Automatic Segmentation of Dermoscopy Images using Saliency Combined with Otsu Threshold

Haidi Fan\textsuperscript{a,b}, Fengying Xie\textsuperscript{a,b,}, Yang Li\textsuperscript{a,b}, Zhiguo Jiang\textsuperscript{a,b}, Jie Liu\textsuperscript{c}

\textsuperscript{a}Image Processing Center, Beihang University, Beijing 100083, China
\textsuperscript{b}Beijing Key Laboratory of Digital Media, Beihang University, Beijing 100191, China
\textsuperscript{c}Department of Dermatology, Peking Union Medical College Hospital, Beijing 100730, China

Abstract

Segmentation is one of the crucial steps for the computer-aided diagnosis (CAD) of skin cancer with dermoscopy images. To accurately extract lesion borders from dermoscopy images, a novel automatic segmentation algorithm using saliency combined with Otsu threshold is proposed in this paper, which includes enhancement and segmentation stages. In the enhancement stage, prior information on healthy skin is extracted, and the color saliency map and brightness saliency map are constructed respectively. By fusing the two saliency maps, the final enhanced image is obtained. In the segmentation stage, according to the histogram distribution of the enhanced image, an optimization function is designed to adjust the traditional Otsu threshold method to obtain more accurate lesion borders. The proposed model is validated from enhancement effectiveness and segmentation accuracy. Experimental results demonstrate that our method is robust and performs better than other state-of-the-art methods.

Keywords: Automatic segmentation, Computer-aided diagnosis, Dermoscopy images, Saliency, Threshold

*Corresponding author

Email address: xfy_73@buaa.edu.cn (Fengying Xie)
1. Introduction

Skin cancer is one of the most rapidly increasing cancers in the world [1]. Invasive melanoma alone has an estimated incidence of 76,380 and an estimated total of 10,130 deaths in the United States in 2016 [2]. Many visual diagnosis procedures have been introduced in order to help the clinical diagnosis of skin cancer, such as ABCD rule [3], 7-point checklist [4], and Menzies method [5]. However, it is still a challenging task even using these procedures due to the subjectivity of clinical interpretation and lack of reproducibility [6]. Dermoscopy, as a non-invasive skin imaging technique which makes subsurface structure more easily visible, is widely used by dermatologists in early skin cancer diagnosis. Dermoscopy images can increase clinical diagnostic accuracy compared with naked-eye examination if interpreted properly [7]. However, studies have indicated that the diagnostic accuracy of inexperienced dermatologists may be decreased by dermoscopy [8, 9]. Therefore, many researchers concentrate on the computer-aided diagnosis (CAD) system for proper interpretation of dermoscopy images. In contrast to visual assessment, CAD can provide quantitative and objective evaluation.

The standard skin cancer CAD procedure includes five aspects: image acquisition, pre-processing, lesion segmentation, feature extraction, and classification. Due to its great influence on the accuracy of the subsequent steps, lesion segmentation is crucial in CAD procedures.

A large number of algorithms have been developed for lesion segmentation in the past two decades [10]. Most automatic segmentation methods can be classified into three main categories: 1) Histogram Thresholding. Yuksel et al. [11] utilized the histogram thresholding method based on type-2 fuzzy logic techniques to segment the dermoscopy images. Cavalcanti and Scharcanski [12] proposed texture, darkness, color channels and employed hybrid threshold method to obtain binary result. Celebi et al. [13] extracted lesion borders by fusing ensembles of thresholding methods. Thresholding methods usually can achieve satisfactory result for the image with simple texture and high contrast
between lesion and skin; 2) Clustering and Region Merging. Xie et al. [14] used self-generating neural network (SGNN) combined with genetic algorithm to obtain stabilized segmentation results. And in [15], Celebi et al. detected the lesion border using the statistical region merging (SRM) algorithm. This kind of methods generally merge pixels or subregions with similar color and texture through some merge rules. They often have trouble in segmenting images with complex texture and variegated color; 3) Active Contour Model. Active contour model obtains object border based on deformable spline. Abbas et al. [16] modified region-based active contours (RACs) for multiple lesion segmentation. Kasmi et al. [17] segmented skin lesion using a biologically inspired geodesic active contour (GAC) technique. Zhou et al. [18] proposed a mean shift based gradient vector flow algorithm to locate the correct borders. This model performs poorly when the lesion border is fuzzy.

Although many lesion segmentation methods have been developed, the problem of finding accurate lesion borders remains inadequately solved due to the complexity of dermoscopy images [13]. Visual saliency, investigated by many disciplines including cognitive psychology, neurobiology, and computer vision, is focused on how we perceive and process visual stimuli [19]. Santos and Pedrini [20] utilized a saliency detection method to reduce the false positive rate in skin segmentation. To extract lesion borders from dermoscopy images, [21] and [22] used sparse-coding-based saliency method to enhance lesion objects. In this paper, a novel automatic segmentation method using saliency combined with Otsu threshold is proposed. The complex conditions, such as low contrast, hairs, incomplete lesion, are common in dermoscopy images, which may cause inaccurate segmentation. In the proposed framework, the saliency theory is used to enhance lesion objects and then the Otsu threshold method is improved to correctly segment dermoscopy images in this paper, which obtains robust segmentation results. The remainder of the paper is organized as follows. In the section 2 we describe our proposed approach in detail. Experimental results are presented and discussed in the section 3. At last, the section 4 gives the conclusion of our work.
2. Method

The proposed segmentation algorithm includes two stages, as shown in Fig. 1. In enhancement stage, the prior information on healthy skin is extracted and then, the color and brightness saliency maps are constructed. By fusing the two saliency maps, the lesion object is enhanced. In segmentation stage, the Otsu threshold is adjusted by a designed optimization function to extract accurate lesion borders from enhanced images, following which a post processing is used to remove the spots and holes to obtain the final segmentation result.

2.1. Image Enhancement Based on Saliency

The saliency of object in an image is measured based on its color, gradient and edges. Some researchers believed that the image boundary is mostly background. Based on this boundary prior, they detected objects through computing the saliency of patches or pixels based on their relevance to the image boundary and achieved outstanding experimental results [23, 24]. In this paper,
a more robust saliency method, combining the boundary prior with the color prior and brightness prior, is proposed for dermoscopy images.

2.1.1. The extraction of prior information

Before constructing saliency maps, prior information needs to be extracted. Dermoscopy images contain lesion regions and healthy skin regions. Compared with lesion, three priors on healthy skin can be concluded in most cases, which are:

- **Boundary Prior**: Healthy skin often distributes in the boundary, whereas lesion locates at the center of the image.
- **Color Prior**: Healthy skin has uniform color, while lesion colors are variegated. The color contrast between healthy skin and lesion is high.
- **Brightness Prior**: The brightness of healthy skin is higher than that of the lesion.

We extract these priors as Fig. 2. Patches of size $N \times N$ are firstly extracted with stride $\tau_1$ along the image boundary. According to Boundary Prior, boundary patches belong to healthy skin regions. For each patch, we extract prior information in the color and brightness spaces respectively.

In the color space, the image is quantized by the minimum variance (MV) method to speed up the saliency calculating process [25]. MV method divides the RGB color cube into a number of boxes. In this way, the pixels of an image are associated into groups based on the minimum intra-cluster variance principle, thus each pixel is mapped to the center value of the corresponding box. In this paper, the number of groups is set to the rounded number after multiplying the entropy value of the image by 5. Let the number of boundary patches be $n$, and their color histograms, labeled as $h_{c1}, h_{c2}, ... , h_{cn}$, are extracted from the quantized image. According to Boundary Prior and Color Prior, the boundary patch is healthy skin and its color histogram is quite different from that of the lesion.
The brightness $I$ of the RGB color image is defined as the minimum of three channels:

$$I(i) = \min\{r(i), g(i), b(i)\}$$  \hspace{1cm} (1)

where $r(i)$, $g(i)$, $b(i)$ are the values of $i$th pixel in three channels respectively.

Similar to the color space, the mean brightness values of boundary patches, labeled as $\mu_1$, $\mu_2$, ..., $\mu_n$, are calculated respectively.

When a dermoscopy image contains an incomplete lesion object, some lesion regions will locate at the image boundary. To measure the possibility that a boundary patch is filled with healthy skin, weighting factor $\omega_i$ for $i$th boundary patch is given as:

$$\omega_i = \frac{\mu_i}{\sum_{j=1}^{n} \mu_j} \hspace{1cm} i = 1, 2, 3, \ldots, n$$  \hspace{1cm} (2)

where $n$ is the number of boundary patches. According to Brightness Prior, the higher the weighting factor of a boundary patch $\omega_i$ is, the more likely it is a healthy skin patch.

The extracted color histogram, brightness mean, and weighting factor describe the prior information of boundary patches in the color and brightness spaces.
spaces respectively. Taking boundary patch as prior patch, its color and brightness prior information will be utilized for enhancing the lesion object.

2.1.2. Saliency calculation in the color space

![Figure 3: Saliency calculation process in the color space.](image)

The saliency calculation process in the color space is illustrated in Fig. 3. We use a window of size $N \times N$ to move on the whole quantized image with stride $\tau_2$. For each sampling position, the color histogram of corresponding window region is calculated.

Assuming there are $L_c$ sampling patches, the color saliency map generated using $i$th prior patch is represented as:

$$S_{ci}(j) = \chi(h_{ci}, h_j) \quad j = 1, 2, ..., L_c$$  \hspace{1cm} (3)

where $h_{ci}$ and $h_j$ are color histograms of $i$th prior patch and $j$th sampling patch respectively, $\chi(\cdot)$ is the chi-square distance between two color histograms defined as:

$$\chi(h_a, h_b) = \sum_{i=1}^{m} \frac{(h_a(i) - h_b(i))^2}{h_a(i) + h_b(i)}$$  \hspace{1cm} (4)

where $m$ is the number of color levels in quantized image. According to Boundary Prior, prior patch (boundary patch) belongs to healthy skin. When a patch is sampled from the healthy skin region, its color histogram is similar to that of prior patch, and the histogram distance between them is small. Contrarily,
when a patch is sampled from the lesion region, the corresponding color histogram distance is large. Therefore, for the color saliency map $S_{ci}$, the lesion region has higher value than the healthy skin region.

For the image with incomplete lesion, lesion regions exist in some prior patches, which may lead to an unsatisfactory saliency result. To deal with this situation, the final color saliency map is obtained through the weighted sum of all the $n$ saliency maps:

$$S_c(j) = \sum_{i=1}^{n} \omega_i S_{ci}(j) \quad j = 1, 2, ..., L_c$$

(5)

where $\omega_i$ is the weighting factor in (2). When the prior patch is filled with lesion, its weighting factor is small and its effect on the final $S_c$ is small. Obviously, the size of $S_c$ is $1/\tau_2^2$ of the original image. We enlarge it to the original image size and normalize the enlarged image $S_c'$ as:

$$S_{c,\text{norm}} = \frac{S_c' - \min (S_c')}{\max (S_c') - \min (S_c')}.$$  

(6)

Figure 4: Color saliency maps of dermoscopy images. (a) Original images; (b) Results of simple average of $S_{ci}$; (c) Results of weighting sum of $S_{ci}$ using (5).

Fig. 4 shows color saliency maps of several dermoscopy images with different conditions. It can be seen that, the lesions in the first two columns, which have
low contrast and fuzzy border, are effectively enhanced, and the hairs in the third
column are greatly weakened because of the patch-based calculation. Lesions
in the first three columns are completely contained in images (i.e. all of the
boundary patches belong to background). Simple average of $S_{ci}$ is close to the
weighting sum of $S_{ci}$. The last column is a case of incomplete lesion. Its prior
patches corresponding to healthy skin has higher weights, thus the weighting
sum of $S_{ci}$ has more satisfactory result than the simple average of $S_{ci}$.

2.1.3. Saliency calculation in the brightness space

According to Brightness Prior, the skin region has higher brightness than
the lesion region. Therefore, the brightness saliency map generated by ith prior
patch can be defined as:

$$S_{bi}(j) = \max \{0, \mu_i - I(j)\} \quad j = 1, 2, ..., L$$  \hspace{1cm} (7)

where $L$ is number of pixels in the image, $\mu_i$ is the mean of the ith prior patch.
The max operator is used to avoid negative value when pixel $j$ has higher bright-
ness than prior patch $i$. Obviously, $S_{bi}$ has the same size as the original image.

The same as the color space, in order to avoid the influence of lesion in the
boundary patch, the final brightness saliency map is calculated as:

$$S_b(j) \sum_{i=1}^{n} \omega_i S_{bi}(j) \quad j = 1, 2, ..., L$$  \hspace{1cm} (8)

and normalized as:

$$S_{b,norm} = \frac{S_b - \min (S_b)}{\max (S_b) - \min (S_b)}.$$  \hspace{1cm} (9)

The brightness saliency maps of Fig. 4(a) are shown in Fig. 5. Compared
with the color saliency maps in Fig. 4(c), brightness saliency maps have clearer
healthy skin background.

2.1.4. Saliency map fusion strategy

We fuse the color saliency map with the brightness saliency map to generate
the final enhanced image by following equation:

$$S(j) = (S_{c,norm}(j))^\alpha \times (S_{b,norm}(j))^\beta \quad j = 1, 2, ..., L \quad \hspace{1cm} (10)$$
where $\alpha$ ranges from 1 to 10 and $\beta$ ranges from 0 to 1.

For the color saliency map $S_{c,norm}$, the saliency value is usually close to 1 in lesion region, and small (usually not close to 0) in healthy skin region, as shown in Fig. 4(c). Therefore, $S_{c,norm}$ is stretched through setting $\alpha$ greater than 1 in (10), which can lower the saliency value of healthy skin greatly and meanwhile, keep the saliency value of lesion almost invariant. $\alpha$ and $\beta$ control the relative importance between $S_{c,norm}$ and $S_{b,norm}$. Considering too large $\alpha$ value will cause $S_{b,norm}$ invalidated, $\alpha$ is limited to less than 10 here. For the brightness saliency map $S_{b,norm}$, the saliency value is usually close to 0 in healthy skin region, and high (usually not close to 1) in lesion region, as shown in Fig. 5. Similarly, we set $\beta$ between 0 and 1 to stretch $S_{b,norm}$.

For a dermoscopy image, many enhancement results are generated when changing $\alpha$ and $\beta$, and their histograms presents bimodal distribution. Among these enhanced images, the one with the minima segmentation error rate $E_{\text{min}}$ is regarded as the best enhanced image:

$$E_{\text{min}} = \min_{i=0,1,\ldots,255} ER_i$$

$$ER_i = \frac{\text{ErrorPixels}_i}{\text{Area}},$$

where $\text{ErrorPixels}_i$ is the number of false segmented pixels when threshold is $i$, and $\text{Area}$ is the total number of pixels in the image.

We determine the optimal $\alpha$ and $\beta$ by grid search method. 50 dermoscopy images are used for verification. Finally, $\alpha = 1.52$ and $\beta = 0.48$ are determined as the optimal parameters in this paper. Fig. 6 is the fusion result of Fig. 10.
Fig. 4(c) and Fig. 5 using (10) with the optimal $\alpha$ and $\beta$. Obviously, objects are enhanced effectively and the backgrounds are suppressed.

Figure 6: Enhancement result using Fig. 4(c) and Fig. 5 based on (10).

2.2. Adjustment on Otsu threshold method

Fig. 7 presents two dermoscopy images and the histograms of their enhanced images obtained by (10). The enhanced images have high contrast and their histograms have two peaks corresponding to lesion and healthy pixels, respectively. For the image with high contrast and two-class objects, Otsu threshold [26] is a simple and effective method to extract object borders. However, when there are fuzzy borders and unbalanced area ratio between object and background, the threshold obtained by Otsu method usually is inclined to the side with larger area, which may cause over-segmentation or under-segmentation. Figure 7(d) is the segmentation results for (a) using Otsu threshold [26] on enhanced images. It can be seen that the first image has balanced area ratio between the lesion and the healthy skin, thus its segmentation result is satisfactory. For the second image, the lesion has much larger area than the healthy skin and its border is fuzzy, under-segmentation happened.

To avoid from over-segmentation and under-segmentation, we design two functions: $F_h$ and $F_w$. $F_h$ is used to measure the altitude difference (frequency difference) between gray level $t$ and the second highest peak of the histogram. $F_w$ is used to measure the closeness degree between $t$ and the traditional Otsu threshold. Based on the two functions, the traditional Otsu threshold method is adjusted and more accurate lesion borders are extracted.
2.2.1. The design of $F_h$

Certain bins in the histogram can be zero when there is no pixel at those gray levels. In order to avoid the influence of 0 values, the histogram is first smoothed through a max filter of size 5 as:

$$H_s(t) = \max_{t-2 \leq i \leq t+2} H_{en}(i)$$

(13)

where $H_{en}$ represents the histogram of the enhanced image. All the subsequent calculations are performed on the smoothed histogram $H_s$.

For a gray level $t$ between the two peaks of $H_s$, the altitude difference between $t$ and the second highest peak of $H_s$ is defined as:

$$F_h(t) = \min \{H_s(t_l), H_s(t_r)\} - H_s(t)$$

(14)

where $t_l$ and $t_r$ are the gray levels corresponding to the left peak and right peak respectively, $t_l \leq t \leq t_r$. Let $t_{otsu}$ be the Otsu threshold on $H_{en}$, the left peak and right peak can be located by searching the maximum from both sides of $t_{otsu}$ on $H_s$. Obviously, when $t$ locates at the lowest point between the two peaks, $H_s$ reaches the minimum and $F_h$ reaches the maximum.
2.2.2. The design of $F_w$

Function $F_w$ is designed to measure the closeness between $t$ and Otsu threshold $t_{otsu}$, which monotonically decreases on both sides of $t_{otsu}$, and reaches the maximum when $t = t_{otsu}$:

$$F_w(t) = \frac{1}{c} (X(t))^a(1-X(t))^b, \quad (15)$$

$$X(t) = \frac{t-t_l}{t_r-t_l}, \quad t_l \leq t \leq t_r, \quad (16)$$

where $c$ is a parameter more than zero, $X(t)$ and $(1-X(t))$ represent the distance from $t$ to the left peak and the right peak, respectively.

To insure $F_w$ reaches its extremum when $t = t_{Otsu}$, its derivative should meet:

$$\frac{dF_w}{dt}(t_{otsu}) = 0. \quad (17)$$

Let

$$F_w(t_{otsu}) = 1 \quad (18)$$

according to (17) and (18), $a$ and $b$ can be represented by the parameter $c$ as follows (the detailed confirmation is given in the appendix):

$$a = \frac{X_{otsu} \log (c)}{X_{otsu} \log (X_{otsu}) + (1-X_{otsu}) \log (1-X_{otsu})} \quad (19)$$

$$b = \frac{(1-X_{otsu}) \log (c)}{X_{otsu} \log (X_{otsu}) + (1-X_{otsu}) \log (1-X_{otsu})} \quad (20)$$

where $X_{otsu} = X(t_{otsu})$. Fig. shows the curves of $F_w$ with different values of $c$, where $t_l = 0$, $t_r = 255$, $t_{otsu} = 120$. $F_w$ is used to measure the closeness between $t$ and $t_{otsu}$, and it should reach the maximum at $t_{otsu}$. Therefore, the value of $c$ is within (0,1).

2.2.3. Threshold optimization

The enhanced image in section 2.1 has bimodal distribution histogram. Its optimal threshold should be close to the lowest point between the two peaks, and
Figure 8: The curves of $F_w(t)$ with different values of $c$.

as well as close to the Otsu threshold. Hence we combine $F_h$ and $F_w$ together to obtain the optimal threshold $T_{opt}$ as:

$$T_{opt} = \arg\max_t F_h(t) F_w(t).$$  \hspace{1cm} (21)

Most segmentation results have spots and holes. We use morphological operations as post-processing to remove these spots and holes. Fig. 8 shows the segmentation results for Fig. 7(a) by traditional and adjusted Otsu threshold methods with different $c$ values, where $t_{otsu}$ is the traditional Otsu threshold, $t_{c=0.3}$ and $t_{c=0.8}$ are the improved thresholds when $c = 0.3$ and $c = 0.8$, respectively. Obviously, $c$ is an adjustment parameter for traditional Otsu threshold, and it can prompt the optimal threshold to move to the correct location. Generally, when $c$ takes range value from 0.5 to 0.9, the segmentation result is satisfactory. For the image with distinct lesion border, the value of $c$ has very slight impact on the results and both the traditional and adjusted Otsu threshold methods can achieve satisfactory segmentation results. However, for the image with fuzzy lesion border and unbalanced area ratio, the adjusted Otsu threshold method (with $c = 0.8$) significantly outperforms the traditional Otsu method, as shown in the second row of Fig. 9.
Figure 9: Segmentation results of different thresholds for Fig. 7 (a). (a) Different thresholds for Fig. 7(a); (b) Segmentation results of Fig. 7(a).

3. Experimental results and analysis

A series of experiments are conducted using MATLAB 2016 on Win10 OS with i7 3.4 GHz quad-core CPU and 4 GB RAM. Three image datasets are used:

- **EDRA**: Selected from the CD resource EDRA-CDROM [27], composed of 566 dermoscopy images including 438 benign and 128 malignant. The ground truth (GT) of Dataset1 are manually marked by Jie Liu, an experienced dermatologist.

- **PH2**: A publicly available dataset named PH2 [28], contains 200 melanocytic lesions (80 benign, 80 atypical and 40 malignant), and the ground truth is available.

- **ISBI 2016**: A publicly available dataset used in ISBI 2016 challenge [29], contains 900 dermoscopy images (727 benign and 173 malignant), and the
ground truth is provided.

All the images in three datasets are not crossed with the 50 dermoscopy images used for the determination of the optimal parameters $\alpha$ and $\beta$ in section 2.1.

Parameters of the proposed approach are set as: $N = 13$, $c = 0.8$, $\tau_1 = \tau_2 = 11$. To evaluate the performance of segmentation methods quantitatively, 4 metrics including precision, recall, error, and dice similarity coefficient (DSC) are used:

\[
Precision = \frac{TP}{TP + FP} \tag{22}
\]

\[
Recall = \frac{TP}{TP + FN} \tag{23}
\]

\[
Error = \frac{FP + FN}{TP + FP + TN + FN} \tag{24}
\]

\[
DSC = \frac{2TP}{2TP + FP + FN} \tag{25}
\]

Their definitions are based on the concepts of true/false positive/negative defined in Tab. 1. A satisfactory segmentation method has high values of precision, recall, DSC, and a low value of error.

<table>
<thead>
<tr>
<th>Pixel in ground truth border</th>
<th>Pixel in segmentation border</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lesion</td>
<td>True Positive (TP)</td>
</tr>
<tr>
<td>Background</td>
<td>False Negative (FN)</td>
</tr>
</tbody>
</table>

3.1. Selection of brightness space

A variety of brightness spaces have been utilized in dermoscopy image segmentation. Garnavi et al. [30] segmented dermoscopy images by Otsu threshold method in different color and brightness spaces and selected the best one for
Abbas et al. [16] extracted and analyzed texture information using the brightness channel J in Jch space. In this paper, brightness space is used to generate weighting factors $\omega_i$ and calculate the saliency map. It is necessary to select an optimal brightness space here.

Six brightness spaces are compared in this experiment, including the classic gray space (gray), L channel in Lab space (Lab-L), I channel in HSI space (HSI-I), V channel in HSV space (HSV-V), L channel in HSL space (HSL-L), and the minimum of three channels in RGB space (minRGB, its definition is given in [I]). Area under curve (AUC) are adopted to evaluate the effectiveness of saliency enhancement for lesion in different brightness spaces. A high AUC value indicates a good algorithm performance. Table 2 is the AUC statistic results of six brightness space on three datasets. It can be seen that minRGB space achieves the highest AUC values. Therefore, minRGB space is chosen as the optimal brightness space for our saliency algorithm.

<table>
<thead>
<tr>
<th>Space</th>
<th>EDRA</th>
<th>PH²</th>
<th>ISBI 2016</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gray</td>
<td>0.9561</td>
<td>0.9615</td>
<td>0.9339</td>
</tr>
<tr>
<td>Lab-L</td>
<td>0.9577</td>
<td>0.9605</td>
<td>0.9350</td>
</tr>
<tr>
<td>HSI-I</td>
<td>0.9565</td>
<td>0.9633</td>
<td>0.9315</td>
</tr>
<tr>
<td>HSV-V</td>
<td>0.9551</td>
<td>0.9436</td>
<td>0.9044</td>
</tr>
<tr>
<td>HSL-L</td>
<td>0.9624</td>
<td>0.9631</td>
<td>0.9251</td>
</tr>
<tr>
<td>minRGB</td>
<td>0.9666</td>
<td>0.9671</td>
<td>0.9391</td>
</tr>
</tbody>
</table>

3.2. Analysis on the saliency algorithms

In section 2.1, a saliency method based on healthy skin priors is proposed to enhance the lesion in the images. We compare it with 6 state-of-the-art saliency algorithms including a frequency approaches (MSSS [31]), a global approach (L-C [32]), a local approach (RC [19]), and three boundary-prior-based approaches (MR [24], wCtr* [33] and SLSS [21]). Fig. 10 shows some example results and corresponding ground truth. Obviously, all the algorithms are able to achieve satisfactory results for the images with simple background and high contrast, as shown in the first column of Fig. 10. However, for the images with fuzzy
border, low contrast, and incomplete lesion, our method outperforms others, as the last three columns in Fig. 10. Table 3 is the AUC statistics of different saliency methods on three datasets. Among six compared saliency algorithms, three boundary-prior-based approaches (MR, wCtr*, SLSS) perform better than other 3 compared methods, which indicates that boundary prior is effective for lesion enhancement. For further improving algorithm performance, we combined boundary prior with color prior and brightness prior. With the highest AUC values on three datasets, our saliency method outperforms the 6 compared methods.

Figure 10: Example results of
Table 3: AUC statistics of different saliency methods on three datasets

<table>
<thead>
<tr>
<th>Saliency method</th>
<th>EDRA</th>
<th>PH²</th>
<th>ISBI 2016</th>
<th>ISBI 2016</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSSS</td>
<td>0.5230</td>
<td>0.7005</td>
<td>0.8058</td>
<td></td>
</tr>
<tr>
<td>LC</td>
<td>0.6870</td>
<td>0.6492</td>
<td>0.7583</td>
<td></td>
</tr>
<tr>
<td>RC</td>
<td>0.7015</td>
<td>0.7953</td>
<td>0.8009</td>
<td></td>
</tr>
<tr>
<td>MR</td>
<td>0.7668</td>
<td>0.9321</td>
<td>0.9041</td>
<td>0.9041</td>
</tr>
<tr>
<td>wCtr*</td>
<td>0.9585</td>
<td>0.9552</td>
<td>0.9135</td>
<td></td>
</tr>
<tr>
<td>SLSS</td>
<td>0.9564</td>
<td>0.9595</td>
<td>0.9286</td>
<td></td>
</tr>
<tr>
<td>Ours</td>
<td><strong>0.9666</strong></td>
<td><strong>0.9671</strong></td>
<td><strong>0.9391</strong></td>
<td></td>
</tr>
</tbody>
</table>

In our segmentation framework, the saliency method is used to improve the segmentation accuracy. For further evaluation, all the saliency maps are segmented by traditional Otsu threshold method. Tables 4, 5 and 6 are the statistics for the 4 segmentation metrics on three datasets respectively. When a segmentation algorithm has high precision and low recall, under-segmentation happens; otherwise, over-segmentation appears. When both precision and recall metrics are high, the algorithm is considered to be satisfactory. Relative to the six compared methods, with a high precision, our method achieves the best recall, DSC and error values. Therefore, our saliency method is more adaptive for enhancing lesion objects.

Table 4: Traditional Otsu segmentation result statistics using different saliency maps on EDRA

<table>
<thead>
<tr>
<th>Saliency method</th>
<th>Precision(%)</th>
<th>Recall(%)</th>
<th>Error(%)</th>
<th>DSC(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSSS</td>
<td>96.89</td>
<td>58.91</td>
<td>17.81</td>
<td>59.16</td>
</tr>
<tr>
<td>LC</td>
<td>89.28</td>
<td>52.00</td>
<td>20.31</td>
<td>67.23</td>
</tr>
<tr>
<td>RC</td>
<td>88.08</td>
<td>65.60</td>
<td>16.54</td>
<td>73.19</td>
</tr>
<tr>
<td>MR</td>
<td>99.50</td>
<td>76.75</td>
<td>7.59</td>
<td>86.18</td>
</tr>
<tr>
<td>wCtr*</td>
<td><strong>99.60</strong></td>
<td>78.85</td>
<td>7.90</td>
<td>87.53</td>
</tr>
<tr>
<td>SLSS</td>
<td>98.95</td>
<td>67.80</td>
<td>13.10</td>
<td>78.36</td>
</tr>
<tr>
<td>Ours</td>
<td>96.85</td>
<td><strong>94.05</strong></td>
<td><strong>2.67</strong></td>
<td><strong>94.08</strong></td>
</tr>
</tbody>
</table>

3.3. The performance of segmentation method

In this paper, the dermoscopy images are first enhanced through the proposed saliency method, and then segmented using the adjusted Otsu threshold. The proposed segmentation framework is compared with six widely-used seg-
Table 5: Traditional Otsu segmentation result statistics using different saliency maps on PH2

<table>
<thead>
<tr>
<th>Saliency method</th>
<th>Precision(%)</th>
<th>Recall(%)</th>
<th>Error(%)</th>
<th>DSC(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSSS</td>
<td>74.67</td>
<td>35.47</td>
<td>26.89</td>
<td>45.81</td>
</tr>
<tr>
<td>LC</td>
<td>42.63</td>
<td>30.78</td>
<td>33.12</td>
<td>33.10</td>
</tr>
<tr>
<td>RC</td>
<td>73.25</td>
<td>64.34</td>
<td>22.55</td>
<td>63.39</td>
</tr>
<tr>
<td>MR</td>
<td>92.67</td>
<td>75.85</td>
<td>13.23</td>
<td>81.04</td>
</tr>
<tr>
<td>wCtr*</td>
<td>96.75</td>
<td>71.14</td>
<td>13.72</td>
<td>79.71</td>
</tr>
<tr>
<td>SLSS</td>
<td>98.21</td>
<td>60.25</td>
<td>17.86</td>
<td>71.30</td>
</tr>
<tr>
<td>Ours</td>
<td>96.62</td>
<td>86.86</td>
<td>6.78</td>
<td>88.75</td>
</tr>
</tbody>
</table>

Table 6: Traditional Otsu segmentation result statistics using different saliency maps on ISBI 2016

<table>
<thead>
<tr>
<th>Saliency method</th>
<th>Precision(%)</th>
<th>Recall(%)</th>
<th>Error(%)</th>
<th>DSC(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSSS</td>
<td>83.55</td>
<td>32.04</td>
<td>21.95</td>
<td>42.76</td>
</tr>
<tr>
<td>LC</td>
<td>68.37</td>
<td>39.30</td>
<td>21.86</td>
<td>46.63</td>
</tr>
<tr>
<td>RC</td>
<td>76.75</td>
<td>59.01</td>
<td>15.40</td>
<td>59.55</td>
</tr>
<tr>
<td>MR</td>
<td>86.25</td>
<td>67.37</td>
<td>15.40</td>
<td>67.39</td>
</tr>
<tr>
<td>wCtr*</td>
<td>92.63</td>
<td>60.62</td>
<td>13.68</td>
<td>69.60</td>
</tr>
<tr>
<td>SLSS</td>
<td>96.46</td>
<td>51.77</td>
<td>15.74</td>
<td>64.13</td>
</tr>
<tr>
<td>Ours</td>
<td>94.01</td>
<td>73.81</td>
<td>10.35</td>
<td>79.26</td>
</tr>
</tbody>
</table>

mentation methods for dermoscopy images, including: threshold-based Otsu [25], active-contour-based RACs [16] and GAC [17], region-based SRM [15], clustering-based SGNN [14], and saliency-based SLSS [21]. Fig. 11 demonstrates the benefits of our algorithm, where red line represents the ground truth and green line indicates the automatic segmentation result. The last two columns in Fig. 11 are the segmentation results by our methods, where Method1 is our saliency method combined with traditional Otsu threshold and Method2 is our final segmentation framework (saliency + adjusted Otsu). The image in the first row of Fig. 11 is simple and the contrast between lesion and background is strong, in this case, all the methods can obtain satisfactory results. From the second row to the last row, the images are complex (low contrast, fuzzy boundaries, hairs, large lesion regions, and variegated color). For these challenging images, our two methods outperform the six compared methods. In addition, comparing the last two columns in the fifth row of Fig. 11 it can be seen that our Method2 can obtain more excellent result than Method1, which indicates
that our adjusted Otsu threshold is more suitable for our saliency method.

Figure 11: Example results of different segmentation methods. (a) Original images; (b) Otsu [26]; (c) RACs [16]; (d) GAC [17]; (e) SRM [15]; (f) SGNN [14]; (g) SLSS [21]; (h) Our Method1; (i) Our Method2.

Tab. 7, 8 and 9 show the average statistics for all segmentation methods on three datasets respectively, where SC [22] is another saliency-based segmentation method for dermoscopy images and its statistics are from the original reference. Because EDRA is not employed in [22], the statistics result on EDRA is not presented in Table 7. Relative to other methods, our final segmentation framework (Method2) achieves high precision and recall metrics with a good balance. The outstanding error and DSC values also indicate the superior performance of our method.
Table 7: Segmentation result statistics of different segmentation methods on EDRA

<table>
<thead>
<tr>
<th>Method</th>
<th>Precision(%)</th>
<th>Recall(%)</th>
<th>Error(%)</th>
<th>DSC(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Otsu [26]</td>
<td>99.16</td>
<td>73.48</td>
<td>8.51</td>
<td>82.39</td>
</tr>
<tr>
<td>RACs [16]</td>
<td><strong>99.32</strong></td>
<td>78.78</td>
<td>6.47</td>
<td>87.21</td>
</tr>
<tr>
<td>GAC [17]</td>
<td>89.26</td>
<td>93.41</td>
<td>7.18</td>
<td>85.42</td>
</tr>
<tr>
<td>SRM [15]</td>
<td>98.69</td>
<td>76.44</td>
<td>7.65</td>
<td>84.18</td>
</tr>
<tr>
<td>SGNN [14]</td>
<td>92.19</td>
<td>91.28</td>
<td>5.55</td>
<td>90.46</td>
</tr>
<tr>
<td>SLSS [21]</td>
<td>95.18</td>
<td>82.86</td>
<td>6.02</td>
<td>88.34</td>
</tr>
<tr>
<td>SC [22]</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Our Method1</td>
<td>96.85</td>
<td>94.05</td>
<td>2.67</td>
<td>94.08</td>
</tr>
<tr>
<td>Our Method2</td>
<td>97.65</td>
<td>93.85</td>
<td><strong>2.50</strong></td>
<td><strong>95.32</strong></td>
</tr>
</tbody>
</table>

Table 8: Segmentation result statistics of different segmentation methods on PH2

<table>
<thead>
<tr>
<th>Method</th>
<th>Precision(%)</th>
<th>Recall(%)</th>
<th>Error(%)</th>
<th>DSC(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Otsu [26]</td>
<td>96.37</td>
<td>73.45</td>
<td>11.45</td>
<td>81.33</td>
</tr>
<tr>
<td>RACs [16]</td>
<td>93.77</td>
<td>82.44</td>
<td>9.21</td>
<td>85.90</td>
</tr>
<tr>
<td>GAC [17]</td>
<td>80.31</td>
<td><strong>90.63</strong></td>
<td>13.49</td>
<td>77.67</td>
</tr>
<tr>
<td>SRM [15]</td>
<td>82.79</td>
<td>88.21</td>
<td>10.90</td>
<td>82.15</td>
</tr>
<tr>
<td>SGNN [14]</td>
<td>87.83</td>
<td>88.44</td>
<td>8.48</td>
<td>84.54</td>
</tr>
<tr>
<td>SLSS [21]</td>
<td>91.53</td>
<td>83.17</td>
<td>9.88</td>
<td>85.30</td>
</tr>
<tr>
<td>SC [22]</td>
<td>86.00</td>
<td>–</td>
<td>–</td>
<td>86.00</td>
</tr>
<tr>
<td>Our Method1</td>
<td>96.62</td>
<td>86.86</td>
<td>6.78</td>
<td>88.75</td>
</tr>
<tr>
<td>Our Method2</td>
<td><strong>96.78</strong></td>
<td>87.03</td>
<td><strong>6.40</strong></td>
<td><strong>89.35</strong></td>
</tr>
</tbody>
</table>

Table 9: Segmentation result statistics of different segmentation methods on ISBI 2016

<table>
<thead>
<tr>
<th>Method</th>
<th>Precision(%)</th>
<th>Recall(%)</th>
<th>Error(%)</th>
<th>DSC(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Otsu [26]</td>
<td>86.18</td>
<td>54.10</td>
<td>15.64</td>
<td>61.43</td>
</tr>
<tr>
<td>RACs [16]</td>
<td>96.68</td>
<td>65.85</td>
<td>11.24</td>
<td>74.69</td>
</tr>
<tr>
<td>GAC [17]</td>
<td>70.46</td>
<td><strong>93.13</strong></td>
<td>14.05</td>
<td>70.94</td>
</tr>
<tr>
<td>SRM [15]</td>
<td>87.15</td>
<td>60.51</td>
<td>14.91</td>
<td>67.93</td>
</tr>
<tr>
<td>SGNN [14]</td>
<td>87.45</td>
<td>72.64</td>
<td>12.78</td>
<td>75.21</td>
</tr>
<tr>
<td>SLSS [21]</td>
<td>96.88</td>
<td>62.77</td>
<td>11.52</td>
<td>73.97</td>
</tr>
<tr>
<td>SC [22]</td>
<td>86.00</td>
<td>–</td>
<td>–</td>
<td>80.00</td>
</tr>
<tr>
<td>Our Method1</td>
<td>94.01</td>
<td>73.81</td>
<td>10.35</td>
<td>79.26</td>
</tr>
<tr>
<td>Our Method2</td>
<td><strong>97.28</strong></td>
<td>74.70</td>
<td><strong>8.20</strong></td>
<td><strong>81.83</strong></td>
</tr>
</tbody>
</table>

4. Conclusion

A novel segmentation method is proposed for dermatoscopy images in this paper, which first enhances the image using saliency method, and then obtains the lesion border through adjusted Otsu threshold. Lesions are salient in der-
moscopy images. Healthy skin, usually distributing in the image boundary, has uniform color and high brightness. According to these priors, color saliency map and brightness saliency map are constructed in the color and brightness spaces respectively. The two maps are fused through a designed fusion function, thus the dermoscopy image is enhanced effectively. The proposed saliency enhancement method is robust to challenging conditions, such as low contrast, incomplete object and hairs. Enhanced images have high contrast and their histograms have bimodal distribution, thus Otsu threshold method is employed to extract lesion borders in segmentation stage. However, over-segmentation or under-segmentation may happen on fuzzy border when the area ratio between the lesion and the healthy skin is unbalanced. In order to improve segmentation accuracy, an optimization function is designed according to the histogram distribution of the enhanced image. This optimization function can adjust the threshold to move to the correct location and obtain more accurate lesion borders. A series of experiments are done on three datasets. Experimental results show that our approach achieve more satisfactory results than other segmentation methods.

Appendix

Equation (15) is a function of gray level $t$. Its derivative can be written as:

$$\frac{dF_w}{dt} = \frac{1}{c} \left[ aX^{a-1}(1-X)^b - bX^a(1-X)^{b-1} \right] dX \, dt. \quad \text{(A.1)}$$

Due to $\frac{dF_w}{dt} (t_{otsu}) = 0$, we have

$$a (1 - X_{otsu}) - bX_{otsu} = 0. \quad \text{(A.2)}$$

Equation (18) shows:

$$F_w (t_{otsu}) = \frac{1}{c} (X_{otsu})^a (1 - X_{otsu})^b = 1 \quad \text{(A.3)}$$

where $X_{otsu} = X (t_{otsu})$. We simplify (A.3) by taking logarithms:

$$a \log (X_{otsu}) + b \log (1 - X_{otsu}) = \log (c) \quad \text{(A.4)}$$
According to (A.2) and (A.4), parameters $a$ and $b$ are obtained as:

$$a = \frac{X_{otsu} \log(c)}{X_{otsu} \log(X_{otsu}) + (1 - X_{otsu}) \log(1 - X_{otsu})} \quad (A.5)$$

$$b = \frac{(1 - X_{otsu}) \log(c)}{X_{otsu} \log(X_{otsu}) + (1 - X_{otsu}) \log(1 - X_{otsu})} \quad (A.6)$$

The proof of (19) and (20) is thus completed.

Acknowledgements

This work was supported by the National Natural Science Foundation of China under Grants 61471016 and 61371134.

References


