

O uso de redes neurais para a segmentação e extração da densidade mamária a partir de imagens mamográficas

Palavras-Chave: REDES NEURAIIS, MAMOGRAFIA, DENSIDADE MAMÁRIA

Autores(as):

Aline Yoshida Machado, IFGW– Unicamp

Prof. Dr. Mario Antônio Bernal (orientador), IFGW-Unicamp

Prof. Dr. Alessandra Tomal, Arthur K. IFGW-Unicamp

Dr. Rodrigo Trevisan Massera, Katholieke Universiteit Leuven

INTRODUÇÃO:

Atualmente, a mamografia é modalidade de imagem recomendada para a detecção precoce do câncer de mama, sendo extensivamente empregada em programas de rastreamento em todo o mundo [1]. Um dos fatores associados com menores sensibilidade e especificidade da mamografia é a alta densidade da mama, uma vez que mamas densas têm uma proporção maior de tecido fibroglandular, que apresenta composição similar ao tecido neoplásico, tornando a identificação de cânceres mais difícil [2]. Além disso, estudos a respeito da composição de tecidos mamários mostram que a densidade da mama é um fator de risco para o câncer de mama [3]. Por fim, há uma forte correlação entre a densidade da mama e a dose depositada pela radiação ionizante [3].

A grandeza comumente usada na estimativa da densidade mamária é a densidade volumétrica da mama (VBD), que quantifica a razão entre o volume de tecido glandular e o volume da mama inteira [4]. A VBD é usualmente estimada na rotina de exames mamográficos de forma qualitativa sendo determinada por meio do uso da classificação *Breast Imaging Reporting Data System* (BI-RADS) [5]. Tal avaliação se baseia em critérios subjetivos como a análise visual de radiologistas e carece de critérios mais rigorosos, como a quantificação exata da VBD e da localização do tecido glandular na mama. Estimativas mais precisas da VBD podem ser obtidas de forma quantitativa, utilizando diversos métodos semiautomáticos ou automáticos computacionais, tais como os softwares comerciais Cumulus (University of Toronto, Canada), Volpara (Volpara Solutions Ltd., New Zealand), Quantra (Hologic Inc., Bedford, MA, USA) and LIBRA (University of Pennsylvania, USA) [6], [7]. Estes softwares possuem uma alta acurácia, porém têm as desvantagens de custo elevado e são pouco acessíveis, além de possuírem limitações operacionais.

Diversas pesquisas científicas estão direcionadas no desenvolvimento de técnicas que otimizam e qualificam os processos de segmentação e quantificação de imagens de raios X e de mamografia, em específico na área de inteligência artificial e *Machine Learning*. Estes campos de estudo possibilitaram uma maior acurácia das técnicas de diagnóstico e passou a fornecer ferramentas computacionais como auxílio para os profissionais de saúde [8]. O objetivo deste trabalho, foi aplicar e otimizar um algoritmo de *deep learning* utilizando redes neurais para determinar a distribuição de tecido glandular em cada região da mama e estimar sua densidade volumétrica. Após o treinamento da rede, foi obtida a proporção de tecido glandular em cada região da imagem e o respectivo valor de VBD foi estimado para cada imagem. Este estudo é derivado de um estudo piloto desenvolvido no Grupo de Física Radiológica Médica (GFRMd) do Departamento de Física Aplicada do IFGW/UNICAMP [9], que tem como objetivo principal a estimação da VBD a partir de imagens mamográficas gerados via simulação de Monte Carlo de modelos de mama antropomórficos e técnicas de redes neurais para executar a segmentação e quantificação da VBD.

METODOLOGIA:

Inicialmente, foram geradas as máscaras de treinamento específicas para cada rede neural. Em seguida, foram simuladas imagens mamográficas para treinar três redes: de segmentação (rede I), a da altura relativa (rede II) e a de altura relativa glandular (III). Cada rede foi submetida a um rigoroso processo de treinamento, que incluiu etapas de validação, teste e otimização. Finalmente, os resultados obtidos foram utilizados para estimar o VBD. A Figura 1 ilustra o fluxograma da metodologia e os detalhes de cada etapa do processo serão descritos nas subseções seguintes.

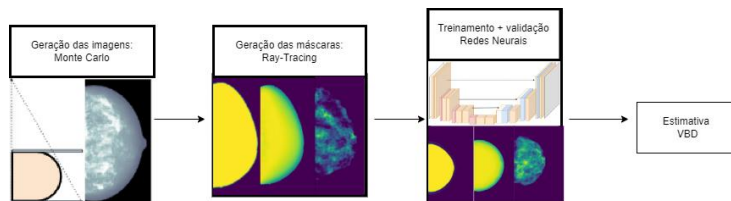


Figura 1: Ilustração das etapas realizadas neste trabalho. Figura adaptada da referência [9]

1- CRIAÇÃO DAS MÁSCARAS DE TREINAMENTO:

As máscaras e suas respectivas imagens mamográficas de treinamento foram extraídas do banco de dados grupo GFRMd, sendo gerado utilizando simulações Monte Carlo (MC) da imagem mamográfica e *ray-tracing* em diferentes etapas. Primeiro, 308 *phantoms* antropomórficos 3D foram criados utilizando o software *BreastPhantom*, projetados para simular mamas com diferentes dimensões, formas e composições de tecidos. Na segunda etapa, utilizou-se as imagens mamográficas desses *phantoms* nas projeções craniocaudal (CC) e médio-oblôquo lateral (MLO) foram simuladas pelo método MC, cobrindo uma gama de espectros de energia. As máscaras foram geradas utilizando estes *phantoms* e algoritmos analíticos de *ray-tracing*. Mais detalhes sobre os métodos utilizados podem ser encontrados na referência [9].

2- DADOS DE TREINAMENTO:

Para gerar as imagens de treinamento, foram realizadas 3485 simulações MC, que serviram como base para os dados de treinamento das redes I, II e III. Para o processamento das imagens de treinamento da rede I, aplicou-se um filtro inicial com o objetivo de ajustar os valores de intensidade dos pixels. Essa etapa foi essencial para compensar as inhomogeneidades causadas pela divergência dos caminhos percorridos pelo feixe de raios X, garantindo uma maior uniformidade nas imagens. Para a rede II, os dados de treinamento são compostos pela mesma imagem usada na rede I segmentada para incluir apenas a região da mama e sua espessura total como entrada. No total, foram utilizadas 2322 imagens para o treinamento dos modelos, 257 para a validação e 906 para os testes.

Para a rede III, foi adotado um conjunto de cinco *perceptrons* multicamadas. A rede é composta por duas camadas totalmente conectadas de 200 neurônios cada e uma saída linear. O procedimento adotado começou a partir da imagem original, onde uma caixa é desenhada ao redor do tecido mamário interno segmentado da máscara da rede I. Em seguida, essa caixa é dividida em *patches* de imagens de 4×4 pixels. Os valores dos pixels são normalizados pelo número de fótons incidentes. No total, 11 parâmetros são incluídos no treinamento: o espectro de raios X, camada semirredutora do espectro e energia média, a espessura da mama, a altura local relativa média, a distância entre o patch e o centro de massa da máscara, a área coberta pelo patch da mama no detector de imagem, a intensidade média e o desvio padrão dos pixels, e duas estatísticas de primeira ordem (curtose e assimetria). Todas as características são normalizadas entre 0 e 1. A rede prevê a altura glandular relativa, definida como o produto da altura glandular pela altura total relativa.

3- REDE I, II, III

Nesta etapa do estudo, três redes neurais foram implementadas, treinadas e otimizadas. A Rede I, uma rede de segmentação, foi projetada para identificar três regiões distintas em imagens de mamografia. A Rede II estima a altura relativa, determinando a espessura da mama para cada pixel em relação à espessura total. A Rede III estima a distribuição da altura glandular relativa. O treinamento das redes foi realizado com um conjunto de dados dividido em treinamento, validação e teste, utilizando arquiteturas como UNET e MLP.

A rede I foi baseada em uma U-Net, que possui uma arquitetura composta por camadas de convolução e camadas de *max-pooling*, permitindo a extração de características detalhadas das imagens e a recuperação de informações espaciais. A rede procura estimar a região da mama, pela e fundo que são regiões com áreas de tamanhos distintos, o que caracteriza uma rede com classes desbalanceadas. Sendo assim, foi aplicada uma função de perda baseada na *Cross Entropy* ponderada, com pesos ajustados para cada classe, penalizando mais a classe menos representada, que é a pele neste contexto. A rede foi

treinada por meio do otimizador Adam, com uma taxa de aprendizado inicial de 0.001, e a técnica de *learning rate scheduler* foi utilizada para ajustar a taxa de aprendizado com base na melhoria da validação.

No treinamento da rede II, foi implementado uma U-Net para estimar a altura relativa da mama em imagens mamográficas. A função de perda escolhida foi o erro médio quadrático (MSE), adequada para tarefas de regressão como esta. O otimizador Adam, com uma taxa de aprendizado inicial de 0.0001, foi utilizado para ajustar os pesos da rede, enquanto o ajuste fino da taxa de aprendizado foi realizado por um *scheduler*, que reduzia a taxa de aprendizado em 50% após quatro épocas sem melhoria. O modelo foi treinado por 10 épocas com batches de tamanho 1.

A rede III foi treinada utilizando uma arquitetura MLP, entretanto o conjunto de treinamento utilizado foi menor com 178 imagens, apenas considerando as projeções CC e espectros menores. Foi utilizado um lr de 0.0001 e 5 epochs. Para garantir a robustez e a generalização do modelo, foi aplicado o método de *cross-validation* com 5 *folds*. Esse método envolveu a divisão dos dados de treinamento em cinco subconjuntos distintos, onde, a cada iteração, um subconjunto foi utilizado para validação enquanto os outros quatro foram usados para treinamento.

Os modelos finais foram avaliados com o conjunto de teste, utilizando a métrica Erro Percentual Médio Absoluto (do inglês *Mean Absolute Percentage Error*, MAPE). O erro MAPE avalia a discrepância relativa entre o valor previsto de uma região de interesse e seu valor real e fornece uma interpretação intuitiva do erro em termos percentuais. Sua fórmula é descrita pela Equação abaixo:

$$MAPE = 100 \cdot \frac{|ref - pred|}{ref}$$

Onde *ref* é o valor de referência e *pred* é o valor estimado.

4- ESTIMATIVA VBD

A última etapa do projeto é estimar o valor do VBD a partir dos outputs das redes I, II e III, implementadas nas etapas anteriores. A VBD corresponde a razão entre o volume glandular total e o volume total da mama. Com a rede I, é possível segmentar a região da mama. Para estes pixels, é conhecido a área correspondente a cada pixel e sua altura total (obtida pela multiplicação da rede II com a espessura da mama), obtendo-se o volume. O mesmo procedimento é feito para o volume glandular com a rede III. Por fim, o VBD é determinado pela razão entre o volume glandular total e o volume total da mama, fornecendo uma medida da densidade glandular.

RESULTADOS E DISCUSSÃO:

Os resultados das criações das máscaras podem ser visualizados na Figura 2. Observa-se que o método de *ray tracing* fornece uma abordagem analítica para obter os rótulos utilizados no treinamento das redes. Foi realizada uma comparação entre as máscaras e os valores nominais de volume e proporção de tecido glandular para cada simulação, com o objetivo de validar o método, permitindo que essas máscaras fossem utilizadas como valores de referência no treinamento. Para o volume da mama, foram obtidos um erro percentual médio de 5,5% e um desvio padrão de 4,4%. Já para a proporção de tecido glandular, o erro percentual médio foi de 5,6% e o desvio padrão de 4,87%.

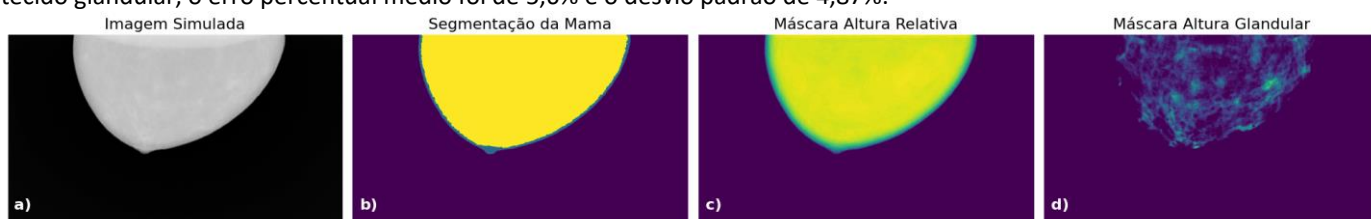


Figura 2: a) Imagem de mamografia CC simulada a partir do código MC; b) Máscara da segmentação, separando a imagem entre fundo, pele e mama; c) Máscara da altura relativa da mama; d) Máscara da altura de tecido glandular;

Para verificar as predições feitas pelas redes I e II, a Figura 3 apresenta uma comparação visual das inferências dessas redes com suas respectivas máscaras para quatro *phantoms* diferentes. No caso da rede I, observa-se que as discrepâncias estão concentradas principalmente na região das bordas da pele e do mamilo. Para a rede II, foi realizada uma comparação pixel a pixel das regiões destacadas em azul e vermelho. Nota-se que o erro foi mais pronunciado nas áreas próximas às bordas, onde a altura da mama apresenta uma diminuição acentuada em relação à altura total.

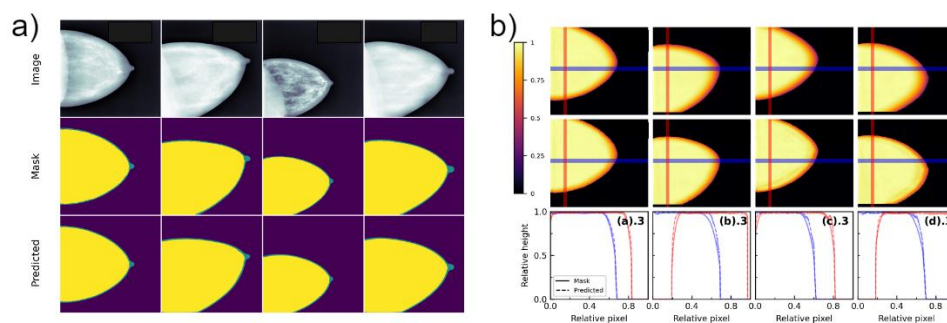


Figura 3: Comparação visual entre imagens, máscaras e previsões da a) rede I b) rede II

Para avaliar o desempenho das redes I e II, foi utilizada a métrica MAPE em um conjunto de teste. A Figura 4 ilustra histogramas das distribuições do erro MAPE para a segmentação da mama, pele e da mama total obtida pela rede I, assim como para a altura relativa obtida pela rede II. No caso da rede I, o maior erro está associado à segmentação da pele e do mamilo, que possuem uma área menor e um formato mais irregular, com uma média de erro e um desvio padrão de 8,6% e 6,5%, respectivamente. Esse resultado pode ser explicado pelo fato de que as máscaras dessa rede foram criadas estabelecendo valores de limiar para a pele projetada e o mamilo, resultando em algumas bordas não suaves e em descontinuidades em certas regiões, o que afetou a predição. Entretanto, a previsão da região interna da mama apresentou resultados satisfatórios, com uma média de erro e desvio padrão de 0,46% e 0,25%, respectivamente. Já a rede II apresentou uma média de erro e desvio padrão de 4,39% e 2,62%, respectivamente, com algumas simulações alcançando erros de até aproximadamente 12%.

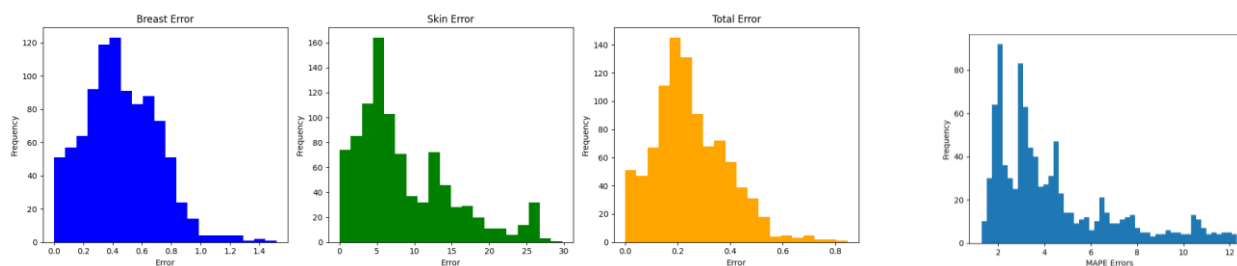


Figura 4: Histograma dos erros MAPE na região da mama, pele, total (mama+pele) e da altura relativa, respectivamente

A rede III apresentou um erro MAPE médio de 34,3% e um desvio padrão de 24,3%. A Figura 5 ilustra um exemplo de inferência com projeção CC da rede III, juntamente com uma comparação pixel a pixel do eixo horizontal da região da mama.

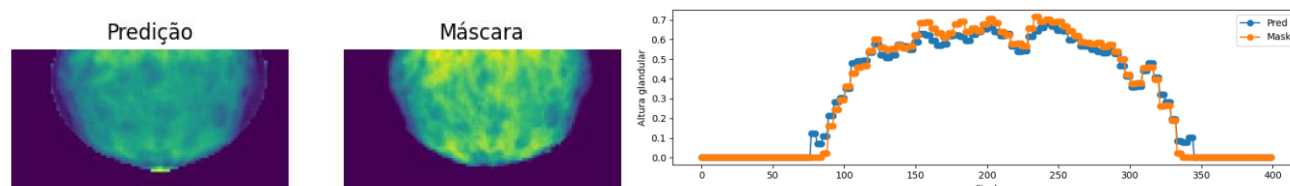


Figura 5: Comparação entre predição e máscara para rede III

Por fim, os resultados das redes I, II e III foram utilizados para calcular a área da mama e da pele, o volume correspondente da mama e o VBD, sendo comparados com os valores calculados a partir das máscaras. Os gráficos de comparação desses valores são ilustrados na Figura 6. Observa-se que tanto a rede I, utilizada no cálculo das áreas, quanto a rede II, no cálculo dos volumes, obtiveram resultados satisfatórios, embora ambas tenham mostrado uma tendência a superestimar as previsões de área e volume.

Vale ressaltar que, para a estimativa do VBD a partir das redes I, II e III, foi utilizado um conjunto de 178 phantoms, com projeções CC, que foram usados no treinamento das redes I e II, e no teste das redes III. Logo, o erro propagado para este valor está exclusivamente associado à rede III, uma vez que este conjunto de imagens pode levar a overfitting nas redes I e II. Assim, mesmo com um erro MAPE alto associado à rede III, os valores de VBD mostraram-se próximos dos valores de

referência, embora ligeiramente subestimados. A média e o desvio padrão do erro MAPE para o cálculo do VBD foram 0,16% e 0,18%, respectivamente.

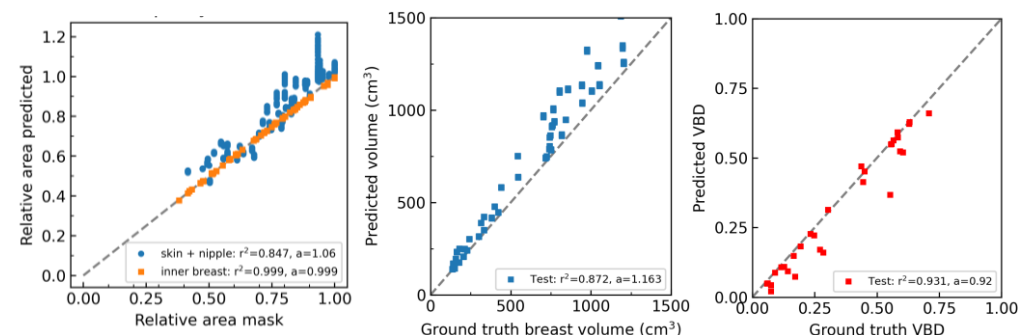


Figura 6: Comparação entre os valores da máscara e previsão para a área estimada da mama e da pele+ mamilo, o volume da região da mama e o VBD, em relação á linha de identidade

CONCLUSÕES E TRABALHOS FUTUROS:

Neste trabalho, foi aplicado e otimizado um *framework* para estimar a densidade volumétrica da mama (VBD) a partir de imagens de mamografia computacionais. O framework foi generalizado através do aumento do número e da variabilidade das imagens de treinamento com a inclusão da projeção MLO e de novos espectros de raios X, bem como pela aplicação de técnicas de otimização e métricas de avaliação para validação do pipeline. Os resultados mostram que o framework é promissor, não apenas na construção de algoritmos de *deep learning* eficazes para a segmentação e estimativa de quantidades relevantes em imagens de mamografia, mas também na introdução de uma metodologia para a estimativa quantitativa da densidade mamária. Para trabalhos futuros, é esperado a generalização da rede III com a inclusão de mais variabilidade nos dados de treinamento, e a inclusão do tecido do músculo peitoral na rede de segmentação como uma nova classe. Além disso, será testada outras grandezas para avaliação da função de perda que classifique de maneira mais precisa a segmentação da pele.

BIBLIOGRAFIA

- [1] European Society of Radiology, ESR, Screening & Beyond: Medical Imaging in the Detection, Diagnosis and Management of Breast Diseases. 2016.
- [2] N. Day and R. Warren, "Mammographic screening and mammographic patterns", Breast Cancer Res, vol. 2, no. 4, pp. 247–51, 2000, Epub 2000 Jun 7.
- [3] D. d. P. Ventorim, "Estudo epidemiológico sobre densidade mamográfica e sua possível influência no risco de câncer de mama", M.S. thesis, UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO, Vitória, 2015
- [4] B.H.M. van der Velden, M.H.A. Janse, M.A.A. Ragusi, et al., "Volumetric breast density estimation on mri using explainable deep learning regression", Sci Rep, vol. 10, p. 18 095, 2020
- [5] Destounis S, Arieno A, Morgan R, Roberts C, Chan A. Qualitative Versus Quantitative Mammographic Breast Density Assessment: Applications for the US and Abroad. Diagnostics (Basel). 2017
- [6] O. Alonzo-Proulx, G. E. Mawdsley, J. T. Patrie, M. J. Yaffe, and J. A. Harvey, "Reliability of automated breast density measurements", Radiology, vol. 275, no. 2, pp. 366–376, 2015.
- [7] E. Ekpo, P. Hogg, R. Highnam, and M. McEntee, "Breast composition: Measurement and clinical use", Radiography, vol. 21, no. 4, pp. 324–333, 2015.
- [8] M. K. Santos, J. R. Ferreira Júnior, D. T. Wada, A. P. M. Tenório, M. H. Nogueira-Barbosa, and P. M. d. A. Marques, "Inteligência artificial, aprendizado de máquina, diagnóstico auxiliado por computador e radiômica: Avanços da imagem rumo à medicina de precisão", Radiologia brasileira, vol. 52, pp. 387–396, 2019
- [9] R. T. Massera and A. Tomal, "Breast glandularity and mean glandular dose assessment using a deep learning framework: Virtual patients study", Physica Medica, vol. 83, pp. 264–277, 2021.