



UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO  
DEPARTAMENTO DE COMPUTAÇÃO

**Kallebe Felipe Pereira Bezerra**

**DESENVOLVIMENTO DE UM ALGORITMO BASEADO EM LÓGICA FUZZY PARA  
SEGMENTAÇÃO DE LESÕES EM IMAGENS DE MAMOGRAFIA DIGITAL**

Recife - PE

2018

**KALLEBE FELIPE PEREIRA BEZERRA**

**DESENVOLVIMENTO DE UM ALGORITMO BASEADO EM LÓGICA FUZZY PARA  
SEGMENTAÇÃO DE LESÕES EM IMAGENS DE MAMOGRAFIA DIGITAL**

Monografia apresentada ao Curso de Bacharelado em Ciência da Computação da Universidade Federal Rural de Pernambuco, como requisito parcial para obtenção do título de Bacharel em Ciência da Computação.

Orientador: Prof . Dr. Filipe Rolim Cordeiro

Recife - PE

2018

Dedico este trabalho a Deus, a meus pais, Pereira e Mércia, a minha esposa Nathalia e a minha irmã Vanessa.



## **AGRADECIMENTOS**

A Deus o criador do universo, que me deu a oportunidade de mudar de cidade pra fazer um curso de bacharelado que sonhei, e que me deu tudo que precisava durante todos esses anos de graduação.

Aos meus pais que me criaram e fizeram todos os esforços para me educar, e me criar em bons caminhos.

A minha esposa Nathalia, que Deus colocou no meu caminho, e esteve comigo desde o começo do curso, sempre me auxiliando e me ajudando.

Ao professor Filipe que me orientou desde da iniciação científica até o trabalho de conclusão de curso, sempre me ajudando quando tive dificuldades.

“Sonhos determinam o que você quer.

Ação determina o que você conquista.”

Aldo Novak

## RESUMO

O câncer de mama tem sido um problema crescente para mulheres do mundo todo. De acordo com a Organização Mundial de Saúde (OMS), ele é o tipo de câncer mais comum entre mulheres, com mortalidade crescente, se tornando um dos tipos mais fatais de câncer em todo o mundo. No Brasil, ele é a principal causa de morte por câncer entre as mulheres, sendo estimado 59.000 casos novos de câncer em 2018, com uma incidência de cerca de 59,70 casos de cada 100 mil mulheres. Vários métodos de prevenção têm sido desenvolvidos, mas um dos métodos mais eficazes para a detecção de lesão é o diagnóstico através da mamografia digital. No entanto, a interpretação da mamografia pode ser uma tarefa difícil até mesmo para um especialista, uma vez que a análise é afetada por diversos fatores, tais como a qualidade da imagem, experiência do radiologista e tipo de lesão. É estimado que 12% a 30% dos casos de câncer de mama deixam de ser detectados devido à má interpretação mamográfica. O trabalho proposto tem como objetivo principal o estudo e desenvolvimento de uma técnica de segmentação de tumores em imagens de mamografia, utilizando lógica *Fuzzy*. Têm-se como objetivo inserir a abordagem *Fuzzy* no algoritmo *Random Walker*, a fim de propor uma nova solução para segmentação de lesões. Por fim, tem-se como objetivo realizar comparação com técnicas do estado da arte. O trabalho proposto utiliza a base de dados Mini-MIAS para realizar a avaliação das técnicas. A base consiste em 322 imagens de mamografia da visão médio lateral oblíqua, obtido de 161 pacientes. Entretanto, apenas 57 das imagens contem lesão. Os resultados mostraram que a abordagem proposta do algoritmo *Random Walker* com lógica *Fuzzy*, usado para segmentação do tumor, obteve melhores resultados para a maioria das métricas, em comparação com o algoritmo *Random Walker* clássico, e diminuiu o esforço do usuário necessário na etapa de inicialização do algoritmo.

**Palavras-chave:** Segmentação – Mamografia Digital – Algoritmo *Fuzzy* – Câncer de Mama – *Random Walker*.

## ABSTRACT

Breast cancer has been a growing problem for women around the world. According to the World Health Organization (WHO), it is the most common type of cancer among women, with increasing participation, making it one of the most fatal types of cancer worldwide. In Brazil, it is the leading cause of cancer death among women, with 59.000 new cases of cancer in 2018, with an incidence of about 59,70 cases per 100,000 women. Several methods of prevention have been developed, but one of the most effective methods for the detection of lesions is the diagnosis through digital mammography. However, the interpretation of mammography can be a difficult task even for a specialist, since the analysis is affected by several factors, such as image quality, radiologist experience and type of lesion. 12% to 30% of breast cancer cases are not detected because of bad mammography interpretation. The main objective of this work is the study and development of a tumor segmentation technique in mammography images using Fuzzy logic. It aims to insert the Fuzzy approach in the algorithm Random Walker, in order to propose a new solution for lesion segmentation. It aims to insert the Fuzzy approach in the algorithm Random Walker, in order to propose a new solution for segmentation of tumors. Finally, this work compare the results with state of the art techniques. The database has 322 mammography images obtained from 161 patients. However, only 57 of the images contain masses. Results showed that the proposed approach of the Random Walker with Fuzzy logic, used for mass segmentation, obtained better results when compared with the classic Walker Random algorithm, besides decreasing the user effort in the algorithm initialization step.

**Key-words:** Segmentation - Digital Mammography - Fuzzy Algorithm - Breast Cancer - Random Walker.



## Lista de Figuras

Figura 1 – Ilustração da composição de imagem por pixels.....	19
Figura 2 – (a) Imagem Original. (b) Imagem Segmentada. (c) Contorno da imagem segmentada sobreposta na imagem original.....	20
Figura 3 – Funções de pertinência do ponto de vista da (a) lógica convencional e (b) da lógica fuzzy.....	22
Figura 4 – Processo de segmentação Random Walker: (a) Imagem Original; (b) Sementes Iniciais; (c) Imagem Segmentada.....	34
Figura 5 – Fluxograma do Fuzzy Random Walker.....	36
Figura 6 – Estágios do Fuzzy Random Walker. (a) Imagem original; (b) Seleção de Semente; (c) Região Fuzzy Gaussiana; (d) Imagem Segmentada.....	37
Figura 7 – Mapeamento do indivíduo no modelo de geração automática de sementes.....	38
Figura 8 – Funções de pertinências implementados: (a) Triangular; (b) Trapezoidal; (c) Gaussiano; (d) Sino Generalizada.....	42
Figura 9 – Comparação da segmentação do RW e FRW para as imagens mdb010, mdb021, mdb028 e mdb190. Coluna(a): RW sementes; Coluna(b): RW segmentação; Coluna(c): FRW sementes; Coluna(d): FRW segmentação.....	45
Figura 10 – Curva ROC com parâmetros variados do RW clássico e da abordagem proposta.....	47
Figura 11 – Comparação do Random Walker e a abordagem proposta, para métricas de sensibilidade, especificidade, BAC, Jaccard e Dice.....	48
Figura 12 – Imagens de resultados.....	49
Figura 13 – Comparação da segmentação do FRW com sementes gerada por DE e PSO Coluna(a): DE sementes; Coluna(b): FRW segmentação; Coluna(c): PSO sementes; Coluna(d): FRW segmentação.....	50
Figura 14 – Curva ROC com parâmetros variados do FRW com inicialização Manual e Automática DE e PSO.....	51

## Lista de tabelas

Tabela 1– Espaço de Exploração de Parâmetros do <i>Random Walker</i> .....	43
Tabela 2– Espaço de Exploração de Parâmetros de Evolução Diferencial.....	49
Tabela 3 – Espaço de Exploração de Parâmetros de Otimização por Enxame de Partículas.....	50
Tabela 4– Comparações de resultados por variações de funções de pertinência nos algoritmos <i>Random Walker</i> e <i>GrowCut</i> .....	52

## Lista de abreviaturas e siglas

OMS	Organização Mundial de Saúde
RW	<i>Random Walker</i>
CAD	<i>Computer Aided Diagnosis</i> (Diagnóstico Auxiliado pelo Computador)
FRW	<i>Fuzzy Random Walker</i>
VP	Verdadeiros Positivo
FN	Falso Negativo
VN	Verdadeiros Negativo
FP	Falso Positivo
SE	Sensibilidade
SP	Especificidade
BAC	Acurácia balanceada
DE	Evolução Diferencial
PSO	<i>Particle Swarm Optimization</i> (Enxame de Partículas)

## SUMÁRIO

<b>1. INTRODUÇÃO</b>	<b>16.</b>
1.1 Objetivos	19.
1.1.1 Geral	19.
1.1.2 Específicos	19.
1.2 Estrutura do Trabalho	20.
<b>2. FUNDAMENTAÇÃO TEÓRICA</b>	<b>18.</b>
2.1 Conceitos Básicos de Imagem	18.
2.1.1 Segmentação	20.
<b>2.2. Conceitos de Inteligência Artificial</b>	<b>20.</b>
<b>2.2.1 <i>Random Walker</i></b>	<b>21.</b>
<b>2.2.1 Lógica Fuzzy</b>	<b>22.</b>
2.3 Algoritmos de Otimização	23.
2.3.1 Evolução Diferencial	23.
2.3.2 Enxame de Partículas	25.
2.4 Soluções Numéricas	26.

2.4.1	Gradiente Conjugado	26.
2.4.1.1	Pré-Condicionador <i>Multgrid</i>	27.
<b>3.</b>	<b>REVISÃO BIBLIOGRÁFICA</b>	<b>28.</b>
<b>4.</b>	<b>MATERIAIS E MÉTODOS</b>	<b>32.</b>
4.1	AMBIENTE EXPERIMENTAL	32.
4.2	PROPOSTA DA PESQUISA	33.
4.2.1	Segmentação de Imagem	33.
4.2.2	Geração Automática de Sementes	37.
4.2.2.1	Experimentos de Geração Automática	39.
4.3	Métricas de Avaliação	39.
4.5	FUNÇÕES DE PERTINÊNCIAS	40.
<b>5.</b>	<b>RESULTADOS</b>	<b>43.</b>
5.1	Otimização de Parâmetros	44.
5.2	<i>Resultados Fuzzy Random Walker</i>	44.
5.3	Resultados Geração Automática de Sementes	48.
5.4	Resultados Variação de Funções de Pertinência	52.
5.5	Discussão dos Resultados	53.

<b>6. CONCLUSÕES</b>	<b>54.</b>
6.1. CONTRIBUIÇÕES	55.
<b>7. REFERÊNCIAS BIBLIOGRÁFICAS</b>	<b>56.</b>

## 1. INTRODUÇÃO

O câncer de mama tem sido um problema crescente para mulheres do mundo todo. De acordo com a Organização Mundial de Saúde (OMS), ele é o tipo de câncer mais comum entre mulheres, com mortalidade crescente, se tornando um dos tipos mais fatais no mundo todo (WORLD HEALTH ORGANIZATION, 2014). Aproximadamente 246.660 de novos casos foram diagnosticados em 2016, setenta e cinco por cento das sobreviventes de câncer de mama (mais de 2,6 milhões de mulheres) têm 60 anos ou mais, enquanto 7% têm menos de 50 anos (DESANTIS et al., 2016). No Brasil, ele é a principal causa de morte por câncer entre as mulheres, sendo estimado 59.000 casos novos de câncer em 2018, com uma incidência de cerca de 59.70 casos de cada 100 mil mulheres, segundo o (INCA, 2018). Em 2016, mais de 3,5 milhões de mulheres, havia tido ou estavam com câncer nos EUA (AMERICAN CANCER SOCIETY, 2017). Mais da metade das mortes ocorreram em regiões menos desenvolvidas, e estima-se que essas proporções irão aumentar até 2025 (WORLD HEALTH ORGANIZATION, 2013).

A baixa sobrevivência em países menos desenvolvidos está relacionada à falta de programas de detecção precoce (HARFORD, 2011). Isto ocorre porque a detecção precoce tem um impacto importante no sucesso do tratamento do tumor, sendo o tratamento mais difícil em estágios avançados (AUTIER et al., 2011). Vários métodos de prevenção têm sido desenvolvidos, mas um dos métodos mais eficazes para a detecção de lesão é a mamografia digital (TORTORA et al., 2012). No entanto, a interpretação da mamografia pode ser uma tarefa difícil até mesmo para um especialista, uma vez que a análise é afetada por diversos fatores, tais como a qualidade da imagem, experiência do radiologista e tipo de lesão. Foi estimado que

12% a 30% dos casos de câncer de mama deixam de ser detectados devido à má interpretação da mamografia (EVANS et al., 2013).

Em geral, regiões de lesões não possuem um contorno bem definido (DHEEBA et al. 2014). Por isso, técnicas baseadas em semente (CORDEIRO et al., 2013; ZHENG et al., 2013; CIESIELSKI et al., 2013), em que o usuário rotula pixels iniciais, apresentam melhor qualidade na segmentação final, porém requerem maior conhecimento do especialista. Técnicas não supervisionadas, tais como as baseadas em Contorno Ativos (RAHMATI, 2012) e Limiar Multinível (CHAKRABORTY, 2012), apresentam boa segmentação para lesões com contorno bem definidos, porém demonstram dificuldades em definir contornos do tipo especulado ou com borda não definida. Métodos de segmentação supervisionados, tais como os baseados em sementes, onde o usuário rotula regiões iniciais com e sem lesão, têm obtido boa qualidade de segmentação quanto à forma dos objetos (WONGTHANAVASU et al., 2011; SAIDIN, 2013).

Existem distintas metodologias terapêuticas para o tratamento do câncer de mama, sendo alguns métodos mais invasivos que outros. No entanto, a definição do tratamento apropriado depende da complexidade e do tipo de câncer de mama (SENKUS *et al.*, 2015). Por isso, é importante a precisão na avaliação da condição do paciente, visto que a correta classificação do câncer pode determinar se um tratamento invasivo é de fato necessário, provendo maior qualidade de vida ao paciente. Um diagnóstico falso-positivo de mamografia é aquele que apresenta resultados anormais quando não existe câncer. Estes resultados demandam exames adicionais para verificar se a anormalidade é um câncer. As mulheres que recebem o resultado falso-positivo podem sofrer de uma significativa ansiedade durante o processo e outros tipos de estresse psicológico, além do tempo, valor investido e



possíveis desconfortos físicos provenientes de alguns exames (ACS, 2016) (WORLD HEALTH ORGANIZATION, 2014). Enquanto um diagnóstico falso-negativo apresenta resultado normal quando há presença de câncer, o que pode provocar atraso no tratamento, diminuindo as chances de recuperação (ACS, 2016). Tanto o diagnóstico falso-positivo quanto o falso-negativo ocorrem devido à má interpretação da mamografia.

Segundo a *American Câncer Society*, no período de 10 anos, metade das mulheres que fazem mamografia anual recebe um resultado falso-positivo; e uma a cada cinco mamografias não apresenta o câncer de mama existente (ACS, 2016). O alto volume de dados da mamografia, distração, tipo de tecido mamário, bordas não definidas e baixa qualidade da imagem são fatores que podem dificultar a interpretação dos especialistas e o diagnóstico pode ser efetuado incorretamente (DESHPANDE; RAJURKAR; MANTHALKAR, 2013). Desta forma, é importante ter uma ferramenta computacional para auxiliar o especialista, melhorando assim a precisão da avaliação das imagens de mamografia.

O *Random Walker (RW)* é um método de segmentação supervisionado, proposto para segmentação de imagem por Grady (GRADY, 2010) e usado por Zheng *et al.* (2016) para segmentar massas (lesão) em mamografia, encontrando regiões de tumor com uma taxa de acerto de 93,4% com teste feito em 15 imagens contendo lesão, onde esse resultado é o estado da arte do método. Cordeiro *et al.* modificaram o algoritmo *GrowCut* (VEZHNEVETS *et al.*, 2005), utilizando lógica *Fuzzy* para reduzir o esforço do especialista, obtendo resultados melhores quando comparado com o estado da arte. A inserção da lógica *Fuzzy* na etapa de inicialização do algoritmo de segmentação melhorou os resultados para o algoritmo *GrowCut* no trabalho de Cordeiro. No entanto, essa análise foi testada apenas no

algoritmo *GrowCut*. Este projeto tem como intuito aprofundar a utilização de lógica Fuzzy em outros algoritmos de segmentação e analisar o comportamento do modelo em diferentes algoritmos de segmentação, propondo novas abordagens para a detecção de lesões. Avaliar técnicas de geração automática de sementes com os algoritmos de Enxame de Partículas (PSO) e Evolução Diferencial (DE). Comparar funções de pertinência do estado da arte na etapa de inicialização do algoritmo de segmentação em relação ao método proposto *Fuzzy*.

## **1.1 Objetivos**

### **1.1.1 Geral**

O trabalho proposto tem como objetivo principal o estudo e desenvolvimento de uma técnica de segmentação de tumores em imagens de mamografia, utilizando lógica *Fuzzy*. Têm-se como objetivo inserir a abordagem *Fuzzy* em um modelo existente, a fim de propor uma nova solução para segmentação de tumores. Por fim, tem-se como objetivo realizar comparação com técnicas do estado da arte.

### **1.1.2 Específicos**

- Desenvolvimento de um algoritmo de segmentação de tumores que seja menos dependente do posicionamento correto das sementes iniciais do que as técnicas do estado da arte;
- Entender a abordagem de modelagem *Fuzzy* Gaussiana para diferentes algoritmos de segmentação de tumores;

- Desenvolvimento de um método de segmentação, que a partir da região de uma imagem de mamografia realize a segmentação da lesão de forma automática;
- Redução da quantidade de marcações iniciais necessárias do algoritmo de segmentação, a fim de auxiliar o especialista e diminuir o impacto da marcação inicial;

Análise e validação dos resultados obtidos de forma qualitativa e quantitativa, utilizando métricas inerentes à forma da lesão segmentada, comparando os resultados obtidos do método proposto com os resultados das técnicas implementadas do estado da arte.

## **1.2 Estrutura do Trabalho**

Os capítulos seguintes abordam conceitos e a metodologia utilizada neste trabalho. O Capítulo 2 aborda os conceitos básicos importantes para compreensão desse trabalho. O Capítulo 3 apresenta o estudo da arte. Os Capítulos 4, 5 e 6 referem-se respectivamente a metodologia utilizada, resultados e conclusão obtida.

## 2. FUNDAMENTAÇÃO TEÓRICA

A computação visual busca replicar os efeitos da visão humana, em perceber e entender uma imagem. Isso não é uma simples tarefa, tendo em vista que o mundo é 3D e a maioria das ferramentas de imagem fornecem apenas uma imagem 2D (SONKA; HLAVAC; BOYLE, 2014). Neste capítulo é abordado as etapas e algoritmos utilizados na computação visual para identificação de imagens.

Uma imagem digital é definida pela integração e amostragem de dados contínuos (amostra do sinal analógico da imagem real) em um domínio espacial. A imagem consiste de um *array* retangular de pixels  $(x; y; v)$ , que é a combinação da localização  $(x; y)$  com o valor de cada amostra  $(v)$  (KLETTE, 2014).

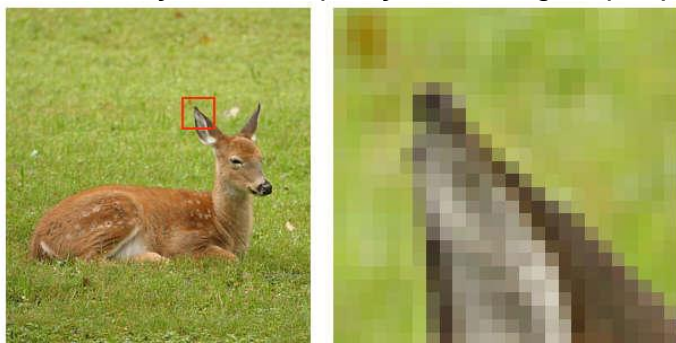
O entendimento de uma imagem por uma máquina pode ser compreendido como a relação entre uma imagem de entrada, e o modelo previamente estabelecido na observação do mundo real. Essa transição da imagem para o modelo, reduz as informações contidas na imagem para produzir informações relevantes. Este processo é geralmente dividido em vários passos e diferentes níveis de representação da imagem, cujo primeiro nível possui a imagem original e o último a imagem interpretada. A computação visual cria algoritmos para definir a relação entre estes níveis (SONKA; HLAVAC; BOYLE, 2014).

### 2.1 Conceitos Básicos de Imagem

Uma imagem pode ser definida como uma função bidimensional,  $f(x, y)$ , onde  $x$  e  $y$  são coordenadas espaciais, e o valor de  $f$  em qualquer par de coordenadas  $(x, y)$  é chamada de intensidade, ou nível de cinza em uma imagem monocromática, da

imagem neste ponto. Quando  $x$ ,  $y$  e os valores de intensidade de  $f$  são finitos e possuem quantidades discretas, chama-se a imagem de imagem digital. Os elementos que compõem a imagem digital, cada um com localização e valor específico, são chamados de pixels. Desta forma, o pixel é o menor elemento que compõe uma imagem digital, e ele possui um valor de intensidade. Cada pixel possui uma cor, representado em um ponto da imagem. A Figura 1 mostra como uma imagem digital é composta por pixels.

Figura 1: Ilustração da composição da imagem por *pixels*.



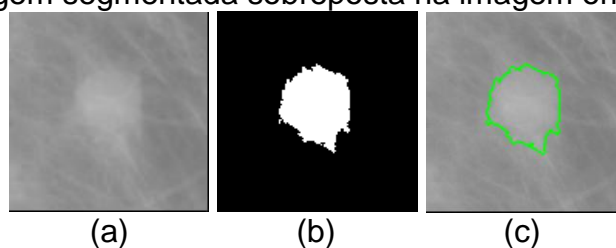
Fonte: (ARAÚJO, 2015).

Uma imagem em tons de cinza é uma imagem composta por pixels que possuem um valor de intensidade correspondente ao nível de cinza em um ponto específico da imagem. Esses tons de cinza variam entre branco e preto em 256 níveis diferentes. A imagem em tom de cinza difere da imagem binária, que é aquela que apresenta apenas duas intensidades correspondentes ao preto e ao branco. A imagem em tons de cinza pode ser representada como uma matriz de inteiros, com valores que variam de 0 a 255, onde o ponto  $(x, y)$  representa a localização espacial do ponto na cena, e o valor dessa coordenada na matriz representa a quantização, em inteiros, do ponto da imagem.

### 2.1.1 Segmentação

A segmentação é o processo de subdividir uma imagem em regiões ou objetos que a compõem. O nível de detalhe e região a ser segmentada depende do problema a ser resolvido. Em imagens de mamografia digital, por exemplo, o interesse do objeto a ser subdividido é a lesão. Uma vez separando o objeto de interesse do restante da imagem, fica mais fácil extrair suas características para posterior análise e diagnóstico. A Figura 2 mostra uma aplicação de segmentação para uma imagem de mamografia. A Figura 2 (a) mostra uma região com lesão, enquanto a Figura 2 (b) mostra a imagem segmentada. Para fins de visualização, muitas vezes é utilizada a representação da Figura 2 (c), onde é apresentado o contorno da imagem binarizada sobre a imagem original.

Figura 2: (a) Imagem Original. (b) Imagem Segmentada. (c) Contorno da imagem segmentada sobreposta na imagem original.



Fonte: O autor.

## 2.2. Conceitos de Inteligência Artificial

Nesta seção, serão descritos alguns fundamentos e algoritmos de inteligência artificial, explicando alguns termos que serão utilizados ao longo deste documento. Muitos algoritmos de processamento de imagem incorporam técnicas de inteligência artificial para fins de otimização ou para reconhecimento de padrões.

### 2.2.1 Random Walker

O *Random Walker (RW)* é um algoritmo para segmentação de imagens. A abordagem do *Random Walker* para segmentar imagens é determinar qual classe um *pixel*  $u$  pertence, começando em  $u$  e terminando ao *pixel* de rótulo determinado. Em seguida, calcula-se a probabilidade de cada rótulo ser alcançado por um passeio aleatório iniciado no *pixel*  $u$ . Os rótulos com maior probabilidade de serem alcançados são associados ao *pixel*  $u$ . A probabilidade de um passeio aleatório de um *pixel*  $u$  até o *pixel*  $j$  é calculado pela Equação 1:

$$p_{uj} = \frac{w_{uj}}{\sum_{k \in N(u)} w_{uk}}, \quad (1)$$

onde  $p_{uj}$  é a probabilidade de um *pixel* deixar  $u$  e chegar ao *pixel*  $j$ ,  $w_{uj}$  é o peso associado entre  $u$  e  $j$ ,  $N(u)$  é o vizinho de  $u$ . O peso de um *pixel* para outro é calculado pela Equação 2:

$$w_{ij} = e^{\frac{-(I_i - I_j)^2}{\sigma^2}}, \quad (2)$$

onde  $I_i$  e  $I_j$  são os valores de intensidade de *pixel*  $i$  e *pixel*  $j$  e  $\sigma$  ajusta o valor do peso. Quanto maior o sigma, mais difícil será a difusão da segmentação. Enquanto os pesos representam as diferenças entre os valores dos pixels, as caminhadas têm maiores probabilidades de correr através de regiões uniformes, evitando bordas. O *Random Walker* minimiza a energia da Equação 3:

$$\text{Energia} = \frac{1}{2} \sum_{(i,j) \in \epsilon} w_{ij} (x_i - x_j)^2 \quad (3)$$

O algoritmo requer três parâmetros para ser otimizado: *sigma*, *mode*, e *tol*. O valor *sigma* regula o coeficiente de penalização para o movimento aleatório do *Random Walker*. O parâmetro *mode* indica o modo de resolução do sistema de equação do *Random Walker*. Usamos três configurações para este parâmetro: força bruta (*bf*), gradiente conjugado (*cg*) e gradiente conjugado com pré-condicionador *multigrid* (*cg\_mg*). A usar *bf*, a fatorização de Laplaciano é computado. Quando *cg* é usado, o sistema linear é resolvido iterativamente usando o método de gradiente conjugado, e usando o *cg\_mg*, um pré-condicionador é computado usando um solucionador *multigrid*, então a solução é calculada com gradiente conjugado. Por fim, os conjuntos de parâmetros *tol* define a tolerância a alcançar quando se resolve o sistema linear do *Random Walker*. O valor de tolerância é aplicado somente quando o sistema linear é resolvido usando os modos *cg* e *cg\_mg*.

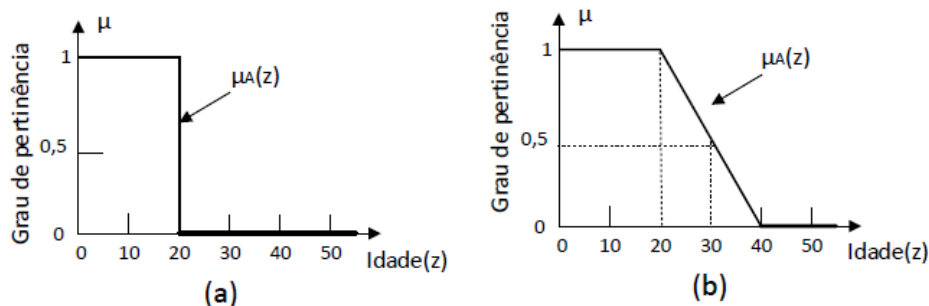
### 2.2.2 Lógica Fuzzy

Segundo Rajasekaran e Pai (2011), a lógica *fuzzy*, também chamada de lógica difusa ou nebulosa, é uma extensão da lógica booleana que permite valores de pertinência intermediários entre 0 e 1.

A teoria clássica dos conjuntos permite o tratamento de classes de objetos e suas inter-relações em um universo definido. Nessa teoria, a pertinência de um dado elemento a um conjunto refere-se ao fato de tal elemento pertencer ou não a esse conjunto. Na teoria clássica dos conjuntos, a pertinência só pode ser verdadeira ou falsa no sentido tradicional da lógica booleana de dois valores, com 1 normalmente indicando verdadeiro e 0 indicando falso. A Figura 3 (a) resume esse conceito, como um exemplo de alto nível, utilizando um limiar de idade de 20 e com  $\mu_A(z)$  expressando a função de pertinência no conjunto dos adultos.



Figura 3 – Funções de pertinência do ponto de vista da (a) lógica convencional e (b) da lógica *fuzzy*.



Fonte:(GONZALEZ; WOODS; EDDINS, 2009).

Na Figura 3 (a), que representa a lógica booleana, o grau de pertinência varia apenas em 0 ou 1, separando as pessoas com idade maior que 20 anos e menor ou igual em duas classes. Já na Figura 3 (b), que representa a lógica *fuzzy*, o grau de pertinência assume valores entre 0 e 1 para as idades entre 20 e 40 anos, o que mostra que nessa região não é assumido que o elemento pertence completamente ao conjunto, mas sim que há um grau de pertinência que pode variar nesse intervalo.

Na lógica *fuzzy* os elementos têm graus de pertinência aos conjuntos *fuzzy* definidos, que variam de 0 a 1. Ao contrário da lógica convencional, a lógica *fuzzy* utiliza a idéia de que elementos de um universo admitem graus de pertinência a um dado conjunto. A Figura 3 (b) mostra os níveis de pertinência em relação à função de pertinência, que varia gradativamente.

### 2.3. Algoritmos de Otimização

Algoritmos de otimização foram desenvolvidos para resolver problemas caracterizados pela dificuldade de encontrar soluções ótimas em um grande espaço de busca: chamados de problemas de otimização. Nesse tipo de problema é inviável testar todos os parâmetros possível devido ao grande espaço de exploração. As

características a serem otimizadas são chamadas de objetivos, e muitas vezes esses objetivos competem entre si.

### 2.3.1 Evolução Diferencial

O algoritmo de Evolução Diferencial (*Diferencia Evolution*, DE) é um algoritmo evolucionário que evolui com operadores genéticos. O algoritmo Evolução Diferencial se baseia em uma população inicial, composta por indivíduos, que são as soluções do problema. Estes indivíduos sofrem alterações ao longo das iterações, a fim de se obter o melhor conjunto de soluções.

#### Pseudocódigo 1 Evolução Diferencial

```

1: Gere um população inicial  $x^{i,0}$ ,  $i=1, \dots, P$  e avalie aptidão
2: Seja  $F$  um valor real entre 0 e 1
3: for  $g = 1, \dots, MAXGEN$  do
4:   for  $k = 1, \dots, P$  do
5:     for  $i = 1, \dots, M$  do
6:       Selecione aleatoriamente  $r_1 \neq r_2 \neq r_3 \neq k \in \{1 \dots P\}$   $\{1 \dots P\}$   $r_1 \neq r_2 \neq r_3 \neq k \in \{1 \dots P\}$ 
7:        $r_{nbr} \leftarrow randint(1, n)$ 
8:       for  $j = 1, \dots, n$  do
9:         if  $rand(0,1) \leq CR$  ou  $r_{nbr}$  then
10:            $filhos_j^{k,i} \leftarrow x_j^{r_3,g} + F * (x_j^{r_1,g} - x_j^{r_2,g})$ 
11:         else
12:            $filhos_j^{k,i} \leftarrow x_j^{k,g}$ 
13:         end if
14:       end for
15:     end for
16:     Seja  $u^k$  a melhor solução dentre  $x^{k,g}$  e  $filhos^{k,i}$ ,  $i = 1, \dots, M$ 
17:      $x^{k,g+1} \leftarrow u^k$ 
18:   end for
19: end for

```

O funcionamento do algoritmo Evolução Diferencial é mostrado no Pseudocódigo 1. No Pseudocódigo 1, a variável  $M$  representa o número de soluções novas por geração,  $CR$  é a taxa de cruzamento,  $MAXGEN$  é o número máximo de gerações e  $n$  é o número de elementos do vetor que representa a solução. De acordo com o Pseudocódigo 1, a geração de novas soluções é feita através de um operador de mutação, que realiza combinação linear entre três soluções da população, um operador de mutação, que realiza combinação linear entre três soluções. De acordo com o Pseudocódigo 1, a geração de novas soluções é feita através de um operador de mutação, que realiza combinação linear entre três soluções da população, um operador de cruzamento, que mistura coordenadas do setor gerado pelo operador de mutação. A última solução competirá com a nova solução gerada pela permanência na população. Se a nova solução preestabelecida. então a solução de cruzamento será substituída pela nova solução gerada. O critério de avaliação das soluções é baseado na função de aptidão, que é definida de acordo com o objetivo da ser otimizado por cada problema. O processo é repetido até atingir o número máximo de gerações.

### 2.3.2 Enxame de Partículas

O algoritmo Enxame de Partículas (PSO), também é um algoritmo de meta de aprendizado e se diferencia do DE por ser um algoritmo inspirado em um comportamento inteligente realizado por pássaros por busca de alimentos, o Pseudocódigo 2 mostra o algoritmo PSO.

---

#### **Pseudocódigo 2** Otimização por Enxame de Partícula

1:InicializarEnxame():

2:for interacao = 1, ...,  $MAXGEN$  do

3:for cadaParticula do

---

---

```
4:for Dimensao = 1, ..., N do
5:atualizeVelocidadePosicao();
6:end for
7:   Atualize líder;
8:   if fitnessAtual < melhorFitnessIndividual then
9:     melhorFitnessIndividual = fitnessAtual;

10:  end if
11:  if fitnessAtual < melhorFitnessVizinhanca then
12:    melhorFitnessVizinhanca = fitnessAtual;
13  end if
14: end for
```

---

O enxame é inicializado com os parâmetros *MAXGEN*, *N*, *Wk*, *C1* e *C2* onde *MAXGEN* é o número máximo de interações, *N* é o tamanho da dimensão, *Wk* é a constante de inércia, *C1* e *C2* cognitiva e social parâmetro respectivamente. Valores dos vetores de velocidade e posição gerados aleatoriamente. A primeira iteração do algoritmo inicia com a atribuição de valores aos parâmetros da equação de velocidade. Definem-se então os valores referentes ao enxame, constantes e o critério de parada. Tendo já definido os valores para posição das partículas e suas respectivas velocidades, aplica-se o cálculo do fitness a cada partícula desta população. O *fitness* avalia o desempenho da partícula. Com as partículas do enxame avaliadas, extrai o líder, isto é, a melhor posição encontrada pela partícula e pelo exame. Depois as velocidades e as posições de cada partícula do enxame são atualizadas. Diante das novas posições, caso o critério de parada tenha sido atingido, a solução do problema encontrada é apresentada. Caso contrário, aplica-se novamente o *fitness* a este enxame, atualiza-se o líder, caso seja apresentada uma solução melhor, seguido da velocidade e posição de cada partícula do enxame. O laço prossegue até o critério de para ter sido atingido.

## 2.4 Soluções Numéricas

Os algoritmos de inteligência artificial, são baseados em cálculos matemáticos, nessa seção são descritos alguns conceitos de algoritmos para resolução de problemas matemáticos no algoritmo *Random Walker*.

### 2.4.1 Gradiente Conjugado

O método do gradiente conjugado foi originalmente proposto por Hestenes *et al* (1952). O método do gradiente conjugado é um algoritmo para a solução numérica de sistemas particulares de equação lineares, aquelas cuja matriz é simétrica e positiva definida. O método do gradiente conjugado é um método iterativo, então ele pode ser aplicado a sistemas esparsos que são grandes demais para ser tratados por métodos diretos. Tais sistemas surgem frequentemente quando se resolve numericamente equações diferenciais parciais.

#### 2.4.1.1 Pré- Condicionador *Multgrid*

Devido ao mal-condicionamento da matriz nas iterações finais do método é necessário o uso de pré-condicionador. O Método *Multgrid* em análise numérica são um grupo de algoritmos para solução de equação diferenciais usando hierarquia de discretizações. A idéia é similar à extrapolação entre malhas mais grossas e mais finas. A aplicação típica para o *multigrid* é na solução de equações diferenciais parciais elípticas em duas ou mais direções. O *multigrid* pode ser aplicado junto com qualquer técnica comum de discretização Ele não depende da separabilidade das equações ou de outras propriedades da equação. O *multigrid* pode ser generalizado numa variedade de formas diferentes e pode ser aplicado diretamente para equações diferenciais parciais dependente do tempo.

### 3. REVISÃO BIBLIOGRÁFICA

Muitas pesquisas têm sido feitas, a fim de melhorar a segmentação de lesão em imagens de mamografia digital. A segmentação baseada na região do tumor é considerada mais adequada para detecção de massa, uma vez que a região do tumor é mais brilhante do que os tecidos vizinhos, tendo uma uniformidade de densidade e limites difusos (SUCKLING et al., 1994).

Recentes estudos feitos em segmentação de imagens de mamografia, têm sido bem sucedidos de segmentação da lesão de mama, Lewis *et. al* (2012) empregam uma técnica baseada em *Watersheds* para segmentar automaticamente regiões candidatas de tumor, este trabalho adotou uma abordagem para limpar imagens para determinar marcadores de primeiro plano e de fundo antes de usar segmentação *Watersheds* que traz resultados mais confiáveis. O experimento com a base de dados MIAS mostrou uma taxa de detecção de 90% para massas de tumores. O diagnóstico por Assistência por Computador (CAD) foi aprovado para automatizar a detecção de câncer de mama com mamografias em 1998. Mas devido à grande variabilidade de tamanho do tumor, forma e tecido mamário, é difícil para algoritmos de reconhecimento de padrões a adaptação a diferentes situações.

Eltoukhy e Faye (2013) usam uma técnica de limiar adaptativo para executar de forma automática a detecção por segmentação de câncer de mama, atingindo 100% de sensibilidade, com média de 1,97% de falsos positivos, quando aplicado a 188 imagens. Contudo, o valor da sensibilidade varia dependendo da taxa de falso positivo, e cada trabalho usa uma taxa diferente. A interpretação mamográfica é uma tarefa difícil que requer muito treinamento e experiência para os médicos.

Consequentemente, há fortes necessidades para o desenvolvimento de sistemas de diagnóstico assistido por computador (CAD).

Hong (2010) propõe uma abordagem topográfica de segmentação, baseado no fato das regiões suspeitas serem mais brilhantes que as regiões vizinhas, com densidades uniformes. Contudo, na maioria dos casos, as áreas de lesão não têm contornos bem definidos. Devido a este fato, as técnicas baseadas em sementes, i.e, técnicas em que os usuários rotulam as sementes iniciais, alcançam maior qualidade no final da segmentação.

Cordeiro et. al (2013) aplicam o *GrowCut* clássico para segmentar massas em mamografia. Posteriormente, Cordeiro et. al (2016) propõem uma abordagem *Fuzzy* do *GrowCut*, com a função de associação *fuzzy*, foi conseguido lidar com a complexidade de tumores com bordas não definidas. A técnica proposta faz com que não seja necessário a marcação de pontos externos a lesão, só sendo necessário a marcação de pontos internos a lesão. A técnica usada é baseada em uma modificação no algoritmo semi supervisionado *GrowCut* usando funções de associação *Fuzzy* Gaussianas, os resultados mostraram que a abordagem proposta teve melhores resultados em média, obtendo uma taxa de 85,83% de classificação, mas os resultados comparados são apenas comparados com métodos não supervisionados.

Zen et. al (2013) usam uma segmentação baseada em caminhos aleatórios, obtendo bons resultados de segmentação, mas ele não fornece análise quantitativa dos resultados dos experimentos. Apesar das técnicas baseadas em sementes demonstrarem uma boa segmentação de massa, a técnica requer um elevado nível de conhecimento do especialista sobre a seleção de sementes.

Cordeiro et. al (2018) propõe uma nova função de pertinência *chamada Multi-Gaussian* e utiliza na técnica de segmentação *Fuzzy GrowCut* melhorando os resultados em relação as funções de pertinência do estado da arte.

Chakraborty et. al (2012) segmenta com um limiar multinível com um abordagem controlada por gradiente de intensidade, onde uma imagem é considerada como um mapa 3D topográfico, sendo a terceira dimensão a intensidade. O método proposto foi testado na base de dados MINI-MIAS e apresentou uma sensibilidade de 98,8%.

Rahmati et. al (2012) apresenta uma abordagem supervisionada para segmentar lesões suspeitas em mamografia digital. Com base em um novo modelo de contorno ativo de semelhança usando conjuntos de níveis, o algoritmo estima o contorno de segmentação que melhor separa a lesão do fundo usando a distribuição *gamma* para modelar a intensidade de ambas as regiões. Apresentando uma taxa de precisão de segmentação de 86,85.

Conforme observado por Raman et. al (2011), os resultados obtidos com os trabalhos relacionados no estado da arte são significativamente diferentes. Muitas vezes baseiam-se em opiniões subjetivas com pouca análise quantitativa. Além disso, a maioria dos estudos descrevem uma precisão das técnicas baseadas apenas na localização do tumor e não na sua forma e contorno, que é um aspecto muito importante para diagnósticos precisos. A maioria dos trabalhos discutem a qualidade da segmentação com base na precisão de sobreposição da área medida e não fornecem uma análise qualitativa da segmentação em termos da forma da lesão. O trabalho proposto estará preocupado com a análise quantitativa e qualitativa dos resultados, onde essas análises são baseadas na forma e contorno da lesão, uma vez que a qualidade do contorno da segmentação é importante para



definir o tipo da lesão, é importante ter melhores algoritmos para otimizar o contorno final da segmentação.

## 4. MATERIAIS E MÉTODOS

### 4.1 Ambiente Experimental

O trabalho proposto utiliza a base de dados Mini-MIAS para realizar a avaliação das técnicas. A Sociedade de Análise de Imagem Mamográfica (MIAS) é uma organização de grupos de pesquisa do Reino Unido, interessados na compreensão de mamografias e gerou uma base de dados Mini-MIAS de mamografia digital, contendo a imagem da mamografia e também a região consideradas por radiologistas como anormais. A base consiste em 322 imagens de mamografia da visão médio lateral oblíqua, obtido de 161 pacientes. Entretanto, apenas 57 das imagens contem lesão. As outras imagens da base dados são classificadas como normal, sem lesão ou contem calcificações ou assimetrias.

As imagens com segmentação com imagem ouro foram obtidas por um software auxiliar supervisionado, com base num limiar adaptativo, que utilizam indicações fornecidas da base de dados. Entre as imagens selecionadas, há diferentes tipos de massas: circunscritas, espiculadas e mal definidas. As regiões de interesse foram manualmente selecionadas para executar a segmentação da técnica proposta.

Para executar o algoritmo *Random Walker* clássico, nós usamos a implementação em *python*, disponível na biblioteca *scikit-image*. As configurações do computador de execução foram ambiente com sistema operacional Ubuntu 16.04, 8 GB de memória (RAM), processados Intel Core i5-3337U, CPU 1.80GHz.

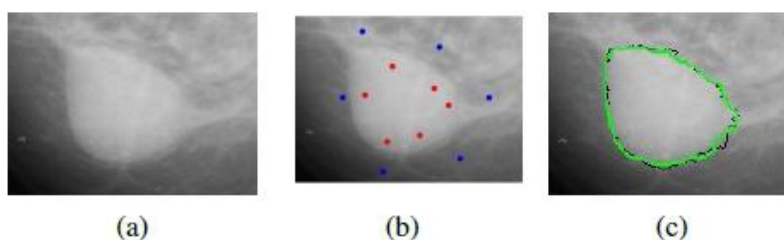
## 4.2 Proposta da Pesquisa

### 4.2.1 Segmentação de Imagem

O algoritmo de segmentação *Random Walker* (RW), que é um método de interativos de segmentação, originalmente proposto por Grady (2016). O RW é um método baseado em grafo onde cada *pixel* é representado como um nó, o qual é conectado ao seu vizinho por arestas.

O algoritmo Random Walker usa o conceito de pixels de semente: o usuário rotula inicialmente um conjunto de pixels de diferentes classes, e tendo em conta os níveis de cinza destas sementes, o algoritmo tenta rotular todos os pixels da imagem. O algoritmo Random Walker recebe como entrada a imagem a ser segmentada e a matriz com os rótulos fornecidos pelo usuário. O rótulo fornecido pelo usuário pode ser de dois tipos: rótulo de fundo ou rótulo de objeto, que corresponde às regiões situadas no exterior e no interior da lesão. Inicialmente, os pixels não estão associados a rótulos até o usuário determinar os rótulos de objeto e de fundo dos pixels. Portanto, é necessário um conhecimento inicial e esforço para realizar a segmentação. A Figura 4 mostra cada estágio da segmentação RW.

Figura 4: Processo de segmentação *Random Walker*. (a) Imagem Original; (b) Sementes Iniciais; (c) Imagem Segmentada.



Fonte - O autor

A Figura 4a mostra como uma imagem de parte da mamografia contendo a lesão. A Figura 4b mostra a marcação inicial executada pelo usuário. Embora o usuário não saiba o contorno exato, ele deve selecionar as sementes corretamente, escolhendo *pixels* que representam áreas não lesionadas e *pixels* dentro da lesão. Esta fase pode ser uma tarefa desafiadora para lesões com bordas mal definidas. Finalmente, a Figura 4c mostra a segmentação.

Como a maioria das técnicas baseadas em sementes, a qualidade da segmentação depende diretamente das posições das sementes iniciais. Portanto, depende do conhecimento do usuário para selecionar as sementes adequadamente próximo da borda do objeto a ser segmentado.

A nossa abordagem proposta baseia-se numa modificação na inicialização do algoritmo *Random Walker*, através do uso da lógica *fuzzy*. Nomeamos nosso método proposto como *Fuzzy Random Walker* (FRW). Com a nossa proposta de modificação do algoritmo *Random Walker*, pretendemos reduzir a necessidade de conhecimento inicial especializado sobre o contorno do objeto de interesse, reduzindo o esforço na seleção de sementes. Ao contrário do RW clássico, a nossa proposta baseia-se na seleção de sementes de apenas uma classe: o objeto de interesse. Descartamos a seleção da classe de fundo, porque, a partir das sementes do objeto de interesse, podemos estimar a região de fronteira separando o objeto e o fundo. Na abordagem proposta, a localização inicial das sementes é suficiente para uma função *fuzzy*-gaussiana. Diferente de conjuntos clássicos, os elementos de conjuntos *fuzzy* têm graus para esse conjunto. O grau de associação de um conjunto *fuzzy* indica a certeza de que o elemento pertence a esse conjunto (DUBOIS et al. 2014).

Na abordagem proposta, o *pixel* será rotulado como fundo se a pertinência de um determinado *pixel* para o fundo é maior que a pertinência do mesmo *pixel* para o objeto. Caso contrário, os *pixels* permanecerão como se não tivesse sido selecionado pelo usuário, ou como objeto, se já tivesse sido selecionado. A função de associação ao fundo e ao primeiro plano são definidos pela Equação 4 e Equação 5:

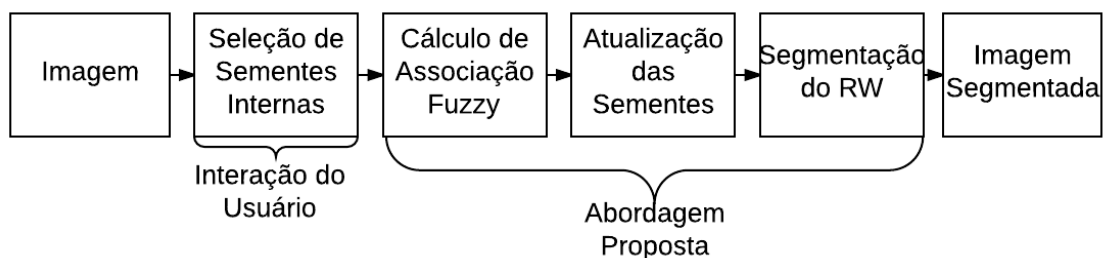
$$\mu_{Obj}(i) = \exp\left(-\frac{(x_i - x_m)^2}{2\alpha_x s_x^2}\right) \exp\left(-\frac{(y_i - y_m)^2}{2\alpha_y s_y^2}\right) \quad (4)$$

$$\mu_{Bkg}(i) = 1 - \mu_{Obj}(i), \quad (5)$$

onde  $\mu_{Bkg}(i)$  é o grau de pertinência *fuzzy* associado a incerteza que o *i*-ésimo *pixel* pertence ao fundo da imagem, enquanto que  $\mu_{Obj}(i)$  é o grau de associação à incerteza do *i*-ésimo *pixel* pertencente ao objeto de interesse. Estas funções de pertinência *fuzzy* são funções Gaussianas cujas variáveis  $x_i$  e  $y_i$  corresponde às coordenadas do *i*-ésimo *pixel* na imagem, enquanto que  $\alpha_x$  e  $\alpha_y$  são coordenadas do centro de massa das sementes inicialmente selecionadas.  $s_x$  e  $s_y$  são obtidos calculando o desvio-padrão da posição dos pontos selecionados pelo usuários. O rótulo de cada *pixel*  $q$ ,  $l_q$ , é atualizado de acordo com a expressão da Equação 6:

$$l_q = \begin{cases} l_{bkg}, & \mu_{Bkg}(q) > \mu_{Obj} \\ l_q, & \mu_{Bkg} \leq \mu_{Obj}(q) \end{cases} \quad (6)$$

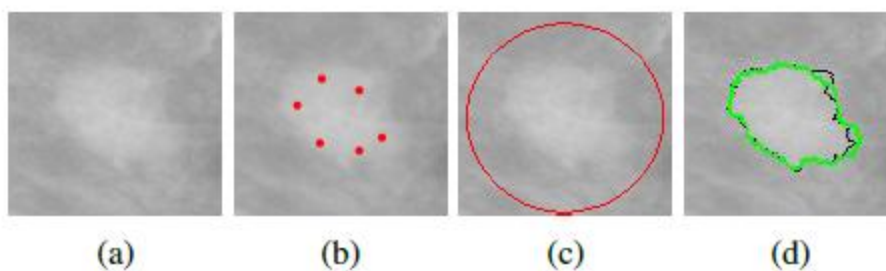
em que  $l_{bkg}$  representa o rótulo de fundo. De acordo com Equação 6, se  $\mu_{Bkg}(q) \leq \mu_{Obj}(q)$  o rótulo continua com sua rotulação inicial, isto é, indefinida ou objeto, se não selecionada ou selecionada pelo usuário, respectivamente. A inicialização proposta será utilizada como entrada no algoritmo RW. O fluxograma da abordagem proposta é mostrado na Figura 5:

Figura 5: Fluxograma do *Fuzzy Random Walker*

Fonte – O autor

Conforme mostrado na Figura 5, o método proposto começa com uma imagem, ou ROI, que é necessário que o usuário selecione sementes internas. Para o processo, destacamos que apenas as sementes internas são necessárias. Estes pontos de sementes serão rotulados como lesão. Em seguida, a associação *fuzzy* é calculada baseada nas sementes selecionadas pelo usuário. Baseado na associação *fuzzy*, as sementes são atualizadas, onde as sementes de fundo serão automaticamente definidas. Em seguida, as sementes atualizadas serão usadas como entrada para o algoritmo RW, gerando a imagem de segmentação final. A Figura 3 mostra cada passo do processo de segmentação do FRW. A Figura 6a mostra a parte da mamografia original, que é submetido à rotulagem inicial pelo especialista. No FRW, só é necessário selecionar sementes dentro da região, como mostrado na Figura 6b, reduzindo o esforço comparado ao RW clássico. A Figura 3c mostra a região gaussiana estimado pela a função de associação. A área externa é considerada como fundo, e os rótulos de fundo são automaticamente definidas. O tamanho da área pode ser ajustada pelo parâmetro  $\alpha$  da Equação 2. Após a rotulagem ser inicializada, a segmentação começa. A Figura 6(d) mostra o resultado da segmentação, em verde, em relação ao contorno de referência, em preto.

Figura 6: Estágios do *Fuzzy Random Walker*. (a) Imagem original; (b) Seleção de Semente; (c) Região *Fuzzy Gaussian*; (d) Imagem Segmentada.



Fonte - O autor

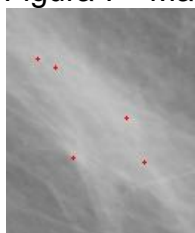
#### 4.2.2 Geração Automática de Sementes

Um modelo de seleção automática de sementes pode ser incorporado ao algoritmo proposto para realizar a etapa de inicialização de sementes de forma automática. No processo de seleção automática de sementes, deseja-se que as sementes estejam localizadas nas regiões mais claras da imagem, pois tendem a ser regiões onde há lesão. Desta forma, observamos que o problema de localização das sementes geradas é um problema de otimização, onde se deseja otimizar o valor de intensidade onde estão localizadas as sementes. Neste trabalho, foi desenvolvido um método que utiliza técnicas de otimização para encontrar sementes iniciais. Para realizar a etapa de seleção automática de sementes foi utilizado o algoritmo de otimização Evolução Diferencial. Esse algoritmo foi escolhido por ser uma técnica de otimização simples, rápida e já validada na literatura. No entanto, outros algoritmos de otimização, tal como *Simulated Annealing*, descrito por Dowsland e Thompson (2012), também poderiam ser utilizados.

Para utilizar o algoritmo Evolução Diferencial no problema de geração automática de sementes, inicialmente foi necessário realizar o mapeamento da

solução para o problema. Para o mapeamento do problema, cada solução é composta por um conjunto de sementes na imagem, sendo cada semente composta por dois atributos: posição no eixo  $x$  e posição no eixo  $y$ . A partir da localização da semente é possível estimar a intensidade do *pixel* correspondente à sua posição. A Figura 7 mostra o exemplo de mapeamento de um indivíduo.

Figura 7 - Mapeamento do indivíduo no modelo de geração automática de sementes.



x	y	x	Y	x	y	x	y	x	Y
120	110	90	105	120	130	90	98	112	140
Semente 1		Semente 2		Semente 3		Semente 4		Semente 5	

Fonte - O autor

Conforme mostrado na Figura 7, uma solução do processo de otimização é um vetor contendo a posição nos eixos  $x$  e  $y$  do conjunto de sementes. No algoritmo Evolução Diferencial define-se uma população de solução candidatas, onde buscam-se soluções que otimizem uma determinada função objetivo. Para realizar a otimização, foi definida uma função de otimização para avaliar a qualidade das soluções, a qual se baseia em maximizar o objetivo de intensidade dos *pixels* na localização das sementes. Desta forma, quanto maior a intensidade do *pixel* da localização das sementes, e quanto mais separadas elas estiverem entre si, melhor a solução.

O outro algoritmo implementado para seleção automática de sementes foi o Otimização por Enxame de Partículas (PSO), onde a estrutura do problema é a mesma explicada para o algoritmo de Evolução Diferencial.



Para avaliar os resultados foi usado a porcentagem de sementes selecionadas internamente ao tumor, baseado na imagem ouro fornecido pela base de dados.

#### **4.2.2.1 Experimentos de Geração Automática**

Na geração de sementes automática com algoritmo de Evolução Diferencial, foi experimentados alguns conjuntos de parâmetros a fim de se obter a melhor configuração de parâmetros para o problema de seleção automática de sementes. O algoritmo requer os seguintes parâmetros:  $N$ ,  $MAXGEN$ ,  $M$ ,  $CR$ ,  $F$ . Onde  $N$  representa o tamanho da população,  $MAXGEN$  o tamanho máximo de gerações, onde foi usado o valor fixo de 3000 como foi mostrado por Cordeiro *et. al* [16] sendo o melhor tamanho máximo de geração para o nosso problema,  $M$  representa o número de soluções novas por geração,  $CR$  é a taxa de cruzamento,  $F$  é o peso diferencial, deixamos esse valor fixo igual a 1 dado o tamanho do nosso problema.

Para cada parâmetro testado, foi tirado a média da porcentagem de sementes corretamente selecionadas entre as 57 imagens contendo tumor, também foi calculado o desvio padrão das médias. O processo foi repetido 10 vezes (pela inicialização ser aleatória e poder impactar no resultado final), e foi obtido a médias final dos resultados e a média dos desvios padrão de cada interação entre as 10.

### **4.3 Métricas de Avaliação**

A avaliação dos resultados é baseada nas seguintes métricas: sensibilidade, especificidade, *Jaccard* e *Dice*. Os escolhidos métodos foram definidos baseados nos valores de Verdadeiros Positivo (TP), Falso Negativo (FN), Verdadeiros Negativo (TN) e Falso Positivo (FP).

A sensibilidade (SE) avalia a taxa de *pixels* corretamente classificadas como lesão. Seja GT a imagem ouro, e S a imagem resultante do algoritmo de segmentação. A métrica SE é definida pela Equação 7:

$$SE = \frac{TN}{FP + TN} \quad (7)$$

A especificidade (SP) mede a taxa de *pixels* corretamente classificados como fundo e é definido pela equação 8:

$$SP = \frac{TP}{TP + FN} \quad (8)$$

A acurácia balanceada (BAC) calcula a média de especificidade e sensibilidade. O índice de Jaccard (J) indica quão similar a segmentação é quando comparada com a imagem ouro. O cálculo é dado pela Equação 9:

$$J = \frac{TP}{TP + FN + FP} \quad (9)$$

A métrica *Dice* é outra métrica comumente usada para avaliação de segmentação e é definida pela Equação 10:

$$Dice = 2 * \frac{TP}{S + GT} \quad (10)$$

#### 4.4 Funções de Pertinências

Após a seleção de sementes internas, é calculado a região de pertinência baseado em uma função de pertinência. Neste trabalho, implementamos as principais funções de pertinência do estado da arte para realização de segmentação com o *Random Walker*, a fim de poder comparar com os resultados da técnica proposta *Fuzzy Random Walker*. Foram implementadas as seguintes funções: Gaussiana, Triangular, Trapezoidal, Sino Generalizada e Multi-Gaussiana (Cordeiro

*et. al*, 2018). Os resultados obtidos também foram comparados com os experimentos feito por Cordeiro *et. al* (2018), com as mesmas funções de pertinências utilizando a técnica de segmentação GrowCut, usando as métricas de sensibilidade, especificidade, acurácia balanceada (BAC), *Jaccard* e *Dice*.

A função de pertinência Triangular é definida na Equação 11:

$$\mu_{tri}(x) = \max(\min(\frac{x-a}{b-a}, \frac{c-x}{c-b}), 0), \quad (11)$$

onde  $a$  e  $c$  representam a posição da base do triângulo, no eixo  $x$ , e  $b$  é a posição também no eixo  $x$ , no topo do triângulo. A função *max* e *min* representam o máximo e mínimo entre os valores, respectivamente. Nessa função, os parâmetros seguem a seguinte ordem:  $a < b < c$ .

A função de pertinência Trapezoidal é definida na Equação 12:

$$\mu_{trap}(x) = \max(\min(\frac{x-a}{b-a}, 1, \frac{d-x}{d-c}), 0), \quad (12)$$

onde  $a$  e  $d$  representam a posição da base mais larga do trapézio, no eixo  $x$ , e  $b$  e  $c$  são as posições também no eixo  $x$ , na base menor do trapézio. Nessa função, os parâmetros seguem a seguinte ordem:  $a < b \leq c < d$ .

A função de pertinência Gaussiana é definida na Equação 13:

$$\mu_{gauss}(x) = \varepsilon^{-\frac{(x-c)^2}{2\beta^2}}, \quad (13)$$

onde  $c$  e  $\beta$  são o centro e largura do conjunto difuso, respectivamente. O número  $x$  no centro da Região Gaussiana é o valor máximo de associação e os vizinhos diminuem gradativamente.

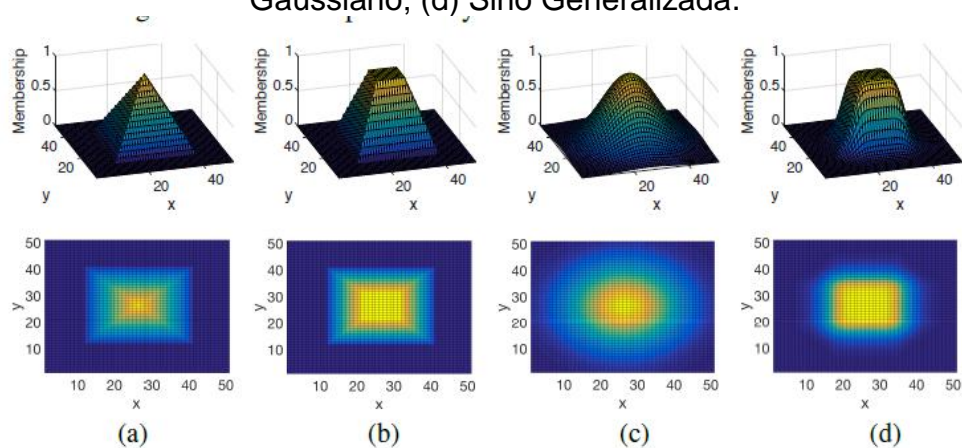
A função de pertinência Sino Generalizada é definida na Equação 14:

$$\mu_{ng}(x) = \frac{1}{1 + |\frac{x-c}{a}|^{2b}}, \quad (14)$$

onde  $c$ ,  $a$  e  $d$  são a largura, o centro e a suavidade da função Sino Generalizada respectivamente.

A Figura 8 mostra o gráfico das funções de pertinências implementadas. A primeira linha mostra cada função, em 3 dimensões, enquanto que a segunda mostra as regiões em 2 dimensões. A forma das regiões de associação influencia no mapeamento da região da lesão a partir da semente fornecida pelo usuário.

Figura 8: Funções de pertinências implementados: (a) Triangular; (b) Trapezoidal; (c) Gaussiano; (d) Sino Generalizada.



Fonte: (Cordeiro et. al, 2018)

## 5. RESULTADOS

### 5.1 Otimização de Parâmetros

Essa seção mostra os resultados obtidas com a segmentação *Fuzzy Random Walker* comparando com a implementação clássica, usado por Wang e Landay. A primeira análise feita com o objetivo de encontrar o melhor parâmetro do Random Walker clássico. A Tabela 1 mostra os valores dos parâmetro utilizados nesta análise:

Tabela 1: Espaço de Exploração de Parâmetros do *Random Walker*.

Parâmetros	Valores
<i>Sigma</i>	5000, 10000, 20000, 30000, 40000
<i>Mode</i>	bf, cg, cg_mg
<i>Tol</i>	0.001, 0.05, 0.1, 0.2

Fonte - O autor

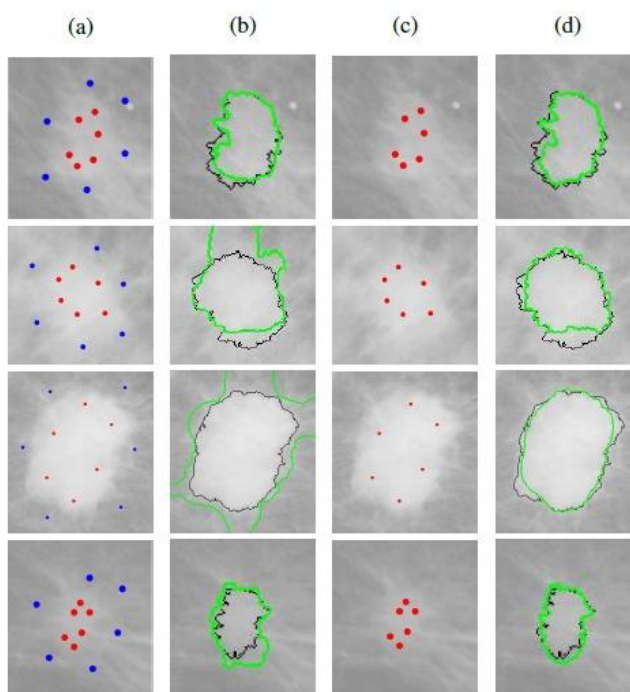
Os experimentos do ajuste de parâmetros mostraram que o *Random Walker* é muito sensível a correta seleção de parâmetros. Os parâmetros padrões fornecidos pela biblioteca *scikit-image* não foram suficientes para realizar uma segmentação precisa de massas em imagens de mamografia. No entanto, testando os parâmetros descritos na Tabela 1, pudemos de obter melhores resultados de segmentação. A melhor configuração encontrada para o RW foi obtida utilizando valores de parâmetros de  $\sigma = 10000$ ,  $mode = cg$  e  $tol = 0.02$ .

### 5.2 Resultados *Fuzzy Random Walker*

Na abordagem proposta há adição do parâmetro  $\alpha$ , que ajusta o tamanho da região gaussiana. Experimentamos o parâmetro  $\alpha$  com os valores 0.5, 1, 2 e 3, a fim

de obter o melhor valor de parâmetro para ser usado na segmentação da lesão com a abordagem proposta. A melhor configuração da abordagem proposta foi obtida utilizando os seguintes parâmetros:  $\alpha = 1$ ,  $\sigma = 30000$ ,  $mode = bf$ . Para o  $mode$  e  $bf$  não há parâmetro de tolerância. Os resultados da segmentação das imagens da base são mostrados na Figura 9. Nós usamos as imagens mdb010, mdb190, mdb132 e mdb312 como resultados de amostra da base de dados.

Figura 9: Comparação da segmentação do RW e FRW para as imagens mdb010, mdb021, mdb028 e mdb190. Coluna(a): RW sementes; Coluna(b): RW segmentação; Coluna(c): FRW sementes; Coluna(d): FRW segmentação.

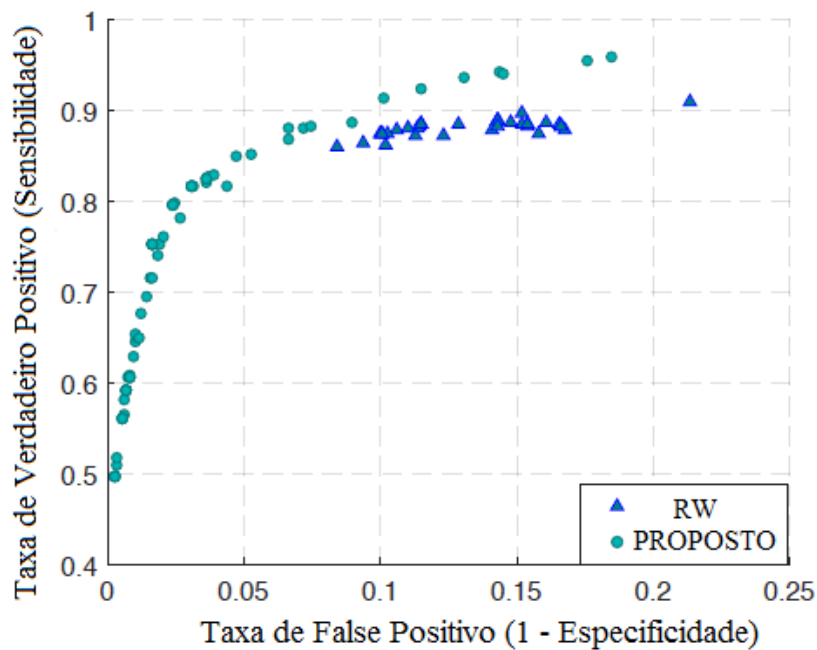


Fonte – O autor

Na Figura 9, a coluna (a) mostra os pontos de sementes usados para segmentação RW. Os pontos azuis são as sementes rotulados como fundo e os pontos em vermelho são rótulos de lesão. A coluna (b) mostra a segmentação obtida usando pontos da coluna (a). O contorno preto representa a segmentação verdadeira, enquanto que o contorno verde representa a segmentação do algoritmo. Como pode se observado, para as mesmas imagens, o algoritmo RW não funciona bem em alguns casos, faltando algumas áreas das bordas. A coluna (c) na Figura 5

mostra as sementes usadas pelo FRW. É importante notar que as sementes internas estão no mesmo local usado no RW clássico. Mas as sementes externas não são necessárias para executar a segmentação. O impacto disso é a redução de esforço e conhecimento necessário para do problema. Finalmente, a coluna (d) mostra a segmentação obtida usando a abordagem proposta. Para as imagens na Figura 9, nos podemos ver que mesmo sem fornecer as sementes de fundo, o FRW foi capaz de realizar uma segmentação próxima da verdade, com melhores resultados para as imagens que o RW teve dificuldades para segmentar. Analisamos a curva ROC de diferentes configurações utilizando o RW e a abordagem proposta. A Figura 10 mostra o comportamento dos parâmetros de configuração de cada técnica. Os resultados da Figura 10 mostram em verde e azul as diferentes configurações de RW e a abordagem proposta respectivamente. A abordagem proposta tem mais configurações porque tem um parâmetro adicional em comparação com o RW. Dando mais opções ao usuário para selecionar uma configuração com maior taxa de verdadeiro positivo ou menor falso positivo. Por outro lado, o processamento da configuração pode ser mais desafiador. No entanto, como pode se visto na Figura 6, as configurações do método proposto superam o RW clássico quando considerado os verdadeiros positivos e falsos positivos. Além disso, o FRW tem a vantagem de seleção de sementes, sendo necessária apenas a rotulagem de sementes internas. Para avaliar todas as métricas, escolhemos as configurações com melhores valores de *Jaccard* e *Dice*, que são as configurações selecionadas, analisamos todas as métricas juntas e calculamos a média dos resultados, para todas as imagens.

Figura 10: Curva ROC com parâmetros variados do RW clássico e da abordagem proposta.

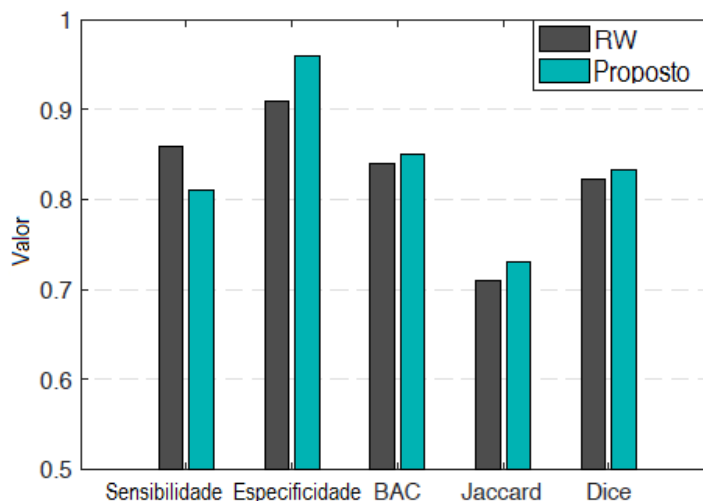


Fonte - O autor

Os resultados das métricas são mostrados na Figura 11. Os resultados demonstram que a abordagem proposta obteve melhores resultados para a maioria das métricas, em comparação com o RW clássico. Embora RW tenha obtido maior valor de sensibilidade, o método proposto atingiu um maior valor de BAC, que mede a média entre sensibilidade e especificidade.



Figura 11: Comparação do *Random Walker* e a abordagem proposta, para métricas de sensibilidade, especificidade, BAC, Jaccard e Dice.



Fonte – O autor

### 5.3 Resultados Geração Automática de Sementes

Na geração de sementes automática com algoritmo de Evolução Diferencial, foi experimentados alguns conjuntos de parâmetros a fim de se obter a melhor configuração de parâmetros para o problema de seleção automática de sementes. Os parâmetros testados são descritos na Tabela 2.

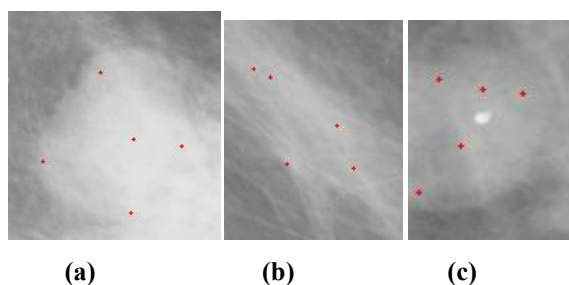
Tabela 2 - Espaço de Exploração de Parâmetros de Evolução Diferencial.

Parâmetros	Variação
Tamanho de População	[10, 20, 30]
CR	[0.5, 0.6, 0.9]
Soluções	[5, 10, 20]

Fonte - O autor

Observamos após experimentos que o conjunto mais adequado de parâmetros para o algoritmo de Evolução Diferencial foi  $F=1$ ,  $CR=0.6$ , Tamanho da População=30, Número de soluções por indivíduo=5. Algumas imagens demonstrativas, podem ser observadas na Figura 12.

Figura 12 -Imagens de resultados



Fonte - O autor

Também foi feita análise de parâmetros no algoritmo Otimização por Enxame de Partícula. O algoritmo requer os seguintes parâmetros: N, MAXGEN, C1 e C2. Onde N representa o tamanho da população, MAXGEN o tamanho máximo de geração, onde foi usado o valor fixo de 3000 como foi mostrado por Cordeiro *et. al* sendo o melhor tamanho máximo de geração para o nosso problema, C1 e C2 são parâmetros de cognitivo e social. Os parâmetros testados são descritos na Tabela 3.

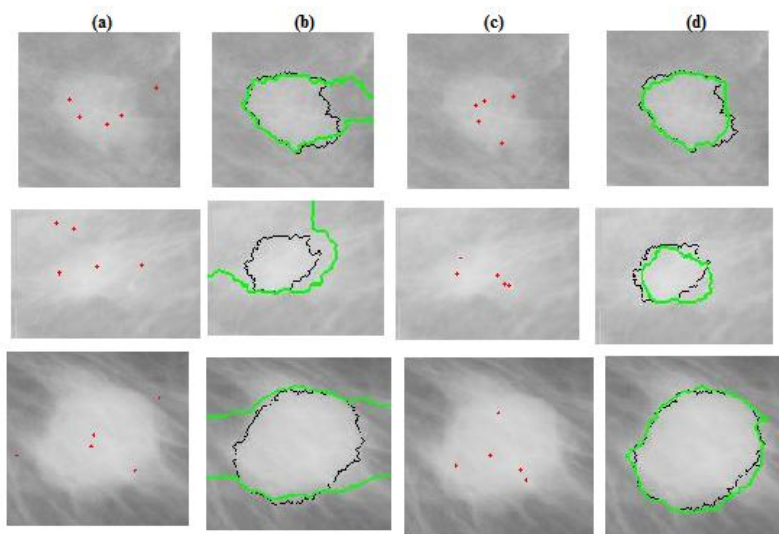
Tabela 3 - Espaço de Exploração de Parâmetros de Otimização por Enxame de Partículas.

Parâmetros	Varição
Tamanho de População	[10, 20, 30]
Soluções	[5, 10, 20]

Fonte - O autor

Foi observado que o melhor conjunto de parâmetro é População=30, Sementes=5. Com a taxa de acerto obtida foi de 65% na base Mini-MIA sem relação a localização interna ao tumor. Algumas imagens exemplo da segmentação feita a partir das sementes geradas com os algoritmos evolutivos podem ser observados na Figura 13.

Figura 13: Comparação da segmentação do FRW com sementes gerada por DE e PSO Coluna(a): DE sementes; Coluna(b): FRW segmentação; Coluna(c): PSO sementes; Coluna(d): FRW segmentação

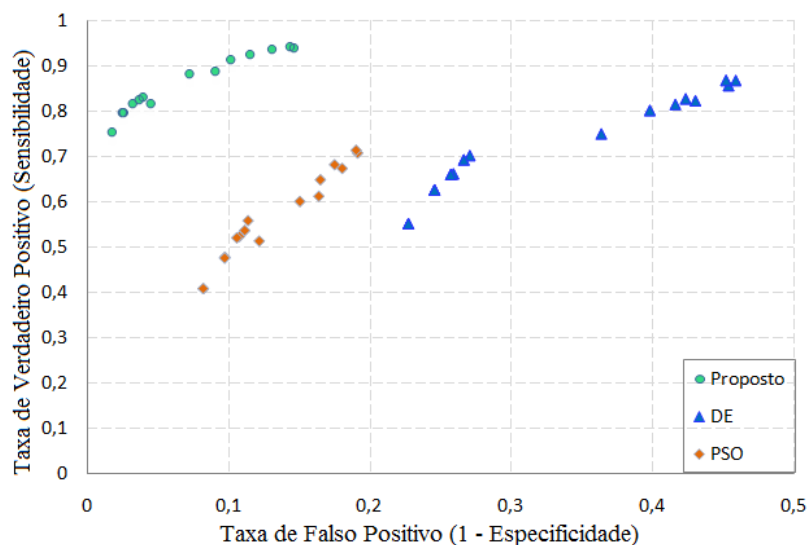


Fonte - O autor

Na Figura 13, a coluna (a) mostra os pontos de sementes gerados pelo algoritmo de Evolução Diferencial (DE). Os pontos vermelho são rótulos de lesão. A coluna (b) mostra a segmentação obtida usando pontos da coluna (a). O contorno preto representa a segmentação verdade, enquanto que o contorno verde representa a segmentação do algoritmo. A coluna (c) na Figura 13 mostra as sementes gerados pelo algoritmo de Otimização por Enxame de Partículas (PSO). É importante notar que as sementes internas estão no mesmo local usado no RW clássico.

Para avaliar o resultado da geração de sementes automáticas, foi feito uma comparação entre os resultados das métricas da segmentação feita com a abordagem proposta *Fuzzy Random Walker* com inicialização manual utilizando o algoritmo de Evolução Diferencial e o de Otimização por Enxame de Partículas, com o FRW com inicialização automática. Podemos observar essa comparação na Figura 14 contendo curva ROC da técnica proposta com inicialização automática e manual.

Figura 14 - Curva ROC com parâmetros variados do FRW com inicialização Manual e Automática DE e PSO



Fonte - O autor

Podemos observar a evolução dos resultados para a abordagem inicializada manualmente e automaticamente. É esperado que o resultados da segmentação com inicialização automática sejam piores dada a complexidade do problema. Entretanto podemos observar que é possível ter uma segmentação razoável, sem que seja necessário a entrada de rótulos fornecidos por um humano. E podemos observar pelos resultados que o algoritmo evolucionário PSO teve melhores resultados que o de Evolução Diferencial.

#### 5.4 Resultados Variação de Funções de Pertinência

Para analisar os resultados dos experimentos com variações de funções de pertinência foi feita uma comparação com os resultados de segmentação com as diferentes funções de pertinência, usando o algoritmo *Random Walker* e o *GrowCut*, os resultados estão descritos na Tabela 5.

Tabela 4: Comparações de resultados por variações de funções de pertinência nos algoritmos Random Walker e GrowCut.

Algoritmo	Função	SE	SP	J
<b>GrowCut</b>	Gaussiana	0.74	0.88	0.62
	Triangular	0.66	0.85	0.55
	Trapezoidal	0.66	0.85	0.55
	Sino Generalizado	0.35	<b>1.00</b>	0.34
	Mult-Fuzzy	0.72	0.98	0.63
<b>Random Walker</b>	Sino Generalizado	0.46	0.99	0.45
	Gaussiana	<b>0.80</b>	0.81	0.62
	Mult-Fuzzy	0.78	0.94	0.69
	Mult-Gaussian	0.72	0.96	0.66
	Trapezoidal	0.74	0.84	0.61
	Triangular	0.79	0.76	0.56
	Fuzzy	0.75	0.98	<b>0.70</b>

Fonte: O autor

Analisando os resultados da tabela acima podemos observar que o *Fuzzy Random Walker* apresentou os melhores resultado em relação as outras funções de pertinência e também em relação ao algoritmo *GrowCut* tendo o maior índice de Jaccard com valor de 0.70, sensibilidade 0.75 e especificidade 0.98.

## 5.5 Discussão dos Resultados

Com tudo isso, podemos observar que o *Random Walker* é muito sensível a seleção de parâmetros, e no problema de segmentação de massas em mamografia os melhores valores de parâmetros foram  $\sigma = 10000$ ,  $mode = cg$  e  $tol = 0.02$ . Os resultados também mostraram que o método proposto FRW obteve melhores resultados em comparação com o RW clássico na maioria das métricas.

Na abordagem da seleção automática de sementes percebemos que o PSO teve melhores resultados em relação a o algoritmo de Evolução Diferencial. Embora, os resultados sejam piores em relação a seleção manual de sementes

dada a complexidade do problema, mas os resultados foram razoáveis. Nos experimentos da função de pertinência o método *Fuzzy* teve os melhores resultados em relação a outras funções de pertinência do estado da arte, tendo maior índice de Jaccard com o valor de 0.70.

É importante observar, que a abordagem proposta funciona muito bem, com imagens com a região do tumor com borda bem definidas, já em imagens mais complexas com lesão sem bordas definidas ou formato espiculado, o algoritmo não segmenta muito bem a região do tumor.

## 6. CONCLUSÃO

Este trabalho propôs um novo método de segmentação baseado no Algoritmo *Random Walker*, aplicado à segmentação de massas em imagens de mamografia. O método proposto, chamado *Fuzzy Random Walker*, usa uma associação fuzzy-gaussiana para atualizar a rotulagem inicial. A avaliação foi realizada com 57 imagens da base de dados do Mini-MIAS, comparado com as métricas de sensibilidade, especificidade, precisão e índice de Jaccard. Os resultados mostraram que a abordagem proposta exige menos esforço do usuário, em comparação com o tradicional RW. Algo inconveniente é que o método proposto tem a adição de um parâmetro que controla a região gaussiana.

Quando comparado com o *Random Walker*, a abordagem proposta obteve valores elevados de precisão e índice de Jaccard. O método proposto mostrou-se uma solução interessante quando considerado as taxas de verdadeiros positivos e taxas de falsos positivos. A técnica proposta obteve resultados melhores que o *Random Walker* clássico e o estado da arte que era o *Fuzzy Grow Cut*. Finalmente Pudemos aumentar o desempenho do RW para segmentar massas em imagens de mamografia, reduzindo o esforço necessário na fase inicial. Também observamos com os experimentos com variações das funções de pertinência que não teve impacto no algoritmo, mas esses experimentos podem ser estendidos para comparação de resultados das diferentes funções de pertinência com geração automática de sementes. Na seleção automática de sementes o algoritmo PSO teve melhores resultados que o de Evolução Diferencial, e apesar dos resultados ser

piores que a seleção manual, o resultado foi razoável dado a complexidade do problema.

Como trabalhos futuros, consideramos importante, o aumento da base de dados de imagens de mamografia com tumor. Também após definir os contornos do tumor, a partir do formato, classificar o tipo de tumor encontrado.

### **6.1. Contribuições**

- Publicamos um artigo no 30 th IEEE *International Symposium on Computer-Based Medical Systems*, qualis B1, que é um congresso realizado na Grécia na Aristotle University of Thessaloniki, contendo nossa nova proposta do método *Fuzzy Random Walker* e os resultados obtidos na comparação do *FRW* com *Random Walker* clássico.



## 7.REFERÊNCIAS BIBLIOGRÁFICAS

World Health Organization. **WHO position paper on mammography screening**. Geneva: 2014.

C. E. De Santis, C. C. Lin, A. B. Mariotto, R. L. Siegel, K. D. Stein, J. L. Kramer, R. Alteri, A. S. Robbins, and A. Jemal, "**Cancer treatment and survivorship statistics**, 2016," *CA: a cancer journal for clinicians*, vol. 66, no. 4, pp. 271--289, 2014.

INCA, **Estimativa Incidencia de câncer no Brasil**, 2018, vol. 1, no. 1.

**American Cancer Society. Breast Cancer Facts & Figures 2017-2018**. Atlanta: American Cancer Society, Inc 2017.

J. B. Harford, "**Breast-cancer early detection in low-income and middle-income countries: do what you can versus one size fits all**," *The lancet oncology*, vol. 12, no. 3, pp. 306--312, 2011.

P. Autier, M. Boniol, R. Middleton, J.-F. Doré, C. Héry, T. Zheng, and A. Gavin, "**Advanced breast cancer incidence following population-based mammographic screening**," *Annals of Oncology*, pp. 600--633, 2011.

G. J. Tortora and B. Derrickson, **Corpo Humano: Fundamentos de Anatomia e Fisiologia**. ArtMed, 2012.

K. K. Evans, R. L. Birdwell, and J. M. Wolfe, "**If you don't find it often, you often don't find it: Why some cancers are missed in breast cancer screening**," *PloSone*, vol. 8, no. 5, pp. 364--366, 2013.

J. Dheeba, N. A. Singh, and S. T. Selvi, "**Computer-aided detection of breast cancer on mammograms: A swarm intelligence optimized wavelet neural network approach**," *Journal of Biomedical Informatics*, vol. 49, pp. 45--52, 2014.

F. R. Cordeiro, W. P. Santos, and A. G. Silva-Filho, "**Segmentation of mammography by applying growcut for mass detection**," *women*, vol. 1, p. 3, 2013.

S.-W. Zheng, J. Liu, and C.-C. Liu, "**A random-walk based breast tumors segmentation algorithm for mammograms**," *Computer, Consumer and Control (IJ3C), International Journal on*, 2013.

K. C. Ciesielski, P. A. Miranda, A. X. Falcão, and J. K. Udupa, "**Join graphcut and relative fuzzy connected ness image segmentation algorithm**," *Medical image analysis*, vol. 17, no. 8, pp. 1046--1057, 2013.

P. Rahmati, A. Adler, and G. Hamarneh, "**Mammography segmentation with maximum like lihood active contours**," *Medical image analysis*, vol. 16, no. 6, pp. 1167--1186, 2012.

J. Chakraborty, S. Mukhopadhyay, V. Singla, N. Khandelwal, and R. M. Rangayyan, "**Detection of masses in mammogram susing region growing controlled by multilevel thresholding**," in *Computer-Based Medical Systems (CBMS), 2012 25th International Symposiumon*. Roma: IEEE, 2012, pp. 1--6.

S. Wongthanavas, **Cellular automata for medical image processing**. INTECH Open Access Publisher, 2011.

N. Saidin, H. A. Mat Sakim, U. K. Ngah, and I. L. Shuaib, "**Computer aided detection of breast density and mass, and visualization of other breast anatomical regions on mammograms using graphcuts**," *Computational and mathematical methods in medicine*, vol. 2013, 2013.

L. Grady, "**Random walks for image segmentation**," *IEEE transactions on pattern analysis and machine intelligence*, vol. 28, no. 11, pp. 1768--1783, 2006.

S.-w. Zheng, J. Liu, and C.-C. Liu, "**A Random-Walk Based Breast Tumors Segmentation Algorithm for Mammograms**," vol. 2, no. 2, pp. 66--74, 2013.

V. Vezhnevets and V. Konouchine, "**GrowCut - Interactive Multi-Label N-D Image Segmentation e By Cellular Automata**," *Graphicon*, pp. 150--156, 2005.

S. Lewis and A. Dong, "**Detection of breast tumor candidates using marker controlled watershed segmentation and morphological analysis**," in *Image Analysis and Interpretation (SSIAI), 2012 IEEE Southwest Symposiumon*, April 2012, pp. 1--4.

F. R. Cordeiro, B. Albuquerque, V. Macario, "**A Multi-Gaussian Fuzzy Membership Function to the Algorithm Fuzzy GrowCut Applied to Segment Lesions in Mammography Images**," xv *Encontro Nacional de Inteligência Artificial e Comutacional (ENIAC)*, em São Paulo, 2018.

Novák, V.; Perfilieva, I.; Močkoř, J., "**Mathematical principles of fuzzy logic**," Dordrecht: Kluwer Academic, 1999.

Hestenes, Magnus R.; Stiefel, Eduard, "**Methods of Conjugate Gradients for Solving Linear Systems**,". Journal of Research of the National Bureau of Standards, 1952

SENKUS, E. et al., "**Primary breast cancer: Esmo clinical practice guidelines for diagnosis, treatment and follow-up**," Annals of Oncology, v. 26, n. suppl 5, p. v8, 2015.

ACS, A. C. S. Limitations of Mammograms. 2016. Acesso em: 16dez.2018. Disponível em: <"<https://www.cancer.org/cancer/breast-cancerscreening-tests-and-early-detection/mammograms/limitations-of-mammograms.html>">.

DESHPANDE, D. S.; RAJURKAR, A. M.; MANTHALKAR, R. M. "**Medical image analysis an attempt for mammogram classification using texture based association rule mining**," In: 2013 Fourth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG). [S.l.: s.n.], 2013. p. 1–5.

SONKA, M.; HLAVAC, V.; BOYLE, R. "**Image Processing, Analysis, and Machine Vision**," Cengage Learning, 2014.

KLETTE, R. "**Concise Computer Vision: An Introduction into Theory and Algorithms**," Springer London (Undergraduate Topics in Computer Science), 2014.

SONKA, M.; HLAVAC, V.; BOYLE, R. "**Image Processing, Analysis, and Machine Vision**," Cengage Learning, 2014.

ARAÚJO, A. P. *Megapixels*. InfoEscola, 2015. Disponível em: <http://www.infoescola.com/fotografia/megapixels/>>. Acesso em: 19 dez. 2018.

RAJASEKARAN, S.; PAI, G. V. "**Neural Networks, Fuzzy Logic and Genetic Algorithms**," PHI Learning Private Limited, 2011.

GONZALEZ, R. C.; WOODS, R. E.; EDDINS, S. L. "**Digital image processing using MATLAB**," Gatesmark Publishing Knoxville, 2009.

DOWSLAND, K. A.; THOMPSON, J. M. "**Simulated annealing**. In: **Handbook of Natural Computing** " Springer, 2012. p. 1623–1655.

F. Wang and D. P. Landau, "**Efficient, multiple-range random walk algorithm to calculate the density of states**," Physical Review Letters, vol. 86, no. 10, p. 2050, 2001.