

# Redes DenseNet com Mecanismos de Atenção Múltipla aplicadas à Classificação Automática de Câncer Peniano em Imagens Histopatológicas

João Guilherme Araujo do Vale<sup>1</sup>, Italo Francyles Santos da Silva<sup>2</sup>,  
Caio Eduardo Falcão Matos<sup>2</sup>, Geraldo Braz Júnior<sup>1</sup>,  
Marcos Gabriel Mendes Lauande<sup>2</sup>

<sup>1</sup>Laboratório de Visão e Processamento de Imagens – Universidade Federal do Maranhão (UFMA), São Luís – MA, Brazil

<sup>2</sup>Núcleo Computação Aplicada – Universidade Federal do Maranhão (UFMA), São Luís – MA, Brazil

jga.vale@discente.ufma.br,

{italo.francyles, geraldo.braz, marcos.lauande}@ufma.br,

caioefalcao@nca.ufma.br

**Abstract.** *Penile carcinoma, also known as penile cancer, is a malignant neoplasm that predominantly affects men over 50 years old. Late diagnosis leads to significant complications that deeply impact patients' lives in psychological, emotional, and social terms. Exams such as histopathological analysis are essential for early diagnosis but require time and highly qualified professionals. The use of deep machine learning, through convolutional neural networks, emerges as a highly applicable approach in this context. In this work, two methods were proposed for the automatic classification of penile cancer, using DenseNet neural networks with attention mechanisms in their architecture. The dataset was provided by the Penile Cancer Project of the Legal Amazon, containing 194 images at 40x and 100x magnifications. As a result, the best model achieved an F1-Score of 93.1% at the 100x magnification.*

**Resumo.** *O carcinoma peniano, conhecido como câncer de pênis, é uma neoplasia maligna que afeta predominantemente homens com mais de 50 anos. O diagnóstico tardio acarreta complicações significativas que impactam profundamente a vida dos pacientes em termos psicológicos, emocionais e sociais. Exames como a análise histopatológica são fundamentais para o diagnóstico precoce, porém, requerem tempo e profissionais altamente qualificados. O uso de aprendizado de máquina profundo, por meio de redes neurais convolucionais, apresenta-se como uma abordagem altamente aplicável neste contexto. Neste trabalho, foram propostos dois métodos para a classificação automática do câncer peniano, utilizando redes neurais DenseNet com mecanismos de atenção em sua arquitetura. Como resultado, o melhor modelo alcançou um F1-Score de 93,1% na ampliação de 100x.*

## 1. Introdução

O câncer de pênis, também conhecido como carcinoma peniano, é uma forma pouco comum de neoplasia que tipicamente afeta homens com mais de 50 anos. Segundo

dados do Instituto Nacional de Câncer [INCA 2022] do ano de 2022, essa malignidade representa cerca de 0,4% de todos os cânceres diagnosticados em homens no Brasil, com uma distribuição geográfica destacada nas regiões Norte e Nordeste do país. O estado do Maranhão, se destaca por apresentar a maior taxa mundial de ocorrências desta patologia. Entre os principais fatores de risco associados estão a fimose, que limita a higiene adequada, a infecção pelo vírus do papiloma humano (HPV) e condições socioeconômicas desfavoráveis. É fundamental ressaltar que o tratamento do câncer de pênis acarreta consideráveis impactos na qualidade de vida dos pacientes, afetando tanto aspectos sexuais quanto urinários. Essas complicações podem resultar em desafios significativos do ponto de vista psicológico, emocional e social para os indivíduos afetados.

As imagens histopatológicas desempenham um papel fundamental no diagnóstico e estadiamento do câncer de pênis [Thomas et al. 2021]. A histopatologia, que consiste no estudo microscópico de tecidos obtidos por meio de biópsias ou cirurgias, é uma etapa crucial no processo diagnóstico, segundo o que foi indicado pelo Ministério da Saúde do Brasil [MS 2022]. No caso específico do câncer de pênis, uma biópsia é realizada para coletar uma amostra do tecido tumoral, a qual é subsequentemente examinada por um patologista especializado. A análise histopatológica não apenas confirma o diagnóstico da doença, mas também fornece informações essenciais para o planejamento do tratamento e a previsão do prognóstico do paciente. Contudo, é importante destacar que a interpretação precisa das imagens histopatológicas requer um investimento significativo de tempo e a experiência de profissionais qualificados [Melo et al. 2020]. Além disso, a variabilidade na interpretação entre patologistas pode representar um desafio adicional, ressaltando a necessidade de protocolos e diretrizes claras para garantir a consistência e confiabilidade dos resultados.

Nesse contexto, a automação da classificação de imagens histopatológicas para identificar a presença de câncer emerge como uma perspectiva promissora [Erickson et al. 2017]. A aplicação de técnicas de aprendizado de máquina na área médica tem demonstrado resultados notáveis, evidenciando êxitos significativos na detecção de lesões [Lauande et al. 2022], classificação de patologias [Machado and Veloso 2021] e interpretação de imagens médicas complexas [Oliveira et al. 2022]. Em particular, o emprego de técnicas de aprendizado profundo, como as redes neurais convolucionais (CNNs), desponta como uma abordagem capaz de não apenas agilizar o processo diagnóstico, mas também de permitir que os profissionais médicos dediquem mais tempo e atenção ao planejamento e execução do tratamento dos pacientes. Essa integração da tecnologia de aprendizado de máquina na análise de imagens histopatológicas oferece uma perspectiva transformadora para a prática médica, promovendo uma abordagem mais eficiente e precisa no diagnóstico e manejo do câncer de pênis.

Embora as redes neurais apresentem um potencial significativo para a classificação de imagens histopatológicas, há desafios substanciais a serem superados a fim de garantir resultados satisfatórios. Entre esses desafios, destaca-se, particularmente, a falta de padronização das cores presentes nas imagens do conjunto de dados [Vahadane et al. 2016]. Esta questão está intrinsecamente ligada à natureza das imagens histopatológicas, as quais são geradas a partir de cortes em tecido celular pigmentado. A variação na utilização de pigmentos e equipamentos para a aquisição das imagens, bem como o desvanecimento do pigmento ao longo do tempo, contribui para a ausência de padronização nas cores das

imagens, o que por sua vez dificulta o processo de aprendizado das redes neurais. Esse desafio ressalta a necessidade de abordagens robustas e estratégias específicas para lidar com as nuances envolvidas na análise de imagens histopatológicas.

Para enfrentar esses desafios, uma variedade de arquiteturas de redes neurais, juntamente com outros métodos de aprendizado de máquina, foram empregados. Os modelos utilizados são adaptações da rede neural DenseNet, originalmente introduzida por [Huang et al. 2017], nos quais foram incorporados mecanismos de atenção inspirados em abordagens utilizadas em tarefas de segmentação de imagens médicas por redes do tipo Unet. Esses mecanismos têm o propósito de filtrar informações mais relevantes, resultando em uma maior precisão nos resultados. Além disso, foi aplicada uma ferramenta de pré-processamento para lidar com a falta de padronização do conjunto de imagens, o método de normalização Vahadane [Vahadane et al. 2016], feito para tratar especificamente imagens histopatológicas.

Portanto, este trabalho visa propor e avaliar comparativamente abordagens de aprendizado profundo para a classificação do câncer de pênis em imagens histopatológicas, visando auxiliar no combate desta patologia ao melhorar a sensibilidade dos patologistas durante o processo de diagnóstico.

Em termos gerais, as contribuições deste trabalho podem ser destacadas nos seguintes pontos:

- Proposição de dois métodos para o diagnóstico automático de câncer de pênis em imagens histopatológicas, utilizando a rede neural DenseNet201 com mecanismos de atenção incorporados em sua estrutura;
- Adaptação e integração de técnica de pré-processamento ainda não utilizada anteriormente neste problema em específico, a normalização Vahadane;

## 2. Trabalhos relacionados

Esta seção apresenta trabalhos relacionados a classificação em imagens histopatológicas de câncer de pênis. Ao final, os resultados obtidos por cada um desses trabalhos estarão organizados na tabela 1.

[Belfort et al. 2023] desenvolve um método em cascata de duas etapas, e ambas etapas com uma CNNs e mecanismos de *Soft-Attention*. A primeira etapa consiste na seleção da ampliação da imagem. A segunda etapa na classificação. Como resultado, o modelo alcançou acurácia de 93% e F1-Score de 94% para a ampliação de 40x, e acurácia de 90% e F1-Score de 94% para a ampliação de

O estudo conduzido por [Lauande et al. 2022] apresenta uma abordagem automatizada para a classificação de câncer de pênis em imagens histopatológicas, utilizando redes neurais convolucionais com transferência de aprendizado e *fine-tuning*. Ao final, a metodologia apresentada neste trabalho alcançou resultados promissores, com F1-Score de até  $97,39\% \pm 2,13\%$  em uma base de dados utilizada pela primeira vez em experimentos para automação de diagnóstico médico baseado em imagens.

Observando os trabalhos relacionados, destaca-se a aplicação de redes neurais convolucionais da família DenseNet [Huang et al. 2017] e dos algoritmos CLAHE e seu predecessor, HE, como ferramentas de pré-processamento, na composição da metodologia proposta pelos trabalhos apresentados. Desta forma, neste trabalho, buscou-se explorar

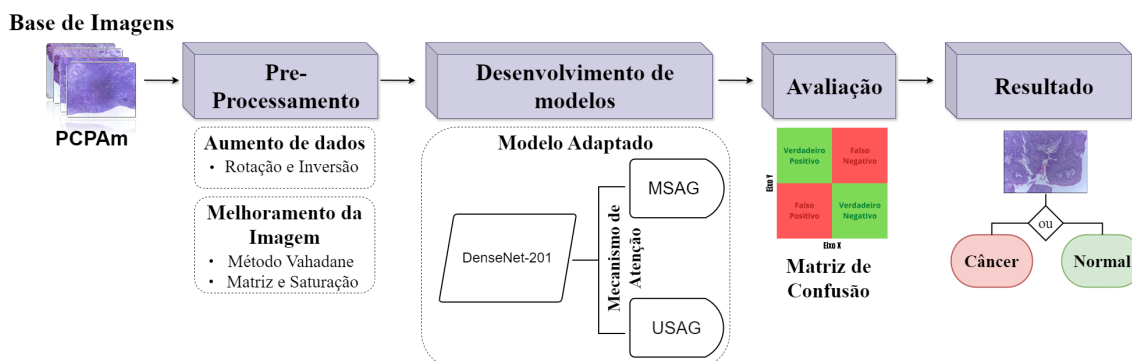
adaptações da arquitetura DenseNet, através da incorporação de mecanismos de atenção em sua estrutura, e novas ferramentas de pré-processamento, como o método de normalização Vahadane [Vahadane et al. 2016].

**Tabela 1. Resultados apresentados pelos trabalhos relacionados.**

Autor	Base de Imagens	Método	Resultados
[Belfort et al. 2023]	Projeto Câncer de Pênis da Amazônia Legal	Modelo em cascata com mecanismos de <i>Soft-Attention</i>	Ampliação 40x: Accuracy de 93% e F1-Score de 94% Ampliação 100x: Accuracy de 90% e F1-Score de 94%
[Lauande et al. 2022]	Projeto Câncer de Pênis da Amazônia Legal	Transferência de aprendizado	Ampliação 40x: F1-Score de 97,39% $\pm$ 2,13% Ampliação 100x: F1-Score de 97,31% $\pm$ 3,62%

### 3. Metodologia

Neste estudo, foram selecionadas e avaliadas abordagens de atenção para a classificação de câncer em imagens histopatológicas. Na presente seção, são apresentados todos os detalhes relacionados a essas abordagens. A Figura 1 resume a metodologia proposta neste trabalho, composta pela aquisição da base de dados, pela etapa de pré-processamento das imagens, pelo desenvolvimento e treinamento de arquiteturas de redes neurais, e finaliza com a avaliação dos resultados obtidos.



**Figura 1. Visão Geral da Metodologia Proposta**

#### 3.1. Base de Imagens

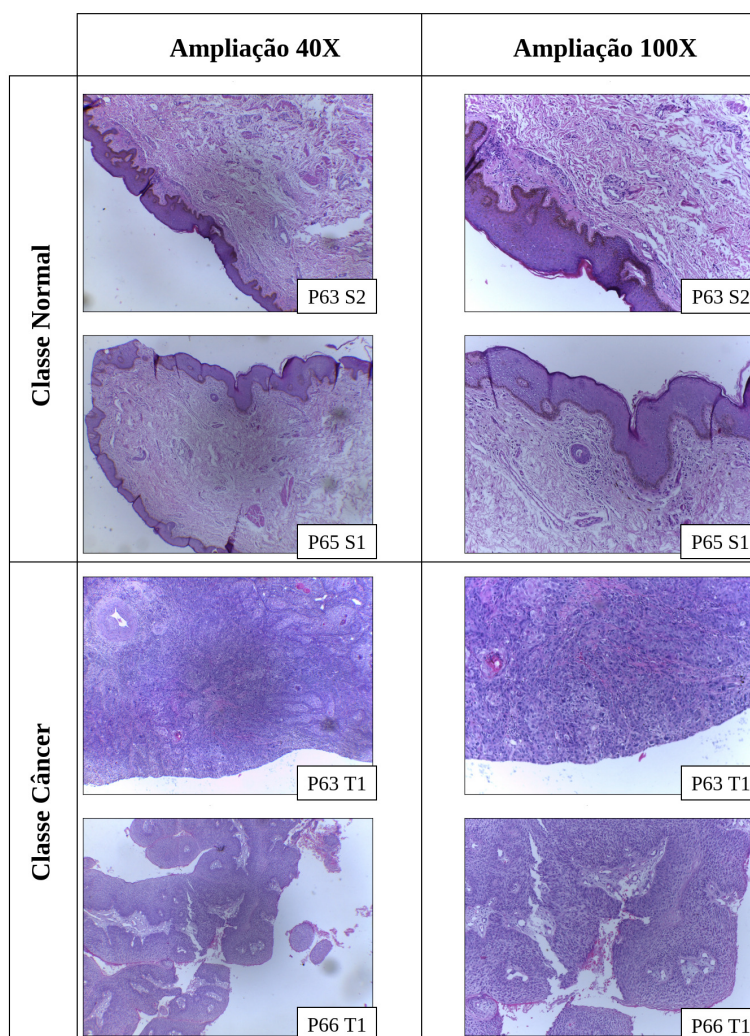
O conjunto de imagens utilizadas para a produção deste trabalho foi organizado pelo Projeto Câncer de Pênis da Amazônia Legal (PCPAm) em colaboração com Núcleo de Computação Aplicada da Universidade Federal do Maranhão. Este consiste em 194 imagens RGB de resolução 2048×1536 pixels, organizadas por ampliação e situação patológica conforme a tabela 2.

Esses registros foram obtidos em 2021, a partir de biópsias penianas. As amostras recolhidas foram coloridas com Eosina e Hematoxilina e fotografadas por uma câmera de alta definição (Leica ICC50 HD) acoplada a um microscópio de campo claro (Leica DM500), em duas ampliações distintas. A análise e classificação das imagens foram realizadas por dois patologistas com o auxílio de um software específico (Leica Aperio

**Tabela 2. Organização do conjunto de imagens.**

Ampliação	Normal	Câncer	Total
40x	40	57	97
100x	40	57	97

ImageScope). A Figura 2 apresenta alguns exemplos de imagens, com identificação do paciente e da lâmina utilizada para fotografar a amostra recolhida, organizadas por classe e ampliação.



**Figura 2. Exemplos de imagens presentes no conjunto de dados**

Para a condução deste estudo, optou-se por utilizar apenas a ampliação de 100x. A decisão de utilizar exclusivamente a ampliação de 100x foi baseada na escolha do método de pré-processamento que requer um número significativo de treinamentos.

### 3.2. Pré-processamento e Augmentation

Um tópico muito importante no desenvolvimento deste trabalho é o pré-processamento e o *Augmentation*.

Para a tarefa de pré-processamento, foi selecionado o método Vahadane, desenvolvido e apresentado em [Vahadane et al. 2016]. Este método é uma técnica avançada utilizada na análise de imagens histopatológicas para corrigir variações de cor e tornar as imagens mais consistentes e comparáveis. O método de Vahadane é especialmente útil em imagens histopatológicas, onde variações na coloração podem surgir devido a diferentes protocolos de coloração, condições de iluminação e características do tecido biológico.

Para solucionar falta de padronização, este método envolve a escolha de uma imagem alvo como parâmetro. A imagem de referência, conhecida como imagem alvo ou modelo, é utilizada para normalizar a cor das outras imagens histopatológicas. A distribuição de cor da imagem alvo é usada como referência para ajustar a distribuição de cor das demais imagens, garantindo assim uma consistência de cor entre elas. A escolha da imagem parâmetro para o pré-processamento é um fator muito importantes e será discutida na seção de treinamento. Após a aplicação do método Vahadane, as imagens foram redimensionadas para a resolução 256x192, em virtude das limitações computacionais.

Além do pré-processamento, também foram aplicadas técnicas de *Augmentation* para lidar com dois problemas presentes no conjunto de imagens: o desbalanceamento de classes e a escassez de imagens para o treinamento.

O *Augmentation* consiste em criar novas imagens por meio de transformações nas imagens disponíveis. Para lidar com o desbalanceamento de classes, novas imagens da classe “Normal” foram geradas no conjunto de treino com a utilização da transformação de matriz e saturação (*Hue Saturation*), uma ferramenta que permite ajustar a tonalidade das cores e sua intensidade de maneira flexível. Nesta ocasião, a transformação foi aplicada de maneira branda, com todos os limites de mudança no intervalo (-10,10), visando manter as características da imagem original, mas com algumas diferenças leves. Para lidar com as poucas imagens do dataset, foram utilizadas transformações geométricas: rotação em 90°, 180° e 270°, inversão horizontal e vertical.

### 3.3. Desenvolvimento de Modelos

Para a tarefa de classificação, foram selecionadas abordagens relacionadas à rede neural DenseNet. Essa rede é composta pela repetição de blocos densos, cujas saídas são concatenadas com a entrada do próximo bloco. Essa concatenação, conhecida como conexão densa, desempenha um papel crucial na resolução do problema de “gradient vanishing” e contribui para o processo de aprendizagem. Interpolados com os blocos densos, encontram-se os blocos de transição, responsáveis por reduzir a resolução do sinal. Ao final, a saída da parte convolucional do modelo é passada por uma camada densamente conectada para gerar a predição.

A Figura 3 apresenta a arquitetura da DenseNet201 utilizada neste estudo. Os blocos densos são compostos por duas repetições da camada de *BatchNormalization*, seguidas pela convolução 2D de 3x3 e ativação ReLU, nessa ordem, e concatenados com o sinal de entrada. Durante as sucessões de blocos densos, eles são repetidos 6, 12, 48 e 32 vezes, respectivamente. Após cada sucessão de blocos densos, há um bloco de transição, composto por uma camada de *BatchNormalization*, seguida de uma ativação ReLU, uma convolução 2D de 3x3 e uma camada de *Average Pooling 2D*. Após todos os blocos densos, o sinal passa por uma camada de *Flatten* e por uma MLP, que consiste em duas camadas densamente conectadas de 256 neurônios, com dropout de 50%. A predição é gerada por

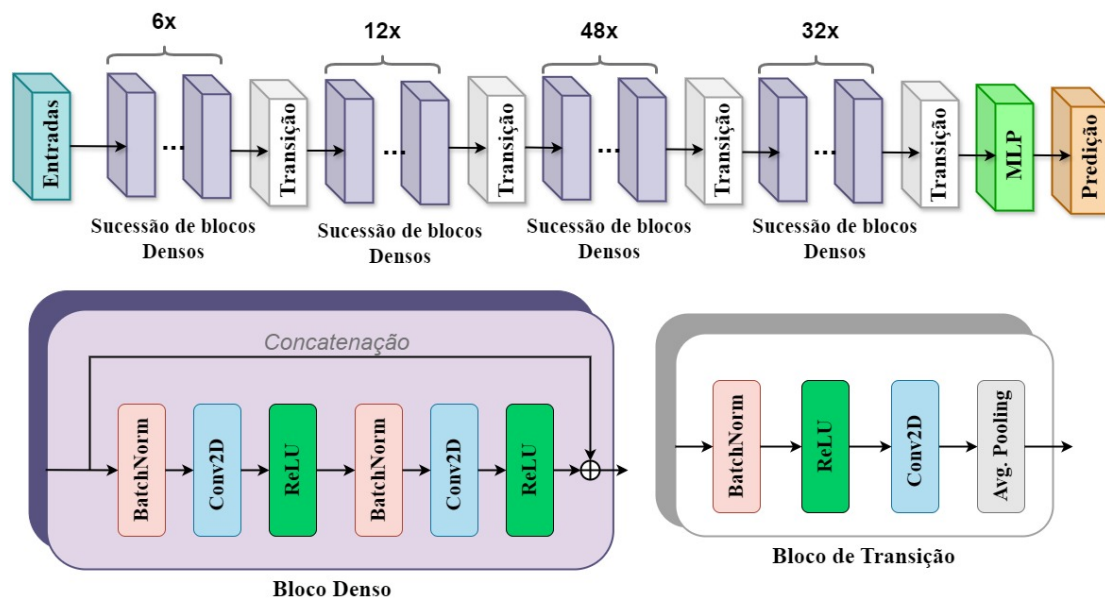


Figura 3. Arquitetura da rede neural DenseNet201.

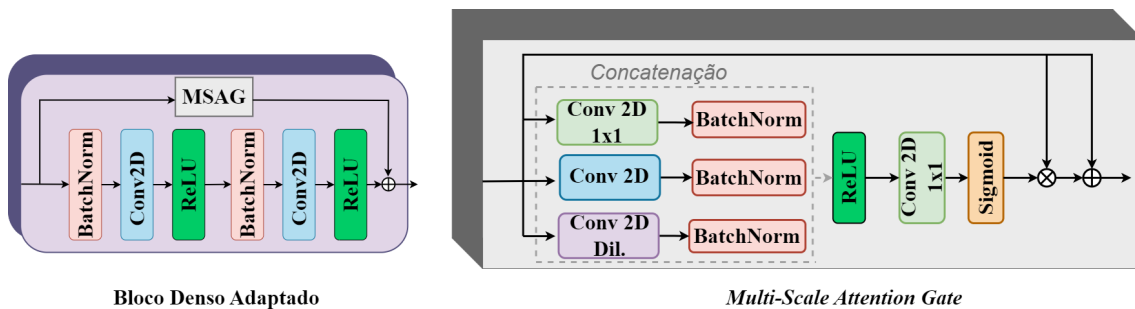
meio de uma camada densa com 2 neurônios, seguida da ativação softmax.

As adaptações realizadas na arquitetura da DenseNet foram aplicadas em seus blocos densos, mais especificamente na concatenação. A primeira adaptação, chamada DenseMSAG, incorpora o *Multi-Scale Attention Gate* (MSAG) em suas concatenações, um mecanismo de atenção apresentado por [Tang et al. 2023]. Este mecanismo consiste em três tipos de convoluções: convolução com filtros de tamanho 1x1, conhecida como *Point Wise*; convoluções padrão com filtros de tamanho 3x3; e convoluções dilatadas com filtros de tamanho 3x3 e uma taxa de dilatação igual a 2.

A utilização dessas três convoluções auxilia na extração de características com vários campos receptivos diferentes, aprimorando a supressão de recursos de baixa importância e a captura de características-alvo de forma mais eficaz. Os resultados dessas convoluções passam por uma camada de *Batch Normalization* e são concatenados. Após a concatenação, as informações ainda passam por uma ativação ReLU, uma convolução 2D de 3x3 e uma ativação Sigmoid. Por fim, o produto dessas operações é multiplicado pela entrada e somado a ela. A Figura 4 demonstra a arquitetura do MSAG e seu local de aplicação.

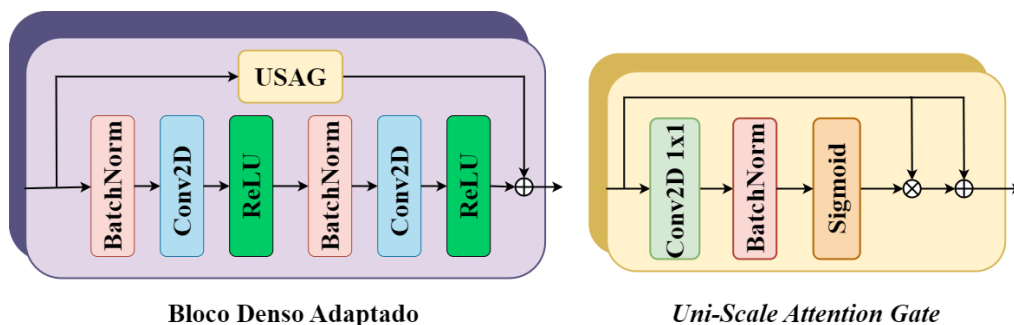
Outra adaptação analisada é chamada de DenseUSAG. Essa abordagem também se inspirou no mecanismo de atenção em múltiplas escalas (MSAG), porém, surgiu como uma solução para um problema enfrentado durante o treinamento da DenseMSAG.

Devido à complexidade do MSAG, o número de parâmetros da rede tornou-se muito elevado, impossibilitando sua utilização em todos os blocos densos da rede, sendo aplicado apenas nos blocos das duas primeiras sucessões. A solução para este problema foi chamada de *Uni-Scale Attention Gate* (USAG). Este mecanismo de atenção surgiu da redução do número de convoluções utilizadas no MSAG, mantendo apenas a convolução 1x1 (*Point Wise*), como ilustrado na Figura 5.



**Figura 4. Arquitetura do *Multi-Scale Attention Gate* e seu local de aplicação nos blocos convolucionais da DenseNet.**

A escolha por essa redução foi feita tendo em vista que à medida que a rede se torna mais profunda, mais características importantes são extraídas. Portanto, utilizar os mecanismos de atenção nos blocos mais profundos auxiliaria a filtrar essas informações de maior relevância. Apesar da grande redução de parâmetros entre o MSAG e o USAG, ainda não foi possível utilizar este mecanismo em toda a rede, mas permitiu sua aplicação nas três primeiras sucessões de blocos densos.



**Figura 5. Arquitetura do *Uni-Scale Attention Gate* e seu local de aplicação nos blocos densos da DenseNet.**

## 4. Resultados

Nesta seção, serão detalhados os resultados dos experimentos realizados para validar o método proposto. Os modelos DenseMSAG e DenseUSAG foram implementados utilizando os *frameworks* Keras e TensorFlow [Joseph et al. 2021]. O pré-processamento Vahadane foi conduzido por meio da biblioteca StainTools. Todos os treinamentos foram iniciados com 50 épocas, com um critério de parada antecipada (*early stop*) de 30 épocas, salvando os pesos do último estágio e do estágio com menor função de perda na validação. O otimizador AdamW [Loshchilov and Hutter 2017] foi utilizado devido sua eficiência computacional e pouca necessidade de memória. Este otimizador consiste em uma extensão do otimizador Adam que incorpora uma penalidade de regularização diretamente à função de perda durante o treinamento, auxiliando na prevenção de *overfitting*. A taxa de aprendizado foi definida como 0,0001, com um decaimento para 0,00001 ao longo do treinamento.

Para avaliação dos resultados, as métricas adotadas para monitorar o progresso do treinamento foram a acurácia, a precisão, a sensibilidade e o *F1-Score*. A utilização



de métricas diferentes é fundamental para uma avaliação abrangente e precisa de uma abordagem de classificação, pois cada métrica fornece uma perspectiva única sobre o desempenho do modelo e destaca diferentes aspectos da capacidade de classificação.

Os resultados deste trabalho foram obtidos a partir de duas sessões distintas de treinamento. Devido à natureza do método de pré-processamento Vahadane, que requer a escolha de uma imagem como parâmetro, a primeira sessão de treinamento foi realizada para determinar quais imagens proporcionam os melhores resultados como parâmetro de pré-processamento para os dois modelos implementados. Inicialmente, o conjunto de dados foi dividido utilizando o método *Hold out*, no qual 50% das imagens foram alocadas para o conjunto de treinamento, 40% para o conjunto de teste e 10% para o conjunto de validação. Para a aplicação do método Vahadane, foram selecionadas todas as imagens da classe “Câncer” presentes no conjunto de treino, totalizando 27 imagens, das quais foram gerados 27 novos conjuntos de imagens pré-processadas para o treinamento dos modelos DenseMSAG e DenseUSAG.

Como resultado da primeira fase, a imagem P49 T1 alcançou as melhores métricas, servindo como parâmetro de pré-processamento para o modelo DenseMSAG. Este experimento alcançou acurácia de 94,8%, precisão de 100,0%, sensibilidade de 91,3% e F1-Score de 95,4% no conjunto de teste. Já para o modelo DenseUSAG, a imagem com melhores métricas foi a P35 T. O treinamento utilizando esta imagem como parâmetro atingiu acurácia de 94,8%, precisão de 100,0%, sensibilidade de 91,3% e F1-Score de 95,4%.

Com as imagens já selecionadas, iniciou-se a segunda fase de treinamento: aplicação da validação cruzada sobre os melhores pré-processamentos. O método de validação cruzada escolhido foi o Monte Carlo, no qual o conjunto de dados é dividido em conjuntos de treinamento, teste e validação várias vezes de forma aleatória, gerando diversas divisões diferentes. Em seguida, o modelo é treinado e avaliado em cada divisão, e os resultados são combinados para calcular estatísticas resumidas, como média e desvio padrão. Foram realizadas 5 partições distintas do conjunto de dados, sempre garantindo que a imagem utilizada para o pré-processamento estivesse presente no conjunto de treinamento.

Os resultados da validação cruzada foram agrupados e, a partir deles, foram calculadas a média e o desvio padrão, os quais estão apresentados na Tabela 3. Nela, também estão disponíveis outras informações importantes, como as imagens utilizadas como parâmetro para o método Vahadane, o modelo com os melhores resultados obtidos, marcado em verde, e o estágio do qual foram obtidos os pesos. O termo ‘Última época’ refere-se aos pesos salvos na última época do treinamento, enquanto ‘Melhor época’ refere-se aos pesos salvos na época com o menor valor da função de perda no conjunto de validação. É importante ressaltar que todos os resultados foram obtidos a partir de treinamentos na ampliação de 100x apenas, conforme explicado na Seção 3.1.

**Tabela 3. Resultados obtidos pela metodologia proposta**

Trabalhos	Estágio	Imagem Parâm.	Acurácia	Precisão	Sensibilidade	F1-Score
DenseMSAG	Melhor época	P49 T1	82,56% ± 4,93%	93,75% ± 4,16%	75,65% ± 9,02%	83,45% ± 5,13%
DenseUSAG	Melhor época	P35 T	92,30% ± 4,44%	95,22% ± 5,05%	92,17% ± 10,83%	93,16% ± 4,63%
DenseMSAG	Última época	P49 T1	80,00% ± 7,11%	93,97% ± 4,77%	70,43% ± 9,91%	80,32% ± 7,85%
DenseUSAG	Última época	P35 T	88,20% ± 6,93%	95,78% ± 5,81%	84,34% ± 13,95%	88,95% ± 7,25%

#### 4.1. Discussão

Esta seção é dedicada à discussão dos resultados obtidos pelos modelos treinados, contrastando-os com os resultados apresentados em estudos relacionados. Essa comparação é apresentada na Tabela 4.

**Tabela 4. Comparação entre as abordagens deste trabalho e dos trabalhos relacionados para a ampliação 100x.**

Trabalhos	Acurácia	Precisão	Sensibilidade	F1-Score
[Belfort et al. 2023]	91,2% ± 1,4%	88,4% ± 3,3%	97,6% ± 5,6%	92,4% ± 1,5%
[Lauande et al. 2022]	96,8% ± 4,2%	96,5% ± 4,2%	98,1% ± 3,6%	97,3% ± 3,6%
DenseUSAG	92,3% ± 4,4%	95,2% ± 5,0%	92,1% ± 10,8%	93,1% ± 4,6%
DenseMSAG	82,5% ± 4,9%	93,7% ± 4,1%	75,6% ± 9,0%	83,4% ± 5,1%

Observa-se que o modelo DenseUSAG demonstra resultados competitivos em relação à literatura, uma vez que esta abordagem apresenta resultados próximos aos do trabalho de [Lauande et al. 2022] e supera o trabalho de [Belfort et al. 2023] em acurácia, precisão e F1-Score.

Em relação à acurácia, o modelo alcançou valores de 92,3% ± 4,4%, indicando uma boa proporção de exemplos classificados corretamente em relação ao total. Considerando a precisão, os resultados do modelo são de 95,2% ± 5,0%, o que demonstra que esta abordagem apresenta poucos casos de falsos positivos. Em relação à sensibilidade, a DenseUSAG alcançou apenas 92,1% ± 10,8%, valor inferior quando comparado aos outros trabalhos. A métrica da sensibilidade é de grande relevância em estudos que tratam do diagnóstico de doenças ao avaliar o modelo quanto à presença de falsos negativos, os quais podem representar uma falsa percepção de segurança. Essa variação em relação às outras métricas indica a necessidade de soluções para mitigar os falsos negativos. Por fim, ao avaliar o F1-Score, o modelo atingiu 93,1% ± 4,6%.

De forma geral, os resultados destacam a aplicabilidade da arquitetura desenvolvida neste trabalho e do método Vahadane como ferramenta de pré-processamento. Também evidenciam os pontos que devem ser revisados para que, no futuro e com as devidas modificações, esta abordagem possa apresentar um desempenho ainda melhor.

Avaliando a rede DenseMSAG, seus resultados foram significativamente inferiores aos demais. Além do número reduzido de mecanismos de atenção em comparação com a DenseUSAG, outra questão identificada ao longo do treinamento desta arquitetura foi sua propensão ao *overfitting*. Essa conclusão foi alcançada ao comparar as métricas obtidas pelo modelo nos conjuntos de treino e teste, que apresentavam uma grande discrepância. Assim, embora o modelo não tenha alcançado os resultados das outras abordagens, ele pôde aprender algumas relações e pode oferecer melhores resultados em trabalhos futuros com a otimização de sua arquitetura, juntamente com a aplicação de outras técnicas de aprendizado de máquina para mitigar o *overfitting*.

#### 5. Conclusão

Neste estudo, foram propostos dois modelos para o diagnóstico automático de câncer de pênis em imagens histopatológicas, juntamente com a utilização de uma ferramenta de pré-processamento ainda não utilizada neste domínio.

Os resultados foram obtidos exclusivamente para a ampliação 100x, devido à natureza do método Vahadane e a alguns testes preliminares. O modelo DenseUSAG demonstrou resultados comparáveis à literatura, evidenciando sua aplicabilidade neste contexto. Ele alcançou um F1-Score de  $93,1\% \pm 4,6\%$ .

Os resultados obtidos são promissores e há espaço para melhorias significativas. A DenseMSAG pode ser aprimorada através da otimização de sua arquitetura e da implementação de estratégias para mitigar o *Overfitting*. Embora a DenseUSAG tenha resultados comparáveis à literatura, ela ainda apresenta algumas limitações, como uma sensibilidade inferior e elevados desvios padrão em suas métricas. Para abordar essas questões, em futuros trabalhos, pretende-se explorar a utilização de otimizadores e técnicas para lidar com o *Overfitting* em ambos os modelos, além de considerar a incorporação de novos mecanismos de atenção em arquiteturas da família DenseNet.

## 6. Agradecimentos

Os autores agradecem à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Código Financeiro 001, ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), à Empresa Brasileira de Serviços Hospitalares (Ebserh) (Concesso número 409593/2021-4) e à Fundação de Amparo à Pesquisa Desenvolvimento Científico e Tecnológico do Maranhão (FAPEMA) pelo apoio financeiro.

## Referências

- Belfort, F., Silva, I., Silva, A., and Paiva, A. (2023). Detecção de câncer peniano em imagens histopatológicas usando redes neurais convolucionais em cascata. In *Anais do XXIII Simpósio Brasileiro de Computação Aplicada à Saúde*, pages 328–339, Porto Alegre, RS, Brasil. SBC.
- Erickson, B. J., Korfiatis, P., Akkus, Z., and Kline, T. L. (2017). Machine learning for medical imaging. *Radiographics*, 37(2):505–515.
- Huang, G., Liu, Z., Van Der Maaten, L., and Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708.
- INCA (2022). Câncer de pênis: Instituto nacional do câncer - inca. Acesso em 13/05/2024.
- Joseph, F. J. J., Nonsiri, S., and Monsakul, A. (2021). Keras and tensorflow: A hands-on experience. *Advanced deep learning for engineers and scientists: A practical approach*, pages 85–111.
- Lauande, M. G. M., Teles, A. M., da Silva, L. L., Matos, C. E. F., Junior, G. B., de Paiva, A. C., de Almeida, J. D. S., da Costa Oliveira, R. M. G., Brito, H. O., Nascimento, A. G., et al. (2022). Classification of histopathological images of penile cancer using densenet and transfer learning. In *VISIGRAPP (4: VISAPP)*, pages 976–983.
- Loshchilov, I. and Hutter, F. (2017). Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*.
- Machado, A. and Veloso, L. (2021). Classificação de distúrbios pulmonares em radiografias de tórax usando redes convolucionais. In *Simpósio Brasileiro De Telecomunicações E Processamento De Sinais*. SBC.

- Melo, R. C., Raas, M. W., Palazzi, C., Neves, V. H., Malta, K. K., and Silva, T. P. (2020). Whole slide imaging and its applications to histopathological studies of liver disorders. *Frontiers in medicine*, 6:310.
- MS (2022). Câncer de pênis: Saúde de a à z, ministério da saúde. Acesso em 13/05/2024.
- Oliveira, M. V., Matos, C. E. F., Júnior, G. B., de Paiva, A. C., de Almeida, J. D. S., Costa, G., Levy, M., and Freitas, M. (2022). Ppm-unet: Uma rede neural convolucional para a segmentação de rins em imagens de tc. In *Anais do XXII Simpósio Brasileiro de Computação Aplicada à Saúde*, pages 268–276. SBC.
- Tang, F., Wang, L., Ning, C., Xian, M., and Ding, J. (2023). Cmu-net: A strong convmixer-based medical ultrasound image segmentation network. In *2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI)*. IEEE.
- Thomas, A., Necchi, A., Muneer, A., Tobias-Machado, M., Tran, A. T. H., Van Rompuy, A.-S., Spiess, P. E., and Albersen, M. (2021). Penile cancer. *Nature Reviews Disease Primers*, 7(1):11.
- Vahadane, A., Peng, T., Sethi, A., Albarqouni, S., Wang, L., Baust, M., Steiger, K., Schlitter, A. M., Esposito, I., and Navab, N. (2016). Structure-preserving color normalization and sparse stain separation for histological images. *IEEE transactions on medical imaging*, 35(8):1962–1971.