

Efetividade de Técnicas de Atenção Espacial em Redes Neurais Profundas para a Predição de Resposta Patológica Completa (pCR) aplicada ao Câncer de Mama

^{1st} Vinícius Alves de Faria Resende
Departamento de Ciência da Computação
Universidade Federal de Minas Gerais (UFMG)
Belo Horizonte, Brasil
viniciusfariaresende@ufmg.br

^{2nd} George Teodoro (Orientador)
Departamento de Ciência da Computação
Universidade Federal de Minas Gerais (UFMG)
Belo Horizonte, Brasil
george@dcc.ufmg.br

Abstract—This project developed an AI model to predict how well Triple-Negative Breast Cancer will respond to chemotherapy by analyzing medical images. The study found that while a model using only standard H&E stained images works, its performance dramatically improves when combined with images of specific biomarkers (Ki-67 and PHH3). The most successful approach used an "attention map" created from these biomarkers to help the AI focus on the most critical regions of the tissue, achieving a high prediction accuracy (AUC of 0.96). Ultimately, the work confirms that this combined strategy is highly effective and suggests that "virtual staining"—using AI to computationally create these valuable biomarker images—is a promising direction for future research to make this technology more accessible and scalable.

Index Terms—pCR, virtual staining, attention maps, biomarkers, breast cancer, neural networks.

Resumo—Este projeto desenvolveu um modelo de IA para prever quão bem o Câncer de Mama Triplo-Negativo responderá à quimioterapia, analisando imagens médicas. O estudo descobriu que, embora um modelo usando apenas imagens padrão coradas com H&E funcione, seu desempenho melhora drasticamente quando combinado com imagens de biomarcadores específicos (Ki-67 e PHH3). A abordagem de maior sucesso usou um "mapa de atenção" criado a partir desses biomarcadores para ajudar a IA a focar nas regiões mais críticas do tecido, alcançando uma alta precisão de predição (AUC de 0.96). Por fim, o trabalho confirma que essa estratégia combinada é altamente eficaz e sugere que a "coloração virtual" (*virtual staining*) — o uso de IA para criar computacionalmente essas valiosas imagens de biomarcadores — é uma direção promissora para pesquisas futuras, a fim de tornar essa tecnologia mais acessível e escalável.

Index Terms—pCR, *virtual staining*, mapas de atenção, biomarcadores, câncer de mama, redes neurais.

I. INTRODUÇÃO

A Organização Mundial da Saúde (OMS) caracteriza o câncer como um grupo amplo de doenças que pode se originar em praticamente qualquer órgão ou tecido do

corpo, resultante de um crescimento descontrolado de células anômalas. O estágio avançado da doença, conhecido como metástase, ocorre quando essas células ultrapassam seus limites habituais, invadindo tecidos adjacentes e se disseminando para outros órgãos, sendo este processo uma das principais causas de mortalidade associada ao câncer. Em 2018, o câncer foi a segunda maior causa de morte mundial, contabilizando aproximadamente 9,6 milhões de óbitos, o que equivale a 1 em cada 6 mortes. Entre os tipos mais comuns de cânceres em mulheres destacam-se o de mama, colorretal, pulmonar, cervical e de tireoide¹.

Segundo a Organização Mundial da Saúde (OMS)², o câncer de mama caracteriza-se pelo crescimento anômalo de células nos ductos de leite e/ou lóbulos produtores de leite da mama, podendo disseminar-se para o tecido mamário adjacente (invasão) e formar tumores que resultam em nódulos ou espessamentos perceptíveis. Este tipo de câncer representa a principal causa de mortalidade por câncer entre as mulheres globalmente, com aproximadamente 1,7 milhões de novos casos diagnosticados anualmente [1].

Os últimos anos presenciaram grandes avanços na captura de imagens histopatológicas, o que motivou o desenvolvimento de diversas ferramentas na área de diagnósticos assistidos por computadores [2], em especial, a predição da resposta patológica completa (pCR) para quimioterapia neoadjuvante (NAC). Nesse cenário, a pCR é definida como a ausência de todos os sinais de câncer, principalmente a ausência de células cancerígenas em imagens patológicas de amostras de tecido dissecadas durante a cirurgia. No âmbito oncológico, por exemplo, houve grande progresso em diversas áreas, incluindo o câncer de mama triplo negativo (TNBC). Esse subtipo de câncer de mama é

¹<https://www.who.int/health-topics/cancer>

²<https://www.who.int/news-room/fact-sheets/detail/breast-cancer>

agressivo e não expressa os receptores de estrogênio (ER), progesterona (PR) e o receptor 2 do fator de crescimento epidérmico humano (HER2) [3]. Isso o torna menos responsivo à terapia hormonal e às terapias-alvo, deixando a quimioterapia como a principal opção de tratamento.

Uma característica do câncer de mama é ser uma doença biologicamente heterogênea. Ele é classificado clinicamente pela expressão do receptor de estrogênio (ER), receptor de progesterona (PR) e proteína HER2. Pacientes com ER+ (um subtipo do câncer) tipicamente são tratadas com quimioterapia após a cirurgia, mas apenas uma parte deste grupo se beneficia de tal procedimento.

Por conseguinte, nem todos os pacientes com TNBC respondem bem à quimioterapia. Na verdade, apenas uma parcela dos pacientes alcança a PCR, que é definida como a ausência de células cancerosas invasivas no tecido mamário após a quimioterapia [4]. Dessa forma, a previsão precoce da PCR é crucial para o planejamento de tratamento ideal em pacientes com TNBC. Uma vez que pacientes que são previstos para alcançar a PCR podem ser capazes de evitar procedimentos cirúrgicos mais agressivos, enquanto aqueles que são previstos para não alcançar a PCR podem se beneficiar de terapias adicionais.

O avanço da área de Inteligência Artificial (IA), mais especificamente da subárea de Aprendizado Profundo (do inglês, *Deep Learning (DL)*) permitiu que imagens médicas pudessem ser analisadas quantitativamente de forma automatizada. Sendo assim, modelos de DL aplicados a histologia são capazes de medir o nível de severidade de tumores, expressões gênicas e até mesmo alterações genéticas [5], [6]. Outras frentes consistem no treinamento de modelos em imagens com biomarcadores H&E para predição da pontuação de recorrência baseada em expressões gênicas do câncer de mama [1]. Essas imagens podem ser usadas para identificar padrões que estão associados à PCR, isso pode ser feito por intermédio da identificação de informações como: o grau de infiltração tumoral; a presença de células cancerosas em estado de repouso; e a atividade da resposta imune. Alguns trabalhos recentes demonstraram ser possível prever a resposta pCR baseando-se em dados de tecidos H&E [7], [8] com certa qualidade, mostrando que este é um caminho promissor.

O estado da arte atual para a previsão de PCR em TNBC utiliza imagens histopatológicas coradas com Ki-67 e PHH3 para alimentar modelos de aprendizado profundo. No entanto, a obtenção de imagens com esses corantes não é trivial, uma vez que os mesmos são caros e muitas vezes não são encontrados em qualquer laboratório. Além disso, a acurácia da previsão de PCR utilizando imagens histopatológicas coradas ainda tem limitações. Uma vez que previsões erradas podem gerar impactos negativos na vida das pessoas, a falta de modelos com alta acurácia também limita o uso desses modelos.

Assim, o objetivo deste projeto é, principalmente, ser suporte para a elaboração de etapas mais avançadas com o intuito de avançar o estado da arte ao superar as limitações

supracitadas. Isso será feito, por meio da avaliação do impacto da utilização de cada etapa das técnicas de atenção espacial em [7], mensurando sistematicamente qual o ganho de performance do modelo ao progredir na sofisticação e complexidade da entrada do modelo, com o propósito de confirmar a hipótese de que a abordagem guiada à atenção espacial apresentará resultados superiores.

Confirmada a hipótese, pode-se então progredir para trabalhos futuros, onde explorar-se-á a utilização de técnicas de *virtual staining* (coloração por IA) [9] como estratégia para confeccionar mapas de atenção sofisticados a partir das imagens histopatológicas H&E. Além disso habilita-se assim, a possibilidade de se combinar estes dados artificialmente gerados com os originais H&E para o desenvolvimento de metodologias de atenção mais robustos [10]. Dessa forma, pode-se integrar a geração de imagens histopatológicas artificialmente coradas, a conversão da predição pCR apresentada no artigo [7]. Por conseguinte, verificar a hipótese de que será possível obter resultados significativos na predição da resposta pCR mesmo com a utilização das imagens coradas artificialmente.

II. REFERENCIAL TEÓRICO

Aprendizado Profundo: O aprendizado profundo (DL do inglês *Deep Learning*) é um subconjunto do aprendizado de máquina que utiliza redes neurais artificiais para aprender padrões em dados. O DL permitiu que múltiplas camadas de processamento pudessem ser concatenadas para criar modelos computacionais robustos capazes de aprender representações de dados mais complexas. O aprendizado profundo desvenda estruturas intrincadas em grandes conjuntos de dados através do algoritmo de backpropagation³, guiando a máquina a ajustar seus pesos internos para melhor representar os dados em cada camada e melhorar a performance da rede. Dessa forma, alguns modelos ganharam popularidade após obterem altas acurácias na execução de tarefas específicas como por exemplo as ResNets⁴ baseadas em redes neurais convolucionais profundas que se destacam na segmentação de imagens e as GANs (Redes Adversariais Generativas) [11] que são um tipo de modelo de aprendizado de máquina que aprende representações profundas sem dados de treinamento extensivamente anotados.

Aprendizado Profundo para Patologia Digital: No contexto da patologia digital, o aprendizado profundo tem demonstrado resultados promissores em uma ampla gama de tarefas, incluindo análise de imagens histopatológicas para detecção de células cancerígenas [12], classificação de tumores [13] e previsão de prognóstico [14]. A sua capacidade de aprender padrões complexos a partir de grandes conjuntos de dados sem a necessidade de recursos pré-fabricados o torna uma ferramenta poderosa para

³<https://www.geeksforgeeks.org/machine-learning/backpropagation-in-neural-network/>

⁴<https://www.geeksforgeeks.org/residual-networks-resnet-deep-learning/>

análise histopatológica. Além disso, ele pode ajudar a padronizar e automatizar essas tarefas de análise, reduzindo a variabilidade interobservador e melhorando a eficiência diagnóstica. Já no contexto do câncer de mama, o padrão-ouro para o diagnóstico e classificação ainda é feito através da análise de imagens histopatológicas, no entanto, a realização dessa tarefa de forma tradicional é demorada e propensa a erros, pois exige uma inspeção visual detalhada e interpretação de uma amostra. A patologia digital visa aliviar os problemas das abordagens tradicionais de exame, fornecendo ferramentas computadorizadas auxiliares que analisam quantitativamente imagens histopatológicas digitalizadas.

Biomarcadores: Biomarcadores são moléculas ou genes que podem ser usados para identificar, diagnosticar ou monitorar vários tipos de doença. Eles podem ser encontrados em fluidos corporais, tecidos ou células. Os biomarcadores são importantes para a medicina porque podem fornecer informações sobre a saúde de um indivíduo ou o progresso de uma doença. A análise histológica de marcadores em tecidos humanos desempenha um papel crucial na identificação de doenças e na avaliação patológica de pacientes. A etapa de marcação dos tecidos fornece contraste e coloração para estruturas específicas, sendo a técnica com hematoxilina e eosina (H&E) a mais comum e amplamente empregada. Outros métodos de coloração incluem o Tricrômico de Masson (*Masson's Trichrome (MT)*) e o Ácido Periódico de Schiff (*Periodic Acid-Schiff (PAS)*), além de uma variedade de colorantes especiais que não utilizam o H&E. A mudança para um diagnóstico mais precoce do câncer de mama devido a métodos de imagem aprimorados e programas de rastreamento destaca a necessidade de novos fatores e combinações de biomarcadores para quantificar o risco residual dos pacientes e indicar o valor potencial de estratégias de tratamento adicionais. Os marcadores de proliferação Ki-67 e PHH3 mostraram associações significativas com alto grau histológico e negatividade de receptores hormonais [15] ganhando relevância em várias aplicações na terapia neoadjuvante, além de seu valor prognóstico moderado.

Transformação de Marcadores: A preparação dos tecidos marcados supracitados é um procedimento que demanda tempo, altos custos e infraestrutura laboratorial avançada. Os marcadores especiais, em particular, exigem maior complexidade de preparo e representam custos procedurais ainda mais elevados [16]. Tendo ciência das limitações do fluxo de obtenção dos tecidos marcados com os colorantes, estudos recentes desenvolveram técnicas computacionais de marcação conhecidas como *Virtual Staining (VS)*. A partir dessa técnica, em [16] foi proposta a aplicação de modelos de DL para transformar os tecidos já marcados com um colorante em outros tipos de colorantes. O mapeamento entre os diferentes tipos de marcadores seria aprendido pelo modelo após treiná-lo com imagens de entrada utilizando um marcador e tendo uma imagem resultante de referência usando outro marcador aplicado

ao mesmo segmento de tecido. Como não é possível fazer a marcação histoquímica usando dois marcadores diferentes para o mesmo segmento de tecido, têm-se a possibilidade de utilizar a técnica de VS para gerar diferentes marcadores para a mesma imagem.

Aplicação de Técnicas Generativas na Síntese de Imagens Histológicas: Como supracitado, as técnicas de VS tem um potencial ainda pouco explorado no contexto da análise histológica, em especial ao se tratar de associar essa estratégia com modelos que se beneficiam de imagens histológicas com biomarcadores diversos. Portanto, recentemente desenvolveram-se trabalhos os quais exploraram a aplicação de diferentes técnicas generativas (*CycleGANs* e *Difussion Models*) para a síntese de imagens histológicas simulando a aplicação de biomarcadores (*Virtual Staining*), isso com o propósito de avaliar o desempenho das diferentes abordagens [9]. Complementarmente, em [17] os autores realizam uma avaliação da performance do modelo de VS por meio da utilização destas imagens geradas artificialmente como entrada para um modelo de classificação que distingue entre células ativas e mortas. Neste mesmo artigo, os autores discutem o potencial de se aplicar as técnicas de VS em *Clinical Workflows* de forma a potencializar os resultados, por meio de métodos de atenção mais robustos por exemplo. Considerando que a técnica de VS é um tipo de Síntese de Imagem, cria-se uma expectativa sobre a possibilidade de que modelos de difusão possam performar melhor do que os tradicionais GANs [18].

Mecanismo de Atenção Espacial: Uma forma mais robusta de mecanismo de atenção encontrada na literatura [7] é implementada para a predição da Resposta Patológica Completa (pCR) à quimioterapia neoadjuvante (NAC) em casos de Câncer de Mama Triplo Negativo (TNBC). O artigo propõe a incorporação de informação de três biomarcadores distintos: H&E, Ki-67 e PHH3 nas imagens a fim de agregar conhecimento do domínio médico na imagem e melhorar o desempenho do modelo de predição. Após a obtenção das áreas marcadas com tumor utilizando a coloração H&E, são encontradas as células tumorais positivas utilizando o biomarcadores Ki-67 e PHH3. Uma vez que as células tumorais foram anotadas pelos biomarcadores, é feita uma estimativa de densidade de *kernel* utilizando o método não-paramétrico *Kernel Density Estimation (KDE)*. O mecanismo de atenção é implementado como a ponderação das amostras marcadas com H&E pelos biomarcadores Ki-67 e PHH3. A imagem de tecido juntamente com o mapa é utilizado como entrada para treinar um modelo de DL, utilizando uma *ResNet-34* como *backbone* para predizer os valores de pCR. As amostras são analisadas a nível de *patch* (subdivisão), região e paciente. Métricas como acurácia, *Area Under Curve (AUC)*, precisão e *recall* foram comparadas para a imagem sem mecanismo de atenção, com atenção no tumor anotado pelo marcador Ki-67, com atenção no tumor anotado pelo PHH3 e com a combinação de atenção de

ambos os marcadores. O resultado mostrou que o modelo com o mecanismo de atenção combinada obteve o melhor desempenho em todas as métricas exceto a precisão a nível de *patch*. Conclui-se com esse artigo que a combinação de marcadores que realçam diferentes características do tecido são capazes, não só de agregar informação, mas de melhorar o desempenho do modelo de predição de pCR de uma forma geral, instigando a validação desta melhora também para o caso de imagens histopatológicas geradas artificialmente (*Virtual Staining*).

III. CONTRIBUIÇÃO

Nas sessões seguintes, será apresentado o desenvolvimento do trabalho, almejando responder as seguintes perguntas:

- O uso de imagens histopatológicas coradas dos tipos Ki-67, PHH3 e H&E fornece melhoria de acurácia para o modelo?
- É razoável utilizar dados gerados a partir de *virtual staining* para previsão da resposta pCR?

O desenvolvimento foi dividido nas sessões de IV à VIII, onde para cada uma especifica-se o contexto, trabalho desenvolvido e resultados obtidos. Dessa forma, cada uma das etapas entre IV e VII detalha o que foi realizado e validado, já a etapa VIII especifica o trabalho confeccionado em paralelo por outros pesquisadores do departamento (onde o referido autor deste trabalho teve participação ativa), que têm profunda ligação com este trabalho, conforme será definido na sessão de Conclusão e Trabalhos Futuros (IX).

Portanto, para responder a primeira pergunta usaremos o conteúdo apresentado de IV à VII, e para a segunda pergunta usaremos o conteúdo da sessão VIII de forma que possa-se dissertar sobre as respostas e suas justificativas na etapa final deste trabalho.

IV. DETECÇÃO DE CÉLULAS TUMOROSAS

Com o objetivo de realizar a confecção de uma pipeline capaz de prever a resposta pCR, seguiu-se a abordagem apresentada em [7] onde utiliza-se uma rede capaz de fazer a detecção e segmentação de células tumorosas. Esta informação é então utilizada para compor os dados de entrada para o algoritmo responsável pela geração de mapas de atenção. Portanto, a obtenção dessa informação é parte crucial da pipeline. A Figura 1 apresenta em alto nível como essa detecção de células tumorosas é parte da pipeline responsável pela geração dos mapas de atenção.

A. Abordagem Implementada

Inicialmente, fez-se contato com os autores do artigo referenciado [7] para obter mais informações sobre a etapa de segmentação, haja vista que nos materiais do artigo esta rede não está disponível. A resposta obtida foi a de que a rede foi construída com base na rede de segmentação Detectron2⁵. Existiu-se uma tentativa para a confecção de

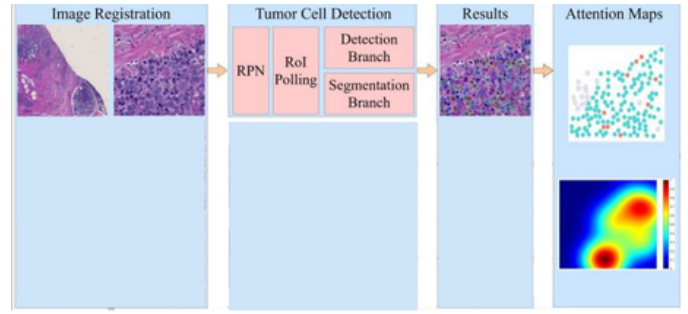


Figura 1. Segmentação de Células Tumorosas

uma rede equivalente de segmentação de células tumorosas utilizando esta rede como base, contudo a rede é altamente complexa e precisa de um grande volume de dados para ser adaptada para o contexto desejado. Nesse contexto, uma nova tentativa de contato com os autores do artigo foi efetuada, com o objetivo de obter a base de dados utilizada para treinar o modelo de segmentação, para tal, faz-se necessário um volume significativo de imagens histopatológicas anotadas com as células tumorosas, em especial, imagens de tecido mamário. Todavia, a tentativa foi frustrada, os autores não foram capazes de fornecer os dados usados para o treinamento da rede, tampouco forneceram a rede já treinada para replicação dos resultados do artigo. Este fato foi inesperado e culminou em uma necessidade de adaptação desta etapa.

Defronte com a situação supracitada, emergiu-se a necessidade de contornar o problema, três possíveis abordagens foram consideradas para tal: obtenção em bases públicas dos dados necessários para o treinamento da rede Detectron2, troca da estratégia de segmentação, troca da rede de segmentação em conjunto com a obtenção de dados que fossem compatíveis com a rede. Todas as três abordagens foram exploradas neste trabalho, a fim de determinar qual a viabilidade de cada uma e como o resultado final será impactado a depender da abordagem.

A primeira abordagem era preferível inicialmente, uma vez que poderia manter-se a mesma rede utilizada no artigo, aumentando as chances de obter um resultado similar. Contudo, esta tarefa se mostrou complexa e não produtiva, mesmo com uma exploração em bases públicas e/ou a transformação de *labels* para o padrão necessário. Assim, os resultados dessa abordagem foram insatisfatórios, levando a adoção de uma metodologia alternativa.

A segunda abordagem é uma troca de estratégia para a segmentação em si, i.e., não necessariamente fazer a segmentação das células tumorosas propriamente ditas. Uma alternativa que foi explorada é a da segmentação de áreas de interesse, essa abordagem mostrou-se promissora uma vez que a segmentação em si é utilizada apenas para direcionar a criação do mapa de atenção, conforme é explicitado na Figura 1, ou seja, contanto que a nova estratégia seja capaz de direcionar o algoritmo de geração dos mapas sem prejuízos de performance da pipeline como um todo, a

⁵<https://detectron2.readthedocs.io/en/latest>

estratégia é válida. No artigo [19] é apresentada uma rede capaz de segmentar regiões das imagens histopatológicas (e.g. tumor, estroma, necrose, linfócito), essa informação poderia ser utilizada em conjunto com a segmentação de células tumorosas e detecção de biomarcadores (V) para dar pesos positivos ou negativos no mapa de atenção gerado a depender da classe da região segmentada. Para este trabalho há integração desta abordagem na pipeline principal, podendo ser revisitada em trabalhos futuros.

A terceira abordagem é a que foi implementada e integrada na pipeline, porém sua eficiência verificada com a pipeline implementada como um todo foi insatisfatória. A implementação consistiu em um algoritmo de segmentação capaz de produzir máscaras binárias para as células consideradas tumorosas nas imagens histopatológicas do tipo (H&E) de tecido mamário. Para a obtenção dos dados de treinamento, foi feita uma exploração e foi encontrada a base de dados nomeada Breast Cancer Cell Segmentation⁶ na plataforma Kaggle⁷, essa base é anotada e foi capaz de guiar o treinamento de um modelo de segmentação.

A primeira tentativa com a nova base, foi de treinar um modelo Detectron2, o que exigiu uma transformação das máscaras binárias que eram dadas como *labels* da base, em descritores no formato COCO (Common Objects in Context)⁸, para tal desenvolveu-se um algoritmo capaz de usar o contraste para separar cada região marcada, e então com a adoção de uma estratégia de Convex Hull, determina-se o perímetro de cada célula. Mesmo com a transformação dos dados, não houve possibilidade de treinar a rede Detectron2 de forma satisfatória, justamente pela pouca quantidade de dados, a precisão do modelo ficou muito aquém do esperado.

Partiu-se então para a exploração de uma nova estratégia para a implementação de um modelo de segmentação, um modelo que fosse mais simples para que pudesse ser treinado com a menor quantidade de dados e ainda conseguisse um resultado satisfatório. Então, desenvolveu-se um modelo de segmentação baseado em uma U-Net [20], com duas camadas convolucionais. O modelo é simples porém foi capaz de lidar com o problema, mostrando resultados promissores inicialmente nos dados de validação, esses resultados não se mantiveram na avaliação da pipeline completa.

Por fim, ainda na etapa de segmentação, foi necessário implementar um algoritmo capaz de separar cada uma das células segmentadas em uma máscara binária individual que identifica apenas uma célula. Essa etapa é necessária dado o padrão de entrada esperado pela etapa de geração dos Mapas de Atenção (VI), etapa esta que foi adaptada para lidar com a ausência da segmentação celular no *pipeline* final. Para realizar essa separação, aproveitou-se o algoritmo desenvolvido para a conversão das máscaras

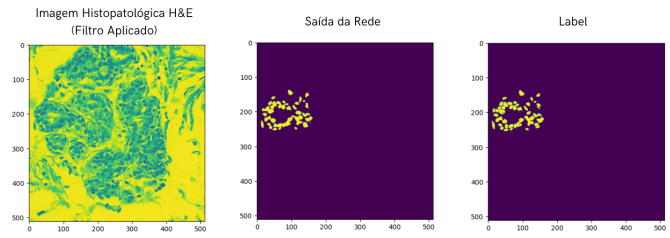


Figura 2. Resultados da Rede de Segmentação

binárias em formato COCO, alterando apenas a segunda etapa que, ao invés de delimitar o perímetro da célula, apenas salva uma nova máscara individual contendo apenas uma célula.

B. Resultados Obtidos

Conforme foi supracitado, a etapa de segmentação não apresentou resultados satisfatórios, uma vez que seu desempenho e aderência ao *pipeline* original do artigo [7] foi abaixo das expectativas. Porém, com o trabalho desenvolvido, foi possível avaliar sua efetividade aplicada às etapas seguintes.

Na Figura 2, temos um exemplo de saída da rede, onde vemos um resultado bem aderente com a realidade. O problema é que existiam muitos cenários mais complexos onde não observava-se tamanha eficiência, especialmente em dados fora do conjunto de treino/validação que podem ter sido obtidas por equipamentos diferentes, o que resulta em uma diferença de *magnification*⁹, nível de detalhes, foco de luz entre outros.

Posteriormente à obtenção da saída exemplificada acima, tem-se o processo de extração de cada um das células identificadas em uma máscara individual para cada uma, i.e., a imagem acima será transformada em uma pilha de imagens, cada uma contendo uma célula identificada.

Sendo capaz de executar o processo descrito acima, partiu-se então para a criação de um *script* capaz de executar em *batch* o processamento de todas as imagens de treino disponibilizadas no artigo [7], ou seja, era preciso de fato executar a segmentação com os dados reais. Esse processo foi feito em máquinas dedicadas com alta capacidade de processamento paralelo (GPUs do tipo RTX 4090), primeiramente fez-se o processamento das imagens originais e o salvamento do resultado intermediário que eram as máscaras binárias contendo todas as células identificadas. Em um segundo momento, um novo *script* foi desenvolvido para realizar o processamento das máscaras binárias, separando-as individualmente por células e gerando o resultado necessário para as próximas etapas.

Por conseguinte, o resultado obtido ao final deste processo, foi a confecção das máscaras binárias individualmente segmentadas por células tumorosas em uma pilha, isso para todas as imagens disponibilizadas pelos autores do artigo [7]. Por infortúnio, essas imagens não foram

⁶<https://www.kaggle.com/datasets/andrewmvd/breast-cancer-cell-segmentation/data>

⁷<https://www.kaggle.com/>

⁸<https://cocodataset.org/>

⁹Magnification refere-se ao nível de ampliação da imagem

utilizadas ao fim do projeto, isso decorreu do fato de que percebeu-se uma ineficiência no processo de segmentação de células tumorosas por falta de dados anotados compatíveis. Portanto, mesmo tendo-se desenvolvido o trabalho supracitado, a rede foi desanexada da *pipeline* final, conforme mostrarão os diagramas seguintes. Contudo avalia-se a possibilidade de obter uma rede mais robusta que possa ser anexada à pipeline em trabalhos futuros.

V. DETECÇÃO DE BIOMARCADORES

Mantendo-se no objetivo de realizar a confecção de uma pipeline capaz de prever a resposta *pCR*, e seguindo o que foi apresentado em [7], paralelamente à segmentação de células tumorosas, também é feito um processo de criação de máscaras binárias por meio da detecção de Biomarcadores, mais especificamente o Ki-67 e o PHH3. Esta informação é então utilizada para compor os dados de entrada para o algoritmo responsável pela geração de mapas de atenção. Portanto, vê-se também como parte crucial da pipeline. A Figura 3 apresenta em alto nível como essa informação compõe-se com a segmentação para a geração dos mapas de atenção.

A. Abordagem Implementada

Diferentemente da etapa anterior, desta vez foi possível obter os dados e códigos necessários para replicar esta etapa do artigo referenciado. Isso fez com que o processo tivesse um comportamento melhor seguindo as expectativas ao reproduzir o resultado da pipeline.

Dessa forma, a primeira etapa caracterizou-se pela análise e entendimento do código dos autores, de forma que fosse possível submeter as imagens contendo os biomarcadores (Ki-67 e PHH3) e obter a máscara binária como saída. O processo não foi tão desimpedido quanto esperava-se inicialmente, isso porque o código disponibilizado não se adaptava facilmente aos dados fornecidos pelos autores, ademais, o código não seguia as melhores práticas de engenharia de software, o que prejudicou a compreensão e retardou o processo.

Estando de frente com a situação supracitada, decidiu-se então pela adoção de um *framework* padronizado para lidar com a segmentação, a detecção de biomarcadores e também toda a pipeline. Escolheu-se, então, o *framework* de ciência de dados *open-source* conhecido como Kedro¹⁰, o que possibilitou a reestruturação do código em uma arquitetura padronizada e de maior compreensibilidade.

Não obstante, a adoção do *framework* por si só não resolveu as questões relacionadas à falta de compreensibilidade intrínseca das funções que compunham o algoritmo, tampouco resolveram as questões de compatibilidade com os dados fornecidos. Para solucionar essas questões, um processo de reescrita das funções mantendo suas funcionalidades foi necessário [21], de forma que fosse mais fácil de compreender o código e adaptá-lo às entradas e saídas necessárias.

Após a execução dos passos mencionados, o processo foi efetivamente integrado à pipeline, produzindo resultados condizentes com os descritos no artigo. O processo em si é uma Deconvolução de Cores, que é um método usado em microscopia diagnóstica de campo claro para transformar imagens coloridas de múltiplas amostras biológicas coradas (*stained*) em imagens que representam as concentrações de manchas [22]. Esse processo é determinístico¹¹, são aplicadas diversas transformações, remoções de ruídos, segmentações, ajuste de intensidade entre outros processos para a obtenção do resultado.

B. Resultados Obtidos

Essa etapa teve resultados extremamente promissores, sendo que a lógica de Deconvolução de Cores do artigo referenciado pôde ser inteiramente adaptada. Isso fez com que fosse possível realizar o processamento das imagens obtendo resultados precisos.

Especialmente para esta etapa, existe uma certa "facilidade" para verificar a assertividade do processo de detecção dos biomarcadores, mesmo sem experiência em micropatologia, isso pode ser visto na Figura 4, onde é possível notar na imagem da esquerda as células demarcadas pelo biomarcador (pontos escuros), e vemos que esses mesmos pontos estão sendo destacados corretamente na imagem da direita.

Tendo avaliado o resultado para alguns dos dados disponíveis e verificado sua assertividade, prosseguiu-se para a configuração da pipeline Kedro para que executasse o processamento em todas as imagens de biomarcadores disponíveis, armazenando os resultados para serem usados em etapas posteriores. Mais uma vez foi necessário que esse processo fosse feito em máquinas dedicadas com alta capacidade de processamento paralelo (GPUs do tipo RTX 4090).

Por fim, o resultado final deste processo, foi a confecção das máscaras binárias que identificam as células detectadas pelos biomarcadores Ki-67 e PHH3, isso para todas as imagens disponibilizadas pelos autores do artigo [7], possibilitando a progressão para as próximas etapas documentadas abaixo.

VI. MAPAS DE ATENÇÃO

Prosseguindo na confecção da pipeline de predição da resposta *pCR*, chega-se a hora de combinar os resultados obtidos nas duas seções anteriores (IV e V) para a criação dos mapas de atenção. Os mapas de atenção tem desempenhado um papel fundamental na confecção de redes neurais modernas, potencializando suas capacidades de maneira abrupta [10]. No contexto deste projeto, o mapa de atenção será combinado com as imagens originais antes de ser submetido a uma Rede Neural Convolutiva de predição, a Figura 5 ilustra de maneira geral qual o processo decorrido até esse ponto, onde combina-se a

¹¹Não envolve técnicas de aprendizado de máquina, inteligência artificial ou redes neurais.

¹⁰<https://kedro.org/>

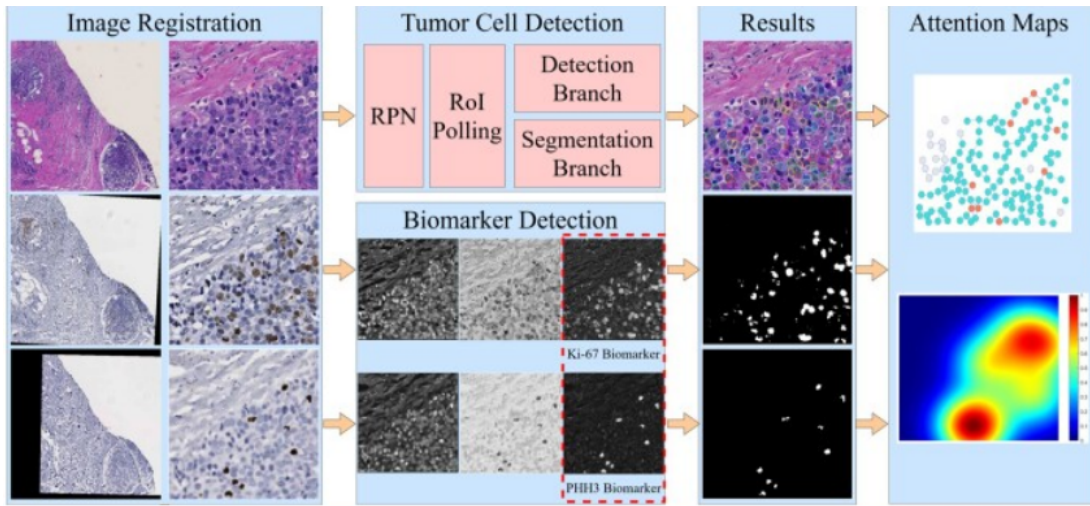


Figura 3. Detecção de Biomarcadores

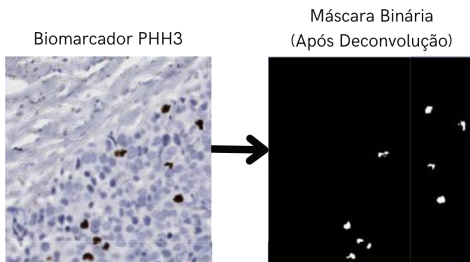


Figura 4. Resultados da Detecção de Biomarcadores

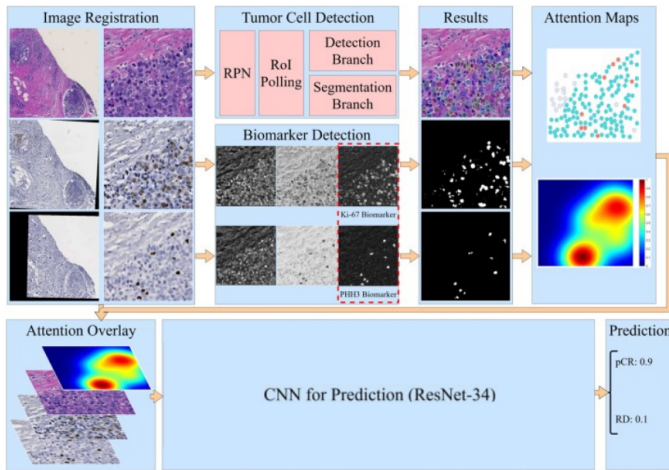


Figura 5. Utilização de Mapas de Atenção

segmentação de células tumorosas, com a detecção de biomarcadores para a criação do mapa de atenção que será adicionado a uma pilha juntamente com as imagens histopatológicas do tipo H&E, Ki-67 e PHH3 para serem submetidas à rede CNN (*Convolutional Neural Network*).

A. Abordagem Implementada

A implementação da etapa de geração dos mapas de atenção precisou de adaptações contínuas comparadas às inicialmente proposta original do artigo de referência. Começando pela necessidade de refatoração completa do código, para que estivesse aderente ao formato moderno proposto pelo Python e tivesse um aumento de compreensibilidade.

De forma similar à etapa anterior, a primeira parte foi caracterizada pela análise e entendimento do código dos autores, de forma que a compreensão de como os dados gerados até esta etapa, eram utilizados para a confecção dos mapas de atenção. Mais uma vez, o processo não ocorreu de maneira desimpedida, dado que o código não era capaz de adaptar-se aos *outputs* das etapas anteriores, somado à falta de padrões de engenharia de software, mais uma vez prejudicando a compreensão do código.

Inicialmente, fez-se a adaptação do código ao modelo esperado pelo *framework* utilizado, de forma que essa nova etapa estivesse de acordo com o modelo das etapas anteriores. Adicionalmente, foi necessário realizar a reestruturação do código de forma a aumentar sua compreensão, bem como a adaptação do mesmo de forma a permitir com que lidasse com os dados intermediários produzidos até então.

Nesse contexto, foi possível de fato testar o algoritmo disponibilizado para a geração dos mapas de atenção, porém os resultados eram completamente baldados, isso porque não passavam de imagens ruidosas sem padrão definido para que pudessem ser usadas na próximas etapas. Mesmo com um processo extenso de depuração, não foi possível solucionar os problemas do algoritmo de forma direta, contudo, durante esse processo uma importante compreensão do código foi obtida.

Dessa forma, decidiu-se por realizar a confecção de um novo algoritmo baseado no fornecido pelo autores, de forma que a lógica fosse mantida mas que fosse feito de

forma gradual e testável. Durante o processo percebeu-se diversas partes do código original improdutivas e enviesadas, após reestruturação da lógica, foi possível testar o código de geração dos mapas de atenção, que, dessa vez, geraram resultados bem condizentes com as expectativas que se tinham.

Por fim, em uma iteração final desta etapa, fez-se necessária a adaptação do processo mais uma vez. Isso decorreu do fato de que a segmentação de células tumorosas se mostrou insatisfatória, o que enviesava e prejudicava o processo da geração dos mapas de atenção. Após a avaliação dos resultados com a implementação dos ditos mapas, evidenciou-se a deficiência do processo.

Por isso, alterou-se a abordagem utilizada, anteriormente usava-se as células tumorosas segmentadas como base, e no caso de existir intersecção com a máscara binária da imagem Ki-67 ou PHH3 deconvoluída adicionava-se então, um valor positivo ao mapa. A alteração simplificou o processo, fazendo uma sobreposição das máscaras binárias e transformando esse resultado em um mapa de calor, i.e., regiões com resultados positivos para ambos os marcadores, teriam sobreposição de gaussianas o que culminaria em uma região "quente" do mapa, regiões onde só um dos marcadores está presente seriam intermediárias e onde nenhum dos marcadores está presente seriam consideradas de baixo interesse.

B. Resultados Obtidos

Ao fim desta etapa, os resultados aparentemente são promissores, uma vez que conseguem combinar os dados intermediários gerados pelas etapas anteriores de forma condizente com o que foi elucidado no artigo referenciado.

Para essa etapa, mais uma vez foi necessária a configuração da pipeline Kedro para que executasse o processamento de todas as imagens disponíveis e efetuasse a criação dos mapas de atenção. Contudo, dada a natureza das entradas anteriores, pilha de imagens de alta resolução (8000x8000 *pixels*), houve um problema no que se refere à alocação de memória, onde, mesmo em máquinas com alto poder de processamento, a memória disponível não foi suficiente para lidar com o processamento de maneira ingênua. Portanto, otimizações em termos de utilização e limpeza de memória foram efetuadas, bem como a segmentação do trabalho para evitar uma sobrecarga ao tratar com entradas muito grandes, particularmente as que apresentavam muita detecção de células tumorosas, uma vez que, para cada detecção uma imagem individual é criada e adicionada à pilha.

Em suma, após a realização das atividades supracitadas, foi possível executar o processo para a geração dos mapas de atenção com base nos dados fornecidos pelos autores do artigo referenciado. Mais uma vez, os dados foram todos armazenados de maneira intermediária, seguindo os padrões do *framework* Kedro para que possam ser usados nas próximas etapas do processo. A Figura 6 mostra exemplos de alguns dos mapas de calor gerados, o primeiro deles é

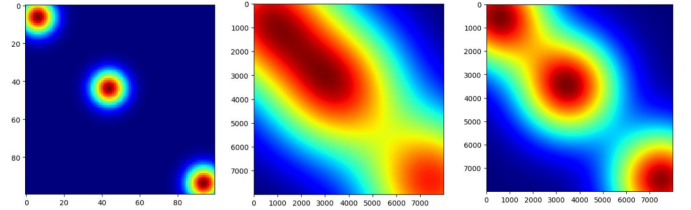


Figura 6. Resultados da Geração de Mapas de Atenção

um mapa de teste, onde intencionalmente se posicionam três concentrações de células tumorosas a fim de validar o funcionamento do algoritmo, já os dois outros mapas são extraídos de execuções com dados reais.

VII. REDE DE CLASSIFICAÇÃO

Esta seção descreve a etapa culminante do projeto: o desenvolvimento e a avaliação de modelos de aprendizado profundo para a predição da Resposta Patológica Completa (pCR). Nesta fase, os artefatos gerados previamente — as imagens histopatológicas e os mapas de atenção espacial — são integrados como uma entrada multimodal para uma Rede Neural Convolutacional (CNN). O objetivo é conduzir uma análise comparativa, ou estudo de ablação, para quantificar o impacto de cada componente de dados no desempenho preditivo final. Todas as configurações do modelo foram treinadas e validadas utilizando os rótulos de classificação (dados-alvo) disponibilizados pelos autores do trabalho de referência [7], assegurando a consistência com a pesquisa original.

A. Trabalho Desenvolvido

O núcleo da arquitetura preditiva é a **ResNet-34**, uma rede neural residual profunda detalhada em [23]. Esta arquitetura foi escolhida por seu reconhecido equilíbrio entre profundidade, capacidade de aprendizado de características complexas e por mitigar o problema de gradientes descendentes, tornando-a um padrão-ouro para tarefas de classificação de imagens médicas. Para acelerar a convergência e melhorar a generalização foi avaliada a possibilidade do modelo ser inicializado com pesos pré-treinados na base de dados ImageNet¹², uma prática consolidada de aprendizado por transferência (*transfer learning*). Contudo, essa abordagem não se mostrou produtiva para o contexto específico, e um treinamento partindo de um cenário sem inicialização prévia se mostrou mais produtivo.

Para avaliar sistematicamente a contribuição de cada fonte de dados, foram definidos três experimentos principais:

- **Modelo Base (H&E Only):** Uma configuração inicial treinada exclusivamente com as imagens de coloração padrão H&E. Este modelo serve como nossa linha de base para medir as melhorias subsequentes.

¹²<https://www.image-net.org/>

- **Modelo com Biomarcadores (H&E + Biomarkers):** Nesta configuração, as imagens H&E foram combinadas com as imagens dos biomarcadores de proliferação Ki-67 e PHH3. Tecnicamente, as imagens foram empilhadas ao longo da dimensão de canais, formando um tensor de entrada mais rico em informações.
- **Modelo com Atenção (H&E + Biomarkers + Attention):** A configuração final e mais completa. Além da pilha de imagens do modelo anterior, foi ponderar todas as imagens com o mapa de atenção espacial. Este mapa guia a rede a focar em regiões de alta relevância biológica, conforme identificado pela sobreposição dos biomarcadores.

B. Resultados Obtidos

Os resultados quantitativos dos três experimentos são apresentados na Tabela I. A análise revela uma melhoria clara e progressiva no desempenho preditivo à medida que informações mais ricas são fornecidas ao modelo.

O **Modelo Base (H&E Only)** estabeleceu um desempenho inicial razoável, com um AUC de 0.8311. No entanto, seu baixo valor de precisão (0.7538) em contraste com um altíssimo recall (0.9859) indica uma forte tendência a gerar falsos positivos.

A adição dos biomarcadores (**H&E + Biomarkers**) provocou um salto de performance substancial e significativo. O AUC aumentou em quase 11 pontos para 0.9420, e a precisão subiu para 0.8546. Isso demonstra que os marcadores Ki-67 e PHH3 fornecem características discriminativas cruciais, resultando em um modelo muito mais equilibrado e confiável, como reflete o F1-Score de 0.9013.

Finalmente, o **Modelo com Atenção** confirmou a hipótese central deste trabalho. Ao utilizar os mapas de atenção para guiar o foco da rede, o modelo atingiu o desempenho máximo em todas as métricas principais, com um **AUC de 0.9598** e um **F1-Score de 0.9091**. Este resultado valida que a aplicação de um mecanismo de atenção, um conceito fundamental em arquiteturas de IA modernas [10], é extremamente eficaz para direcionar o processo de aprendizado para as características patológicas mais relevantes, maximizando a capacidade preditiva do modelo.

VIII. VIRTUAL STAINING

A etapa de *Virtual Staining* não é foco deste trabalho, sendo essa desenvolvida no projeto de pares no Departamento de Ciência da Computação da UFMG. Contudo, o autor deste artigo teve participação ativa no projeto em questão, sendo ambos os projetos correlatos, tendo por objetivo sua integração em trabalhos futuros.

O *Virtual Staining* é desenvolvido de forma completamente independente ao trabalho apresentado em [7], uma vez que não há utilização de imagens histopatológicas coradas artificialmente neste artigo. De forma geral, a ideia

é incrementar a pipeline de predição da resposta *pCR* de forma que as imagens de biomarcadores utilizadas (Ki-67 e PHH3) sejam geradas artificialmente por meio da técnica de *Virtual Staining*. Sendo isso feito a partir das imagens de Eosina e Hematoxilina (H&E), visto que essas são de mais fácil obtenção [24].

A. Trabalho Desenvolvido

Para esta etapa usou-se uma, Cycle-GAN [17] que é uma arquitetura de rede generativa adversária (GAN) projetada para a tradução de imagens entre dois domínios distintos de forma não supervisionada, ou seja, sem a necessidade de pares de imagens correspondentes. O modelo é composto por dois geradores e dois discriminadores (sendo um para cada domínio), que são treinados simultaneamente para traduzir imagens de um domínio para o outro e para distinguir as imagens reais das geradas.

A principal inovação da Cycle-GAN é a "perda de ciclo de consistência", uma função que assegura que uma imagem traduzida de um domínio para o outro e, em seguida, de volta para o domínio original, permaneça estruturalmente semelhante à imagem original. Isso garante a preservação de conteúdo e características importantes, evitando distorções.

A implementação específica utilizada neste projeto é baseada no estado da arte apresentado em [17] e aprimora a arquitetura padrão com duas adições cruciais: o uso do SSIM na função de perda para melhorar a preservação de detalhes anatômicos e a incorporação de conexões de atalho (*skip-connections*), que ajudam a manter a fidelidade de estruturas celulares finas. Essas características tornam o modelo uma ferramenta poderosa para tarefas como a normalização e coloração virtual (*virtual staining*) de imagens histopatológicas.

B. Resultados Obtidos

Para a avaliação do trabalho pretende-se conduzida duas etapas principais: uma **qualificação intrínseca** para medir a qualidade das imagens virtualizadas e uma **qualificação extrínseca** para avaliar a utilidade dessas imagens na tarefa de predição da *pCR* para pacientes com TNBC.

A análise intrínseca, em particular, utilizou três métricas distintas (PSNR, SSIM e MI) para realizar uma avaliação multifacetada. Esta avaliação consistiu em três comparações principais:

- 1) As imagens virtualizadas foram comparadas com as imagens H&E para aferir a qualidade da tradução de domínio.
- 2) A similaridade estrutural (via SSIM) foi analisada comparando-se tanto as imagens IHC originais quanto as virtualizadas com as imagens H&E.
- 3) A distribuição de informação foi estudada comparando as imagens IHC (originais e virtualizadas) com as de H&E, e também entre si.

Tabela I
COMPARAÇÃO DAS MÉTRICAS DE PERFORMANCE DOS MODELOS

Modelo	Acurácia	AUC	Precisão	Revocação	F1-Score
<i>H&E Only</i>	0.8031	0.8311	0.7538	0.9859	0.8544
<i>H&E + Biomarkers</i>	0.8776	0.9420	0.8546	0.9533	0.9013
<i>H&E + Biomarkers + Attention</i>	0.8923	0.9598	0.8986	0.9198	0.9091

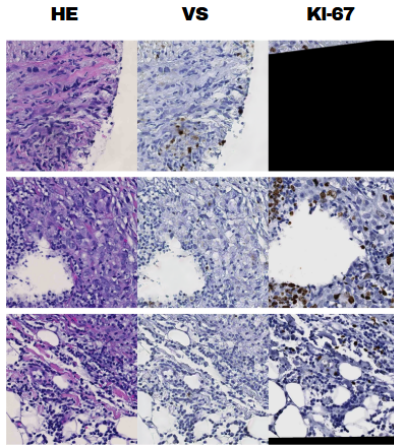


Figura 7. *Virtual Staining* aplicado à geração de imagens Ki-67

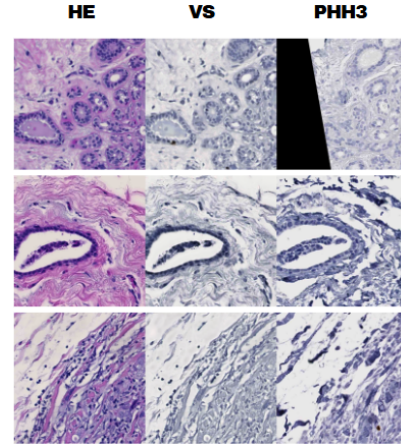


Figura 8. *Virtual Staining* aplicado à geração de imagens PHH3

Os resultados da análise intrínseca podem ser vistos nas Figuras 7 e 8. De forma geral, é possível ver como a representação espacial das imagens de *Virtual Staining* é muito superior as imagens de referência, isso acontece pois o processo de coleta das lâminas deforma o decido, em alguns casos gerando uma grande distorção. Isso se apresenta como um grande desafio para o treinamento das redes neurais, uma vez que não há uma compatibilidade total entre as imagens de referência (H&E) e as imagens *label* (Ki-67 e PHH3).

Já a validação extrínseca será feita justamente em uma etapa de trabalhos futuros, onde a *pipeline* descrita neste trabalho será integrada com a etapa de geração de imagens coradas artificialmente, conforme explicitado na sessão de Conclusão e Trabalhos Futuros (IX).

IX. CONCLUSÃO E TRABALHOS FUTUROS

Este trabalho conclui com o desenvolvimento e a validação bem-sucedida de uma pipeline de aprendizado profundo para a predição da resposta patológica completa (pCR), confirmando a hipótese de que a integração de dados multimodais melhora significativamente a acurácia da classificação. Os resultados demonstraram uma melhoria clara e progressiva, na qual o modelo de base, treinado apenas com imagens H&E, foi substancialmente superado ao adicionar os dados dos biomarcadores Ki-67 e PHH3. O modelo mais sofisticado, que incorporou um mecanismo de atenção derivado destes biomarcadores, alcançou o mais alto desempenho em todas as métricas-chave, incluindo um AUC de 0.9598. Este fato ressalta o imenso valor de

guiar o foco de uma rede neural em direção a regiões de alta relevância biológica para alcançar uma predição mais robusta e precisa, um passo fundamental para o desenvolvimento de ferramentas de apoio à decisão clínica.

O processo de desenvolvimento exigiu uma adaptação metodológica significativa, notavelmente a superação da falha da etapa inicial de segmentação de tumores através da reengenharia do mecanismo de atenção para depender unicamente dos dados dos biomarcadores. Essa reorientação pragmática não apenas assegurou o sucesso do projeto, mas também destacou a robustez de uma abordagem centrada em biomarcadores. Os resultados conclusivos sustentam fortemente o potencial de trabalhos futuros em *virtual staining* (coloração virtual). Se estes biomarcadores são tão eficazes, a capacidade de gerá-los computacionalmente a partir de lâminas H&E padrão poderia tornar esta poderosa técnica preditiva mais acessível, econômica e escalável, avançando o estado da arte em patologia digital automatizada.

Em termos de trabalhos futuros, o roteiro se concentra em duas frentes principais que são ilustradas na Figura 9, sendo essas a integração do *virtual staining* e da etapa de segmentação de células tumorosas.

Primeiramente, planeja-se integrar uma etapa de *virtual staining* no início da pipeline, utilizando técnicas como as descritas em [17] para gerar sinteticamente os marcadores Ki-67 e PHH3 a partir de imagens H&E. Em segundo lugar, pretende-se implementar uma técnica de segmentação tumoral mais robusta, possivelmente adotando uma abordagem diferente da do artigo original, como a que é

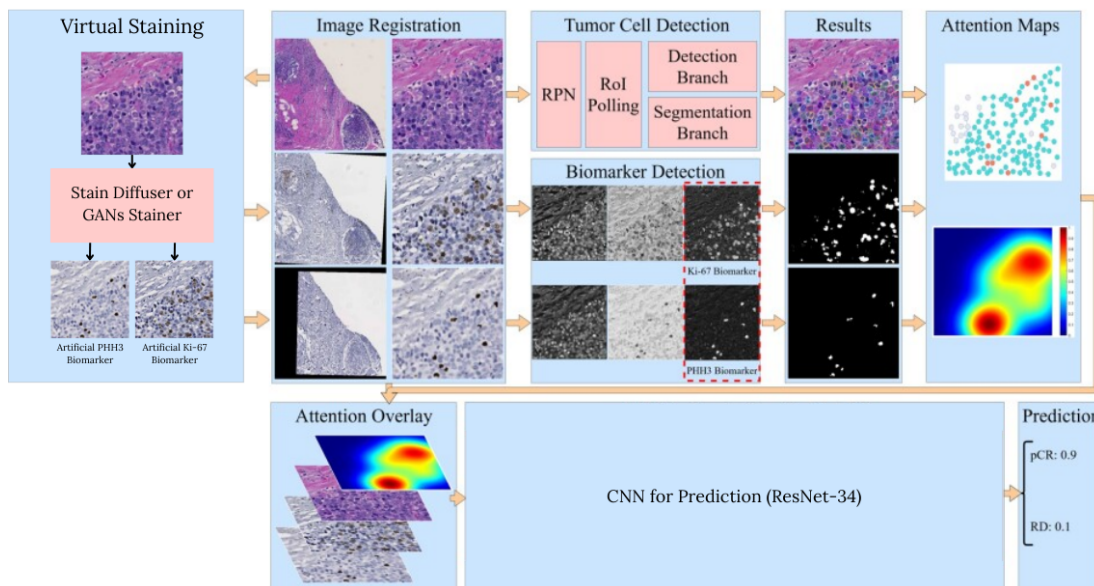


Figura 9. *Virtual Staining* e Segmentação de Regiões Tumorosas integrados à pipeline de predição da *pCR*

descrita em [19]. Este método segmenta diferentes regiões do microambiente tumoral (e.g., estroma, linfócitos), o que poderia fornecer um contexto ainda mais rico para a geração de mapas de atenção. Assim, a pipeline de trabalhos futuros imaginada, que integra a coloração virtual e a segmentação de regiões tumorais à pipeline de predição de *pCR* representaria uma potencial evolução da pipeline original, configurando um avanço no estado da arte atual superando o problema de obtenção de imagens histopatológicas coradas com Ki-67 e PHH3, isso enquanto aplica novas técnicas de atenção espacial com o propósito de potencializar a performance da rede.

APÊNDICE A REPOSITÓRIO DO CÓDIGO FONTE

O código fonte completo desenvolvido para este trabalho, incluindo os scripts para pré-processamento dos dados, treinamento dos modelos e a pipeline de análise, está publicamente disponível em um repositório no GitHub para garantir a reprodutibilidade e transparência desta pesquisa.

O repositório pode ser acessado no seguinte endereço: <https://github.com/Lab-of-Applied-Computer-Science/breast-cancer-segmentation>

Recomenda-se consultar o arquivo `README.md` contido no repositório para obter instruções sobre a configuração do ambiente de desenvolvimento e a execução dos experimentos descritos neste artigo.

APÊNDICE B BASE DE DADOS UTILIZADA

Os modelos de aprendizado profundo descritos neste trabalho foram treinados e validados utilizando um conjunto de dados composto, proveniente das fontes detalhadas abaixo.

A. Fonte dos Dados

A base de dados principal, que contém as imagens histopatológicas (H&E, Ki-67, PHH3) e os respectivos rótulos de classificação da *pCR*, foi a mesma utilizada e gentilmente disponibilizada pelos autores do trabalho de referência [7]¹³. Agradecemos aos autores por tornarem seus dados acessíveis, o que foi fundamental para a validação comparativa de nossos resultados.

Adicionalmente, para a etapa exploratória de segmentação de células tumorais, foi utilizada a base de dados pública "Breast Cancer Cell Segmentation", disponível na plataforma Kaggle¹⁴.

B. Características da Base de Dados

O conjunto de dados principal para a predição de *pCR* consiste em:

- Imagens de lâminas inteiras (*Whole Slide Images - WSI*) coradas com H&E.
- Imagens de seções de tecido correspondentes, coradas com os biomarcadores Ki-67 e PHH3.
- Rótulos de classificação binária (*pCR* ou não-*pCR*) para cada paciente, que serviram como o alvo (*ground truth*) para o treinamento supervisionado.

REFERÊNCIAS

- [1] HOWARD, F. M. et al. Integration of clinical features and deep learning on pathology for the prediction of breast cancer recurrence assays and risk of recurrence. *npj Breast Cancer* 2023 9:1, Nature Publishing Group, v. 9, p. 1–6, 4 2023. ISSN 2374-4677. Disponível em: <<https://www.nature.com/articles/s41523-023-00530-5>>.

¹³Os dados não são de acesso público, portanto não estão referenciados neste documento.

¹⁴<https://www.kaggle.com/datasets/andrewmvd/breast-cancer-cell-segmentation>

- [2] WU, Y. et al. Recent advances of deep learning for computational histopathology: Principles and applications. *Cancers 2022*, Vol. 14, Page 1199, Multidisciplinary Digital Publishing Institute, v. 14, p. 1199, 2 2022. ISSN 2072-6694. Disponível em: <<https://www.mdpi.com/2072-6694/14/5/1199/htm>>.
- [3] RUIJTER, T. C. D. et al. Characteristics of triple-negative breast cancer. *Journal of Cancer Research and Clinical Oncology*, v. 137, p. 183, 2 2010. ISSN 01715216. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/articles/PMC3018596/>>.
- [4] POGODA, K. et al. Analysis of pattern, time and risk factors influencing recurrence in triple-negative breast cancer patients. *Medical Oncology*, Med Oncol, v. 30, 3 2013. ISSN 13570560. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/23292831/>>.
- [5] KATHER, J. N. et al. Pan-cancer image-based detection of clinically actionable genetic alterations. *Nature cancer*, Nat Cancer, v. 1, p. 789–799, 8 2020. ISSN 2662-1347. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/33763651/>>.
- [6] COUTURE, H. D. et al. Image analysis with deep learning to predict breast cancer grade, er status, histologic subtype, and intrinsic subtype. *NPJ breast cancer*, NPJ Breast Cancer, v. 4, 12 2018. ISSN 2374-4677. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/30182055/>>.
- [7] DUANMU, H. et al. A spatial attention guided deep learning system for prediction of pathological complete response using breast cancer histopathology images. *Bioinformatics*, Oxford Academic, v. 38, p. 4605–4612, 9 2022. ISSN 1367-4803. Disponível em: <<https://dx.doi.org/10.1093/bioinformatics/btac558>>.
- [8] LI, H. et al. Deep learning-based pathology image analysis enhances magee feature correlation with oncotype dx breast recurrence score. *Frontiers in Medicine*, v. 9, 2022. ISSN 2296-858X. Disponível em: <<https://www.frontiersin.org/articles/10.3389/fmed.2022.886763>>.
- [9] ABRAHAM, T.; LEVENSON, R. *A comparison of diffusion models and CycleGANs for virtual staining of slide-free microscopy images*. 2023.
- [10] VASWANI, A. et al. Attention is all you need. In: *Advances in Neural Information Processing Systems*. [S.l.: s.n.], 2017. v. 2017-December. ISSN 10495258.
- [11] SALEHI, P.; CHALECHALE, A.; TAGHIZADEH, M. Generative adversarial networks (gans): An overview of theoretical model, evaluation metrics, and recent developments. 5 2020. Disponível em: <<https://arxiv.org/pdf/2005.13178>>.
- [12] CHEN, J. et al. Deep learning models for cancer stem cell detection: a brief review. *Frontiers in Immunology*, Frontiers Media SA, v. 14, 2023. ISSN 16643224. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/37441078/>>.
- [13] YADAVENDRA; CHAND, S. A comparative study of breast cancer tumor classification by classical machine learning methods and deep learning method. *Machine Vision and Applications*, Springer, v. 31, p. 1–10, 9 2020. ISSN 14321769. Disponível em: <<https://link.springer.com/article/10.1007/s00138-020-01094-1>>.
- [14] ZHU, W. et al. The application of deep learning in cancer prognosis prediction. *Cancers*, MDPI AG, v. 12, 3 2020. ISSN 20726694. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/32150991/>>.
- [15] KNUTSVIK, G. et al. Evaluation of ki67 expression across distinct categories of breast cancer specimens: A population-based study of matched surgical specimens, core needle biopsies and tissue microarrays. *PLoS ONE*, Public Library of Science, v. 9, 11 2014. ISSN 19326203. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/25375149/>>.
- [16] HAAN, K. de et al. Deep learning-based transformation of h&e stained tissues into special stains. *Nature Communications 2021 12:1*, Nature Publishing Group, v. 12, p. 1–13, 8 2021. ISSN 2041-1723. Disponível em: <<https://www.nature.com/articles/s41467-021-25221-2>>.
- [17] BAI, B. et al. Deep learning-enabled virtual histological staining of biological samples. *Light: Science & Applications 2023 12:1*, Nature Publishing Group, v. 12, p. 1–20, 3 2023. ISSN 2047-7538. Disponível em: <<https://www.nature.com/articles/s41377-023-01104-7>>.
- [18] DHARIWAL, P.; NICHOL, A. Diffusion models beat gans on image synthesis. In: *Advances in Neural Information Processing Systems*. [S.l.: s.n.], 2021. v. 11. ISSN 10495258.
- [19] AMGAD, M. et al. Structured crowdsourcing enables convolutional segmentation of histology images. *Bioinformatics*, Oxford Academic, v. 35, p. 3461–3467, 9 2019. ISSN 1367-4803. Disponível em: <<https://dx.doi.org/10.1093/bioinformatics/btz083>>.
- [20] WENG, W.; ZHU, X. U-net: Convolutional networks for biomedical image segmentation. *IEEE Access*, Institute of Electrical and Electronics Engineers Inc., v. 9, p. 16591–16603, 5 2015. ISSN 21693536. Disponível em: <<https://arxiv.org/abs/1505.04597v1>>.
- [21] KAUR, A.; KAUR, M. Analysis of code refactoring impact on software quality. *MATEC Web of Conferences*, EDP Sciences, v. 57, 5 2016. ISSN 2261236X. Disponível em: <<https://www.researchgate.net/publication/302977540>>.
- [22] HAUB, P.; MECKEL, T. A model based survey of colour deconvolution in diagnostic brightfield microscopy: Error estimation and spectral consideration. *Scientific Reports 2015 5:1*, Nature Publishing Group, v. 5, p. 1–15, 7 2015. ISSN 2045-2322. Disponível em: <<https://www.nature.com/articles/srep12096>>.
- [23] HE, K. et al. Deep residual learning for image recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, v. 2016-December, p. 770–778, 12 2015. ISSN 10636919. Disponível em: <<https://arxiv.org/pdf/1512.03385>>.
- [24] ALVES, A. P. N. N. et al. Is immunohistochemistry more sensitive than hematoxylin-eosin staining for identifying perineural or lymphovascular invasion in oral squamous cell carcinoma? a systematic review and meta-analysis. *Medicina Oral Patologia Oral y Cirugia Bucal*, Medicina Oral S.L., v. 27, p. e238–e247, 5 2022. ISSN 16986946. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/35420066/>>.