



Disciplina:

Análise de Algoritmos e Estrutura de Dados

Discente Especial:

Wagner Lopes Cardozo

Docente Orientadora da UC:

D.Sc. Lilian Berton

Tema da Pesquisa:

Implementação da Segmentação em Imagens de Mamografia Por Meio do Método Superpixel Grafo Com a Lógica Clusterização Interativa Linear Simples (SLCI): Comparação Assintótica com Outros Métodos de Superpixel Grafo.

Sumário:

- ☐ Motivação.
- ☐ TAD Grafo e Grafo Superpixel.
- ☐ Clusterização Linear Simples (SLIC) e Felzenszwalb, Watershed e Connected Components
- ☐ Objetivo, Metodologia, Resultados e Discursão.
- ☐ Análise Assintótica Big (O) e Conclusão.
- ☐ Referência Bibliográfica.
- ☐ Considerações finais.

Motivação:

Cenário do Problema na Sociedade
(Justificativa do Desenvolvimento e Implementação)

- ☐ Estimativa elevada para o **surgimento de novos casos de câncer de mama**.

- ☐ Corresponde a um **risco estimado de 66,54 novos casos a cada 100 mil mulheres**, totalizando **73.610** novos casos no Brasil.

- ☐ **Dificuldade de manter e implementar um programa de QC em Mamografia** (exigências da Portaria MS 453/98 e 531/2012 e RDC 330/2019).

Fonte: INCA-MS 2023 -

<https://www.inca.gov.br/sites/ufu.sti.inca.local/files//media/document//estimativa-2023.pdf>

Motivação:

Quem Sofre:

- ☐ A **população** que faz uso de Serviços de Mamografia.
- ☐ O **Médico** Radiologista ou Mastologista.
- ☐ O **Serviço** de Mamografia.

O Que Está Sendo Feito:

- ☐ Pesquisas científicas no universo da Ciência da Computação e Ciência de Dados.
- ☐ Gerando tecnologias inovadoras que podem contribuir para solucionar ou amenizar o problema.
- ☐ Aumentando a precisão e agilidade na conclusão do diagnóstico médico.

TAD Grafo:

- ❑ A TAD (Tipo Abstrato de Dados) Grafo é uma estrutura conceitual que modela relações entre entidades, representadas por vértices (nós) conectados por arestas (linhas).
- ❑ Baseada nos fundamentos da Teoria de Grafos, desenvolvida por Leonhard Euler no século XVIII.
- ❑ A TAD Grafo abstrai as propriedades de conectividade e interdependência em diversos sistemas, permitindo o estudo de estruturas complexas de forma sistemática.

Fonte: KAVEH, 1991

TAD Grafo:

- ❑ Sob uma perspectiva filosófico-matemática, os grafos transcendem seu uso puramente técnico, oferecendo uma linguagem universal para descrever a interação entre elementos em contextos.

Fonte: KAVEH, 1991

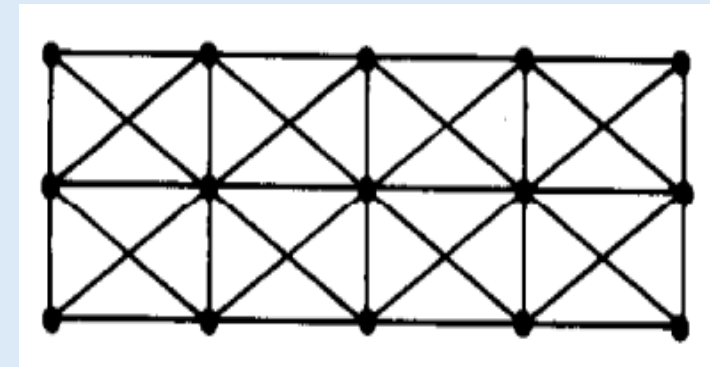
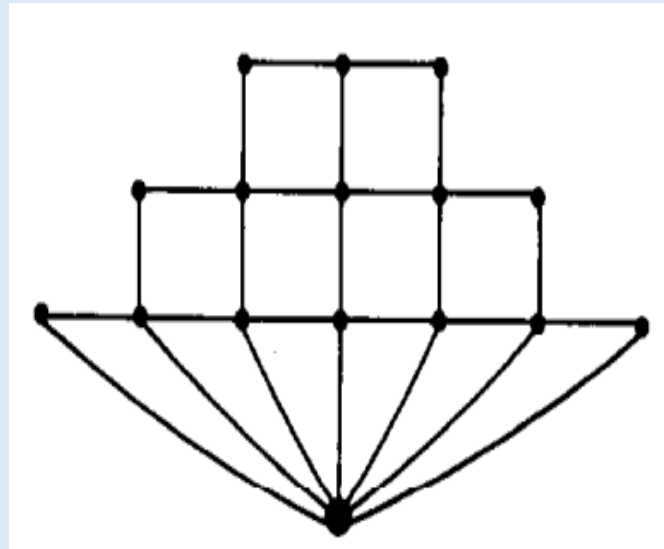
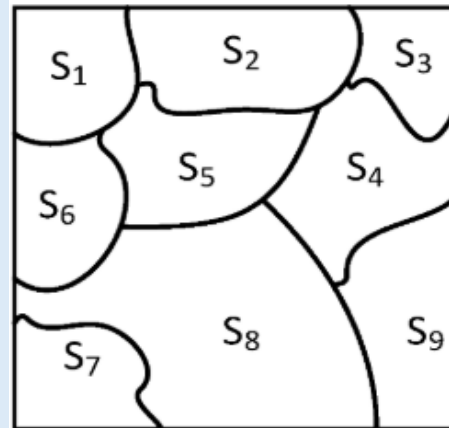


Figura-Fonte: KAVEH, 1991

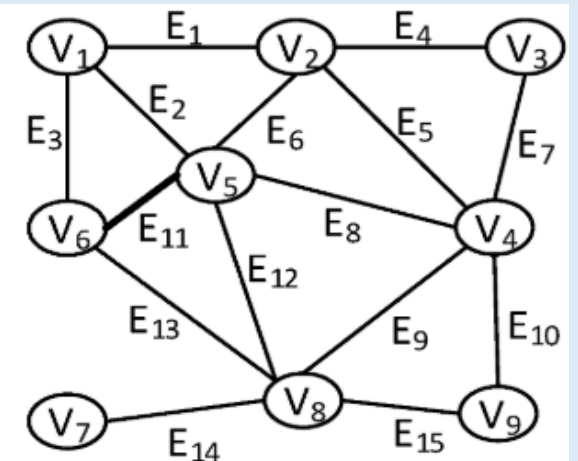
Superpixel Grafo:

- ❑ Superpixel grafo combina dois elementos fundamentais no processamento de imagens e análise de dados: superpixels e grafos.
- ❑ São agrupamentos de pixels que compartilham características semelhantes, como cor, textura ou intensidade.
- ❑ Reduz a complexidade de uma imagem ao organizar pixels em regiões mais significativas.

Fonte: YI, 2019



(a)



(b)

Superpixel Grafo:

- ❑ Os grafos fornecem uma estrutura matemática para representar relações entre essas regiões.
- ❑ Com vértices correspondendo a superpixels e arestas indicando conexões baseadas em proximidade ou similaridade.
- ❑ Então, a integração desses dois conceitos dá origem a um modelo eficaz para tarefas de segmentação de imagens, permitindo a análise mais eficiente e hierárquica de padrões complexos.

Algoritmo SLIC (Clusterização Interativa Linear Simples):

- ❑ Particiona a imagem em regiões de tamanho aproximadamente uniforme e de características homogêneas. **Fonte:** ACHANTA, 2020
- ❑ As regiões são representadas como vértices no grafo, com as arestas conectando superpixels vizinhos.
- ❑ As propriedades do grafo são então utilizadas para identificar e agrupar superpixels pertencentes a objetos ou regiões similares na imagem, reduzindo ruídos e melhorando a consistência da segmentação.

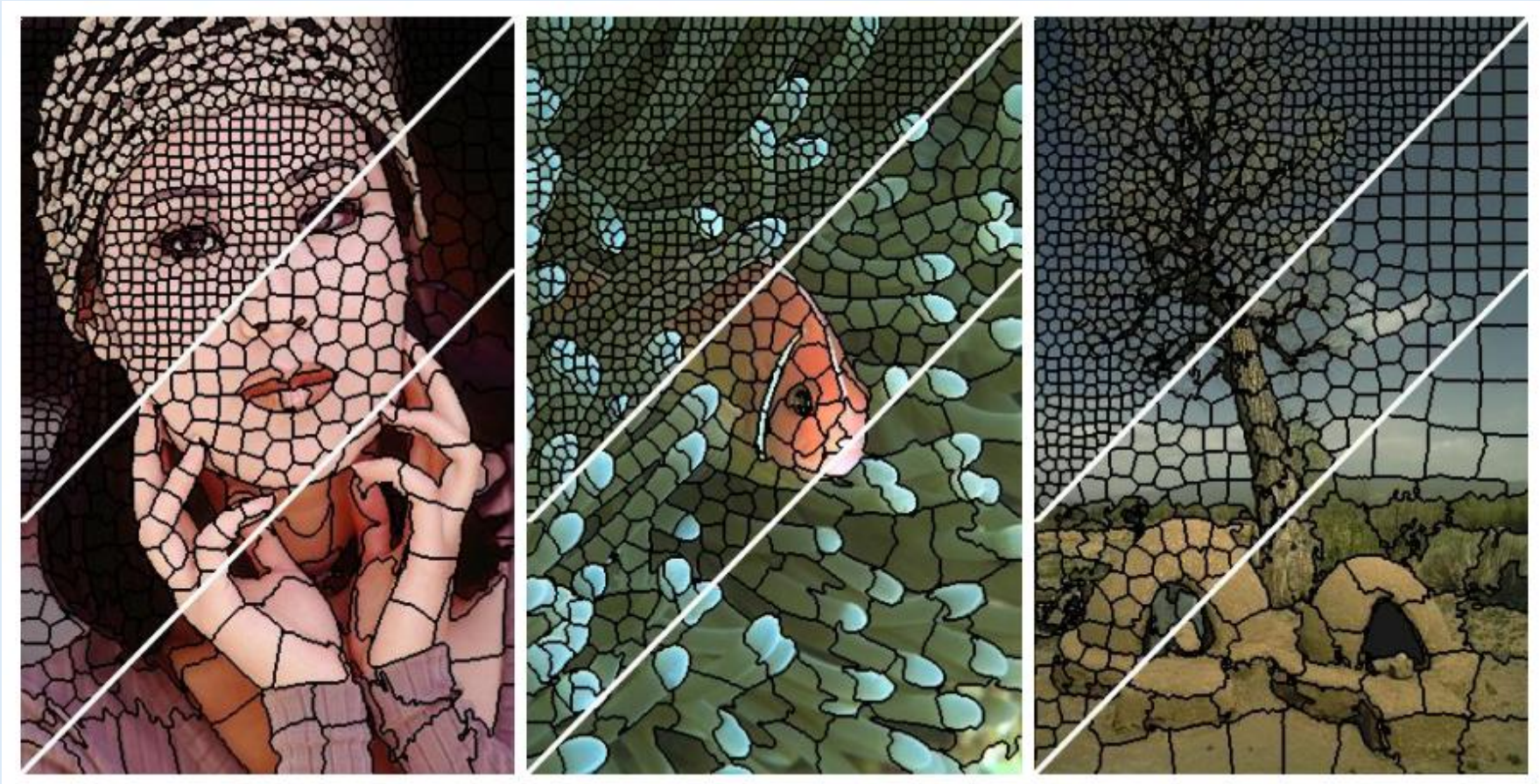
Fonte: HSU, 2013

Algoritmo SLIC (Clusterização Interativa Linear Simples):

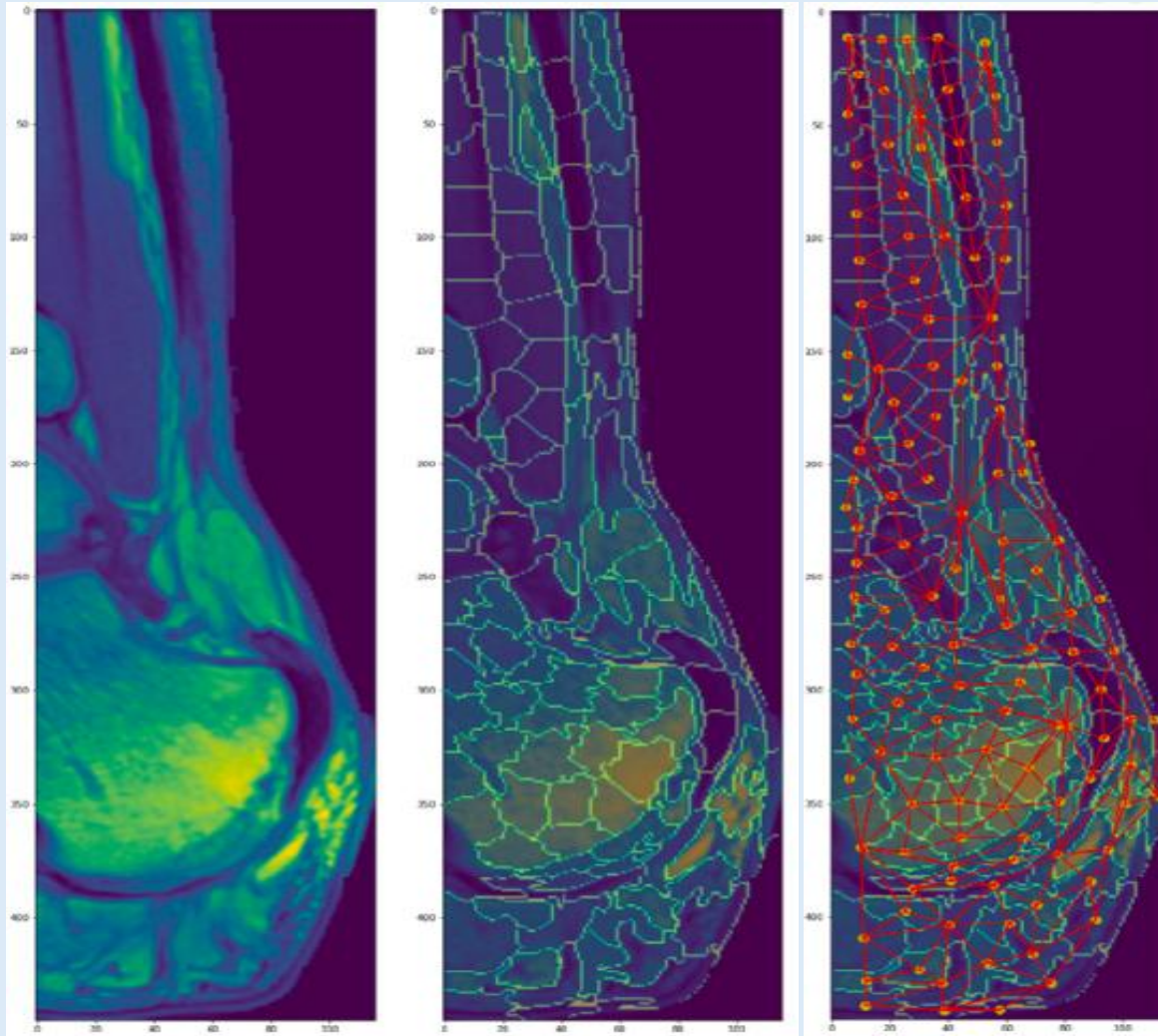
- ❑ Na técnica SLIC, os centróides desempenham um papel central no agrupamento e na definição de superpixels.
- ❑ Cada superpixel é representado por um centróide, que é inicialmente posicionado de forma uniforme na imagem e ajustado iterativamente durante o processo de clusterização.
- ❑ Esses centróides não apenas indicam a localização espacial do agrupamento, mas também encapsulam as características médias de cor e posição dos pixels pertencentes ao superpixel.

Fonte: GAMMOUDI, 2020.

AlgoritmoSLIC (Clusterização Interativa Linear Simples):



AlgoritmoSLIC (Clusterização Interativa Linear Simples):



- ❑ **RAGs** (Region Adjacency Graphs) são estruturas que representam relações entre regiões adjacentes de uma imagem, formadas por segmentação inicial, como superpixels ou clusters.
- ❑ Os nós representam regiões, e as arestas indicam conexões entre regiões vizinhas.
- ❑ São usadas para calcular propriedades como área, forma ou média de cores, reduzindo a complexidade ao analisar apenas regiões relevantes em vez de todos os pixels da imagem.

Felzenszwalb, Watershed e Connected Components:

- ❑ O algoritmo de **Felzenszwalb** se destaca por sua abordagem hierárquica baseada em grafos, na qual os pixels são tratados como nós conectados por arestas ponderadas.
- ❑ Sua segmentação ocorre pela análise iterativa dos componentes conectados, utilizando critérios de fusão que consideram as propriedades locais e globais da imagem.
- ❑ Essa característica torna o método especialmente eficiente para imagens com padrões complexos, fornecendo uma segmentação em tempo quase linear em relação ao número de pixels.

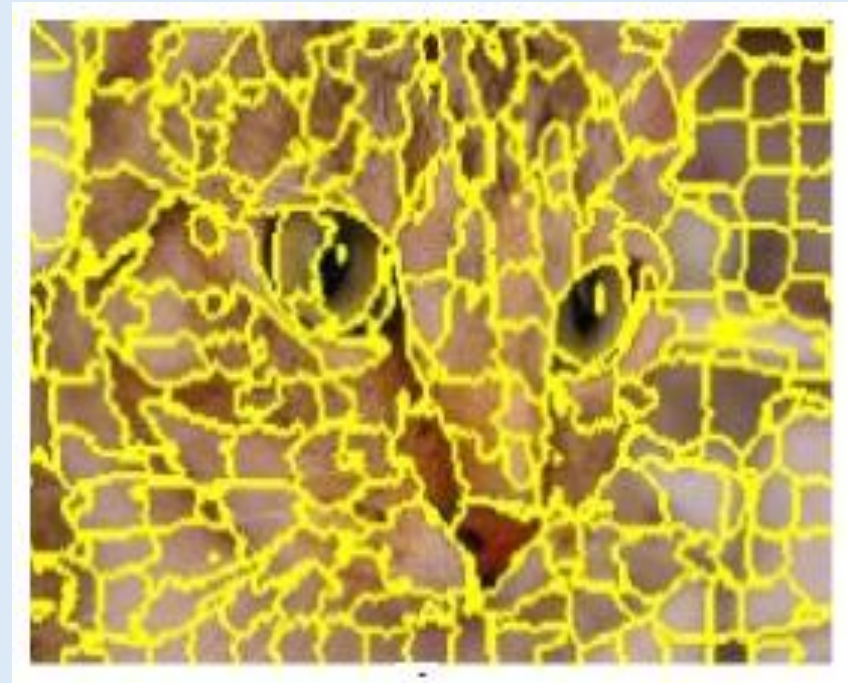
Felzenszwalb, Watershed e Connected Components:

- ❑ O algoritmo **Watershed** utiliza conceitos de morfologia matemática para tratar a imagem como uma topografia tridimensional.
- ❑ Segmentando com base na propagação de "águas" a partir de pontos de mínima intensidade.
- ❑ Este método é amplamente utilizado em aplicações biomédicas, devido à sua alta sensibilidade a detalhes e à capacidade de identificar bordas precisas.
- ❑ No entanto, sua sensibilidade ao ruído pode introduzir desafios em imagens não pré-processadas, exigindo técnicas adicionais de suavização.
- ❑ Apesar disso, sua eficiência computacional o torna uma escolha robusta para segmentação em imagens com gradientes bem definidos.

Felzenszwalb, Watershed e Connected Components:



Felzenszwalb



Watershed

Felzenszwalb, Watershed e Connected Components:

- ❑ O algoritmo de **Connected Components**, fundamentado na análise de componentes conectados, oferece uma abordagem simplificada para segmentação, identificando regiões contíguas com base em critérios binários ou multi-níveis.
- ❑ Este método é amplamente utilizado em cenários em que a conectividade espacial é um critério fundamental, como na detecção de objetos em imagens binárias ou segmentações preliminares em pipelines mais complexos.
- ❑ Sua implementação em tempo linear garante desempenho satisfatório mesmo em imagens de alta resolução, tornando-o uma ferramenta essencial em diversos sistemas de processamento de imagem.

Objetivo:

- ❑ Implementar o algoritmo de segmentação **SLIC** (Simple Linear Iterative Clustering) utilizando a Teoria de TAD Grafo.
- ❑ Posicionar o SLIC como uma alternativa assintoticamente mais eficiente (Big O) em comparação a métodos baseados em redes neurais profundas, como U-Net.
- ❑ Ao mesmo tempo realizar uma análise comparativa com outros algoritmos clássicos de segmentação, como Felzenszwalb, Watershed e Connected Components.

Objetivo:

- ❑ Para isso, serão avaliados critérios de tempo de execução e memória consumida, considerando a escalabilidade e a aplicabilidade prática em contextos clínicos.
- ❑ Tendo por objetivo também nortear um pipeline integrado que pode aprimorar a acurácia e a generalização dos modelos de CNNs e GCNs na tarefa de classificação, com especial atenção ao diagnóstico assistido por computador em imagens de mamografia.
- ❑ O trabalho visa demonstrar que essa abordagem pode reduzir significativamente os custos computacionais, mantendo altos níveis de precisão e escalabilidade.

Metodologia e Materiais:

- ❑ Implementação dos algoritmos de segmentação de imagens SLIC, Felzenszwalb, Watershed e Connected Components foi realizada utilizando a linguagem de programação Python 3.10.12.
- ❑ Integrando diversas bibliotecas científicas, como NumPy 1.26.4, Matplotlib 3.8.0, Scikit-Image 0.24.0, Tabulate 0.9.0, NetworkX 3.4.2, SciPy 1.13.1 e Time.

Metodologia e Materiais:

- ❑ O ambiente de desenvolvimento escolhido foi o Google Colab.
- ❑ Ambiente em computação em nuvem.
- ❑ Que inclui 100 núcleos de processamento, armazenamento de 100 GB, 12 GB. de memória RAM, e uma GPU NVIDIA T4 com 16 GB de memória dedicada para processamento gráfico.
- ❑ A biblioteca Google Colab File foi utilizada para facilitar o gerenciamento de arquivos e a integração com o dataset de imagens.

Metodologia e Materiais:

- ❑ O dataset utilizado foi o Mini-DDSM (Digital Database for Screening Mammography), disponível publicamente na plataforma Kaggle.
- ❑ O uso deste dataset foi realizado em conformidade com a Lei Geral de Proteção de Dados (LGPD).
- ❑ Os algoritmos de segmentação foram implementados para segmentar as estruturas anatômicas das imagens de mamografias.

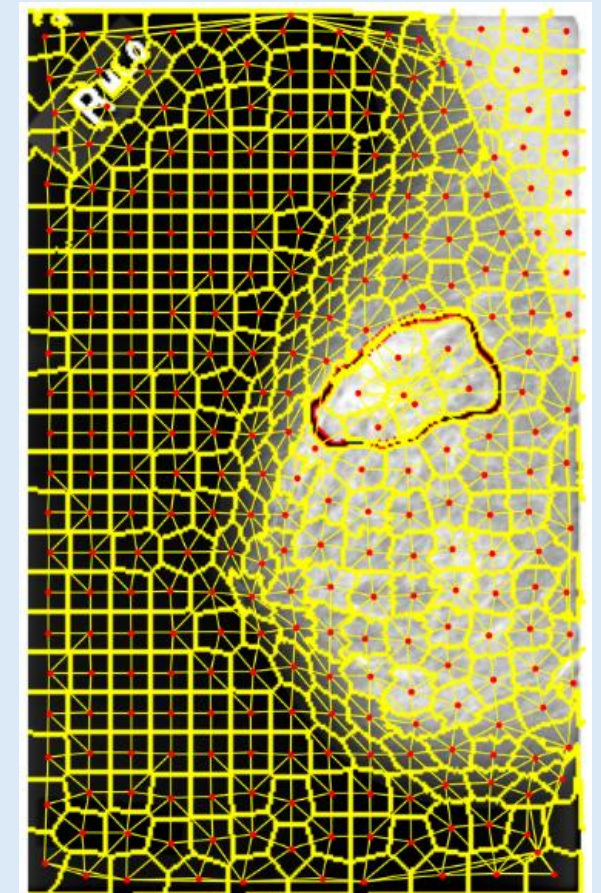
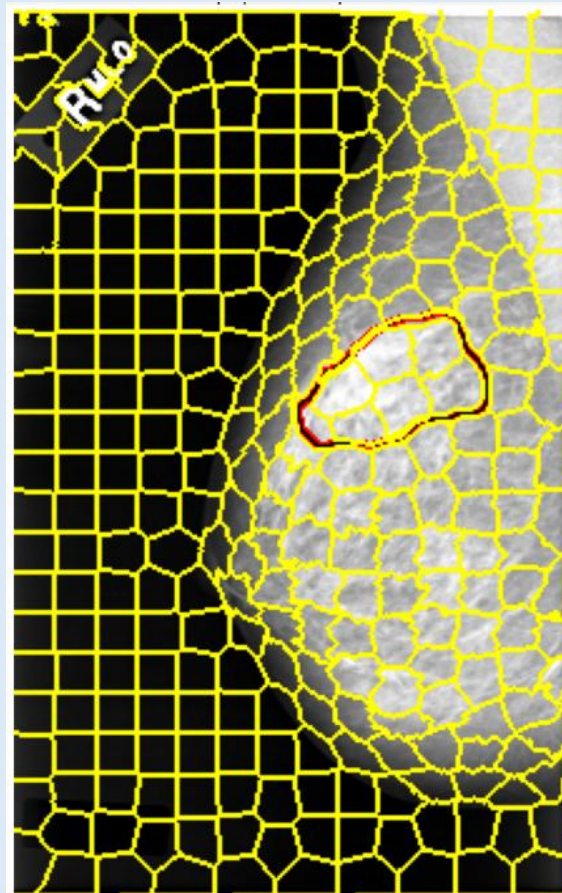
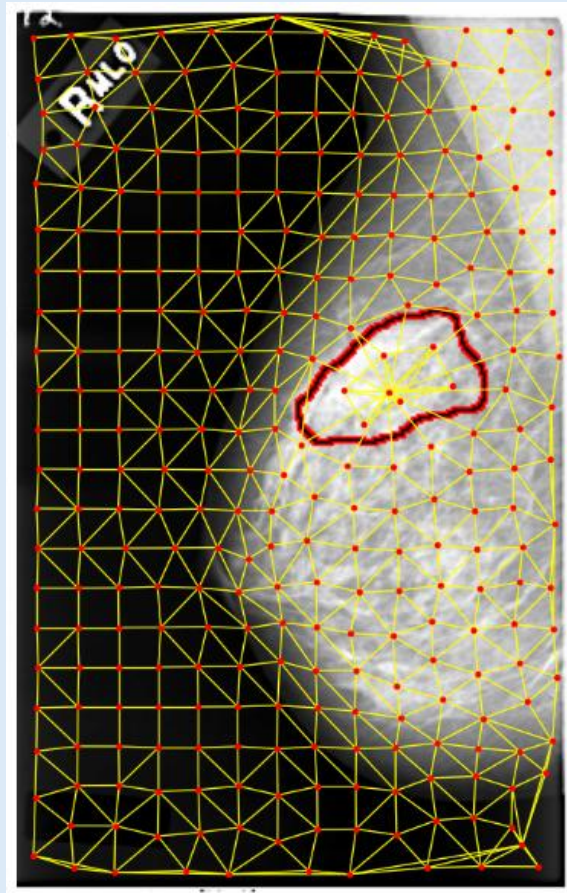
Fonte-Kaggle: <https://www.kaggle.com/datasets/cheddar/miniddsm>

Metodologia e Materiais:

- ☐ Medições precisas do tempo de execução e da memória consumida durante o processamento de cada imagem.
- ☐ Essas medições foram realizadas para avaliar o custo computacional dos métodos.
- ☐ Com foco no entendimento do comportamento assintótico dos algoritmos implementados.
- ☐ Podendo ser uma alternativa aos métodos de segmentação usando modelos de arquiteturas U-Net.

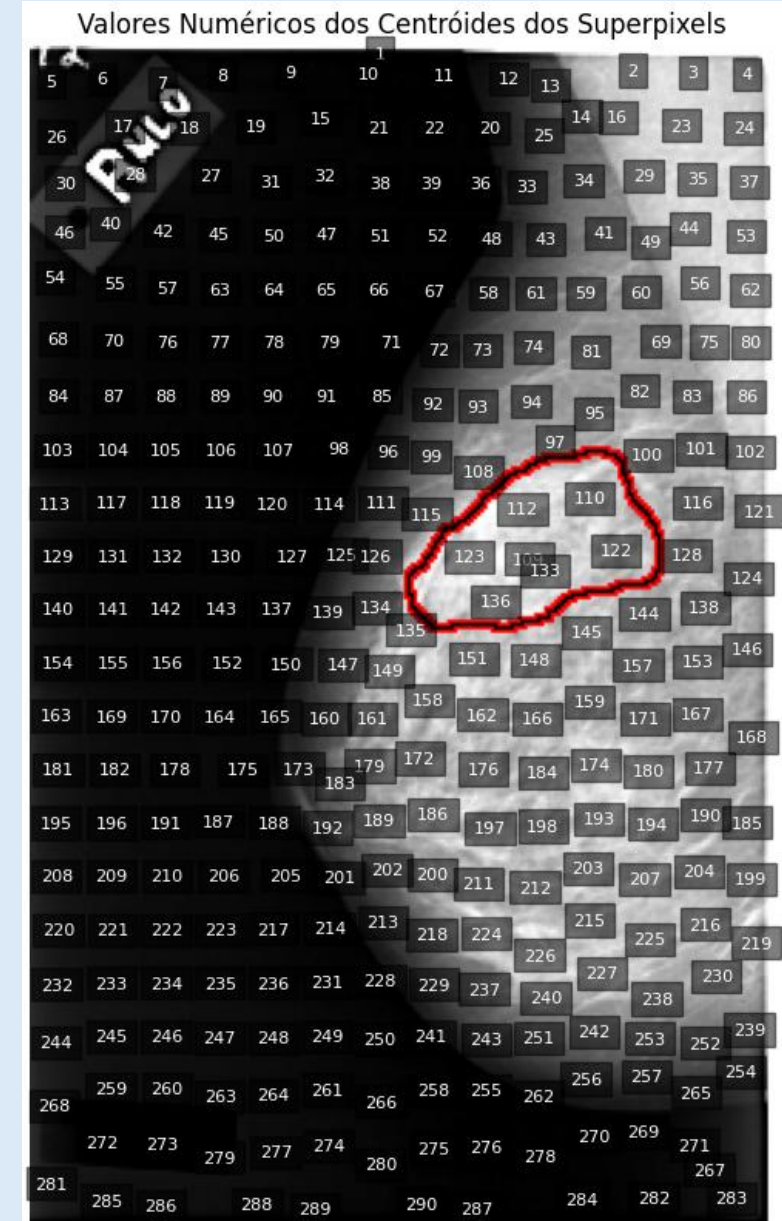
Resultados e Discussão:

- ❑ O algoritmo **SLIC** demonstrou ser uma solução robusta, gerando superpixels que preservaram a coerência espacial e estrutural das imagens.



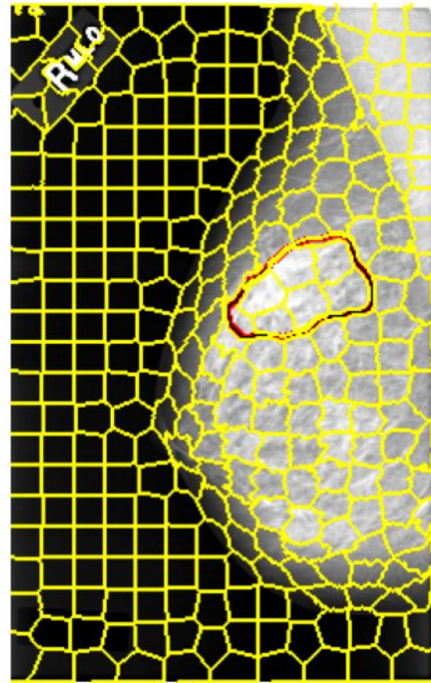
Resultados e Discussão:

Propriedade	Quantidade
Nós (Vértices)	290
Arestas	780



Resultados e Discussão:

Massa Tumoral Por Meio da Técnica de Segmentação SLIC e Grafo Superpixel



Máscara da Região Com a Massa Tumoral

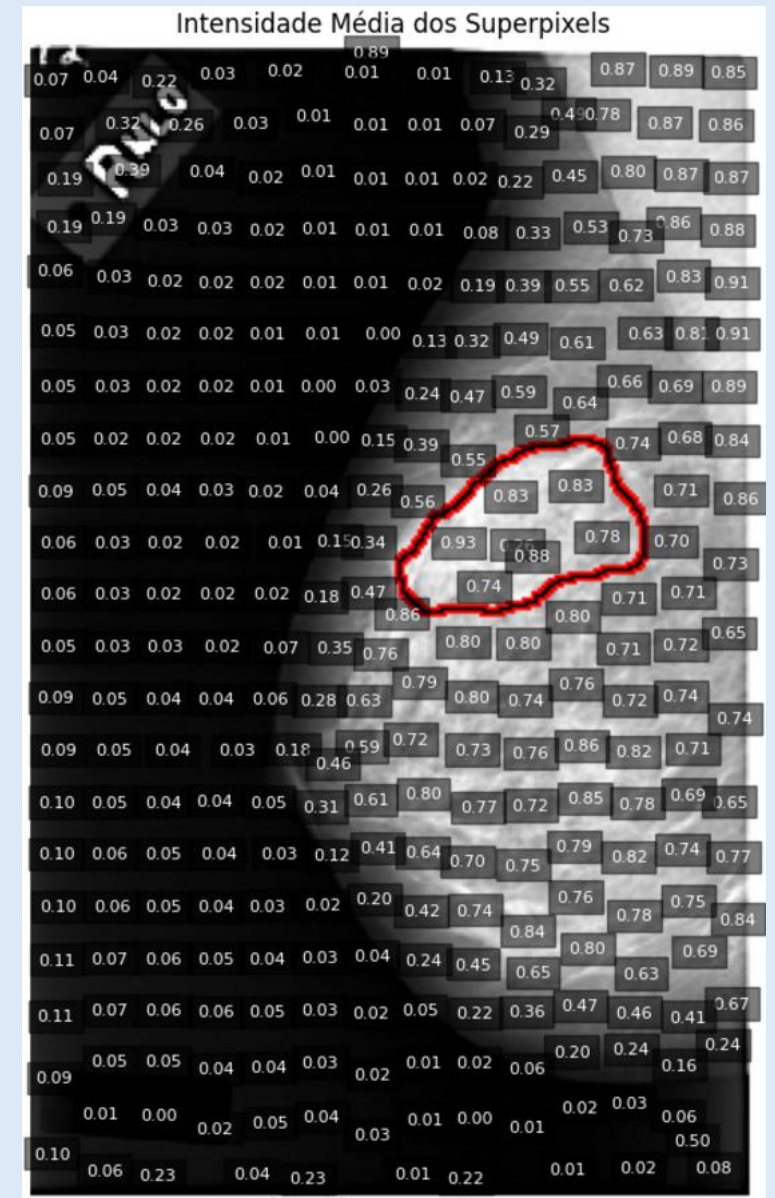


Massa Tumoral Segmentada



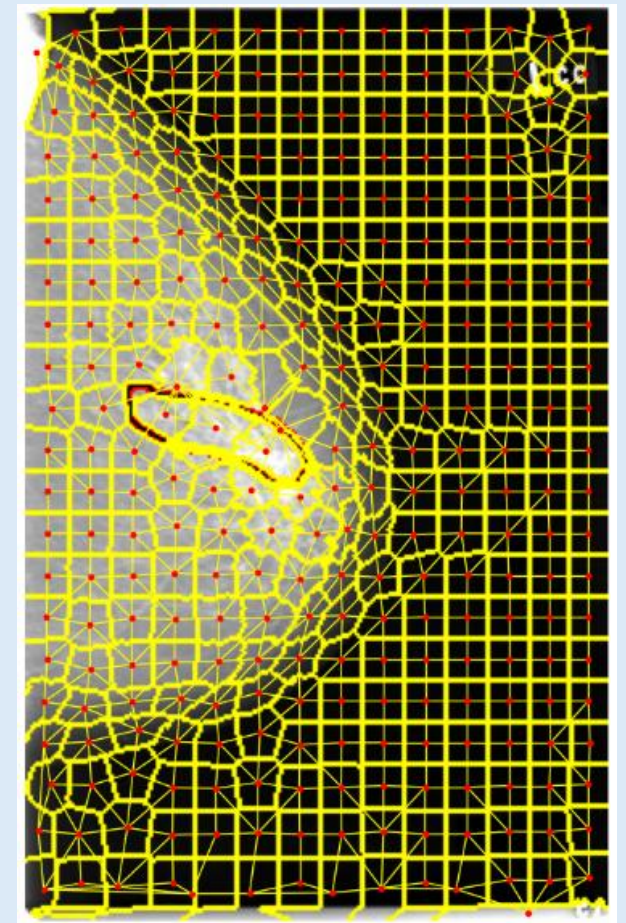
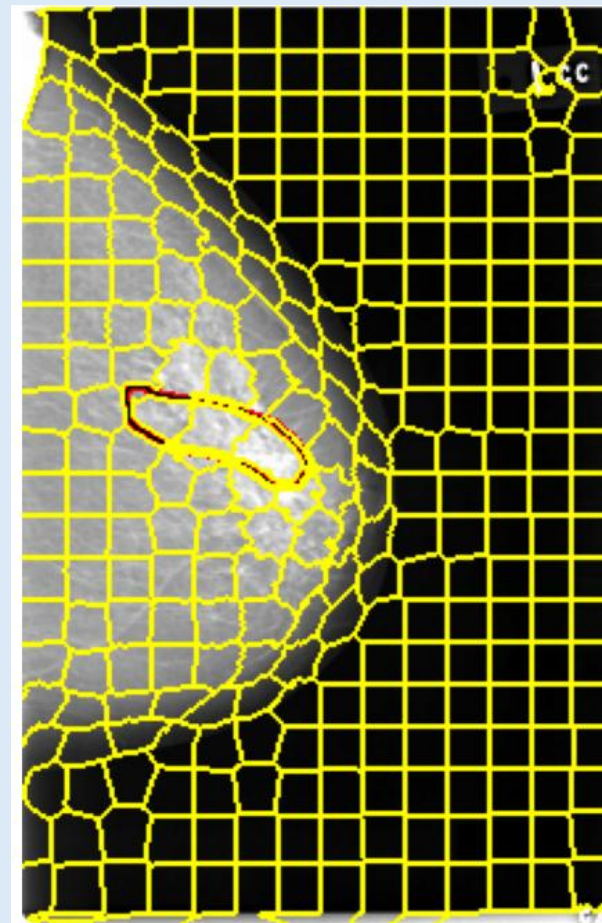
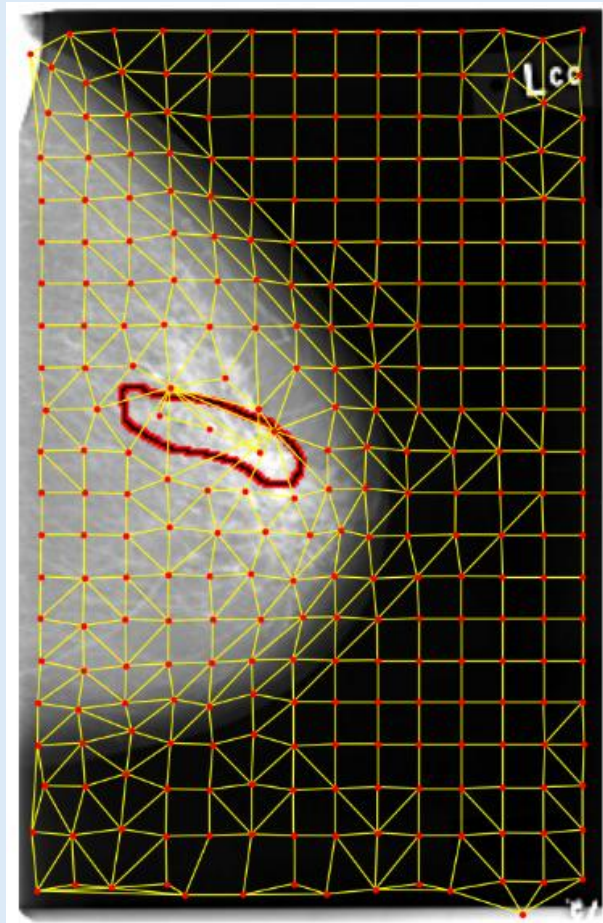
Resultados e Discussão:

- ❑ Extração de características das imagens das regiões com achado patológico que podem auxiliar os modelos de arquiteturas de redes neurais convolucionais CNNs ou redes de grafos convolucionais GCNs.



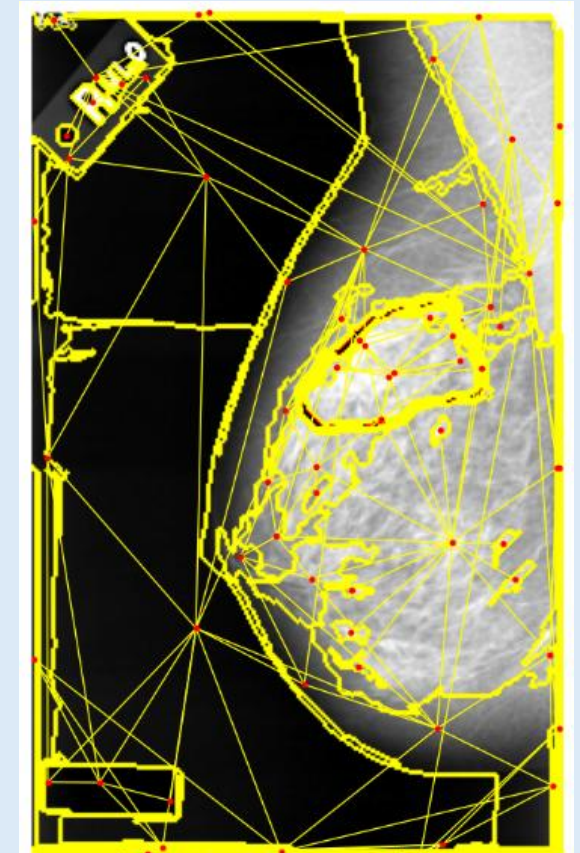
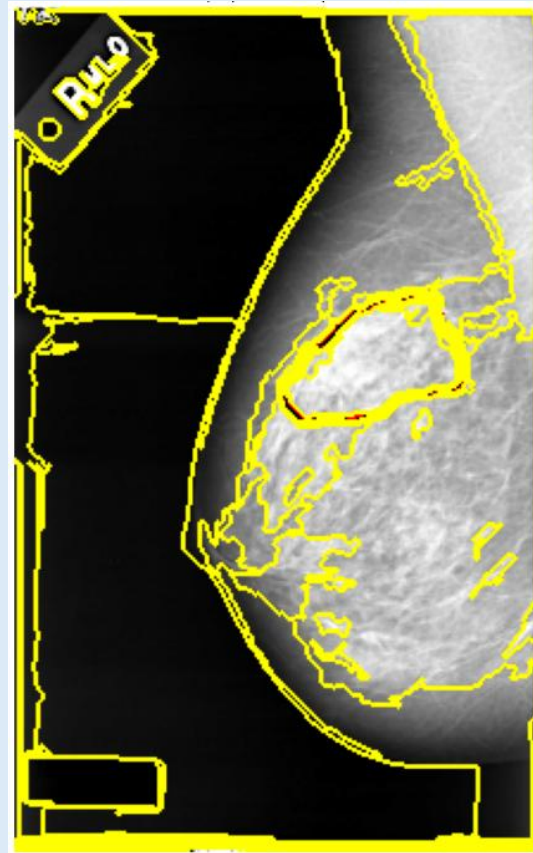
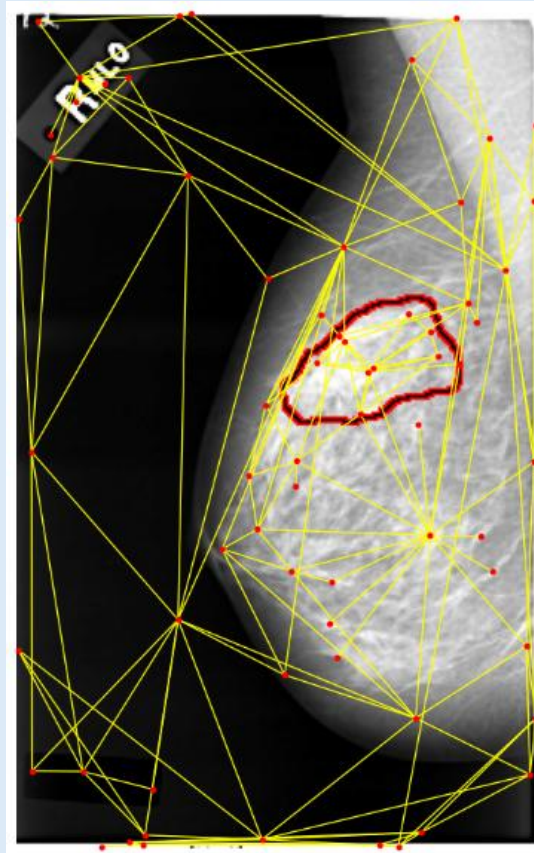
Resultados e Discussão:

- ❑ SLIC Para Outra Imagem do Dataset



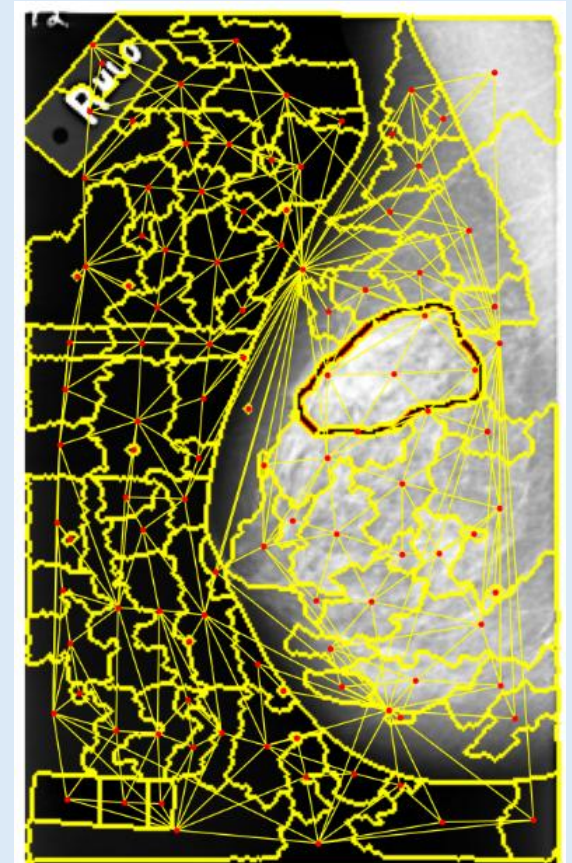
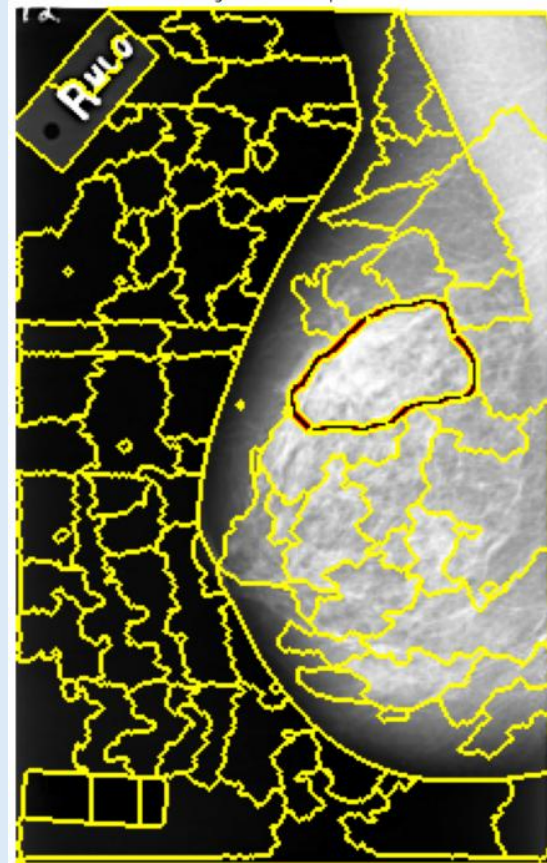
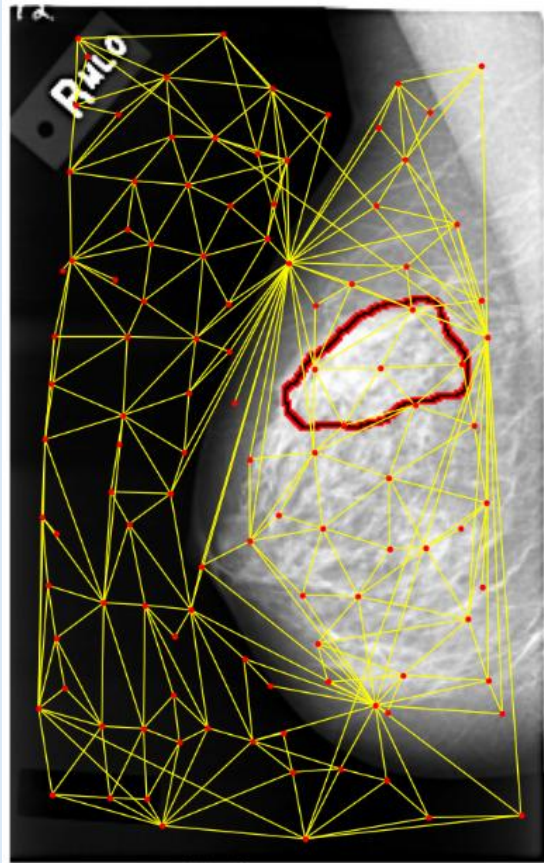
Resultados e Discussão:

- ❑ algoritmo **Felzenszwalb** produziu segmentações detalhadas, baseando-se em um modelo hierárquico de agrupamento.



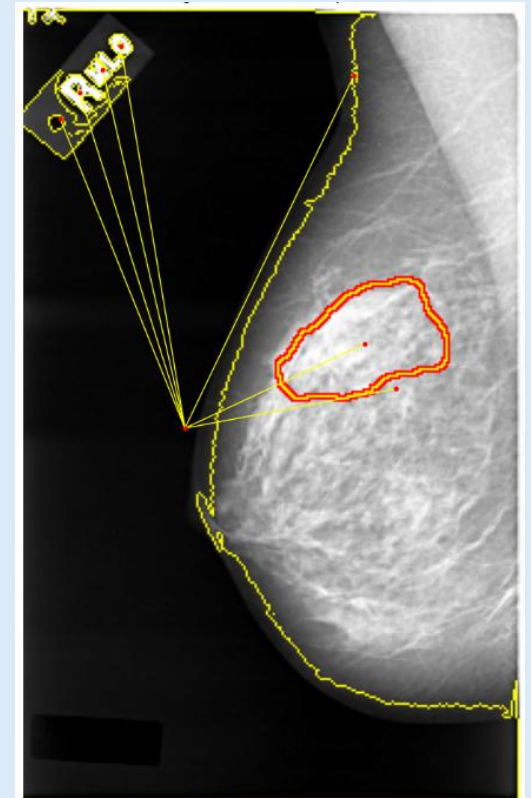
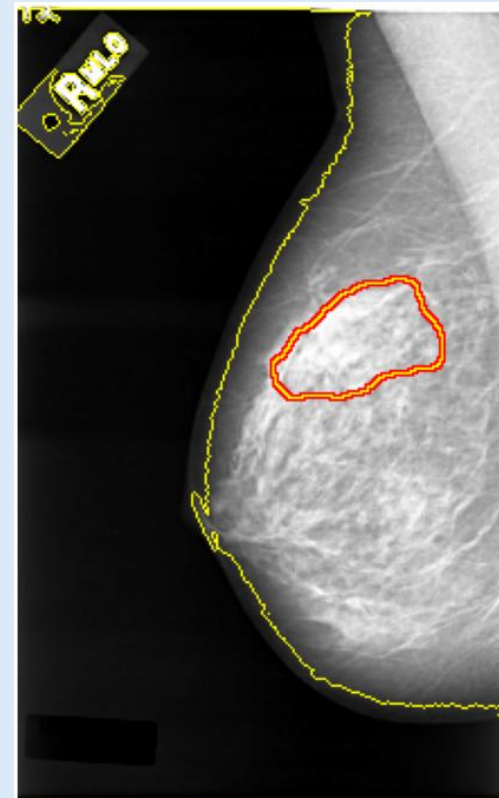
Resultados e Discussão:

- ❑ Por sua vez, o método **Watershed** apresentou segmentações precisas e bem adaptadas às bordas naturais das imagens de mamografia.



Resultados e Discussão:

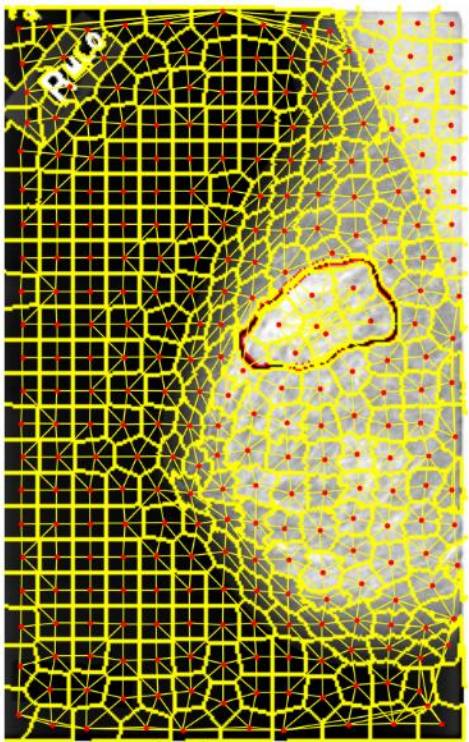
- ❑ O método **Connected Components** gerou segmentações claras e bem delimitadas, demonstrando eficiência na separação de regiões conectadas em áreas homogêneas.



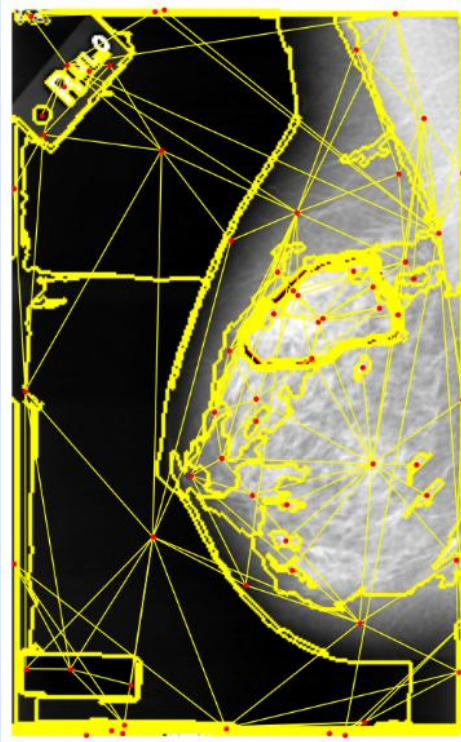
Resultados e Discussão:

- ❑ As imagens de mamografia segmentadas pelos algoritmos implementados apresentaram características específicas que evidenciam as particularidades de cada método.

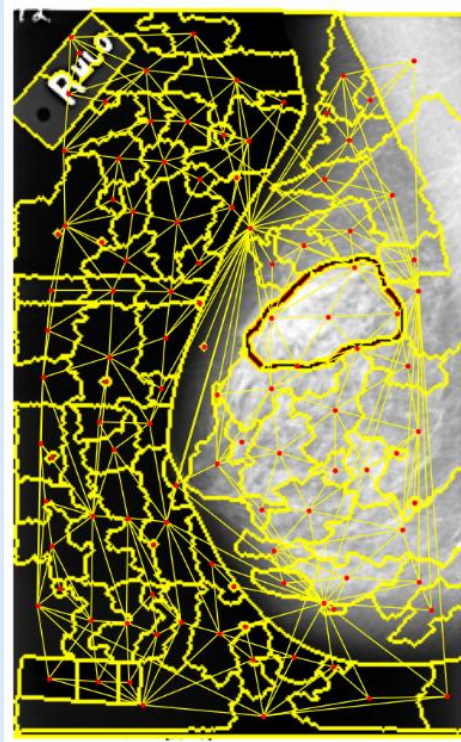
SLIC



Felzenszwalb



Watershed



Connected Components



Resultados e Discussão:

☐ Tempo para 1 imagem do dataset

Método	Tempo (segundos)	Memória (MB)
SLIC	0.44673919677734375	0.6839561462402344
Felzenszwalb	0.48841381072998047	0.7021217346191406
Watershed	0.41265225410461426	0.7007484436035156
Connected Components	0.09787392616271973	0.6844062805175781

☐ Tempo médio para 500 imagens do dataset

Método	Tempo (segundos)	Memória (MB)
SLIC	0.5005476474761963	0.7043571472167969
Felzenszwalb	0.6444883346557617	0.7220039367675781
Watershed	0.5368821620941162	0.7189826965332031
Connected Components	0.16281723976135254	0.718170166015625

Análise Assintótica Big (O):

- ☐ O algoritmo **SLIC** utiliza uma abordagem baseada em K-means para gerar superpixels.
- ☐ Sua complexidade assintótica pode ser analisada em relação ao número de pixels na imagem (N) e ao número de superpixels (K).
- ☐ Inicializa K centróides.
- ☐ Realiza T iterações, onde cada pixel é associado ao centróide mais próximo.
- ☐ Calcula novos centróides baseados nos pixels associados.

Análise Assintótica Big (O):

- ❑ Durante cada iteração, o algoritmo verifica cada pixel N e compara com os K centróides próximos.
- ❑ Como o algoritmo SLIC limita a busca a uma região (janela $m \times n$) ao redor do centróide, reduz o número de comparações de $N \cdot K$ para algo proporcional a N , dependendo da janela.
- ❑ Atualizar os K centróides após cada iteração também é $O(N)$.
- ❑ Logo, a **Complexidade Total** considerando T iterações será:

$$O(N \cdot T)$$

Análise Assintótica Big (O):

- ❑ A **U-Net** é um modelo de arquitetura de rede neural convolucional, sua complexidade depende de:
 - ✓ O tamanho da imagem de entrada ($H \times W$).
 - ✓ O número de canais de entrada (C).
 - ✓ O número de filtros por camada (F).
 - ✓ O número de camadas (L).

- ❑ Consiste em L camadas convolucionais que reduzem progressivamente as dimensões da imagem, geralmente por fatores de 2.

Análise Assintótica Big (O):

- ❑ A complexidade de cada camada convolucional é:

$$O(H \times W \times C \times F \times k^2)$$

- ❑ Onde k é o tamanho do kernel.
- ❑ Somando todas as camadas e considerando uma redução de dimensão ($H/2^L \times W/2^L$) em cada camada: $O(H \times W \times C \times F \times k^2)$.
- ❑ Se L cresce linearmente com H ou W , o custo pode ser

$$O(H^2 \times W^2) - \text{no pior caso}$$

Análise Assintótica Big (O):

Algoritmo (Segmentação)	Complexidade Assintótica
SLIC	$O(N \times T)$
Felzenszwalb	$O(N \log N)$
Watershed	$O(N \times T)$
Connected Components	$O(N \times T)$
U-Net	$O(H^2 \times W^2)$

Conclusão:

- ❑ Os algoritmos superpixel grafo se mostraram bem eficientes como método alternativo para segmentação de imagens de mamografia.
- ❑ São eficientes assintoticamente Big (O) em relação a modelos de arquiteturas de redes neurais U-Nets.
- ❑ Cada algoritmo superpixel grafo fazendo uso da TAD grafo, tem características semelhantes com aplicabilidades recomendadas dependendo do universo de imagens a ser segmentada.

Conclusão:

- ❑ Pode ser aplicado em outros universos de imagens para o diagnóstico clínico como ultrassonografias e imagens histológicas.
- ❑ Podendo ser implementado também para outros estudos com imagens, como por exemplo imagens climatológicas (satélite, radar, topográfica, monitoramento climático).
- ❑ Pode ser uma alternativa para auxiliar a classificação de modelos de arquiteturas de CNNs e GCNs (sugestão de trabalhos futuros).
- ❑ Podem contribuir para o processamento de imagens que irá compor o pipeline na criação de IAs éticas, responsáveis e democráticas, baseadas em CNNs ou GCNs.

Referências Bibliográficas:

- ❑ ACHANTA, R., et al., **SLIC Superpixels**, EPFL Technical Report 149300, June 2010
- ❑ CARTER, S. M., et al., **The ethical, legal and social implications of using artificial intelligence systems in breast cancer care**, The Breast 49 (2020) 25e32, <https://doi.org/10.1016/j.breast.2019.10.001>
- ❑ GAMMOUDI, I., MAHJOUB, M. A., GUERDELLI, F., **Unsupervised Image Segmentation based Graph Clustering Methods**, *Computación y Sistemas*, Vol. 24, No. 3, 2020, pp. 969–987, doi: 10.13053/CyS-24-3-3059
- ❑ HSU, C. Y., DING, J. J., **Efficient Image Segmentation Algorithm Using SLIC Superpixels and Boundary-focused Region Merging**, *2013 9th International Conference on Information, Communications & Signal Processing*, IEEE Xplore: 2014, DOI: [10.1109/ICICS.2013.6782861](https://doi.org/10.1109/ICICS.2013.6782861)
- ❑ INCA (Brasil), Estimativa 2023, **Incidência do Câncer no Brasil**, Instituto Nacional do Câncer, Rio de Janeiro: INCA, 2022
- ❑ KAVEH, A., **Graphs and structures**, *Computers & Structures*, *Volume 40, Issue 4*, 1991, Pages 893-901, [https://doi.org/10.1016/0045-7949\(91\)90319-H](https://doi.org/10.1016/0045-7949(91)90319-H)
- ❑ KUPPILI, V. N. S., CHEELU, D., BABU, M. R., KRIDHNA, P.V., **Image Segmentation based on Connected Component Analysis**, *International Journal of Advances in Science and Technology*, Vol. 2, No.3, 2011, ISSN 2229 5216
- ❑ MIHAILA, A. F., et al., **On Image Segmentation Using A Combination Of Felzenszwalb, Slic And Watershed Methods**, *Journal Of Information Systems & Operations Management*, Vol. 14.1, May 2020
- ❑ PACILÈS, Serena et al., **Improving breast cancer detection accuracy of mammography with the concurrent use of an artificial intelligence tool**. *Radiology: Artificial Intelligence*, v. 2, n. 6, p. e190208, 2020.

Muito Obrigado!!!

Duvidas, Perguntas ou Questionamentos

E-mail: wagner.cardozo72@gmail.com

LinkedIn: www.linkedin.com/in/wagner-lopes-cardozo-8b4a031ab