



**Universidade Federal de São Paulo**  
**Campus São José dos Campos**  
**Instituto de Ciência e Tecnologia**

**Implementação de um Sistema Híbrido Utilizando  
Mapas Auto-Organizáveis de Kohonen e GrabCut para  
Segmentação de Imagens**

**Aluno:** Rodolfo Freitas Alves Borges  
**Orientadora:** Profa. Dra. Regina Célia Coelho

**São José dos Campos, 21 de Abril de 2017.**

## **Resumo**

Normalmente a segmentação utiliza informações locais na imagem para obter o melhor resultado. Um dos melhores métodos semiautomático para segmentação de imagens é o GrabCut. Para utilizá-lo, o usuário deve selecionar o menor retângulo possível que envolva o objeto que se deseja segmentar na imagem. No entanto, se esse retângulo não for justo o suficiente para eliminar o máximo possível de fundo, o resultado pode não isolar o objeto corretamente. O objetivo geral deste projeto foi a construção um sistema híbrido composto por Mapa Auto-Organizáveis de Kohonen e GrabCut de forma que o Kohonen forneça um agrupamento próximo ao objeto que se deseja segmentar, melhorando a segmentação final do GrabCut. O sistema foi desenvolvido inteiramente em MatLab. Em testes com imagens de ecocardiogramas de baixo contraste realizados no laboratório, o sistema híbrido Kohonen-GrabCut apresentou significativa melhora da segmentação em relação ao GrabCut sozinho. Em imagens em que o GrabCut retornou o retângulo inteiro de entrada, o sistema híbrido, utilizando a mesma entrada, apresentou segmentação muito próxima ao objeto desejado.

**Palavras Chave:** Segmentação de Imagens, GrabCut, Mapas Auto-Organizáveis de Kohonen.

## **Sumário**

<b>1. INTRODUÇÃO .....</b>	<b>3</b>
<b>2. MATERIAS E MÉTODOS .....</b>	<b>4</b>
<b>2.1. Bancos de Dados Utilizados .....</b>	<b>4</b>
<b>2.2. Considerações sobre Mapas Auto-Organizáveis de Kohonen (K-SOM)...</b>	<b>4</b>
<b>2.3. Considerações sobre o método de segmentação GrabCut .....</b>	<b>5</b>
<b>2.4. Desenvolvimento do Sistema Híbrido.....</b>	<b>6</b>
2.4.1. <i>Implementação do Kohonen .....</i>	<i>6</i>
2.4.2. <i>Modificações no GrabCut.....</i>	<i>8</i>
<b>2.5. Descrição de uso .....</b>	<b>9</b>
<b>3. RESULTADOS .....</b>	<b>12</b>
<b>3.1. Teste empírico.....</b>	<b>12</b>
<b>4. CONCLUSÃO .....</b>	<b>15</b>
<b>5. REFERÊNCIAS .....</b>	<b>16</b>

## 1. INTRODUÇÃO

Segmentação de imagens tem por objetivo dividir uma imagem em regiões que podem ser mais representativas e mais fáceis de analisar. Estas regiões podem corresponder a superfícies, objetos ou partes de objetos individuais. Normalmente, a segmentação de imagem é utilizada para localizar objetos e/ou limites em imagens [Gonzalez, 2006]. Ela também pode ser definida como um processo de rotulação de pixels em uma imagem, sendo que todos os pixels com o mesmo rótulo compartilham características em comum [Lalitha, 2013]. Normalmente a segmentação utiliza informações locais na imagem para obter o melhor resultado, como por exemplo, informações de cor, ou relacionadas à borda, ou ainda, de textura [Sharma, 2012].

Existem diversos métodos de segmentação de imagens na literatura, dentre os quais se destaca o GrabCut [Rother, 2004], considerado um dos melhores métodos semi-automático para segmentação de imagens. Para utilizá-lo, o usuário deve selecionar o menor retângulo possível que envolva o objeto que se deseja segmentar na imagem. No entanto, se esse retângulo não for justo o suficiente para eliminar o máximo possível de fundo, o resultado pode não isolar o objeto corretamente.

Por outro lado, algoritmos de agrupamento, como o Mapa Auto-Organizável de Kohonen (K-SOM – Kohonen Self-Organizing Map) [Kohonen 1990; Kohonen 1996] são muito indicados para auxiliar na segmentação de imagens por apresentar como resultado diferentes níveis de agrupamento da imagem de forma que um destes grupos contenha o objeto de interesse. Uma grande vantagem deste método é que o agrupamento é realizado de forma totalmente automática.

Tendo em vista o exposto acima, este projeto teve como objetivo construir um sistema híbrido unindo os métodos de Kohonen e GrabCut de forma que o Kohonen forneça uma pré-segmentação para o GrabCut, resultando, dessa forma, em uma segmentação do objeto com menos necessidade de correções. No entanto, para que o resultado seja satisfatório, a proposta é que o usuário escolha qual dos agrupamentos apresentados pelo Kohonen melhor representa a região em que o objeto de interesse está contido.

Com isso, é esperado que o resultado da segmentação final seja muito melhor do que quando usamos apenas o GrabCut com retângulo, uma vez que o agrupamento do Kohonen fornecerá como semente do GrabCut um contorno mais próximo ao contorno do objeto de interesse, eliminando, assim, muito mais partes contendo fundo da imagem.

## **2. MATERIAS E MÉTODOS**

Neste capítulo serão apresentados os bancos de dados e as técnicas utilizadas neste trabalho.

### **2.1. Bancos de Dados Utilizados**

O conjunto de imagens utilizadas para os testes é formado por imagens provenientes do banco de imagens Berkeley Segmentation Data Set and Benchmarks 500 (BSDS500) do Computer Vision Group da Universidade de Berkeley, de imagens ecocardiográficas oriundas do Youtube e por imagens coletadas na internet. Tal conjunto permite testar o método híbrido Kohonen-GrabCut com imagens de diferentes tipos.

### **2.2. Considerações sobre Mapas Auto-Organizáveis de Kohonen (K-SOM)**

O K-SOM [Kohonen, 1990; Kohonen, 1996] é um método muito utilizado para reduzir a dimensionalidade de um conjunto de dados. Ele é constituído por um arranjo de neurônios que procura estabelecer e preservar noções de vizinhança. Como estes mapas apresentam propriedades de auto-organização, eles podem ser aplicados a problemas de agrupamentos. Dessa forma, o objetivo é obter um mapeamento do espaço original para o espaço em que o arranjo de neurônios é definido. Este mapeamento deve ser tal que minimize a perda de informação.

Para o processamento de imagens, o K-SOM geralmente é composto por duas camadas de neurônios artificiais: uma de entrada que recebe os dados e outra de saída que é o próprio mapa. Cada neurônio do mapa é conectado a todos os outros de entrada. Os pesos dessas conexões formam o vetor  $w_j(t)$  no tempo  $t$ , que possui a mesma dimensão que um vetor de entrada

$x(t) = (x_1, x_2, \dots, x_m)$  também no tempo  $t$ . Em  $w_j(t)$ ,  $j$  é o índice do neurônio da camada de saída.

O treinamento de uma rede de Kohonen é feito de modo competitivo e não supervisionado por meio da repetição de três passos básicos (amostragem, casamento por similaridade e atualização) até a formação do mapa estar completa. A seguir é apresentado um resumo do funcionamento do algoritmo.

a) *Inicialização*: Escolha valores aleatórios para os vetores de peso iniciais  $w_j(0)$ . A única restrição aqui é que os  $w_j(0)$  sejam diferentes para todos os neurônios da rede. Pode ser desejável manter a magnitude dos pesos pequena.

b) *Amostragem*: Retire uma amostra  $x$  do espaço de entrada com uma certa probabilidade; o vetor  $x$  representa o padrão de ativação que é aplicado à grade. A dimensão do vetor  $x$  é igual a  $m$ .

c) *Casamento por similaridade*: Encontre o neurônio com o melhor casamento (vencedor)  $i(x)$  no passo de tempo  $n$  usando o critério da mínima distância euclidiana, conforme a Eq. 1,

$$i(x) = \arg \min_j \|x(n) - w_j\|, j = 1, 2, \dots, l. \quad (1)$$

d) *Atualização*: Ajuste os vetores de peso sináptico de todos os neurônios usando a fórmula de atualização, de acordo com a Eq. 2,

$$w_j(n+1) = w_j(n) + c(n)h_{j,i(x)}(n)(x(n) - w_j(n)), \quad (2)$$

em que  $c(n)$  é o parâmetro da taxa de aprendizagem e  $h_{j,i(x)}(n)$  é a função de vizinhança centrada em torno do neurônio vencedor  $i(x)$ ; ambos  $c(n)$  e  $h_{j,i(x)}(n)$  variam dinamicamente durante a aprendizagem para melhorar os resultados.

e) *Continuação/Parada*: Continue com o passo b) até que não sejam observadas modificações significativas no mapa de características.

### 2.3. Considerações sobre o método de segmentação GrabCut

O método GrabCut é uma evolução do método de segmentação de imagens utilizando corte iterativo em grafo [Boykov, 2001; Boykov 2006; Rother, 2004]. Nele o usuário especifica um retângulo ao redor do objeto que se deseja segmentar. A seguir, é realizado o corte em grafo para inferir a

segmentação [Bokov, 2006]. As etapas de adistribuição de cores e corte são repetidas até que a função de energia utilizada para separar o objeto do fundo alcance seu mínimo. A convergência do método é sempre garantida.

Mais especificamente, a segmentação começa construindo um grafo da imagem em que cada nó representa um pixel na imagem, e incorporando dois nós extras que representam o objeto (O) e o fundo (F). Inicialmente todos os nós são conectados aos nós O e F. Os pesos entre os nós são determinados considerando a informação de borda na imagem (quanto maior a diferença entre as intensidades dos pixels, menor será o peso). O retângulo selecionado pelo usuário define um *trimap* em que cada pixel é marcado como O, F ou D (desconhecido). Os pixels que estiverem fora do retângulo serão permanentemente marcados como fundo (F) e os outros serão marcados como desconhecidos (D). A seguir, utilizando o k-média e corte em grafos, os pixels marcados como D são iterativamente reagrupados de forma que, ao final, todos os pixels estejam marcados apenas com O ou F. As informações de O e F são usadas para calcular uma função de energia. Quando esta função atinge seu mínimo global, o algoritmo é finalizado.

## **2.4. Desenvolvimento do Sistema Híbrido**

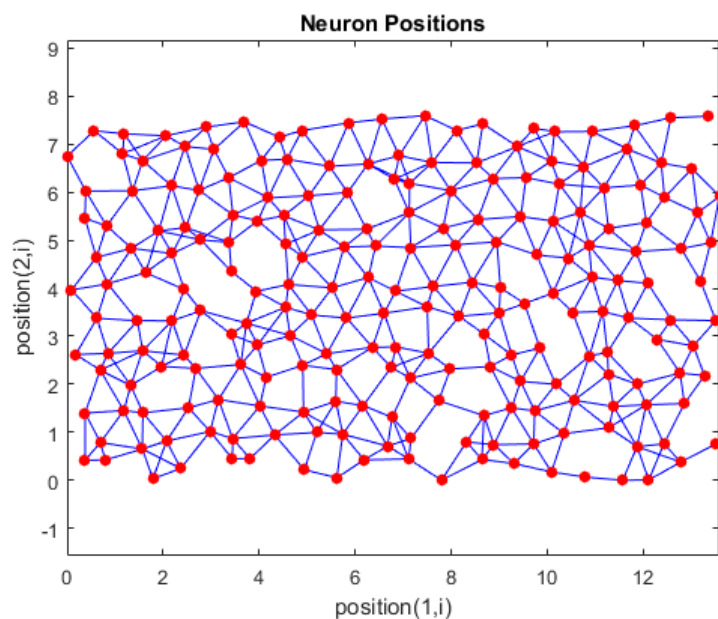
Primeiramente, foi desenvolvido um sistema que, utilizando Mapas Auto-Organizáveis de Kohonen, recebe uma imagem (ou apenas parte da imagem) como entrada e retorna os agrupamentos feitos à partir de informações de níveis de cinza e vizinhança. Sequencialmente, modificou-se uma implementação do GrabCut para MatLab de Olga Veksler (2005) para aceitar como semente do GrabCut agrupamentos retornados pelo Kohonen. Por fim, foram feitos ajustes na interface gráfica e no algoritmo do sistema, para que fossem aceitas modificações na segmentação retornada pelo GrabCut. A seguir será detalhado como o sistema híbrido foi implementado.

### **2.4.1. Implementação do Kohonen**

Para a criação dos Mapas Auto-Organizáveis de Kohonen foi utilizada a função *selforgmap* do Matlab. Para o treinamento da rede foi utilizada a função

*train* do Matlab, que modifica a rede baseada no conjunto de dados passado como parâmetro. Por se tratar de um sistema para segmentar imagens de diversos tipos, sem um padrão definido, optou-se por treinar a rede com a mesma imagem que se deseja segmentar.

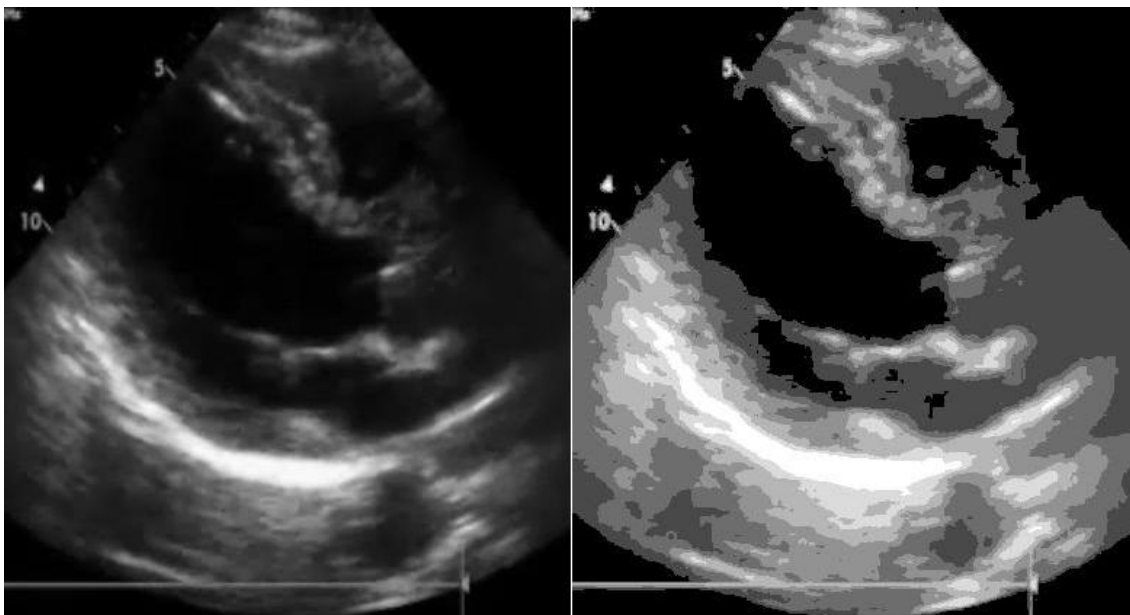
Utilizando testes realizados no Laboratório de Computação Biomédica do Instituto de Ciência e Tecnologia de São José dos Campos, definiu-se que a melhor configuração do mapa de Kohonen, que mais se adaptou as necessidades buscadas foi uma rede neural com apenas uma camada de 7 neurônios, onde cada neurônio considera os outros 6 como vizinhos. A distância entre os neurônios foi definida como distância randômica e o número de iterações da etapa inicial de treinamento foi definido como 100. Desta maneira, cada entrada realizada retornará no máximo 7 agrupamentos distintos. A Figura 1 exemplifica a configuração da malha de neurônios com distância randômica.



**Figura 1:** Neurônios de um Mapa Auto-Organizável de Kohonen arranjados em um padrão randômico N-dimensional.

Desta maneira, cada imagem processada pelos mapas de Kohonen terá como saída no máximo 7 agrupamentos distintos, como está exemplificado na Figura 2.





**Figura 2:** Esquerda: Imagem original. Direita: Agrupamentos feito pelos Mapas de Kohonen

#### 2.4.2. Modificações no GrabCut

Para a implementação do GrabCut, optou-se por utilizar uma versão já implementada para Matlab, disponibilizada por Olga Veksler através no link <http://www.csd.uwo.ca/faculty/olga/software.html>. A versão utilizada implementa métodos de minimização de energia em cortes em grafos descritos em [Boykov, Veksler, Zabih, 2001].

Além disso utiliza-se as bibliotecas *energy.h* para implementar técnicas de minimização binária de energia, descritas em [Kolmogorov, Zabih, 2002] e os códigos *graph.h*, *block.h*, *maxflow.cpp* para implementar o algoritmo de fluxo máximo (maxflow) descrito em [Boykov, Kolmogorov, 2004].

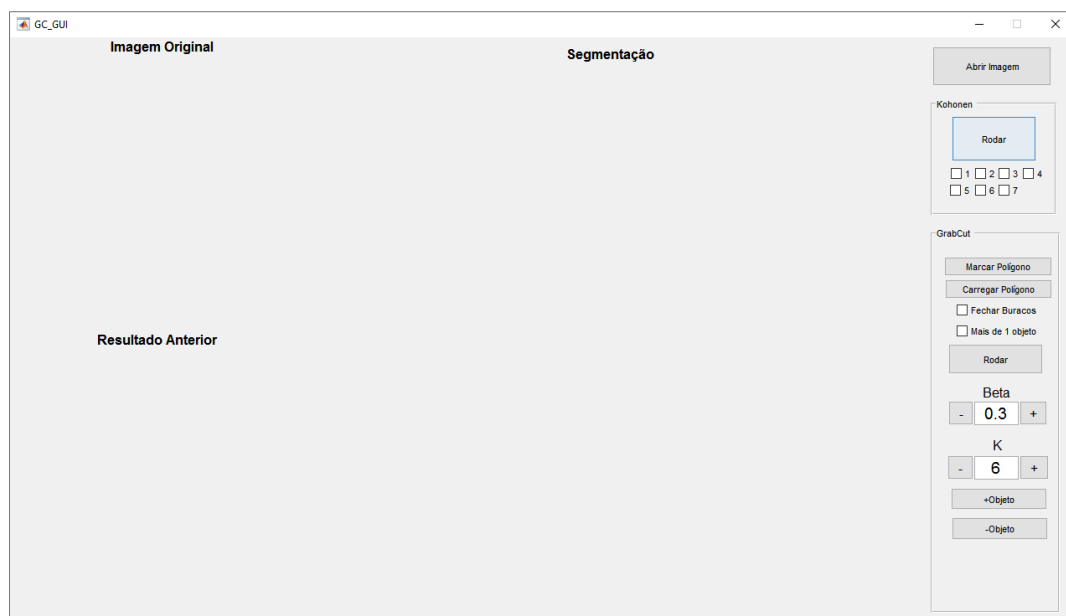
Para o GrabCut aceitar como semente um ou mais grupos do agrupamento feito pelo Mapa de Kohonen, foram feitas modificações no código de modo que a entrada para o GrabCut seja a imagem original e uma máscara composta por um ou mais grupos retornados pelo Kohonen, escolhidos pelo usuário. Todos os pixels da imagem original com posição diferente das posições contida na máscara são reconhecidos como fundo e os que tem posição igual às contidas na máscara são processados pelo GrabCut que decide quais pixels pertencem ao fundo e quais pertencem ao objeto. Foi ainda desenvolvido ferramentas de pós processamento para possibilitar o fechamento eventuais buracos nas segmentações, possibilidade de reconhecer

como objeto mais de um componente conexo e a possibilidade de adicionar à segmentação partes dos objetos que foram reconhecidas como fundo e/ou remover fundo que foi reconhecido como objeto.

## 2.5. Descrição de uso

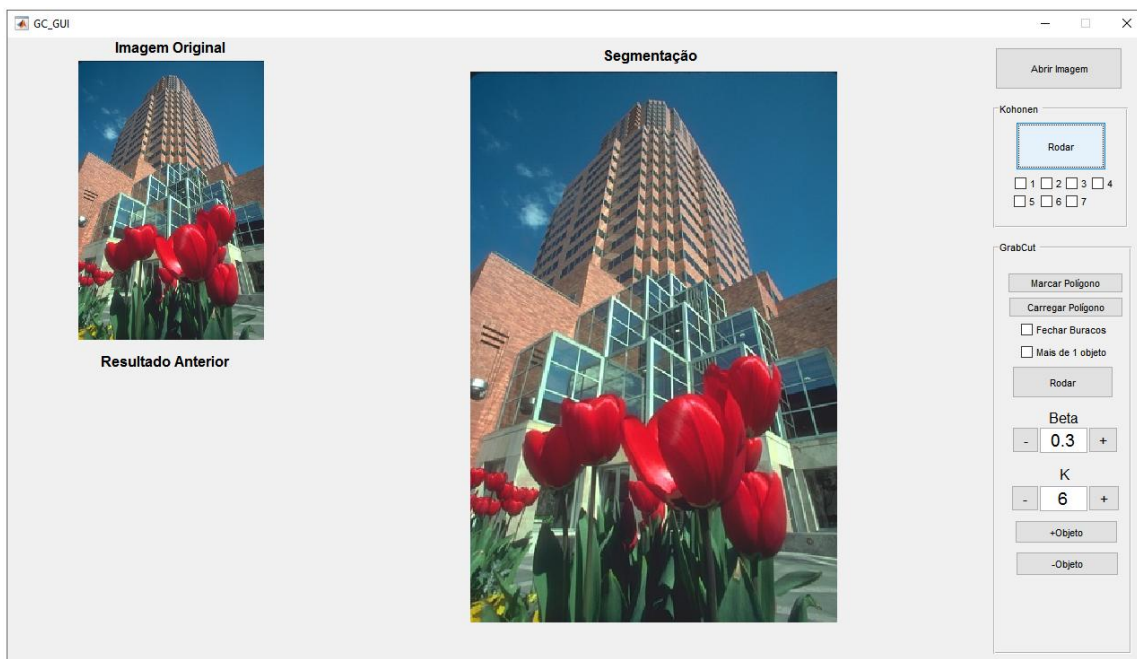
Nesta sessão serão detalhadas as funcionalidades do Sistema Híbrido Kohonen-GrabCut.

Utilizando o botão <Abrir Imagem> Abra a imagem que deseja segmentar.



**Figura 3:** Instruções de uso - 1: Interface Gráfica do Sistema Híbrido Kohonen-GrabCut

Na guia "Kohonen" o botão <Rodar> habilita a função para desenhar polígono na imagem que isolará a região de interesse. O objetivo aqui é apenas isolar um pouco o que se deseja segmentar, pois há várias informações na imagem e o usuário deve marcar o que se deseja segmentar. No exemplo da Figura 4, o objetivo é segmentar os botões vermelhos das tulipas. É possível corrigir o polígono desenhado arrastando os seus pontos. Depois de completar o desenho, com um duplo clique dentro do polígono, o algoritmo das redes de Kohonen é acionado.

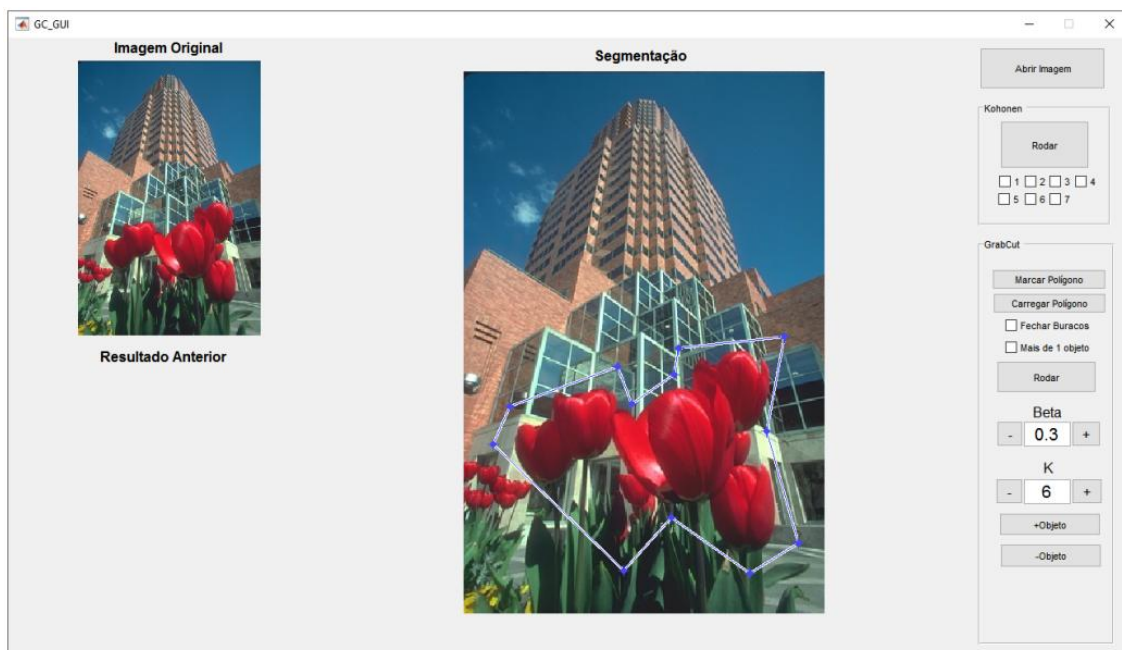


**Figura 4:** Instruções de uso - 2: ao abrir a imagem que se deseja segmentar a mesma será apresentada na janela Imagem Original. Ao clicar em <Rodar> na guia Kohonen a imagem é apresentada em tamanho maior na janela Segmentação, com a função de desenhar polígono já habilitada.

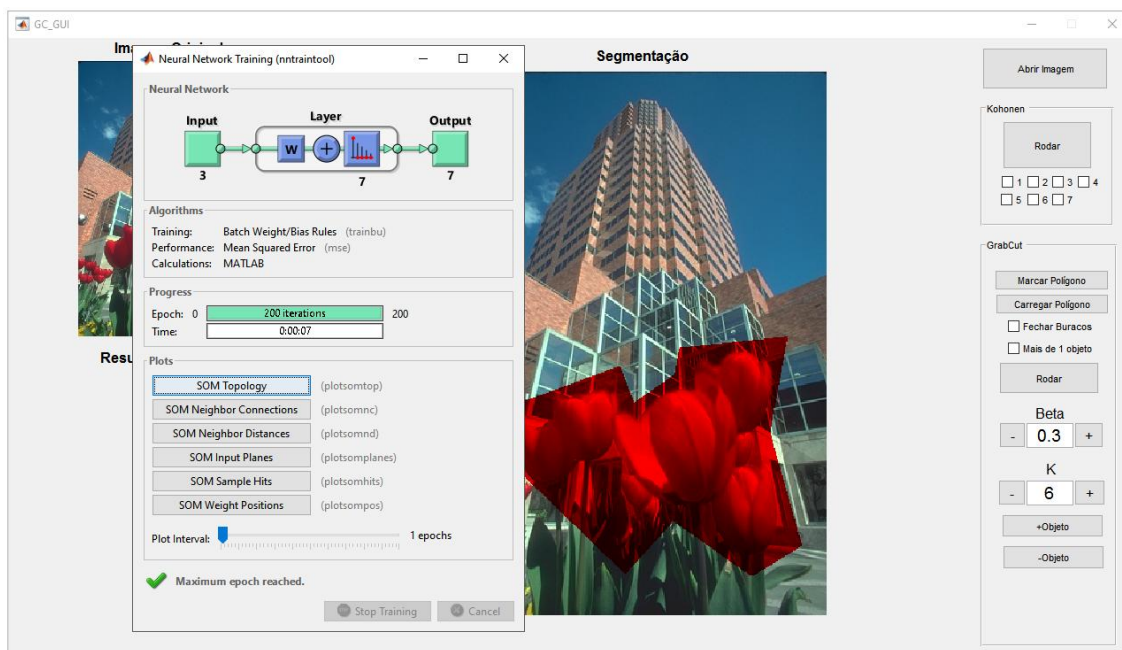
Utilizando checkbox na guia Kohonen, é possível selecionar os grupos que melhor contém o objeto que se deseja segmentar. Também é possível selecionar se deseja mais de um objeto conexo e se deseja que sejam fechados os buracos no interior do objeto. Clicando no botão <Rodar> na guia GrabCut, o algoritmo do GrabCut se iniciará. Neste exemplo vamos deixar marcado para fechar buracos e desmarcada a opção de mais de um objeto conexo.

Ao final, segmentação será apresentada. Caso algum ajuste precise ser feito, pode-se adicionar partes do objeto à segmentação e retirar partes de fundo que foram reconhecidas como objeto utilizando os botões <+Objeto> e <-Objeto> respectivamente, e rodando o GrabCut novamente.

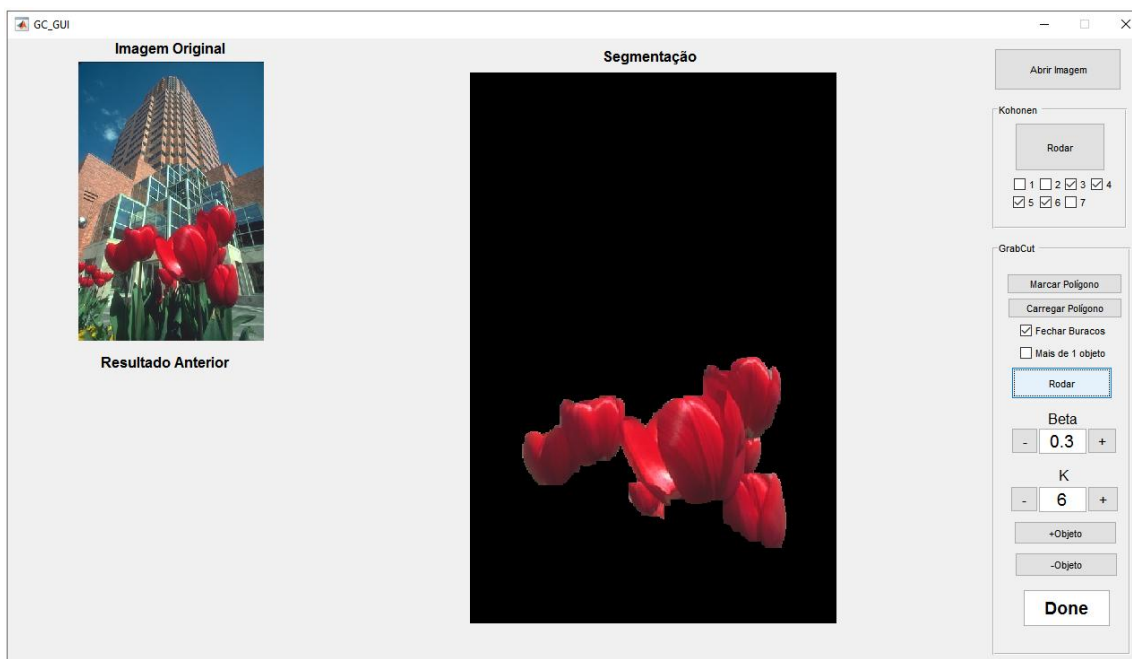
As Figuras 5, 6 e 7 mostram o polígono enviado ao algoritmo de Kohonen, contendo os objetos que se deseja segmentar, a janela de treinamento da rede neural que se abre após um duplo clique dentro do polígono desenhado e a segmentação final obtida, respectivamente.



**Figura 5:** Instruções de uso – 3: nesta imagem vemos o polígono que será enviado à rede neural de Kohonen, contendo os botões maiores das tulipas.



**Figura 6:** Instruções de uso - 4: Após um duplo clique no interior do polígono, a janela de treinamento da rede neural deve se abrir. Está janela pode ser fechada após o termino do treino da rede.



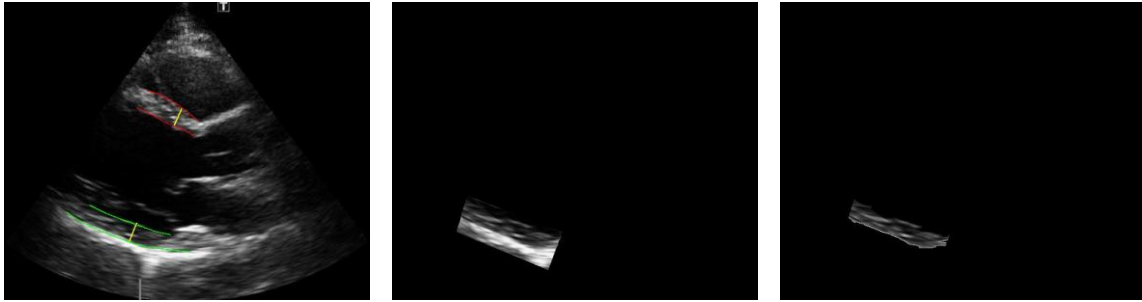
**Figura 7:** Instruções de uso - 7: Segmentação final apresentada. Caso alguma correção precise ser feita. Pode se ainda clicar nos botões <+Objeto> e <-Objeto> para adicionar ou remover partes do objeto da segmentação final.

### 3. RESULTADOS

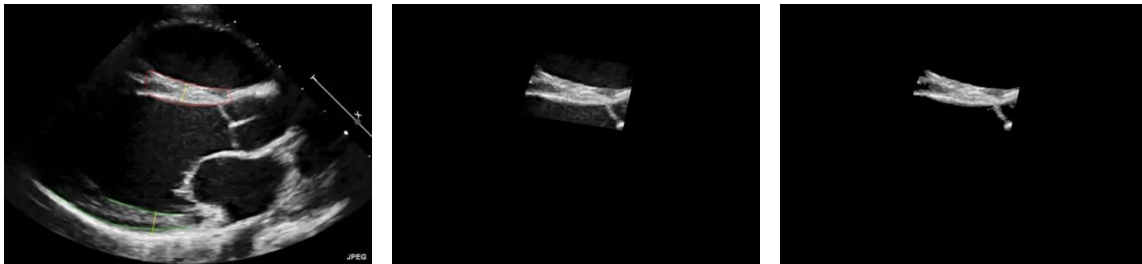
#### 3.1. Teste empírico

Para a validação da técnica descrita neste trabalho foi feita a comparação visual entre segmentações feitas pelo sistema híbrido e a segmentação feita apenas pelo GrabCut. Apesar do sistema híbrido permitir que sejam feitas correções na segmentação final, para fins de teste foi considerada apenas a primeira segmentação gerada, sem correções.

Nos testes realizados com imagens ecocardiográficas, o GrabCut muitas vezes retornou todo o polígono de entrada, e o Sistema Híbrido apresentou maior estabilidade, conseguindo isolar melhor a região de interesse, como é mostrado nas Figuras 8, 9 e 10.



**Figura 8:** Esquerda: Imagem original com a marcação (em verde) da parede posterior que se deseja segmentar. Meio: Segmentação com GrabCut retornou todo o polígono de entrada. Direita: Segmentação com Sistema Híbrido apresentou um contorno mais próximo ao desejado.

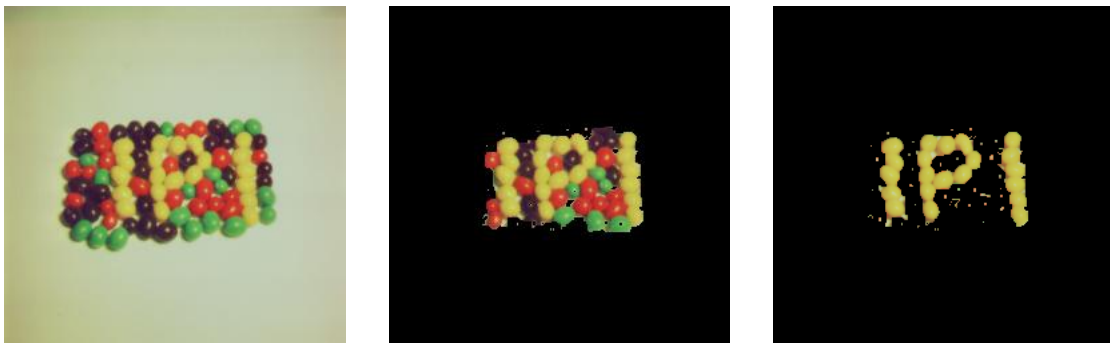


**Figura 9:** Esquerda: Imagem original com marcação (em vermelho) das paredes do septo que se deseja segmentar. Meio: Segmentação com GrabCut retornou todo o polígono de entrada. Direita: Segmentação com Sistema Híbrido conseguiu isolar o septo de maneira mais eficiente.



**Figura 10:** Esquerda: Imagem original com a marcação (em verde) da parede posterior que se deseja segmentar. Meio: Segmentação com GrabCut retornou todo o polígono de entrada. Direita: Segmentação com Sistema Híbrido apresentou um contorno mais próximo ao desejado.

O Sistema Híbrido apresentou grande vantagem para segmentar imagens onde os objetos, apesar de bem definidos, estão misturados com outros objetos. Como mostra a Figura 11.





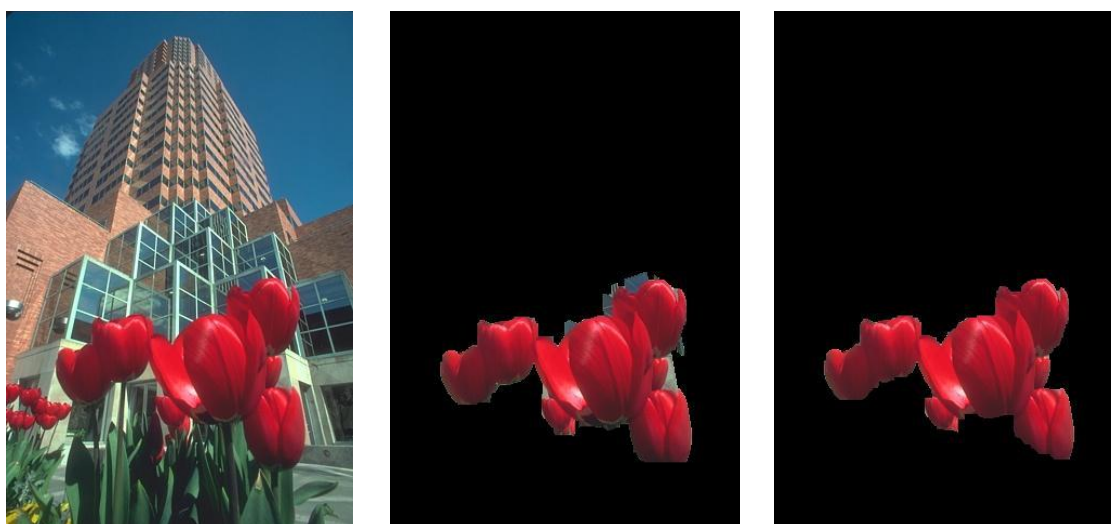
**Figura 11:** Esquerda: Imagem original onde deseja-se segmentar os m&m's amarelos. Meio: Segmentação com GrabCut onde não foi possível especificar o objeto que se deseja segmentar. Direita: Segmentação com Sistema Híbrido permite pela seleção de grupos especificar bem o que se deseja segmentar.

Tal vantagem se deve ao fato do Sistema Híbrido apresentar a possibilidade de seleção de grupos que podem ser compostos por grupos não conexos e apresentar formas difíceis de se marcar usando polígonos.

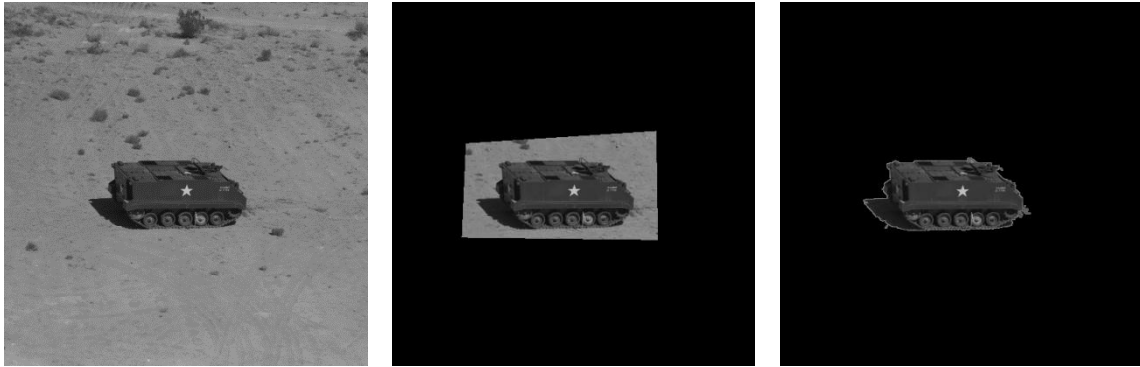
As Figuras 12, 13 e 14 apresentam outros exemplos de comparações entre os dois métodos



**Figura 12:** Esquerda: Imagem original onde deseja-se segmentar o pimentão verde localizado no meio da imagem. Meio: A segmentação obtida com o GrabCut reconheceu as sombras ao lado do pimentão como parte do objeto. Direita: Segmentação com Sistema Híbrido conseguiu eliminar mais fundo, obtendo uma segmentação melhor.



**Figura 13:** Esquerda: Imagem original onde se deseja segmentar os botões vermelhos das tulipas em primeiro plano. Meio: Segmentação com GrabCut reconheceu partes da janela do edifício ao fundo como parte do objeto. Direita: Segmentação com Sistema Híbrido apresenta melhor segmentação dos botões das tulipas.



**Figura 14:** Esquerda: Imagem original onde deseja-se segmentar o tanque de guerra., Segmentação com GrabCut retornou todo o polígono de entrada. Direita: Segmentação com Sistema Híbrido apresentou melhor aproximação do objeto desejado.

#### 4. CONCLUSÃO

Levando em consideração os testes realizados, pode-se dizer que o objetivo geral do trabalho foi atingido. Na maior parte dos testes, o Sistema Híbrido apresentou maior eficácia em separar o fundo da imagem do objeto que se deseja segmentar. Tal melhora ocorreu pelo fato do sistema híbrido permitir que a semente para o GrabCut seja algo mais próximo do objeto.

A melhoria é mais acentuada em imagens de baixo contraste, em que por vezes o GrabCut retornou todo o polígono de entrada, enquanto que o sistema híbrido conseguiu uma segmentação bem próxima ao objeto.



## 5. REFERÊNCIAS

[Rother, 2004] Rother, C.; Kolmogorov V.; Blake A. "Grabcut": interactive foreground extraction using iterated graph cuts. ACM Transactions on Graphics (TOG) – Proceedings of ACM SIGGRAPH 2004. Volume 23 Issue 3, August 2004. Pages 309-314.

[Sharma, 2012] Sharma, N., Mishra, M., Shrivastava, M. Colour image segmentation techniques and issues: an approach. International Journal of Scientific & Technology Research 1, 9–12, 2012.

[Gonzalez, 2006] Gonzalez, R.C.; Woods, R.E. Digital Image Processing, 3rd edition, Prentice-Hall, Inc, 2006.

[Lalitha, 2013] Lalitha, M., Kiruthiga, M., Loganathan, C. A Survey on Image Segmentation through Clustering Algorithm. International Journal of Science and Research (IJSR) 2, 348–358, 2013.

[Kohonen 1990] Kohonen, T. The self-organizing map, Proceedings of the IEEE, vol. 78, no. 9, pp. 1464–1480, 1990.

[Kohonen, 1996] Kohonen, T., Oja, E., Simula, O., Visa, A., Kangas, J. Engineering applications of the self-organizing map. Proceedings of the IEEE 84, 1358–1384, 1996.

[Boykov, 2001] Boykov, Y. Y.; Jolly, M. Interactive Graph Cuts for Optimal Boundary & Region Segmentation of Objects in N-D Images. Proceedings of International Conference on Computer Vision, Vancouver, Canada, July 2001.

[Boykov, 2006] Boykov, Y. Y.; FUNKA-LEA G. Graph Cuts and Efficient N-D Image Segmentation. International Journal of Computer Vision 70(2), 109–131, 2006

[Rother, 2004] Rother, C.; Kolmogorov V.; Blake A. "Grabcut": interactive foreground extraction using iterated graph cuts. ACM Transactions on Graphics (TOG) – Proceedings of ACM SIGGRAPH 2004. Volume 23 Issue 3, August 2004. Pages 309-314.

[Boykov, Veksler, Zabih, 2001] Boykov, Y. Y., Veksler, O, Zabih, R.; Efficient Approximate Energy Minimization via Graph Cuts IEEE transactions on PAMI, vol. 20, no. 12, p. 1222-1239, November 2001.

[Kolmogorov, Zabih, 2002] Kolmogorov, V, Zabih, R; What Energy Functions can be Minimized via Graph Cuts? To appear in IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI). Earlier version appeared in European Conference on Computer Vision (ECCV), May 2002.

[Boykov, Kolmogorov, 2004] Boykov, Y. Y., Kolmogorov, V; An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision. In IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), September 2004