

Aprendizado profundo para assistência histopatológica: um estudo de classificação de micrometástases em câncer de mama

Gabriela Kuhn¹, Felipe André Zeiser¹, Adriana Roehé²,
Cristiano André da Costa¹, Gabriel de Oliveira Ramos¹

¹Programa de Pós-Graduação em Computação Aplicada
Universidade do Vale do Rio dos Sinos
São Leopoldo – RS – Brasil

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

Abstract. *Deep Learning algorithms for detecting micrometastasis in breast cancer can improve the gold standards and the efficiency of pathologists' routines. An analysis of the literature indicates that the models still need to improve their performance in identifying micrometastases and tumor cells. In the case of isolated tumor cells, the detection rates were below 40%. There are also opportunities to improve on positive rates, since many models still detect nerves or contaminations as false micrometastases. In this study, we investigated the implementation to perform the classifications of metastasis. The goal is that this classification neural network become the first layer of a two-layer final neural network model - classification and segmentation. As partial results, this work achieved an AUC = 0.998 for the classification task on the slide patch.*

Resumo. *Algoritmos de Deep Learning para detecção de micrometástases em câncer de mama podem melhorar os padrões-ouro e a eficiência das rotinas dos patologistas. Uma análise da literatura indica que os modelos ainda precisam melhorar seu desempenho na identificação de micrometástases e células tumorais. No caso de células tumorais isoladas, as taxas de detecção ficaram abaixo de 40%. Também há oportunidades para melhorar as taxas de falsos positivos, pois muitos modelos ainda detectam nervos ou contaminações como falsas micrometástases. Neste estudo, investigamos a implementação de uma rede neural para a classificação de metástases. O objetivo é que esta rede neural de classificação seja a primeira camada de um modelo final composto por duas redes neurais - classificação e segmentação. Como resultados parciais, este trabalho obteve uma AUC = 0,998 para a tarefa de classificação nos fragmentos da lâmina.*

1. Introdução

O câncer é um dos principais problemas de saúde pública em todo o mundo. O câncer de mama é um dos cânceres mais comuns e mortais em mulheres [Ferlay et al. 2021]. Apesar do prognóstico para pacientes com câncer de mama ter melhorado nos últimos anos, este prognóstico é significativamente prejudicado quando o câncer de mama apresenta metástase [Howlander et al. 2018]. Detectar a presença de células metastáticas é

um fator importante no prognóstico do câncer de mama [Siegel et al. 2021]. No entanto, este processo é meticuloso e demorado e pode levar à perda de micrometástases [Litjens et al. 2018a].

Os avanços na digitalização de imagens histopatológicas e a disponibilização de conjuntos de dados relevantes intensificaram estudos para o uso clínico de algoritmos de visão computacional relacionados ao auxílio no diagnóstico de metástases em câncer de mama. A avaliação da presença de células tumorais nos linfonodos sentinelas axilares através de exames histopatológicos figura como um dos métodos mais precisos para determinar a localização e a extensão de uma metástase de câncer de mama, permanecendo como um dos mais importantes fatores prognósticos [Cserni et al. 2021]. Essa tarefa no entanto é exaustiva ao patologista e suscetível a falhas, principalmente para a análise de micrometástase, que são áreas metastáticas na escala de menos de 2 mm [Cardoso et al. 2019].

Estudos apontam o potencial benefício das aplicações de Deep Learning em áreas de patologia digital e aplicações médicas relacionadas, embora ciente das limitações desta tecnologia em termos de raciocínio e interpretação humana [Dwivedi et al. 2021]. Trabalhos sugerem que alguns algoritmos podem até exceder a sensibilidade de um patologista para detectar focos de câncer individuais em imagens digitais, por isso, aliar esse processo à rotina patológica pode aumentar o sucesso do diagnóstico [Ehteshami Bejnordi et al. 2017].

Na literatura atual existe uma limitação em relação ao conjunto de dados e a complexidade computacional de se manipular imagens de tamanho na escala de Gigapixel, no qual se apresentam as imagens histopatológicas digitalizadas. Quanto às limitações dos modelos já implementados, muitos ainda falham na identificação de micrometástases e falham em gerar falsos positivos devido a confusão com outras células de características semelhantes presentes no linfonodo.

Este trabalho tem como objetivo auxiliar na investigação da aplicação de algoritmos de Deep Learning para a detecção de micrometástases em câncer de mama. Para isso, este artigo apresenta os resultados encontrados para a rede neural implementada, realizando a classificação das imagens metastáticas. O objetivo é que esta rede seja a primeira de um modelo final composto de duas camadas de redes neurais de aprendizado profundo. A primeira, apresentada neste trabalho, para realizar a classificação das metástases, seguida de uma segunda rede neural para realizar a segmentação das células tumorais. Para a tarefa de classificação, o algoritmo proposto alcançou valores para a Área sob a Curva (AUC) = 0.998 nos teste a nível de fragmentos da lâmina, enquanto a baseline definida apresenta valor para a tarefa de classificação AUC = 0.994.

As principais contribuições deste trabalho são:

- Investigar uma rede de classificação de imagens para que a mesma seja a primeira rede de um modelo composto por duas etapas para detectar micrometástase no linfonodo sentinela com base em imagens histopatológicas.
- Implementar um novo pipeline de pré-processamento para Whole-Slide-Images (WSI) para lidar com a complexidade computacional de se manipular essas imagens de alta resolução e para realizar o rotulamento das imagens para o treinamento.

- Uma revisão de literatura buscando verificar as oportunidades e limitações do uso de modelos de visão computacional para a classificação de metástases no câncer de mama.

2. Trabalhos Relacionados

O estudo [Lee and Paeng 2018] propõe realizar uma classificação do linfonodo baseado no sistema *pN-staging* o qual é uma estratégia para entender o risco atual para o estadiamento nodal no câncer de mama. O estudo também tem foco na previsão do pN-staging das metástases como potencial para reduzir a carga de trabalho e ajudar o patologista utilizando um modelo baseado em uma Rede Neural Convolucional (CNN) *Random Forest*. O método alcançou uma pontuação kappa de 0,9203 no conjunto de dados Camelyon17.

No trabalho realizado por [Liu et al. 2019] foi utilizado o conjunto de dados de treinamento Camelyon16 para treinar um modelo de detecção denominado LYNA. Uma contribuição importante deste trabalho encontra-se no fato de os achados terem sido comparados com conjunto de dados de diferentes laboratórios independentes para medir a reprodutibilidade de treinamento do Camelyon16. O LYNA alcançou uma AUC de 99.6%.

Também utilizando uma CNN, [Wang et al. 2021] criou um modelo para prever e classificar metástases nos linfonodos sentinelas. Ao observarem que as resoluções de imagem para este treinamento são geralmente menores do que a resolução de entrada do modelo projetado, [Wang et al. 2021] propuseram um método de aumento de dados chamado Random Center Cropping (RCC) que amplia a escala dos conjuntos de dados, mas também preserva a resolução e a área central das imagens, alcançando uma precisão de até 97.96% para detectar metástases.

Detectando que a maioria dos trabalhos usa o patch como um indivíduo independente para treinar o modelo, [Zhang et al. 2019] desenvolveu uma rede de contexto espacial múltiplo (MSC-Net) com o objetivo de mitigar a falta de precisão causada pelo corte, considerando as correlações espaciais entre patches vizinhos.

Uma análise de imagens de alto rendimento por apresentar rápida velocidade de processamento foi apresentada por [Wollmann and Rohr 2017]. O método executa a classificação em pequenos patches e usa a média do modelo para diminuir a variância e o bias.

No estudo realizado por [Turki et al. 2021], foram testadas diferentes arquiteturas de aprendizado profundo, incluindo *DenseNet121*, *ResNet50*, *VGG16*, *Xception*. Também foram utilizados e avaliados algoritmos de aprendizado como *Random Forest*, *AdaBoost* e *SVM*. Neste estudo, os resultados experimentais mostraram que DenseNet121 gerou melhores resultados de desempenho com base nas medidas usadas em problemas de desequilíbrio de classe.

O método proposto por [Valkonen et al. 2017] é baseado em *feature engineering* e aprendizado supervisionado com um modelo *Random Forest*. Assim, o método extrai das imagens de entrada características importantes para a detecção de metástases que incluem textura da imagem, estrutura espacial e distribuição de núcleos. Os resultados apresentados demonstram que o método detecta áreas metastáticas em linfonodos com boa precisão (AUC 0,97–0,98 para detecção de tumor em toda a área da imagem, AUC 0,84–0,91 para tumor *versus* detecção de tecido normal).

Usando uma técnica de normalização de cores genérica, [Chen et al. 2016] treinou uma rede neural convolucional profunda baseada em patches para realizar uma avaliação de metástases baseada em lâminas. O trabalho alcançou resultados de sensibilidade 0,96, especificidade 0,89 e pontuação AUC de 0,90 o que indica que o modelo pode ser aplicado a WSI para regiões metastáticas sem calibração adicional para os perfis de coloração.

2.1. Limitações e oportunidades

A análise dos resultados e a matriz de confusão dos mesmos revelam que os trabalhos ainda precisam melhorar a performance para identificar micrometástases e células tumorais isoladas de forma confiável. Neste último caso em [Bandi et al. 2018] as taxas de detecção ficaram abaixo de 40%.

Há também espaço para explorar categorias de falsos positivos, visto que muitos modelos ainda detectam nervos ou contaminações como falsas micrometástases. Se averigua também que é preciso melhorar a pontuação de falsos positivos e explorar a carência da amplitude de conhecimento e experiência de patologistas que os algoritmos possuem para, por exemplo, detectar e distinguir as micrometástases de outras anormalidades que o modelo não foi explicitamente treinado para classificar.

A arquitetura proposta objetiva amenizar o problema de confusão das células metastáticas com outras células ou contaminações, desta forma, neste estudo as imagens não são redimensionadas, utilizando as lâminas na resolução original e fragmentando as mesmas para realizar a classificação. Assim, pretende-se que a rede neural aprenda um maior número de características das células metastáticas. Para mitigar o problema de falsos negativos, a classificação é feita a nível de fragmentos das imagens, denominados de *patches*, penalizando para erros de falsos negativos.

3. Abordagem Proposta

Este trabalho visa o estudo de uma rede neural de classificação de metástases a nível de patches de lâminas histopatológicas. Esta abordagem é composta por uma etapa de pré-processamento das imagens WSI, para amenizar problemas causados pelas técnicas histopatológicas, e também amenizar problemas de processamento computacional dessas imagens, dividindo a imagem original em patches. Em seguida esses patches alimentam a entrada da rede de classificação onde os mesmos são classificados como positivo ou negativo para a classe metástase. O objetivo futuro é que, após essa classificação, os patches positivos serão a entrada de dados de uma segunda rede neural que fará a segmentação a nível de pixel dos patches positivos, localizando a região metastática, criando a máscara da classe metástase e dimensionando o tamanho da mesma, realizando a correspondência pixel-mm correspondente ao conjunto de dados, produzindo o resultado final.

3.1. Conjunto de dados

O conjunto de dados escolhido para os experimentos iniciais é o Camelyon16 [Litjens et al. 2018b]. Dois fatores amenizados na etapa de pré-processamento foram a dificuldade de processar essas imagens devido ao seu tamanho - cerca de 2 GB cada - e a variabilidade de cores nas imagens que a tonalização em diferentes laboratórios pode causar.

3.2. Pré-processamento dos dados

Na primeira etapa do pré-processamento das imagens foi realizado o processo de normalização das cores. Para isso, implementou-se uma função para normalizar as tonalizações, tendo como referência o trabalho proposto por [Macenko et al. 2009]. Este processo pode ser visto na Figura 1.

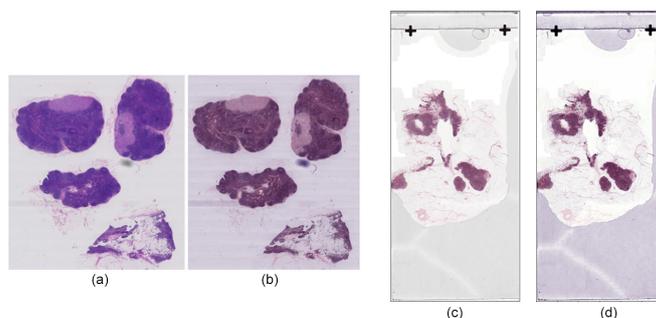


Figura 1. A imagem mostra o antes, (a) e (c), e depois, (b) e (d), da aplicação do filtro para corrigir a variabilidade de cores em duas lâminas histopatológicas de dois laboratórios diferentes no conjunto de dados Camelyon16. [Litjens et al. 2018b]

Após a normalização, para reduzir o tempo e o processamento computacional e melhorar a ativação das redes para as metástases, o fundo da imagem foi removido, focando nas regiões da lâmina que contém tecido, preenchendo com valores $[0, 0, 0]$.

Em seguida, a imagem original e a máscara anotada foram fragmentadas em patches, armazenando a posição matricial do patch em relação a lâmina original para posterior correlação espacial e reconstrução. Este processo foi realizado para lidar com os problemas relacionados a complexidade computacional sem necessariamente sacrificar a resolução e, conseqüentemente, características importantes da imagem.

As imagens foram divididas em patches de 224×224 pixels, movendo exatos 224 pixels da esquerda para a direita - *stride* - com a imagem na máxima resolução. O tamanho 224 foi determinado através da proporção matemática com a relação pixel-mm das lâminas que é $1792px = 1mm$. O mesmo processo foi realizado para as máscaras de anotações, criando patches de máscaras exatos correspondentes aos patches das imagens.

Os patches relevantes para o treinamento foram selecionados baseando-se na quantidade de tecido presente, sendo essa igual ou maior a 25% do patch, determinado empiricamente. A última etapa do pré-processamento consiste na rotulação dos patches como positivo ou negativo para metástase. Para averiguar a classe de cada patch, foi analisado o centro 56×56 pixels da imagem, se ao menos um pixel dentro dessa região central era positivo o patch, o foi rotulado como positivo. Este processo é apresentado na Figura 2

3.3. Arquitetura do modelo

O modelo proposto utiliza uma arquitetura de Rede Neural Convolutiva para realizar a análise das características extraídas das imagens. Através das operações de convolução e de subamostragem, é realizada a análise das imagens para a extração das características. Posteriormente, através das camadas totalmente conectadas definimos as relações que diferenciam as células metastáticas das não-metastáticas.

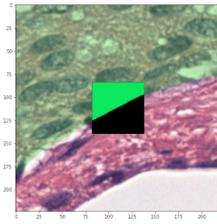


Figura 2. Na imagem, podemos ver o patch fragmentado. A transparência verde corresponde à máscara anotada pelo patologista e o recorte analisado se encontra no centro da imagem. A região do corte central que contém metástase está na cor verde.

Neste estudo, foram analisados os resultados para diferentes valores de hiperparâmetros com o objetivo de alcançar um melhor desempenho e também evitar problemas com *overfitting*. Assim, diferentes configurações da arquitetura da rede neural para classificação, bem como funções para melhorar o treinamento, foram investigadas. A fim de facilitar essa tarefa, utilizamos os recursos disponíveis na API Keras. A Figura 3 mostra uma visão geral da estrutura do modelo de classificação.

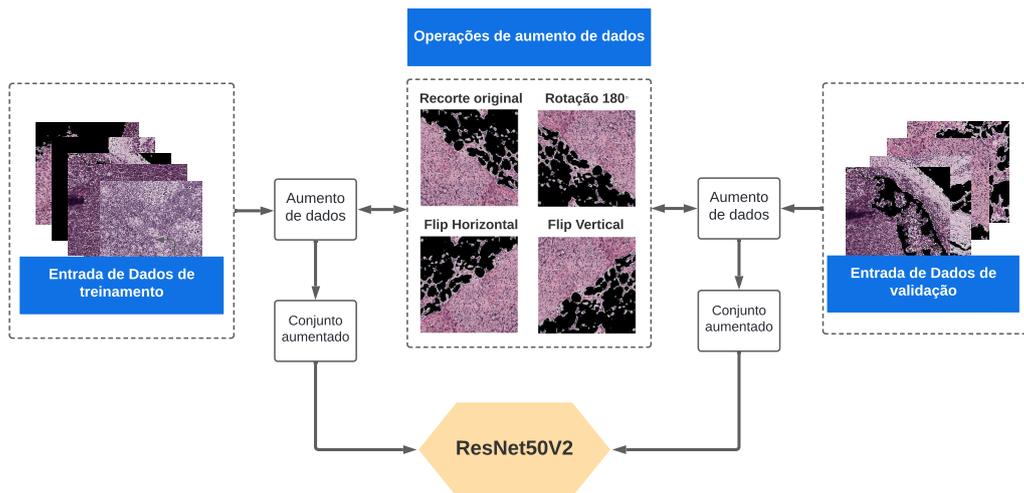


Figura 3. Treinamento para a classificação de patches positivos para metástases

3.3.1. Dados de entrada

Os dados de entrada da rede são os patches pré-processados e rotulados através da metodologia abordada na seção anterior. Esses dados são representados por uma matriz de valores de pixel na forma de $[largura, altura, canais]$. Como os patches possuem dimensão de 224×224 e o canal de cores da imagem pertence ao espaço RGB (*red, green, blue*), a entrada é uma matriz de valores $[224 \times 224 \times 3]$. Para evitar *bias*, foi utilizado lâminas diferentes para gerar os patches para o treinamento, validação e teste.

As quantidades de patches negativos e positivos em cada uma das etapas foram equilibradas para que o conjunto não apresente desbalanceamento, gerando um modelo fraco. No modelo proposto, o conjunto de treinamento e validação foi balanceado através

da remoção aleatória de patches negativos até que a classe negativa não seja maior que a classe positiva em um limite de 10%.

3.3.2. Aumento de dados

O processo de aumento de dados foi aplicado com a finalidade de melhorar a eficácia do treinamento. Neste contexto a ideia é criar um modelo mais robusto e preciso. Para realizar esta tarefa, as seguintes operações foram selecionadas: rotação nos ângulos de 90°, 180° e 270°, *flip* vertical e horizontal.

3.3.3. ResNet50V2

Para a tarefa de classificação de patches foi utilizado a arquitetura de redes neurais ResNet50V2, que obteve bom desempenho nos trabalhos relacionados, iniciando o classificador com os pesos da ImageNet [Deng et al. 2009]. Foi aplicado o otimizador Adam e testado diferentes valores de *batch size* que, ao final, foi definido em $BatchSize = 64$. As funções *Early Stopping*, *ModelCheckpoint* e *ReduceLROnPlateau* foram implementadas durante o treinamento. Estas possuem o intuito de definir o número ideal de épocas e mitigar o problema de *overfitting*, salvar uma cópia do modelo de melhor desempenho e reduzir a taxa de aprendizado quando a métrica *val-accuracy* para de melhorar.

4. Resultados

Com a avaliação dos resultados encontrados, pretende-se validar se a rede neural de classificação estudada está aprendendo as características relevantes para que seja possível detectar as regiões de metástases e micro-metástases em uma lâmina WSI; se as etapas de pré-processamento das lâminas estão de acordo com as necessidades apresentadas pelo problema; e se esta rede de classificação apresenta potencial para integrar um futuro modelo de duas camadas - classificação e segmentação dos patches - o qual se objetiva implementar para contribuir com este domínio de problema.

4.1. Linhas de base e métricas de avaliação

Para validar a eficácia deste modelo de classificação, foram empregadas como *baseline* dos 5 melhores resultados dentre todos os métodos documentados no desafio Camelyon16 citados em [Ehteshami Bejnordi et al. 2017] para a tarefa de classificação de metástase bem como as métricas utilizadas neste desafio. O objetivo do estudo base é similar aos deste trabalho: auxiliar na identificação de micrometástases com o intuito de auxiliar o patologista no diagnóstico das mesmas. Este trabalho também utiliza o conjunto de dados Camelyon16, desta forma fomos capazes de equiparar os resultados. As métricas utilizadas nesta baseline são AUC e F1-Score. Para analisar a incidência de falsos negativos e positivos, estamos utilizando a matriz de confusão dos resultados.

4.2. Classificação

Para a tarefa de classificação de patches, a nível de fragmentos da imagem, a rede proposta atingiu um resultado de $AUC = 0.998$, sendo o $F1 - Score = 1.00$ para a classe negativa e $F1 - Score = 0.99$ para a positiva. Este resultado supera a baseline definida, que possui

Tabela 1. Resultados do método proposto, baselines e trabalhos relacionados e do patologista especialista sem restrição de tempo

Método	AUC
Harvard & MIT	0.994
HMS & MGH III	0.976
HMS & MGH I	0.964
Y. Liu et al. 40X	0.967
Patologista especialista sem restrição de tempo	0.966
Nosso método	0.998

Fonte: [Ehteshami Bejnordi et al. 2017] e [Liu et al. 2017]

$AUC = 0.994$. A comparação do resultado obtido com as baselines e com o patologista especialista sem restrição de tempo é apresentada na Tabela 1. Outra métrica positiva para a tarefa de classificação de patches é que, analisando a matriz de confusão dos testes, podemos averiguar que o modelo não gerou falsos negativos. A Figura 4 apresenta as curvas de precisão e perda obtidas durante o treinamento. Na Figura 5 podemos observar a curva ROC (valor AUC) da rede de classificação e na Figura 6 a matriz de confusão do modelo.

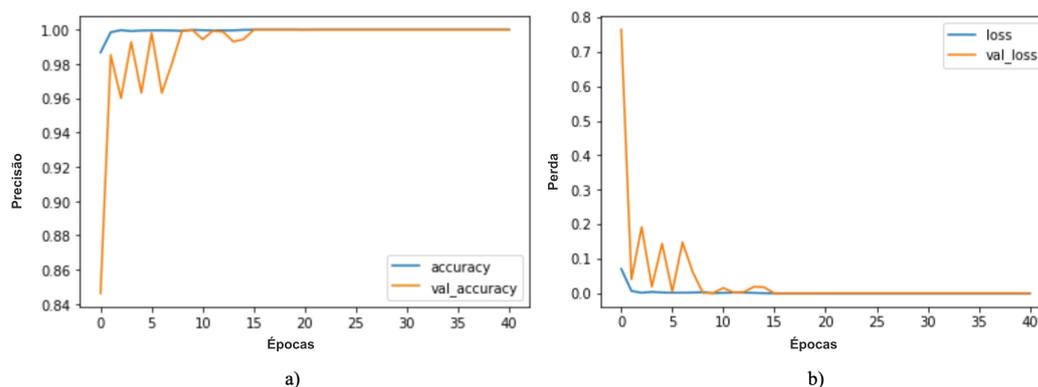


Figura 4. Gráfico de precisão (a) e da perda (b) para o treinamento (azul) e para a validação (laranja)

4.3. Discussão

Para avaliar a métrica AUC durante o treinamento da classificação, os dados de teste foram fragmentados e pré-processados das lâminas de teste utilizando as mesmas técnicas de subdivisão e rotulamento utilizadas durante o treinamento. Conjectura-se que a técnica influencie o resultado e por este motivo esta técnica de pré-processamento e rotulamento precisa ser validada ao fragmentar as lâminas em sua totalidade.

Ainda, uma análise das características dos patches classificados como falsos positivos em busca de fatores confundidores consistentes pode indicar caminhos para ajustar o conjunto de dados e melhorar esta métrica.

Embora é preciso ter em vista esses pontos relevantes, analisando os resultados obtidos durante os experimentos a nível de fragmentos da imagem, em relação a tarefa de classificação de metástases, este trabalho supera a baseline proposta com $AUC = 0.998$.

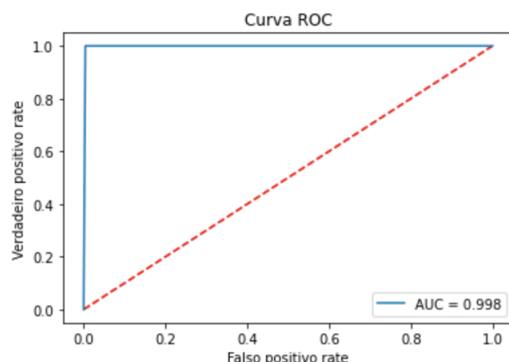


Figura 5. Curva ROC (AUC)

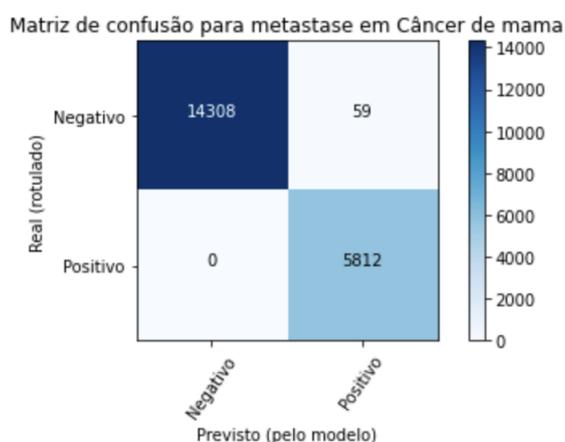


Figura 6. Matriz de confusão para metástase em Câncer de mama

O resultado da classificação é especialmente importante pois o modelo não gerou falsos negativos o que, considerando o campo de estudo deste trabalho e o objetivo do mesmo, e sabendo que esta é uma lacuna listada pelos trabalhos relacionados, este representa um potencial resultado.

Além do potencial ganho pelo modelo não estar gerando falsos negativos, esta rede será a primeira de um modelo composto por duas camadas de redes neurais - classificação-segmentação. Desta forma, esta rede neural apresenta resultados significativos para integrar o modelo, pois conjectura-se que esta composição ajudará a filtrar os falsos negativos, visto que os patches classificados como falso positivo nesta primeira rede serão posteriormente reanalisados a nível de pixel pela segunda rede de segmentação, diminuindo também, a incidência de falsos positivos no resultado final.

5. Conclusão

A área de estudo do auxílio computacional para detecção de micrometástases em câncer de mama tem potencial de gerar resultados expressivos quando aplicados para melhorar a eficiência da rotina dos patologistas de forma complementar, diminuindo a incidência de falsos negativos.

Neste trabalho se investigou a implementação de uma rede de classificação para

futuramente compor um modelo de duas camadas de redes neurais. A primeira para realizar a classificação da metástase e a segunda para realizar a segmentação das células tumorais, realizando a análise do tamanho da região metastática, gerando o resultado final. Espera-se que este método auxilie a diminuir a incidência de falsos negativos, tendo uma rede de classificação com pesos ajustados para penalizar o erro negativo, e, em contrapartida, também diminuindo a incidência de falsos positivos pois, embora a primeira rede tenda a gerar resultado falsos positivos, os patches erroneamente classificados como positivos serão reavaliados pela segunda rede.

O modelo final ainda encontra-se em fase de construção. Todavia, este estudo apresenta resultados significativos e dá direções para os trabalhos futuros. Como pontos de melhoria, posteriormente pretende-se realizar um maior número de experimentos de modo a desempenhar uma melhor análise das métricas para a classificação. Pretende-se, também, implementar a segunda rede do modelo final, realizando a segmentação das áreas metastáticas e produzindo um resultado a nível da lâmina. Ainda pode-se melhorar o conjunto de dados, aumentando e diversificando o mesmo, agregando o conjunto de dados Camelyon17 aos estudos e analisando diferentes formas de pré-processar e rotular as imagens.

Agradecimentos

Os autores agradecem às seguintes agências brasileiras pelo apoio a este trabalho: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, código financeiro 001) e Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, bolsas 309537/2020–7 e 305263/2021–8).

Referências

- Bandi, P., Geessink, O., Manson, Q., Van Dijk, M., Balkenhol, M., Hermsen, M., Bejnordi, B. E., Lee, B., Paeng, K., Zhong, A., et al. (2018). From detection of individual metastases to classification of lymph node status at the patient level: the camelyon17 challenge. *IEEE transactions on medical imaging*, 38(2):550–560.
- Cardoso, F., Kyriakides, S., Ohno, S., Penault-Llorca, F., Poortmans, P., Rubio, I., Zackrisson, S., and Senkus, E. (2019). Early breast cancer: Esmo clinical practice guidelines for diagnosis, treatment and follow-up. *Annals of oncology*, 30(8):1194–1220.
- Chen, R., Jing, Y., and Jackson, H. (2016). Identifying metastases in sentinel lymph nodes with deep convolutional neural networks. *arXiv preprint arXiv:1608.01658*.
- Cserni, G., Maguire, A., Bianchi, S., Ryska, A., and Kovács, A. (2021). Sentinel lymph node assessment in breast cancer—an update on current recommendations. *Virchows Archiv*, pages 1–13.
- Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, pages 248–255. Ieee.
- Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., Eirug, A., et al. (2021). Artificial intelligence (ai): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, 57:101994.

- Ehteshami Bejnordi, B., Veta, M., Johannes van Diest, P., van Ginneken, B., Karssemeijer, N., Litjens, G., van der Laak, J. A. W. M., , and the CAMELYON16 Consortium (2017). Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer. *JAMA*, 318(22):2199–2210.
- Ferlay, J., Colombet, M., Soerjomataram, I., Parkin, D. M., Piñeros, M., Znaor, A., and Bray, F. (2021). Cancer statistics for the year 2020: An overview. *International journal of cancer*, 149(4):778–789.
- Howlander, N., Noone, A., Krapcho, M., et al. (2018). Seer cancer statistics review (csr) 1975–2014. national cancer institute website.
- Lee, B. and Paeng, K. (2018). A robust and effective approach towards accurate metastasis detection and pn-stage classification in breast cancer. In *International conference on medical image computing and computer-assisted intervention*, pages 841–850. Springer.
- Litjens, G., Bandi, P., Ehteshami Bejnordi, B., Geessink, O., Balkenhol, M., Bult, P., Halilovic, A., Hermsen, M., van de Loo, R., Vogels, R., et al. (2018a). 1399 h&e-stained sentinel lymph node sections of breast cancer patients: the camelyon dataset. *GigaScience*, 7(6):giy065.
- Litjens, G., Bandi, P., Ehteshami Bejnordi, B., Geessink, O., Balkenhol, M., Bult, P., Halilovic, A., Hermsen, M., van de Loo, R., Vogels, R., et al. (2018b). 1399 h&e-stained sentinel lymph node sections of breast cancer patients: the camelyon dataset. *GigaScience*, 7(6):giy065.
- Liu, Y., Gadepalli, K., Norouzi, M., Dahl, G. E., Kohlberger, T., Boyko, A., Venugopalan, S., Timofeev, A., Nelson, P. Q., Corrado, G. S., et al. (2017). Detecting cancer metastases on gigapixel pathology images. *arXiv preprint arXiv:1703.02442*.
- Liu, Y., Kohlberger, T., Norouzi, M., Dahl, G. E., Smith, J. L., Mohtashamian, A., Olson, N., Peng, L. H., Hipp, J. D., and Stumpe, M. C. (2019). Artificial intelligence–based breast cancer nodal metastasis detection: Insights into the black box for pathologists. *Archives of pathology & laboratory medicine*, 143(7):859–868.
- Macenko, M., Niethammer, M., Marron, J. S., Borland, D., Woosley, J. T., Guan, X., Schmitt, C., and Thomas, N. E. (2009). A method for normalizing histology slides for quantitative analysis. In *2009 IEEE international symposium on biomedical imaging: from nano to macro*, pages 1107–1110. IEEE.
- Siegel, R. L., Miller, K. D., Fuchs, H. E., Jemal, A., et al. (2021). Cancer statistics, 2021. *Ca Cancer J Clin*, 71(1):7–33.
- Turki, T., Al-Sharif, A., and Taguchi, Y. (2021). End-to-end deep learning for detecting metastatic breast cancer in axillary lymph node from digital pathology images. In *International Conference on Intelligent Data Engineering and Automated Learning*, pages 343–353. Springer.
- Valkonen, M., Kartasalo, K., Liimatainen, K., Nykter, M., Latonen, L., and Ruusuvoori, P. (2017). Metastasis detection from whole slide images using local features and random forests. *Cytometry Part A*, 91(6):555–565.

- Wang, J., Liu, Q., Xie, H., Yang, Z., and Zhou, H. (2021). Boosted efficientnet: Detection of lymph node metastases in breast cancer using convolutional neural networks. *Cancers*, 13(4):661.
- Wollmann, T. and Rohr, K. (2017). Automatic breast cancer grading in lymph nodes using a deep neural network. *arXiv preprint arXiv:1707.07565*.
- Zhang, W., Zhu, C., Liu, J., Wang, Y., and Jin, M. (2019). Cancer metastasis detection through multiple spatial context network. In *Proceedings of the 2019 8th International Conference on Computing and Pattern Recognition*, pages 221–225.