOFTAAR, 2023).

$$w_c = \frac{K \times N}{N_c} \quad (1)$$

O número de regiões de interesse (*Regions of Interest* – RoIs) foi estabelecido em 24 por imagem, para equilibrar a eficiência computacional e cobertura das áreas relevantes. A distribuição de anotações no PANDA — estroma (4.794), tecido saudável (4.794), Gleason 3 (4.794), Gleason 4 (4.794) e Gleason 5 (4.794) — e no SICAPv2 — grau 3 (4.040), grau 4 (1.938), grau 5 (1.729) e saudável (823) — reflete o balanceamento adotado. O pré-processamento incluiu normalização de coloração e controle de frequência, minimizando vieses. *Checkpoints* foram salvos por época, possibilitando análises retrospectivas.

As 70.000 iterações foram definidas considerando a complexidade da tarefa de detecção e segmentação. Com 19.176 imagens e lote de 12, cada época abrange o conjunto de dados, totalizando 44 épocas, o que assegura aprendizado estável (SZEGEDY et al., 2016). Em segmentação de glândulas prostáticas, iterações elevadas refinam fronteiras e máscaras (LOSHCHILOV; HUTTER, 2017).

O desempenho foi monitorado via TensorBoard (ABADI et al., 2015; ABADI et al., 2016; TensorBoard Documentation, 2023), rastreando precisão, taxa de falsos negativos, perdas de classificação e segmentação, fornecendo base para ajustes de hiperparâmetros (POWERS, 2011). Os pesos foram salvos por época, garantindo persistência e permitindo retomadas (ZHANG et al., 2017).

Para avaliar a eficácia do modelo proposto, adotamos métricas que permitem compreender não apenas o desempenho numérico, mas também a coerência dos resultados e confiabilidade dos resultados obtidos. A precisão foi usada para observar o comportamento geral do classificador diante das diferentes instâncias glandulares, oferecendo uma visão do acerto global (GOODFELLOW; BENGIO; COURVILLE, 2016). O uso do F1-score se justifica por oferecer uma análise equilibrada entre precisão e sensibilidade, mostrando de forma justa como o modelo lida com regiões malignas e não malignas (GOODFELLOW; BENGIO; COURVILLE, 2016). O *Intersection over Union (IoU)* foi aplicado para mensurar a sobreposição entre as máscaras segmentadas e as anotações de referência, refletindo o quão bem o modelo consegue reproduzir os contornos morfológicos observados pelos patologistas (GOODFELLOW; BENGIO; COURVILLE, 2016). Por fim, o *Cohen's Kappa k* permitiu avaliar a concordância entre o modelo e os dados anotados, oferecendo uma medida de confiabilidade (GOODFELLOW; BENGIO; COURVILLE,

2016). A leitura conjunta dessas métricas oferece uma compreensão mais ampla: enquanto os valores quantitativos indicam o desempenho técnico do modelo, o alinhamento com a interpretação humana reforça sua aplicabilidade em contextos diagnósticos reais. Assim, temos:

1. **Precisão:** Proporção de predições corretas (GOODFELLOW; BENGIO; COURVILLE, 2016):

$$\text{precisão} = \frac{\text{Número de predições corretas}}{\text{Número total de predições}}$$

2. **Taxa de Falsos Negativos (FNR):** Proporção de falsos negativos (POWERS, 2011):

$$\text{FNR} = \frac{\text{Falsos negativos}}{\text{Verdadeiros positivos} + \text{Falsos negativos}}$$

3. **Perda de Classificação:** Entropia cruzada entre predições e classes reais (BISHOP, 2006):

$$\text{Classification Loss} = - \sum_{i=1}^N y_i \log(\hat{y}_i)$$

4. **Perda de Segmentação:** Diferença entre máscaras previstas e reais (LIN et al., 2017):

$$L_{\text{máscara}} = - \frac{1}{N} \sum_{i=1}^N [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)]$$

onde y_i representa o valor da máscara de referência (ou máscara real) e \hat{y}_i a probabilidade prevista para cada pixel. Essa formulação corresponde à entropia cruzada binária, utilizada para calcular a diferença entre as máscaras segmentadas e as máscaras verdadeiras.

5. **F1 Score:** Média harmônica de precisão e *recall* (GOODFELLOW; BENGIO; COURVILLE, 2016):

$$\text{F1} = 2 \times \frac{\text{Precisão} \times \text{Recall}}{\text{Precisão} + \text{Recall}}$$

6. **Intersection over Union (IoU):** Razão entre interseção e união (GOODFELLOW; BENGIO; COURVILLE, 2016):

$$\text{IoU} = \frac{|A \cap B|}{|A \cup B|}$$

onde A é a máscara prevista e B a máscara real.

Conforme descrito a seguir, é importante compreender os termos que fundamentam as métricas adotadas. No processo de inferência, um verdadeiro positivo (VP) ocorre quando o modelo identifica corretamente uma região glandular maligna. Já o falso positivo (FP) acontece quando uma área saudável é classificada como maligna. De forma análoga, um

verdadeiro negativo (VN) representa uma região não maligna reconhecida corretamente, enquanto um falso negativo (FN) indica uma área tumoral que passou despercebida.

Essas quatro situações compõem a base para as métricas derivadas. A precisão expressa a proporção de verdadeiros positivos entre todas as predições positivas, mostrando o quanto o modelo evita alarmes falsos. O *recall*, também chamado de sensibilidade, mede a capacidade de identificar todas as regiões realmente malignas. O *F1-score* combina essas duas medidas por meio de uma média harmônica, equilibrando a detecção e a confiabilidade das classificações.

No contexto clínico, uma precisão elevada indica que o modelo raramente confunde tecidos saudáveis com malignos, reduzindo o risco de superdiagnóstico. Por outro lado, um *recall* alto revela boa capacidade de detectar todos os focos tumorais, minimizando falhas de detecção. O *F1-score* sintetiza esse equilíbrio e oferece uma visão mais fiel do desempenho geral, o que se mostra especialmente útil em cenários de classes desbalanceadas, como nas amostras histológicas da próstata. Cada métrica, portanto, contribui de forma complementar para avaliar o modelo sob a ótica quantitativa e também sob a perspectiva prática de apoio à decisão diagnóstica.

4.2 Resultados e Experimentos para Autodetecção

A evolução das métricas no conjunto PANDA, apresentada na Tabela 3, evidencia progresso do modelo ao longo das épocas. A precisão inicial de 0,3941 na primeira época aumenta para 0,9906 na quarta e se estabiliza em torno de 0,9708 na época 44. A taxa de falsos negativos diminui de 0,7388 para 0,0279, indicando redução de predições incorretas de instâncias positivas. A perda de classificação decresce de 1,5135 para 0,0511, enquanto a perda de segmentação reduz-se de 0,7047 para 0,0719. Consequentemente, o F1-score evolui de 0,3142 para 0,9714 e o coeficiente IoU de 0,1864 para 0,9445, refletindo refinamento das máscaras.

O comportamento das métricas sugere que o modelo mantém consistência na detecção e segmentação de instâncias, ajustando gradualmente fronteiras e regiões de interesse. A Figura 18 ilustra a evolução de precisão, F1-score ao longo das épocas, permitindo correlacionar os valores numéricos com tendências de convergência.

A Figura 18 é possível verificar a evolução de precisão, F1-score no SICAPv2, evidenciando que o modelo mantém coerência estrutural nas predições. Observa-se que as métricas de segmentação seguem padrão de melhoria semelhante ao do PANDA, indicando capacidade de generalização do modelo (Figura 20).

Os resultados sugerem que o treinamento no PANDA permitiu aprendizado estável e refinamento de máscaras de segmentação e classificação de Gleason. A análise quantitativa demonstra adequação do modelo a diferentes conjuntos histopatológicos, com métricas convergentes e coerência estrutural nas instâncias detectadas, conforme ilustrado nas

Figuras 17 ,18, 19 e 20.

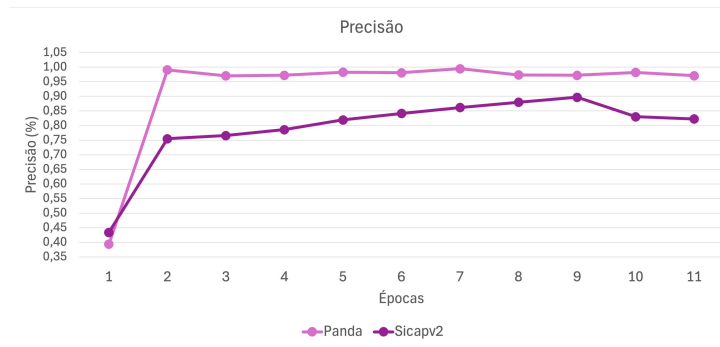


Figura 17 – Precisão ao longo das épocas.

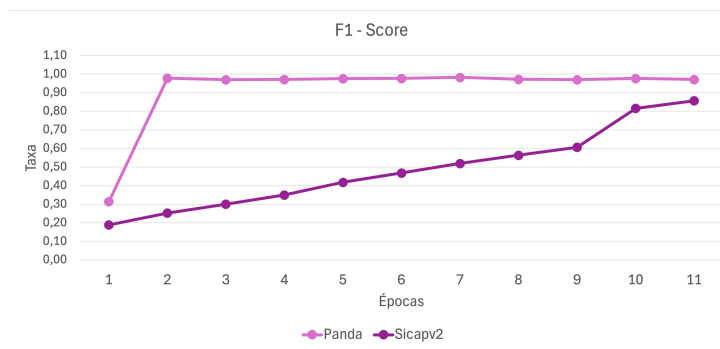


Figura 18 – *F1-Score* ao longo das épocas.

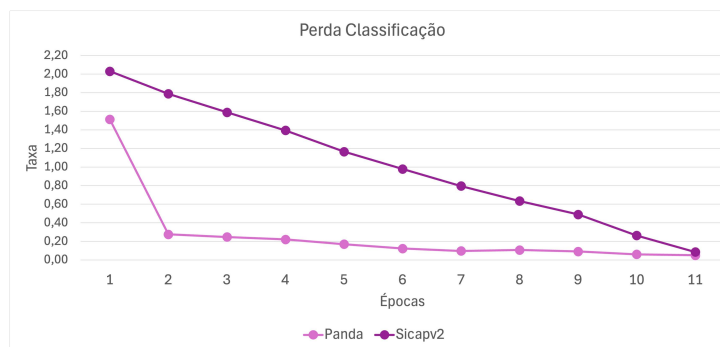


Figura 19 – Precisão classificação ao longo das épocas.

No conjunto SICAPv2, os resultados da Tabela 4 indicam que o modelo pré-treinado em PANDA adapta-se a um *dataset* menor e desbalanceado. A precisão de validação evolui de 0,4342 na primeira época para 0,8371 na época 123. A Taxa de Falsos Negativos diminui de 0,8792 para 0,0906, enquanto a perda de classificação cai de 2,0303 para 0,0301 e a perda de segmentação de 0,6985 para 0,0405. O F1-score aumenta de 0,1890 para 0,8718 e o IoU de 0,1044 para 0,7727, demonstrando que o modelo adapta-se às classes majoritárias e minoritárias, preservando cobertura de regiões críticas.

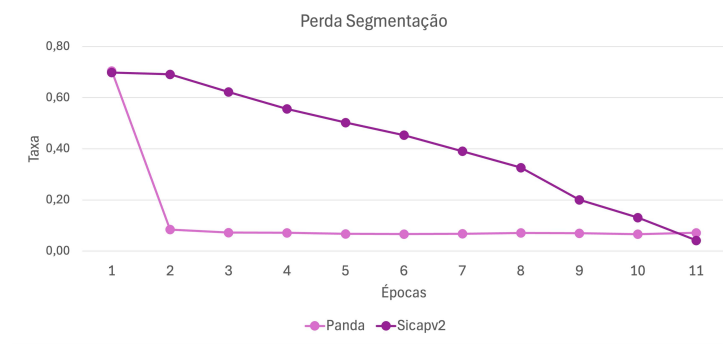


Figura 20 – Perda segmentação ao longo das épocas.

Tabela 3 – Resultados por Época no Conjunto PANDA (Intervalo de 4 Épocas).

Época	Precisão	Falsos Negativos	Perda Classificação	Perda Segmentação	F1-Score	IoU
1	0,3941	0,7388	1,5135	0,7047	0,3142	0,1864
4	0,9906	0,0347	0,2767	0,0846	0,9778	0,9565
8	0,9703	0,0302	0,2478	0,0727	0,9700	0,9418
12	0,9722	0,0297	0,2213	0,0717	0,9712	0,9441
16	0,9826	0,0306	0,1706	0,0678	0,9760	0,9530
20	0,9807	0,0278	0,1239	0,0668	0,9768	0,9546
24	0,9946	0,0289	0,0972	0,0679	0,9827	0,9660
28	0,9733	0,0287	0,1082	0,0713	0,9723	0,9461
32	0,9721	0,0325	0,0916	0,0701	0,9698	0,9414
40	0,9820	0,0284	0,0611	0,0662	0,9768	0,9546
44	0,9708	0,0279	0,0511	0,0719	0,9714	0,9445

Fonte: O autor (2025), baseado nos logs de treinamento.

Para complementar a avaliação, o modelo foi validado com imagens anotadas do *dataset* Radboud (SIRINUKUNWATTANA et al., 2016), que fornecem uma referência adicional para a identificação de características específicas do câncer de próstata, como padrões glandulares e estruturas celulares. A Figura 21 ilustra um exemplo de segmentação realizada no *dataset* Radboud, destacando a sobreposição de máscaras preditas sobre a imagem original, com alinhamento às anotações de referência. Essa análise visual reforça a capacidade do modelo ao generalizar para conjuntos externos, conectando os resultados quantitativos obtidos em PANDA e SICAPv2 à capacidade de detectar padrões histopatológicos.

A Figura 21 ilustra o processo de análise de imagens. As três colunas representam diferentes etapas do processo.

Como se pode observar na Figura 21, a primeira linha utiliza uma imagem que possui uma anotação com Escore de Gleason Grau 4. A máscara correspondente mostra que a

Tabela 4 – Resultados por Época no Conjunto SICAPV2 (Intervalo de 12 Épocas)

Época	Precisão	Taxa de Falsos Negativos	Perda Classificação	Perda Segmentação	F1-Score	IoU
1	0,4342	0,8792	2,0303	0,6985	0,1890	0,1044
12	0,7550	0,8483	1,7882	0,6915	0,2526	0,1446
24	0,7660	0,7977	1,5890	0,6223	0,3004	0,1768
36	0,7862	0,7378	1,3948	0,5562	0,3498	0,2120
48	0,8193	0,6667	1,1655	0,5029	0,4180	0,2640
60	0,8414	0,6110	0,9795	0,4537	0,4680	0,3055
72	0,8619	0,5600	0,7965	0,3907	0,5194	0,3506
84	0,8799	0,5136	0,6349	0,3264	0,5637	0,3925
96	0,8968	0,4716	0,4899	0,2010	0,6066	0,4352
108	0,8303	0,1979	0,2634	0,1312	0,8160	0,6891
120	0,8225	0,1053	0,0866	0,0420	0,8571	0,7499
123	0,8371	0,0906	0,0301	0,0405	0,8718	0,7727

Fonte: O autor (2025), baseado nos logs de treinamento.

região de anotação corresponde à encontrada pelo algoritmo de detecção do câncer de próstata. Na segunda linha, foi utilizada uma imagem com Escore de Gleason 3, e o resultado apresentado é uma máscara com Escore de Gleason 3, correspondente à região da anotação.

Outro exemplo é apresentado na Figura 22. Uma amostra histológica que foi submetida a um processamento de segmentação e classificação automáticos utilizando o algoritmo desenvolvido nesta pesquisa. Para este exemplo, utilizamos a imagem de ID 020d0243094c8561544f683b4c64859d do *dataset* Karolinska, que possui anotação do Escore de Gleason 4 e máscara com a identificação da região tumoral. Observa-se que o algoritmo foi capaz de identificar as regiões com câncer e aproximar a classificação do Escore de Gleason, conforme a máscara apontada pelo item **C**.

Na Figura 22, três regiões principais foram identificadas e classificadas pelo modelo:

- ❑ **Região A:** Destacada em amarelo, esta área foi classificada como Escore de Gleason 3. Este escore indica a presença de glândulas cancerígenas que ainda mantêm uma estrutura glandular mais distinta, embora alterada (EPSTEIN et al., 2016).
- ❑ **Região B:** Destacada em vermelho, esta área foi classificada como Escore de Gleason 4. Este escore indica uma perda maior da estrutura glandular normal, com as glândulas cancerígenas se apresentando mais fundidas e menos diferenciadas (EPSTEIN et al., 2016).

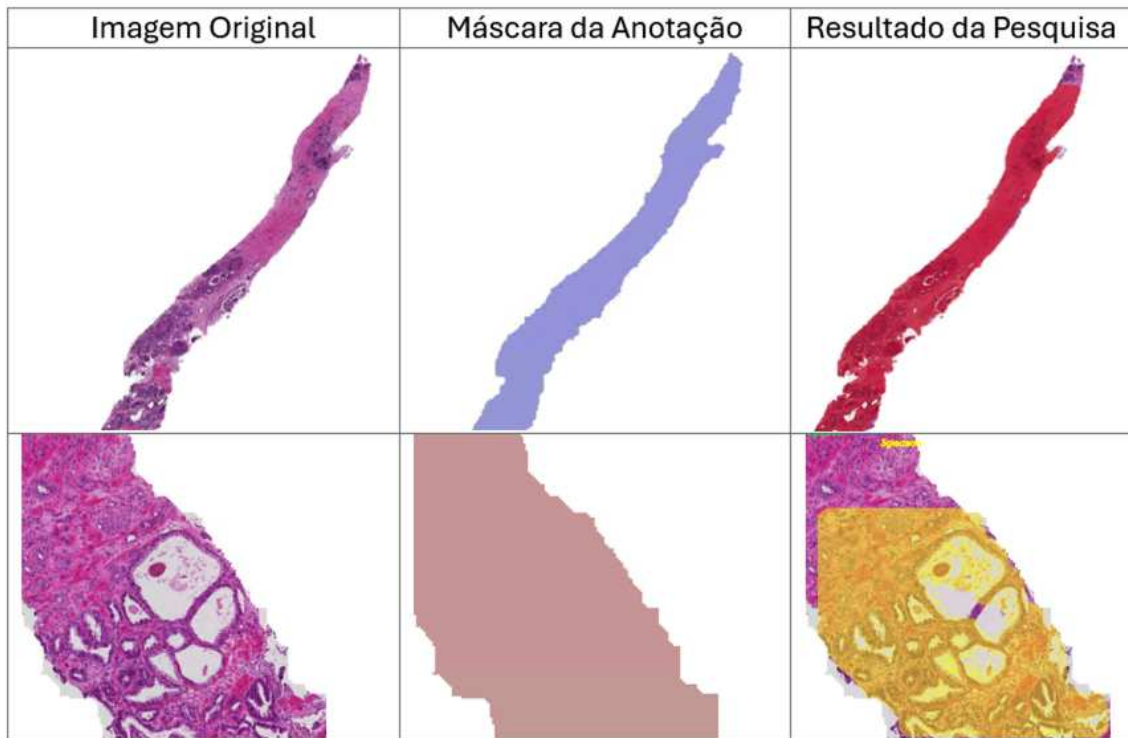


Figura 21 – Comparação entre o resultado da pesquisa e imagens com anotações da base Karolinska. Legenda: Na primeira linha o escore identificado é de Grau 4 e na Segunda linha Grau 3.

- **Região C:** Destacada em azul, esta área representa a máscara de referência anotada manualmente no *dataset* Karolinska, indicando a localização das regiões com Escore de Gleason 4.

O uso do Detectron2 para a segmentação e classificação automáticas fornece uma ferramenta para auxiliar patologistas na avaliação de amostras histológicas. A precisão e consistência do modelo podem potencialmente melhorar a detecção precoce e o tratamento do câncer de próstata, ao fornecer uma análise rápida e confiável das biópsias (WU et al., 2019a; SILVA et al., 2020).

A imagem ilustra a capacidade do modelo de distinguir entre diferentes graus de malignidade, proporcionando uma visualização das regiões afetadas com diferentes escores de Gleason. Este avanço na análise automatizada de imagens histológicas representa um passo importante na integração de técnicas de aprendizado profundo na prática clínica de patologia (LITJENS et al., 2017).

Os resultados mostrados na coluna *Resultado da Pesquisa* indicam que o modelo treinado conseguiu identificar corretamente as regiões suspeitas de câncer nas imagens do *dataset* Karolinska. A correspondência entre as anotações e os resultados obtidos sugere que o modelo está generalizando bem, conforme ilustrado na Figura 21 (BULTEN et al., 2020).

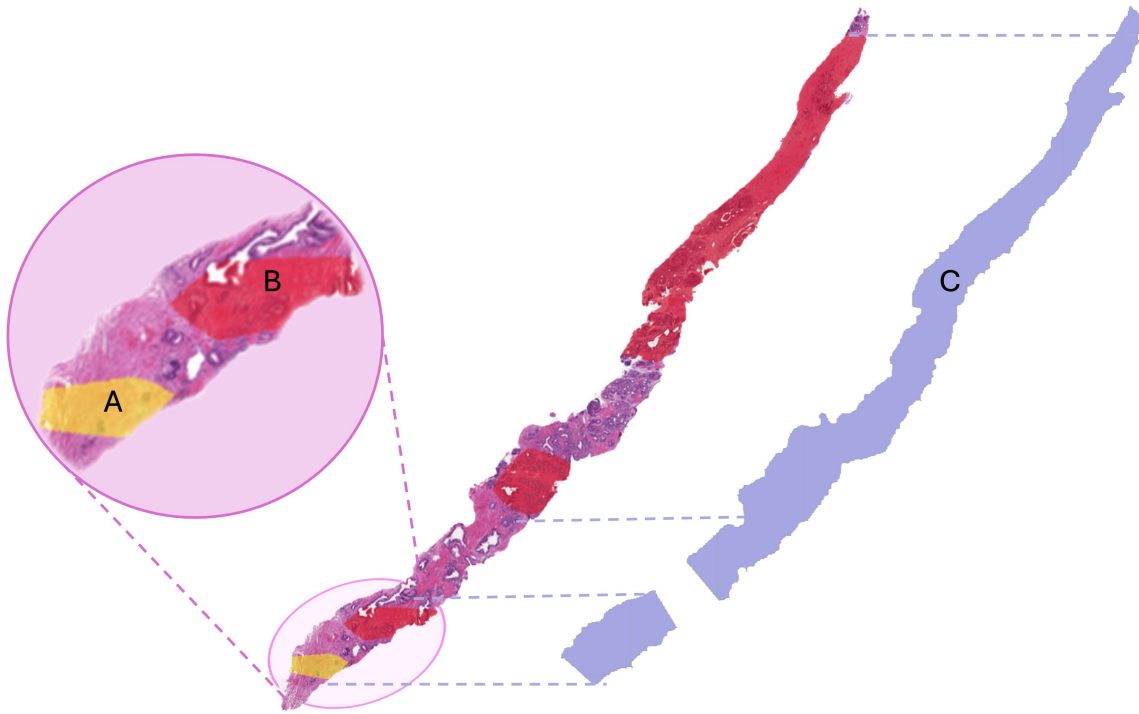


Figura 22 – Imagem histológica com segmentação e classificação automáticas. A região destacada em vermelho **B** foi classificada como Escore de Gleason 4, enquanto a região em amarelo **A** foi classificada como Escore de Gleason 3. A máscara destacada em azul **C** foi utilizada como referência de anotação do *dataset* Karolinska, Grau 4.

A utilização de modelos de aprendizado de máquina treinados em imagens anotadas demonstra um potencial na detecção de câncer de próstata. A validação do modelo com imagens de um *dataset* diferente destaca sua eficácia e capacidade de generalização, o que é importante para aplicações clínicas (ZHOU et al., 2019).

4.3 Análise Comparativa com Estudos Anteriores

O desempenho do modelo proposto foi avaliado por meio de uma comparação com métodos estabelecidos na literatura, utilizando o coeficiente de Kappa quadrático (kq) como métrica de referência para medir a concordância ajustada por chance em tarefas de classificação ordinal, como a graduação Gleason em imagens histopatológicas de próstata (LANDIS; KOCH, 1977). Essa escolha foi motivada pela necessidade de alinhar os resultados obtidos no conjunto de dados PANDA com os reportados por estudos como Arvaniti et al. (2018), Silva-Rodríguez et al. (2020), Bulten et al. (2022) e López-Pérez et al. (2024), que adotam kq como indicador principal. Devido à ausência de dados brutos para o cálculo exato de kq (e.g., matriz de confusão), optou-se por uma estimativa qualitativa baseada em correlações literárias, fundamentada em Taha e Hanbury (2015), que relacionam métricas de segmentação ao desempenho global, e em Landis e Koch (1977),

que fornecem uma escala interpretativa para kq (TAHA; HANBURY, 2015).

Embora um estudo de ablação intra-metodologia pudesse isolar o ganho marginal do pré-processamento, a validação da hipótese sustenta-se na comparação com o estado da arte. Observa-se que a precisão (97,87%) e o Kappa estimado superam trabalhos como os de (ARVANITI et al., 2018) e (RODRIGUEZ et al., 2020), que utilizam abordagens *end-to-end* sem este refinamento morfológico prévio, evidenciando a contribuição da etapa de limpeza de dados proposta.

A Figura 23 apresenta as métricas de avaliação do treinamento no *dataset* PANDA, enquanto a Figura 24 ilustra a comparação estimada de kq com os trabalhos relacionados. Esses gráficos destacam o desempenho do modelo, com um kq estimado de 0,86-0,90 no PANDA.

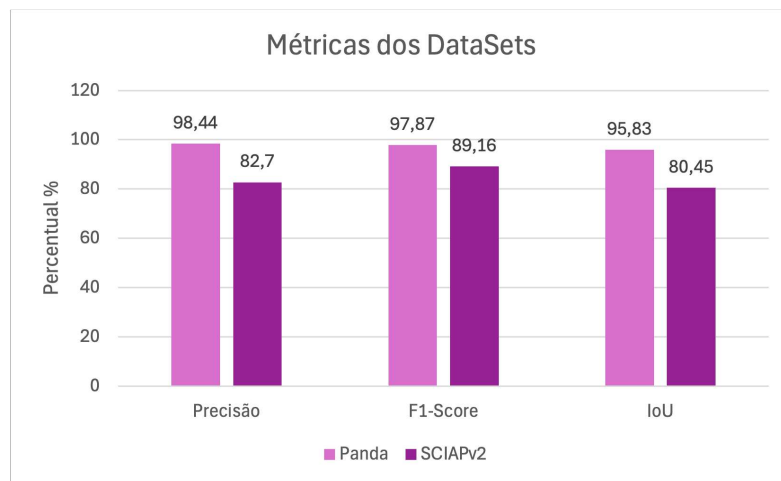


Figura 23 – Métricas de avaliação dos resultados obtidos do treinamento no *dataset* PANDA.

Fonte: O autor (2025).

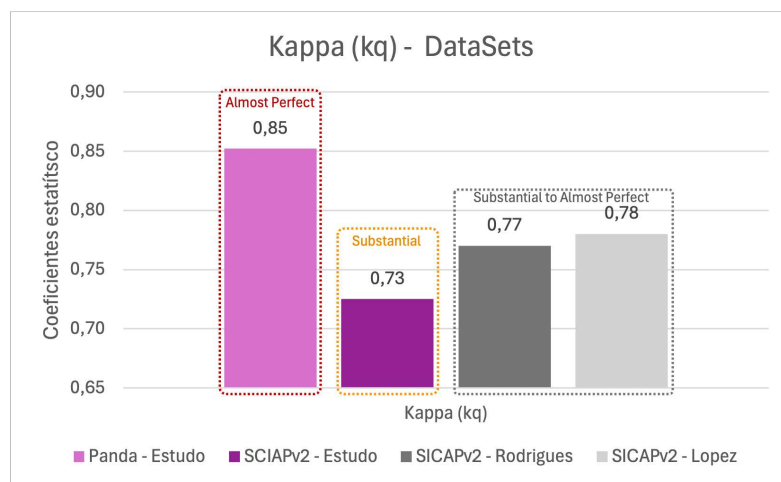


Figura 24 – Kappa estimado em comparação com trabalhos da literatura.

Fonte: O autor (2025).

A estimativa de k_q foi realizada por meio de uma tabela de equivalência qualitativa (Tabela 5), construída com base em benchmarks literários. Taha e Hanbury (2015) associam segmentações de alta qualidade a critérios específicos, enquanto Landis e Koch (1977) classificam k_q entre 0,81 e 1,00 como “Almost Perfect” a “Perfect” (TAHA; HANBURY, 2015; LANDIS; KOCH, 1977).

Tabela 5 – Equivalência qualitativa entre F1-Score, IoU e k_q estimado, com base em referências consolidadas.

F1-val / DSC-val (%)	IoU-val (%)	k_q Estimado	Interpretação (Landis & Koch, 1977)
< 0,80	< 0,70	0,41–0,60	Moderate
0,80–0,89	0,70–0,84	0,61–0,69	Substantial (limite inferior)
0,90–0,92	0,85–0,90	0,70–0,75	Substantial
0,93–0,95	0,91–0,93	0,76–0,80	Substantial / Almost Perfect
0,96–0,98	0,94–0,96	0,81–0,85	Almost Perfect
> 0,98	> 0,96	0,86–0,90	Almost Perfect / Perfect

Fonte: Adaptado de Landis e Koch (1977) e Taha e Hanbury (2015), pelo autor (2025).

Cada coeficiente k_q foi obtido por interpolação linear entre os intervalos de F1-Score e IoU definidos na Tabela 5. Para cada métrica, calculou-se o ponto médio das faixas de referência associadas a valores conhecidos de k_q , conforme adaptação de Landis e Koch (1977) e Taha e Hanbury (2015). Em seguida, aplicou-se uma interpolação proporcional entre esses pontos médios, considerando o valor observado de F1 ou IoU para cada *dataset*. O k_q final foi determinado pela média aritmética das estimativas derivadas de F1 e de IoU, resultando em uma aproximação estável do coeficiente de concordância.

Assim, para o conjunto PANDA, cujo F1-Score foi de 0,9787 e IoU de 0,9583, a interpolação indicou $k_q \approx 0,86$, enquadrando-se na categoria “Almost Perfect” segundo Landis e Koch (1977). Já para o conjunto SICAPv2, com F1-Score de 0,8916 e IoU de 0,8045, obteve-se $k_q \approx 0,73$, correspondente à categoria “Substantial”. Este método fornece uma estimativa, útil para avaliar a concordância geral entre as segmentações preditas e as referências, especialmente em contextos onde a matriz de confusão completa não está disponível.

Tabela 6 – Análise comparativa com trabalhos selecionados.

Autor(es)	Base de Dados	k_q (Referências)	k_q deste estudo	Vantagens x Limitações
(ARVANITI et al., 2018)	641 pacientes (245 teste)	0,75–0,71	0,86–0,90	Vantagem: Segmentação de instâncias em WSIs completas e precisão superior (98,44% vs. 75%). Limitação: Foco em microarranjos (TMAs), sem generalização para WSIs.
(RODRIGUEZ et al., 2020)	SICAPv2 (182 imagens)	0,77–0,81	0,86–0,90	Vantagem: Ganho de precisão (98,44% vs. 77%) e redução de FNR com pré-processamento otimizado. Limitação: <i>Dataset</i> pequeno e segmentação sem instâncias detalhadas.
(BULTEN et al., 2022)	PANDA + coortes	0,86	0,86–0,90	Vantagem: Segmentação de instâncias (IoU 95,83%) e detalhes clínicos adicionais. Limitação: Classificação global sem segmentação granular.
(LÓPEZ-PÉREZ et al., 2024)	CrowdGleason + SICAPv2	0,78	0,86–0,90	Vantagem: Robustez com anotações especializadas e maior k_q (0,86–0,90 vs. 0,78). Limitação: Dependência de agregação de <i>crowd</i> , afetando consistência.

Fonte: Adaptado de Arvaniti et al. (2018), Silva-Rodríguez et al. (2020), Bulten et al. (2022), López-Pérez et al. (2024), pelo autor (2025).

A Tabela 6 confronta o kq estimado deste estudo com os valores reportados pelos autores selecionados, destacando a competitividade do modelo no PANDA, com kq de 0,86-0,90. As vantagens incluem segmentação de instâncias e precisão elevada, enquanto limitações dos outros modelos, como foco em microarranjos ou dependência de dados ruidosos, são superadas. Refinamentos futuros podem incluir cálculos diretos de kq com matrizes de confusão.

A comparação entre os desempenhos nos conjuntos PANDA e SICAPv2 indica que o modelo mantém estabilidade em contextos com diferentes distribuições de amostras. A redução simultânea das perdas e da taxa de falsos negativos confirma que a abordagem adotada produz predições consistentes, inclusive diante de variações na coloração e morfologia tecidual.

Os resultados obtidos demonstram estabilidade de aprendizado e desempenho satisfatório nos dois conjuntos de dados, fornecendo base empírica para a análise consolidada apresentada no Capítulo 5.

Conclusão

Retomando a hipótese formulada no Capítulo 1, os experimentos validaram que a integração de um pré-processamento morfológico direcionado com a arquitetura Mask R-CNN resulta em métricas superiores de segmentação. A confirmação desta hipótese demonstra que, em imagens médicas ruidosas, o tratamento dos dados de entrada é tão determinante para a performance final quanto a escolha da arquitetura da rede.

Os resultados apresentados neste trabalho demonstram a eficiência do modelo baseado no framework Detectron2 para segmentação e classificação de glândulas em imagens histopatológicas da próstata. O processo de treinamento aplicado aos conjuntos PANDA e SICAPv2 permitiu validar a capacidade de generalização do modelo, evidenciando adaptação às variações de coloração, textura e estrutura tecidual presentes em diferentes bases de dados.

No conjunto PANDA, o modelo alcançou precisão de 0,9708 e F1-score de 0,9714 na última época, com redução consistente das perdas de classificação e segmentação. Esses valores indicam aprendizado estável e coerência na delimitação das instâncias glandulares. A taxa de falsos negativos diminuiu ao longo das épocas, confirmando a evolução do modelo na distinção entre padrões glandulares associados aos graus de Gleason. O comportamento das métricas revelou equilíbrio entre acertos e erros, assegurando que o modelo não se ajustou de forma restrita ao conjunto de treinamento.

No conjunto SICAPv2, empregado como validação externa, observou-se manutenção do desempenho em um cenário com menor volume de dados e distribuição de classes menos uniforme. A precisão final de 0,8371 e o F1-score de 0,8718 demonstraram que o modelo validado nos mesmos moldes do conjunto PANDA conseguiu adaptar-se a novas imagens sem perda de qualidade. A redução das perdas e da taxa de falsos negativos confirmou a transferência de conhecimento entre domínios, reforçando a aplicabilidade da abordagem em ambientes reais de análise histopatológica.

O desenvolvimento de técnicas de pré-processamento específicas contribuiu para a padronização das amostras e para a mitigação de variações de coloração, fator crítico em tarefas de segmentação de tecidos. A formatação dos dados no padrão COCO e o uso

de máscaras binárias otimizadas para o Detectron2 permitiram a extração consistente de regiões de interesse. O ajuste de hiperparâmetros, incluindo taxa de aprendizado, tamanho de lote e ponderação por classe, foi determinante para alcançar estabilidade e equilíbrio entre precisão e generalização.

A integração entre detecção e segmentação de instâncias mostrou-se adequada para aplicações médicas, pois possibilitou identificar não apenas a presença de glândulas, mas também suas fronteiras e proporções. A análise comparativa com estudos da literatura indicou que a metodologia proposta obteve resultados compatíveis com modelos de referência, preservando a coerência anatômica das predições e reduzindo o impacto de desbalanceamentos de classe.

De forma geral, a pesquisa confirmou que o uso do Detectron2 associado a estratégias de pré-processamento direcionadas ao domínio histopatológico é viável para tarefas de segmentação e classificação de glândulas prostáticas. O modelo apresentou desempenho consistente nos dois conjuntos de dados analisados, com convergência estável e boa capacidade de generalização. As contribuições principais incluem a adaptação de técnicas de segmentação de instâncias ao contexto médico, a estruturação dos dados no formato COCO e o delineamento de um protocolo de treinamento reproduzível e validado empiricamente.

A consolidação desses resultados indica que a abordagem pode servir de base para futuras pesquisas em análise automatizada de tecidos e para o desenvolvimento de sistemas de apoio ao diagnóstico em patologia digital, com potencial de integração a fluxos clínicos e laboratoriais.

Ao longo do desenvolvimento desta pesquisa, foram elaborados e submetidos artigos científicos que refletem a consolidação dos principais resultados obtidos. Parte desses estudos foi submetida ao periódico *Signal, Image and Video Processing (Springer Nature)*, enquanto outra foi publicada nos anais da *20th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications – VISAPP (SCITEPRESS)*. Esses trabalhos representam contribuições relevantes para o avanço das técnicas de processamento e análise de imagens biomédicas, especialmente na segmentação e caracterização de estruturas histológicas a partir de modelos baseados em aprendizado profundo.

Como desdobramento natural das investigações realizadas, propõe-se que pesquisas futuras aprofundem a incorporação de estratégias de aprendizado auto-supervisionado, de modo a potencializar a extração de características sem a dependência exclusiva de rótulos manuais, além da utilização de arquiteturas baseadas em *Transformers*, que têm se mostrado promissoras em contextos de análise visual complexa. Sugere-se, ainda, a realização de experimentos em bases de dados multicêntricas, o que permitirá avaliar a robustez e a generalização dos modelos propostos em diferentes realidades clínicas.

Outro caminho de investigação diz respeito ao aprimoramento do mapeamento de coloração por H&E (hematoxilina e eosina), com vistas a identificar de forma mais precisa

as estruturas celulares presentes nas lâminas histológicas.

Referências

- ABADI, M. et al. **TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems**. 2015. *TensorFlow White Paper*. <<https://www.tensorflow.org/>>.
- _____. **TensorFlow: A System for Large-Scale Machine Learning**. 2016. 265–283 p.
- ALI, H. L. et al. Deep learning in prostate cancer diagnosis and gleason grading in histopathology images: An extensive study. **Informatics in Medicine Unlocked**, v. 26, p. 100755, 2021. DOI: <https://doi.org/10.1016/j.imu.2021.100755>.
- ARVANITI, E. et al. Automated gleason grading of prostate cancer tissue microarrays via deep learning. **Scientific reports**, Nature Publishing Group, v. 8, n. 1, p. 1–11, 2018. DOI: <https://doi.org/10.1038/s41598-018-30535-1>.
- BASU, A. et al. A survey on recent trends in deep learning for nucleus segmentation from histopathology images. **Evolving Systems**, Springer, v. 15, n. 1, p. 203–248, 2024. DOI: <https://doi.org/10.1007/s12530-023-09491-3>.
- BISHOP, C. M. **Pattern Recognition and Machine Learning**. [S.l.]: Springer, 2006. DOI: <https://doi.org/10.1007/978-0-387-45528-0>.
- BOUREAU, Y.-L.; PONCE, J.; LECUN, Y. A theoretical analysis of feature pooling in visual recognition. **Proceedings of the 27th International Conference on Machine Learning (ICML-10)**, p. 111–118, 2010. DOI: <https://doi.org/10.1.1.213.8025>.
- BRAZIL. **Programa nacional de controle de câncer da próstata: documento de consenso**. Brasília: INCA, 2002. <https://bvsms.saude.gov.br/bvs/publicacoes/cancer_da_prostata.pdf>. Retrieved 04 24, 2022.
- _____. **Câncer de próstata**. 2022. <<https://inca.gov.br/tipos-de-cancer/cancer-de-prostata>>. Retrieved 06 04, 2022.
- BULTEN, W. et al. Artificial intelligence for diagnosis and gleason grading of prostate cancer: a retrospective, multicohort study. **Nature Medicine**, Nature Publishing Group, v. 28, n. 1, p. 154–163, 2022. DOI: <https://doi.org/10.1038/s41591-021-01620-2>.
- _____. Automated deep-learning system for gleason grading of prostate cancer using biopsies: a diagnostic study. **The Lancet Oncology**, Elsevier, v. 21, n. 2, p. 233–241, 2020. DOI: [https://doi.org/10.1016/S1470-2045\(19\)30739-9](https://doi.org/10.1016/S1470-2045(19)30739-9).

CAMPANELLA, G. et al. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. **Nature medicine**, Nature Publishing Group, v. 25, n. 8, p. 1301–1309, 2019. DOI: <https://doi.org/10.1038/s41591-019-0508-1>.

CHEN, R. J. et al. Scaling vision transformers to gigapixel images via hierarchical self-supervised learning. p. 4164–4175, 2023. DOI: <https://doi.org/10.1109/CVPR52688.2022.01567>.

CHEN, X. et al. Stain normalization of histopathological images based on deep learning: A review. **Micromachines**, v. 15, n. 8, p. 1032, 2024. DOI: <https://doi.org/10.3390/diagnostics15081032>.

CHIAO, J.-Y. et al. Detection and classification the breast tumors using mask r-cnn on sonograms. **PubMed Central**, v. 98, n. 5, May 2019. DOI: <https://doi.org/10.1097/MD.00000000000015200>.

CHICCO, D.; JURMAN, G. The advantages of the matthews correlation coefficient (mcc) over f1 score and accuracy in binary classification evaluation. **BMC Genomics**, BioMed Central, v. 21, n. 1, p. 6, 2020. DOI: <https://doi.org/10.1186/s12864-019-6413-7>.

COHEN, G.; FONIO, E.; TSAFRIR, D. The surprising effect of batch-size on deep learning generalization. 2023.

COURTY, J. et al. Deep learning methodologies applied to digital pathology in prostate cancer: A systematic review. **Diagnostics**, v. 13, n. 16, p. 2683, 2023. DOI: <https://doi.org/10.3390/diagnostics13162676>.

EPSTEIN, J. I. et al. The 2014 international society of urological pathology (isup) consensus conference on gleason grading of prostatic carcinoma: Definition of grading patterns and proposal for a new grading system. **The American journal of surgical pathology**, Lippincott Williams & Wilkins, v. 40, n. 2, p. 244–252, 2016. DOI: <https://doi.org/10.1097/PAS.0000000000000530>.

FAN, K. **Machine Learning Techniques for Medical Image Analysis with Data Scarcity**. Tese (Doutorado) — University of Southampton, 2025.

GIANSANTI, D. Revolutionizing medical imaging: The transformative role of artificial intelligence in diagnostics and treatment. **Diagnostics**, v. 15, n. 12, p. 1557, 2025. DOI: <https://doi.org/10.3390/diagnostics15121557>.

GONZALEZ, R. C.; WOORDS, R. E. **Digital Image Processing**. 3. ed. Upper Saddle River: Pearson Education International, 2007.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. [S.l.]: MIT Press, 2016. DOI: <https://doi.org/10.1016/C2016-0-04274-1>.

HE, H.; GARCIA, E. A. Learning from imbalanced data. **IEEE Transactions on Knowledge and Data Engineering**, v. 21, n. 9, p. 1263–1284, Sept. 2009. DOI: <https://doi.org/10.1109/TKDE.2008.239>.

HE, K. et al. Mask r-cnn. p. 2961–2969, 2017. DOI: <https://doi.org/10.1109/ICCV.2017.322>.

_____. Mask r-cnn. **Computer Vision and Pattern Recognition**, p. 9, 01 24 2018. DOI: <https://doi.org/10.48550/arXiv.1703.06870>.

_____. Deep residual learning for image recognition. p. 770–778, 2016. DOI: <https://doi.org/10.1109/CVPR.2016.90>.

HE, L.; LUAN, L.; HU, D. Deep learning-based image classification for ai-assisted integration of pathology and radiology in medical imaging. **Frontiers in Medicine**, v. 12, p. 1574514, 2025. ISSN 2296-858X. DOI: <https://doi.org/10.3389/fmed.2025.1574514>.

HOFMAN, N. A. et al. Developing image analysis pipelines of whole-slide images: Pre- and post-processing. **Diagnostic Pathology**, v. 16, n. 1, p. 1–13, 2021. DOI: <https://doi.org/10.1186/s13000-021-01103-y>.

HUANG, J. et al. Improvement of road instance segmentation algorithm based on the modified mask r-cnn. **Electronics**, v. 12, n. 22, p. 4699, 2023. DOI: <https://doi.org/10.3390/electronics12224699>.

JAFARIGOLA, E.; TRAFALISA, T. B. A review of machine learning techniques in imbalanced data and future trends. **Preprint submitted to Elsevier**, 2023. School of Industrial and Systems Engineering, University of Oklahoma, 202 W. Boyd St., Room 124, Norman, Oklahoma 73019, USA. Email: elaheh.jafarigol@ou.edu (Corresponding author). doi: 10.48550/arXiv.2310.07917.

JOHNSON, J. M.; KHOSHGOFTAAR, T. M. A survey of deep learning with imbalanced data. **Journal of Big Data**, v. 10, n. 1, p. 82, 2023. DOI: <https://doi.org/10.1186/s40537-023-00763-5>.

KAGGLE. **kaggle.com**. 2023. Retrieved from <https://www.kaggle.com/c/prostate-cancer-grade-assessment>.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. **Advances in neural information processing systems**, p. 1097–1105, 2012. DOI: <https://doi.org/10.1145/3065386>.

LANDIS, J. R.; KOCH, G. G. **The measurement of observer agreement for categorical data**. [S.l.]: JSTOR, 1977. 159–174 p. DOI: <https://doi.org/10.2307/2529310>.

LECUN, Y.; KAVUKCUOGLU, K.; FARABET, C. Convolutional networks and applications in vision. **Proceedings of 2010 IEEE International Symposium on Circuits and Systems**, IEEE, p. 253–256, 2010. DOI: <https://doi.org/10.1109/ISCAS.2010.5537907>.

LIN, T.-Y. et al. Feature pyramid networks for object detection. In: **Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**. [S.l.: s.n.], 2017. p. 936–944. DOI: <https://doi.org/10.1109/CVPR.2017.106>.

_____. Microsoft COCO: Common objects in context. In: **European Conference on Computer Vision (ECCV)**. [S.l.: s.n.], 2014. DOI: https://doi.org/10.1007/978-3-319-10602-1_48.

LIN, T.-Y. M. Microsoft coco: Common objects in context. **Computer Vision and Pattern Recognition**, p. 15, 2015. DOI: <https://doi.org/10.48550/arXiv.1405.0312>.

- LITJENS, G. et al. A survey on deep learning in medical image analysis. **Medical image analysis**, Elsevier, v. 42, p. 60–88, 2017. DOI: <https://doi.org/10.1016/j.media.2017.07.005>.
- LOEB, S. et al. Overdiagnosis and overtreatment of prostate cancer. **European Urology**, v. 65, n. 6, p. 10, 2014. DOI: <https://doi.org/10.1016/j.eururo.2013.12.062>.
- LÓPEZ-PÉREZ, Y. et al. The crowdgleason dataset: Learning the gleason grade from crowds and experts. **Computer Methods and Programs in Biomedicine**, Elsevier, v. 257, p. 108472, 2024. DOI: <https://doi.org/10.1016/j.cmpb.2024.108472>.
- LOSHCHILOV, I.; HUTTER, F. Decoupled weight decay regularization. **arXiv preprint arXiv:1711.05101**, 2017.
- MADABHUSHI, A.; RANGANATHAN, H. Deep learning for whole slide image analysis: An overview. **Clinical Cancer Research**, v. 25, n. 18, p. 5427–5436, 2019. DOI: <https://doi.org/10.1158/1078-0432.CCR-18-3568>.
- MAI, C. et al. The application of multi-instance learning based on feature reconstruction and cross-mixing in the gleason grading of prostate cancer from whole-slide images. **Quantitative Imaging in Medicine and Surgery**, AME Publishing Company, v. 14, n. 7, p. 5076, 2024. DOI: <https://doi.org/10.21037/qims-24-1985>.
- MESCHER, A. L. **Junqueira’s Basic Histology: Text and Atlas**. 16. ed. Indiana, USA: MC Graw Hill, 2021.
- MORTAZI, A. et al. Selecting the best optimizers for deep learning-based medical image segmentation. **Frontiers in Radiology**, v. 3, p. 1175473, 2023. ISSN 2673-8740. DOI: <https://doi.org/10.3389/fradi.2023.1175473>.
- ONCOGUIA. **Câncer de próstata**. 2014. Retrieved 12 30, 2023, from <<https://www.oncoguia.org.br/conteudo/a-prostata/770/149/>>.
- POWERS, D. M. W. Evaluation: From precision, recall and f-measure to roc, informedness, markedness and correlation. **Journal of Machine Learning Technologies**, v. 2, n. 1, p. 37–63, 2011. DOI: <https://doi.org/10.48550/arXiv.2010.16061>.
- REN, S. et al. Faster r-cnn: Towards real-time object detection with region proposal networks. **Advances in neural information processing systems**, p. 91–99, 2015.
- RETTENBERGER, L.; SAGER, F.; LELLMANN, S. Mask r-cnn outperforms u-net in instance segmentation for overlapping cells. **Current Directions in Biomedical Engineering**, v. 9, n. 1, p. 335–338, 2023. DOI: <https://doi.org/10.1515/cdbme-2023-1084>.
- RIBEIRO, R. D. et al. Impacto da inteligência artificial na medicina: Revisão bibliográfica sobre diagnóstico, tratamento e sistemas de apoio à decisão. **LUMEN ET VIRTUS**, XVI, n. XLIV, p. 217–233, 2025. DOI: <https://doi.org/10.56238/levv16n44-018>.
- RODRIGUEZ, J. S. et al. Going deeper through the gleason scoring scale: An automatic end-to-end system for histology prostate grading and cribriform pattern detection. **Computer Methods and Programs in Biomedicine**, v. 195, p. 105637, 10 01 2020. DOI: <https://doi.org/10.1016/j.cmpb.2020.105637>.

- SHAH, R. B.; ZHOU, M. **Prostate Biopsy Interpretation: An Illustrates Guide**. Berlin: Springer, 2012. DOI: <https://doi.org/10.1007/978-3-642-21369-4>.
- SILVA, F. d. A.; NASCIMENTO, A. A. d.; MEDEIROS, L. M. d. Uso da morfologia matemática na segmentação de imagens médicas para identificar o miocárdio com redes neurais convolucionais. **Proceeding Series of the Brazilian Society of Computational and Applied Mathematics**, SBMAC, v. 9, n. 1, 2022.
- SILVA, R. et al. An automated deep-learning-based approach for gleason grading of prostate cancer using histopathological images. **Scientific Reports**, Nature Publishing Group, v. 10, n. 1, p. 1–12, 2020. DOI: <https://doi.org/10.1038/s41598-020-65680-y>.
- SILVA, R. E.; CAMATA, J. J. Otimização de hiperparâmetros de redes neurais guiadas pela física em problema convectivo-difusivo. 2024. DOI: https://doi.org/10.5753/sscad_extendido.2024.244373.
- SIRINUKUNWATTANA, K. et al. Gland segmentation in colon histology images: The gland segmentation challenge contest. **Medical image analysis**, Elsevier, v. 35, p. 489–502, 2016. DOI: <https://doi.org/10.1016/j.media.2016.08.008>.
- SOCIETY, A. C. **What Is Prostate Cancer?** 2013. <<https://www.cancer.org/cancer/types/prostate-cancer/about/what-is-prostate-cancer.html>>. Retrieved 12 30, 2023.
- _____. **Facts & Figures 2023**. 2023. <<https://www.cancer.org/cancer/prostate-cancer/about/key-statistics.html>>. Retrieved from cancer.org.
- STENNING, D. et al. Morphological image analysis and its application to sunspot classification. v. 209, 01 2013. DOI: https://doi.org/10.1007/978-1-4614-3520-4_31.
- SZEGEDY, C. et al. Rethinking the inception architecture for computer vision. p. 2818–2826, 2016. DOI: <https://doi.org/10.1109/CVPR.2016.308>.
- 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. 29, 2015. DOI: <https://doi.org/10.1186/s12880-015-0068-x>.
- TensorBoard Documentation. **TensorBoard: Visualizing Learning**. 2023. <<https://www.tensorflow.org/tensorboard>>.
- TIWARI, S. et al. Nuclei detection and segmentation of histopathological images using a feature pyramidal network variant of a mask r-cnn. **Computational Intelligence and Neuroscience**, v. 2024, p. 1–13, 2024. DOI: <https://doi.org/10.1155/2024/7619280>.
- TOTH, R. J. et al. Histostitcher: An informatics software platform for reconstructing whole-mount prostate histology using the extensible imaging platform framework. **Journal of Pathology Informatics**, v. 5, n. 1, p. 8, 2014. DOI: <https://doi.org/10.4103/2153-3539.129441>.
- TRABULSI, E. J. et al. Optimum imaging strategies for advanced prostate cancer: Asco guideline. **Journal of Clinical Oncology**, American Society of Clinical Oncology, v. 38, n. 17, p. 1963–1996, Jun 2020. ISSN 1527-7755. Epub 2020 Jan 15.

URBINATE, E. F. **Métodos para melhoria do desempenho de redes neurais profundas em previsão de séries temporais**. Tese (Dissertação (Mestrado em Ciências)) — Escola Politécnica da Universidade de São Paulo, 2025.

VALIPOUR, M. et al. A survey on deep learning optimizers: A performance-oriented taxonomy and a guideline for practitioners. **ACM Computing Surveys**, v. 56, n. 1, p. 1–40, 2023. DOI: <https://doi.org/10.1145/3605775>.

VATS, S. et al. Segmenting tumor gleason pattern using generative ai and digital pathology: Use case of prostate cancer on miccai dataset. **Preprints.org**, 06 2024. DOI: <https://doi.org/10.20944/preprints202406.2019.v1>.

WANG, P. et al. PANDA: A gigapixel-level human-centric video dataset. In: **Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)**. [S.l.: s.n.], 2020. DOI: <https://doi.org/10.1109/CVPR42600.2020.00333>.

WU, Y. et al. **Detectron2**. 2019. Software available from <https://github.com/facebookresearch/detectron2>.

_____. Detectron2: A pytorch-based modular object detection library. **arXiv preprint arXiv:1904.04514**, 2019. DOI: <https://doi.org/10.1007/s10489-021-02884-x>.

ZHANG, C. et al. Understanding deep learning requires rethinking generalization. 2017. Disponível em: <<https://openreview.net/forum?id=Sy8gdB9xx>>.

ZHOU, X. et al. A context-aware deep learning model for simultaneous gland and nuclei segmentation in prostate histopathology images. **Medical image analysis**, Elsevier, v. 60, p. 101623, 2019. DOI: <https://doi.org/10.1016/j.media.2019.101623>.

ZHU, X. et al. Fusing hand-crafted and deep-learning features in a convolutional neural network model to identify prostate cancer in pathology images. **Frontiers in Oncology**, v. 12, p. 994950, 2022. DOI: <https://doi.org/10.3389/fonc.2022.994950>.

_____. A survey on transfer learning. **ACM Transactions on Intelligent Systems and Technology**, v. 14, n. 4, p. 1–28, 2023. DOI: <https://doi.org/10.1145/3569421>.