

# Aplicando MultiInstance Learning (MIL) para o Diagnóstico de Câncer de Mama em Imagens Histopatológicas

Mario Pinto Freitas<sup>1</sup>, Marcos Gabriel Mendes Lauande<sup>1</sup>, Geraldo Braz Júnior<sup>1</sup>,  
Marcus Vinicius Oliveira<sup>1</sup>, Gabriel Costa<sup>1</sup>, Matheus Levy<sup>1</sup>,  
Anselmo Cardoso de Paiva<sup>1</sup>, João D. Sousa de Almeida<sup>1</sup>

<sup>1</sup>Núcleo de computação aplicada – Universidade Federal do Maranhão (UFMA)  
65.080-805 – São Luís – MA – Brasil<sup>1</sup>

**Abstract.** *Breast cancer is one of the most common cancers among women because it accounts for 29.07% of cases. Among all the variations of the disease, breast cancer is the most frequent in Brazil. And for this reason it is imperative that techniques are developed that speed up the detection process of these tumors to reduce the rate of terminal cases. Deep learning has become a strong ally for pathologists in the analysis of histopathological images making decisions quickly and reliably. In this work we present an approach based on MIL - Multi Instance Learning that has a different approach from the traditional one due to the fact that it works with several instances of the same image. We used the BreakHis breast cancer dataset to evaluate this method. In the experiments carried out, an accuracy of 90% and up to 98% in sensitivity were achieved in this binary classification problem (Benign or Malignant).*

**Resumo.** *O câncer de mama é um dos cânceres mais comuns entre as mulheres pelo fato de este contabilizar 29,07% dos casos. Entre todas as variações da doença, o câncer de mama é o mais frequente no Brasil. E por esse motivo é imperativo que sejam desenvolvidas técnicas que agilizem o processo de detecção destes tumores para diminuir a taxa de casos terminais. O aprendizado profundo tem se tornando um forte aliado dos patologistas na análise de imagens histopatológicas tomando decisão de maneira rápida e confiável. Neste trabalho apresentamos uma abordagem baseada em MIL - Multi Instance Learning que tem uma abordagem diferente da tradicional devido ao fato de trabalhar com várias instâncias de uma mesma imagem. Utilizamos para avaliar esse método o dataset de câncer de mama BreakHis. Nos experimentos realizados, foram alcançados uma acurácia de 90% e 98% de sensibilidade para classificação binária (Benigno ou Maligno).*

## 1. Introdução

O câncer, caracteriza-se pelo crescimento desordenado de células que invadem os tecidos adjacentes ou órgãos à distância. Esta doença pode variar em mais de 100 tipos diferentes [INCA 2021]. Um deles é o câncer de mama que mundialmente se trata de um dos cânceres mais comuns entre as mulheres [WHO 2020], contabilizando no mundo um total de 2,261,419 casos no ano de 2020 superando o câncer de pulmão. Uma realidade parecida é vivenciada no Brasil, representando 29,07% dos casos de câncer [INCA 2021].

Por esse motivo, é imperativo o desenvolvimento de técnicas que agilizem o processo de detecção destes tumores para diminuir a taxa de casos terminais. Atualmente, são

encontrados métodos comumente empregados no diagnóstico como o autoexame, exame clínico das mamas, mamografia, ultrassonografia e a ressonância magnética. Após a detecção do tumor é necessário verificar se o mesmo é maligno ou benigno. Este processo é realizado através da biópsia [Pereira et al. 2021]. Após este procedimento é colhido o tecido e mandado para a análise a ser realizada por um patologista. Usualmente esse tecido é analisado através de imagens histopatológicas.

Por conta da evolução do aprendizado de máquina, junto com o processamento de imagens aplicadas às imagens médicas, várias pesquisas têm sido desenvolvidas com foco em imagens histopatológicas, [Srinidhi et al. 2021]. Entre elas, trabalhos direcionados ao diagnóstico de câncer, como [Cruz-Roa et al. 2014, da Silva et al. 2019] para câncer de mama, leucemia [Vogado et al. 2019], renal [Santos et al. 2019] ou [Linkon et al. 2021] para próstata. Logo, as redes neurais convolucionais (CNNs) tem sido frequentemente utilizadas para a classificação pois, diferente dos métodos tradicionais de extração de características, as CNNs se mostram superiores em diversos aspectos [Spanhol et al. 2017]

Este trabalho, tem como objetivo avaliar uma metodologia de diagnóstico de câncer de mama, baseada em aprendizado profundo e Multi Instance Learning (MIL) [Hering Jan 2020]. Avaliamos que a utilização de técnicas fracamente supervisionadas, como o MIL, podem contribuir de maneira significativa no treinamento de redes neurais convolucionais ao necessitar de um dataset rotulado menor. O avanço em pesquisas nessa área, pode ajudar a habilitar tecnologias onde a quantidade de informações disponíveis é escassa e aumentar o poder de generalização dos algoritmo de redes neurais convolucionais.

## 2. Metodologia

O método proposto neste trabalho é composto pelas etapas de aquisição da base de imagens, pré-processamento, preparação para aplicação da técnica de multi instance learning (MIL), construção da técnica de MIL com o cálculo de atenção e por fim a avaliação. Essas etapas são ilustradas na Figura 1

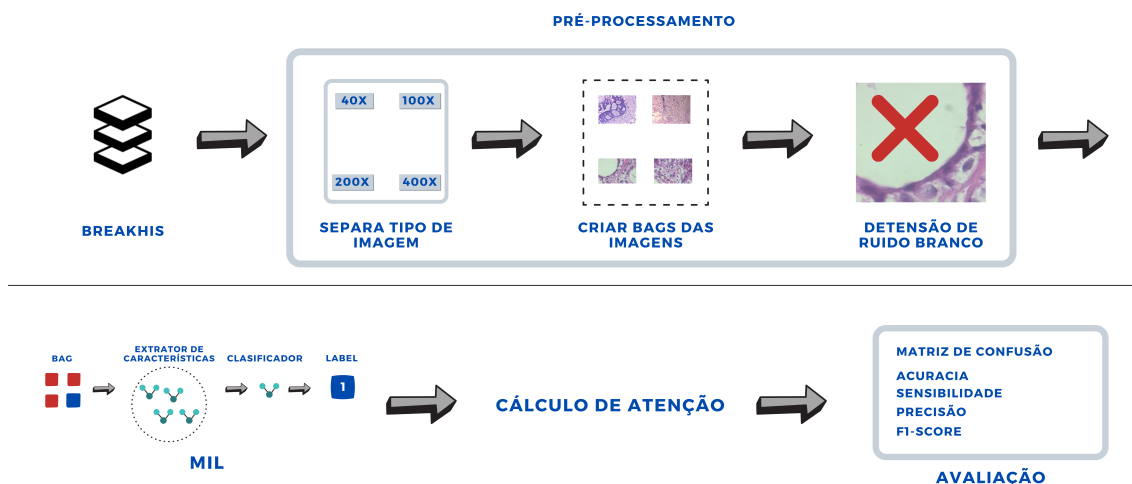
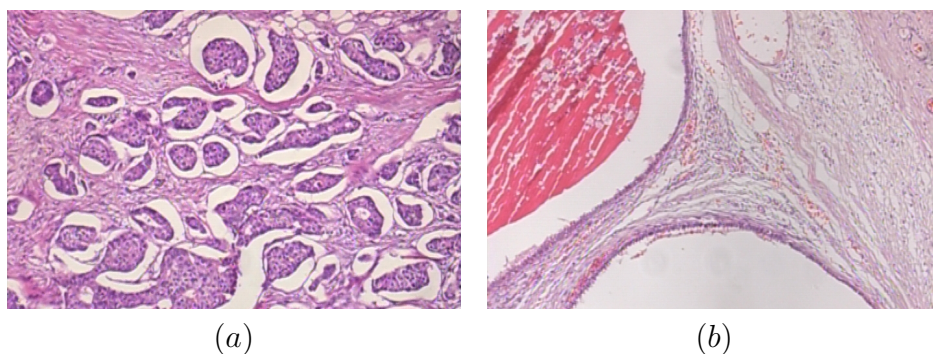


Figura 1. Etapas da metodologia proposta

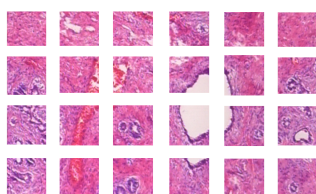
## 2.1. Aquisição e Pré-processamento das imagens

O dataset usado para os experimentos foi o BreakHis [Spanhol et al. 2016], que consiste em 9.109 imagens de tecidos tumorais mamários, coletados de 81 paciente e foram usados diferentes fatores de ampliação (40X, 100X, 200X, 400X). O dataset contém 2.480 amostras benignas e 5.429 malignas. E utilizou-se para o treinamento e validação do modelo o fator de ampliação 40X. Alguns exemplos são apresentados na Figura 2.



**Figura 2. Amostras presentes na base. Em (a) uma amostra maligna. Em (b) uma amostra benigna**

Dos fatores de ampliação, este trabalho utiliza apenas as imagens 40x. Foi utilizado para o treinamento, 75% dos pacientes. O restante foi usado para o teste. Após a separação por paciente, foram criadas as Bags que são compostas por  $N$  instâncias e cada uma faz parte da imagem original. Neste trabalho, dividimos as imagens em 16 instâncias, sendo elas de tamanho iguais e compondo uma imagem como um todo (Figura 3).

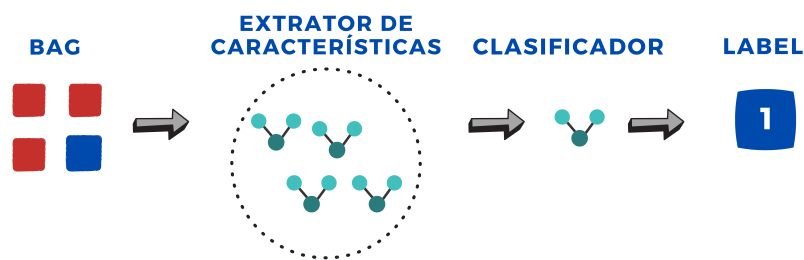


**Figura 3. Exemplo de uma bag**

Após a definição das bags, é feita uma estimativa da área do fundo (placa de petri), transformando cada pixel em 0 ou 1. As zonas mais claras se transformam em 0 referentes à placa de Petri e o tecido torna-se o 1. Em sequência, é verificada em cada instância a porcentagem de área branco. Se for acima de 60%, aquela instância é substituída por outra instância da mesma bag, que contém uma porcentagem de ruído branco menor que a porcentagem pré-determinada, passando-a por um augmentation. O dataset final, portanto, possui  $E = I$  entradas, onde  $E$  seria as entradas da rede e  $I$  o número de instâncias por bag.

## 2.2. MIL (Multi instance learning)

O Multi Instance Learning - MIL [Dietterich et al. 1997] é uma técnica de aprendizado de máquina que usa aprendizagem fracamente supervisionada. MIL trabalha com bags, em



**Figura 4. Diagrama das etapas básicas presentes no MIL**

que dentro de cada bag existirá  $N$  instâncias, para uma bag ser positiva,  $Y = 1$  se existe um  $y_i$  tal que  $y_i = 1$  e para uma bag ser negativa,  $Y = 0$  se somente se para todo  $y_i$  tal que  $y_i = 0$ . Dando assim a label da bag Figura:4.

Tem sido desenvolvidas técnicas diversas como o [Hering Jan 2020], que através do cálculo de atenção da MIL, buscam a região cancerígena em uma imagem WSI. O [Ilse et al. 2018], tem uma abordagem baseada em atenção em que agrupa várias instâncias de uma bag, classificando-as de acordo com o nível de atenção de cada uma. O [Xu et al. 2019] aborda sobre o desenvolvimento de um framework de imagens histopatológicas, para segmentação e classificação em que usa-se o MIL para ambas as funções.

Logo o MIL pode ser usado tanto para segmentação, quanto para a classificação das mesmas, tratando-se de um processo de aprendizagem fracamente supervisionado. Nota-se que a interpretabilidade é um problema recorrente em redes convolucionais [Patil et al. 2019], e trata-se de entender o porquê da rede ter dado aquela resposta para determinada imagem. Com isso, [Patil et al. 2019] produziu um algoritmo baseado em MIL, onde ele identifica as regiões cancerígenas através das instâncias e para solucionar o problema de interpretabilidade ele mostra na bag cada instância, e as diferencia entre si, de maneira totalmente visual, mostrando por onde a rede aprendeu e quais instâncias teriam material cancerígeno.

### 2.2.1. Cálculo de atenção

Na formulação do problema de classificação binário supervisionado, o objetivo do modelo é prever um valor alvo  $y \in (0,1)$  para determinada instância(imagem). No caso do MIL, temos não um valor alvo por instância e sim um valor alvo por bag, que dentro dela existe  $N$  instâncias  $X \in (x_1, x_2, \dots, x_n)$  não ordenadas [Ilse et al. 2018].

A abordagem MIL define a probabilidade da bag  $\theta(X)$ , deve ser invariante de permutação, uma vez que assumimos uma não orientação e não dependências das instâncias dentro da bag. Portanto, o problema MIL pode ser considerado na fórmula do teorema fundamental da função simetria monômios dado pelo seguinte teorema: [Zaheer et al. 2017].



$$S(X) = g\left(\sum_{x \in X} f(x)\right) \quad (1)$$

Um teorema semelhante com *max* em vez de soma é dado por Teorema 2.

$$|S(X) - g(\max_{x \in X} f(x))| < \epsilon \quad (2)$$

Os Teoremas 1 e 2, indicam que, para uma classe de funções flexível o suficiente, podemos modelar qualquer função de pontuação invariante de permutação [Ilse et al. 2018]. Consideramos uma classe de parametrizadas por redes neurais  $f_\psi(\cdot)$  com parâmetros  $\psi$  que transformam o  $k$ -ésima instância em uma instância de baixa dimensão,  $h_k = f_\psi(X_k)$  onde  $h_k \in H$ ,  $H \in [0,1]$  para abordagem baseada em instância e  $H = R^M$  para abordagem baseadas em incorporação [Ilse et al. 2018].

A parametrização de todas as transformações usando redes neurais, se torna muito interessante deixando a abordagem arbitrariamente flexível, podendo ser treinada ponta a ponta. A única restrição é que o pool da MIL deve ser diferenciável [Ilse et al. 2018]. A formulação do problema da MIL requer o pooling da MIL, para ser invariante de permutação.

Conforme mostrado no Teorema 1 e 2, existem dois operadores de pooling da MIL que garantem a função de pontuação (a probabilidade da bag) para ser uma função simétrica [Ilse et al. 2018]. Todos os operadores de pool da MIL mencionados possuem a desvantagem de serem predefinidos e não podem ser treinados. Assim, aplica-se mecanismo de atenção para contornar o problema. Neste se usa de uma média ponderada de instâncias (incorporação de baixa dimensão) onde os pesos são determinados por uma rede neural. Além disso, os pesos devem somar **1** ser invariante para o tamanho de uma bag, a média ponderada cumpre os requisitos do Teorema 1 como mostra a Equação 3 [Ilse et al. 2018].

$$z = \sum_{k=1}^K a_k h_k \quad (3)$$

$$a_k = \frac{\exp(\mathbf{w}^T \tanh(\mathbf{V}\mathbf{h}_k^T))}{\sum_{j=1}^K \exp(\mathbf{w}^T \tanh(\mathbf{V}\mathbf{h}_j^T))} \quad (4)$$

onde  $W \in R_{L \times 1}$  e  $V \in R_{L \times M}$ , além disso utilizamos a tangente hiperbólica,  $\tanh(\cdot)$  não linear, para incluir valores negativos e positivos. A construção da proposta permite descobrir semelhanças entre as instâncias [Ilse et al. 2018].

Notamos que a  $\tanh(\cdot)$  pode ser ineficiente para aprender relações complexas. Isso decorre do fato de que a  $\tanh(x)$  é aproximadamente linear para  $x \in [-1,1]$ , o que pode limitar a expressividade final das relações aprendidas entre instâncias. Portanto, propomos o uso do mecanismo de bloqueio [Dauphin et al. 2016]. junto com a  $\tanh(\cdot)$  não linear [Ilse et al. 2018] produzindo Figura 5.

$$a_k = \frac{\exp(\mathbf{w}^T \tanh(\mathbf{V}\mathbf{h}_k^T) \text{sigm}(\mathbf{U}\mathbf{h}_k^T))}{\sum_{j=1}^K \exp(\mathbf{w}^T \tanh(\mathbf{V}\mathbf{h}_j^T) \text{sigm}(\mathbf{U}\mathbf{h}_j^T))} \quad (5)$$

MIL com base no cálculo de atenção proposto na Equação 5, permite atribuir pesos diferentes a instâncias dentro de uma mesma bag, portanto, a representação final da bag pode ser altamente informativa para o classificador. Em outras palavras, ele deve ser capaz de localizar **instâncias-chave** dentro da bag.

### 2.2.2. Definição do Modelo

Com o processo de atenção e funcionamento do MIL sendo definido, podemos aplicá-lo ao treinamento de uma rede neural convolucional com a finalidade de classificação. Neste trabalho, usamos como rede base um DenseNet201 pré-treinada com a imageNet [Huang et al. 2017]. Dentro do modelo cada entrada da rede tem uma camada de pré-processamento seguida da DenseNet pré-treinada, normalização, flatten, seguidas de duas camadas densas e por último a saída da rede (0 ou 1). A arquitetura da rede é demonstrada através do diagrama na Figura 5. O otimizador usado foi adam com learning rate de 0.01. Os Hiperparâmetros escolhidos foram ,20 épocas durante o treinamento e batch size igual a 8, sendo os mesmos escolhidos através de um otimizador de Hiperparâmetros.

### 2.3. Avaliação

Neste trabalho foram usadas 4 técnicas de avaliação sendo elas acurácia, sensibilidade, precisão e f1-score. A acurácia se baseia na quantidade de acertos positivos que o modelo conseguiu efetivar durante a avaliação do mesmo. A sensibilidade é dada pela porcentagem de dados classificados como positivos comparado com a quantidade real de positivos que existem, dada pela Equação 6.

$$Sensibilidade = \frac{PositivoVerdadeiro}{FalsoNegativo + PositivoVerdadeiro} \quad (6)$$

A Precisão é dada pela quantidade de acertos positivos dividido pela quantidade total de positivos, dado pela Equação 7

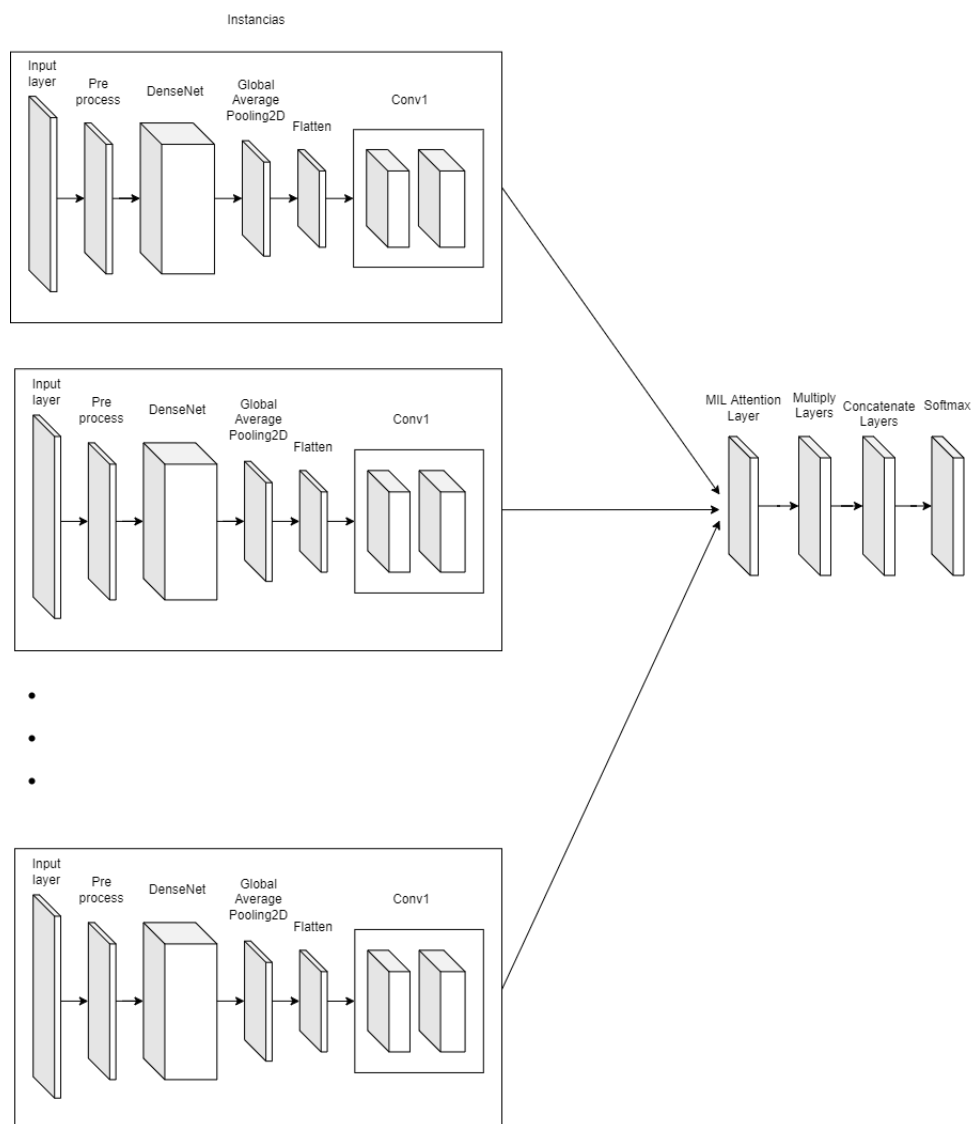
$$Precisao = \frac{PositivoVerdadeiro}{FalsoPositivo + PositivoVerdadeiro} \quad (7)$$

F1-score é uma média harmônica entre precisão e sensibilidade, muito usada para dataset com classes desbalanceados, dada pela Equação 8

$$F1Score = 2 * \frac{Precisao * Sensibilidade}{Precisao + Sensibilidade} \quad (8)$$

## 3. Resultados

Inicialmente, utilizamos um otimizador para estimar os hiperparâmetros dos modelos. Este processo, nos indicou que deveriam ser usadas 16 instâncias por bag. Avaliamos dois tamanhos de instância: de 32x32 e 64x64 pixels. Para o treinamento do otimizador e obtenção das métricas foram escolhidas as imagens com ampliação 40X. Estes resultados são apresentados na Tabela 1.



**Figura 5. Arquitetura da rede com o MIL**

Entre os resultados apresentado na Tabela 1, pode-se notar a discrepância entre os dois tipos de imagem em relação a sensibilidade e precisão, mostrando que com imagens maiores a precisão aumenta ou seja ele consegue identificar melhor a classe positiva, já em imagens menores percebe-se que a sensibilidade é a que se destaca. Contudo é interessante notar que a média harmônica (F1 Score) de ambas se mantém a mesma, apenas a acurácia consegue se sobressair em 1% mostrando que em imagens maiores não só a precisão aumenta como a acurácia.

É interessante notar que as instâncias dentro de uma bag, vão aprendendo entre si e gerando um valor de atenção para cada uma, contudo todo esse cálculo de atenção só é válido por uma bag, a bag subsequente terá outros valores para suas instâncias. Seria relevante pensar na possibilidade de todas as bags levarem em conta o aprendizado da bag anterior, algo parecido acontece em [Huang et al. 2017]. Esse método criaria não só instâncias ligadas entre si, mas também , bag ligadas entre si, podendo trazer melhores resultados. Tendo em vista que anteriormente o cálculo de atenção não levava em contas

**Tabela 1. Resultados**

Px	Acurácia	Sensibilidade	Precisão	F1-score
32x32	0,89	0,98	0,88	0,92
64x64	0,90	0,89	0,96	0,92

todas as instâncias de uma bag [Dietterich et al. 1997], com os avanços na área, surgir métodos que levam em conta todas as instâncias dentro de uma bag [Ilse et al. 2018], trazendo resultados bem mais relevantes. Fazer um calculo de atenção que interliga as bags, levando em conta a bag anterior, poderia melhorar bastante o resultado final do modelo.

Outra observação interessante sobre o modelo é o fato de que embora a estrutura presente em cada entrada não seja tão robusta, se torna extremamente pesada quando reunidas todas as entradas, para consolidação do mesmo. E essas entradas requerem que suas camadas densas não sejam tão profundas, pelo fato de cada uma ser uma imagem N vezes menor que a imagem original, fazendo isso para que haja melhor aproveitamento da extração de características. Isso pôde ser notado durante o treinamento dos hiperparâmetros, quando se usava camadas densas muito profundas como (128, 256, 512), o modelo alcançava uma acurácia de até 80%. Em contra ponto utilizando camadas menos profundas como (16,32), foi possível conseguir uma acurácia de até 90%.

Em comparação com trabalhos relacionados, apresentados na Tabela 2, todos baseados em MIL e breakhis, destacamos que o trabalho [Patil et al. 2019] usam outra fórmula de atenção onde não temos a sigmóide para refinar o cálculo e fazer com que ele aprenda semelhanças mais complexas. Já [Das et al. 2020] usa imagens (224,244) pixel para suas instâncias, nesse trabalho devido limitações de hardware não podemos utilizar imagens muito grandes para o tamanho das instâncias, por isso usamos instâncias de tamanhos (32,32) e (64,64). Vemos na Tabela 2 que o f1-score se manteve o mesmo, contudo entre os dois tamanhos de imagens temos uma grande diferença entre a sensibilidade e a precisão. instâncias maiores fazem com que o modelo consiga determinar melhor a classe positiva, podemos afirmar que imagens maiores podem vir a aumentar a taxa de acerto, imagens 64px tem uma acurácia maior de até 1% em comparação às de 32px.

**Tabela 2. Comparações com trabalhos relacionados**

	Acurácia	Sensibilidade	Precisão	F1-score	AUC
[Alharbi et al. 2021]	0,892	0,889	0,946	0,918	0,935
[Das et al. 2020]	0,93	0,94	-	-	-
[Sudharshan et al. 2019]	0,921 ± 0,59	-	-	-	-
[Rony et al. 2019]	0,9240 ± 0,282	-	-	-	-
[Patil et al. 2019]	0,86	„	„	„	„
Este trabalho	0,90	0,89	0,96	0,92	-

#### 4. Conclusão

Este trabalho descreveu um método baseado em MIL, proposto por [Dietterich et al. 1997], com cálculos de atenção baseados em [Ilse et al. 2018]

para torná-lo mais eficiente. Os resultados obtidos são de 0,89 de sensibilidade e 0,96 de precisão, indicando que embora não sejam o estado da arte, o MIL pode ser aplicado adequadamente para auxiliar no processo de aprendizado em redes neurais convolucionais.

A contribuição deste trabalho consiste na adaptação de um processo de aprendizado dentro do contexto de redes neurais aplicado ao reconhecimento em imagens histopatológicas. Evidente que esta é uma área de pesquisa recente e que necessita de maiores formulações para atingir os melhores resultados. Entretanto, os resultados iniciais indicam uma tendência promissora. Importante destacar que esta pesquisa pretende continuar na direção de avaliar e desenvolver novas formas de atenção e pré-processamento das Bags para o MIL, sem descartar a possibilidade de realizar esse processo através de heurísticas.

## Agradecimentos

Os autores agradecem às instituições brasileiras que ajudaram no desenvolvimento desta pesquisa: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e Fundação de Amparo à Pesquisa e ao Desenvolvimento Científico e Tecnológico do Maranhão (FAPEMA) pelo financiamento Apoio.

## Referências

- Alharbi, A., Wang, Y., and Zhang, Q. (2021). *Trans-Attention Multiple Instance Learning for Cancer Tissue Classification in Digital Histopathology Images*, page 79–84. Association for Computing Machinery, New York, NY, USA.
- Cruz-Roa, A., Basavanthally, 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. In Gurcan, M. N. and Madabhushi, A., editors, *Medical Imaging 2014: Digital Pathology*, volume 9041, pages 1 – 15. International Society for Optics and Photonics, SPIE.
- da Silva, A., Júnior, I., and Amora, M. (2019). Breast cancer detection in histopathological images using convolutional neural networks. In *Anais do XIX Simpósio Brasileiro de Computação Aplicada à Saúde*, pages 1–9, Porto Alegre, RS, Brasil. SBC.
- Das, K., Conjeti, S., Chatterjee, J., and Sheet, D. (2020). Detection of breast cancer from whole slide histopathological images using deep multiple instance cnn. *IEEE Access*, 8:213502–213511.
- Dauphin, Y. N., Fan, A., Auli, M., and Grangier, D. (2016). Language modeling with gated convolutional networks. *CoRR*, abs/1612.08083.
- Dietterich, T. G., Lathrop, R. H., and Lozano-Pérez, T. (1997). Solving the multiple instance problem with axis-parallel rectangles. *Artificial intelligence*, 89(1-2):31–71.
- Hering Jan, K. J. (2020). *Multiple Instance Learning Via Deep Hierarchical Exploration for Histology Image Classification*. International Symposium on Biomedical Imaging (ISBI), 17th edition.

- 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.
- Ilse, M., Tomczak, J., and Welling, M. (2018). Attention-based deep multiple instance learning. In *International conference on machine learning*, pages 2127–2136. PMLR.
- INCA (2021). Types of cancer — national cancer institute - josé alencar gomes da silva - inca. <https://www.inca.gov.br/o-que-e-cancer>.
- Linkon, A. H. M., Labib, M. M., Hasan, T., Hossain, M., and Marium-E-Jannat (2021). Deep learning in prostate cancer diagnosis and gleason grading in histopathology images: An extensive study. *Informatics in Medicine Unlocked*, 24:100582.
- Patil, A., Tamboli, D., Meena, S., Anand, D., and Sethi, A. (2019). Breast cancer histopathology image classification and localization using multiple instance learning. In *2019 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE)*, pages 1–4.
- Pereira, caroline Silva, F., Pereira, L. R. M., and Couto, S. D. (2021). Procedimentos para diagnóstico do câncer de mama: Revisão de literatura. *Repositório Universitário da Anima (RUNA)*.
- Rony, J., Belharbi, S., Dolz, J., Ayed, I. B., McCaffrey, L., and Granger, E. (2019). Deep weakly-supervised learning methods for classification and localization in histology images: a survey. *CoRR*, abs/1909.03354.
- Santos, J., Veras, R., Silva, R., Aldeman, N., Aires, K., and Bianchi, A. (2019). Classificação de imagens de biópsias renais com glomeruloesclerose segmentar e focal ou com lesões mínimas utilizando transfer learning em cnn. In *Anais do XIX Simpósio Brasileiro de Computação Aplicada à Saúde*, pages 82–93, Porto Alegre, RS, Brasil. SBC.
- Spanhol, F. A., Oliveira, L. S., Cavalin, P. R., Petitjean, C., and Heutte, L. (2017). Deep features for breast cancer histopathological image classification. In *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 1868–1873.
- Spanhol, F. A., Oliveira, L. S., Petitjean, C., and Heutte, L. (2016). A dataset for breast cancer histopathological image classification. *IEEE Transactions on Biomedical Engineering*, 63(7):1455–1462.
- Srinidhi, C. L., Ciga, O., and Martel, A. L. (2021). Deep neural network models for computational histopathology: A survey. *Medical Image Analysis*, 67:101813.
- Sudharshan, P., Petitjean, C., Spanhol, F., Oliveira, L. E., Heutte, L., and Honeine, P. (2019). Multiple instance learning for histopathological breast cancer image classification. *Expert Systems with Applications*, 117:103–111.
- Vogado, L., Veras, R., Araujo, F., Silva, R., and Aires, K. (2019). Rede neural convolucional para o diagnóstico de leucemia. In *Anais do XIX Simpósio Brasileiro de Computação Aplicada à Saúde*, pages 46–57, Porto Alegre, RS, Brasil. SBC.
- WHO (2020). Cancer today — world health organization.
- Xu, G., Song, Z., Sun, Z., Ku, C., Yang, Z., Liu, C., Wang, S., Ma, J., and Xu, W. (2019). Camel: A weakly supervised learning framework for histopathology image

segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 10682–10691.

Zaheer, M., Kottur, S., Ravanbakhsh, S., Póczos, B., Salakhutdinov, R. R., and Smola, A. J. (2017). Deep sets. *Advances in neural information processing systems*, 30.