

Modelagem baseada em grafo para detecção de regiões anômalas em termografias da mama

Stefano W. P. Pontes¹, Tiago B. Borchardt¹, Francisco G. N. Clímaco¹,
Darlan B. P. Quintanilha¹

¹Programa de Pós-Graduação em Ciência da Computação (PPGCC)
Universidade Federal do Maranhão (UFMA) – São Luís – MA – Brazil

stefanowalkerpontes@gmail.com

{tiago.bonini, francisco.glaubos, darlan.quintanilha}@ufma.br

Abstract. *The most common type of cancer among women occurs in the breasts, and it is the second leading cause of cancer deaths worldwide. Therefore, much effort is dedicated to improving diagnostic methods, enhancing the quality, accuracy, and cost of diagnostic imaging exams. This work aims to offer a methodology for detecting anomaly regions in breast thermography. This methodology combines superpixel segmentation, graph construction, and analysis of complex network measures. The evaluation of the detection performed in the methodology's experiments provided results that can still be improved, regarding the detection of the exact region of the anomalies, with a percentage of 59.1%. However, it yielded significant results in indicating which breast has an anomaly, with an accuracy rate of 86.4%.*

Resumo. *O tipo de câncer mais comum entre as mulheres ocorre nas mamas, sendo este o segundo tipo de câncer que mais faz vítimas no mundo. Por isso são dedicados muitos esforços para aprimorar os métodos de diagnósticos, melhorando a qualidade, precisão e custo dos exames de diagnóstico por imagem. Esse trabalho tem como objetivo oferecer uma metodologia de detecção de região de anomalias em termografia de mama. Nesta metodologia são combinadas segmentação de superpixel, construção do grafo e análise de medidas de redes complexas. A avaliação da detecção realizada nos experimentos da metodologia forneceu resultados que ainda podem ser aprimorados, em relação à detecção da exata região das anomalias, com percentual de 59,1%. Contudo a mesma propiciou resultados expressivos, no sentido de indicar em qual das mamas há anomalia, com percentual de acerto em 86,4%.*

1. Introdução

O câncer de mama é o segundo tipo de câncer que mais faz vítimas em todo o mundo e é o tipo mais comum entre as mulheres [Bray et al. 2018, World Health Organization 2026]. A prevenção desta doença é ainda difícil por causa da ampla variedade de fatores de risco e características genéticas que a causam. Na maioria dos casos o tratamento desta doença é eficaz quando é detectado em estágios iniciais, aumentando a possibilidade de tratamentos menos agressivos e ampliando as taxas de sucesso.

Dentre a variedade de exames, a mamografia é o principal meio usado para detectar o câncer precocemente. Contudo, a mamografia apresenta algumas limitações para

a detecção de tumores em pacientes jovens por causa da densidade da mama, que interfere na identificação de massas e microcalcificações por raio-X [Wishart et al. 2010, Usuki et al. 1998, Yasmin et al. 2013].

A termografia infravermelha de mama (também chamada de imagem térmica de mama) tem tomado certa atenção nesse campo de estudo, por oferecer a capacidade de ser ferramenta alternativa ou complementar à mamografia em determinados casos [Sattarov et al. 2015, Wang et al. 2014, Luna and Delgado 2010]. Contudo, esses dois tipos de exames têm características distintas. Enquanto a mamografia é um exame estrutural, a termografia é um exame fisiológico e funcional [Bezerra et al. 2013].

Visão computacional e detecção assistida por computador são ferramentas viáveis que desempenham um importante papel na análise de termografias. Uso de técnicas de processamento de imagens atreladas a métodos de aprendizado de máquina têm ajudado a reduzir a taxa de diagnósticos falso-positivos e evitar o viés associado à análise do médico das imagens em infravermelho. Análises estatísticas, por assimetria, por classificação de redes neurais artificiais e reconhecimento automático de tumor são tecnologias exploradas para esta tarefa atualmente.

O objetivo deste artigo é a proposição de uma metodologia de auxílio a detecção de áreas suspeitas de anomalias em termografias de mama, a partir da modelagem das regiões de imagem em forma de grafo e posterior análise de medidas de redes complexas. Entende-se aqui como área anômala no termograma, uma região que a distribuição de temperatura possa ser um indício para o diagnóstico de alguma doença, como fibroadenoma, carcinoma, inflamações ou cistos, não pretende-se realizar a classificação de regiões conforme a doença ou neoplasia apresentada.

2. Trabalhos Relacionados

Nesta seção, são apresentados estudos que fundamentam o uso de redes complexas, grafos e processamento de imagens médicas para a detecção de anomalias, com foco particular na termografia de mama e no uso de grafos.

Recentemente, o uso de modelos computacionais avançados para a análise de termogramas de mama tem apresentado avanços significativos, especialmente na integração de estruturas baseadas em grafos e redes neurais profundas. O trabalho de [Veerlapalli and Dutta 2025] introduziu o framework BCDGAN, que combina Redes Adversárias Generativas (GAN) com modelos híbridos de deep learning. A principal contribuição deste estudo reside na utilização de GANs para sintetizar regiões de interesse (ROIs) anômalas, mitigando o problema da escassez de dados anotados. Utilizando o extrator InceptionResNetV2, os autores alcançaram uma acurácia de 98,56% no dataset DMR-IR, demonstrando a eficácia de técnicas generativas na melhoria da generalização de modelos de detecção.

No âmbito da segmentação de regiões críticas, [Veerlapalli and Das 2024] propuseram o ROISegNet, um framework focado na segmentação automática de ROIs em imagens térmicas. Diferente das abordagens tradicionais de superpixel, este método utiliza convoluções atrosas e um módulo decodificador otimizado para segmentação semântica. O algoritmo resultante, LbBROIS, superou arquiteturas clássicas como VGG19 e ResNet50, atingindo 98,63% de acurácia. Este avanço é particularmente relevante para me-

tecnologias que dependem de uma delimitação precisa da ROI para análises subsequentes de textura ou medidas de rede.

A aplicação da teoria dos grafos para aprimorar a interpretabilidade clínica foi explorada por [Aidossov et al. 2023]. Os autores desenvolveram um sistema inteligente que integra Redes Bayesianas e modelagem baseada em grafos para a detecção precoce de anomalias. O foco na explicabilidade (XAI) permite que o modelo mapeie a funcionalidade do tecido a partir de sua representação térmica, oferecendo uma ponte entre as medidas estruturais de redes complexas e o diagnóstico médico. Esta abordagem corrobora a premissa de que a topologia das variações térmicas pode ser modelada como uma rede de interdependências funcionais.

Por fim, o estado da arte em modelagem geométrica foi apresentado por [Parshionikar et al. 2025], que introduziu o uso de Redes de Atenção em Grafos Hiperbólicos (Hyperbolic Graph Attention) integradas a CNNs bidirecionais. Ao utilizar a geometria hiperbólica, o modelo consegue capturar estruturas hierárquicas e relações espaciais complexas inerentes aos padrões de calor na mama, as quais muitas vezes não são totalmente representadas em espaços euclidianos. Este trabalho destaca a evolução das técnicas baseadas em grafos para arquiteturas de atenção, proporcionando uma detecção mais robusta em ambientes de computação em nuvem.

Apesar dos avanços recentes, a literatura ainda apresenta lacunas significativas, especialmente no que diz respeito à dependência de grandes volumes de dados para o treinamento de modelos de *deep learning* e à natureza de "caixa-preta" de muitas dessas arquiteturas, que limitam a interpretabilidade clínica. Nesse cenário, a metodologia proposta neste artigo se destaca ao utilizar a análise de redes complexas sobre grafos construídos a partir de *superpixels*, oferecendo uma alternativa de baixo custo computacional que não prescinde do rigor analítico.

3. Metodologia

As termografias utilizadas neste trabalho pertencem à base DMR-IR, que possui exames de termografia dinâmica, sendo selecionados para este estudo 35 exames com presença confirmada de alguma anomalia e 35 exames saudáveis [Silva et al. 2014]. Essa base é a que contém a maior quantidade de termografias, possui aquisição de boa resolução, apresenta um protocolo mais rígido e padronizado, e é amplamente utilizada nos trabalhos que apresentaram os melhores resultados na literatura. Por tais motivos, ela foi eleita para ser utilizada nesta pesquisa. Dos 35 exames selecionados para compor a classe com presença de anomalia, posteriormente foi identificado que apenas 22 exames possuíam a descrição detalhada da localização do tumor, informação importante para etapa final dessa metodologia.

A metodologia proposta consiste em 6 etapas consecutivas, apresentadas na Figura 1. Cada etapa é descrita, respectivamente, nas subseções seguintes, juntamente com as técnicas utilizadas.

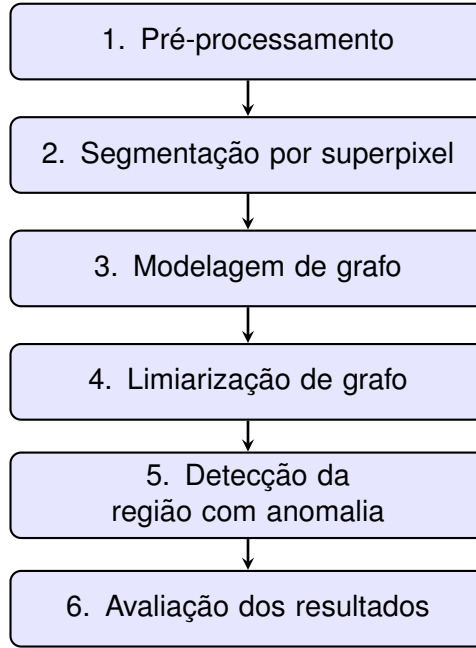


Figura 1. Fluxograma sequencial da metodologia proposta.

3.1. Pré-processamento

3.1.1. Normalização dos termogramas

Para os resultados desta metodologia não sofrerem com perdas de dados causados pela conversão de temperatura para valores em níveis de cinza, a posterior análise de medidas de grafo é feita com o uso das temperaturas. Por outro lado, a fim de realizar a etapa de modelagem de grafo, é necessário usar imagens em escala de cinza, resultantes da conversão das matrizes. As imagens foram geradas pela adaptação da conversão de escalas termométricas, conforme apresentado na Equação 1:

$$p_{ij} = 255 \cdot \frac{t_{ij} - t_{min}}{t_{max} - t_{min}} \quad (1)$$

onde p_{ij} representa os valores dos pixels da imagem a serem obtidos na conversão dos valores de temperatura t_{ij} da matriz. Já t_{max} e t_{min} são, respectivamente, a maior e menor temperaturas encontradas entre todos os exames e 255 é o maior valor de cinza possível. O valor de t_{max} encontrado foi de 35,92 °C. Já o valor de t_{min} foi de 22,56 °C, mas foi ajustado para 22,40 °C, para evitar que, na conversão, a menor temperatura resulte em um pixel com valor 0, se confundindo com o fundo da imagem.

3.1.2. Extração da região de interesse

A *Region of Interest* (ROI) de cada termografia foi extraída por meio da subtração entre o termograma em escala de cinza e a máscara da ROI, que é fornecida pela base. Antes de ser extraída a ROI, as imagens em escala de cinza passam por um melhoramento de contraste com o método *Contrast-Limited Adaptive Histogram Equalization* (CLAHE) e,

logo em seguida, por um filtro da mediana com *kernel* igual a 5 para suavizar os ruídos provocados pela aplicação do CLAHE [Pizer et al. 1990, Gonzalez and Woods 2012].

O resultado do pré-processamento pode ser visto na Figura 1.

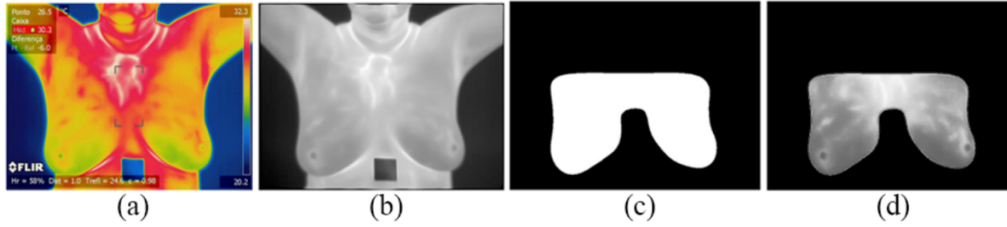


Figura 2. Resultado do pré-processamento: (a) o termograma original obtido pela câmera térmica; (b) a representação da imagem em escala de cinza obtida após a normalização; (c) a máscara da ROI; (d) a aplicação da máscara da ROI sobre a imagem em escala de cinza.

3.2. Segmentação de regiões de superpixels

Nesta etapa, o termograma em escala de cinza passa pelo processo de tesselação, sendo dividido em regiões de *superpixels*. Optou-se por realizar uma tesselação irregular, por *superpixels*, uma vez que prioriza o relacionamento geográfico das regiões dos termogramas e valoriza as diversas formas das sub-regiões das imagens. Então, foi aplicado o algoritmo SLIC, proposto por [Achanta et al. 2011].

O SLIC executa um agrupamento local de *pixels* no espaço 5-D definido pelos valores L , a , b do espaço de cores CIELAB e coordenadas x , y dos *pixels*. O SLIC recebe como parâmetro o número de sementes K e o grau de compactação m . O número de sementes corresponde à mesma quantidade de *superpixels* gerados ao final. Considerando que uma imagem tem N *pixels*, então cada *superpixel* terá aproximadamente N/K *pixels*. O tamanho do diâmetro aproximado de cada um desses *superpixels* seria aproximadamente $S = \sqrt{N/K}$. Em relação à dimensão, os *clusters* são definidos por $C_k = [l_k, a_k, b_k, x_k, y_k]$, com $k \in [1, K]$, sendo l_k , a_k , b_k , x_k e y_k as representações dos componentes respectivos l , a , b , x e y , contidos no espaço do *cluster* C_k . A extensão espacial de qualquer *cluster* é aproximadamente S^2 , inicialmente.

As medidas de distância utilizadas no algoritmo são definidas conforme as Equações 2, 3 e 4:

$$d_{lab} = \sqrt{(l_k - l_i)^2 + (a_k - a_i)^2 + (b_k - b_i)^2} \quad (2)$$

$$d_{xy} = \sqrt{(x_k - x_i)^2 + (y_k - y_i)^2} \quad (3)$$

$$D_S = d_{lab} + \frac{m}{S} d_{xy} \quad (4)$$

onde D_S é a soma da distância *lab* (d_{lab}) e a distância do plano *xy* (d_{xy}) multiplicada por m/S . A variável m permite controlar a compactação do *superpixel*. Quanto maior o valor de m , a proximidade espacial é valorizada, sendo mais compacto o *cluster*. O valor de m pode assumir um valor no intervalo $[1, 20]$, por padrão, foi utilizado o valor 10.

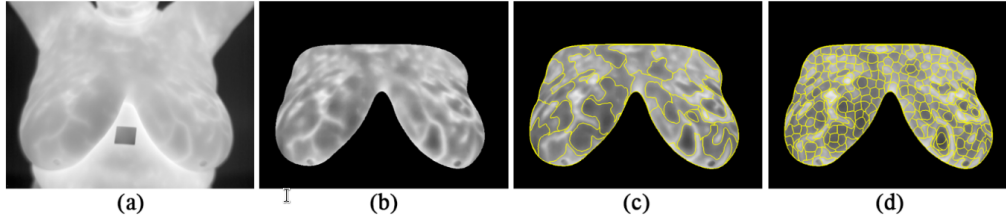


Figura 3. Resultado da segmentação em superpixels: (a) imagem normalizada; (b) a ROI segmentada e com ajuste de contraste e filtro da média; (c) aplicação de SLIC com $k = 30$ e (d) com $k = 250$.

3.3. Modelagem de grafo

Nesta etapa, para cada exame é construído um grafo ponderado e não-direcionado, cujos nós representam os *superpixels* e as arestas representam as relações de adjacência entre os *superpixels*. Os pesos das arestas são calculados por função de similaridade de temperatura.

Na modelagem desenvolvida, nomeada modelagem direta, cada *superpixel* é representado por um nó. A relação de vizinhança se dá quando um *superpixel* é fronteiro a outro, representada por uma aresta não-direcionada. O peso w_{uv} da aresta entre um nó u e um nó v é definido como o valor absoluto da diferença entre as maiores temperaturas de cada *superpixel*, conforme definido pela Equação 5:

$$w_{uv} = \gamma - |t_{max}(u) - t_{max}(v)| \quad (5)$$

onde γ é uma constante que assume o valor igual a 10, que corresponde a um valor seguro, superior à diferença entre a maior e a menor temperaturas máximas de todos os *superpixels*. Quando u e v possuem temperaturas máximas aproximadas, o peso será alto. E quando há uma diferença muito grande entre as temperaturas máximas de u e v , o peso w_{uv} terá valor baixo. Ou seja, essa fórmula provê que sejam valorizadas as conexões entre regiões que mais se assemelham. Após a geração do grafo, os pesos de todas as arestas são normalizados para valores entre 0 e 1.

Inicialmente, o grafo obtido contém um conjunto de arestas que conectam todos os nós entre si, demonstrando que a rede tem certa regularidade, pois a maioria dos nós possui quase a mesma quantidade de ligações. Se o grafo continuar dessa maneira, apresentará poucas propriedades relevantes que permitam observar padrões distintivos entre as regiões. Por isso, é preciso transformar o grafo para que este ganhe mais aspectos de redes de pequeno mundo.

A fim de obter esse resultado, é aplicada a técnica de limiarização de grafo para criar variações de um mesmo grafo, usando eliminação de arestas de determinada faixa de peso. A limiarização, aqui definida, elimina as arestas menores que o limiar L . Em termos de matriz de adjacência A , a equação que condiciona a limiarização é definida da seguinte maneira:

$$a_{uv} = \begin{cases} 0, & \text{se } w_{uv} \leq L \\ a_{uv}, & \text{caso contrário} \end{cases} \quad (6)$$

As arestas com peso baixo (que representam ligações entre regiões muito distintas) são eliminadas, restando somente as ligações entre as regiões mais homogêneas, que são as arestas com peso alto. Então, essa modelagem, atrelada à limiarização, é ideal para fazer comparações entre componentes conectados e/ou comunidades, usando como critério as medidas de conectividade local, estrutura de rede e correlação.

Por meio de testes preliminares, foram escolhidos 4 valores distintos para o limiar L . Como os pesos das arestas são normalizados, L assumirá valores 0,2, 0,4, 0,6 e 0,8. Os resultados desta etapa podem ser vistos na Figura 3.

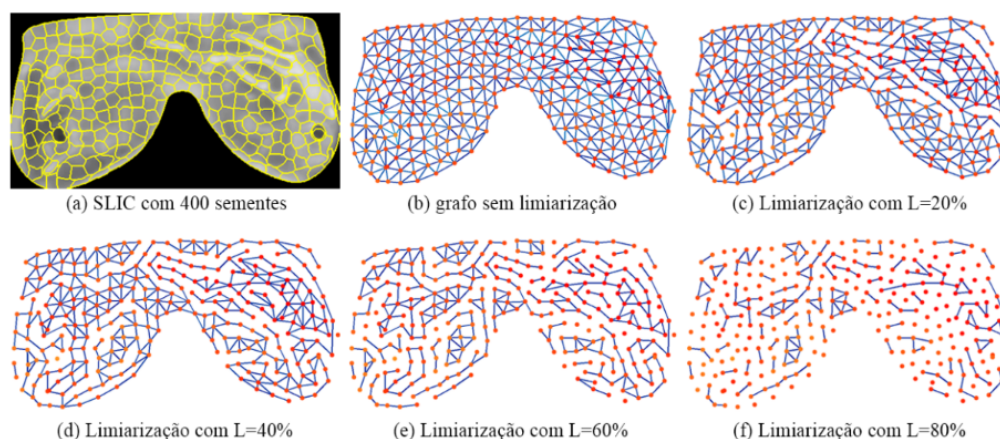


Figura 4. Diferentes graus de limiarização com a modelagem direta.

3.4. Detecção da região da mama com anomalia

Primeiramente, os grafos são divididos em comunidades e/ou componentes conectados, formando subgrafos. Cada um dos subgrafos é submetido à análise de 14 medidas ([Jorge-Hernandez et al. 2014]):

- **grau máximo:** representa a maior conectividade, ou seja, o maior número de arestas ligadas a um único nó dentro do grafo;
- **coeficiente de clusterização:** considerada uma medida da conectividade local, indica o nível de agrupamento do subgrafo formado por um nó e a sua vizinhança;
- **energia:** calculada a partir da soma da probabilidade de ligação para cada grau, elevada ao quadrado;
- **entropia:** medida frequentemente associada ao grau de desordem do grafo, calculada por meio das probabilidades de distribuição dos graus;
- **grau médio da vizinhança de cada nó:** consiste na média aritmética dos graus (conectividade) de todos os nós do subgrafo;
- **coeficiente de Pearson (assortatividade):** mede o nível de correlação das conexões, verificando o quanto vértices com um alto grau tendem a ser conectados a outros vértices que também possuem alto grau;
- **número de vértices (ordem):** também chamado de cardinalidade, representa a quantidade total de nós presentes no grafo;
- **densidade:** definida por meio da razão (divisão) entre a quantidade total de arestas e a quantidade de vértices do grafo;

- **centralidades (de grau, de intermediação, de proximidade e de autovetor):** medidas que determinam a relevância relativa de um nó. A centralidade de grau é o número de ligações incidentes; a de intermediação mede o fluxo nos caminhos mais curtos; a de proximidade mede o inverso do afastamento (rapidez de difusão de informação); e a de autovetor mede a importância do nó em função da influência de seus vizinhos;
- **tamanho do clique máximo:** refere-se ao tamanho do maior subconjunto possível dentro do grafo em que todos os nós estão totalmente conectados entre si;
- **número de cliques maximais:** a quantidade de subconjuntos de nós totalmente conectados que não podem ser estendidos pela adição de novos nós adjacentes.

Para um determinado grafo, uma medida é aplicada a cada um de seus subgrafos, a fim de saber qual subgrafo se destaca para tal medida, e posteriormente é verificado se esse subgrafo coincide com a região da anomalia.

3.5. Avaliação da detecção de região anômala

A base DMR-IR utiliza o esquema de divisão de região das mamas em quadrantes para indicar a região com anomalia, não sendo possível saber as coordenadas XY da localização da anomalia e nem a sua extensão ou forma. Assim, não é possível avaliar com precisão, em termos numéricos, os índices de performance da metodologia usando métricas de segmentação conhecidas. A fim de resolver o problema de avaliação da detecção, foi preciso definir um método de avaliação que pudesse ser aplicado a todos os exames neste trabalho.

Os acertos de detecção foram avaliados conforme o grau de proximidade entre o quadrante sinalizado pela metodologia e o quadrante correto, denotados pelos conceitos A, B, C e D, detalhados na Tabela 1.

Tabela 1. Níveis de aproximação de acerto para avaliação da detecção.

Nível	Significado do nível de aproximação de acerto
A	A região sinalizada pela detecção é a mesma região descrita no diagnóstico.
B	A região sinalizada pela detecção é adjacente ao quadrante correto, em relação às posições horizontais e verticais. Indica que houve um acerto de aproximação, com pouca exatidão.
C	A região sinalizada pela detecção é um quadrante oposto ao quadrante correto, em relação à diagonal, indicando que houve acerto apenas da mama com anomalia, mas sem exatidão da localização quanto aos quadrantes.
D	Sinaliza que houve erro de detecção da mama com anomalia.

4. Resultados e Discussões

Neste trabalho, a etapa de detecção da região com anomalia foi realizada por meio da comparação de comunidades, utilizando a *modelagem direta*. Cada grafo foi dividido em comunidades, às quais foram aplicadas as 14 medidas de redes complexas, sendo a comunidade com o maior índice escolhida como a região detectada para aquela característica.

Para a configuração deste experimento, a quantidade de sementes escolhida para o algoritmo SLIC foi $k = 600$, pois apresentou altos níveis de modularidade da rede

aliada a um menor custo computacional. Os percentuais de limiarização (L) definidos foram de 40% e 60%, originando duas variações do experimento (I-a e I-b), uma vez que esses limiares produziram uma boa quantidade de componentes sem gerar excesso de nós desconectados.

No **Experimento I-a** (parâmetros $k = 600$ e $L = 40\%$), as medidas que alcançaram o maior índice de acerto foram o *grau máximo* e o *grau médio de vizinhança*. Ambas obtiveram um percentual de acerto de 36,4% para o Nível A de avaliação. Ao analisar os níveis agrupados de proximidade, $A \cup B$ e $A \cup B \cup C$, os índices de acerto subiram para 63,6% e 68,2%, respectivamente.

No **Experimento I-b** (parâmetros $k = 600$ e $L = 60\%$), as medidas que mais se destacaram para o Nível A de proximidade foram a *centralidade de proximidade* e o *tamanho do clique máximo*, atingindo também 36,4% de acerto. Para os níveis de proximidade agrupados $A \cup B$ e $A \cup B \cup C$, a medida de *centralidade de intermediação* apresentou os melhores resultados, com 63,6% e 68,2%, respectivamente.

Observa-se pela Tabela 2 que as medidas que alcançaram, neste experimento, o maior índice de acerto foram o *grau máximo* e o *grau médio de vizinhança*. Ambas, empatadas, alcançaram o percentual de acerto de 59,1% para o Nível A. Os níveis agrupados $A \cup B$ e $A \cup B \cup C$ obtiveram percentuais de nível de proximidade de 63,6% e 86,4%, respectivamente, demonstrando que há uma razoável tendência para que as comunidades identificadas apontem para a proximidade da região de anomalia.

Tabela 2. Resultados de avaliação do Experimento I-a ($k = 600$; $L = 40\%$).

Medidas	Avaliações em Níveis				Níveis Agrupados (%)		
	A	B	C	D	A	$A \cup B$	$A \cup B \cup C$
Grau máximo	8	6	1	7	36,4	63,6	68,2
Coef. de agrupamento	7	6	1	8	31,8	59,1	63,6
Energia	7	4	1	10	31,8	50,0	54,5
Entropia	7	4	1	10	31,8	50,0	54,5
Grau médio de vizinhança	8	6	1	7	59,1	63,6	86,4
Coeficiente de Pearson	2	2	2	16	9,1	18,2	27,3
Número de vértices	5	6	1	10	22,7	50,0	54,5
Densidade	1	2	2	17	4,5	13,6	22,7
Centralidade de grau	5	4	0	13	22,7	40,9	40,9
Central. de intermediação	6	5	0	11	27,3	50,0	50,0
Central. de proximidade	6	4	0	12	27,3	45,5	45,5
Central. de autovetor	6	1	2	13	27,3	31,8	40,9
Tamanho do clique máximo	6	3	1	12	27,3	40,1	45,5
Número de cliques maximais	6	5	1	10	27,3	50,0	54,5

De forma geral, observou-se que ambas as variações resultaram nos mesmos percentuais máximos de acerto para os níveis de proximidade, embora atingidos por medidas diferentes. Contudo, constatou-se que a modelagem direta contribuiu negativamente em diversos cenários. Isso ocorreu porque, em alguns exames, as regiões de anomalia correspondiam a partes do grafo com poucas conexões ou tiveram suas arestas eliminadas durante a limiarização, resultando em falsas detecções (avaliação de Nível D).

5. Conclusão

Este trabalho apresentou uma metodologia de detecção de regiões suspeitas de anomalias em termografias de mama, a partir da modelagem das regiões em forma de grafo e posterior análise de medidas de redes complexas.

Como parte dos objetivos, foi desenvolvida uma modelagem de grafo a partir de imagens utilizando: técnicas de pré-processamento de imagens; técnicas de segmentação de regiões com agrupamento de *pixels*; e construção de representação de imagem por meio de grafo. Através de investigação de trabalhos relacionados na literatura, pôde-se aplicar técnicas de análise de grafo pertinentes ao objetivo do estudo. E, por meio de experimentos, foi possível explorar a influência de diversas medidas na realização de uma análise que trouxesse bons resultados de detecção.

Foi desenvolvido um método de avaliação de detecção baseado nos níveis de proximidade em termos de quadrante. Com esse método, foi possível confrontar a região de detecção com a região de diagnóstico fornecida pela base de termografias, validando e medindo a performance desta metodologia.

Por se tratar de uma abordagem que utiliza modelagem de grafo, a representação de regiões *superpixels* destacou as formas e valorizou as correlações geográficas das regiões da termografia, além de reduzir a complexidade computacional em comparação ao cenário onde cada *pixel* fosse representado por um nó de grafo.

Atualmente, o estado da arte na análise de imagens médicas é dominado por Redes Neurais Convolucionais (CNNs). Para mitigar o problema da "caixa-preta" desses modelos, técnicas de visualização como o Grad-CAM (*Gradient-weighted Class Activation Mapping*) são amplamente utilizadas para gerar mapas de calor que destacam as regiões mais influentes para a decisão da rede.

A abordagem baseada em grafos deste trabalho propõe um paradigma alternativo e vantajoso em termos de interpretabilidade intrínseca:

- **Explicabilidade Baseada em Princípios (A priori):** Enquanto o Grad-CAM reconstrói caminhos de gradiente de forma post-hoc (dependendo de heurísticas de retropropagação que podem sofrer com instabilidade ou falta de fidelidade matemática local), as medidas de redes complexas (como centralidade de intermediação ou proximidade) oferecem uma justificativa geométrica e topológica exata do porquê uma comunidade de *superpixels* foi selecionada.
- **Independência de Dados e Recursos:** Mapas Grad-CAM exigem o treinamento prévio de modelos massivos em bases de dados extensas sob o risco de viés de sobreajuste (*overfitting*). O método por grafos opera de forma não supervisionada diretamente sobre a topologia do exame do paciente, demandando um custo computacional significativamente menor.

Muito ainda pode ser explorado com relação à linha de trabalhos que utilizam modelagem em grafo para processamento de imagem combinado com detecção ou segmentação de regiões. Quanto à detecção em termografia de mama, há a necessidade de elaboração de métodos que aumentem a eficiência da detecção.

Agradecimentos

O presente trabalho foi realizado com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e da Fundação de Amparo à Pesquisa e ao Desenvolvimento Científico e Tecnológico do Maranhão (FAPEMA).

Em conformidade com o Código de Conduta para Autores em Publicações da Sociedade Brasileira de Computação (SBC), os autores explicitam que a ferramenta de Inteligência Artificial (IA) generativa Gemini, foi utilizada como suporte na revisão textual e formatação do manuscrito na linguagem *LaTeX*. Os autores reiteram que assumem total responsabilidade pela verificação de fatos, edição final e garantia da integridade ética e científica de todo o conteúdo assistido.

Referências

- Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., and Süsstrunk, S. (2011). Finding objects of interest in images using saliency and superpixels. Technical report, EPFL, Lausanne.
- Aidossov, N., Zarikas, V., Zhao, Y., Mashekova, A., et al. (2023). An integrated intelligent system for breast cancer detection at early stages using ir images and machine learning methods with explainability. *SN Computer Science*, 4(153).
- Bezerra, L., Oliveira, M., Araújo, M., Viana, M. J., Santos, L. C., Santos, F., Rolim, T., Lyra, P., Lima, R. C., Borchardt, T. B., Resmini, R., and Conci, A. (2013). Infrared imaging for breast cancer detection with proper selection of properties: From acquisition protocol to numerical simulation. In Ng, E., Acharya, U., Rangayyan, R., and Suri, J., editors, *Multimodality Breast Imaging: Diagnosis and Treatment*, pages 285–322. Washington, USA.
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 68(6):394–424.
- Gonzalez, R. C. and Woods, R. E. (2012). *Digital Image Processing*. Pearson, London, England.
- Jorge-Hernandez, F., Chimeno, Y. G., Garcia-Zapirain, B., Zubizarreta, A. C., Beldarrain, M. A. G., and Fernandez-Ruanova, B. (2014). Graph theory for feature extraction and classification: A migraine pathology case study. *Bio-Medical Materials and Engineering*, 24(6):2979–2986.
- Luna, J. and Delgado, F. (2010). Feasibility of new-generation infrared imaging screening for breast cancer in rural communities. *US Obstetrics & Gynecology*, 1:52–55.
- Parshionikar, S. et al. (2025). Breast cancer detection using optimized hyperbolic graph attention with bidirectional convolutional neural network. *Discover Oncology*, 16.
- Pizer, S. M., Johnston, R. E., Ericksen, J. P., Yankaskas, B. C., and Muller, K. E. (1990). Contrast-limited adaptive histogram equalization: speed and effectiveness. In *Proceedings of the First Conference on Visualization in Biomedical Computing*, pages 337–345, Nova Jersey, EUA. IEEE.

- Sattarov, A., McIntyre, P., and Motowidlo, L. (2015). High-field open mri for breast cancer screening. *IEEE Transactions on Applied Superconductivity*, 25(3):1–5.
- Silva, L. F., Saade, D. C. M., Sequeiros, G. O., Silva, A. C., Paiva, A. C., Bravo, R. S., and Conci, A. (2014). A new database for breast research with infrared image. *Journal of Medical Imaging and Health Informatics*, 4(1):92–100.
- Usuki, H., Ikeda, T., Igarashi, Y., Takahashi, I., Fukami, A., Yokoe, T., Sonoo, H., and Asaishi, K. (1998). What kinds of non-palpable breast cancer can be detected by thermography? *Biomed Thermology*, 18(4):8–12.
- Veerlapalli, P. and Das, K. (2024). Roisegnet: A deep learning framework for automatic segmentation of region of interest from breast thermogram imagery. *International Journal of Intelligent Systems and Applications in Engineering*, 12(3):2248–2261.
- Veerlapalli, P. and Dutta, S. R. (2025). A hybrid gan-based deep learning framework for thermogram-based breast cancer detection. *Scientific Reports*, 15(19665).
- Wang, L., Bohler, T., Zohrer, F., Georgii, J., Rauh, C., Fasching, P., Brehm, B., Schulz-Wendtland, R., Beckmann, M., Uder, M., and Hahn, H. (2014). A hybrid method towards automated nipple detection in 3d breast ultrasound images. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 2869–2872, Chicago, Illinois, EUA. IEEE EMBC.
- Wishart, G., Campisi, M., Boswell, M., Chapman, D., Shackleton, V., Iddles, S., Hallett, A., and Britton, P. (2010). The accuracy of digital infrared imaging for breast cancer detection in women undergoing breast biopsy. *European Journal of Surgical Oncology*, 36(6):535–540.
- World Health Organization (2026). Global health observatory database: Cancer. disponível em: <https://www.who.int/news-room/fact-sheets/detail/cancer>. Acesso em: 25 maio 2026.
- Yasmin, M., Sharif, M., and Mohsin, S. (2013). Survey paper on diagnosis of breast cancer using image processing techniques. *Research Journal of Recent Sciences*, 2(10):88–98.