# Clustering segmentation using different methods and color spaces

Saul Gómez
Universidad de Los Andes
sc.gomez11@uniandes.edu.co

Diego Valderrama
Universidad de Los Andes
df.valderrama@uniandes.edu.co

## Abstract

*Image segmentation is an important task in computer vision. Many algorithm have been developed in order to address this problem. In the current study, four different algorithms are analyzed and compared in order to determined the better performance in different color spaces. Watershed algorithm implemented in RGB color space obtained the higher performance among the others, however, this value was not as high as expected, so the limitations of the algorithm are described and discussed*

## 1. Introduction

Understanding the image and extracting information from the image to accomplish some works is an important area of application in computer vision [6]. Often it is not necessary to focus on the image as a whole, but only for some certain areas which provides us information of interest. For this reason, image segmentation has become a crucial task in image processing and pattern recognition. It is a process based on certain criteria to divide an input image into different regions in order to extract the area we interested in [6], [3].

A broad family of approaches to segmentation involve integrating features such as brightness, color, or texture over local image patches and then clustering those features based on, such as, fitting mixture models mode-finding, or graph partitioning [1]. Based on this, there are many commonly used algorithms for image segmentation, among the main we can find the threshold segmentation, which is based on the determination of an optimal threshold to differentiate the structures of the background, the regional growth segmentation, whose main idea is to group pixels that in similar characteristics in regions, or the segmentation by clustering algorithm, based on the similarity between things as the criterion of class division, that is, it is divided into several subclasses according to the internal structure of the sample set, so that the same kind of samples are as similar as possible, and the different are not as similar as possible, and many others more [6].

Likewise, image segmentation plays a crucial role in many medical-imaging applications, by automating or facilitating the delineation of anatomical structures and other regions of interest [4]. Therefore, it is necessary to identify the methods that present a higher performance when segmenting images with the objective of developing optimal algorithms that present meaningful results when developing this task. Keeping this in mind, the current study employed a fraction of Berkeley's Segmentation Dataset (BSDS500) and made a comparison of different clustering methods, with the aim of implementing an algorithm that would have the best performance in image segmentation. The results were presented, the relevant analysis of the results was carried out and concluded in the matter.

## 2. Materials & Methods

### 2.1. Clustering algorithms

#### 2.1.1 K-Means

The K-Means algorithm clusters data by trying to separate samples in a certain number of groups of equal variance, minimizing a criterion known as the inertia or within-cluster sum-of-squares. The algorithm has three steps: The first step chooses the initial centroids randomly. Then, assigns each sample to its nearest centroid. Further, creates new centroids by taking the mean value of all of the samples assigned to each previous centroid. The difference between the old and the new centroids are computed and the algorithm repeats these last two steps until this value is less than a threshold, that is, until the centroids have a significant displacement [2].

#### 2.1.2 Gaussian Mixture Model

In mixtures of Gaussians model (GMM), each cluster center is augmented by a covariance matrix whose values are re-estimated from the corresponding samples. Instead of using nearest neighbors to associate input samples with cluster centers, a Mahalanobis distance is used. Additionally, samples are softly assigned to several nearby clusters, that is, a pixel has an associated probability to belong to one cluster or another, and that probability has a Gaussian distribution around each of the centroids of the clusters [5].

#### 2.1.3 Hierarchical

Hierarchical clustering is a general family of clustering algorithms that build nested clusters by merging or splitting the image into successively finer regions. This process is repeated until regions are either fairly uniform or below a certain size [5].

#### 2.1.4 Watershed

This technique segments an image into several catchment basins, which are the regions of an image, interpreted as a height field, where rain would flow into the same lake. An efficient way to compute such regions is to start flooding the landscape at all of the local minima and to label ridges wherever differently evolving components meet. Unfortunately, watershed segmentation associates a unique region with each local minimum, which can lead to over-segmentation. Therefore, it is often necessary to first marks seed locations that correspond to the centers of different desired components, process commonly named minimum imposition [5].

### 2.2. Image pre-processing

In order to maximize the performance of the image segmentation algorithm to be developed, a normalization was performed in the different color spaces, mainly those containing spatial coordinates. For this, x and y channels were assigned less weight, compared to the original channels of the color space, making the segmentation mainly based on the intensity of the pixel. On the other hand, taking into account the computational resources required for the execution of hierarchical clustering algorithms, a reprogramming of the images equivalent to 0.125 of the original image was implemented. Therefore, this pre-processing represents a factor to be considered when analyzing the performance of the segmentation algorithm because the resized image does not have the same information as the original image

### 2.3. Evaluation

The most evaluation metric used in a supervised segmentation problem with a single annotation is the Jaccard index [4]. For the case in which there are multiple annotations, the researcher must take a decision on how to carry out the evaluation form. This means that the researcher can choose to use only one annotation per image, or take the whole set of annotations, process them and obtain a metric. However, in unsupervised problems such as clustering, there is no established metric due to usually there is no a set of annotations.

For the evaluation of our algorithm, we have established a metric that uses the mathematical definition of the Jaccard index. In this metric, it is necessary to have one image for the annotation and other for segmentation. These images must be two 2D matrices composed of labels of the same size. Then, with these matrices the Jaccard index is calculated as the intersection divided the union of the two images.

### 2.4. Experiments

The experiments were made in order to obtain the hyperparameters with which the segmentation is optimum for each clustering method allowed us to select the best method. To do this, we change the number of clusters in all methods and some others parameters were taken into account according with clustering method

First of all, it was selected the first anotation of all of them for all of the images because of all images do not have the same number of human labels. In the next step, the number of clusters were varyed for all method in order to find the value that maximizes the performance. Once the optimal number of clusters was set, some hyperparameters were changed. With GMM the covariance type and the number of maximum iterations before the convergence of the method was varyed. The last hyperparameter mentioned also was optimized for k-means. Furthermore, the linkage parameter was varyed using herarchical method and the numbers of peaks per label and compactness that optimize watershed method was found. Finally, we select the best method to evaluate it in differente color spaces.

### 2.5. Algorithm

The first step was to load the images and the training labels. Then, the train set was segmented with the four methods of clustering mentioned above. To do this, the hyperparameters found with the experiments was used. In the next step, the results of the segmentations was evaluated. Finally, test set was segmented and evaluated using watershed.

## 3. Results

First, it was selected the optimal number of clusters for the four clustering methods used in the rgb color space keeping the other parameters by default. According to the figures 1, 3, 6, 8 the values associated with the clusters are 2,2,13 and 48 for K-means, GMM, herarchical and watershed, respectively. Subsequently, different experiments were carried out for each method in order to find some hyperparameters that maximize performance. For K-means, the maximum number of iterations performed by the method before converging to a value was optimized. As can be seen in Fig 2, the value found was 162 iterations because of the jaccard index of approximately was obtained. For GMM, the first step was optimized the convergence type in which it was obtained that the covariance with better performance is the full type (see Fig 4.) Likewise, as in K-mean for this method the maximum number of iterations was also optimized and as shown in Fig 5 its value is 400 and the jaccard index associated was approximately 26%.

In the same way, an experiment was carried out for the hierarchical segmentation method in which the linkage type was optimized. According to the Fig 7 the linkage that obtained the highest Jaccard index (approximately 18 %) was the complete linkage so this type was selected as a hyperparameter. For watershed method the compactness and the number of peaks per label were varied and it was obtained that the values that maximize the performance of this method are 1 for the two hyperparameters (see Fig 9, 10). After comparing the performances obtained for all the clustering methods used, it was found that the method with a higher Jaccard index was watersheds with approximately 31 %. Due to this, this method was subjected to an additional experiment that consisted of varying the color spaces used. It means that this method with the hyperparameters found (48 clusters, 1 peak per label and compactness of 1) was evaluated in 6 different color spaces (rgb, lab, hsv, rgbxy, labxy, hsvxy).

Finally, according to the results presented in Fig 11 the jaccard index obtained in rgb is the highest, so these parameters (clusters, peaks per label, compactness and rgb as color space) were used to evaluate the algorithm in the test set where a jaccard index of approximately 32 % was obtained. Fig 12 shows a test image segmentation using our algorithm and Fig 13 shows its human anotation.

## 4. Discussion

Firstly, it is consistent to use different color spaces since RGB, although it is the most compatible with electronic devices, is limited only to color intensities and does not take into account aspects such as spatial location, saturation or brightness, which can provide relevant information when grouping pixels. However, as evidenced by the results obtained, the highest performance was presented using RGB color space.

Now, as mentioned above, RGB color space does not discriminate between channels, each being a different color intensity, so there is no one more discriminative than another, all offering the same information regarding an image. On the other hand, when analyzing the case of the spaces Lab and HSV, the opposite effect is presented as in RGB. For these two color spaces, there is a channel associated with luminescence while the others are associated with chromminicence, with the channel L for Lab representing the intensities of the objects and the channel V for HSV representing the brightness. These two channels are more discriminative than their counterparts a, b, H, and S, as they provide more information about the structures of the objects in the image. Also, by looking more closely at the HSV channel, the S channel has to be more discriminatory than the H channel, since the first one can offer specific information of the object associated with its composition, while the second one is related to the tonality of the image

According to the results obtained, the method of segmentation by Watershed in the RGB color space, presented the best performance for the set of images used in this study. This is because the nature of the algorithm contemplates a set of points representing the regional minimums of the image, which may be associated with the center of the objects. This can be understood in parallel with the definition of image segmentation addressed by Cheng *et al*., where it is a process of dividing an image into different regions such that each region is, but the union of any two adjacent regions is not, homogeneous [3]. This homogeneous regions corresponds to the catchment basins in watershed algorithm and, therefore, image features. Consequently, it represents an advantage compared to the other clustering methods used, which mechanism is based on the grouping of similar characteristics shared between pixels.

As mentioned in the results section and as evidenced in figures 12 and 13, the implemented algorithm is able to segment the images in regions similar to those annotated in the groundtruth, however, the allocation of the tags for the pixels of those regions varies significantly, which generates confusion in the used metric and decreases the performance of the algorithm. Additionally, since it is a segmentation problem and you do not know what you are looking for, the algorithm is not properly trained so its performance is considerably reduced. Similarly, by using a database with such varied images, the algorithm implemented is not in the ability to learn to recognize patterns in objects as they possess very different visual characteristics. All of these factors mentioned in conjunction with a metric that can be improved to better match the problem addressed and provide a better evaluation of the performance of the algorithm make up the set of conditions that seem to be associated with the low performance of the implemented algorithm

## 5. Conclusion

As noted, the problem of image segmentation is a difficult task to address because of the lack of information available in the algorithm in its training and what is expected of it is not clearly defined. In addition to the limitations presented in the previous sections, the need to employ an appropriate metric to perform the performance evaluation of the algorithm in the segmentation of images is highlighted. In order to improve the evaluation strategy and, therefore, the best method it is necessary to develop a mechanism that takes into account the results of segmentation as regions regardless of the tag assigned to the set of pixels that compose it, so that the end result is not affected by a mismatching between the segmented image and the groundtruth labels. A good strategy can be to consider the contours of the regions in which the image was divided and compare the overlapping of these borders with the borders of the annotations, thus obtaining a more accurate measurement of the accuracy of the algorithm.

## References


[1] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik. Contour detection and hierarchical image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(5):898–916, May 2011.

[2] D. Arthur and S. Vassilvitskii. k-means++: The advantages of careful seeding. In *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms, Society for Industrial and Applied Mathematics*, 2007.

[3] H. Cheng, X. Jiang, Y. Sun, and J. Wang. Color image segmentation: advances and prospects. *Pattern Recognition*, 34(12):2259 – 2281, 2001.

[4] D. L. Pham, C. Xu, and J. L. Prince. Current methods in medical image segmentation. *Annual Review of Biomedical Engineering*, 2(1):315–337, 2000. PMID: 11701515.

[5] R. Szeliski. *Computer vision*. Springer, 2011.

[6] S. Yuheng and Y. Hao. Image segmentation algorithms overview. *CoRR*, abs/1707.02051, 2017.


## 6. Images

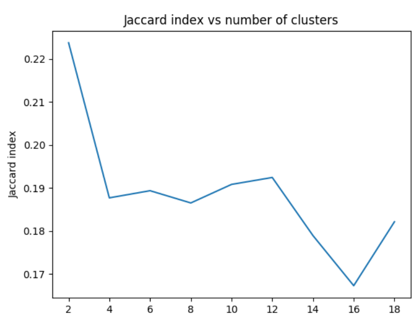


Figure 1. Metric performance in train set varying the number of clusters in rgb colorspace using K-means method

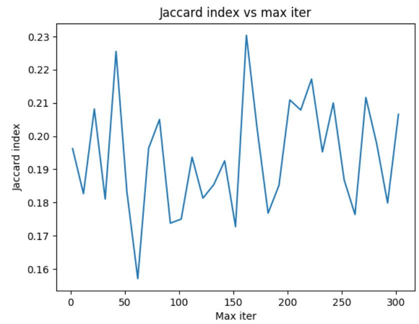


Figure 2. Metric performance in train set varying the maximum number of iterations before the converge of K-means method in rgb colorspace setting the number of clusters in 2

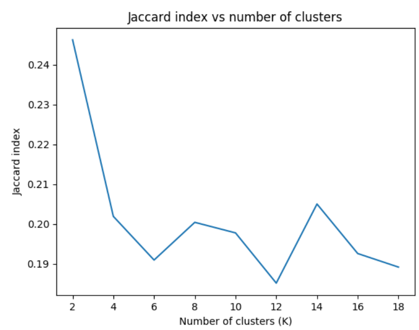


Figure 3. Metric performance in train set varying the number of clusters in rgb colorspace using GMM method

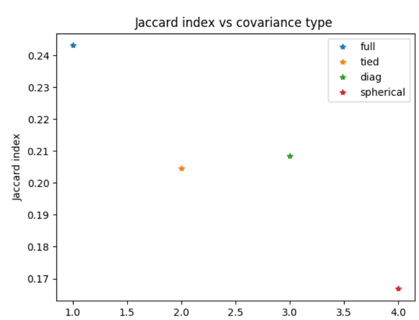


Figure 4. Metric performance in train set varying the covariance type in rgb colorspace setting the number of clusters in 2 using GMM method

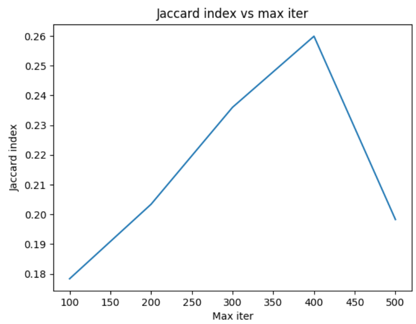


Figure 5. Metric performance in train set varying the maximum number of iterations before the converge of GMM method in rgb colorspace setting the number of clusters in 2 and covariance type as full

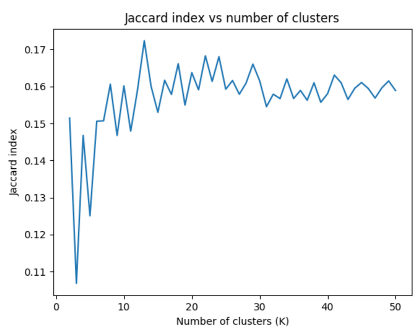


Figure 6. Metric performance in train set varying the number of clusters in rgb colorspace using hierarchical method

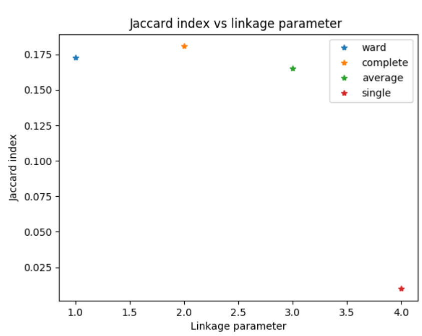


Figure 7. Metric performance in train set varying the linkage type in rgb colorspace setting the number of clusters in 13 using hierarchical method

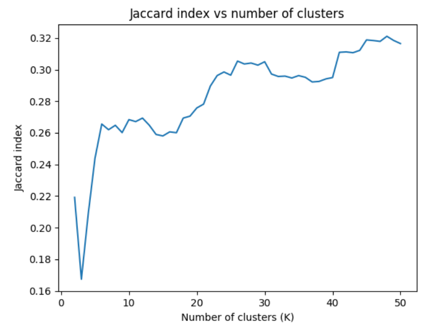


Figure 8. Metric performance in train set varying the number of clusters in rgb colorspace using Watershed method

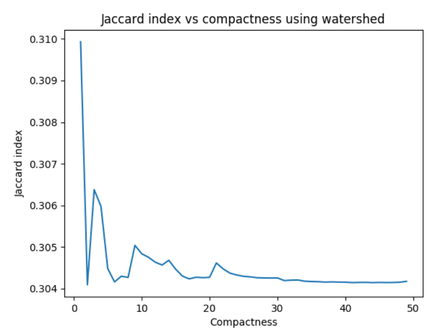


Figure 9. Metric performance in train set varying the compactness in rgb colorspace setting the number of clusters in 48 using watershed method

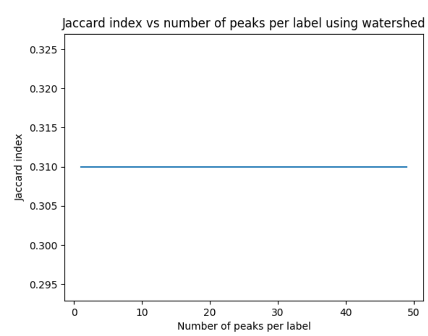


Figure 10. Metric performance in train set varying the compactness in rgb colorspace setting the number of clusters in 48 and the compactness in 1 using watershed method

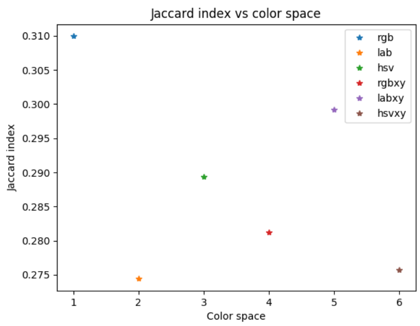


Figure 11. Metric performance in train set varying the colorspace setting the number of clusters in 48, the compactness in 1 and the peaks per label in 1 using watershed method

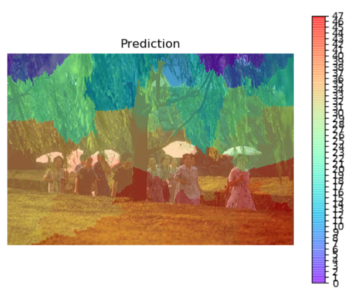


Figure 12. Image segmented of test set obtain with watershed method setting the number of clusters in 48, the compactness in 1 and the peaks per label in 1

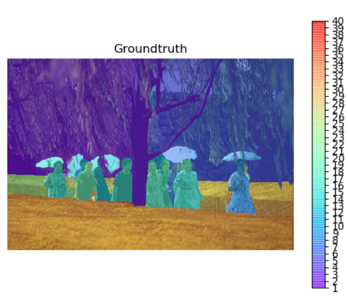


Figure 13. Test image anotation