

Análise comparativa da influência dos espaços de cores na segmentação multi-classe de *Whole Slide Imaging* do câncer de mama utilizando *deep learning*

Felipe André Zeiser¹, Cristiano André da Costa¹, Gabriel de Oliveira Ramos¹, Adriana Vial Roehé², Henrique C. Bohn¹, Ismael Santos¹, Rodrigo da Rosa Righi

¹Software Innovation Laboratory - SOFTWARELAB
Applied Computing Graduate Program
Universidade do Vale do Rio dos Sinos (UNISINOS)
São Leopoldo, Brazil

²Departamento de Patologia e Medicina Legal
Universidade Federal de Ciências da Saúde de Porto Alegre
Porto Alegre, Brasil

Abstract. *Early diagnosis usually results in a better prognosis in cancer cases. Currently, histopathological analysis is the gold standard for the diagnosis, staging and definition of the treatment of breast neoplasms. However, the technique has some restrictions that hinder the process of investigation by the pathologist, such as different staining protocols, color variations on the slides, and tissue overlap. CNNs have been shown to be adequate to deal with this type of problem. Thus, the computational representation of colors can have a significant influence on the behavior of convolutional models. Therefore, this study proposes a comparative analysis of the influence of four color spaces (RGB, HSV, YCrCb, and LAB) for the multi-class segmentation of Whole Slide Imaging (WSI) of breast cancer. The proposed methodology consists of a pre-processing step for WSI, data augmentation, and multi-segmentation using the U-Net convolutional architecture with a pre-trained ResNet-50 as an encoder. The results obtained demonstrate that the HSV color space presented, in general, better indexes for the segmentation of WSI using the methodology proposed in this study.*

Resumo. *O diagnóstico em estágios iniciais normalmente resulta em um prognóstico melhor nos casos de câncer. Atualmente, a análise histopatológica é o padrão ouro para o diagnóstico, estadiamento e definição do tratamento de neoplasias de mama. Contudo, a técnica possui algumas restrições que dificultam o processo de análise pelo patologista, como diferentes protocolos de coloração, variações de coloração nas lâminas, e sobreposição de tecidos. Assim, a representação computacional das cores pode exercer uma influência significativa no comportamento dos modelos convolucionais. Desta forma, este estudo propõe uma análise comparativa da influência de quatro espaços de cores (RGB, HSV, YCrCb e LAB) para a segmentação multi-classe de Whole Slide Imaging (WSI) do câncer de mama. A metodologia proposta é composta por uma etapa de pré-processamento das WSI, data augmentation e multi-segmentação utilizando a arquitetura convolucional U-Net com uma ResNet-50 pré-treinada como codificador. Os resultados obtidos demonstram que o espaço*

de cores HSV apresentou de maneira geral índices melhores para a segmentação das WSI utilizando a metodologia proposta neste estudo.

1. Introdução

Estimativas apontam a neoplasia de mama como o câncer com a maior taxa de morbidade e mortalidade no mundo [World Health Organization 2019a]. As taxas de incidência e mortalidade para o câncer de mama são ainda mais preocupantes em países com o Índice de Desenvolvimento Humano (IDH) elevado [World Cancer Research Fund 2019]. Embora alguns fatores hereditários estejam relacionados ao desenvolvimento da doença (cerca de 5% a 10%), os principais fatores de risco são relacionados a: menstruação (precoce ou menopausa tardia), reprodução (idade avançada), ingestão hormonal (contraceptivos ou reposição hormonal), excesso de peso, consumo de bebidas alcoólicas e tabagismo [Bray et al. 2018, World Cancer Research Fund 2019].

No Brasil, as taxas de incidência do câncer de mama são ainda mais preocupantes se comparadas às médias globais. Para um grupo de 100 mil mulheres, aproximadamente 63 possuem câncer de mama, enquanto a média global é em torno de 46 [World Health Organization 2019a]. As estimativas apontavam para quase 60 mil novos casos e cerca de 15 mil mortes para o ano de 2018 no Brasil em decorrência do câncer de mama [INCA 2019].

Métodos não-invasivos como a mamografia, ultrassonografia ou ressonância magnética são utilizados no Brasil para a detecção do câncer de mama, seguindo uma orientação da Organização Mundial da Saúde (OMS) [INCA 2019]. Contudo, para confirmação do diagnóstico é necessária uma análise histopatológica de uma seção da área suspeita [World Health Organization 2019b].

As seções de tecido coletadas da área suspeita são geralmente coradas com Hematoxilina e Eosina (H&E) [Roy et al. 2019]. A hematoxilina é responsável pela coloração do DNA do núcleo celular, do RNA do citoplasma e da cartilagem, produzindo uma cor azulada nas estruturas [Roy et al. 2019]. Enquanto a eosina tingem estruturas do citoplasma e o colágeno de tons rosas [Yan et al. 2019]. O processo de coloração é necessário por que são a maior parte das células e material extracelular serem incolores [Mescher 2018].

Contudo, algumas características na análise histológica provocam variações intra e inter observador no diagnóstico [Mittal and Saraswat 2019]. Estas variações ocorrem principalmente pelas sutis diferenças entre as células cancerígenas dos tipos de câncer de mama. Além disso, as propriedades químicas dos reagentes, diferenças nos protocolos utilizados e as habilidades dos patologistas provocam variações nas análises das lâminas histopatológicas [Khan et al. 2015].

As cores presentes nas lâminas histopatológicas ou em imagens digitalizadas de lâminas inteiras (*Whole Slide Imaging* — WSI) transmitem grande quantidade de informações e desempenham um papel importante no diagnóstico. Na computação, as cores são representadas por espaços de cores, que objetivam representar as cores visíveis pelo ser humano. Contudo, existem diferenças significativas na quantidade de cores do espectro eletromagnético visível que cada modelo é capaz de representar [Gonzalez and Woods 2018].

Portanto, este trabalho tem como objetivo analisar o comportamento da multi-

segmentação de WSI do câncer de mama para quatro espaços de cores (RGB, HSV, YCbCr e LAB) utilizando uma *Convolutional Neural Network* (CNN). A arquitetura de rede do artigo foi baseada na U-Net [Ronneberger et al. 2015], utilizando como codificador uma ResNet50 pré-treinada para o conjunto de imagens ImageNet. Nesse âmbito, a principal contribuição científica do trabalho é a análise da influência dos espaços de cores no comportamento de uma arquitetura de segmentação de WSI do câncer de mama.

O artigo está dividido em seis seções principais. Na Seção 2, são apresentadas algumas características da análise histopatológica do câncer de mama. Enquanto na Seção 3, são discutidos os trabalhos relacionados. A Seção 4 descreve os materiais e métodos. Na Seção 5, são apresentados os resultados e a discussão. Finalmente, a Seção 6 destaca as principais conclusões e orientações para trabalhos futuros.

2. Histopatologia Aplicada ao Câncer de Mama

A técnica de coloração de lâminas por H&E já é empregada há mais de um século, sendo, hoje, amplamente difundida pelos patologistas por ser uma técnica barata, simples e, principalmente, funcional [Chan 2014].

A coloração H&E cora as estruturas por relacionar afinidades químicas do corante às estruturas celulares simples. Por seu caráter ácido e de carga negativa (aniônico), a Eosina, corante de tom rosa, é denominada um corante ácido, corando estruturas celulares de caráter básico, e sendo assim, estruturas celulares denominadas acidófilas. Como exemplos de estruturas acidófilas podemos citar o citoplasma celular (também chamado de citosol) e as mitocôndrias [Kumar et al. 2016]. Já a Hematoxilina, corante de tom azulado, tem caráter básico e carga positiva (catiônico), corando estruturas celulares de caráter ácido denominados basófilos [Roy et al. 2019]. Como exemplo de estrutura basófila podemos citar o próprio núcleo celular, que é composto pelos seus grupamentos ácidos.

A diversidade de cores produzidas pela interação dos corantes ácidos e básicos na lâmina, torna possível deduzir e avaliar até mesmo questões funcionais da célula em questão. A grande gama de tons produzidos, a intensidade do corante no meio e a textura realçada pelo pigmento que se apresenta nas lâminas expõe informações celulares importantes ao meio médico [Chan 2014], capazes de comunicar e diferenciar, inclusive, processos celulares normais ou patológicos [Mittal and Saraswat 2019].

A análise base para diagnóstico e tipagem do câncer de mama, embora todo avanço tecnológico de pesquisas de base molecular, é a pesquisa de material biológico coletado e analisado via microscopia óptica [Chan 2014]. O câncer de mama possui diversas subclassificações, que são determinadas pelo tipo de estrutura acometida [Makki 2015]. Assim, a anatomia e histologia mamárias como ductos e lóbulos, tipos de células epiteliais (luminal e miopitelial) e tipos de estroma (inter ou intralobular) se faz importante para o diagnóstico e classificação [Kumar et al. 2016].

Contudo os processos de preparação, coloração e digitalização de lâminas de tecidos podem sofrer com fixações inadequadas, que podem gerar alterações na morfologia do tecido e induzir a morfometria incorreta desse tecido pela análise de imagens [Roy et al. 2019]. Outra dificuldade encontrada são seções que não foram adequadamente coradas (podendo ser essas super ou sub-manchadas) e montadas, isto poderá resultar em regiões fora de foco e/ou peças ausentes [Li et al. 2018]. Além disso, pequenas

variações nas condições de coloração podem levar a variações na aparência do tecido [Yan et al. 2019].

3. Trabalhos Relacionados

As imagens histopatológicas do câncer de mama são objeto de estudo há alguns anos [Spanhol et al. 2016]. As metodologias propostas nos estudos podem ser subdivididas em dois grupos: baseadas em extração de recursos e em CNN. Os métodos baseados em extração de recursos buscam representar uma imagem em características significativas para o problema. Neste contexto, o trabalho de [Balazsi et al. 2016] segmenta regiões de imagens histopatológicas e extrai características, como padrões binários locais e histogramas de gradientes orientados, para identificar carcinoma ductal invasivo utilizando Random Forest. Outra forma de representar uma imagem histológica do câncer de mama é com a utilização de descritores wavelet de cores [Issac Niwas et al. 2012].

Após a etapa de extração de recursos, pode se utilizar o Principal Component Analysis (PCA) para identificar as características mais representativas para então realizar a classificação, conhecido também como feature selection [Sabeena Beevi et al. 2019]. Contudo, na maioria dos estudos, a classificação é realizada sem a etapa de feature selection. A classificação é realizada utilizando técnicas de machine learning, como k-Nearest Neighbors (kNN) [Khan et al. 2015], Support Vector Machines [Tashk et al. 2015], ou Decision Trees [Baker et al. 2018].

No entanto, a utilização de metodologias baseadas na extração de recursos é dependente do conhecimento do problema para obter recursos que representem de forma ideal a imagem [Spanhol et al. 2016]. Em contraponto, as CNN aprendem de forma automática as características necessárias para classificação, detecção ou segmentação [Li et al. 2018].

As CNN utilizam um grande poder computacional por estas imagens conterem bilhões de pixels o que dificulta a utilização de WSI diretamente nas redes [Guo et al. 2019]. A forma mais utilizada na literatura para contornar este problema, até o momento, é a utilização de patches (segmentos das WSI) [Cruz-Roa et al. 2014].

Na classificação das WSI utilizando patches, a imagem é dividida geralmente em uma matriz de regiões não sobrepostas que são classificadas individualmente pelas CNN. Alternativamente, as lesões classificadas como positivas podem então ser segmentadas, por exemplo, utilizando uma rede Fully Convolutional Network (FCN) [Guo et al. 2019]. Outra alternativa é a utilização de FCN empilhadas realizando progressivamente a segmentação de regiões doentes de ampliações maiores para as menores [Gecer et al. 2018].

Além do tamanho, existem as variações de cores presentes nas WSI devido à propriedades químicas dos reagentes, diferenças sutis entre os tecidos, e a presença de artefatos nas imagens [Khan et al. 2015]. Desta forma, as cores e texturas podem influenciar diretamente no comportamento dos modelos de classificação, detecção ou segmentação. Computacionalmente, as cores são representadas apenas por números. Contudo, existem diversos formatos para realizar esta representação, sendo os mais comuns RGB, HSV, YCrCb e LAB. À vista disso, abre-se a oportunidade de estudo dos espaços de cores e a influência desses sobre a segmentação utilizando CNN.

4. Materiais e Métodos

Este trabalho segue o método apresentado na Fig. 1. A metodologia, basicamente, está dividida em quatro etapas: pré-processamento, *data augmentation*, treinamento e teste. O pré-processamento das imagens consiste em conversão dos espaços de cores e normalização das imagens. A etapa *data augmentation* descreve os métodos utilizados para criar imagens sintéticas. Já na etapa de treinamento, são definidos a CNN e os parâmetros utilizados. Finalmente, a etapa de teste apresenta os resultados e o comportamento da CNN para cada espaço de cor.

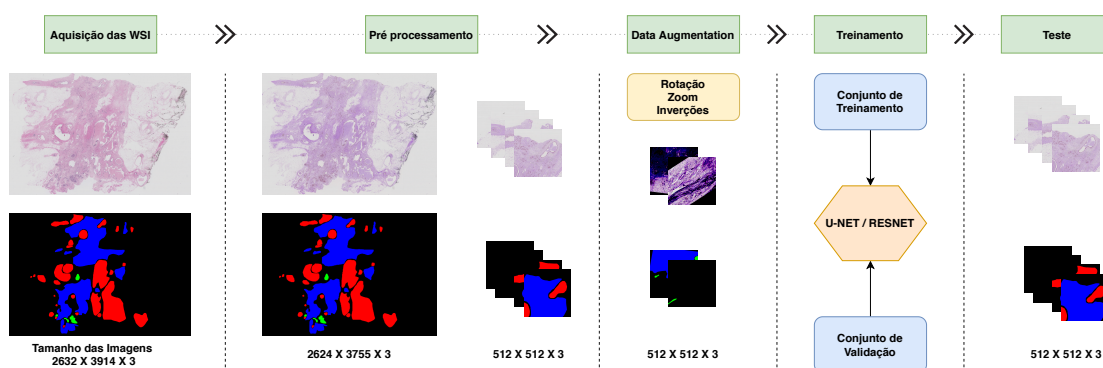


Figura 1. Diagrama da metodologia proposta, com base em um modelo U-Net e Resnet para a segmentação de WSI do câncer de mama.

4.1. Aquisição das Imagens

A base de imagens utilizada neste trabalho foi a *Grand Challenge on Breast Cancer Histology images* (BACH) [Aresta et al. 2019]. Esta base consiste em dois conjuntos denominados A e B. O primeiro consiste de 400 imagens de secções de WSI com quatro classes: normal, benigno, carcinoma in situ e carcinoma invasivo. O conjunto B consiste de 40 WSI de alta resolução contendo a totalidade do tecido amostrado. As 40 WSI do conjunto B, são divididas em dois subconjuntos, sendo 30 para treinamento e 10 para teste. Das 30 WSI fornecidas para treinamento, somente 10 são rotuladas a nível de pixel em quatro classes: normal, benigno, carcinoma in situ e carcinoma invasivo. Desta forma, neste trabalho são utilizadas somente 10 WSI, pela necessidade da rotulagem a nível de pixel.

As WSI da base foram adquiridas entre 2013 e 2015 em pacientes de Portugal utilizando um scanner de slide digital Leica SCN400 e disponibilizadas em formato SVS. Ainda, cada arquivo SVS fornece dez níveis de ampliação para as amostras, sendo o mínimo de 0,04x e o máximo de 20x. Já quanto aos rótulos, foi fornecido um arquivo XML contendo as coordenadas para cada região rotulada [Aresta et al. 2019].

Apesar da base de imagens ser relativamente pequena é importante destacar o tamanho destas imagens com bilhões de pixels [Aresta et al. 2019]. Desta forma, as técnicas de *transfer learning* e *data augmentation* são úteis para auxiliar o processo de aprendizado dos modelos com *datasets* de poucas imagens. A metodologia adotada na utilização destas técnicas será abordadas nas próximas subseções.

4.2. Pré-processamento

Nesta etapa, utilizamos as WSI numa ampliação de 2,5 vezes. A utilização das WSI nesta ampliação é devida à capacidade computacional disponível. Como forma de acelerar o processo de segmentação das WSI removemos as bordas brancas presentes nas WSI utilizando a segmentação adaptativa de Otsu [Otsu 1979] para separar o fundo do tecido. Após a segmentação, são realizadas as normalizações das imagens pela média e desvio padrão da intensidade global dos pixels [Spanhol et al. 2016].

As imagens segmentadas foram divididas em três conjuntos, treinamento (70%), validação (10%) e teste (20%). As CNN necessitam de uma quantidade significativa de imagens no processo de treinamento para obter a generalização do problema. Desta forma, cada conjunto passou pela extração de *patches* não sobrepostos de 256x256 pixels.

4.3. Data augmentation

Uma técnica que auxilia na convergência e aprendizado das CNN é a utilização de *data augmentation*. Neste estágio, utilizamos a classe *ImageDataGenerator* do *framework Keras* aplicada aos conjuntos de treinamento e validação. A classe em questão gera em tempo real dados sintéticos com base nas imagens extraídas das WSI. As transformações aplicadas aos *patches* nesta etapa foram: rotações, zoom e inversões horizontais e verticais.

4.4. Modelo convolucional

O modelo CNN foi baseado na arquitetura conhecida como U-NET [Ronneberger et al. 2015]. Esta arquitetura pode ser dividida em duas partes principais: *downsampling* e *upsampling*. O *downsampling* extrai características e reduz gradualmente o tamanho espacial das imagens. Enquanto o *upsampling* reconstrói as imagens utilizando o processo de deconvolução e concatenação com as características extraídas no processo de *downsampling*.

Como forma de acelerar o processo de aprendizado da rede utilizamos a ResNet50 [He et al. 2016] pré-treinada para o conjunto ImageNet para a etapa de *downsampling*. Os pequenos saltos entre as camadas convolucionais da arquitetura ResNet50 permite uma convergência mais rápida e com menor possibilidade de ocorrer a fuga de gradiente [He et al. 2016].

Na Fig. 2 é apresentada a arquitetura convolucional utilizada neste trabalho. A taxa de aprendizado utilizada no treinamento foi de 0.0001, um *batch size* de 16 imagens, *padding same* para as camadas convolucionais e de *maxpooling* e o algoritmo *Adam* para otimização dos pesos.

5. Resultados e Discussão

Nesta seção serão apresentados e avaliados os resultados obtidos utilizando o modelo convolucional proposto para cada um dos espaços de cores selecionados (RGB, HSV, YCrCb e LAB).

5.1. Treinamento e validação dos modelos

Durante o treinamento dos modelos para cada espaço de cor, a métrica utilizada para mensurar a perda a cada época foi o *Dice Index*. Esta métrica mede a similaridade entre

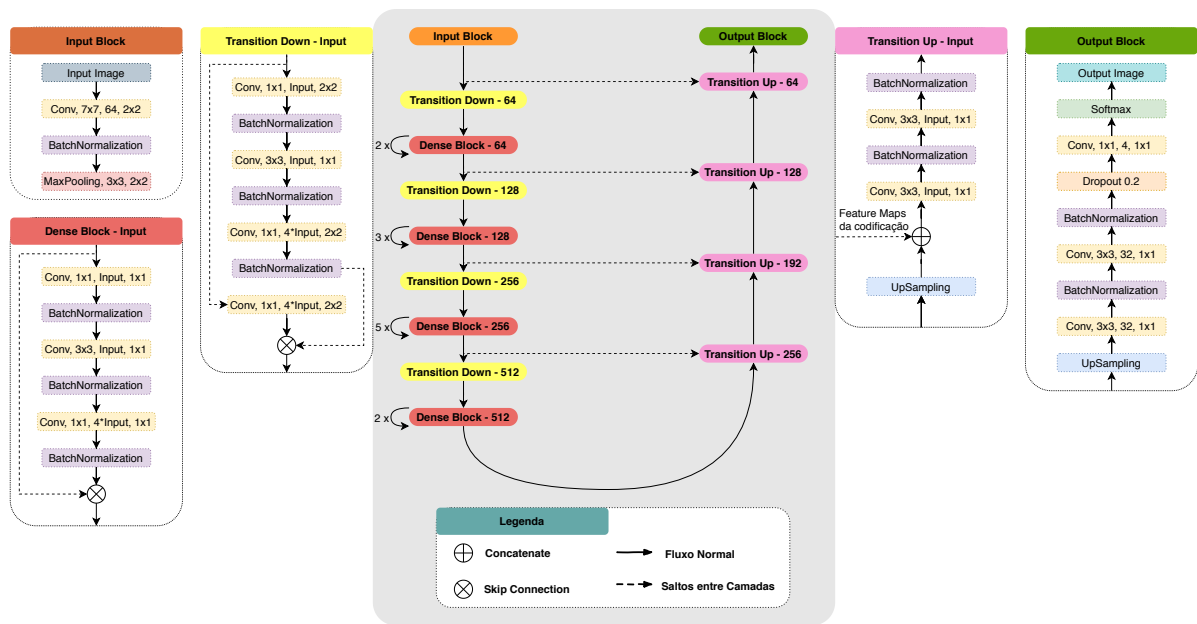


Figura 2. Arquitetura convolucional para a segmentação das WSI. A parte em cinza da figura representa o fluxo resumido da rede. Ao lado de cada *Dense Block* está especificado a quantidade de vezes que este será repetido na arquitetura. No lado esquerdo, estão representados os blocos detalhados do *downsampling* das imagens. Na parte direita, estão representados os blocos detalhados do *upsampling* das imagens. Cada camada convolucional está representada da seguinte forma: *kernel*, *feature maps* e *stride*.

dois conjuntos de dados, com base na sobreposição dos dois conjuntos. Esta métrica pode ser definida como:

$$DSC(A, B) = \frac{2|A \cap B|}{|A + B|} \quad (1)$$

onde A e B são dois conjuntos quaisquer. Ainda, quanto mais próximos de 1.0 o valor obtido pela métrica, melhor é a generalização do modelo.

Os modelos foram treinados por 50 épocas com base no conjunto de treinamento, e validados ao final de cada época no conjunto de validação para obter o melhor conjunto de pesos. Na Fig. 3 são apresentados os índices obtidos para cada época no conjunto de treinamento e validação.

Todos os modelos apresentaram comportamentos estatisticamente similares para o conjunto de treinamento e validação. Para os quatro espaços de cores testados podemos ver uma relativa estabilidade no crescimento do *Dice Index* no conjunto de treinamento e pequenas variações no conjunto de validação. Este comportamento indica que não ocorreu *overfitting* durante o treinamento.

5.2. Validação dos modelos

Os resultados para etapa de validação dos modelos são com base no conjunto de testes para cada espaço de cor. A escolha do melhor conjunto de pesos foi realizada automaticamente com base na perda em cada época para o conjunto de validação. A Tab. 1 apresenta

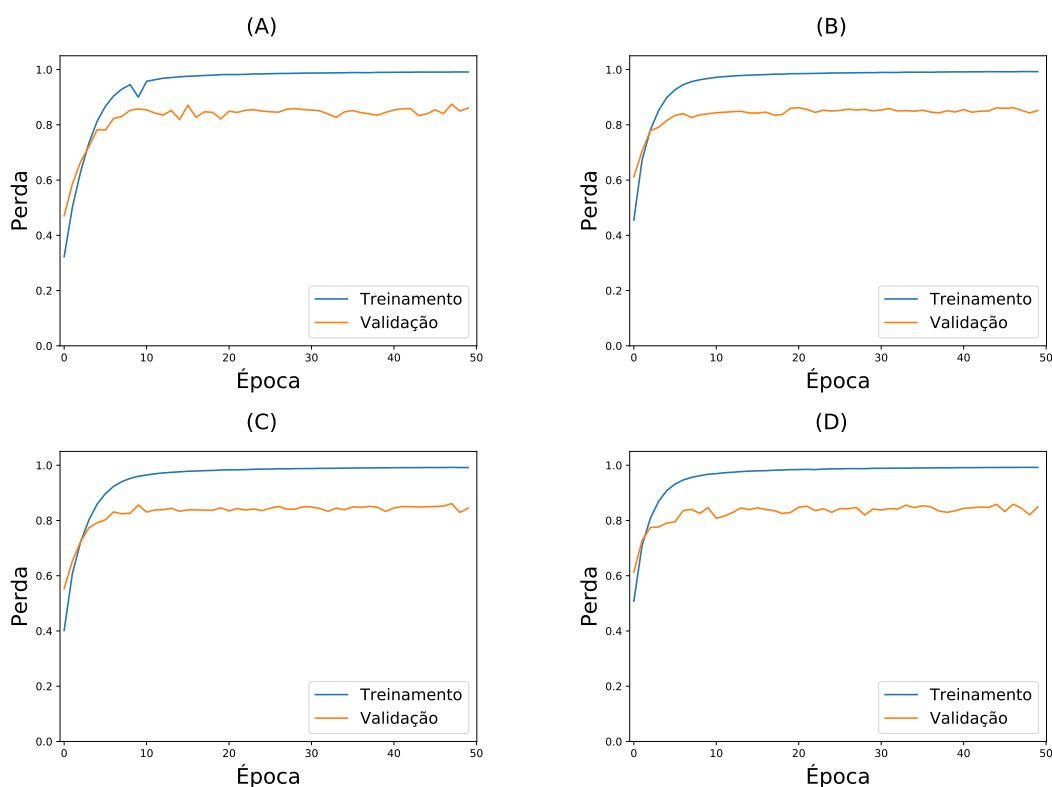


Figura 3. Gráficos de perda para os dados de treinamento e validação para cada modelo: (A) RGB; (B) HSV; (C) YCrCb; (D) LAB.

os valores obtidos para as métricas a nível de pixel e para o Dice Index no conjunto de testes.

Tabela 1. Resultados para o conjunto de teste para cada espaço de cor: (A) RGB; (B) HSV; (C) YCrCb; (D) LAB. Em negrito os melhores resultados. (Acu. pixel é a acurácia a nível de pixel. Já as colunas *Benigno*, *In situ*, *Invasivo* e *Fundo* correspondem ao percentual de pixels classificados corretamente para cada uma destas classes).

Modelo	Dice Index	Acu. pixel	Benigno	In situ	Invasivo	Fundo
A	0.6923	84.67%	78.22%	82.63%	76.64%	91.49%
B	0.7085	83.85%	80.95%	72.73%	89.83%	90.51%
C	0.6881	84.85%	71.56%	74.17%	81.84%	93.70%
D	0.6979	83.80%	73.45%	75.03%	80.81%	91.75%

Analisando os resultados obtidos para a métrica de *Dice Index* podemos, novamente, notar a semelhança estatística entre os índices. A maior variação entre o *Dice Index* é de 2.04%, entre o YCrCb e o HSV. Esta semelhança se reflete nos percentuais obtidos para acurácia a nível de pixel, que apresenta uma variação de no máximo 1.05% entre os espaços de cores.

Contudo, quando analisamos os percentuais a nível de pixel para cada classe esta similaridade não ocorre. Onde, por exemplo, para a classificação de regiões benignas a variação percentual é 9.39% entre os espaços de cores. Já para o tecido acometido pelo

carcinoma invasivo, verifica-se uma melhor classificação para o modelo convolucional do espaço HSV com um percentual de 89.83% de precisão. Na classificação de áreas com carcinoma in situ, o espaço RGB se demonstrou mais efetivo para a detecção de pixels pertencentes à regiões doentes.

De maneira geral, o modelo para o espaço de cores HSV apresentou os melhores resultados quantitativos, contudo demonstra alguma dificuldade para detectar o tecido benigno. Outra conclusão que podemos ter é que, mesmo apresentando a melhor acurácia à nível de pixel, o modelo YCrCb apresentou resultados significativamente inferiores para as regiões doentes. Isto é reflexo da alta precisão para a classificação de pixels como fundo, que são áreas proporcionalmente maiores nas imagens, refletindo um peso maior para a acurácia global.

5.3. Avaliação qualitativa

Nesta seção, apresentamos os resultados para a multi-segmentação de uma WSI (Fig. 4) do conjunto de testes para cada um dos espaços de cores. Na Fig. 4 a cor preta indica o fundo, ou tecido saudável. A cor vermelha indica a presença de tecido com características benignas. A cor verde indica a presença de carcinoma in situ. Já a cor azul indica a presença de carcinoma invasivo.

Analisando o desempenho qualitativo novamente, o espaço de cor que apresenta o melhor resultado é o modelo para o espaço de cores HSV (modelo (B)). A melhor representação, seja quantitativa ou qualitativa, pode estar associada à forma que são representadas as cores no espaço HSV, onde um canal só representa a cor base, e os outros dois canais pelo brilho e a saturação da cor [Gonzalez and Woods 2018]. Desta forma, o modelo pode ser mais robusto para pequenas variações, pois o canal *Hue* apresenta a cor em sua forma primária, sem variações de brilho ou saturação diminuindo.

Especificamente para o caso A01, é possível perceber que os modelos A, C e D detectam de maneira mais eficiente regiões com carcinoma in situ. Enquanto isso, o modelo B apresenta uma precisão significativamente superior para detecção de áreas benignas e carcinoma invasivo.

6. Conclusão

Neste trabalho, comparamos o desempenho de quatro espaços de cores, RGB, HSV, YCrCb e LAB, na multi-segmentação de WSI do câncer de mama utilizando CNN. A multi-segmentação foi realizada utilizando a arquitetura U-Net com uma ResNet-50 pré-treinada como codificador. Como forma de contornar a pequena quantidade de imagens disponíveis utilizamos *data augmentation*, aplicando zoom, rotações e inversões horizontais e verticais em *patches* não sobrepostos extraídos das WSI.

Esta pesquisa limitou-se a estudar a influência dos espaços de cores na multi-segmentação de WSI do câncer de mama do BACH dataset utilizando a arquitetura convolucional apresentada na Seção 4.4. Além disto, os resultados apresentados neste estudo são baseados em uma ampliação de 2.5 vezes.

A principal contribuição científica deste estudo foi a análise da influência de espaços de cores na segmentação de WSI do câncer de mama. Além disso, os resultados indicam que é possível um estudo para a utilização da metodologia como um mecanismo de segunda opinião.

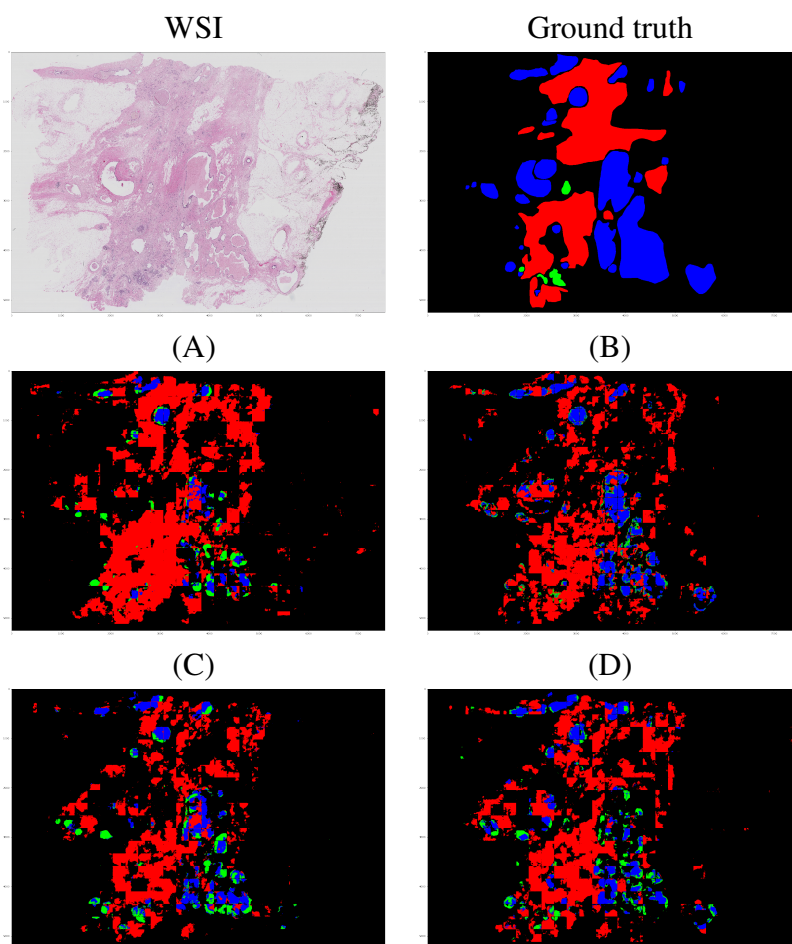


Figura 4. Caso A01 do BACH com a segmentação fornecida pela base e os resultados para os modelos: (A) RGB; (B) HSV; (C) YCrCb; (D) LAB.

Como trabalhos futuros, prevê-se o estudo de outros métodos convolucionais para segmentação de WSI. Uma alternativa a ser estudada é a utilização de segmentação binária e posterior classificação das regiões doentes por outro método. Além disso, pretendemos estudar o comportamento da arquitetura proposta em um conjunto de dados maiores em parceria com outras instituições. Ainda, uma arquitetura híbrida combinando a segmentação obtida pelos quatro espaços de cores pode direcionar trabalhos futuros.

Agradecimentos

Os autores gostariam de agradecer à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES (Código Financeiro 001), Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq (Números de concessão 303640 / 2017-0 e 405354 / 2016-9) e ao NVIDIA GPU Grant Program pelo seu apoio neste trabalho.

Referências

Aresta, G., Araújo, T., Kwok, S., Chennamsetty, S. S., Safwan, M., Alex, V., Marami, B., Prastawa, M., Chan, M., Donovan, M., Fernandez, G., Zeineh, J., Kohl, M., Walz, C., Ludwig, F., Braunewell, S., Baust, M., Vu, Q. D., To, M. N. N., Kim, E., Kwak,

- J. T., Galal, S., Sanchez-Freire, V., Brancati, N., Frucci, M., Riccio, D., Wang, Y., Sun, L., Ma, K., Fang, J., Kone, I., Boulmane, L., Campilho, A., Eloy, C., Polónia, A., and Aguiar, P. (2019). Bach: Grand challenge on breast cancer histology images. *Medical Image Analysis*, 56:122 – 139.
- Baker, Q. B., Zaitoun, T. A., Banat, S., Eaydat, E., and Alsmirat, M. (2018). Automated Detection of Benign and Malignant in Breast Histopathology Images. In *15th International Conference on Computer Systems and Applications*, pages 1–5. IEEE.
- Balazsi, M., Blanco, P., Zoroquiain, P., Levine, M. D., and Burnier, M. N. (2016). Invasive ductal breast carcinoma detector that is robust to image magnification in whole digital slides. *Journal of Medical Imaging*, 3(2).
- Bray, F., et al. (2018). Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*.
- Chan, J. K. C. (2014). The wonderful colors of the hematoxylin–eosin stain in diagnostic surgical pathology. *International Journal of Surgical Pathology*, 22(1):12–32.
- Cruz-Roa, A., Basavanhally, A., González, F., Gilmore, H., Feldman, M., Ganesan, S., Shih, N., Tomaszewski, J., and Madabhushi, A. (2014). Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. *Medical Imaging 2014: Digital Pathology*, 9041:904103.
- Gecer, B., Aksoy, S., Mercan, E., Shapiro, L. G., Weaver, D. L., and Elmore, J. G. (2018). Detection and classification of cancer in whole slide breast histopathology images using deep convolutional networks. *Pattern Recognition*, 84:345–356.
- Gonzalez, R. C. and Woods, R. E. (2018). *Digital Image Processing*. Pearson, New York, 4 edition.
- Guo, Z., Liu, H., Ni, H., Wang, X., Su, M., Guo, W., Wang, K., Jiang, T., and Qian, Y. (2019). A Fast and Refined Cancer Regions Segmentation Framework in Whole-slide Breast Pathological Images. *Scientific Reports*, 9(1):882.
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778.
- INCA (2019). Câncer de mama.
- Issac Niwas, S., Palanisamy, P., Chibbar, R., and Zhang, W. J. (2012). An expert support system for breast cancer diagnosis using color wavelet features. *Journal of Medical Systems*, 36(5):3091–3102.
- Khan, A. M., Sirinukunwattana, K., and Rajpoot, N. (2015). A Global Covariance Descriptor for Nuclear Atypia Scoring in Breast Histopathology Images. *IEEE Journal of Biomedical and Health Informatics*, 19(5):1637–1647.
- Kumar, V., Abbas, A. K., and Aster, J. C. (2016). *Robbins & Cotran Patologia — Bases Patológicas das Doenças*. Elsevier, São Paulo, 9 edition.
- Li, C., Wang, X., Liu, W., and Latecki, L. J. (2018). DeepMitosis: Mitosis detection via deep detection, verification and segmentation networks. *Medical Image Analysis*, 45:121–133.

- Makki, J. (2015). Diversity of breast carcinoma: histological subtypes and clinical relevance. *Clinical Medicine Insights: Pathology*, 8:CPath–S31563.
- Mescher, A. L. (2018). *Junqueira's Basic Histology: Text and Atlas, Thirteenth Edition: Text and Atlas, Thirteenth Edition*. McGraw-Hill Education, New York, 15 edition.
- Mittal, H. and Saraswat, M. (2019). An automatic nuclei segmentation method using intelligent gravitational search algorithm based superpixel clustering. *Swarm and Evolutionary Computation*, 45:15–32.
- Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66.
- Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015*, pages 234–241, Cham. Springer International Publishing.
- Roy, K., Banik, D., Bhattacharjee, D., and Nasipuri, M. (2019). Patch-based system for Classification of Breast Histology images using deep learning. *Computerized Medical Imaging and Graphics*, 71:90–103.
- Sabeena Beevi, K., Nair, M. S., and Bindu, G. R. (2019). Automatic mitosis detection in breast histopathology images using Convolutional Neural Network based deep transfer learning. *Biocybernetics and Biomedical Engineering*, 39(1):214–223.
- Spanhol, F. A., Oliveira, L. S., Petitjean, C., and Heutte, L. (2016). Breast cancer histopathological image classification using Convolutional Neural Networks. In *Proceedings of the International Joint Conference on Neural Networks*, volume 2016-Octob, pages 2560–2567. Institute of Electrical and Electronics Engineers Inc.
- Tashk, A., Helfroush, M. S., Danyali, H., and Akbarzadeh-jahromi, M. (2015). Automatic detection of breast cancer mitotic cells based on the combination of textural, statistical and innovative mathematical features. *Applied Mathematical Modelling*, 39(20):6165–6182.
- World Cancer Research Fund (2019). Breast cancer statistics.
- World Health Organization (2019a). Cancer.
- World Health Organization (2019b). Cancer - diagnosis and treatment.
- Yan, R., Ren, F., Wang, Z., Wang, L., Ren, Y., Liu, Y., Rao, X., Zheng, C., and Zhang, F. (2019). A Hybrid Convolutional and Recurrent Deep Neural Network for Breast Cancer Pathological Image Classification. *IEEE International Conference on Bioinformatics and Biomedicine*, pages 957–962.