Automated Saliency-based Lesion Segmentation in Dermoscopic Images

Euijoon Ahn, Lei Bi, Youn Hyun Jung, Jinman Kim, Member, IEEE, Changyang Li, Michael Fulham and David Dagan Feng, Fellow, IEEE

Abstract—The segmentation of skin lesions in dermoscopic images is considered as one of the most important steps in computer-aided diagnosis (CAD) for automated melanoma diagnosis. Existing methods, however, have problems with over-segmentation and do not perform well when the contrast between the lesion and its surrounding skin is low. Hence, in this study, we propose a new automated saliency-based skin lesion segmentation (SSLS) that we designed to exploit the inherent properties of dermoscopic images, which have a focal central region and subtle contrast discrimination with the surrounding regions. The proposed method was evaluated on a public dataset of lesional dermoscopic images and was compared to established methods for lesion segmentation that included adaptive thresholding, Chan-based level set and seeded region growing. Our results show that SSLS outperformed the other methods in regard to accuracy and robustness, in particular, for difficult cases.

I. INTRODUCTION

Dermoscopy is a non-invasive diagnosis technique, used in dermatology, for the in vivo observation of pigmented skin lesions [1]. Dermoscopic images provide a detailed view of morphological structures and play an important role in the early diagnosis of malignant melanoma [1, 2]. The manual visual inspection made by dermatologists, however, is time consuming, subjective, and not reproducible; for example, well trained dermatologists produce widely varying delineation of the same skin lesions [3]. This is attributed to the complexity of lesion segmentation due to a variation in lesion size and shape, fuzzy lesion boundaries, different skin color types and presence of hair. Motivated by this difficulty, there has been a great interest in the development of computer-aided diagnosis (CAD) systems that can assist the dermatologists’ clinical evaluation [1]. Lesion segmentation is a fundamental requirement for melanoma CAD. A number of segmentation methods have been proposed recently to automatically segment lesions on dermoscopic images [1] and these include adaptive thresholding (AT) [1], adaptive snakes [4], level set proposed by Chan et al (C-LS) [5] and a Seeded Region growing (SRG) [6]. These methods are variants of all the most applicable image segmentation approaches for dermoscopic images. The AT method attempts to segment lesion regions by comparing the color of each pixel with a threshold and classified as a lesion if it is darker than the threshold. It uses the entropy to select the best channel from RGB for best discrimination. The threshold is then automatically determined as the local minimum between the maxima, plus a small offset to account for quantization issues. Although this method is easily adopted due to its simplicity, it is often usually limited by the luminance distribution and may fail if there are multiple peaks in its luminance histogram. The adaptive snake model uses the contour of the regions in the image to detect the lesions’ boundaries. The major drawback with this approach, however, is its reliance on the optimal selection of the segmentation parameters, which makes it problematic to be adopted in clinical environment. C-LS and SRG are region-based approaches. C-LS extends on existing level set method with a mean curvature motion technique to robustly fit the distribution of the pixels. SRG attempts to iteratively find neighboring pixels from the seed pixels that share similar features until a predefined criterion is met. Unfortunately, region-based approaches inherently require manual initialisation and are prone to ‘leakage’ problems in the region growing process [7, 19].

For this study, we propose a new automated saliency-based skin lesion segmentation (SSLS) method that we designed to exploit the inherent properties of dermoscopic images, which have a focal central region and subtle contrast discrimination with the surrounding regions. Our novel approach was motivated by findings in cognitive psychology [8, 9], where ‘saliency’ stems from the human visual system which locates important regions or objects in images by observing uniqueness, unpredictability, rarity and surprise. Technically, saliency-based methods focus on a central region and its surroundings, locally or globally, using features such as intensity, contrast, color and orientation [10]. In this study, we demonstrate that the subtle contrast discrimination from surrounding regions can be overcome via saliency detection through measuring sparse reconstruction errors against image backgrounds. It is a robust approach to detect skin lesions via propagating multi-scale sparse reconstruction information. We evaluated our SSLS method by comparing it to fully- and semi-automated segmentation methods such as AT, C-LS and SRG on a public dermoscopic lesion dataset.

E. Ahn, L. Bi, Y. Jung, J. Kim, C. Li, and D. Feng are with the School of Information Technologies, University of Sydney, NSW, Australia. E. Ahn (e-mail: eahn4614@uni.sydney.edu.au); L. Bi (e-mail: lei@it.usyd.edu.au); Y. Jung (e-mail: yju6175@uni.sydney.edu.au); J. Kim (e-mail: jinman.kim@sydney.edu.au); C. Li (e-mail: changyang.li@sydney.edu.au).

M. Fulham is with Department of Molecular Imaging, Royal Prince Alfred Hospital, NSW, Australia. He is also with the Sydney Medical School, University of Sydney, NSW, Australia (e-mail: michael.fulham@sydney.edu.au).

D. Feng is also with Med-X Research Institute, Shanghai Jiao Tong University, China (e-mail: dagan.feng@sydney.edu.au).

This research was funded in part by Australia Research Council grants.
II. METHODS

A. Overview of Saliency-based Skin Lesion Segmentation (SLSS) Method

Figure 1 is an overview of our SLSS method. Initially, hairs on dermoscopic images were removed as a pre-processing step. A background template was then created by using a superpixels algorithm [11] and images were reconstructed by measuring the sparse reconstruction errors against the background template that is used as an indication of saliency [12]. Context-based error propagation was then applied to smooth the reconstruction error. Finally, a pixel level saliency map was created by integrating multi-scale reconstruction errors and refined by a lesion biased Gaussian model, followed by its conversion to a binary segmentation result.

![Image](image_url)

Figure 1. Overview of our SLSS method.

B. Hair Removal

Artifacts such as hair are crucial obstacles that degrade the performance of the segmentation methods. The presence of hairs on the skin may occlude parts of the lesion, making accurate skin lesion segmentation difficult. To resolve this issue, we adopted a well-known hair removal algorithm proposed by Lee et al. [13] to remove hair visible in the dermoscopic images. Hair was detected by using multiple hair templates with various directions (0°, 45° and 90°) followed by replacement of detected hair pixels with non-hair pixels derived from neighborhood pixels.

C. Background Template Creation

A simple linear iterative clustering (SLIC) [11] algorithm was used to partition the dermoscopic image into $N$ segments where $N$ is the total number of the segments in the image. This helps to capture structural information that pixel-level features may not be able to represent. It also has been shown to be fast method with good performance [11]. The input image was then partitioned into segments with different sizes of $N$ to create multi-scale background templates. A regional mean feature consisting of CIELAB color spaces, RGB color values, and spatial location defined by $[L, a, b, R, G, B, x, y]$ were used to represent each of the segments for their proven performance in saliency detection [14]. An entire image can be represented as $X = [X_1, X_2, ..., X_M]$ where $X_i$ is a regional feature representing the $i^{th}$ segment. Since image boundaries are good visual cues to separate foreground from background where the salient object (lesion) is usually inside the boundary, we extracted the boundary segments to construct the background template $B = [b_1, b_2, ..., b_M]$, where $M$ is the number of boundary segments. Figure 1 illustrates examples of boundary templates extracted at various scales (different size of superpixels) where the non-background regions are masked out.

D. Saliency Measure via Sparse Reconstruction Error

In general, a sparse appearance model can create concrete and unique representations of the object, in particular, for images with cluttered scenes. Sparse reconstruction errors were used to measure the probability of where the segment belongs. Basically, a segment with a larger reconstruction error against the background template is more likely to be in the foreground. Since skin lesions (foreground segments) in dermoscopic images are unlikely located in the background template (e.g. lesions appear at the image boundaries), we can expect that their reconstruction errors are relatively high.

Sparse reconstruction error is computed as the residual-based on the sparse representation of the background templates $B$. The segment $X_i$ is encoded as follows:

$$
\alpha_i = \arg\min_{\alpha_i} \|X_i - B\alpha_i\|^2 + \lambda \|\alpha_i\|_1 ,
$$

and the sparse reconstruction error $\varepsilon_i$ is computed as:

$$
\varepsilon_i = \|X_i - B\alpha_i\|^2 .
$$

E. Context-based Error Propagation

Context-based error propagation method is proposed to smooth the sparse reconstruction errors by considering other segments nearby. As shown in Figure 2, this allows each segment in lesion regions to be evenly emphasized. A $K$-means algorithm [15] was used to cluster all the segments in the image into the $K$ number of clusters. The new propagated reconstruction error of $\varepsilon_i$ was estimated by proportionally combining the weighted averaging error of all other segments in the same cluster, and the initial sparse reconstruction error of $\varepsilon_i$ at segment $i$. This can be defined as:

$$
\tilde{\varepsilon}_i = \tau \sum_{j=1}^{N_i} w_{ikj} \varepsilon_{kj} + (1 - \tau)\varepsilon_i ,
$$

where $[k_1, k_2, k_3, ..., k_N_i]$ indicate the $N_i$ segment labels in cluster $k$ and $\tau$ is a weight parameter. The weight of each segment $i$ was then estimated by its normalised similarity according to:
where the $\sigma_X^2$ denotes the sum of the variance in each of X and $\delta(k_i - i)$ is the indicator function.

\[
W_{k_j} = \frac{\exp\left(-\frac{||x_k-x_j||^2}{2\sigma_X^2}\right)(1-\delta(k_j-i))}{\sum_{j=1}^{N} y_j^N_c \exp\left(-\frac{||x_l-x_j||^2}{2\sigma_X^2}\right)} ,
\]

G. Skin Lesion Segmentation via Saliency Map

The conversion of the generated saliency map to a binary image was conducted via fuzzy similarity thresholding proposed by Huang et al [16]. This method can help effectively locate the deep valley of gray-scale (i.e. saliency map) histogram, which particularly performs well in maintaining overall shape of the segmented area (skin lesions in our setting) [16]. Given a certain threshold value, the membership function of a pixel is defined by the absolute difference between its gray level and the averaged gray level of the region that it belongs to - the lesion or skin. Then the optimal threshold can be found by minimising the measure of fuzziness of an image [16]. The fuzziness of each pixel can be measured by using Shannon's entropy or the Yager's measure and we used Shannon's entropy to measure the fuzziness in our experiment.

III. RESULTS AND DISCUSSION

A. Materials and Experiments Setup

The PH2 public dataset [2] provides 200 dermoscopic image studies, including 80 common nevi, 80 atypical nevi, and 40 melanomas. To avoid studies where the lesion in the image was incomplete, i.e., lesion(s) connected to the boundary edge of the images, 40 studies were excluded. All studies were 8-bit RGB color images with 768x560 pixel resolution. Manually annotated lesions from expert dermatologists were available from the PH2 and used as the ground truth data.

We compared our SLSS method to three well-known segmentation algorithms for dermoscopic images: AT [1], C-LS [5], and SRG [6]. Hair removal, as a pre-processing step, was applied with the following specific settings for all the methods. For SLSS, there were two main parameters: number of clusters K and the weight factor $\tau$ in Eq. 3 and 4. We empirically used $K = 8$ and $\tau = 0.5$ respectively in all our experiments. Different parameters ($K = 4$, 6, 8 and $\tau = 0.3, 0.5, 0.7$) were also used in the experiment and it was observed that the results were insensitive to the either parameter. The parameter $\lambda$ in Eq. 1 was empirically set to 0.01. The superpixels at $N_s = 8$ different scales were created ranging from 50 to 400 superpixels. For C-LS and SRG, an initial starting seed point was necessary: C-LS was initialised with a number of small circular masks across whole image, which generates most fast and accurate results [17]. The maximum number of iteration for C-LS was empirically set as 500, which ensured that it iterated until convergence. For SRG, the seed point was set to be the center of the skin lesion (not the center of image). C-LS and SRG also had to be pre-processed to remove the dark corner regions to have a reasonable comparison with our approach. There were inconsistent corner sizes in the image studies; we manually removed all the corners via a connected thresholding.

The segmentation results from all the methods were compared with the ground truth from the PH2 dataset. The dice similarity coefficient (DSC) [18] method was used to quantitatively evaluate the similarity of the segmentation results to their corresponding ground truth, which was defined as:

\[
DSC(A, B) = \frac{2N(AB)}{N(A)+N(B)} ,
\]
where \( A \) denotes the segmentation result and \( B \) is the ground truth; \( N \) represents the number of pixels in the corresponding set.

B. Results and Discussion

Segmentation results from three study examples, where (a) – (f) represent the original image in column 1, ground truth in column 2 and the segmentation results in columns 3 from AT, C-LS, SRG and our SSLS methods. A summary of the segmentation results for all the methods is shown in Table 1. Figure 3 illustrates the segmentation images for three example studies: a simple case (left), an average case (middle) and a difficult case (right), which was defined based on the average DSC value of all algorithms for each case. Compared with other methods, our SSLS generated significantly higher results, compared to the other methods, and it had the best overall average (91.66%), as well as best maximum (97.52%) / minimum (39.18%) accuracy. The results indicate that SSLS is more capable of identifying and separating lesions, especially the difficult studies. For the difficult cases SSLS had the highest minimum accuracy which was 35.66% higher than the second best method (AT) and had the lowest standard deviations. AT performed second best in minimum accuracy but it was not able to accurately identify the edges of the lesions across all the studies, which resulted in the second lowest average (see Figure 3(c), column 3). The high maximum accuracy for all the methods was mainly attributed to several studies having relative simple regions with high contrast to the background regions (see Figure 3, column 1). The C-LS method was the closest approach to our SSLS in overall DSC measures. In our evaluation, however, it required the manual removal of the dark corner regions (See section 3.1) and this task is challenging to automate reliably. Further, C-LS was not able to segment lesions if they had a heterogeneous background (see Figure 3(d), column 3). Although simple border smoothing or dilation may be applied to C-LS (and others), this will be an additional post-processing which may bias the segmentation results. Our SLSS method, however, was fully automated and was insensitive to such image conditions.

### Table 1: Average, Maximum (Max), minimum (Min) and standard deviation (Std) of the segmentation results measured by DSC for SSLS, AT, C-LS, and SRG methods.

<table>
<thead>
<tr>
<th>Method</th>
<th>( AT )</th>
<th>( C-LS )</th>
<th>( SRG )</th>
<th>( SSLS )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg.</td>
<td>80.51</td>
<td>85.30</td>
<td>61.23</td>
<td>91.66</td>
</tr>
<tr>
<td>Max.</td>
<td>96.68</td>
<td>97.47</td>
<td>96.80</td>
<td>97.52</td>
</tr>
<tr>
<td>Min.</td>
<td>25.21</td>
<td>7.44</td>
<td>0.19</td>
<td>39.18</td>
</tr>
<tr>
<td>Std.</td>
<td>13.81</td>
<td>11.44</td>
<td>28.93</td>
<td>6.54</td>
</tr>
</tbody>
</table>

IV. Conclusion and Future Work

In this work, we proposed a fully automated saliency-based skin lesion segmentation method for dermoscopic images. Experiments using a public dataset of 160 images showed that our method achieved the highest accuracy when compared to other segmentation methods. In future work, we will evaluate our method on a greater number of dermoscopic images and skin lesion types.

### References