

Avaliação de superpixels para segmentação de imagens

Felipe Augusto Lima Reis¹

¹PUC Minas - Pontifícia Universidade Católica de Minas Gerais

`falreis@sga.pucminas.br`

Abstract. *Superpixels are structures that group similar pixels into sets that reflect aspects of the image. This article evaluates the use of SLICO superpixels and partition hierarchy for segmentation. Using **neural networks** for segmentation, the results of superpixels images with different levels of granularity and untreated images were compared to the ground-truth. The article also evaluates the training time of neural networks for superpixels based images. For training and evaluation, the Berkeley Segmentation Data Set (BSDS500) [?] image was used.*

Resumo. *Superpixels são estruturas que agrupam pixels semelhantes em conjuntos que refletem aspectos da imagem. Este artigo avalia a utilização de superpixels SLICO e EGB para segmentação. Avalia também os benefícios da combinação dos métodos para produção de segmentação. Por fim, o artigo mostra uma possível implementação do método hierárquico para a segmentação combinada dos métodos SLIC e EGB, que pode ser utilizada para detalhamento e zoom sobre segmentações. Para avaliação foram utilizadas o conjunto de treinamento de imagens do Berkeley Segmentation Data Set (BSDS500) [?]. Os resultados obtidos pelos métodos foram comparados em relação ao ground-truth, utilizando o método de precisão e revocação.*

1. Introdução

A segmentação de imagens consiste em dividir uma imagem em um conjunto de regiões logicamente agrupadas, de modo a reunir áreas que contém informação relevante dentro dos grupos [?]. Nessa tarefa, tomamos os *pixels* como unidades básicas de processamento [?]. O agrupamento de pixels em unidades maiores permite um tipo de segmentação chamado de *oversegmentation* [?], ilustrado na figura ???. O uso de superpixels possibilita o aumento da velocidade de processamento posterior, uma vez que a quantidade de pixels diminui consideravelmente em relação a imagem original.

A utilização de superpixels possibilita a redução de itens a serem processados, entretanto pode causar perda de informação importante. No entanto, para alguns casos, a perda de qualidade pode se justificar em relação ao ganho de velocidade obtido utilizando esse tipo de operação. Essa relação consiste então em um *trade-off* entre ambas as características, sendo viáveis em alguns cenários de processamento em tempo real ou para dispositivos com baixo desempenho.

Alguns métodos de geração de superpixels são utilizados para segmentação de imagens e detecção de bordas, como os métodos EGB [?] e SLIC [?]. Esse trabalho investiga se a utilização de métodos segmentação baseados em superpixels, a composição

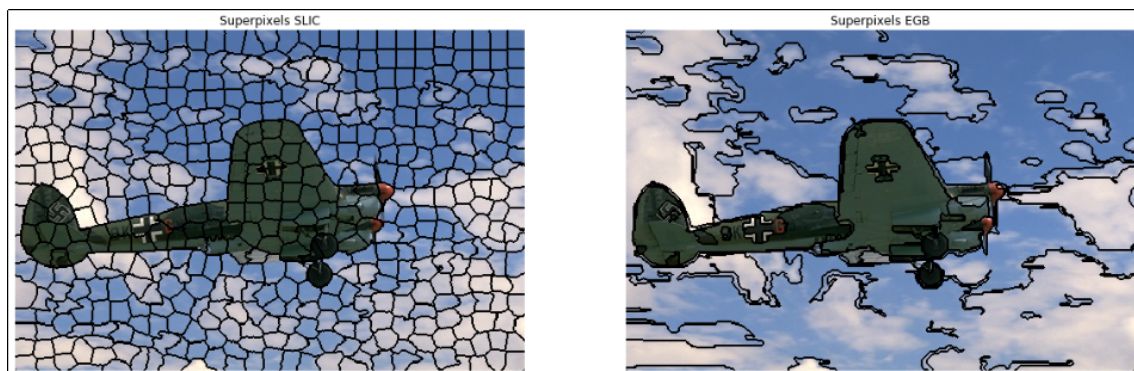


Figura 1. Imagens segmentadas utilizando superpixels SLIC e EGB

de métodos de *oversegmentation* e a utilização de redes neurais convolucionais para segmentação.

Em relação aos métodos de segmentação utilizando redes neurais, o trabalho também tentará identificar se o treinamento utilizando imagens pré processadas obtêm resultados semelhantes àqueles utilizando imagens originais, na etapa de validação.

O presente trabalho apresenta a seguinte estrutura: a Seção ?? mostra o referencial teórico para construção do trabalho, a Seção ??, exibe os materiais e métodos utilizados nos testes; a Seção ?? mostra os resultados obtidos nos testes realizados e a discussões dos mesmos; a Seção ?? contém a conclusão do artigo, com as considerações finais.

2. Referencial Teórico

2.1. Superpixels

Superpixels são estruturas que agrupam pixels semelhantes em conjuntos. O agrupamento possibilita a redução de complexidade das tarefas de processamento [?], ao reduzir a quantidade de itens a serem processados. Os superpixels são utilizados na área de visão computacional para solução de vasto número de problemas, como detecção de contorno [?], segmentação [?] e localização de objetos [?].

Superpixels, segundo [?], devem capturar importante grupos ou regiões, refletindo aspectos da imagem. Devem também ser executados em tempo próximo ao linear em relação a quantidade de pixels. Existem diversas abordagens para a geração de superpixels [?]. Dentre elas, podemos classificá-las, segundo o método de agrupamento em:

- *Algoritmos baseados em grafos*: utilizam abordagem baseadas em grafos para correlação entre pixels e criação dos conjuntos. Dentre os algoritmos baseados em grafos podemos citar o *Efficient Graph-Based Image Segmentation* (EGB) [?];
- *Algoritmos baseados em gradiente ascendente*: utilizam métodos de gradiente ascendente iterativamente até que os critérios de convergência correspondam a forma de um superpixel. Nesse conjunto, podemos citar as abordagens *watersheds* [?] [?]
- *Algoritmos de clusterização iterativo*: utilizam métodos de clusterização, como o *k-means*, para produção de superpixels. Um exemplo desse algoritmo é o SLIC (*Simple Linear Iteravite Clustering*) [?]

2.1.1. Superpixels SLIC

O algoritmo SLIC utiliza um único parâmetro k , correspondente a quantidade aproximada de superpixels. A fim de produzir tamanhos semelhantes, o intervalo analisado é $S = \sqrt{N/k}$, onde N é o número de pixels da imagem. Os centros são movidos para o local das sementes correspondentes a posição mais baixa do gradiente em uma vizinhança de 3×3 , evitando que superpixels sejam centrados nas bordas ou em uma posição de ruído [?]. Em seguida, cada pixel é associado com o centro do cluster mais próximo, de modo que a região de busca se sobreponham [?]. A fim de aumentar o desempenho do algoritmo, a região de busca é limitada em 2 vezes o tamanho aproximado do superpixel S , gerando busca em uma área $2S \times 2S$ [?]. Em seguida, um passo de atualização dos centros dos clusters e computa o erro residual E [?]. O algoritmo disponível na figura ??, resume as informações descritas nesse parágrafo.

SLIC superpixel segmentation

```
/* Initialization */
Initialize cluster centers  $C_k = [l_k, a_k, b_k, x_k, y_k]^T$  by
sampling pixels at regular grid steps  $S$ .
Move cluster centers to the lowest gradient position in a
 $3 \times 3$  neighborhood.
Set label  $l(i) = -1$  for each pixel  $i$ .
Set distance  $d(i) = \infty$  for each pixel  $i$ .

repeat
  /* Assignment */
  for each cluster center  $C_k$  do
    for each pixel  $i$  in a  $2S \times 2S$  region around  $C_k$  do
      Compute the distance  $D$  between  $C_k$  and  $i$ .
      if  $D < d(i)$  then
        set  $d(i) = D$ 
        set  $l(i) = k$ 
      end if
    end for
  end for

  /* Update */
  Compute new cluster centers.
  Compute residual error  $E$ .
until  $E \leq \text{threshold}$ 
```

Figura 2. Algoritmo SLIC - Adaptado de [?]

Para o algoritmo descrito na figura ??, é necessário compreender o método para cálculo da medida de distância D entre os conjuntos. Devido ao algoritmo trabalhar no *colorspace* CIELAB, com o espaço-plano *labxy*, a posição do pixel pode assumir um intervalo de valores de acordo com o tamanho da image [?]. Com isso, o cálculo da distância não pode ser feito utilizando uma distância euclidiana, sendo necessária uma prévia normalização da proximidade espacial e de cor. Para isso é utilizado a fórmula $D = \sqrt{d_c^2 + (d_s/S)^2 \cdot m^2}$, onde D corresponde a distância em 5 dimensões do espaço *labxy*, d_c e d_s correspondem à proximidade de cores e espacial; e m^2 corresponde a distância

máxima entre cores no cluster [?]. Para o cálculo da distância em imagens em escala cinza, é utilizada a distância Euclidiana.

Uma etapa extra no processo de pós processamento é a união de *pixels orfãos*. Esses pixels são adicionados ao cluster mais próximo usando o algoritmo de componentes conexos [?].

Devido a limitação do espaço de pesquisa do algoritmo SLIC, a complexidade do algoritmo é $O(n)$, enquanto outros algoritmos que utilizam k-means para segmentação tem custo $O(k^N)$ [?].

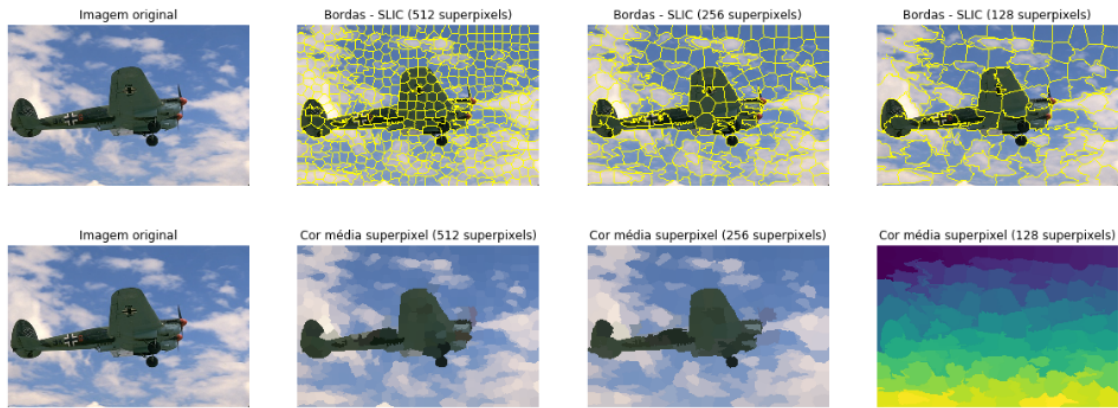


Figura 3. Fronteiras e coloração pelo valor médio dos superpixels SLIC/SLICO, para diferentes quantidades de superpixels

2.1.2. Superpixels EGB

Os superpixels EGB (*Efficient Graph-Based Image Segmentation*) utilizam uma abordagem baseada em grafos não direcionados. Nessa abordagem, cada pixel corresponde a um nó do grafo e a ligação entre eles ocorre por meio de arestas, com pesos não negativos, correspondente a medida de dissimilaridade [?].

Na abordagem baseada em grafos, a segmentação S é uma partição dos vértices V em componentes, no qual cada região $C \in S$ corresponde a um componente conectado em um grafo $G' = (V, E')$, onde $E' \subseteq E$, ou seja, a segmentação é induzida por um conjunto de vértices em arestas E [?].

No algoritmo foi definido um predicado D para avaliação da evidência de bordas entre dois componentes de uma segmentação. O algoritmo avalia a dissimilaridade entre elementos de dois componentes e os compara com elementos vizinhos em um mesmo componente, de modo que o algoritmo possa se adaptar em relação as características dos dados [?].

Para a comparação entre as regiões é utilizada uma função de corte (*threshold*) τ , a fim de medir o grau de diferença entre os componentes. Esse grau deve ser superior a diferença interna mínima, evidenciando uma borda [?]. A função de corte, no algoritmo, é utilizada baseado no tamanho do componente $\tau = k/|C|$, onde $|C|$ corresponde ao tamanho do componente C e k corresponde a um parâmetro do algoritmo [?].

O algoritmo EGB, utilizando pesos inteiros e ordenação por contagem, pode ser executado com custo linear, com complexidade $O(n \log n)$, para qualquer método de ordenação [?].

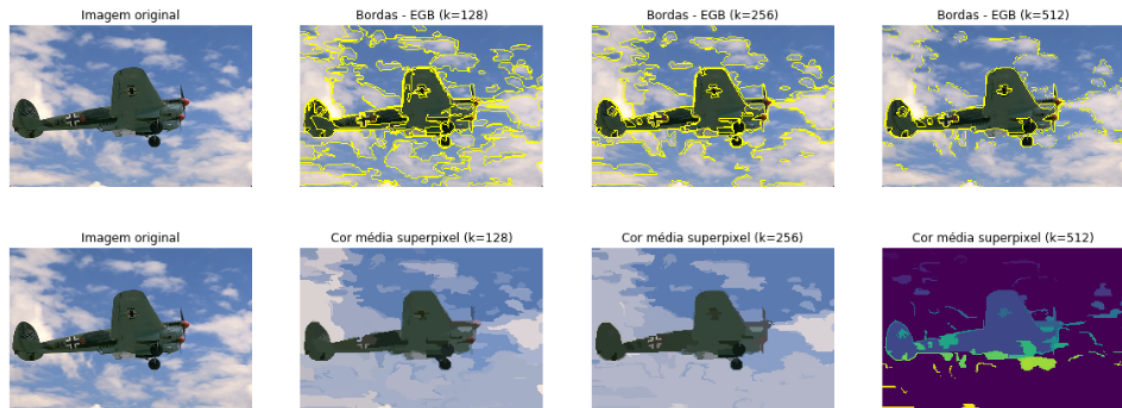


Figura 4. Fronteiras e coloração pelo valor médio dos superpixels EGB, para diferentes quantidades de superpixels

2.2. Clusters

Análise de cluster é a tarefa de agrupar conjunto de objetos de forma que um grupo (cluster) tenha características em comuns a outros grupos, ou clusters [?]. Clusters hierárquicos é um método de análise de clusters a fim de buscar hierarquias entre eles. As estratégias para hierarquização de cluster se dividem em dois grupos [?]:

- Aglomerativo - abordagem "*bottom up*", em que a observação inicia-se no próprio cluster e os pares de clusters são unidos na medida em que se sobe na hierarquia.
- Divisivo - abordagem "*top down*" onde as observações iniciam-se em um cluster e são realizadas recursivamente a medida em que se desce na hierarquia.

Os algoritmos originais de cluster possuem complexidade $O(n^3)$. Alguns algoritmos, entretanto, como o SLINK, ou *single-linkage*, possuem complexidade $O(n^2)$ [?]. O SLINK utiliza como critério de ligação (*linkage*) entre os clusters a distância mínima. A fórmula $\min\{d(a, b) : a \in A, b \in B\}$, corresponde a distância mínima para dois pares de clusters A e B observados, onde d é a métrica de distância escolhida [?].

A partir da aglomeração de clusters é possível construir uma árvore hierárquica, onde os clusters são agrupados por sua similaridade. Um método de visualização dessas características é o dendrograma.

2.3. Segmentação de Imagens

Segmentação de imagens consiste em separar uma imagem em regiões, idealmente correspondente a objetos reais [?]. Esse passo é utilizado como passo de processamento de imagens, vídeos e aplicações de visão computacional. Também consiste em um importante passo na tentativa de explicar uma imagem por meio de algoritmos [?].

Extensiva pesquisa é realizada e muitas abordagens e algoritmos são utilizados, com bons resultados para um conjunto ou classes de imagens [?]. A fim de facilitar a

pesquisa, alguns trabalhos foram desenvolvidos para criação de um conjunto de imagens com suas respectivas segmentações manuais. A base BSDS500 (*Berkeley Segmentation Data Set and Benchmarks 500*) provê uma base de imagens para pesquisa em segmentação e detecção de bordas. [?]. Essa base de dados é uma extensão da base BSDS300, com 200 novas imagens para avaliação [?].

A fim de avaliar a efetividade das segmentações, tradicionalmente são utilizados métodos subjetivos, como a visualização humana, responsável por comparar a qualidade da segmentação ou métodos supervisionados, onde uma segmentação é comparada em relação a uma imagem manualmente segmentada [?]. Projetos de detecção automática, como o SEISM (*Supervised Evaluation of Image Segmentation Methods*) permitem a avaliação da segmentação, usando a recuperação de precisão para bordas e a recuperação de precisão para objetos e peças [?]

3. Materiais e Métodos

Para confecção do trabalho foram escolhidos os algoritmos de *oversegmentation* SLIC e EGB. Os algoritmos possuem características diferentes: o SLIC é capaz de produzir superpixels em formas regulares, porém não é tão preciso ao separar os conjuntos por similaridade; por outro lado, o EGB, produz pixels irregulares, porém é mais aderente às diferenças de cores entre pixels.

Nos testes realizados foi realizada hierarquização dos resultados obtidos após a geração de superpixels SLIC e também sobre a composição de superpixels SLIC+EGB. Os resultados das hierarquizações foram comparados aos melhores resultados de segmentação obtidos pela segmentação SLIC, EGB e SLIC+EGB sem hierarquização. O método de hierarquização está descrito na seção ??, enquanto o método de avaliação de resultados está descrito na seção ?. Os resultados dos testes estão disponíveis na seção ?.

3.1. Hierarquia de Segmentação

3.1.1. Hierarquia Segmentação SLIC

A hierarquização de partições SLIC tem como objetivo possibilitar a utilização do algoritmo para produzir segmentação em diferentes níveis de detalhamento: desde um nível mais detalhado até um nível macro. Os diferentes níveis hierárquicos podem ser combinados em termos do contorno das hierarquias, permitindo a representação das hierarquias indexadas como bordas suaves, chamadas Mapas de Contorno Ultramétrico (UCM - *Ultrametric Contour Map*) [?].

Para construção das hierarquias foram gerados *clusters* com as correlações entre os grupos de superpixels, pós segmentação inicial SLIC. Foi utilizado o algoritmo SLINK (*single-linkage*) para agrupamento dos superpixels próximos, gerando uma árvore com as correlações entre os superpixels.

Para verificação da estrutura da árvore foi utilizado um dendrograma, como representado na figura ?. Dendrograma é um tipo de árvore utilizado para ilustração de uma clusterização hierárquica [?]. Frequentemente é utilizado nas áreas de biologia para indicar a mudanças evolucionárias entre ancestrais e descendentes, baseados em suas características [?].

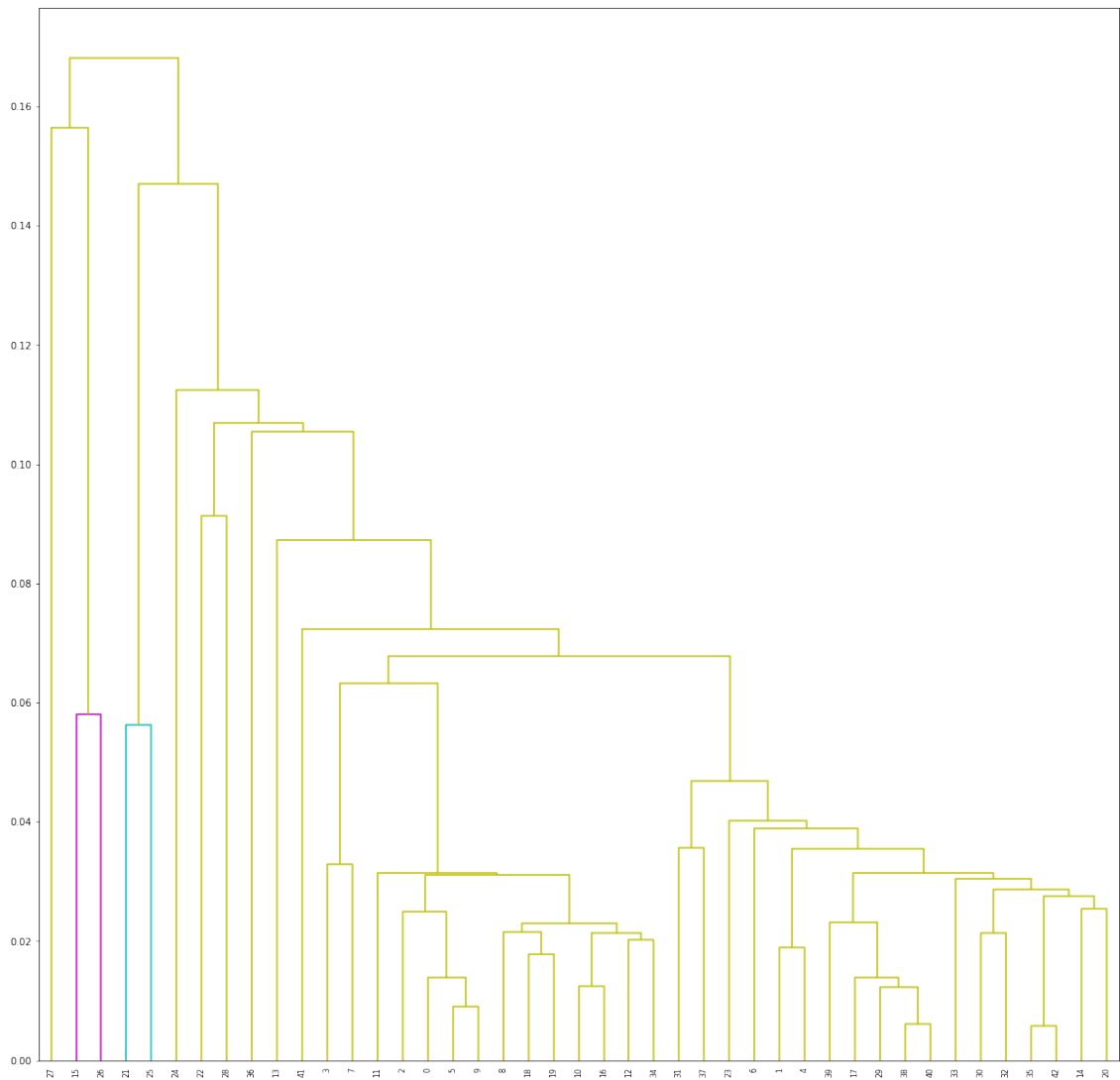


Figura 5. Dendrograma com a correlação entre os superpixels

Em um dendrograma é possível realizar cortes horizontais ou verticais na hierarquia, a fim de obter tipos de agrupamentos diferentes. No trabalho realizado somente foram realizados cortes horizontais. Os cortes foram feitos em diversos níveis da hierarquia.

Devido às características da segmentação, quando comparados ao *groundtruth*, não foram validados todos os níveis hierárquicos, somente níveis intermediários que apresentaram melhores resultados em testes empíricos. Com isso, foi estabelecido um nível máximo e mínimo para os cortes na hierarquia, de modo que os cortes foram feitos dentro de um intervalo, evitando nós com baixa probabilidade adequação ao *groundtruth*.

3.1.2. Hieraquia Segmentação SLIC+EGB

Uma outra abordagem realizada nesse trabalho foi a segmentação conjunta utilizando o método SLIC e em sequência, o método EGB. A utilização de ambos os algoritmos busca

obter um melhor desempenho em relação a utilização dos algoritmos de forma isolada.

O primeiro algoritmo aplicado às imagens foi o algoritmo SLIC, gerando superpixels de formas regulares. Em seguida, foi realizada recoloração da imagem utilizando o valor do pixel médio do superpixel, a fim de tornar a imagem mais simples, com características mais evidentes. Em seguida, foi aplicado às imagens o método EGB.

A partir do resultado do processo conjunto de segmentação, foi realizado o mesmo procedimento de hierarquização utilizado para o algoritmo SLIC. Nele, utilizou-se o algoritmo SLINK (*single-linkage*) para agrupamento dos superpixels próximos. Em seguidas cortes horizontais foram feitos na hierarquia, gerando diferentes níveis de visualização.

3.2. Análise Multiescala

Para que os níveis hierarquicos possam garantir a manutenção da informação, agrupando apenas características semelhantes, as hierarquias devem seguir os princípios de análise multiescala [?]. Esses princípios asseguram manutenção de duas características principais [?]:

- Causalidade - um contorno presente em uma escala $k1$ deve estar presente em qualquer escala $k2 < k1$ [?];
- Localidade - a medida em que o número de regiões diminui, os contornos devem ser estáveis (não devem se mover). A união corresponde a manutenção das bordas dos grupos que se fundiram [?].

No método de hierarquização descrito nesse trabalho manteve as características da análise multiescala. A figura ?? mostra os possíveis cortes na hierarquia de uma segmentação SLIC+EGB, sem recoloração da imagem. Nesse modelo, a segmentação é capaz de unir pixels por similaridade.

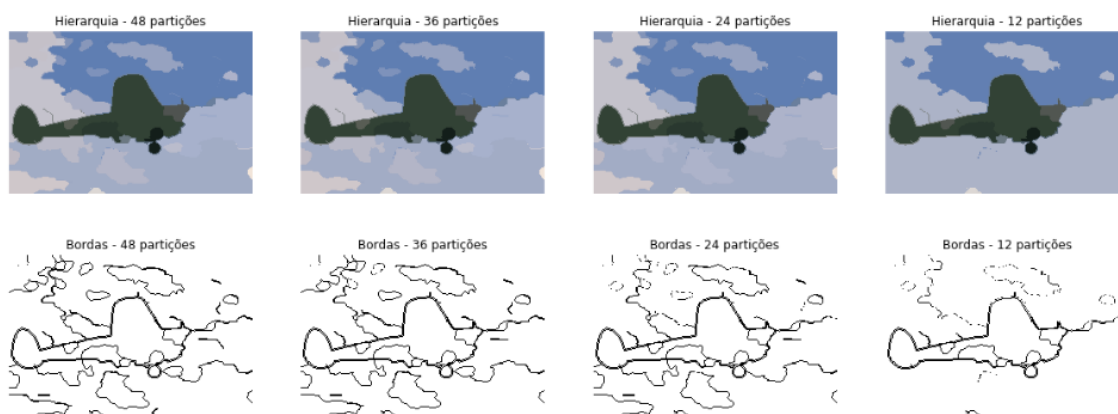


Figura 6. Hierarquia de partições utilizando superpixel SLIC+EGB.

É importante salientar que os métodos originais (SLIC e EGB), quando apenas variados os parâmetros de criação de superpixels não mantém as características de análise multiescala, conforme podemos ver nas figuras ?? e ??.

3.3. Avaliação de Resultados

Para avaliação das imagens, foi primeiramente utilizada a classificação visual, comparando a qualidade dos resultados. Em seguida utilizou-se um método de comparação de precisão e revocação.

O método de precisão é um método de classificação binária, onde a precisão (precision) corresponde à “fração de instâncias recuperadas que são relevantes”, enquanto revocação (*recall*), corresponde “a fração de instâncias relevantes que são recuperadas” [?]. Ambos os métodos são avaliados juntos para que possam identificar 4 tipos de valores possíveis:

- Verdadeiros Positivos - correspondem aos valores que foram corretamente classificados como positivos;
- Falsos Negativos - correspondem aos valores que foram classificados incorretamente como negativos;
- Falsos Positivos - correspondem aos valores que foram classificados incorretamente como positivos;
- Verdadeiros Negativos - correspondem aos valores que foram classificados corretamente como negativos [?];

A classificação descrita acima está ilustrada na figura ??.

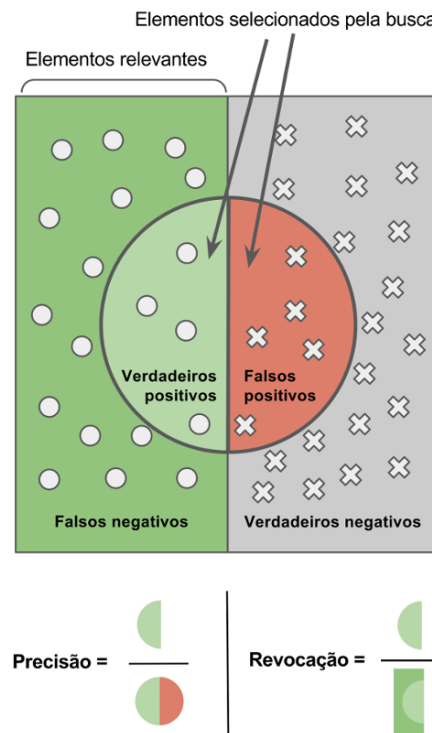


Figura 7. Precisão e revocação. Adaptado de [?]

A classificação de precisão e revocação pode ser dada por uma média harmônica, chamada de *F-measure* ou *F-score* balanceada. Essa medida é dada pela fórmula [?]:

$$F = 2 \cdot \frac{precis \cdot revoc}{precis + revoc} \quad (1)$$

A medida é aproximadamente a média quando seus valores estão próximos, porém o valor é baixo quando as médias estão distantes, favorecendo métodos com baixo número de falsos positivos e verdadeiros negativos [?]

3.4. Código Fonte e Bibliotecas Utilizados

Os algoritmos SLIC e EGB utilizados para confecção desse trabalho foram obtidos pela biblioteca Scikit-Image¹. Os algoritmos de geração de cluster SLINK e representação de dendrogramas foram obtido na biblioteca Scipy.org². O algoritmo de avaliação do resultado, utilizando método de previsão e revocação foi obtido no projeto Image-Segmentation, no Github³.

Os códigos fontes gerados, realizando as segmentações, hierarquias e avaliação dos resultados estão disponíveis publicamente na página pessoal do autor no Github ⁴.

4. Testes, Resultados e Discussões

Para avaliação do desempenho dos algoritmos foi executada a segmentação das imagens do conjunto de treinamento da base de dados BSD500 (*Berkeley Segmentation Data Set*). O conjunto possui 200 imagens naturais com seus respectivos *ground-truth* feito por anotações humanas [?].

Os algoritmos foram aplicados a base de testes e comparados com os *ground-truths* disponíveis na base de dados. A adequação do algoritmo em relação ao *ground-truth* foi feita por meio do mecanismo de precisão e revocação, utilizando a medida *F-measure*, conforme descrito na seção ??.

5. Conclusão

Os algoritmos de geração de superpixels avaliados nesse trabalho tiveram resultados baixos para tarefas de segmentação e detecção de contornos, mesmo quando combinados. As versões hierarquicas dos algoritmos apresentou resultados superiores àqueles obtidos sem hierarquias.

O algoritmo SLIC, apesar das vantagens de produzir superpixels de tamanhos regulares, apresentou a menor eficiência para segmentação de imagens, em relação aos demais algoritmos estudados. A combinação dos algoritmos SLIC+EGB para detecção de contornos não apresentou resultado ligeiramente superior em relação ao algoritmo EGB utilizado separadamente. Além disso, o custo computacional para execução dos dois algoritmos em sequência, bem como os custos para recoloração da imagens mostraram-se injustificáveis frente aos ganhos obtidos.

A utilização de hierarquias após a aplicação dos métodos de geração de *superpixels* aumentou consideravelmente os resultados dos métodos. Apesar da utilização não resultar em alto grau de desempenho na detecção de bordas e segmentação, justifica-se a utilização das versões hierárquias em relação às versões originais dos algoritmos quando houver necessidade de *oversegmentation*. Os parâmetros descritos nesse trabalho podem ser utilizados como passo inicial para refinamento dos parâmetros em trabalhos futuros.

¹<http://scikit-image.org/docs/dev/api/skimimage.segmentation.html>

²<https://docs.scipy.org/doc/scipy-0.18.1/reference/generated/scipy.cluster.hierarchy.linkage.html>

³https://github.com/Htiango/Image-Segmentation/blob/master/main/eval_boundary.py

⁴<https://github.com/falreis/image-segm>

Os métodos hierarquicos apresentados no trabalho também permitem a utilização em diferentes cenários, principalmente para redução e agrupamento de informação relevante em imagens, atendendo aos preceitos de análise multiescala.

Como trabalho futuro sugere-se a utilização dos algoritmos mostrados nesse trabalho como passo de pré processamento de outras técnicas de segmentação, como redes neurais convolucionais. Esses passos de segmentação iniciais podem permitir treinamento mais rápido e menor quantidade de informação a ser processadas, além de identificação mais rápida de bordas e contornos.