

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) [Arbelaez et al. 2011] 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) [Arbelaez et al. 2011]. 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 [Domínguez and Morales 2016]. Nessa tarefa, tomamos os *pixels* como unidades básicas de processamento [Wang et al. 2017]. O agrupamento de pixels em unidades maiores permite um tipo de segmentação chamado de *oversegmentation* [Wang et al. 2017], ilustrado na figura 1. 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.

O presente trabalho tem como objetivo investigar a utilização de superpixels como passo de pré processamento para segmentação de imagens. Esse trabalho investiga se a utilização de métodos de *oversegmentation* podem facilitar ou dificultar o processo de

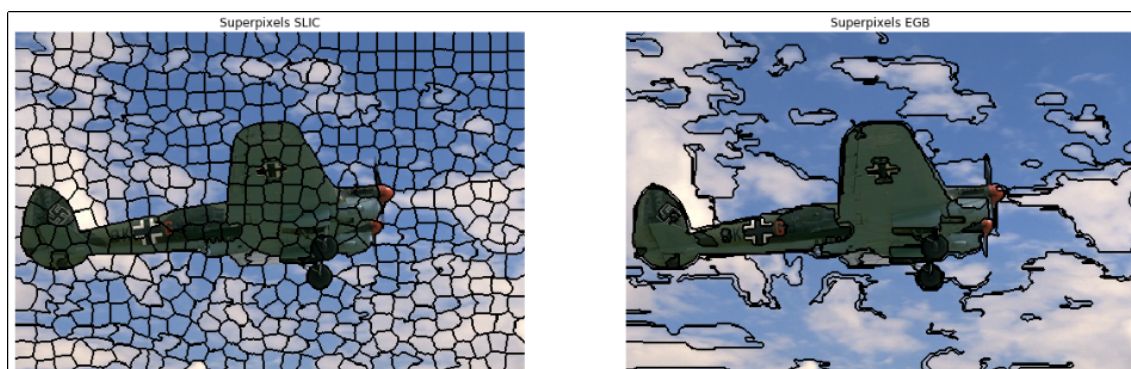


Figura 1. Imagens segmentadas utilizando superpixels SLIC e EGB

treinamento. 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 2 mostra o referencial teórico para construção do trabalho, a Seção 3, exibe os materiais e métodos utilizados nos testes; a Seção 4 mostra os resultados obtidos nos testes realizados e a discussões dos mesmos; a Seção 5 contém a conclusão do artigo, com as considerações finais.

2. Referencial Teórico

2.1. Superpixels

Superpixels, segundo [Felzenszwalb and Huttenlocher 2004], 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 [Achanta et al. 2012]. 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) [Felzenszwalb and Huttenlocher 2004];
- *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* [Vincent and Soille 1991] [Achanta et al. 2012]
- *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*) [Achanta et al. 2012]

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 3x3, evitando que superpixels sejam centrados nas bordas ou em uma posição de ruído [Achanta et al. 2012]. Em seguida, cada pixel é associado com o centro do cluster mais próximo, de modo que a região de busca se sobreponham [Achanta et al. 2012]. 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 x 2S [Achanta et al. 2012]. Em seguida, um passo de atualização dos centros dos clusters e computa o erro residual E [Achanta et al. 2012]. O algoritmo disponível na figura 2, 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 [Achanta et al. 2012]

Para o algoritmo descrito na figura 2, 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 imagem. 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 utilizada 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 [Achanta et al. 2012]. 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 [Achanta et al. 2012].

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 $O(k^N)$ [Achanta et al. 2012].

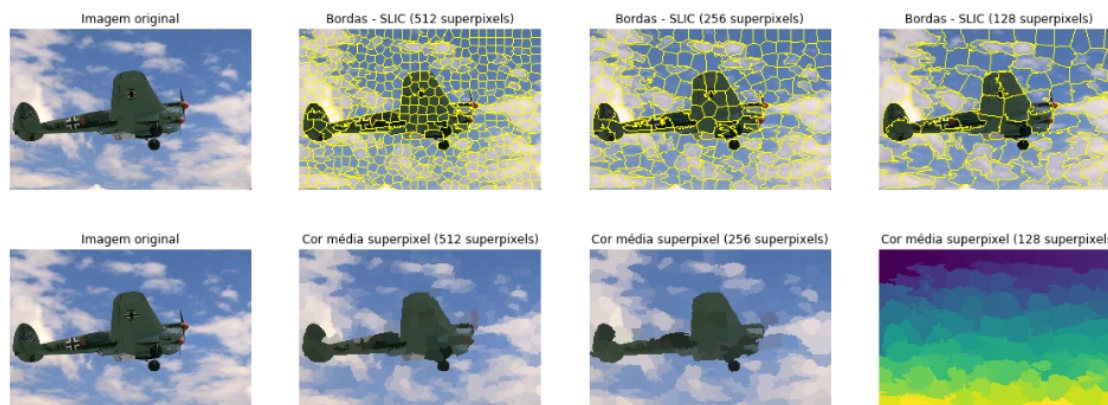


Figura 3. Fronteiras e coloração pelo valor médio dos superpixels SLIC/SLICO, para diferentes quantidades de superpixels

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 [Felzenszwalb and Huttenlocher 2004].

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 [Felzenszwalb and Huttenlocher 2004].

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 [Felzenszwalb and Huttenlocher 2004].

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 [Felzenszwalb and Huttenlocher 2004].

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 [Wikipedia 2018a]. 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 [Rokach and Maimon 2005]:

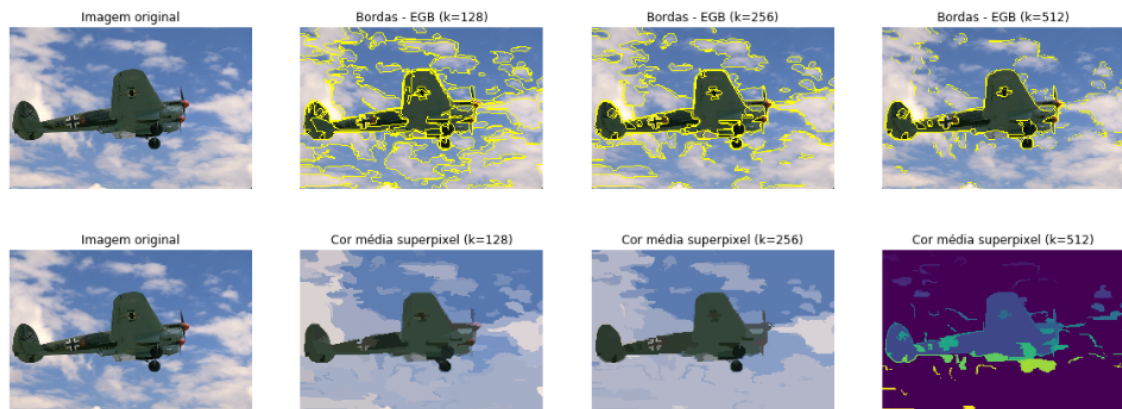


Figura 4. Fronteiras e coloração pelo valor médio dos superpixels EGB, para diferentes quantidades de superpixels

- 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 $O(n^2)$ [Sibson 1973]. 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 [Wikipedia 2018c].

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 [Zhang et al. 2008]. 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 [Zhang et al. 2008].

Extensiva pesquisa é realizada e muitas abordagens e algoritmos são utilizados, com bons resultados para um conjunto ou classes de imagens [Zhang et al. 2008]. 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. [Arbelaez et al. 2011]. Essa base de dados é uma extensão da base BSDS300, com 200 novas imagens para avaliação [Arbelaez et al. 2011].

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 [Zhang et al. 2008]. 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 [Pont-Tuset and Marques 2016]

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.

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*) [Arbelaez 2006].

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 5. Dendrograma é um tipo de árvore utilizado para ilustração de uma clusterização hierárquica [Wikipedia 2018b]. Frequentemente é utilizado nas áreas de biologia para indicar as mudanças evolucionárias entre ancestrais e descendentes, baseados em suas características [Science Dictionary 2018].

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*, em uma abordagem emírica, não foram validados todos os níveis hierárquicos, somente níveis intermediários que apresentaram melhores resultados. 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. Hierarquia Segmentação SLIC+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

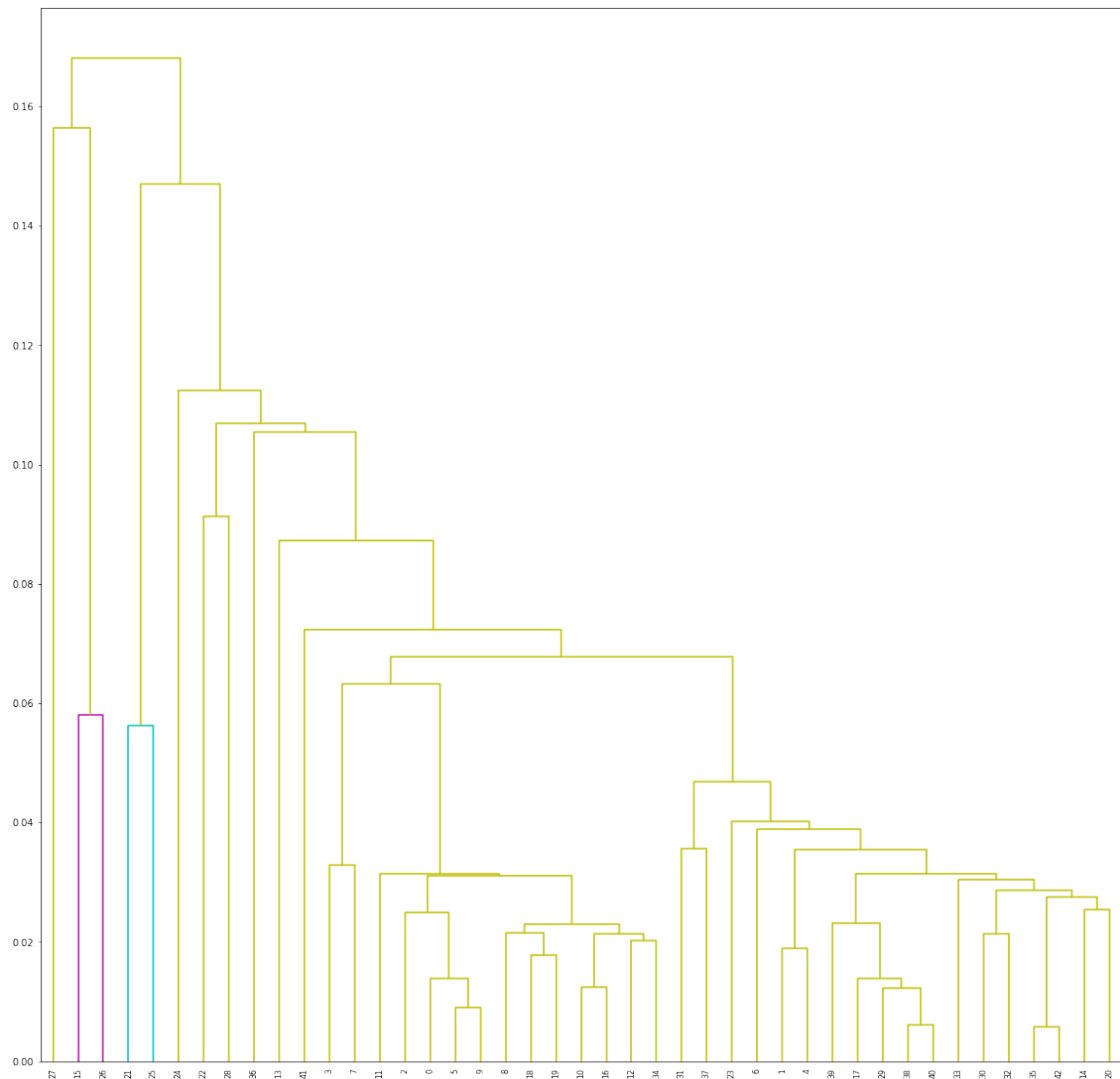


Figura 5. Dendrograma com a correlação entre os superpixels

recoloração da imagem utilizando o valor do pixel médio do superpixel. Em seguida, foi aplicado às imagens o método EGB.

Os resultados dos algoritmo utilizando a composição entre os métodos SLIC e EGB foram comparados aos melhores resultados obtidos para os algoritmos executados separadamente. O método descrito foi aplicado à base de dados BSD500 e comparados com as imagens manualmente segmentadas (*ground-truth*).

3.2. 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 (textitprediction) corresponde à “fração de instâncias recuperadas que são relevantes”, enquanto revocação (*recall*), corresponde “a fração de instâncias relevantes que são recu-



Figura 6. Fronteiras das Hierarquia de partições utilizando superpixel SLICO.

peradas” [Wikipedia 2018d]. 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 [Wikipedia 2018d];

A classificação descrita acima está ilustrada na figura 7.

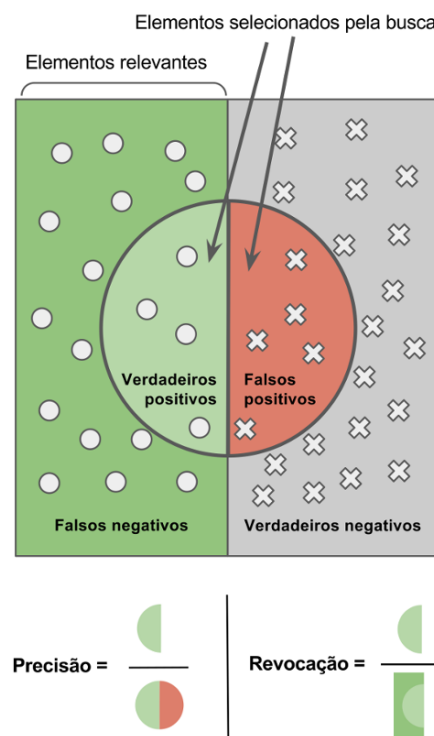


Figura 7. Precisão e revocação. Adaptado de [Wikipedia 2018d]

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 [Wikipedia 2018d]:

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

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 [Wikipedia 2018d]

3.3. 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 Github⁴.

4. Testes, Resultados e Discussões

5. Conclusão

Referências

- Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., and Süsstrunk, S. (2012). Slic superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11):2274–2282.
- Arbelaez, P. (2006). Boundary extraction in natural images using ultrametric contour maps. In *2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'06)*, pages 182–182.
- Arbelaez, P., Maire, M., Fowlkes, C., and Malik, J. (2011). Contour detection and hierarchical image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(5):898–916.
- Domínguez, D. and Morales, R. R. (2016). *Image Segmentation: Advances*, volume 1.
- Felzenszwalb, P. F. and Huttenlocher, D. P. (2004). Efficient graph-based image segmentation. *International Journal of Computer Vision*, 59(2):167–181.
- Pont-Tuset, J. and Marques, F. (2016). Supervised evaluation of image segmentation and object proposal techniques. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 38(7):1465–1478.
- Rokach, L. and Maimon, O. (2005). *Clustering Methods*, pages 321–352. Springer US, Boston, MA.

¹<http://scikit-image.org/docs/dev/api/skimage.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>

Science Dictionary (2018). What is dendrogram?

Sibson, R. (1973). Slink: An optimally efficient algorithm for the single-link cluster method. *Comput. J.*, 16:30–34.

Vincent, L. and Soille, P. (1991). Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(6):583–598.

Wang, M., Liu, X., Gao, Y., Ma, X., and Soomro, N. Q. (2017). Superpixel segmentation: A benchmark. *Signal Processing: Image Communication*, 56:28 – 39.

Wikipedia (2018a). Cluster analysis.

Wikipedia (2018b). Dendrogram.

Wikipedia (2018c). Hierarchical clustering.

Wikipedia (2018d). Precision and recall.

Zhang, H., Fritts, J. E., and Goldman, S. A. (2008). Image segmentation evaluation: A survey of unsupervised methods. *Computer Vision and Image Understanding*, 110(2):260 – 280.