



DOUTORADO EM CIÊNCIA DA COMPUTAÇÃO
ASSOCIAÇÃO UFMA-UFPI

WORKSHOP DCCMAPI 2024 - RELATÓRIO

Classificação de tumores prostáticos baseado em exames de Biópsia com DeepLearning

Este trabalho aborda a utilização de técnicas de aprendizagem profunda para a classificação de tumores prostáticos a partir de imagens de biópsias histológicas, com o objetivo de aprimorar a precisão, velocidade e consistência no diagnóstico do câncer de próstata.

Motivação e objetivo do trabalho

O câncer de próstata é uma das principais causas de mortalidade entre homens em todo o mundo, frequentemente diagnosticado em estágios avançados da doença. O processo tradicional de avaliação da gravidade do câncer é baseado na análise visual de imagens de biópsia realizadas por patologistas, sendo um procedimento demorado, sujeito a erros humanos e inconsistências, especialmente diante de grandes volumes de amostras.

A avaliação do câncer de próstata é realizada com base no padrão arquitetônico do tumor, de acordo com a pontuação de Gleason. Esse sistema caracteriza diferentes padrões histológicos do tecido prostático, inicialmente proposto por Donald Gleason. Originalmente, os padrões histológicos eram classificados de 1 a 5, com a pontuação final sendo a soma dos dois padrões mais frequentes observados no exame. No entanto, revisões posteriores ao sistema, realizadas em 2005 e 2014, levaram a importantes mudanças na classificação.

Atualmente, os escores de 2 a 5 não são mais utilizados, e certos padrões que anteriormente recebiam escore 6 foram reclassificados como 7. Esse ajuste resultou em uma classificação mais consistente e prognóstica. Assim, a classificação foi reorganizada em cinco grupos de grau, simplificando e melhorando a comunicação sobre a gravidade da doença: **Grupo 01 - Escore Gleason 6; Grupo 02 - Escore Gleason 3 + 4 = 7; Grupo 03 - Escore Gleason 4 + 3 = 7; Grupo 04 - Escore Gleason 8; Grupo 05 - Escore Gleason 9-10.**

Essa nova classificação foi proposta para abordar limitações do sistema anterior, reduzir o medo dos pacientes e minimizar o excesso de tratamento para casos de câncer de baixo grau detectados pelo rastreamento de PSA Epstein et al. [2016].

Nesse contexto, a aplicação de técnicas de aprendizagem de máquina surge como uma solução promissora para aprimorar a precisão, velocidade e consistência na análise dessas imagens. Algoritmos avançados podem identificar padrões sutis que

Centro de Ciências Exatas e Tecnologia (CCET),

Cidade Universitária Dom Delgado,

Av. dos Portugueses, 1966,

Bacanga, São Luís, Maranhão, Brasil

Programa de Pós-Graduação em Ciência da Computação

Campus Universitário Ministro Petrônio Portella,

Núcleo de Computação de Alto Desempenho (NCAD),

Ininga, Teresina, PI, Brasil



DOUTORADO EM CIÊNCIA DA COMPUTAÇÃO ASSOCIAÇÃO UFMA-UFPI

frequentemente escapam à análise humana, fornecendo métricas objetivas e reproduzíveis para avaliar a gravidade do câncer.

O objetivo deste trabalho é aplicar técnicas de aprendizado de máquina para aprimorar a precisão, a velocidade e a consistência na análise das imagens de biópsia prostática, utilizando o sistema de pontuação de Gleason.

Atividades realizadas

Nesta seção, detalhamos as etapas conduzidas para desenvolver um modelo de classificação baseado em técnicas de aprendizagem profunda, voltado para a classificação de biópsias prostáticas de acordo com o sistema ISUP Grade, as etapas podem ser vista na Figura 1:

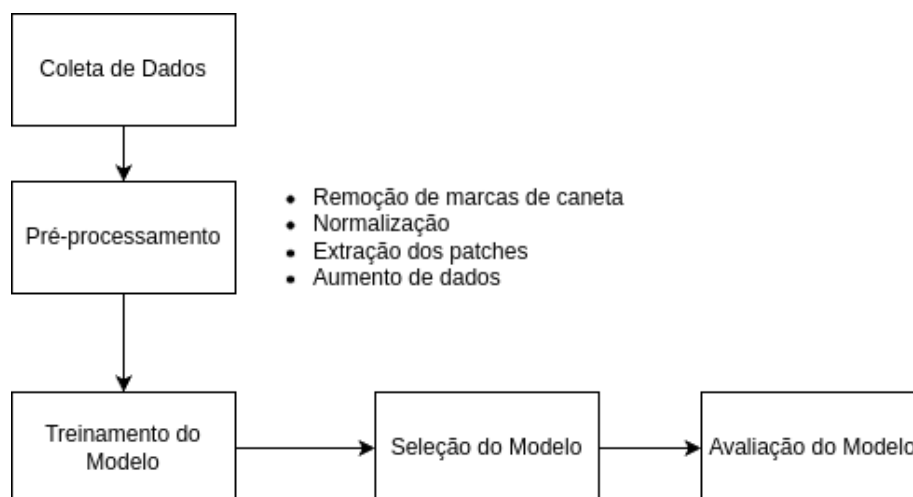


Figura 1: Fluxograma de atividades

Coleta de Dados

Para a realização dos experimentos, utilizamos o dataset da competição do Kaggle Prostate Cancer Grade Assessment (PANDA) Challenge. Este conjunto de dados contém, ao todo, 12.625 imagens digitais de lâminas histológicas que representam amostras de tecido prostático, conforme descrito em Bultenet al. [2022]. No entanto, apenas 10.616 dessas imagens estão publicamente disponíveis no repositório oficial da competição. O restante das imagens foi reservado para a validação dos algoritmos durante o desafio. Portanto, apenas as imagens públicas foram utilizadas neste trabalho.

Pré-processamento

Centro de Ciências Exatas e Tecnologia (CCET),

Cidade Universitária Dom Delgado,

Av. dos Portugueses, 1966,

Bacanga, São Luís, Maranhão, Brasil

Programa de Pós-Graduação em Ciência da Computação

Campus Universitário Ministro Petrônio Portella,

Núcleo de Computação de Alto Desempenho (NCAD),

Ininga, Teresina, PI, Brasil



DOUTORADO EM CIÊNCIA DA COMPUTAÇÃO

ASSOCIAÇÃO UFMA-UFPI

Implementamos técnicas de processamento de imagens para melhorar a qualidade visual e destacar características relevantes para a classificação. O pré-processamento incluiu:

- Remoção de imagens com marcas de caneta - No conjunto de dados, há uma série de imagens com marcas de canetas que podem atrapalhar o treinamento dos modelos;
- Normalização da coloração - Avaliamos os esquemas de cores RGB e HSV;
- Extração de patches - Devido às grandes dimensões das imagens originais, foi necessário realizar uma divisão em blocos menores (patches) de tamanho 256×256 . Esse processo visou reduzir a complexidade computacional e facilitar a análise das imagens. Para selecionar os 36 patches mais informativos de cada imagem, utilizamos um critério baseado na quantidade de informação presente em cada bloco. Após a seleção, os patches foram concatenados em uma única imagem composta com dimensões finais de 1536×1536 .
- Aumento de dados (data augmentation) - Para aumentar a robustez do modelo e evitar sobreajuste (overfitting), aplicamos Espelhamento Horizontal (Horizontal Flip) e Espelhamento Vertical (Vertical Flip), essas transformações ajudaram a ampliar a variabilidade das amostras durante o treinamento.

Treinamento, Seleção e Avaliação

Para a classificação de biópsias, exploramos diversas arquiteturas de redes neurais convolucionais, incluindo ResNet, VGG16, EfficientNet-B0 e EfficientNet-B1, além de modelos baseados em Vision Transformers. Os melhores resultados foram obtidos utilizando o modelo EfficientNet-B0 pré-treinado.

A avaliação da eficácia dos modelos foi realizada por meio do Coeficiente Kappa Ponderado, a mesma métrica de referência empregada no desafio do Kaggle e em estudos prévios aplicados ao mesmo conjunto de dados.

Nos experimentos, reservamos 15% do conjunto de dados para testes, enquanto os 85% restantes foram destinados ao treinamento e validação. Aplicamos a técnica de validação cruzada com 5 folds durante o treinamento, utilizamos 30 épocas para cada fold. Após o treinamento, selecionamos o melhor modelo entre os folds com base no desempenho nos dados de validação. Este modelo final foi utilizado para avaliar o desempenho no conjunto de testes.

Resultados e Discussão

O melhor desempenho nos experimentos foi alcançado utilizando o modelo EfficientNet-B0 pré-treinado, especificamente no terceiro fold da validação cruzada,

Centro de Ciências Exatas e Tecnologia (CCET),

Cidade Universitária Dom Delgado,

Av. dos Portugueses, 1966,

Bacanga, São Luís, Maranhão, Brasil

Programa de Pós-Graduação em Ciência da Computação

Campus Universitário Ministro Petrônio Portella,

Núcleo de Computação de Alto Desempenho (NCAD),

Ininga, Teresina, PI, Brasil



DOUTORADO EM CIÊNCIA DA COMPUTAÇÃO ASSOCIAÇÃO UFMA-UFPI

empregando imagens com esquema de cores RGB. Este experimento obteve um coeficiente Kappa Quadrático Ponderado de 92,4%, destacando-se como a abordagem mais eficaz entre todas as testadas.

Esse resultado demonstra a capacidade do EfficientNet-B0 em extrair características discriminativas relevantes do conjunto de dados, mesmo em um contexto desafiador, como o de classificação de biópsias. A combinação do modelo pré-treinado com imagens em cores (RGB) foi fundamental para capturar nuances visuais presentes nas amostras, contribuindo diretamente para o alto desempenho observado.

A métrica de Kappa Quadrático Ponderado reforça a robustez do modelo ao levar em conta não apenas as classificações corretas, mas também o grau de concordância entre as previsões e os rótulos verdadeiros, considerando a severidade dos erros.

Dentro da competição o melhor resultado alcançado foi de 94%, obtido com os dados privados da competição. Com os dados públicos, o desempenho foi de 90%.

Esse resultado foi obtido utilizando 5-fold cross-validation para o treinamento. Além disso, foi empregada a técnica de extração de patches ordenados pela quantidade de informação, o que contribuiu para melhorar a qualidade dos dados processados. A função de perda CrossEntropy foi utilizada durante o treinamento.

Para lidar com a remoção de ruídos, foi utilizada a EfficientNetB1, enquanto o modelo de avaliação final foi baseado em um ensemble formado pela combinação de EfficientNetB0 e EfficientNetB1.

Outro estudo importante sobre o mesmo conjunto de dados obteve um desempenho de 88,1% ao utilizar 80% dos dados para treinamento. Os autores propuseram duas abordagens: a primeira, um ensemble de CNNs, e a segunda, um Bag de CNN Features. Esse resultado foi alcançado utilizando a primeira proposta, o ensemble de modelos CNN. Bhattacharyya et al. [2024].

Contextualização do ambiente de experimentação

Neste trabalho, diversas tecnologias e bibliotecas foram empregadas para desenvolver, treinar e avaliar o modelo de classificação de câncer de próstata a partir de Imagens de Lâmina Inteira (WSIs). A seguir estão as principais ferramentas e abordagens utilizadas: PyTorch - utilizado como a principal framework para o desenvolvimento e treinamento do modelo de rede neural convolucional (CNN); EfficientNet - como modelo de aprendizagem utilizado no treinamento e avaliação; Pandas - para manipulação, análise e pré-processamento dos dados; Scikit-learn - para calcular a métrica de Kappa; skimage para leitura das imagens; Albumentations - para pré-processamento e *data augmentation*.

Centro de Ciências Exatas e Tecnologia (CCET),

Cidade Universitária Dom Delgado,

Av. dos Portugueses, 1966,

Bacanga, São Luís, Maranhão, Brasil

Programa de Pós-Graduação em Ciência da Computação

Campus Universitário Ministro Petrônio Portella,

Núcleo de Computação de Alto Desempenho (NCAD),

Ininga, Teresina, PI, Brasil



DOUTORADO EM CIÊNCIA DA COMPUTAÇÃO
ASSOCIAÇÃO UFMA-UFPI

Cronograma para os próximos 12 meses

Atividades / Mês	01	02	03	04	05	06	07	08	09	10	11	12
Combinar classificadores baseado na dissimilaridade	X	X	X	X								
Publicar artigo	X	X	X	X	X	X						
Iniciar a escrita da tese							X	X	X	X	X	X
Workshop II			X	X	X	X	X					
Workshop III								X	X	X	X	X

Principais veículos para publicação na área

IEEE Transactions on Medical Imaging; Nature Medicine; SBCAS; IEEE Symposium on Computers and Communications; IEEE International Conference on E-health Networking, Application & Services;

Outros aspectos relevantes

Na execução deste trabalho, diversos desafios se mostraram significativos e continuam a ser obstáculos importantes. O primeiro deles refere-se ao fato de que a base de dados utilizada foi disponibilizada na plataforma Kaggle em formato de competição, mas com os dados de classificação ocultos. Isso torna a comparação direta com os trabalhos de referência algo complexo, uma vez que não é possível avaliar o desempenho dos modelos com os mesmos dados de teste.

Outro desafio considerável está relacionado ao tamanho das imagens. A base original ocupa aproximadamente 400 GB, o que torna o processo de treinamento extremamente custoso em termos de recursos computacionais. Para mitigar esse problema, foram necessárias técnicas de pré-processamento, como a redução da dimensionalidade ou o descarte de informações que, embora possam ser relevantes, tornam o treinamento inviável em sua totalidade.

Além disso, a base de dados apresenta ruídos que podem afetar negativamente o desempenho dos modelos. Esses ruídos incluem diagnósticos inconclusivos e anotações feitas com caneta nas lâminas, o que exige um cuidado adicional no processo de limpeza e preparação dos dados.

Centro de Ciências Exatas e Tecnologia (CCET),

Cidade Universitária Dom Delgado,

Av. dos Portugueses, 1966,

Bacanga, São Luís, Maranhão, Brasil

Programa de Pós-Graduação em Ciência da Computação

Campus Universitário Ministro Petrônio Portella,

Núcleo de Computação de Alto Desempenho (NCAD),

Ininga, Teresina, PI, Brasil



DOUTORADO EM CIÊNCIA DA COMPUTAÇÃO
ASSOCIAÇÃO UFMA-UFPI

Ademais, a expectativa de aplicar técnicas de combinação de modelos representa uma oportunidade interessante. Acredito que treinar múltiplos classificadores e combinar aqueles que exibem comportamentos dissimilares pode gerar resultados inovadores, ainda não explorados na literatura, oferecendo uma abordagem mais robusta para a classificação de biópsias prostáticas.

Referências

Riddhasree Bhattacharyya, Paromita Roy, Sugata Banerji, and Sushmita Mitra. Efficient grading of prostate cancer WSI with deep learning. In John E. Tomaszewski and Aaron D. Ward, editors, *Medical Imaging 2024: Digital and Computational Pathology*, volume 12933, page 1293305. International Society for Optics and Photonics, SPIE, 2024. doi: 10.1117/12.3006835.

Wouter Bulten, Kimmo Kartasalo, Po-Hsuan Cameron Chen, Peter Ström, Hans Pinckaers, Kunal Nagpal, Yuannan Cai, David F Steiner, Hester Van Boven, Robert Vink, et al. Artificial intelligence for diagnosis and gleason grading of prostate cancer: the panda challenge. *Nature medicine*, 28(1):154–163, 2022.

Jonathan I. Epstein, Michael J. Zelefsky, Daniel D. Sjoberg, Joel B. Nelson, Lars Egevad, Cristina Magi-Galluzzi, Andrew J. Vickers, Anil V. Parwani, Victor E. Reuter, Samson W. Fine, James A. Eastham, Peter Wiklund, Misop Han, Chandana A. Reddy, Jay P. Ciezki, Tommy Nyberg, and Eric A. Klein. A contemporary prostate cancer grading system: A validated alternative to the gleason score. *European Urology*, 69(3):428–435, 2016. doi: <https://doi.org/10.1016/j.eururo.2015.06.046>.

Teresina (PI), 07 de dezembro de 2024

Woshington Valdeci de Sousa Rodrigues
(Discente)

Antonio Oseas de Carvalho Filho
(Orientador)

Centro de Ciências Exatas e Tecnologia (CCET),

Cidade Universitária Dom Delgado,

Av. dos Portugueses, 1966,

Bacanga, São Luís, Maranhão, Brasil

Programa de Pós-Graduação em Ciência da Computação

Campus Universitário Ministro Petrônio Portella,

Núcleo de Computação de Alto Desempenho (NCAD),

Ininga, Teresina, PI, Brasil