SKIN LESION SEGMENTATION

Elizaveta Genke, Mladen Rakic, Liliana Valencia Rodriguez, Daria Zotova
Supervised by Alessandro Bria, Post-doc Researcher

University of Cassino and Southern Lazio, Medical Imaging and Applications EMJMD

ABSTRACT

The paper describes methods for automated skin lesion segmentation. Two methods are proposed, based namely on Mean-shift segmentation and watershed segmentation. First sections serve as a brief overview of the problem and introduce the thorough database description and materials used for the research. Later, detailed description of both approaches is given, followed by the obtained results and the qualitative and quantitative comparison of the methods. Jaccard index is used as a performance evaluation metric. Besides, a comprehensive insight into the complete pipeline is given, including all the necessary steps to reach the final outcome, as well as the illustrative examples of the notable milestones obtained, such as pre-processing and hair removal steps.

***Index Terms*—** segmentation, skin lesion, mean shift, watershed, hair removal

1. INTRODUCTION

According to epidemiology studies [1], during next decades the overall number of diagnosed patients with melanoma will increase, especially in white population where the prevalence of this disease is higher than in ethnic groups. Considering this information, health services need to prepare and determine accurate and fast diagnostic procedures.

Melanoma is defined by the Skin Cancer Foundation as the unrepairable DNA damage of the skin cells that leads to their rapid growth and formation of malignant tumors. Different techniques can be used to diagnose skin cancer, one of which is the dermatoscopy. It is an imaging technique that allows the observation of thickening of layers, epidermal organization, and borders of a lesion [2].

A pipeline generalization in CAD algorithms to perform skin lesion diagnosis includes: image acquisition, artifact detection, lesion segmentation, feature extraction and classification [3]. The segmentation step is a key milestone towards obtaining a good classification, due to several varying parameters in the images such as [4] (i) low contrast between the lesion and the surrounding skin; (ii) irregular and fuzzy lesion borders; (iii) artifacts, such as skin lines, air bubbles and hairs; (iv) non-homogeneous coloring inside the lesion.

In the following sections, the automatic segmentation method for the skin lesions is proposed, based on either Mean-shift or watershed segmentation as a main approach. Various steps in the segmentation process are further explained and developed, such as the pre-processing (which includes contrast enhancement, filtering, hair removal, etc.) and region extraction.

Performance measurement is evaluated on both approaches and is quantified by using Jaccard index as metrics. Proposed algorithms are discussed, including comparison for both good and bad segmentation cases.

2. MATERIALS

The data set used consists of 200 images taken from the ISIC (International Skin Imaging Collaboration) open source public access archive, together with the corresponding 200 ground truths, obtained by an expert. Images include both benign and malignant examples, and do not share any particular clinical attribute (e.g. patients' age, longest diameter of the lesion, type of diagnosis, family and personal history of melanoma, etc.). In other words, even though it is small, the data set is considered relatively various.

According to the ISIC, both benign and malignant lesions were confirmed by histopathology reports or a clinical follow-up.

However, all the images used for this study consist of the close-up lesion present in the central part, and do not include surroundings other than the neighbouring skin tissue. This makes the segmentation task somewhat easier to achieve in comparison to the whole database provided by the ISIC, which is greatly more diverse.

In order to accomplish the segmentation challenge, Python programming language is used as a base. Functionalities that help structure the methods proposed mainly come from OpenCV library and Skimage module. Well structured and documented code is available in the form of a Jupyter Notebook script, because it provides a nice readability and a more natural insight to the pipeline. Different sections of the code and a logical separation of the tasks give the possibility for undemanding modifications and further development.

3. METHODOLOGY

Two segmentation methods based on either Mean-shift or Watershed are considered in our approach. Prior to that, images are pre-processed in order to remove noise and hairs, as well as to reduce the impact of dark background pixels.

3.1. Hair removal method

The presence of hair in dermatoscopy images might hide important information about skin lesions and have influence on the result of segmentation. Therefore, it is necessary to include a technique which will tackle this problem. Bottom-hat based algorithm is chosen as a hair removal method, as proposed by Maglogiannis and Delibasis [5]. In this approach, the initial RGB image is converted into grayscale image and Gaussian filter with 5x5 kernel is applied, in order to smoothen the image. The Laplacian image is then subtracted from previously blurred one, in order to receive a sharpened result.

Hairs are usually long and thin structures, and in order to detect them, Bottom-hat transformation is performed: the linear structuring element is rotated in different directions, performing morphological closing that is then subtracted from the input. The results of Bottom-hat transformation are summed up and the output image is converted into binary using Otsu's thresholding. The resulting binary image serves as a mask that contains the outlines which are filled by performing morphological dilation with linear structuring element. As a last step, the final mask is used in order to perform the inpainting technique and retrieve RGB image without hair. Figure 1 demonstrates intermediate and final results of hair removal.

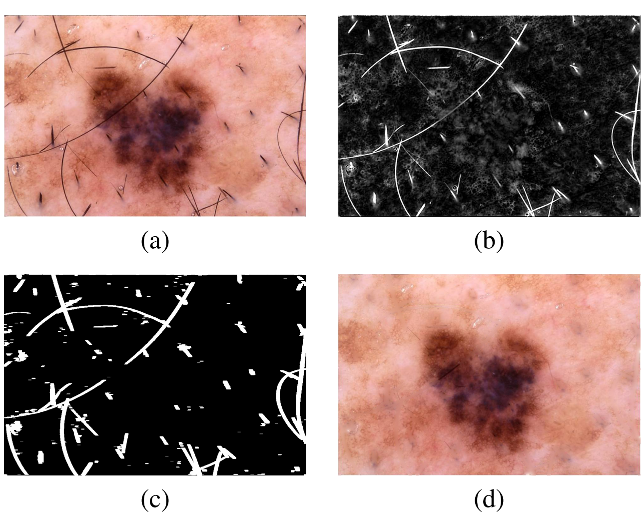

Fig. 1: Notable hair removal stages: a) original RGB image b) bottom-hat transformation c) binary mask d) final image

3.2. Mean-shift segmentation

This method is based on Mean-shift filtering, which is an iterative algorithm that divides image into clusters. For every pixel, its neighbouring pixels within certain radius and color distance are chosen, and new color for cluster is calculated. This process is repeated until convergence.

For every image in data set, OpenCV implementation of Mean-shift filtering is used. After this, filtering image is converted to grayscale and contrast is enhanced. If image had big dark regions, they are subsequently removed. Finally, Otsu's thresholding is applied and output binary image is inverted, to be suitable for further processing.

3.3. Watershed segmentation

Watershed is a morphological tool based on region growing [6]. It considers any gray tone as a topographical surface that has three dimensions: two spatial coordinates and one intensity. Each surface fills from its minima determined by the gradient image. In order to prevent oversegmentation of the regions, a set of markers must be defined. Thus, the segmentation process with watershed consists of two steps: (i) finding the markers and (ii) defining the segmentation criterion that will be used to split the regions. This is also referred to as marker-controller watershed. Two markers are used: one is inside of the region to be segmented (internal) and the other corresponds to the background of the image (external). Both markers are obtained using morphological operations. Then, in order to find the foreground area, distance transform is used (its output is the distance from every pixel to the nearest non-zero valued pixel).

3.4. Region extraction

In some cases, several regions are detected after segmentation. To correct for this behaviour, region extraction function has been created. First, small objects in the picture are removed. Then, segmented regions are extracted and region closest to the center of image is chosen. This condition prevents choosing wrong region because all lesions in our data set are well-centered. Ultimately, holes inside segmented region are filled.

3.5. Dilation

The result of region extraction in general yields in an output region with complex borders, and that imposes a problem when comparing results with ground truth, which is approximated with the bounding polygon in some cases. To overcome this issue, dilation is applied. After this operation, we have smoother region borders, which positively affects the overall result.

3.6. Jaccard index

In order to quantify obtained results and evaluate the performance, Jaccard index (1) is calculated as a metric for every image, and mean index is computed for the whole data set.

$$JC = \frac{TP}{TP + FP + FN} \tag{1}$$

where TP - true positive, FP - false positive and FN - false negative pixels.

4. RESULTS

As a representative example, figure 2 shows two original images, as well as the results of segmentation algorithm and the corresponding groundtruths.

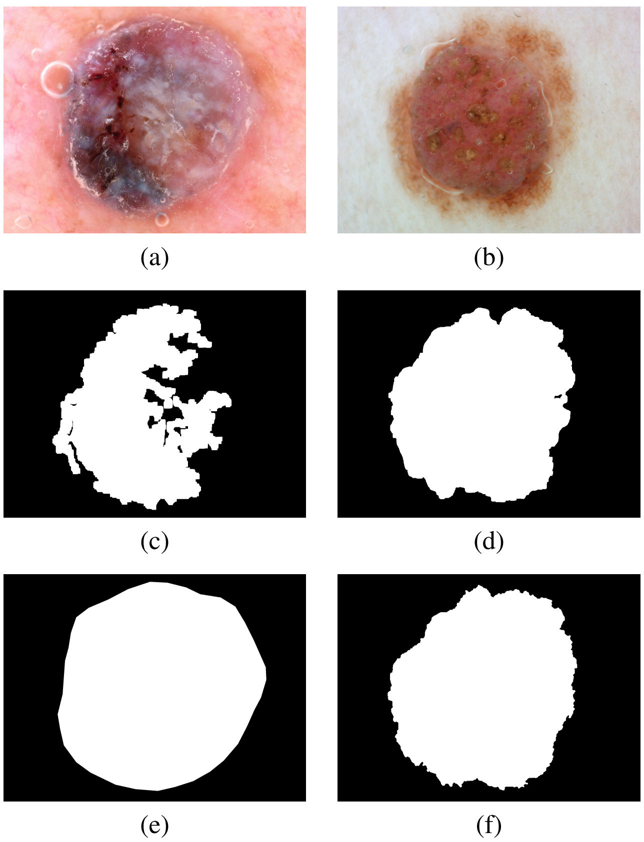

Fig. 2: Sample results: a) and b) original RGB images c) bad segmentation example d) good segmentation example e) and f) groundtruths

As previously stated, Jaccard index is chosen as a metrics for the performance evaluation. In order to quantify the overall performance on the whole data set, mean index is computed. The obtained values for the average Jaccard indices and the standard deviations are shown in table 1.

5. DISCUSSION

As stated in the previous section, both Mean-shift and watershed segmentation methods yield in the relatively similar performance, quantified by Jaccard index. However, the two algorithms perform differently on dissimilar image samples.

Method	mean	std
Mean-shift	0.8195	0.1545
Watershed	0.7924	0.2059

Table 1: Mean values and standard deviations for Jaccard indices

In case of high contrast image, containing homogeneous lesion with relatively uniform texture, both Mean-shift and watershed segmentation work well. When the lesion consists of several dark areas connected with paler regions, Mean-shift proved to be the better choice.

Yet, if the lesion is very pale and the overall image contrast is significantly low, the methods fail to segment the region of interest properly. Even with the contrast enhancement techniques, the overall performance in these cases does not improve notably.

It has to be noted though, that if we consider the data set as a whole, the watershed segmentation has a somewhat more extreme performance than the Mean-shift, if we consider the Jaccard index as a performance metric (i.e. the index is generally either very high or very low). Ultimately, the overall performance ends up being similar and it is not suitable to be compared as a whole, but has to be considered on separate cases instead.

To illustrate this, figures 3 and 4 show two distinct cases of segmentation, one with Mean-shift method yielding in better result, other with watershed as a better option.

Throughout the development, several notable difficulties came to light, all boiling down to the principal one - coming up with a unique algorithm which will perform automatically and well enough on the whole data set. As explained in previous sections, most of these issues were relatively successfully tackled. These include:

- field of view problem, where, as a consequence of the different acquisition modality, several samples contain dark background, often connected to the region of interest;
- low contrast samples, which often lead to a missegmentation, even with contrast enhancement techniques applied;
- long computational time, which at this stage only suits the offline analysis (this still remains an optimization challenge);
- other individual problems, as a consequence of the samples being highly diverse in color, shapes, sizes and other descriptive parameters.

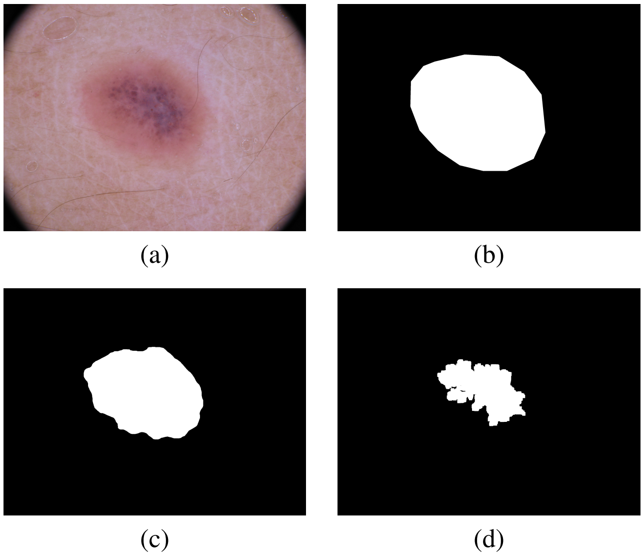

Fig. 3: Sample result with Mean-shift method as a better approach : a) original RGB image b) groundtruth c) Mean-shift segmentation d) watershed segmentation

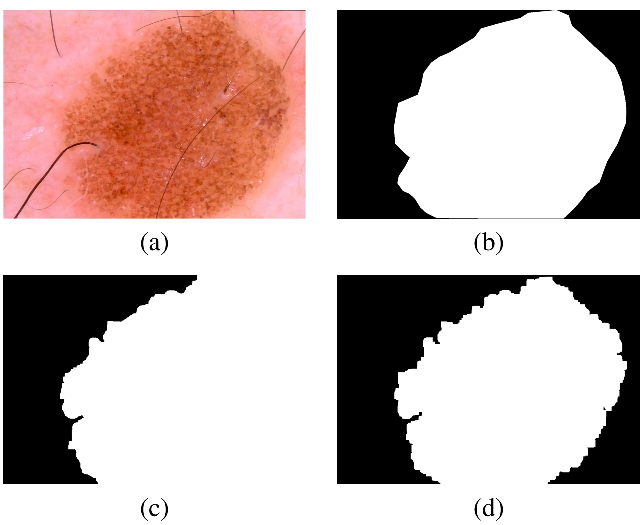

Fig. 4: Sample result with watershed method as a better approach : a) original RGB image b) groundtruth c) Mean-shift segmentation d) watershed segmentation

6. CONCLUSION

We proposed a completely automated method for the skin lesion segmentation using Mean-shift and watershed as two main approaches. As previously shown, it performs relatively well on a very diverse data set. And even though modern approaches in solving this problem often include deep learning models, our perspective is easily understandable and offers a great interpretability.

Future improvements can be focused on one or more of several key points:

- computation time optimization;
- correction for the low contrast samples, which end up being segmented badly;
- correction for different backgrounds, as a consequence of various acquisition modalities;

Furthermore, the algorithm can be tested on a bigger portion of the ISIC archive, in order to evaluate the performance on a larger data set. It is arguable that the performance measure would decrease in that case, since the archive includes vast samples, but it can open other possibilities and improvement ideas.

7. REFERENCES

[1] Zoe Apalla, Aimilios Lallas, Elena Sotiriou, Elizabeth Lazaridou, and Demetrios Ioannides, "Epidemiological trends in skin cancer," *Dermatology practical & conceptual*, vol. 7, no. 2, pp. 1, 2017.

[2] Zoe Apalla, Dorothée Nashan, Richard B Weller, and Xavier Castellsagué, "Skin cancer: Epidemiology, disease burden, pathophysiology, diagnosis, and therapeutic approaches," *Dermatology and therapy*, vol. 7, no. 1, pp. 5–19, 2017.

[3] Paul Wighton, Tim K Lee, Harvey Lui, David I McLean, and M Stella Atkins, "Generalizing common tasks in automated skin lesion diagnosis," *IEEE Transactions on Information Technology in Biomedicine*, vol. 15, no. 4, pp. 622–629, 2011.

[4] M Emre Celebi, Hassan A Kingravi, Hitoshi Iyatomi, Y Alp Aslandogan, William V Stoecker, Randy H Moss, Joseph M Malters, James M Grichnik, Ashfaq A Marghoob, Harold S Rabinovitz, et al., "Border detection in dermoscopy images using statistical region merging," *Skin Research and Technology*, vol. 14, no. 3, pp. 347–353, 2008.

[5] Ilias Maglogiannis and Kostantinos Delibasis, "Hair removal on dermoscopy images," pp. 2960–2963, 2015.

[6] Sameer Ruparelia, "Implementation of watershed based image segmentation algorithm in fpga," M.S. thesis, 2012.

8. APPENDIX

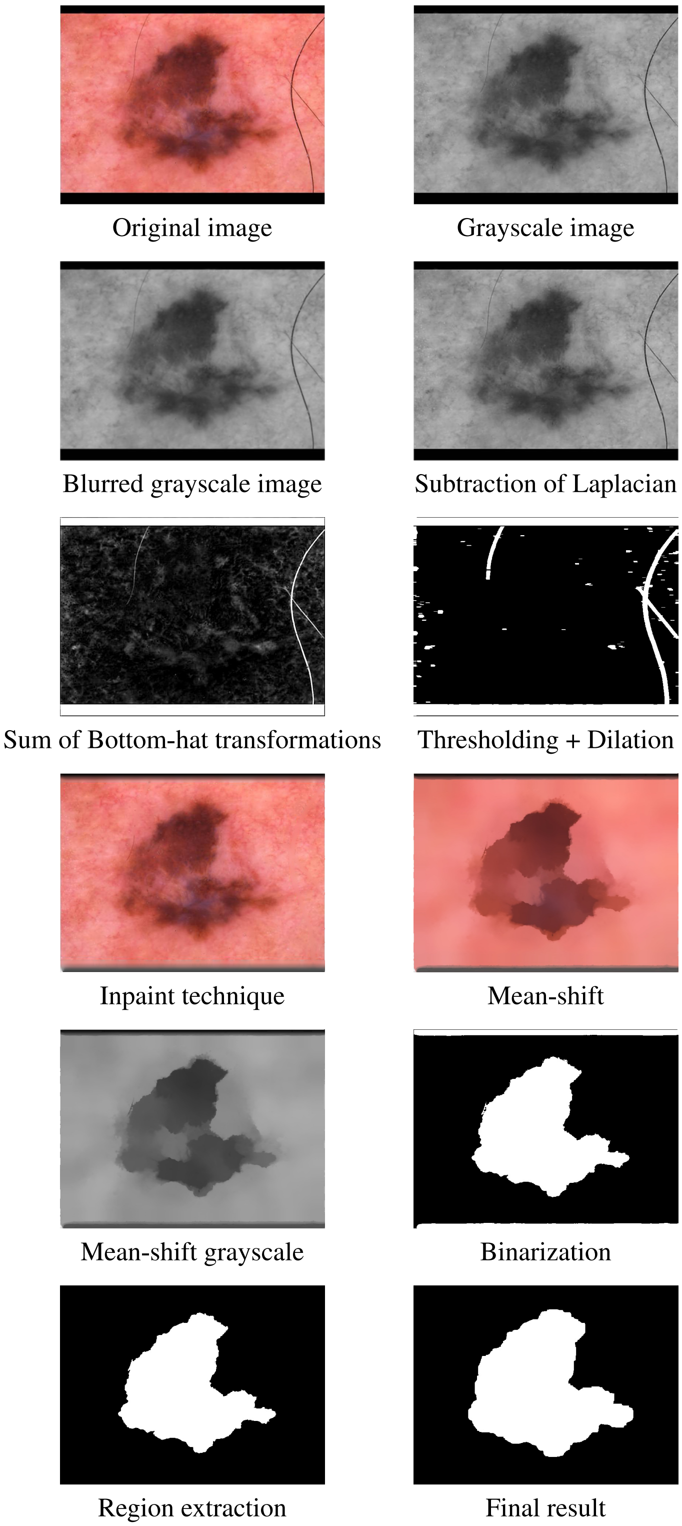

Fig. 5: Intermediate results of Mean-shift segmentation

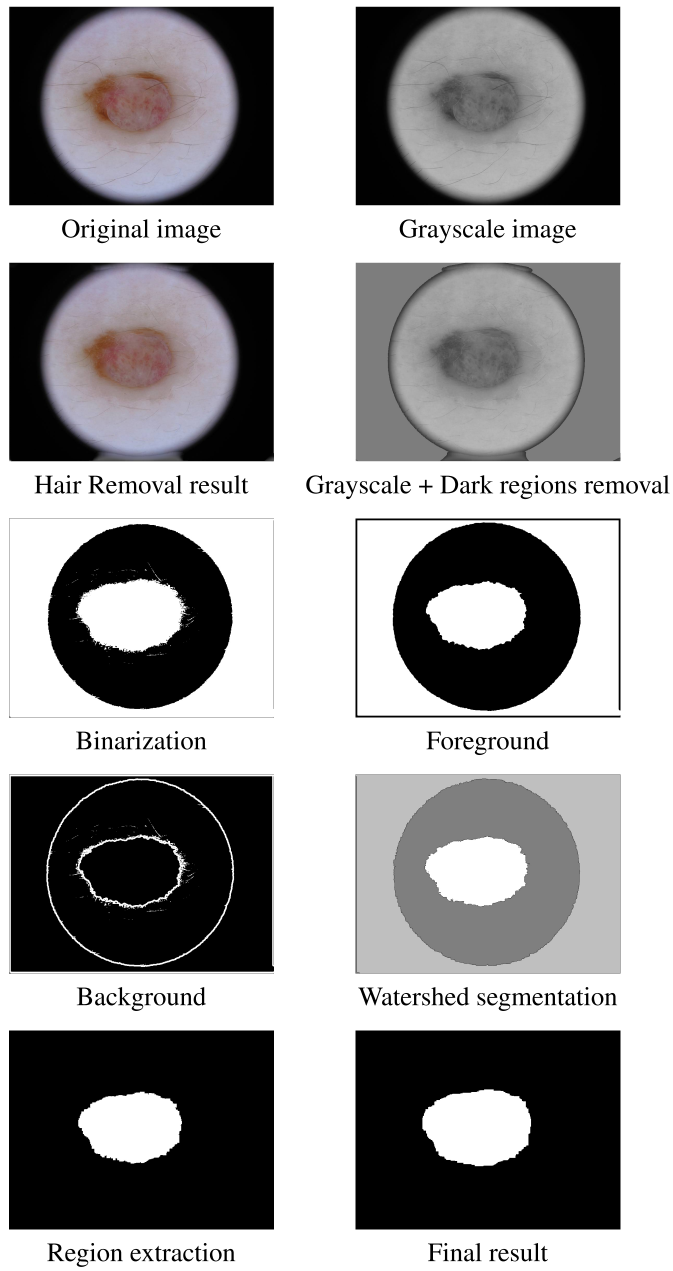

Fig. 6: Intermediate results of Watershed segmentation