Comparative Studies of Region-Based Segmentation Algorithms on Natural and Remote Sensing Images

Asim Shoaib1*, Mogana Vadiveloo2, and Seng Poh Lim3

1,2,3Department of Computer Science, Faculty of Information and Communication Technology, Universiti Tunku Abdul Rahman, Jalan Universiti, Bandar Barat, 31900 Kampar, Perak, Malaysia.

Abstract. Region-based segmentation algorithms are used as a preprocessing approach to generate over-segmented regions. Over-segmented regions refer to the creation of small regions in an image that represent no meaningful object regions. It has been observed that there are limited works on the performance comparison of the region-based segmentation algorithms on both natural and remote sensing (RS) images. Hence, the objective is to compare the performance of region-based segmentation algorithms on natural and RS images with different complexity of object regions of interest (ROIs). There are four algorithms (Felzenszwalb and Huttenlocher (FH), Quick Shift (QS), Compact Watershed (CW), and Simple Linear Iterative Clustering (SLIC)) being compared using two public datasets. The adapted rand error (ARE) and variation of information (VOI) are used for the segmentation evaluations. Generally, the experiments showed that the SLIC achieved better results as compared to the other algorithms for both images with different complexities of ROIs. This is mainly because the over-segmented regions produced by the SLIC adhered to the image object boundaries well than the over-segmented regions generated by other algorithms. However, CW achieved better average ARE than SLIC for RS images because CW has compactness and marker parameters which influence it to achieve better results.

1 Introduction

Image segmentation is the process of partitioning an image into non-intersecting regions [1]. These regions are grouped based on features of image pixels, such as grey level, texture, intensity, and colour, to form object regions of interest (ROIs) [2]. Many image segmentation algorithms rely on two fundamental characteristics which are discontinuity and similarity of the pixels feature values between the object regions in an image to perform the delineation [3]. Boundary- or edge-based algorithms are based on the pixels features values discontinuity. In contrast, region-based algorithms perform segmentation based on pixels feature similarity to build homogeneous and visually distinct regions [4].

*Corresponding author: asimshoaib@1utar.my

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).
Generally, the region-based segmentation algorithms start from the pixels with similar feature values and expands until it reaches the boundaries pixels forming the ROIs for delineation [5]. Though these algorithms delineate the ROIs in an image accurately, however they cause the over-segmentation. Over-segmentation is the generation of tiny regions that do not convey any meaningful segmented ROIs in an image [6 - 8].

Recently, researchers have tested four existing region-based algorithms performance in segmenting medical [9] and natural [10] images. These algorithms are the Felzenszwalb and Huttenlocher (FH) [11], Quick Shift (QS) [12], Compact Watershed (CW) [13], and SLIC [14]. The results show that SLIC algorithm has produced higher accuracy as compared to the other three mentioned algorithms. While in [15], the author has used remote sensing images and compared the performance among the Felzenszwalb and Huttenlocher (FH), Quick Shift (QS), and SLIC algorithms by using evaluation measures of adapted rand error (ARE). The author concluded that the SLIC algorithm achieved better average values of ARE as compared to the other algorithms and efficiently generated the over-segmented regions that adhered to the images object regions boundaries.

As demonstrated in [10] and [15], the works have compared the performance of the abovementioned region-based segmentation algorithms on either natural or remote sensing images. Further to elaborate, the study in [10] exclusively tested the efficacy of the proposed algorithms using only natural images, while the research conducted in [15] evaluated the performance of the proposed algorithms, solely with remote sensing images. Precisely, the performance of an algorithm is determined by the complexity of the ROIs in a chosen image for delineation [16]. Thus, the goal of this paper is to compare the performance among the Felzenszwalb and Huttenlocher (FH), Quick Shift (QS), Compact Watershed (CW), and Simple Linear Iterative Clustering (SLIC) algorithms on natural and remote sensing images. Generally, these two domains of images have different complexity of ROIs, where natural images have lower complexity [17] while remote sensing images have higher complexity of ROIs for segmentation [18].

The structure of this paper includes section 2, which discusses the chosen region-based segmentation algorithms, section 3 explains the chosen algorithms parameters along with the experimental setup. While section 4 includes the analysis and discussion of the results. Finally, section 5 concludes the paper with future works.

### 2 Region-Based Segmentation Algorithms

In this paper, four region-based segmentation algorithms that fall into four prominent categories are discussed. These categories are the graph-, density-, watershed-, and clustering-based segmentation algorithms. These algorithms were chosen since in [10] and [15], the authors have compared the performance of aforementioned region-based segmentation algorithms on natural or remote sensing images, respectively.

According to the author in [19], baseline of graph-based segmentation algorithms is that, an image is represented as a weighted undirected graph, or \( G = (V, E) \). \( E \) stands for edge between two pixels and \( V \) stands for the pixel values as nodes. Moreover, \( G \) is divided into the subgraphs \( A \) and \( B \). While on the basis of this graph theory, graph-based segmentation algorithms segment the images into desired regions. Felzenszwalb and Huttenlocher (FH) algorithm [11] is a graph-based segmentation algorithm, where an image is represented as an undirected weight graph. This FH algorithm performs segmentation based on pairwise region comparison. It relies on selecting edges from a graph where each pixel in the graph corresponds to a node and the neighbouring pixels are connected by undirected edges. Weights on each edge calculate the dissimilarity of the neighbouring pixels nodes from their feature values for merging. Furthermore, FH algorithm generates the over-segmented regions
that adhere to the image boundaries; however, these regions are irregular in size and shape [11].

As for Quick Shift (QS) [12] algorithm, it is a density-based segmentation algorithm, in which each pixel in the feature space is shifted to its closest neighbour for delineation. This shifting process is based on the weight derived from two neighbouring pixels in the feature space. The weight is the density estimation of the similarity and dissimilarity between the two neighbouring pixels. QS uses the 5-Dimension (5D) feature space that are of colour information and pixels position. In QS, a tree is created where all the pixels become the nodes. Then, these nodes are recovered by severing the tree branches that have larger weights than a pre-determined threshold to form segmented regions. Since in QS, the weight between two neighbouring pixels is the density estimation of 5D feature space among the pixels, therefore the over-segmented regions produced adhere to the image boundaries well. Furthermore, the authors in [20] mentioned four benefits of QS which include speed, universality, simplicity, and tuning the parameter that are used to balance the over- and under-segmentation regions generation.

Watershed-based segmentation algorithms, mainly starts with the labelled pixels as seeds while the neighbouring pixels are sequentially labelled with the various priorities. The determination of the seeds and the priorities definition for segmentation are the major factors that differentiate the watershed algorithms performances [21]. The watershed-based segmentation algorithm known as compact watershed (CW), was modified by [13]. The CW algorithm is an extension of marker-controlled watershed [22]. In [22], the authors presented the idea of marker-controlled watershed based on an ordered set of priority of queues. First, empty queues are considered for the image pixels. Then these pixels are inserted into a priority queue using the pixels’ intensity values. The pixels with higher intensity are placed in Q1 while pixels having lower intensity than Q1 are placed in Q2. This process iterates until all the pixels are inserted into the priority queues. By using this data structure, the pixel with lowest intensity value is chosen as the local minima and becomes the marker. This process repeats for all the queues and stops when all the pixels were taken into consideration. Then the non-marker pixels get assigned to the respective marker pixels based on the distance value between the non-marker pixels and the marker pixels. The distance value is determined from the pixels intensity value and each non-marker pixel is assigned to a marker pixel that has the shortest distance. The over-segmented regions generated from the mentioned marker-controlled watershed are in highly irregular size, shape, and have no control of compactness. Therefore in [13], the authors have used the data structure approach of the mentioned marker-controlled watershed by [22] and included the Euclidean distance function and compactness feature for assigning the non-marker pixels to the marker pixels. By this, the over-segmented regions generated are compact and regular in shape [13].

Furthermore, the simple linear iterative clustering (SLIC) algorithm [14] is the clustering-based algorithm. According to the authors in reference [23], SLIC is the adoption of k-means clustering algorithm. The algorithm works by assuming that an image has N pixels and is pre-segmented into R equal-sized regions. Each equal-sized region is approximately N/R pixels. The SLIC algorithm starts by initialising cluster centres values within an approximate equal-sized region. Then, it finds the similar pixels in 2 x 2 neighbourhood to assign these pixels to the respective cluster centres. The algorithm then computes the new cluster centres values. Thus, the cluster centre is derived at each iteration and computed from similar pixels that comprised of each roughly equal-sized region. The average distance between two cluster centres is defined as \( D = \sqrt{N/R} \). This process repeats iteratively until all the pixels and equal-sized regions are taken into consideration. Once all the pixels are traversed, the process stops. In SLIC, the clustering of pixels is based on the pixels intensity and is performed in 5D feature space. In addition, in [14], the authors stated that SLIC algorithm has only one parameter k to be defined. This parameter is the value of the number
of segments to be generated. Moreover, the Simple Linear Iterative Clustering (SLIC) algorithm generates the over-segmented regions that adhere to the object regions boundaries efficiently in an image. However, the right value of parameter $k$ depends on the specific application and implementation [14].

3 Experimental Setup

In this paper, the experiments were conducted using two public datasets images. They are natural images from the Berkeley Segmentation dataset (BSD300) [24] and remote sensing images from WHU Building dataset [25]. Total of 200 test images from the BSD300 dataset with the dimension of 481 x 321 and 100 test images from the WHU Building dataset with the dimension of 512 x 512 were chosen for the experiments. The experiments were conducted using Python V3.10.4 on Spyder IDE for qualitative evaluation and Google Colab for quantitative evaluation, with scikit-image [26] open-source library. Moreover, a laptop that has Intel Core i5 with Windows 10 operating system was used to conduct the experiments.

In addition, the parameters to be defined for FH are scale, sigma, and minimum size. While for QS, kernel size is to be defined as the parameter. In this paper, scale is set as 50, sigma as 0.5, and minimum size is defined as 50 [15]. Moreover, value 3 is used as the kernel size parameter value for QS [15].

As for SLIC, the parameter $k$ known as n_segments has to be defined. Moreover, for CW, the parameters are markers and compactness. In this paper, $k$ in SLIC is set as 100 [26]. While marker and compactness in CW are stated as 100 [16] and 0.001 [26], respectively. In addition, Sobel filter [27] is used to run the CW algorithm since generally, watershed algorithm requires a gradient image as an input image to perform the segmentation.

In this paper, two metrics were chosen for performance evaluation, which are adapted rand error (ARE) and variation of information (VOI). These metrics were selected for comparison of the segmentation results in relation to the ground truth (GT) images. Moreover, these metrics allow the comparison in terms of object boundaries error and overall information gain or loss during segmentation process [15]. They are described as follows where equation (1) refers to adapted rand error (ARE) while equation (2) is the variation of information (VOI):

$$Adapted\ Rand\ Error = 1 - \frac{2pr}{p+r}$$  \hspace{1cm} (1)

where $p$ and $r$ are the precision and recall. Precision is referred as the positive predictive value while recall is the true positive rate of segmented regions in relation to the ground truth image [15]. Adapted rand error value of 0 indicates the perfect segmentation in relation to the ground truth (GT) [28].

$$VOI = FalseSplit + FalseMerge$$  \hspace{1cm} (2)

where FalseSplit are the pixels missed, while FalseMerge refers to the pixels that are merged incorrectly during the segmentation [29]. The lower is the value of VOI, the better is the segmentation result obtained [30].

4 Analysis and Discussion

Fig. 1 and Fig. 2 demonstrates the segmentation results of Felzenszwalb and Huttenlocher (FH), Compact Watershed (CW), Quick Shift (QS) and SLIC algorithms on the Berkeley Segmentation (BSD300) [24] and WHU Building datasets [25].
From Fig. 1 (from left to right), the FH algorithm (third row) has produced 581, 473 and 711 regions. While the CW algorithm (fourth row) has segmented the images into 96 regions, respectively. The QS algorithm (fifth row) has delineated the images into 387, 425 and 442 regions. While the SLIC algorithm (sixth row) has segmented the images into 79, 78 and 69 regions, respectively. While in Fig. 2 (from left to right), the FH algorithm (third row) has delineated the images into 1111, 1154, and 1221 regions. Furthermore, the CW algorithm (fourth row) has produced 100 segments of regions, respectively. In addition, the images were segmented into 3054, 2733, and 2361 regions by the QS algorithm (fifth row). Moreover, the SLIC algorithm (last row) has generated 56, 50, and 64 regions, respectively.

Table 1 shows the quantitative evaluations of the four region-based algorithms in segmenting the images from Berkeley Segmentation dataset (BSD300) and WHU Building dataset [25]. The SLIC algorithm has average ARE (0.9682) on natural images and ARE (0.2426) on remote sensing images. In addition, the CW algorithm has obtained an average ARE (0.9686) on natural images and average ARE (0.1811) on remote sensing images. Moreover, the average VOI of the CW algorithm for natural images is (7.5248) and for remote sensing images is (5.2066), respectively. The SLIC produces better average VOI than CW in natural and remote sensing images. In contrast to SLIC and CW algorithms, FH algorithm achieves better results only than QS on natural and remote sensing images. Thus, QS ranked last as it has the highest average ARE and VOI in segmenting the natural and remote sensing images indicating poor results.

<table>
<thead>
<tr>
<th>Input Image</th>
<th>GT</th>
<th>FH</th>
<th>CW</th>
<th>QS</th>
<th>SLIC</th>
</tr>
</thead>
</table>

Fig. 1. Results of over-segmentation obtained from the FH [11], QS [12], CW [13], SLIC [14] algorithms on Berkely Segmentation (BSD300) dataset [24].
This paper compared the performance of four region-segmentation algorithms on WHU Building dataset [25].

Fig. 2. Results of over-segmentation obtained from the FH [11], QS [12], CW [13], SLIC [14] algorithms on WHU Building dataset.

Table 1. Performance comparison of the FH, CW, QS, SLIC algorithms on Berkely Segmentation (BSD300) and WHU Buildings dataset. Bold entries indicate best performance results for each category of dataset.

<table>
<thead>
<tr>
<th>Algo.</th>
<th>Average Adapted Rand Error</th>
<th>Average Variation of Information</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Natural Images</td>
<td>Remote Sensing Images</td>
</tr>
<tr>
<td>FH</td>
<td>0.9828</td>
<td>0.6714</td>
</tr>
<tr>
<td>CW</td>
<td>0.9686</td>
<td>0.1811</td>
</tr>
<tr>
<td>QS</td>
<td>0.9880</td>
<td>0.7597</td>
</tr>
<tr>
<td>SLIC</td>
<td><strong>0.9682</strong></td>
<td>0.2426</td>
</tr>
</tbody>
</table>
The results of qualitative evaluations in Fig. 1 and Fig. 2 indicate that the SLIC algorithm adheres to image object boundaries far better than other three algorithms in the final segmentation results. The performance of these region-based segmentation algorithms are dependent on the selection of the parameters values, such as scale for FH, compactness for CW, kernel size for QS and n_segments for SLIC. This is because according to the authors in [16], the parameter values greatly affect the segmentation results of an algorithm. If the parameter values are not chosen effectively, then this may cause either severe over- or under-segmentation in the final results.

From the quantitative evaluations, it was observed that the natural images segmentation achieved highest average ARE and VOI values. This indicates that poor segmentation results being generated for natural images segmentation by all the chosen region-based algorithms in comparison to the remote sensing images (RSI) segmentation. The main reason is because the natural images dataset used in this work focuses to segment the entire image as ROIs. While, the RSI dataset used in this work has focused to delineate only the building regions as ROIs. Although severe over-segmentation is observed in the RSI segmentation, it still achieved better segmentation results for all the chosen algorithms in this paper. Even though in general, RSI has higher complexity of ROIs as compared to the natural images [17,18].

In addition, SLIC performed better than other three region-based segmentation algorithms by average ARE (0.9682), VOI (7.4811) for natural images, and VOI (4.9297) for remote sensing images. This algorithm has the ability of utilizing memory efficiently while segmenting the high-resolution images other than generating over-segmented regions that adhere to image object boundaries [31]. Furthermore, the over-segmented regions generated by SLIC algorithm are uniform and regular in shape [14]. In contrast, the CW algorithm also generated over-segmented regions that adhered well to the images objects boundaries. The CW algorithm achieves higher average ARE (0.9686) and VOI (7.5248) than SLIC indicating poor segmentation results for natural images. While for the remote sensing images, it shows better segmentation results with lower average ARE (0.1811) than SLIC and lower average VOI (5.2066) than QS and FH. In this case, the CW demonstrated better segmentation results only for remote sensing images. This is mainly because of the compactness and marker parameters in the CW algorithm [13], in which have the ability to influence the assignment of non-marker pixels to object regions, effectively [32,33]. Furthermore, the FH and QS algorithms show poor adherence to image boundaries and have higher values of average ARE and VOI for natural as well as remote sensing images. It is because as mentioned in [10] and [34], the FH and QS algorithms has no compactness parameter being defined to control the generation of the over-segmented regions. Moreover, in this work, only four region-based segmentation algorithms were chosen for comparison which can be addressed as the limitation of this study.

<table>
<thead>
<tr>
<th>Input</th>
<th>SLIC</th>
<th>CW</th>
<th>QS</th>
<th>FH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natural Images</td>
<td>0.9682</td>
<td>0.9880</td>
<td>0.9686</td>
<td>0.9828</td>
</tr>
<tr>
<td>Remote Images</td>
<td>0.1811</td>
<td>0.6714</td>
<td>0.1811</td>
<td>0.6714</td>
</tr>
</tbody>
</table>

5 Conclusion

This paper compared the performance of four region-based segmentation algorithms on natural and remote sensing images, respectively. It has been concluded that, SLIC algorithm has generated lower number of the over-segmented regions with regular shape and better adherence to image object boundaries. In addition, SLIC has also produced lower average ARE and VOI indicating better segmentation results for natural and remote sensing images than the other three region-based segmentation algorithms. In future, these algorithms performance will be further analysed with the convolutional neural network (CNN) based segmentation. The authors in [35] have compared the FH and the SLIC algorithms with their proposed CNN-based segmentation approach. They have concluded that, by introducing CNN for image segmentation, significant results can be achieved. Furthermore, according to [36], over-segmented regions generated by region-based segmentation algorithm are used to
train the CNN for validation. This trained CNN can be used for generation of the over-segmented regions for large images data sets to achieve the significant segmentation results. Additionally, the comparable research will be performed on the other region-based segmentation algorithms in continuation to the results produced in this paper.

The research has been carried out under UTAR Research Fund project IPSR/RMC/UTARRF/2021-C1/M01 provided by Universiti Tunku Abdul Rahman, Malaysia.

References


33. M. Vadiveloo, R. Abdullah, and M. Rajeswari., *A graph-based watershed merging using fuzzy C-means and simulated annealing for image segmentation*. Seventh International
