This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2017.2653179, IEEE Journal of Biomedical and Health Informatics

Saliency-based Lesion Segmentation via Background Detection in Dermoscopic Images

Euijoon Ahn, Jinman Kim*, *Member, IEEE*, Lei Bi, Ashnil Kumar, *Member, IEEE*, Changyang Li, Michael Fulham and David Dagan Feng, *Fellow, IEEE*

***Abstract*— The segmentation of skin lesions in dermoscopic images is a fundamental step in automated computer-aided diagnosis (CAD) of melanoma. Conventional segmentation methods, however, have difficulties when the lesion borders are indistinct and when contrast between the lesion and the surrounding skin is low. They also perform poorly when there is a heterogeneous background or a lesion that touches the image boundaries; this then results in under- and over-segmentation of the skin lesion. We suggest that saliency detection using the reconstruction errors derived from a sparse representation model coupled with a novel background detection can more accurately discriminate the lesion from surrounding regions. We further propose a Bayesian framework that better delineates the shape and boundaries of the lesion. We also evaluated our approach on two public datasets comprising 1100 dermoscopic images and compared it to other conventional and state-of-the-art unsupervised (i.e. no training required) lesion segmentation methods, as well as the state-of-the-art unsupervised saliency detection methods. Our results show that our approach is more accurate and robust in segmenting lesions compared to other methods. We also discuss the general extension of our framework as a saliency optimisation algorithm for lesion segmentation.**

***Index Terms*— Dermoscopic Image, Lesion Segmentation, Saliency Detection, Computer-aided diagnosis**

I. Introduction

Malignant melanoma is a skin disease that has a significant morbidity and mortality worldwide [1]. Malignant melanoma can be cured, however, by simple lesion excision if diagnosed at an early stage. Dermoscopy is a non-invasive image diagnosis technique for the in vivo observation of pigmented skin lesions in dermatology [2]. This technique allows dermatologists to detect early stage melanoma that are not visible to the human eye. Manual visual interpretation by dermatologists, however, can be time consuming, subjective and non-reproducible [3]. Well-trained dermatologists show inter-observer variability and produce different analyses when delineating the same skin lesion [4]. Therefore, an accurate and automatic detection of melanoma from dermoscopic images has been of great interest in computer aided diagnosis (CAD) systems. Automatic lesion segmentation is a fundamental requirement for melanoma CAD as its accuracy determines the overall analysis of skin lesion including disease classification and quantification. Automated skin lesion segmentation, however, is challenging due to variations in lesion size, shape, fuzzy boundaries and different skin colours. Complicated background regions (e.g. heterogeneous background [5] or lesion touching image boundaries) in dermoscopic images also make it difficult to identify the lesions and separate them from surrounding regions. Artifacts such as hair, specular reflections and colour calibration charts are additional obstacles that lower the accuracy of the segmentation.

A number of different algorithms have been proposed to effectively segment lesion regions from background skin. The most common and simple automatic segmentation methods are based on histogram thresholding [6-12], which creates a histogram using RGB channels, luminance or principal component analysis. Silveira *et al.* [6] proposed an adaptive thresholding (AT) method to segment lesions by comparing the colour of each pixel with a threshold and classified it as a lesion if it is darker than the threshold. It used image entropy to select the best channel from RGB to facilitate the discrimination; the threshold was then automatically determined as the local minimum between the maxima, plus a small offset to account for quantization issues. While thresholding methods are easily adopted because of their simplicity, they are limited by the luminance distribution and may fail if there are multiple peaks in the luminance histogram.

Deformable models, such as active contours, have also been used [6, 13-17]. The key concept with these methods is to minimize the internal forces defined within the curve and external (image-based) energies to find an optimum boundary of the skin lesion. A major drawback is the reliance on the optimal selection of the segmentation parameters and deformable models also often require an ineffective stopping criterion that generates unnecessary computations. Ma *et al.* [13] proposed a new deformable model (NDM) approach to address these limitations. In their work, the CIE L*a*b* and L*u*v* colour models were applied to effectively differentiate normal skin to the lesion. The differences in the colour channels were combined together to define the stopping criterion [13]. However, this work was not fully automated and its performance varied widely depending on the initial parameter estimation [11] and morphological post-processing [6, 18].

E. Ahn, J. Kim, L. Bi, A. Kumar, 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); J. Kim (e-mail: jinman.kim@sydney.edu.au); L. Bi (e-mail: lei@it.usyd.edu.au); A. Kumar (e-mail: ashnil.kumar@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.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2017.2653179, IEEE Journal of Biomedical and Health Informatics

Region based segmentation is another well-known approach for dermoscopic images [6, 18, 19]. These techniques divide an image into different regions based on pre-defined unique features such as colour, intensity, objects or wavelets. For example, Celebi *et al.* [20] expanded the region growing with a statistical region merging (SRM) [21], which treated an input as an observed instance of an unknown theoretical image and reconstructed the input image based on statistical colour distribution of lesion regions. Silveira *et al.* [6] used a Chan-based level set segmentation algorithm (C-LS) [22] with the traditional active contour model (using mean curvature motion techniques), and the Mumford-Shah model. Sadri *et al.* [23] proposed a fixed-grid wavelet network approach. Here, the RGB channels of the images were considered as the input wavelets and an orthogonal least squares algorithm was used to compute network weights and optimize the network structure. Bi *et al.* [24] recently proposed a cellular automata (CA) based approach for lesion segmentation. They used image-wise learning technique to derive a probabilistic map for automated seed selection. In our prior work, we proposed a saliency based method through measuring sparse reconstruction errors from skin lesions against a background template of a set of regions located at the image boundaries [25]. Although the reconstruction error improved the lesion segmentation, it assumed that the lesions were located at the centre of the image and subsequently used image boundaries to separate the lesions from background normal skin regions. However, this approach could not localise lesions at the image boundaries. Apart from the new and widely studied approaches above, several other related methods have previously been compared and evaluated and they can be found in the references below [6, 26].

Recently, supervised learning approaches such as convolutional neural networks (CNNs) have been used for lesion segmentation in dermoscopic images. These methods produced good results in a challenge held in International Symposium on Biomedical Imaging (ISBI) 2016 [27]. CNNs are deep neural networks that require extensive learning and make a prediction at every pixel; each deeper layer or the network learns a semantically 'higher' level representation of the image data [28]. The successful learning-based algorithms such as CNNs require a very large number of parameters and thus needs a large volume of labelled training images, e.g. 1.2 million images for ImageNet (general image). Data acquisition, however, in medical domain is difficult and quality annotation is time consuming due to the complexity of manual annotation and subjectivity among clinicians [29].

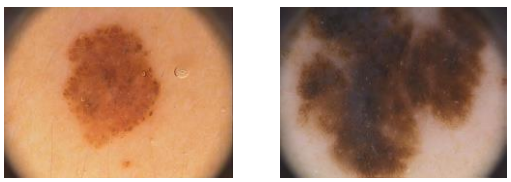
To address these limitations, we suggest a robust saliency-based skin lesion segmentation (RSSLS) framework, that uses reconstruction errors from a sparse representation model coupled with novel background detection. Saliency detection models identify visually salient regions or objects in an image by simulating the human visual and cognitive systems [30]. Saliency models have been applied to many computer vision applications including image segmentation and object-based image retrieval [30-32]. These studies underline that salient regions are visually more distinctive by virtue of their contrast. Hence, we hypothesize that the subtle contrast discrimination of the lesion from surrounding regions can be overcome via saliency detection. Unlike supervised learning algorithms, our RSSLS is an unsupervised approach which does not require any training. Instead, it makes use of robust image processing techniques to leverage the colour, spatial location, and structural cues, which are known to have a strong impact on visual attention [33].

In this work we present a major extension of our previous preliminary data where we first used saliency for lesion segmentation in dermoscopic images [25]. We now introduce: 1) A method to detect the background by analysing the spatial layout of the image regions, the boundaries and the inherent colour characteristics of dermoscopic images (see Fig 1). This allows better separation of the lesion from the background, in particular, for when the skin lesions touch the image boundaries. 2) A Bayesian framework to detect lesion boundaries more accurately and efficiently by refining the reconstruction error of each pixel by its similarities to pixels of lesion regions as well as its difference from those in the background regions using a colour histogram. 3) The results from a comprehensive evaluation of our approach compared to 'state-of-the-art' algorithms on a large test dataset.

A. Saliency Detection from Dermoscopic Images

Research from many cognitive neuroscientists has shown that the human visual system easily detects distinctive, or salient regions, in the visual fields [31-36]. These findings have now been recently applied to computer vision to find objects or regions that help to understand and represent an image. Itti *et al.* reported on one of the earliest saliency detection models using psychological theories of bottom-up attention based on centre-surround mechanisms [31]. Cheng *et al.* proposed a global histogram-based contrast (HC) algorithm [34] that used a pixel-wise colour separation, to detect salient regions and region-based contrast (RC) [35] algorithm. The RC algorithm improved on the HC by taking spatial distances into account at the cost of reduced computational efficiency. Yan *et al.* reported on a hierarchical model to analyse salient regions from multiple levels of structure [36]. These studies underlined that salient regions are more likely to be located centrally and are more conspicuous due to their contrast. Centre-surround contrast is problematic when a salient region is located at the image boundary and when a foreground region is globally compared to the rest of the image.

Some recent approaches use a background prior, rather than centre-surround contrast, to improve saliency detection and they have recorded impressive results [37-40]. Yang *et al.* used a graph-based manifold ranking (MR) that exploited the four boundaries of the input image as background queries to obtain background seeds, and then used the derived seeds to extract a foreground saliency map [37]. Jiang *et al.* formulated the saliency detection via absorbing Markov Chain where



(a) General Case (b) Extreme Case

Fig 1. Examples of skin lesions from dermoscopic images. (a) is a general case with the lesion entirely within the image and (b) is an extreme case when the lesion has multiple contacts with the image boundary.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2017.2653179, IEEE Journal of Biomedical and Health Informatics

boundary nodes are chosen as the absorbing nodes in a Markov chain [40]. Unlike the centre-surround approaches, the background prior method measures the local optimal saliency.

The skin lesions in dermoscopic images are more conspicuous, by colour and contrast, when compared to surrounding regions. This relates to the unique contrast between the lightness and saturation of skin lesions and normal surrounding tissue [13]. In dermoscopy skin lesions are often connected to the image boundary and fill the entire image (Fig 1b). Hence our hypothesis is that skin lesions can be detected from the surrounding regions using saliency detection by incorporating background detection coupled with inherent colour characteristics in the images.

II. Method

A. Overview of a Robust Saliency-based Skin Lesion Segmentation (RSSLS) Framework

Fig 2 is an overview of our RSSLS framework. Initially, hairs on the images are removed as a pre-processing step. A superpixel algorithm [41] partitions the image into a number of segments. A background template is then created by measuring the boundary connectivity of each image segment [39]. The images are reconstructed by measuring the sparse reconstruction errors against the background template as an indication of saliency. Context-based error propagation is then applied to smooth the reconstruction error. A pixel level saliency map is created by integrating multi-scale reconstruction errors, followed by refinement with a Bayesian framework.

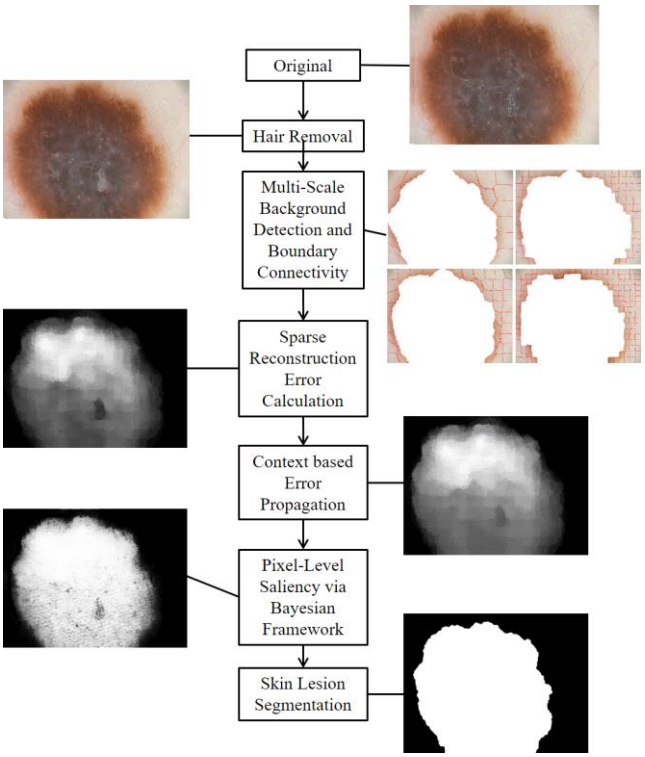


Fig 2. Overview of our RSSLS method.

B. Hair Removal

Visual artifacts degrade the performance of segmentation methods, e.g., skin hair may occlude parts of the lesion. We adopted the algorithm of Lee et al. [42] to remove visible hair from the images. This algorithm first identifies dark hair by applying a generalized gray-scale morphological closing operation with three structure elements at different directions. The low intensity values, the thick dark hair pixels, are located. The hair pixels are replaced by the nearby non-hair pixels and the final result is smoothed.

C. Boundary Connectivity

Our assessment is that skin lesions in the images can be divided into general and extreme cases based on the specific image conditions and spatial layout between the lesions and background (i.e. surrounding regions) (Fig 1). We define 'general' cases when the skin lesion has a general uniform color and is located at the image centre and not connected to the boundary (Fig 1a). We define 'extreme' cases as those where the lesion is connected to image boundaries, on at least two sides or 50% of one side of the lesion connected to the image boundary (Figure 1b). While the image boundaries can be used as good visual cues to separate general from the background it is inappropriate to simply treat all image boundaries as the background for the extreme cases. To address this problem, we propose a method that quantifies the boundary connectivity (*BC*) of an image region (i.e. area), according to [39]:

$$BC(A) = \frac{|\{\mathrm{X}|\mathrm{X} \in A,\ \mathrm{X} \in BndS\}|}{\sqrt{|\{\mathrm{X}|\mathrm{X} \in A\}|}}, \quad (1)$$

where X is an image segment and $BndS$ (Boundary Segments) is the set of image segments located at the image boundary within image area (A) (i.e. set of segments with same colour features). The Eq.1 explains the conceptual definition of boundary connectivity which quantifies the extent that an area (A) is connected to image boundary.

D. Background Detection

We used the simple linear iterative clustering (SLIC) [41] algorithm to partition the image into N segments ($\mathbf{X} = [\mathrm{X}_1, \mathrm{X}_2, \ldots, \mathrm{X}_N]$) where N is the total number of image segments. This helps to capture structural information that is difficult to be seen from the pixel-level features. SLIC is computationally efficient and accurate [41]. The input image is partitioned into segments with different sizes of N to create multi-scale background templates.

An undirected graph was constructed by connecting all adjacent segments (X, Y) with their Euclidean distance $d_e(\mathrm{X}, \mathrm{Y})$ values computed by average colour differences in the CIELAB colour space. The geodesic distance between any two segments $d_g(\mathrm{X}, \mathrm{Y})$ was obtained as the sum of node d_e values along their shortest path on the graph. This can be achieved by solving following minimization problem:

$$d_g(\mathrm{X}, \mathrm{Y}) = \min_{\mathrm{X}_1 = \mathrm{X}, \mathrm{X}_2 \cdots \mathrm{X}_N = \mathrm{Y}} \sum_{i=1}^{N-1} d_e(\mathrm{X}_i, \mathrm{X}_{i+1}) \quad (2)$$

When X and X_i share the same color features located in same

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2017.2653179, IEEE Journal of Biomedical and Health Informatics

region (i.e. $d_e(\mathrm{X}, \mathrm{X}_i) = 0$, set of segments), we can assume that $d_g(\mathrm{X}, \mathrm{X}_i) = 0$. The soft spanning area of each segment X can be then defined as:

$$Area(\mathrm{X}) = \sum_{i=1}^{N} \exp\left(-\frac{d_g^2(\mathrm{X}, \mathrm{X}_i)}{2\sigma_{clr}^2}\right) = \sum_{i=1}^{N} S(\mathrm{X}, \mathrm{X}_i) \quad (3)$$

This equation quantifies the contribution of the segments X_i to X's area. For example, when X and X_i are located in the same region with same color features, $d_g(\mathrm{X}, \mathrm{X}_i) = 0$ and $S(\mathrm{X}, \mathrm{X}_i) = 1$, ensuring that X_i adds a unit area to the area of X. However, when X_i and X are in different regions with different color features, there is always at least one distinct node with $d_e(\mathrm{X}, \mathrm{X}_i) \gg 3\sigma_{clr}$ on their shortest path and $S(\mathrm{X}, \mathrm{X}_i) \approx 0$, ensuring that X_i does not contribute to X's area [39]. σ_{clr} is a Gaussian kernel parameter. Similarly, the boundary connectivity of segments located at the boundaries from area of X can be computed as:

$$\mathrm{LenBndS}(\mathrm{X}) = \sum_{i=1}^{N} S(\mathrm{X}, \mathrm{X}_i) * \delta(\mathrm{X}_i \in BndS), \quad (4)$$

where δ is 1 for segments on the image boundary and 0 otherwise. Final boundary connectivity of each segment is calculated as (below equation is the actual calculation of the concept introduced in Eq. 1):

$$BC(\mathrm{X}) = \frac{LenBndS(\mathrm{X})}{\sqrt{Area(\mathrm{X})}} \quad (5)$$

To make image segment resolution-invariant, the square root of the area was used for the computation.

E. Background Template Creation

Optimal segments representing background regions (i.e. normal skin area) were selected to create a background template using the boundary connectivity and lightness *L** from CIELAB. We used the *K*-means algorithm to group segments into several clusters (K_1). Segments belonging to a cluster whose boundary connectivity and lightness were the largest amongst other clusters, were then considered as the optimal segments. We computed the average value of the boundary connectivity and lightness of segments in each cluster and sorted the clusters according to these values. We defined the largest cluster as the one with the highest average of boundary connectivity and lightness. Optimal segments from the cluster were extracted to construct the background template $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_M]$, where M is the number of segments that belong to the cluster. This template was used as the basis (i.e. dictionary) for the sparse representation model in the following step. Several background templates were then extracted at various scales (different sizes of segments) to understand and learn the structural information represented by each scale.

An entire image can be then represented as $\mathbf{X} = [\mathbf{B}\ \mathbf{L}]$, where $\mathbf{L} = [L_1, L_2, \ldots, L_{N-M}]$ represents lesion segments. A regional mean feature consisting of CIELAB colour spaces, RGB colour values, and spatial location defined by [*L, a, b, R, G, B, x, y*], were used to represent each of all segments. This regional mean feature has proven value in saliency detection and dermoscopic image segmentation [13, 43]. *Lab* and *RGB* colour features also generate more accurate saliency maps [43]. The lightness component *L** provides an advantage for skin lesion segmentation (e.g. detecting background segments) as the perceptual difference between colours and contrast is often affected by lightness variations [13].

F. Saliency Measure via Sparse Reconstruction Error

The sparse representation model is used widely in computer vision and has proven effectiveness in representing, acquiring, and compressing high-dimensional signals [44]. This is attributed to the characteristics of image signals, which have generally sparse representation in regards to fixed bases (i.e. Fourier). Hence, a number of researchers have successfully applied the sparse representation model to a number of applications including medical image segmentation [45, 46], face recognition [47], image super-resolution [48] and image classification [49, 50].

For our application, 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 likely to be classified as a lesion. The sparse reconstruction error was computed as the residual-based on the sparse representation of the background templates **B**. The segment $\mathbf{X}_i$ is encoded as follows:

$$\widehat{\alpha_i} = \underset{\alpha_i}{\operatorname{argmin}} \|\mathbf{X}_i - \mathbf{B}\alpha_i\|_2^2 + \lambda \|\alpha_i\|_1 , \quad (6)$$

where $\widehat{\alpha_i}$ is the sparse representation coefficient. This problem is equivalent to a Lasso linear regression problem. Using Lasso, the sparse reconstruction error = ε_i is then computed as:

$$\varepsilon_i = \|\mathbf{X}_i - \mathbf{B}\alpha_i\|_2^2 , \quad (7)$$

where λ is a regularization parameter which controls sparsity of α. For example, the average sparse reconstruction errors of segments representing skin lesions and background regions in Fig 2 are shown in Table I. It shows that the segments representing lesions consistently have larger sparse reconstruction errors compared with those in background regions.

G. Context-based Error Propagation

We propose a context-based error propagation method to smooth the sparse reconstruction errors by considering other nearby segments. This allows each segment within a lesion to be evenly emphasized (See Fig 2). We used the *K*-means algorithm to cluster all the segments from the image into the *K* clusters (K_2). The new reconstruction error of $\tilde{\varepsilon}_i$ was estimated by proportionally combining the weighted averaging error of all other segments in the same cluster, and the initial sparse

TABLE I. The average sparse reconstruction errors of segments for lesion and background regions.

Mean/Total Number of Segments	***50***	***100***	***150***
Background Regions	$1.84 * 10^{-16}$	$2.84 * 10^{-16}$	$2.04 * 10^{-16}$
Skin Lesion	**0.6195**	**0.5708**	**0.4645**

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2017.2653179, IEEE Journal of Biomedical and Health Informatics

reconstruction error of ε_i at segment i. This can be defined as:

$$\tilde{\varepsilon}_i = \tau \sum_{j=1}^{N_c} w_{ik_j} \tilde{\varepsilon}_{k_j} + (1-\tau)\varepsilon_i \,, \tag{8}$$

where $[k_1, k_2, k_3, \cdots, k_{N_c}]$ indicate the N_c segment labels in cluster k and τ is a weight parameter. The weight of each segment i was then estimated by its normalized similarity according to:

$$w_{ik_j} = \frac{\exp\left(-\frac{\left\|\mathbf{x}_i - \mathbf{x}_{k_j}\right\|^2}{2\sigma_{\mathbf{X}}{}^2}\right)\left(1-\delta(k_j - i)\right)}{\sum_{j=1}^{N_c} \exp\left(-\frac{\left\|\mathbf{x}_i - \mathbf{x}_{k_j}\right\|^2}{2\sigma_{\mathbf{X}}{}^2}\right)} \,, \tag{9}$$

where the $\sigma_{\mathbf{X}}{}^2$ denotes the sum of the variance in each of X and $\delta(k_j - i)$ is the indicator function.

H. Pixel-Level Sparse Reconstruction Error Creation

Instead of using the segment as a whole, pixel level sparse reconstruction errors are calculated by integrating results from multi-scale (i.e. N_s different sizes of segments) sparse reconstruction errors. This increases the detail of the boundary since the segments may not be optimally 'cut' for the boundaries. The pixel-level sparse reconstruction error of $E(z)$ is computed as:

$$E(z) = \frac{\sum_{s=1}^{N_S} \omega_{zn(s)} \tilde{\varepsilon}_{n(s)}}{\sum_{s=1}^{N_S} \omega_{zn(s)}} \,, \tag{10}$$

$$\omega_{zn^{(s)}} = \frac{1}{\left\|f_z - \mathbf{x}_{n(s)}\right\|_2} \,, \tag{11}$$

where, $n^{(s)}$ indicates the label of the segment containing pixel z scale s and f_z denotes the features of pixel z and we used the same feature space as mentioned in Section II.E.

I. Saliency Map Creation via Bayesian Framework

We used a Bayesian model to create the final saliency map. The Bayes formula is used to measure the saliency value of each pixel by the posterior probability [51, 52]:

$$p(L|f(z)) = \frac{p(L)p(f(z)|L)}{p(L)p(f(z)|L) + \left(1-p(L)\right)p(f(z)|B)} \,, \tag{12}$$

where the $p(L|f(z))$ is a probability of predicting a pixel being a lesion and the $p(L)$ and $p(B)$ are the prior probabilities of a pixel z being a lesion and the background respectively. We computed prior probabilities of $p(L)$ and $p(B)$ by binarising an initial saliency map created from pixel-level sparse reconstruction error using the adaptive thresholding method proposed by Huang *et al.* [53]. This method can help effectively locate the deep valley of the gray-scale (i.e. saliency map) histogram, which performs well in maintaining the overall shape of the lesion area [53]. The observation likelihood probabilities can be then computed as follows:

$$p(f(z)|L) = \prod_{f(z)} \frac{N_L(f_z)}{N_L} \,, \tag{13}$$

$$p(f(z)|B) = \prod_{f(z)} \frac{N_B(f_z)}{N_B} \,, \tag{14}$$

where each pixel z is represented by colour histogram using $f(z) = [L, a, b]$, $N_L(f_z)$ indicates the count that lesion region N_L contains $f(z)$, and N_L is the total number of pixels in the lesion region. Similarly, we compute colour distribution histogram for background region N_B . Therefore, the reconstruction error is refined by its similarities to pixels of lesion region L as well as its difference from those from background B using the colour distribution histogram. The final saliency map is converted into a binary segmentation result via a thresholding method proposed by Li *et al.* [54]. We followed common post-processing refinement. We used a morphological dilation process with a disc radius of 5 pixels, to fill small holes

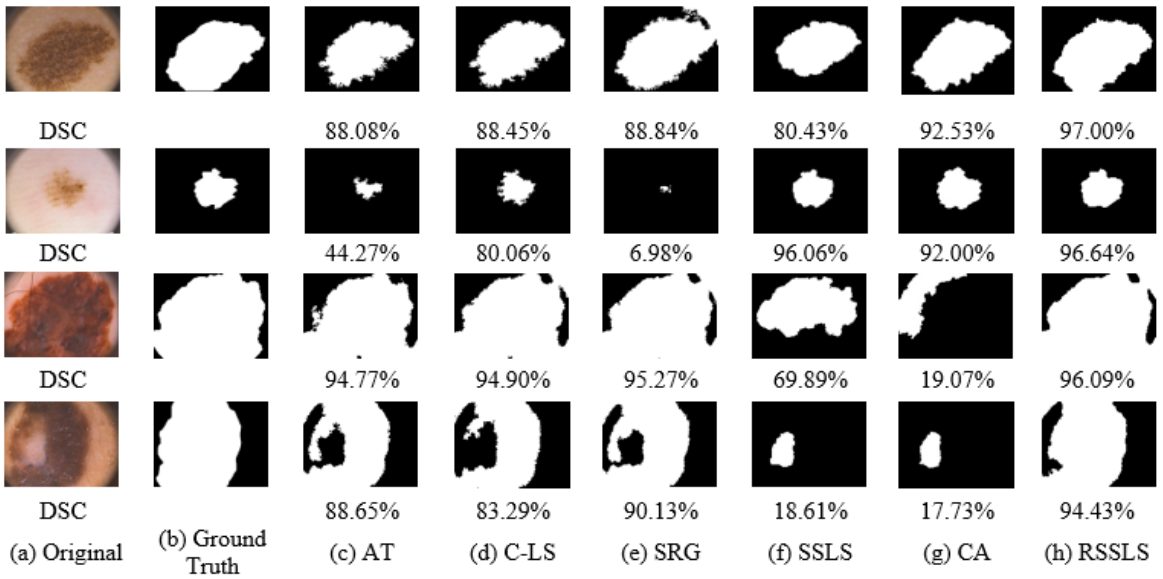


Fig 3. Segmentation results from four study examples (top two – general cases and bottom two – extreme cases), where (a) – (f) represent the original image in column 1, ground truth in column 2 and the segmentation results from column 3 to column 7 for AT, C-LS, SRG, SSLS, CA and RSSLS methods.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2017.2653179, IEEE Journal of Biomedical and Health Informatics

and remove small isolated single pixels.

III. Experimental Setup and Results

A. Materials

We used 2 large public image datasets, the PH2 [55] and International Skin Imaging Collaboration (ISIC) public dataset [56], for the comparative evaluation. The PH2 public dataset provides 200 dermoscopic image studies, including 80 common nevi, 80 atypical nevi and 40 melanomas. The PH2 dataset was divided into two groups based on their image layout. There were 160 general cases: 152 images diagnosed as benign and 8 images as melanoma that capture the complete region of skin lesions. There were also 40 extreme cases, as defined in Section II.C. We selected the same 160 general cases used in the recent study of Ma and Tavaras [13].

The ISIC data archive contains over 10,000 dermoscopic images retrieved from several international clinical centres, acquired from various imaging devices [57]. A new challenge was hosted by ISIC in International Symposium on Biomedical Imaging (ISBI) 2016. For this study we used 900 sample dermoscopic image studies used in the challenge [57]. There were 726 images diagnosed as benign and 174 as melanomas. ISIC dataset contains images with various artifacts such as colour calibration chart and ruler.

Compared with another common dataset [2], both of these datasets contain a greater number of images with complex backgrounds and complicated skin conditions [13]. Manually annotated lesions from expert dermatologists were available from both datasets and were used as the ground truth for the segmentation results. There were 1100 dermoscopic images that we used for quantitative evaluation.

B. Experiments Setup and Implementation Details

We compared our RSSLS with six 'state-of-the-art' unsupervised lesion segmentation methods including: C-LS [6], SRG [25], NDM [13] (semi-automated), AT [6], SSLS [25], and CA [24] (fully-automated). We applied hair removal, as a pre-processing step for PH2 and ISIC datasets. We used the following settings for all the segmentation methods. For RSSLS, there were three main parameters: the number of clusters K_1, K_2 and the weight factor τ in Section II.E, Eq. 8, and Eq. 9. We empirically derived K_1 =4, K_2 = 8, and τ = 0.5. The *K*-mean algorithm was initialized using random seeds and the best performance was achieved using K_1= 4 (K_1 = [2-10]). Variations to other two main parameters, (K_2 = [2-10], and τ = [0.1-0.9]) on the other hand, were insensitive to the final results. The parameters σ_{clr} in Eq. 3 and λ in Eq. 6 were set to 10 and 0.01, consistent with other research [25, 39]. The superpixels at N_s = 8 different scales were created ranging from 50 to 400 segments. For C-LS and SRG, an initial starting seed point was necessary: C-LS was initialized with a number of small circular masks across the whole image which generated the most accurate results [58]. The maximum number of iterations for C-LS was empirically set at 500, which ensured the convergence. For SRG, the seed point was manually set to be the centre of the skin lesion. For AT and C-LS, we initially used the same parameters for each method as used by Silveira et al [6] and empirically fine-tuned the parameters to obtain the best performance. For NDM, we used the results reported in the paper [13]. The CA method is one of our previous segmentation approach; we also tested it with the extreme cases from the PH2 and ISIC datasets. We replaced the colour calibration charts in ISIC dataset with the neighbourhood pixels as a pre-processing step for all of the methods. We applied basic post-processing such as noise removal (noise with isolated objects or pixels) and region filling (dark pixels completely-surrounded by lighter pixels) to all methods.

We also compared our RSSLS to seven unsupervised state-of-the-art saliency detection methods including GC[59], RC[35], HSD[36], (foreground-based), MC[40], MR[37], RBD[39], and RRWR[38] (background-based). All of these approaches were implemented using the source code provided from the authors' websites. The thresholding method proposed by Li *et al.* [54] was used to convert the gray-scale saliency map to a final binary image for all the methods.

We conducted 4 different comparative experiments. At baseline: 1) The 160 general cases from the PH2 dataset were used. 2) The 40 extreme cases from PH2 were used to evaluate the effectiveness of our background detection. 3) We evaluated our RSSLS on the ISIC dataset (900 images). 4) We compared our RSSLS to other state-of-the-art unsupervised saliency detection methods. We implemented our algorithm in Matlab on a standard PC (Intel i5 CPU 2.4Ghz RAM 8GB). A Matlab implementation takes, on average, 28.7 seconds to process each image. Most (26.2 seconds) of the computation time was taken to generate superpixels at eight different scales.

C. Evaluation Metrics

We used the 3 standard metrics that have been reported by many other researchers [13, 24, 63] to assess the performance of the segmentation results. We defined these metrics as follows:

The Dice similarity coefficient (DSC) [60]:

$$DSC(A,B) = \frac{2N(A \cap B)}{N(A)+N(B)}, \tag{15}$$

where *A* denotes the segmentation result and *B* is the ground truth; *N* represents the number of pixels in the corresponding set. A higher DSC score corresponds to a better result.

The Hammoude distance (HM) [61] and XOR [62] measure the dissimilarity between the result of segmentation and the ground truth, defined as:

$$HM(A,B) = \frac{N(A \cup B) - N(A \cap B)}{N(A \cup B)}, \tag{16}$$

$$XOR(A,B) = \frac{N(A \cup B) - N(A \cap B)}{N(B)}. \tag{17}$$

Lower scores for HM and XOR correspond to better results. We also evaluated the robustness of our algorithm by measuring its ability to correctly detect the lesion among all the images in the datasets.

D. Results

The results for the 160 general cases from PH2 dataset for all methods are shown in Table II. We show the results for four selected images of varing image conditions in Figure 3 with general cases in the 1st and 2nd rows and the extreme cases in the 3rd and 4th rows. Our RSSLS generated higher scores, and

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2017.2653179, IEEE Journal of Biomedical and Health Informatics

had the best overall DSC average (93.12%) and the best HM (12.52%) accuracy.

The results from the extreme 40 cases are shown in Table III and they show that our approach had the lowest XOR value (28.21%) and the best DSC (82.75%) and HM (27.34%) scores. We show a summary of the results from all methods over the 1100 imges in Table IV. For the PH2 dataset our approach had the highest DSC score with an overall average of 91.05%; it consistently had the lowest values for HM (15.49%) and XOR (16.45%). For the ISIC dataset our scores were DSC of 83.41% with the lowest HM of 25.67% and XOR scores of 36.21%.

IV. Discussion

A. Evaluation with the PH2 dataset

Our findings indicate that our RSSLS approach improves segmentation accuracy, is robust, and can accurately locate the lesion and its proximity to image boundaries (See Fig 3). Over the 160 general cases the next best method was our own CA-based method. The CA method, however, is expensive computationally because of the time needed to learn the model for seed selection. The high accuracy for all the methods can be attributed to several images having relative simple lesions with high contrast to the background regions (see Fig 3, row 1). The C-LS and NDM methods were the next closest to ours in overall DSC and XOR indices. C-LS was not able, however, to segment lesions when the lesion borders and the surrounding regions were indistinct (Fig 3(d), row 2). NDM also required manual initialization.

Upon review of the recent literature, recent segmentation approaches [13, 24] only used general cases from the PH2 dataset and they excluded the 40 extreme cases. Our approach was the best technique for the extreme cases with a 20% improvement when compared with our previous approach, which further supports this contribution to the body of research. Our new approach was able to identify the overall shape and edge of the lesions, which we attribute to the Bayesian framework. Over-segmentation of lesions was often evident with AT and SRG. Further, AT and C-LS were unable to identify lesion edges across all the images. AT, with a DSC of 80.1% and a HM of 31.9%, had the closest results to ours, which we suggest is because AT is prone to over-segmentation and it benefited from 7 images in the PH2 dataset where the entire image was a lesion.

TABLE II. The segmentation results measured by DSC, HM, and XOR– 160 general cases from PH2 dataset.

Mean%	*AT*	*C-LS*	*SRG*	*SSLS*	*NDM*	*CA*	*RSSLS*
DSC	82.36	87.02	61.24	91.33	N/A	92.49	**93.12**
HM	27.78	21.65	50.28	15.28	N/A	13.71	**12.52**
XOR	41.09	26.76	51.57	17.48	13.92	14.99	**13.52**

TABLE III. The segmentation results measured from the extreme 40 cases from PH2 dataset by DSC, HM, and XOR

Mean%	*AT*	*C-LS*	*SRG*	*SSLS*	*CA*	*RSSLS*
DSC	80.13	76.05	64.10	61.39	74.46	**82.75**
HM	31.91	36.32	46.77	53.49	36.53	**27.34**
XOR	33.09	37.62	47.51	56.40	39.31	**28.21**

TABLE IV. The segmentation results measured by DSC, HM, and XOR– total 200 images and the 900 images from PH2 dataset and ISIC dataset respectivley.

Mean%	*AT*	*C-LS*	*SRG*	*SSLS*	*CA*	*RSSLS*
DSC[1]	81.91	84.82	61.23	85.34	88.89	**91.05**
DSC[2]	70.90	79.61	77.55	80.61	76.30	**83.41**
HM[1]	28.61	24.58	49.30	22.92	19.86	**15.49**
HM[2]	41.12	30.00	32.95	29.78	45.97	**25.67**
XOR[1]	39.49	28.93	49.98	25.27	18.27	**16.45**
XOR[2]	63.44	50.58	45.60	38.40	35.36	**36.21**

[1]Total 200 images from the PH2 dataset.
[2]The 900 images from the ISIC dataset.

B. Evaluation with the ISIC dataset

The ISIC dataset images were obtained from different imaging devices and so, in comparison to the PH2 dataset, they had more complex variations in lesion location, lighting condition and were subject to non-uniform vignetting [27]. As a consequence, the overall performance of all the algorithms was reduced when compared to the performance on the PH2 dataset. Nevertheless, our approach still performed well with a DSC of 83.41% and the lowest HM of 25.67% and XOR of 36.21% (see Table IV).

C. Evaluation of the different saliency detection models

Table V, VI, and VII show the segmentation results for all the methods with both datasets. Our approach consistently outperformed the other approaches. Although the other methods were not designed for lesion segmentation of dermoscopic images, our findings indicate the value of saliency-based models in this setting.

D. Extending saliency based algorithms for lesion segmentation

Our results suggest that other saliency-based methods would benefit from our applications of a Bayesian framework. We suggest that our framework could serve as a saliency optimisation algorithm. Hence we applied our framework to three comparable and high-performing approaches – the

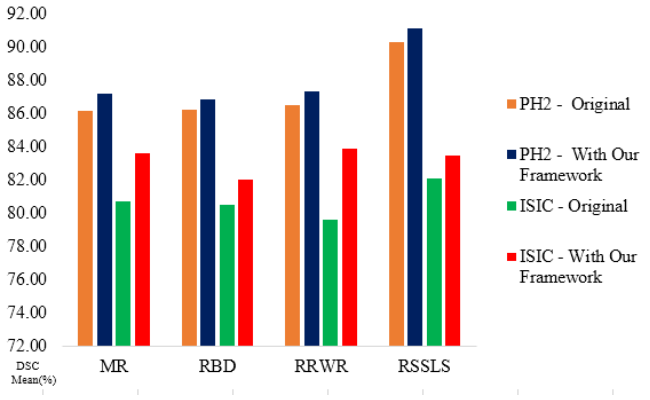


Fig 4. Average DSC scores of original MR, RBD, RRWR, and RSSLS methods and improved DSC scores with our Bayesian framework using PH2 dataset and ISIC dataset.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2017.2653179, IEEE Journal of Biomedical and Health Informatics

TABLE V. The segmentation results measured by DSC, HM, and XOR for saliency based methods using general cases from PH2 (FG = Foreground, BG = Background).

Mean	***HSD***	***RC***	***GC***	***MC***	***MR***	***RBD***	***RRWR***	***RSSLS***
Type	FG	FG	FG	BG	BG	BG	BG	BG
DSC	69.43	81.86	75.92	89.18	91.18	91.63	91.28	**93.12**
HM	41.78	25.05	33.39	17.95	15.68	15.02	15.52	**12.52**
XOR	71.14	29.75	132.3	32.92	19.02	15.84	18.60	**13.52**

TABLE VI. The segmentation results measured by DSC, HM, and XOR for saliency based methods using extreme cases from PH2

Mean	***HSD***	***RC***	***GC***	***MC***	***MR***	***RBD***	***RRWR***	***RSSLS***
Type	FG	FG	FG	BG	BG	BG	BG	BG
DSC	60.76	36.76	77.52	76.32	74.84	64.45	75.00	**82.75**
HM	51.42	74.63	33.87	36.63	37.81	49.33	37.67	**27.34**
XOR	57.06	85.90	35.07	38.52	39.83	50.91	39.56	**28.21**

TABLE VII. The segmentation results measured by DSC, HM, and XOR for saleincy based methods using both PH2 and ISIC dataset.

Mean	***HSD***	***RC***	***GC***	***MC***	***MR***	***RBD***	***RRWR***	***RSSLS***
Type	FG	FG	FG	BG	BG	BG	BG	BG
DSC[1]	65.81	72.84	76.24	86.61	86.11	86.19	86.46	**91.05**
DSC[2]	52.48	69.72	63.15	77.33	80.71	80.45	79.58	**83.41**
HM[1]	46.41	40.98	33.49	34.04	22.56	21.88	21.75	**15.49**
HM[2]	60.42	39.28	46.82	127.0	29.54	29.97	30.22	**25.67**
XOR[1]	71.68	34.97	112.8	21.68	28.80	22.85	32.41	**16.45**
XOR[2]	133.3	337.6	367.1	32.26	69.34	34.12	98.09	**36.21**

[1]Images from the PH2 dataset.
[2]Images from the ISIC dataset.

RRWR, RBD, and MR. The results, which are shown in Table VII, indicate that the segmentation performances of three methods were notably improved as shown in Fig 4. MR, with a DSC score of 83.59%, and RRWR with a DSC of 83.83%, both had a > 3% improvement in segmentation performance using ISIC dataset. We attribute this to the Bayesian framework allowing better detection of the shape and boundaries of a lesion. We note that MR, RBD, and RRWR had a greater improvement when using the more challenging ISIC dataset.

E. Limitation

Our approach detects skin lesion accurately, however, it has some limitations, as we show in images where the skin lesion is very small, not visually distinctive and located at the image boundary. In these instances our background estimation can mistakenly detect skin lesion as part of normal background regions. Nevertheless, such cases are uncommon and our approach prevails over other state-of-the-art methods.

V. Conclusions

In this work we have implemented a major advance on our previous work using a saliency-based segmentation framework for the identification and characterisation of skin lesions in dermoscopic images. We compared our approach to other methods on 2 large public datasets and show that our approach outperformed the other methods and, further, our framework can be used as a saliency optimisation algorithm for lesion segmentation in dermoscopic images.

References

[1] D. S. Rigel, *et al.*, "The incidence of malignant melanoma in the United States: issues as we approach the 21st century," *Journal of the American Academy of Dermatology*, vol. 34, pp. 839-847, 1996.

[2] G. Argenziano, *et al.*, "Dermoscopy, An Interactive Atlas," *EDRA Medical Publishing*, 2000.

[3] M. Binder, *et al.*, "Epiluminescence microscopy: A useful tool for the diagnosis of pigmented skin lesions for formally trained dermatologists," *Archives of Dermatology*, vol. 131, pp. 286-291, 1995.

[4] P. Wighton, *et al.*, "A fully automatic random walker segmentation for skin lesions in a supervised setting," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2009*, ed: Springer, 2009, pp. 1108-1115.

[5] J. C. Nascimento and J. S. Marques, "Improved Gradient Vector Flow for robust shape estimation in medical imaging," in *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, 2010, pp. 4809-4812.

[6] M. Silveira, *et al.*, "Comparison of Segmentation Methods for Melanoma Diagnosis in Dermoscopy Images," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 3, pp. 35-45, 2009.

[7] D. D. Gómez, *et al.*, "Independent histogram pursuit for segmentation of skin lesions," *Biomedical Engineering, IEEE Transactions on*, vol. 55, pp. 157-161, 2008.

[8] F. Peruch, *et al.*, "Simpler, faster, more accurate melanocytic lesion segmentation through meds," *Biomedical Engineering, IEEE Transactions on*, vol. 61, pp. 557-565, 2014.

[9] M. Emre Celebi, *et al.*, "Lesion border detection in dermoscopy images using ensembles of thresholding methods," *Skin Research and Technology*, vol. 19, pp. e252-e258, 2013.

[10] K. A. Norton, *et al.*, "Three-phase general border detection method for dermoscopy images using non-uniform illumination correction," *Skin Research and Technology*, vol. 18, pp. 290-300, 2012.

[11] R. Garnavi, *et al.*, "Border detection in dermoscopy images using hybrid thresholding on optimized color channels," *Computerized Medical Imaging and Graphics*, vol. 35, pp. 105-115, 2011.

[12] H. Iyatomi, *et al.*, "An improved internet-based melanoma screening system with dermatologist-like tumor area extraction algorithm," *Computerized Medical Imaging and Graphics*, vol. 32, pp. 566-579, 2008.

[13] Z. Ma and J. M. R. S. Tavares, "A Novel Approach to Segment Skin Lesions in Dermoscopic Images Based on a Deformable Model," *Biomedical and Health Informatics, IEEE Journal of*, vol. 20, pp. 615-623, 2016.

[14] H. Zhou, *et al.*, "Mean shift based gradient vector flow for image segmentation," *Computer Vision and Image Understanding*, vol. 117, pp. 1004-1016, 2013.

[15] H. Zhou, *et al.*, "Gradient vector flow with mean shift for skin lesion segmentation," *Computerized Medical Imaging and Graphics*, vol. 35, pp. 121-127, 2011.

[16] B. Erkol, *et al.*, "Automatic lesion boundary detection in dermoscopy images using gradient vector flow snakes," *Skin Research and Technology*, vol. 11, pp. 17-26, 2005.

[17] Q. Abbas, *et al.*, "Unsupervised skin lesions border detection via two-dimensional image analysis," *Computer Methods and Programs in Biomedicine*, vol. 104, pp. e1-e15, 12// 2011.

[18] M. E. Celebi, *et al.*, "Lesion border detection in dermoscopy images," *Computerized medical imaging and graphics*, vol. 33, pp. 148-153, 2009.

[19] M. Emre Celebi, *et al.*, "Unsupervised border detection in dermoscopy images," *Skin Research and Technology*, vol. 13, pp. 454-462, 2007.

[20] M. Emre Celebi, *et al.*, "Border detection in dermoscopy images using statistical region merging," *Skin Research and Technology*, vol. 14, pp. 347-353, 2008.

[21] R. Nock and F. Nielsen, "Statistical region merging," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 26, pp. 1452-1458, 2004.

[22] T. F. Chan, *et al.*, "Active contours without edges for vector-valued images," *Journal of Visual Communication and Image Representation*, vol. 11, pp. 130-141, 2000.

[23] A. R. Sadri, *et al.*, "Segmentation of Dermoscopy images using wavelet networks," *Biomedical Engineering, IEEE Transactions on*, vol. 60, pp. 1134-1141, 2013.

[24] L. Bi, *et al.*, "Automated Skin Lesion Segmentation via Image-wise Supervised Learning and Multi-Scale Superpixel Based Cellular

Automata," in *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on*, Prague, Czech Republic, 2016, pp. 1059-1062.
[25] E. Ahn, *et al.*, "Automated Saliency-based Lesion Segmentation in Dermoscopic Images," in *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, 2015, pp. 3009-3012.
[26] M. E. Celebi, *et al.*, *A State-of-the-Art Survey on Lesion Border Detection in Dermoscopy Images*. Dermoscopy Image Analysis: CRC Press, 2015.
[27] N. Mishra and M. E. Celebi, "An Overview of Melanoma Detection in Dermoscopy Images Using Image Processing and Machine Learning," in *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium on*, 2016.
[28] J. Long, *et al.*, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 3431-3440.
[29] H.-C. Shin, *et al.*, "Stacked autoencoders for unsupervised feature learning and multiple organ detection in a pilot study using 4D patient data," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 35, pp. 1930-1943, 2013.
[30] A. Borji and L. Itti, "State-of-the-art in visual attention modeling," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 35, pp. 185-207, 2013.
[31] L. Itti, *et al.*, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 20, pp. 1254-1259, 1998.
[32] R. Achanta, *et al.*, "Salient region detection and segmentation," in *Computer Vision Systems*, ed: Springer, 2008, pp. 66-75.
[33] J. M. Wolfe and T. S. Horowitz, "What attributes guide the deployment of visual attention and how do they do it?," *Nat Rev Neurosci*, vol. 5, pp. 495-501, 06//print 2004.
[34] M. M. Cheng, *et al.*, "Global contrast based salient region detection," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, 2011, pp. 409-416.
[35] M. M. Cheng, *et al.*, "Global Contrast Based Salient Region Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, pp. 569-582, 2015.
[36] Y. Qiong, *et al.*, "Hierarchical Saliency Detection," in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, 2013, pp. 1155-1162.
[37] Y. Chuan, *et al.*, "Saliency Detection via Graph-Based Manifold Ranking," in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, 2013, pp. 3166-3173.
[38] C. Li, *et al.*, "Robust Saliency Detection via Regularized Random Walks Ranking," in *Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on*, 2015.
[39] W. Zhu, *et al.*, "Saliency optimization from robust background detection," in *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, 2014, pp. 2814-2821.
[40] B. Jiang, *et al.*, "Saliency detection via absorbing markov chain," in *Proceedings of the IEEE International Conference on Computer Vision*, 2013, pp. 1665-1672.
[41] R. Achanta, *et al.*, "SLIC superpixels compared to state-of-the-art superpixel methods," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 34, pp. 2274-2282, 2012.
[42] T. Lee, *et al.*, "Dullrazor< sup>®</sup>: A software approach to hair removal from images," *Computers in Biology and Medicine*, vol. 27, pp. 533-543, 1997.
[43] A. Borji and L. Itti, "Exploiting local and global patch rarities for saliency detection," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, 2012, pp. 478-485.
[44] J. Wright, *et al.*, "Sparse Representation for Computer Vision and Pattern Recognition," *Proceedings of the IEEE*, vol. 98, pp. 1031-1044, 2010.
[45] S. Zhang, *et al.*, "Deformable segmentation via sparse representation and dictionary learning," *Medical Image Analysis*, vol. 16, pp. 1385-1396, 2012.
[46] S. Zhang, *et al.*, "Towards robust and effective shape modeling: Sparse shape composition," *Medical image analysis*, vol. 16, pp. 265-277, 2012.
[47] J. Wright, *et al.*, "Robust face recognition via sparse representation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 31, pp. 210-227, 2009.
[48] J. Yang, *et al.*, "Image super-resolution via sparse representation," *Image Processing, IEEE Transactions on*, vol. 19, pp. 2861-2873, 2010.
[49] Z. Jiang, *et al.*, "Learning a discriminative dictionary for sparse coding via label consistent K-SVD," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, 2011, pp. 1697-1704.
[50] J. Mairal, *et al.*, "Supervised dictionary learning," in *Advances in neural information processing systems*, 2009, pp. 1033-1040.
[51] Y. Qin, *et al.*, "Saliency detection via cellular automata," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 110-119.
[52] Y. Xie, *et al.*, "Bayesian saliency via low and mid level cues," *Image Processing, IEEE Transactions on*, vol. 22, pp. 1689-1698, 2013.
[53] L.-K. Huang and M.-J. J. Wang, "Image thresholding by minimizing the measures of fuzziness," *Pattern recognition*, vol. 28, pp. 41-51, 1995.
[54] C. Li and P. K.-S. Tam, "An iterative algorithm for minimum cross entropy thresholding," *Pattern Recognition Letters*, vol. 19, pp. 771-776, 1998.
[55] T. Mendonça, *et al.*, "PH 2-A dermoscopic image database for research and benchmarking," in *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*, 2013, pp. 5437-5440.
[56] ISIC. (2016). *ISIC Archieve: The International Skin Imaging Collaboration: Melanoma Project*. Available: https://isic-archive.com/
[57] D. Gutman, *et al.*, "Skin Lesion Analysis toward Melanoma Detection: A Challenge at the International Symposium on Biomedical Imaging (ISBI) 2016, hosted by the International Skin Imaging Collaboration (ISIC)," *arXiv preprint arXiv:1605.01397*, 2016.
[58] L. A. Vese and T. F. Chan, "A multiphase level set framework for image segmentation using the Mumford and Shah model," *International journal of computer vision*, vol. 50, pp. 271-293, 2002.
[59] K. Sohn, *et al.*, "Efficient learning of sparse, distributed, convolutional feature representations for object recognition," in *2011 International Conference on Computer Vision*, 2011, pp. 2643-2650.
[60] W. R. Crum, *et al.*, "Generalized overlap measures for evaluation and validation in medical image analysis," *Medical Imaging, IEEE Transactions on*, vol. 25, pp. 1451-1461, 2006.
[61] A. Hammoude, "Computer-assisted endocardial border identification from a sequence of two-dimensional echocardiographic images," 1988.
[62] F. Xie and A. C. Bovik, "Automatic segmentation of dermoscopy images using self-generating neural networks seeded by genetic algorithm," *Pattern Recognition*, vol. 46, pp. 1012-1019, 2013.