IEEE Transactions on Medical Imaging, Vol. 36, Issue 1, January 2017, pp. 51-63
DOI:10.1109/TMI.2016.2593725

Intensity and Compactness Enabled Saliency Estimation for Leakage Detection in Diabetic and Malarial Retinopathy

Yitian Zhao, Yalin Zheng, Yonghuai Liu, Jian Yang*, Yifan Zhao, Duanduan Chen and Yongtian Wang

***Abstract*—Leakage in retinal angiography currently is a key feature for confirming the activities of lesions in the management of a wide range of retinal diseases, such as diabetic maculopathy and paediatric malarial retinopathy. This paper proposes a new saliency-based method for the detection of leakage in fluorescein angiography. A superpixel approach is firstly employed to divide the image into meaningful patches (or superpixels) at different levels. Two saliency cues, intensity and compactness, are then proposed for the estimation of the saliency map of each individual superpixel at each level. The saliency maps at different levels over the same cues are fused using an averaging operator. The two saliency maps over different cues are fused using a pixel-wise multiplication operator. Leaking regions are finally detected by thresholding the saliency map followed by a graph-cut segmentation. The proposed method has been validated using the only two publicly available datasets: one for malarial retinopathy and the other for diabetic retinopathy. The experimental results show that it outperforms one of the latest competitors and performs as well as a human expert for leakage detection and outperforms several state-of-the-art methods for saliency detection.**

***Index Terms*—leakage, diabetic, malarial, retinopathy, fluorescein angiogram, saliency, segmentation.**

I. Introduction

Fundus fluorescein angiography (FA) is a valuable imaging modality that provides a map of retinal vascular structure and function by highlighting blockage of, and leakage from, retinal vessels [1]. Although FA is invasive and expensive, and exposes patients with rare but potentially serious side effects, it is indispensable in differential diagnosis of retinal diseases such as diabetic retinopathy (DR), age-related macular degeneration (AMD), malarial retinopathy (MR), and so on [2]–[4]. Incarnated as useful signal of high intensity, retinal leakage in angiography is currently a key feature for clinicians to determine the activities and development of lesions in the retina. Fig. 1 shows the appearance of leakages in MR and DR respectively. MR is believed to be important for the differential diagnosis of cerebral malaria, while DR is a leading cause of vision loss in the working age population. Identification of sites and evaluation of the extent of leakage enable decision-making for treatment and monitoring of disease activities. More specifically, the detection of retinal lesions in general is important for automated diagnosis of retinal disease while the leakage detection is important for therapy planning and treatment outcome monitoring.

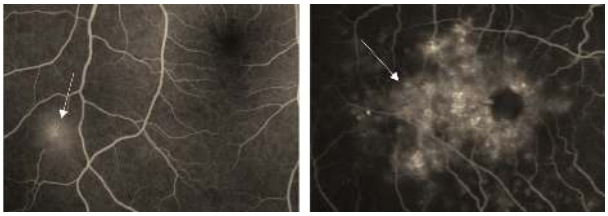


Fig. 1: Illustration of focal leakages in two retinal diseases. Left: malarial retinopathy. Right: diabetic retinopathy. There is a large increase in brightness in leaking regions (white arrows) compared to surrounding non-leaking regions.

Current practical approaches for quantitative analysis of FA features require extensive manual delineation by experienced graders. In eye and vision science research the requirement for such intervention usually introduces human errors, and slows down the process, which makes it impractical to process the vast amount of data collected during routine clinics. There is an increasing demand for the automated detection of the leakage in FA.

In this paper we present a new, unsupervised technique to detect and quantify leakage in FA images with the following contributions.

First, we propose a novel efficient way to enhance leakage regions by using the concept of saliency [5]. Saliency indicates the relative importance of visual features, and is closely related to the characteristics of human perception and processing of visual stimuli [5]–[7]. Saliency emerges from such characteristics in features of the image as visual uniqueness, unpredictability, or rarity, and is often attributed to variations in specific image attributes such as color, gradient, edges, and boundaries [7]–[9]. Such attributes are also characteristics

This work was supported by the National Basic Research Program of China (2013CB328806), Key Projects in the National Science & Technology Pillar Program (2013BAI01B01), National Hi-Tech Research and Development Program (2015AA043203), and the National Science Foundation Program of China (81430039, 61501030).(*Corresponding author: Jian Yang)*

Y. Zhao, J. Yang, and Y. Wang are with Beijing Engineering Research Center of Mixed Reality and Advanced Display, School of Optics and Electronics, Beijing Institute of Technology, Beijing, China. Correspondence email: jyang@bit.edu.cn.

Y. Zheng is with the Department of Eye and Vision Science, University of Liverpool, Liverpool, UK.

Y. Liu is with the Department of Computer Science, Aberystwyth University, Aberystwyth, UK.

Yifan Zhao is with EPSRC Centre for Innovative Manufacturing in Through-life Engineering Services, Cranfield University, UK.

D. Chen is with School of life Science, Beijing Institute of Technology, Beijing, China.

Published by IEEE. This is the Author Accepted Manuscript issued with:
Creative Commons Attribution Non-Commercial License (CC:BY:NC 4.0). Permission from IEEE must be obtained for all other users.
The final published version (version of record) is available online at DOI:10.1109/TMI.2016.2593725. Please refer to any applicable publisher terms of use.

of retinal leakage in FA images. For example, leakage of fluorescent dye causes a large increase in brightness in leaking regions when compared to surrounding non-leaking regions. For this application, leaking regions can be defined as those of high salience. In consequence, we are motivated to firstly identify the leaking regions in FA images through a saliency detection method, and then estimate their areas from the obtained saliency map.

Second, we have proposed a new way to generate multiscale saliency maps with integration of the intensity and compactness cues of superpixels for this specific application. More specifically, traditional saliency extraction methods usually compute the salience of an image in a pixel-by-pixel manner, and ignore the neighborhood and edge information of the objects of interest. Inspired by the fact that human vision is usually more concerned with objects than with individual pixels and the objects of interest may vary in size, in this paper we firstly propose to use patches) at different levels to represent the given images, and the powerful simple linear iterative clustering (SLIC) method [10] is employed for this task.

The reminder of this paper is structured as follows. Section II briefly reviews the related work on leakage detection and saliency detection. Section III details the proposed approach. Section IV describes the datasets and metrics for the evaluation of the proposed technique. In Section V we first described our experiments on different datasets in comparison to those previous proposed methods and report the experimental results. Section VI experimentally investigates the selection of saliency cues and the setting of some hyperparameters used in the proposed method. Section VII concludes the paper.

II. Related work

In this section, the most common leakage and saliency detection methods with application to medical images will be briefly reviewed.

A. Leakage detection

An extensive literature review shows that automated retinal image analysis of FA images, especially for leakage detection, is relatively unexplored. In contrast to the large number of studies on detecting various retinal lesions (i.e. drusen, exudates, hemorrhage, and so on) in colour fundus photograph, relatively few methods have been proposed on automated detection or quantification of leakage. Zhao et al. [11] recently proposed a method to detect three types of leakage (large focal, punctate focal, and vessel segment leakage) on images from eyes with MR. This method can count the number of leakage sites and measures their sizes and has a reasonable performance over only 10 images of MR. However, it only uses the intensity information to generate the saliency map for the detection, which may suffer when some non-leakage areas also have high intensities. Rabbani et al. [12] proposed a method to detect leakage in FA images of subjects with diabetic macular edema. They employed an active contour segmentation model to detect the boundaries of leaking areas. This method is designed to detect areas of leakage in a circular region centered at the fovea with a radius of 1500μm, and has a relatively low sensitivity of 0.69 on 24 images. Martinez-Costa et al. [13] suggested that any pixels with statistically high increments in gray level along the FA sequence close to the foveal centre could be segmented as leakage, and applied this criterion to detect the leakage in the macua due to retinal vein occlusion. However, this method requires manual detection of the foveal center. Phillips et al. [14] calculated the gradient of fluorescence intensity, and then thresholded the gradient values only to determine leakage regions in DR images. This method was applied to only six cases. Saito et al. [15] proposed a detection framework of choroidal neovascularization (CNV) featured by leakage. However, detection of CNV involves analysis of a small area of the retina only instead of the entire image. Trucco et al. [16] and Tsai et al. [17] applied AdaBoost methods to classify the leakage regions of FA images based on multiple handcrafted features. However, these supervised methods are limited by their dependence on training datasets derived from manual annotation. The performance of the classifier will be inherently dependent on the quality of this annotation.

B. Saliency detection as applied to medical images

The application of saliency information for detecting abnormalities from different modalities of medical images is relatively unexploited when compared to other applications. Yuan et al. [18] proposed a saliency-based ulcer detection method from the wireless capsule endoscopy (WCE) images. It uses a multi-level superpixel representation as the preprocessing step for saliency detection, and the saliency map is generated from different levels by integrating all obtained saliency maps according to the color and texture features. This method is capable of accurately representing the contours of the ulcerated regions, and these regions are located through an image feature encoding and recognition method. The limitation of this method is that neither its effectiveness nor its potential is well demonstrated, because the dataset used for validation is too small. Mahapatra and Sun [19] used the saliency and gradient information in a Markov random field for non-rigid registration of dynamic MRI cardiac perfusion images. This approach attempts to address the problem that most nonrigid registration algorithms fail to give satisfactory results in the presence of intensity changes. Although the saliency provides high quality contrast-enhanced images, the gradient information can still be influenced by noise. This method cannot accurately register the boundary of the left ventricle. A visual saliency-based bright lesion detection and classification method was introduced by Deepak et al [20]. The spectral residual saliency model [21] was first employed to compute the saliency map of the color fundus retinal images. The saliency computation leads to a sparse generalized motion patterns representation of the images, and an image is then classified as normal or abnormal (having bright lesions) using the k-nearest neighbour classifier based on the texture. Jampani et al. [22] analyzed the relevance of saliency models in detecting abnormalities in two types of medical images. The experimental results show that the Graph Based Visual

Algorithm 1 Pseudo Code of Saliency and Leakage Detection

Input: An FA image I with focal leakage.

Saliency Detection:

for each level **do**
  Cluster the image to n superpixels;
  for each superpixel **do**
    Compute the intensity-based saliency using Eq. 6, and compactness-based saliency using Eq. 9;
  end for
end for
for each cue **do**
  Fuse the saliency map based on the same cue at different levels using an averaging operator;
end for
Fuse the saliency maps based on different cues using the Hadamard product

Leakage Detection:

Normalize $\mathcal{S}$ to $[0, 1]$, threshold ($T = 0.65$) it to obtain ROIs.
for ROIs **do**
  Graph cut segmentation;
end for
Mask the vessel region from $\mathcal{S}$, and remove optic disc regions and small/isolated objects.

Output: The detected focal leaking areas in the given image.

Saliency method [23] performs best on the chest X-ray images, while the Spectral Residual method [21] performs best over the retinal images. These two methods have been thus selected for further extension for even better performance based on domain knowledge and multiscale analysis respectively.

III. The proposed method

The entire framework for detecting leakages in FA images is summarized in **Algorithm 1**. It includes two main steps: saliency detection and leakagle detection. In the following subsections, each step will be detailed.

A. Saliency Detection

'Salient' regions are those regions of a medical image that contain meaningful information for diagnostic purposes. Typically, the intensities and/or shapes of these regions are significantly different from their surroundings or neighbors [6], [23]–[26]. As shown in Fig. 1, the leaking regions in an FA image are conspicuous objects, and can easily be distinguished *visually* by their intensity or shape. The intensity based approach seems to be a natural choice for *computational* leakage area detection [11]. However, large vessels and the optic disc might also be falsely detected as salient regions for similar reasons in this application. Consequently, the vessel extraction and optic disc detection are essential in this framework: simply masking them will help to improve the accuracy of leakage detection. In this paper, for convenience we define all the aforementioned regions that might be assigned a high saliency value as the regions of interest (ROIs). After the whole process, the false ones such as large vessels and the optic disc will be removed while only the leakage regions will be retained. In the following subsections, the superpixel based saliency detection method will be detailed.

1) Superpixel Segmentation: A region-based approach is well established in saliency measurement: for example, Cheng et al. [8] have used a histogram-based contrast method: the saliency value of each pixel relative to the others in the entire image is estimated and then smoothed in the color space, and further improved through partitioning the given image into regions and assigning saliency values to such regions through considering both their global contrast score and local spatial coherence. This is a two-step method and the first step may assign different saliency values to similar colors due to color quantization. In our method, superpixels are employed to avoid discontinuities at the bin edges of the histogram.

A state-of-the-art superpixel algorithm, called Simple Linear Iterative Clustering (SLIC) [10], is employed in this work to generate a desired number n of regular, compact superpixels to replace the rigid structure of the pixel grid, at a low computation cost, where the default value of 10 for the compactness term is adopted. The SLIC is a k-means clustering method, and is able to assign each pixel to a superpixel according to its intensity and spatial location. The SLIC is capable of grouping meaningful entities into a superpixel by assembling spatially neighboring pixels with similar properties. It not only provides fine segmentation results, but also generates a suitable number of segments for leakage image analysis. Similar research using different method has also been reviewed [27]

In this work, a multi-level superpixel method is proposed. The input image is segmented into L ($L = 3$) levels of superpixels independently, and the corresponding number n of superpixels is set to be 333, 666, and 1000 at each levels, respectively. Fine tuning of the values for these parameters: L and n will be discussed later in Section VI.

2) Intensity-based Saliency Detection: Let $\mathcal{P}_i \in I$ be a viable local representation as a superpixel i ($i = 1, 2, \cdots, n$), and let I indicate the input image. The superpixels may be seen as samples of a multivariate probability density function (PDF) of the imaged objects. A kernel density estimator (KDE) is chosen, as, being non-parametric, it will permit the estimation of any PDF. The probability of a patch $\mathcal{P}_i$ may now be defined as:

$$p(\mathcal{P}_i) = \frac{1}{nh} \sum_{j=1}^{n} K\left(\frac{d(\mathcal{P}_i, \mathcal{P}_j)}{h}\right), \tag{1}$$

where d is a distance function that will be discussed later, K is a kernel, and h is a smoothing parameter. The KDE method has the capacity to average out the contribution of each sample $\mathcal{P}_i$ by spreading it over a certain area [28], which is defined by K. The multivariate distribution will have a higher probability if a superpixel is in dense and similar areas. From our experience, the most commonly used and appropriate kernel is the Gaussian function with zero mean and standard deviation σ. In this case, the probability of a superpixel $p(\mathcal{P}_i)$ can be defined as:

$$p(\mathcal{P}_i) = \frac{1}{n\Gamma} \sum_{j=1}^{n} \exp\Big(-\frac{d^2(\mathcal{P}_i, \mathcal{P}_j)}{2\sigma^2}\Big). \tag{2}$$

The estimated probabilities $p(\mathcal{P}_i)$ can be normalized to become an actual PDF $H(\mathcal{P}_i)$ by setting a proper constant Γ. $\sigma = 0.2$ is chosen to substitute for h. The relative distance d is used in case the distribution of the superpixels is not uniform, and the distance metric mainly focuses on the relationships between similar superpixels. The relative average difference of a pair of superpixels $\mathcal{P}_i$, $\mathcal{P}_j \in W$ in intensity is defined as:

$$d(\mathcal{P}_i, \mathcal{P}_j) = \frac{|a(\mathcal{P}_i) - a(\mathcal{P}_j)|}{\text{ave}_{\mathcal{P}_k \in W}(|a(\mathcal{P}_i) - a(\mathcal{P}_k)|)} \tag{3}$$

where $W = \{\mathcal{P}_1, \mathcal{P}_2, \cdots, \mathcal{P}_n\}$ and $\text{ave}_{\mathcal{P}_k \in W}(|\mathcal{P}_i - \mathcal{P}_k|)$ is the average difference between the average intensity $a(\mathcal{P}_i)$ of pixels inside $\mathcal{P}_i$ and those $a(\mathcal{P}_k)$ of other superpixels $\mathcal{P}_k$ in W. Compared to the absolute difference, the relative difference is more consistent for two sets of pixels with similar neighboring relationships but different resolutions and scales [29].

After determining the probabilities of the superpixels, the dissimilarity measure $\text{dis}_I(\mathcal{P}_i, \mathcal{P}_j)$ between $\mathcal{P}_i$ and $\mathcal{P}_j$ is defined as:

$$\text{dis}_I(\mathcal{P}_i, \mathcal{P}_j) = \frac{(H(\mathcal{P}_i) - H(\mathcal{P}_j))^2}{H(\mathcal{P}_i) + H(\mathcal{P}_j)}. \tag{4}$$

The larger the relative difference of a superpixel from another, the less the similar they are, and the more dissimilar it is.

The distinctness value of each superpixel can be estimated using the dissimilarity measurement above. Superpixel $\mathcal{P}_i$ is considered salient when it is highly dissimilar to other superpixels. The saliency value of $\mathcal{P}_i$ is defined as:

$$\mathcal{S}_I(\mathcal{P}_i) = 1 - \exp\Big(- \frac{1}{n-1} \sum_{j=1, j\neq i}^{n} \text{dis}_I(\mathcal{P}_i, \mathcal{P}_j) \Big). \tag{5}$$

However, in order to reduce computational complexity, we note that it is unnecessary to evaluate the uniqueness of a superpixel by computing its dissimilarity to all the others. For instance, if the most similar superpixels $\mathcal{P}_j$ are significantly different from superpixel $\mathcal{P}_i$, then it follows logically that all the other superpixels are also highly different from superpixel $\mathcal{P}_i$. Therefore, for superpixel $\mathcal{P}_i$, only the M most similar superpixels $\{Q_m\}_{m=1}^{M}$ ($M = 10$ in this paper) need to be found and processed. Hence, the saliency value of superpixel $\mathcal{P}_i$ can be rewritten as:

$$\mathcal{S}_I(\mathcal{P}_i) = 1 - \exp\Big(- \frac{1}{M} \sum_{m=1}^{M} \text{dis}_I(\mathcal{P}_i, Q_m) \Big). \tag{6}$$

The final intensity-based saliency is obtained by fusing the saliency maps $\mathcal{S}_I(\mathcal{P}_i^l)$ of different superpixels $\mathcal{P}_i$ at different levels l . More specifically, all the pixels u within a superpixel will have the same value at each level (the same for fusing the compactness based saliency maps over all the levels). The fusion is performed pixel by pixel as: $\mathcal{S}_I(u) = \frac{1}{L} \sum_{l=1}^{L} \mathcal{S}_I(\mathcal{P}_i^l | u \in \mathcal{P}_i^l)$.

3) Compactness-based saliency detection: Intuitively, the leakage region in an FA image will present different intensity information when compared with the others. However, it is observed in practice that using the intensity feature alone to detect salient regions is not always successful. For example, the red rectangle region of the top row of Fig. 2 (c) shows that non-vessel regions in the middle of the image with high brightness due to uneven illumination have also been detected as highly salient, whereas a human observer perceives only the leakage regions and vessels as more salient. Therefore, this section proposes another feature - compactness. Normally, human observers pay more attention to a more compact object than to a more diffuse object. The measure of compactness of an object might therefore be of use as a complementary feature to intensity for saliency measurement, with the aim of reducing the number of falsely-detected salient regions.

For superpixel $\mathcal{P}_i$, its compactness $c(\mathcal{P}_i)$ is defined as

$$c(\mathcal{P}_i) = \exp\Big(- \alpha \frac{\sigma_{x,i} + \sigma_{y,i}}{\sqrt{\text{X}^2 + \text{Y}^2}} \Big), \tag{7}$$

where $\sigma_{x,i}$ and $\sigma_{y,i}$ are the standard deviations of the x and y coordinates of the pixels inside the superpixel $\mathcal{P}_i$, and α is a constant factor that is empirically set to 15. X and Y are the width and height of the input image. By incorporating the compactness feature with the intensity feature of a given image, the measure $\text{dis}_C(\mathcal{P}_i, \mathcal{P}_j)$ of dissimilarity in compactness between $\mathcal{P}_i$ and $\mathcal{P}_j$ is defined as:

$$\begin{aligned} \text{dis}_C(\mathcal{P}_i, \mathcal{P}_j) = |a(\mathcal{P}_i) - a(\mathcal{P}_j)| \times \Big(1 + \frac{c(\mathcal{P}_i) - c(\mathcal{P}_j)}{2} \Big) \\ \times \exp\Big(- \frac{\beta d(\mathcal{P}_i, \mathcal{P}_j)}{\sqrt{\text{X}^2 + \text{Y}^2}} \Big), \end{aligned} \tag{8}$$

where term $|a(\mathcal{P}_i) - a(\mathcal{P}_j)|$ calculates the difference of the average intensity (a) characteristic of superpixels $\mathcal{P}_i$ and $\mathcal{P}_j$. $d(\mathcal{P}_i, \mathcal{P}_j)$ is the relative average difference between superpixels $\mathcal{P}_i$ and $\mathcal{P}_j$, as proposed in Eq. (3). The constant factor β is empirically set to 300. The larger the dissimilarity $\text{dis}_C(\mathcal{P}_i, \mathcal{P}_j)$, the higher the probability that human attention will be paid from superpixel $\mathcal{P}_j$ to $\mathcal{P}_i$. Hence, the following rules in TABLE I can be used to assist in estimating the saliency value $\mathcal{S}_C(\mathcal{P}_i)$ of superpixel $\mathcal{P}_i$.

TABLE I: Some useful rules for the determination of saliency from compactness

Condition	Expected dissimilarity	Expected salience
$\mathcal{P}_i$ is distinct from $\mathcal{P}_j$	large $\text{dis}_C(\mathcal{P}_i, \mathcal{P}_j)$	$\mathcal{S}_C(\mathcal{P}_i) > \mathcal{S}_C(\mathcal{P}_j)$
$\mathcal{P}_i$ is similar to $\mathcal{P}_j$	small $\text{dis}_C(\mathcal{P}_i, \mathcal{P}_j)$	$\mathcal{S}_C(\mathcal{P}_i) \approx \mathcal{S}_C(\mathcal{P}_j)$
$c(\mathcal{P}_i) > c(\mathcal{P}_j)$	large $\text{dis}_C(\mathcal{P}_i, \mathcal{P}_j)$	$\mathcal{S}_C(\mathcal{P}_i) > \mathcal{S}_C(\mathcal{P}_j)$
$c(\mathcal{P}_i) < c(\mathcal{P}_j)$	small $\text{dis}_C(\mathcal{P}_i, \mathcal{P}_j)$	$\mathcal{S}_C(\mathcal{P}_i) < \mathcal{S}_C(\mathcal{P}_j)$

Similar to Eq. (6), the compactness-based saliency value $\mathcal{S}_C(\mathcal{P}_i)$ of $\mathcal{P}_i$ is defined as

$$\mathcal{S}_C(\mathcal{P}_i) = 1 - \exp\Big(- \frac{1}{M} \sum_{m=1}^{M} \text{dis}_C(\mathcal{P}_i, R_m) \Big) \tag{9}$$

where $R_m (m = 1, 2, \cdots, M)$ is the M most similar superpixels to $\mathcal{P}_i$ in the sense of compactness.

Again, we calculate the final compactness-based saliency based on the mean value of the saliency maps $\mathcal{S}_C(\mathcal{P}_i^l)$ of different superpixels $\mathcal{P}_i$ at different levels l, and the fusion is performed pixel by pixel as well: $\mathcal{S}_C(u) = \frac{1}{L} \sum_{l=1}^{L} \mathcal{S}_C(\mathcal{P}_i^l | u \in \mathcal{P}_i^l)$.

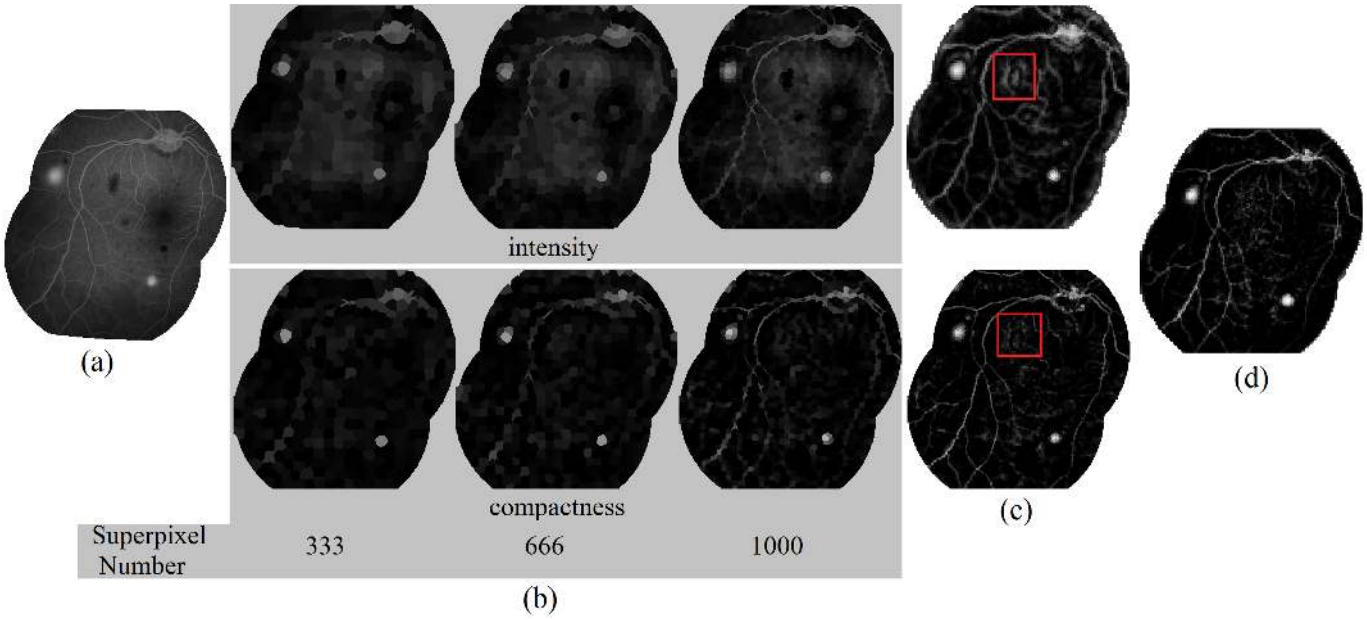


Fig. 2: An example to illustrate a saliency map generated by the proposed method: (a) an example FA image; (b) saliency maps estimated using intensity and compactness features with different numbers of superpixels; (c) fused saliency maps of intensity and compactness across different levels of superpixels; (d) final saliency map.

4) Saliency map fusion: Two bottom-up approaches in our proposed superpixel based saliency detection method have been described so far. It is likely that each of them has its own drawbacks if used alone in real applications. Therefore, an overall saliency map by fusing the saliency maps based on intensity and compactness is expected to provide better performance. linear summation [5] or pixel-wise multiplication (also known as the matrix Hadamard product) [30] are two commonly used methods to fuse the saliency maps. In this work, the intensity and compactness saliency maps are fused by applying the pixel-wise multiplication method so as to force only the regions with higher values in both intensity and compactness channels to be assigned higher values in the final saliency map $\mathcal{S}$. By integrating the two saliency measures, the property of human vision by which attention declines as the edge of the area of interest is approached may be mimicked. That is, the final saliency map highlights salient object regions of interest and suppresses background regions, as illustrated in Fig. 2 (d).

B. Graph cut for leakage detection

The proposed superpixel-based saliency detection approach has successfully enhanced the contrast between vessels/leakages and background. Some example results are shown in Fig. 3 (b). The appearances of these leakages are highlighted, while the background regions are suppressed, when compared to the original images. Once the saliency map is computed and normalized to [0, 1], a threshold value $T = 0.65$ is applied to the saliency map to obtain the ROIs. The thresholding approach itself cannot guarantee the boundaries of the segmented structures are smooth and often generates isolated fragments. In light of this inadequacy, more sophisticated segmentation methods [31]–[36] will be needed for better results. On the other hand, the computational cost is also an important factor for a segmentation tool to be taken into account for potential real applications. For these two reasons, we advocate here a graph cut based segmentation method [33], [34] on the obtained ROIs to identify the leakage. This method imposes the constraint that the neighboring pixels tend to belong to the same class and thus penalizes the isolated pixels in different classes.

Let $\mathcal{N}$ be a set of edges $\{(u, v)\}$ where a pixel u is connected to its 8 nearest neighbors v, and $\mathcal{M}$ denote the set of pixels in the given image I, the discrete energy function is defined as:

$$E(x) = \sum_{u \in \mathcal{M}} E_u(x_u) + \sum_{(u,v) \in \mathcal{N}} E_{uv}(x_u, x_v), \qquad (10)$$

where $x = \{x_1, \cdots, x_{\mathcal{N}}\}$ is the binary labelling where the x_u is either 0 or 1 depending on whether the pixel u belongs to the background or foreground. The first term here approximates the region terms while the second term approximates the regularization term. The unary term E_u is defined as:

$$E_u^0(x_u) = \lambda_1 (I_u - c_1)^2, E_u^1(x_u) = \lambda_2 (I_u - c_2)^2 \qquad (11)$$

where E_u^0, E_u^1 denote the weights between the node u and the two terminals, λ_1 and λ_2 are the non-negative region weighting parameters, I_u is the intensity of the pixel u, and c_1 and c_2 indicate the average intensities of the background and foreground respectively. The binary term E_{uv} is defined as:

$$E_{u,v}(x_u, x_v) = \begin{cases} \mu w_{uv}, & \text{if } x_u \neq x_v \\ 0, & \text{otherwise,} \end{cases} \qquad (12)$$

where w_{uv} denotes the weight between neighboring pixels u and v, as suggested in [37]:

$$w_{uv} = \frac{\delta^2 \cdot \Delta\phi_{uv}}{2 \cdot |e_{uv}|}, \tag{13}$$

where δ is the cell-size of the grid, $|e_{uv}|$ is the Euclidean length of the edge e_{uv}, and $\Delta\phi_{uv}$ is the difference between the angular orientations ϕ_u and ϕ_v of the pixels u and v and is restricted to the interval $[0, \pi]$. In this work, we set $\lambda_1 = \lambda_2 = \lambda = 0.5$ (see Sec. VI for the parameter tuning), and μ is empirically set as 0.2.

C. Final refinement

After the graph cut segmentation, some vessels, the optic disc and some small objects may still remain as they may also have been enhanced during the saliency detection steps. It is important to remove them in order to improve the leakage detection performance. To this end, the following steps are applied: (i) The infinite perimeter active contour with hybrid region (IPACHR) method [38] is used to segment retinal vessels for its good performance. In brief, this method uses an infinite perimeter active contour model for its effectiveness in detecting objects (e.g. vessels) with irregular and oscillatory boundaries. Moreover, this method considers hybrid region information (local phase based vesselness map and intensity) in an image in order to achieve further improved performance compared to the standard infinite perimeter active contour model [38]. For more details, we refer readers to the original paper [38]. (ii) Any small and/or isolated objects are eliminated by the use of a disk-shaped opening operation with a radius of 2 pixels. (iii) In most cases, the optic disc remains as leakage regions after the graph cut based segmentation and should be removed. It has been well observed that the number of vessels surrounding the optic disc is much larger than that close to large focal leaking sites [11], [39]. Thus, any region with a number of surrounding vessels greater than a threshold of 5 will be assumed to be the optic disc, and will be removed. In our experiments this method is found to be efficient and effective. However, other sophisticated methods may work equally well.

IV. Datasets and Evaluation Metrics

Our method will be evaluated on two FA image datasets with two different retinal diseases: DR and MR respectively. To the best of our knowledge, these two datasets are the only FA datasets available in the literature for the evaluation of leakage detection algorithms.

A. DR dataset

The FA images of the DR set [12] were collected by the Vision and Image Processing Laboratory, Duke University, USA and are currently publicly accessible. All images were acquired using a Heidelberg Spectralis 6-mode HRA/OCT unit (Heidelberg Engineering, Heidelberg, Germany). Each image has 768×768 pixels. The study was approved by the Duke University Health System Institution Review Board (IRB). The tenets of the Declaration of Helsinki were adhered to. It contains images of 24 eyes taken from 24 subjects. All subjects had signs of DR on admission. All the images were categorized into three types according to their leakage conditions: predominantly focal, predominantly diffuse, and mixed pattern leakage. Focal leakage manifests as discrete foci of leakage on early FA frames and corresponds to microaneurysms. Diffuse leakage is characterized by generalized leakage prominent on late FA frames without discretely identifiable source.

B. MR dataset

The MR dataset contains 25 FA images and all had signs of MR on admission. These images were randomly chosen from images systematically sorted and graded for quality by the Liverpool Reading Center at St Paul's Eye Unit, Royal Liverpool University Hospital and the Department of Eye and Vision Science, University of Liverpool. The FA images were taken after pupil dilation with Tropicamide 1% and Phenylephrine 2.5%, using a Topcon 50-EX optical unit (Topcon, Japan) and a Nikon E1-H digital camera. The tenets of the Declaration of Helsinki were adhered to. Ethical approval for retinal examination and imaging was given by committees in Blantyre, Malawi and at collaborating institutions. Consent was given by the parents/guardians of children before examination and imaging. 50-degree images were taken after pupil dilation, using a Topcon 50-EX fundus camera (Topcon, Japan). All macula-centered images have a resolution of 3008×1960 pixels and were re-sized to 752×490, similar to the size of the DR images above, so that a single set of parameters can be tuned over both datasets. This is important for the test of the proposed technique whether it is generalisable to different datasets.

C. Evaluation Metrics

In this paper sensitivity (*Se*), specificity (*Sp*), false negative rate (*Fnr*), accuracy (*Acc*), area under the receiver operating characteristic curve (*AUC*), and the Dice coefficient (*DC*) are used to measure the performance of the proposed method. Sensitivity (resp. specificity) is a measure of effectiveness in identifying pixels or regions with positive (resp. negative) classifications. Both accuracy, *AUC* and *DC* measure the overall segmentation performance.

In essence, leakage detection may be seen as an imbalanced data classification problem: there are typically much fewer leakage pixels than others. In such a case *AUC* can better reflect the trade-offs between sensitivity and specificity. In particular, the *AUC* proposed by Hong et al. [40] is employed here, as it was proposed to evaluate the segmentation (or classification) performance at a specific operating point.

More specifically, the chosen metrics are defined as follows: *sensitivity (Se)* $= tp/(tp + fn)$; *specificity (Sp)* $= tn/(tn + fp)$; *accuracy (Acc)* $= (tp + tn)/(tp + fp + tn + fn)$; *true positive rate (Tpr)* $=$ *Se*; *false positive rate (Fpr)* $= 1 -$ *Sp*; *false negative rate (Fnr)* $= 1 -$ *Se*; and *AUC* $= (Se + Sp)/2$ where *tp*, *tn*, *fp* and *fn* indicate the true positive (correctly identified leakage pixels or regions), true negative (correctly

identified background pixels or regions), false positive (incorrectly identified leakage pixels or regions), and false negative (incorrectly identified background pixels or regions), respectively, and all the pixels are equally treated towards their counting without considering the severity of the symptoms they depict. $DC = 2(|A \cap B|)/(|A| + |B|)$, where A is the ground truth region, B indicates the segmented region, and $|A \cap B|$ denotes the number of pixels in the intersected region between A and B.

We grouped all the experimental results together first and then calculate reliably the final scores of performance measurements. For each dataset, two graders (the inter-observers) were invited to manually annotate the leakage regions, one of them was asked to repeat the annotations after 4 weeks (the intra-observer), and the ground truth was obtained from the consensus between the two graders. Statistical analysis is performed as appropriate in order to evaluate the relative performance of different segmentation methods. Due to the relatively small number of images, $p < 0.01$ is considered to be statistically significant. All the experiments were carried out in MATLAB2015a on a PC with an Intel Core i7-4790K CPU, 4.00GHz, and 16GB RAM.

V. Experimental results

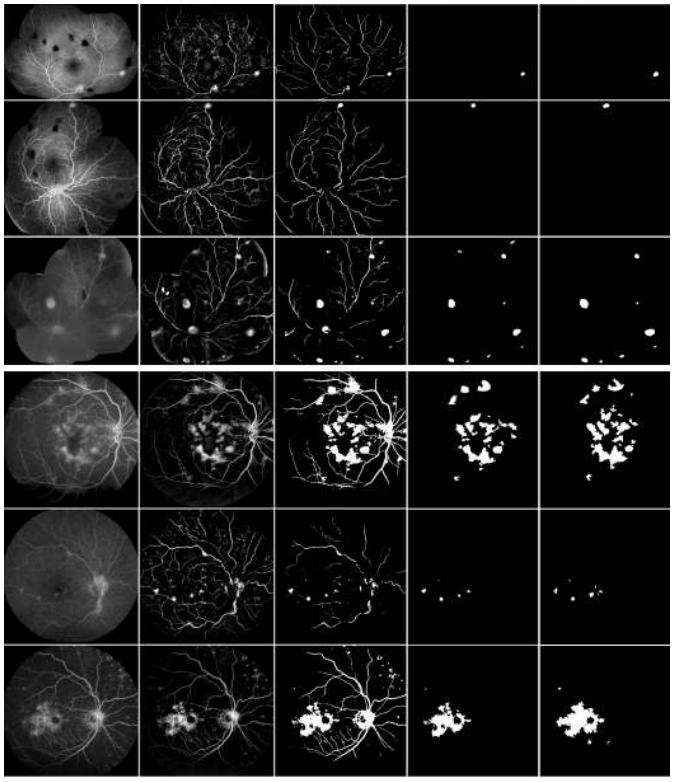
The proposed saliency-guided leakage detection method is evaluated from two aspects: leakage detection over different datasets, and the comparison with existing state-of-the-art saliency detection methods. An experimental investigation will also be carried out in the next section on the effectiveness of different saliency cues and parameter setting - i.e., the level of superpixel maps and the number of superpixels in each level, the threshold value for the generation of the ROIs from the final saliency map and the region weight λ.

A. Results on different datasets

In this section, we quantitatively evaluate the performance of our algorithm in direct comparison with both the leakage detection performance of human graders and that of an existing alternative method. To this end, the method proposed by Rabbani et al. [12] was re-implemented in our study, and applied to the MR dataset. For the DR dataset, however, we directly quoted the results reported in their paper in the hope that their results are the best achievable.

TABLE II: The performances of different methods in detecting focal leakages over the MR dataset at the site level. The number in the brackets indicates the relative performance measurement. 41 leaking sites were manually annotated as ground truth.

	Rabbani et al. [12]	**Proposed**
#(detected focal leakages)	35	40
tp (Se)	32 (0.78)	40 (0.98)
fp (1-Sp)	3 (0.07)	0 (0)
fn (Fnr)	6 (0.14)	1 (0.02)



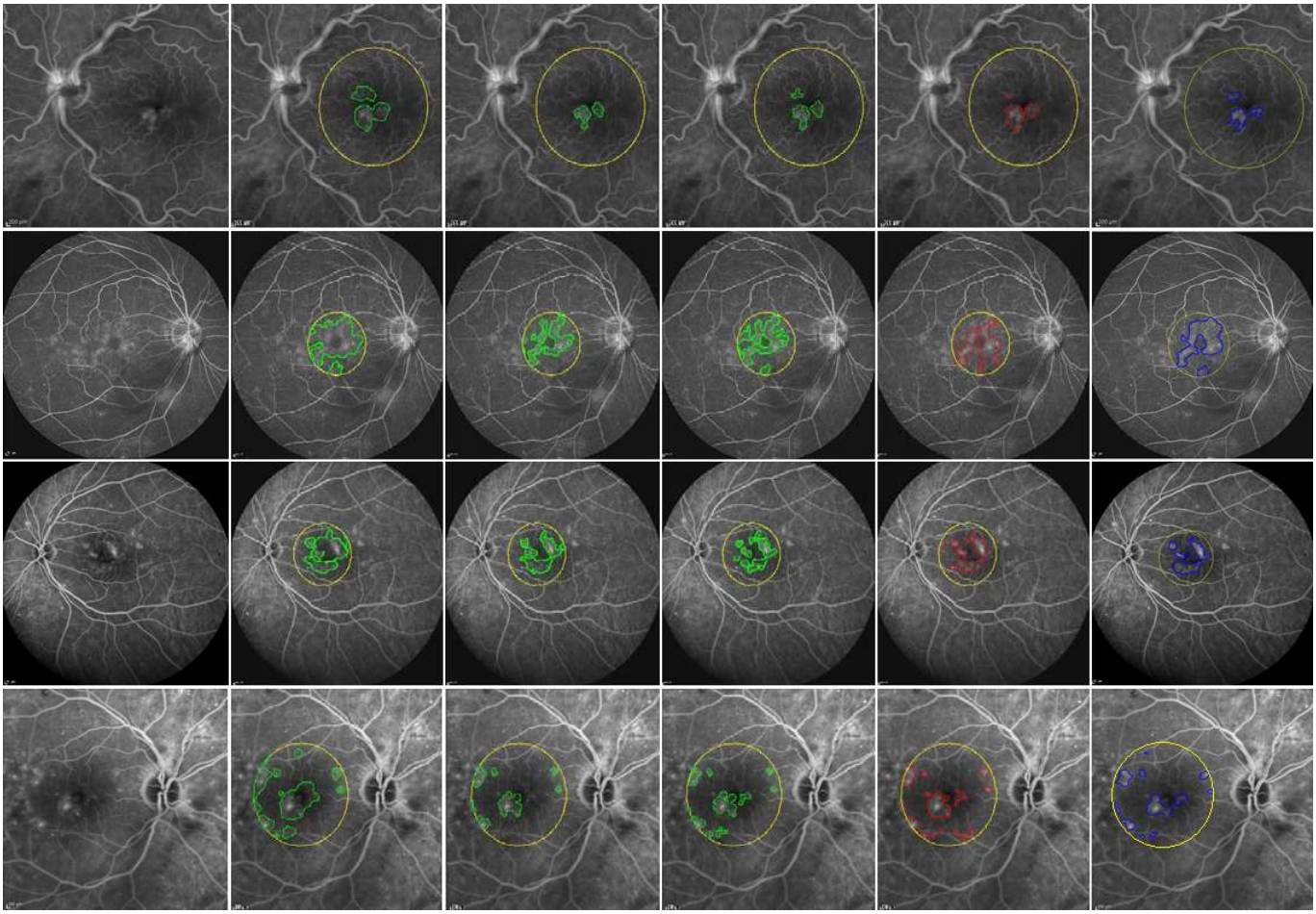
(a) (b) (c) (d) (e)

Fig. 3: Examples illustrating the main steps of our algorithm for detecting leakages. The images in top three rows are from MR dataset, and the images in bottom three rows are from DR dataset. (a) Example FA images; (b) saliency maps of (a). The bright regions indicate the more salient regions, and the dark areas show the less salient regions; (c) binary images of (b) obtained by applying the threshold value T; (d) the detected leakage regions after masking vessels and optic disc; (e) expert's annotations.

TABLE III: The performances (average $\pm$ one standard deviation) of different methods on detecting the focal leakages over the MR dataset at the pixel level.

	Intra obs.	**Inter obs.**	**Rabbani et al. [12]**	**Proposed**
Se	0.96±0.02	0.91±0.04	0.81±0.08	0.93±0.03
Sp	0.97±0.03	0.94±0.05	0.87±0.08	0.96±0.02
Acc	0.96±0.03	0.89±0.04	0.83±0.10	0.91±0.03
AUC	0.96±0.02	0.92±0.04	0.84±0.08	0.94±0.02
DC	0.92±0.04	0.80±0.05	0.74±0.05	0.82±0.03

1) MR Dataset: The leakage detection results over the MR dataset are illustrated in the top three rows of Fig. 3. It can be seen that most of the leaking areas were correctly identified by our automated method. TABLE II shows the performances of different methods in detecting the focal leakage sites. According to the human reference standard there were 41 sites of large focal leakage in 25 images (one image per patient). Our method failed to detect only one out of all these sites, achieving a sensitivity of as high as 0.98, and the false negative rate of as low as 0.02. It is interesting to note that our method produces a false positive rate of 0, which means there were no



(a) (b) (c) (d) (e) (f)

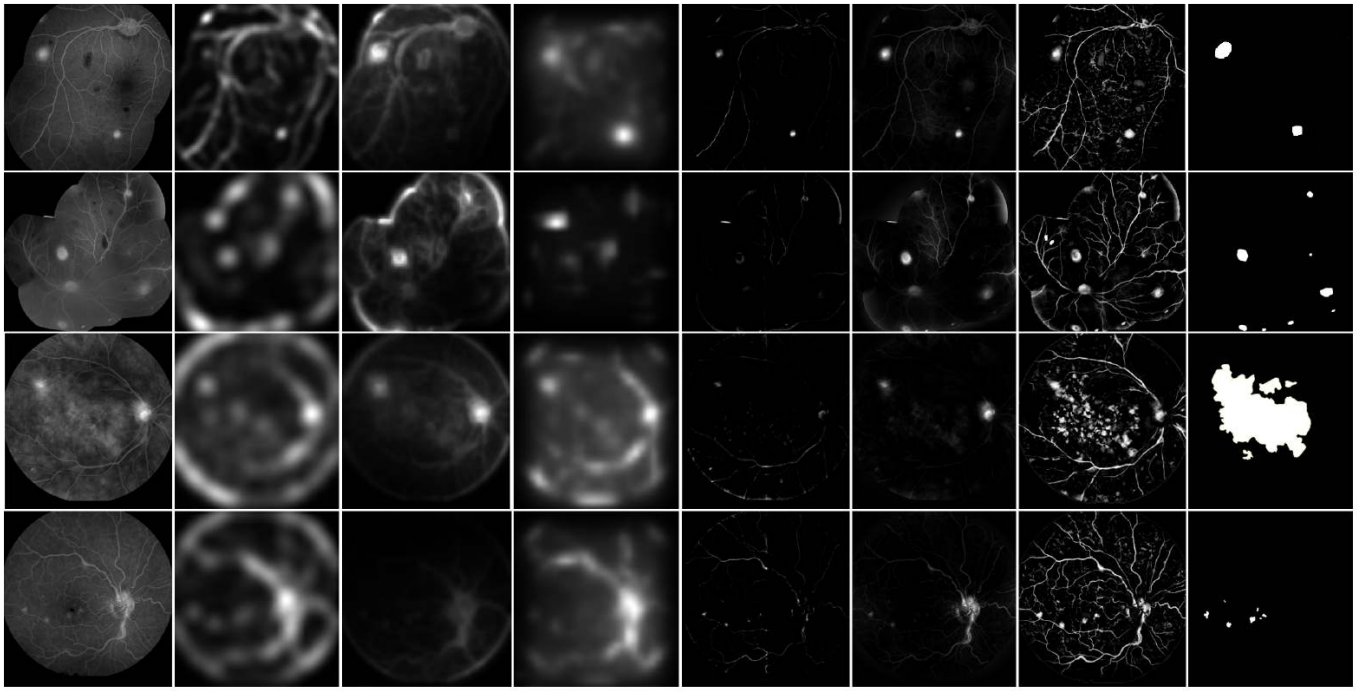
Fig. 4: Leakage segmentation results by experts (green labels), existing method [12] (red labels), and the proposed automated method (blue labels). For fair comparison, only the regions (yellow circle) centered at the fovea with a diameter of $3000\mu m$ were considered as regions of interest. Note: (a)-(e) were quoted from [12] for convenience. (a) Example FA images. (b) Annotation of Expert 1. (c) Annotation of Expert 2. (d) Re-annotation of expert 2 after 4 weeks. (e) Leakage detected by Rabbani et al.'s method. (f) Leakage detected by our proposed method.

regions falsely identified as large focal leakage sites. In sharp contrast, the method proposed by Rabbani et al [12] produces relatively poorer results. It has only successfully detected 32 focal leakage sites, which gives a sensitivity of only 0.78. In addition, 3 non-leaking regions were falsely detected as focal leakage sites. This is because it used only intensity information for the task, which failed to distinguish leaking sites from non-leaking ones with high intensity values. TABLE III shows the performances of different methods in detecting the focal leakage sites at the pixel level. It shows that the proposed method achieves competitive results to human experts: the mean accuracy of 0.96 ± 0.03 for the manual intra-observers; 0.92 ± 0.04 for the manual inter-observers; 0.83 ± 0.10 for Rabbani's method; and 0.91 ± 0.03 for the proposed method. The statistical analysis shows that the performance of the proposed method is significantly higher than that of the method proposed by Rabbani et al [12] (2-tailed t-test, all $p < 0.0001$).

2) DR Dataset: The proposed method was also tested on the DR dataset with the aim of detecting the leakage areas caused by diabetic macular edema. As suggested in [12], quantitative analysis of a circular region centred at the fovea with a radius of 1500 μm is of greatest significance for clinical diagnosis and treatment. For a fair comparative study, we also limited our method in detecting the leakages in this area. Fig. 4 shows the results of different methods in detecting the leaking regions abound the fovea. It can be seen that most of the leaking areas have been detected by both methods, and that the segmentation results are very similar to those of manual annotations.

TABLE IV: The performances (average $\pm$ one standard deviation) of different methods on detecting the focal leakages over the DR dataset at the pixel level.

	Intra obs.	**Inter obs.**	**Rabbani et al. [12]**	**Proposed**
Se	0.95±0.05	0.78±0.09	0.69±0.16	0.78±0.06
Sp	0.73±0.27	0.94±0.08	0.91±0.09	0.94±0.02
Acc	0.83±0.16	0.90±0.08	0.86±0.08	0.89±0.06
AUC	0.84±0.16	0.91±0.08	0.80±0.12	0.86±0.04
DC	0.80±0.08	0.82±0.03	0.75±0.05	0.81±0.02



(a) Input (b) IT (c) GB (d) CA (e) SR (f) FT (g) IC (h) Groundtruth

Fig. 5: Saliency detected on single images from the MR and DR datasets by different algorithms. (a) Original FA images. (b)-(g) Saliency maps generated using different methods. (h) Ground truth. Note, the ground truth only indicates the leaking regions. In the saliency detection step, the large vessels and optic disc are also assigned as salient regions but will be removed later.

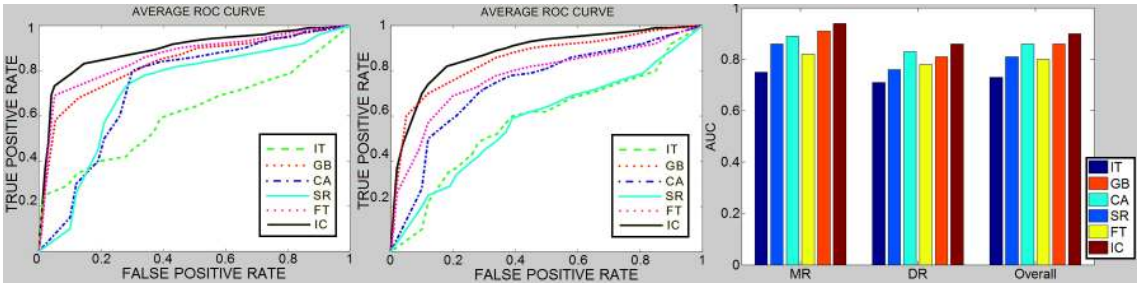
It is difficult to distinguish visually between the two methods. Quantitative results are thus provided in TABLE IV, in terms of sensitivity, specificity, accuracy, area under curve, and Dice coefficient. It can be clearly seen that our automated method outperforms Rabbani's method again, and has relatively better stability. To be more specific, the mean accuracy was 0.89 ± 0.06 for our method; 0.86 ± 0.08 for Rabbani's method; 0.83 ± 0.16 for the manual intra-observers; and 0.90 ± 0.08 for the manual inter-observers. It is interesting to note that the accuracy for intra-observer annotations was lower than that of our automated method, which is very close to the accuracy of the inter-observer annotations. This implies that the DR dataset suffers from noise and other distortions common in real-world clinical imaging: a finding that was also reported in [12]. These results show that humans are prone to variability. The statistical analysis shows that the performance of our proposed method is significantly better than the method proposed by Rabbani et al. (2-tailed t-test, all $p < 0.0001$).

In summary, based on both the quantitative and qualitative comparisons on two different datasets, it can be seen that our automated method is effective, and is superior to the existing automated method in detecting the focal leakages. When compared with the annotations of human observers, it can be seen that on one hand, our method can perform as well as a human expert. On the other hand, human observers are prone to variability (relatively lower *Sp*, *Acc*, and *AUC* scores than the proposed method).

B. Saliency detection

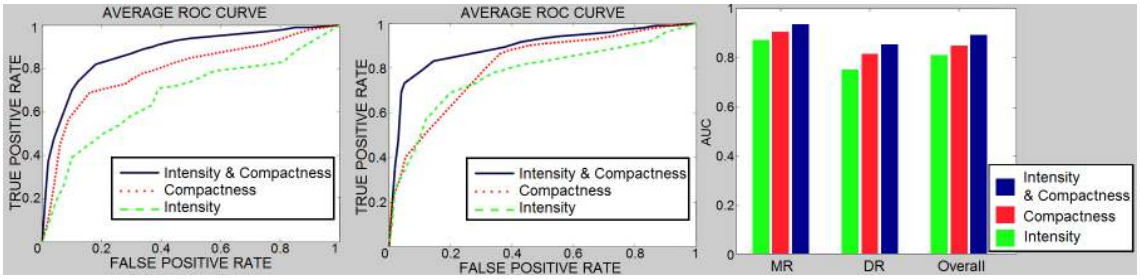
In this section, we carry out a comparative study between the proposed method and the state-of-the-art ones for the detection of salient objects, including vessels, focal leakage, and optic disc, over the MR and DR datasets. Since the detection of salient objects is an immediate step of our method, such comparative study will help further explain its superior performance reported in the last section. To this end, five saliency detection methods were selected: the classic method [5], spectral residual saliency [21], frequency-tuned saliency [6], graph-based visual saliency [25] and context-aware saliency [41]. These competitors and the proposed method are referred to as IT, SR, FT, GB, CA, and IC, respectively. The source codes with default parameter settings provided by the authors for these methods were used, and all generated saliency maps were normalized into the same range of [0,1] with a full resolution of the original images.

The saliency maps estimated by the six different methods are presented in Fig. 5. It can be clearly seen that the proposed method has successfully detected all main objects, including the focal leakages and vessels, which are largely consistent with the results of visual inspection. The SR method has the poorest performance, since the spatial information is lost in the Fourier representation. This means that the spectral energies derived from frequency bands in Fourier domain alone may not be sufficient. Compared with other models, the proposed method preserves salient object boundaries more accurately and highlights the complete salient objects more effectively.



(a) MR (b) DR (c) *AUC* values

Fig. 6: (The reader is referred to the color version of this figure) ROC curves and *AUC* values of different methods over two different datasets. (a) ROC curves over the MR dataset. (b) ROC curves over the DR dataset. (c) *AUC* values.



(a) MR (b) DR (c) *AUC* values

Fig. 7: ROC curves and *AUC* values achieved by the proposed method with different saliency cues over two different datasets. (a) ROC curves over the MR dataset. (b) ROC curves over the DR dataset. (c) *AUC* values.

It is not only capable of suppressing background, but also highlights all salient regions (e.g. leakage area, vessels, and optic disc) with well-defined boundaries. By utilizing the intensity and compactness information, the proposed method can better handle heterogeneous objects (row 2 in Fig. 5), low contrast between objects and background (row 4 in Fig. 5), large-scale salient regions (row 3 in Fig. 5) and small-scale salient objects (row 4 in Fig. 5) more effectively compared with other saliency detection methods.

In order to objectively measure the performance, the false positive rate and true positive rate of the saliency maps derived by different methods were then calculated, by sweeping a threshold from 0 to 1 over the final saliency map. The averaged results over different images of our method and its competitors are plotted as ROC curves in Fig. 6 (a) and (b). It can be seen that our method achieves the best performance over both MR and DR datasets. The *AUC* values were also calculated from the ROC curves of all these methods and are illustrated in Fig. 6 (c). It can be seen that the proposed method consistently outperforms its competitors. It can also be seen that the AUC value achieved by each method is higher over the MR dataset than over the DR dataset. This suggests that the leakage detection on the DR dataset is relatively more challenging than over the MR dataset.

VI. Discussions

Our proposed method includes a number of free features and parameters: the effectiveness of each saliency cue, the numbers of superpixels and levels of superpixel partition; the threshold value for the generation of ROIs from the final saliency map; and the weighting parameter λ for the graph-cut segmentation. In this section, we experimentally investigate their effect on the segmentation.

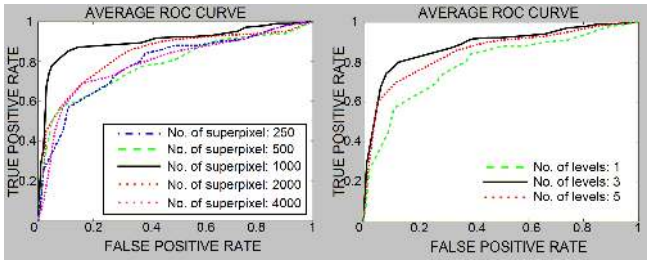
A. The effectiveness of each saliency cue

In our method, two cues were employed to measure salience in each FA image: intensity and compactness. To validate the effectiveness of each cue, we generated three ROC curves and calculated corresponding AUC values over the MR and DR datasets separately: intensity cue only, compactness cue only, and combined intensity and compactness cues.

Fig. 7 shows the experimental results. It can be seen that the ROC curve using the saliency map based on both cues is higher than that using either the intensity cue or the compactness cue alone. Essentially, the proposed method utilizes both the global intensity and compactness information to constrain the saliency detection problem: the leakage regions in FA images have particular color (intensity cue) and shape (compactness) characteristics. The intensity cue is able to reveal important regions at different scales in the image. The compactness cue is effective for distinguishing salient regions against background. Our results indicate that intensity and compactness cues have a complementary effect for the definition of saliency.

B. The numbers of the superpixel maps and superpixels in each map

Choosing a suitable number of superpixels is usually empiric and case-specific for most of segmentation methods. On one hand, too large a number of superpixels leads to



(a) (b)

Fig. 8: The ROC curves of the proposed method with (a) different numbers of superpixels: 250, 500, 1000, 2000, and 4000; (b) different numbers of levels, after setting the optimal number of the superpixels to 1000.

TABLE V: The average $\pm$ standard deviation of various parameters and computational time **t** in seconds of the proposed method with the region weight λ taking different values over different datasets.

dataset	λ	**Se**	**Sp**	**Acc**	**AUC**	**t** (sec)
MR	**0.1**	0.93±0.02	0.96±0.02	0.91±0.02	0.94±0.02	13.1±3.1
	0.5	0.93±0.03	0.96±0.02	0.91±0.03	0.94±0.02	9.6±2.8
	1	0.91±0.04	0.95±0.02	0.90±0.04	0.93±0.03	7.1±2.2
	1.5	0.91±0.03	0.95±0.01	0.90±0.03	0.93±0.02	6.5±1.8
	2	0.90±0.04	0.94±0.03	0.90±0.04	0.92±0.02	5.3±1.1
DR	**0.1**	0.78±0.02	0.94±0.02	0.89±0.04	0.86±0.03	15.2±3.2
	0.5	0.78±0.06	0.94±0.02	0.89±0.06	0.86±0.04	10.9±3.0
	1	0.77±0.04	0.94±0.01	0.88±0.05	0.85±0.04	8.4±2.8
	1.5	0.77±0.04	0.94±0.02	0.88±0.06	0.85±0.04	7.9±2.2
	2	0.76±0.05	0.93±0.03	0.88±0.06	0.85±0.04	6.5±1.2

over-segmentation and over-intensive computation. On the other hand, too few superpixels result in a loss of the edge information of the objects of interest. In this section, we experimentally tune this parameter. To this end, it was set to be successively 250, 500, 1000, 2000 and 4000. Fig. 8 (a) shows the performances of the proposed method under these test numbers, and reveals that the proposed method achieves the best performance when the superpixel number is 1000.

Multiscale analysis is useful to reveal the saliency at different scales. Too few scales may miss the saliency at small scales. Too many scales may detect unimportant objects at the cost of intense computation. Thus, in this section, the optimal number of superpixel levels is evaluated, and results are shown in Fig. 8 (b), where the number of superpixels was set to 1000 at the finest level. It can be clearly seen that our method achieves the best performance when the number of levels is 3. This means that with the optimal number of superpixels being 1000, those at the other two levels were set to $1000 \times \frac{1}{3} = 333$ and $1000 \times \frac{2}{3} = 666$ respectively. It is worth mentioning that the specified combination of the number of superpixels and the number of levels may not be the best choice for our proposed method.

C. Threshold value for the generation of ROIs

The binary segmentation of ROIs from the final saliency map can be obtained by using a simple threshold-based method. In this case, there is a concern that the segmentation results may be affected by the chosen threshold. In this section, we experimentally investigate how to set up such threshold. To this end, we varied the threshold T from 0 to 1 with steps of 0.05. Fig. 9 shows the *AUC* values of our method with the threshold T taking different values over two different datasets. It can be seen that it achieves the highest *AUC* value when $T = 0.65$, for all cases. These results thus justify our choice of $T = 0.65$ throughout this paper.

D. The effect of the region weight λ

In order to demonstrate the robustness of the graph-cut based leakage segmentation, the effect of the region weighting factor λ in the energy minimization function (Eq. 11) are evaluated in this section. λ balances the smoothness of the detected boundary and the uniformity of the detected regions. Usually, the smaller the value of λ, the smoother the boundaries and the larger the regions the segmentation will produce, while a larger λ obtains more complex boundaries and larger regions. To reliably compare how significantly various values of λ affect the detected leakages, a range of values were tested to show the sensitivity of our method.

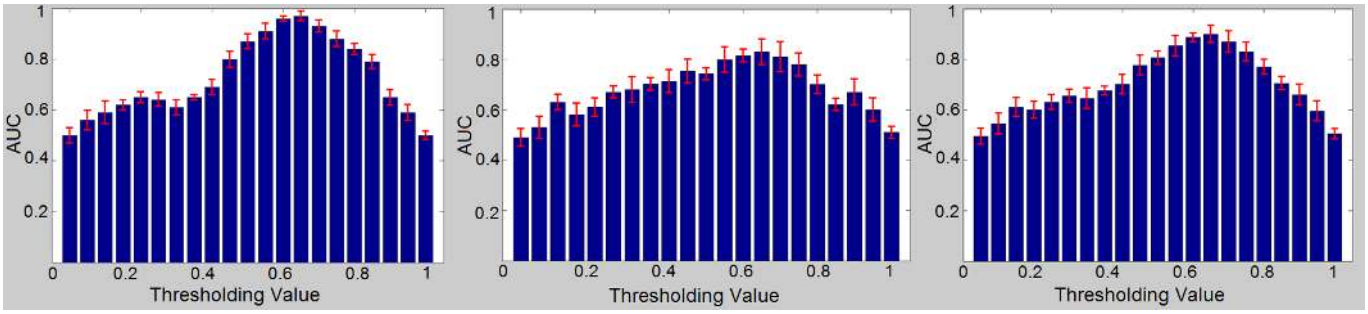
TABLE V shows the experimental results on the performance of the proposed method in terms of pixel-wise sensitivity, specificity, accuracy, and area under curve. In addition, the elapsed time of each trial was also measured. From this table, it can be seen that $\lambda = 0.5$ made the best trade-off between the accurate detection of the leaking regions and the computational time. These results justify our choice of $\lambda = 0.5$ in our method.

VII. Conclusions

It is important to distinguish between leakage in FA and retinal lesions (e.g. drusen, exudate, microaneurysm, pigment abnormalities) commonly seen in colour fundus photograph. Leakage shows activities of retinal diseases while lesions reveals existence or absence of certain types of disease. An extensive literature review shows that automated retinal image analysis of FA images, especially for leakage detection, is relatively unexplored. To the best of our knowledge, this is the first report on the automated detection of the leakage over both DR and MR datasets with the largest number of cases.

In this paper, we have proposed a multiscale saliency detection method for the detection of focal leakages in FA images. The proposed method is based on two saliency cues: intensity and compactness features under multi-level superpixels. Then the saliency values of the superpixels at different levels are estimated in the intensity and compactness channels respectively. While the intensity cue characterizes the intensity contrast among different superpixels, the compactness cue characterizes how densely (or sparsely) the salient pixels distribute inside a superpixel. The superpixel representation helps capture large objects of interest but at a low computational cost, and multiscale analysis helps capture the objects of interest with different sizes.

The saliency maps over the same cues at different levels are fused using a pixel-wise multiplication operator, so that only



(a) MR (b) DR (c) Overall

Fig. 9: The *AUC* values of the proposed method with the threshold T taking different values over two different datasets. (a) MR. (b) DR. (c) Averages of *AUC* values over MR and DR.

such regions that are salient in both channels are detected as salient. The saliency detection step can generate accurate saliency maps with well-highlighted leakage sites and areas. Thus, it can provide both the qualitative and quantitative information for the analysis of the FA images. The regions of interest (ROIs) in the given image are detected through thresholding the saliency map. However, such simple thresholding method usually renders the detected ROIs rugged and isolated. To avoid such shortcoming, the powerful graph-cut method [33], [34] is employed to segment the thresholded image so that the neighbouring pixels tend to belong to the same class of either foreground (ROIs) or background.

The experimental results based on two publicly accessible MR and DR datasets show our method outperforms one of the latest competitors and performs as well as a human expert for leakage detection and outperforms several state-of-the-art methods for saliency detection: it is not only capable of identifying the location of leaking regions, but also has the ability to measure the size of such regions.

To further demonstrate the merits of our proposed method and justify its remarkable performance, we carried out an extensive comparative study with other methods for saliency detection. The experimental results based on the MR and DR datasets show that our method is superior for the detection of salient objects and structures in the FA images. We plan to apply this new tool to assist the management of retinal diseases such as DR and MR. We also plan to automate the parameter tuning process. We also plan to apply the proposed saliency detection method to other types of images, and evaluate its performance on other benchmark datasets [42].

REFERENCES

[1] G. Richard, G. Soubrane, and L. Yannuzzi, *Fluorescein angiography: textbook and atlas (2nd rev. and expanded ed.)*. Thieme, 1998.

[2] T. Chui, M. Dubow, A. Pinhas, N. Shah, A. Gan, R. Weitz, Y. Sulai, A. Dubra, and R. Rosen, "Comparison of adaptive optics scanning light ophthalmoscopic fluorescein angiography and offset pinhole imaging," *Biomed. Opt. Express*, vol. 5, no. 4, pp. 1173–1189, 2014.

[3] N. Patton, T. Aslam, T. MacGillivray, A. Pattie, I. J. Deary, and B. Dhillon1, "Retinal vascular image analysis as a potential screening tool for cerebrovascular disease: a rationale based on homology between cerebral and retinal microvasculatures," *J Anat.*, vol. 206, pp. 319–348, 2005.

[4] M. Potchen, S. Kampondeni, K. Seydel, G. Birbeck, C. Hammond, and W. Bradley, "Acute brain mri findings in 120 malawian children with cerebral malaria: new insights into an ancient disease," *Am. J Neuroradiol.*, vol. 33, pp. 1740–1746, 2012.

[5] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. on Pattern Anal. Mach.Intell.*, vol. 20, pp. 1254–1259, 1998.

[6] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency-tuned salient region detection," in *Proc. IEEE CVPR*, 2009, pp. 1597–1604.

[7] C. Koch and C. Ullman, "Shifts in selective visual attention: towards the underlying neural circuitry," *Human Neurbiol.*, pp. 219–227, 1985.

[8] M. Cheng, G. Zhang, N. Mitra, X. Huang, and S. Hu, "Global contrast based salient region detection," in *Proc. IEEE CVPR*, 2011, pp. 409–416.

[9] X. Chen, H. Zhao, P. Liu, and W. Ren, "Automatic salient object detection via maximum entropy estimation," *Opt. Lett.*, vol. 38, pp. 1727–1729, 2013.

[10] R. Achanta, A. Shaji, K. Smith, A. Lucchi, and P. Fua, "Slic superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach.Intell.*, vol. 34, pp. 2274–2282, 2012.

[11] Y. Zhao, I. J. MacCormick, D. Parry, S. Leach, N. Beare, S. P. Harding, and Y. Zheng, "Automated detection of leakage in fluorescein angiography images with application to malarial retinopathy," *Sci. Rep.*, vol. 5, p. e10425, 2015.

[12] H. Rabbani, M. Allingham, M. P., S. Cousins, and S. Farsiu, "Fully automatic segmentation of fluorescein leakage in subjects with diabetic macular edema," *Invest. Ophthalmol. Vis. Sci.*, vol. 56, no. 3, pp. 1482–1492, 2015.

[13] L. Martinez-Costa, P. Marco, G. Ayala, E. D. Ves, J. Domingo, and A. Simo, "Macular edema computeraided evaluation in ocular vein occlusions," *Comput. Biomed. Res.*, vol. 31, pp. 374–384, 1998.

[14] R. Phillips, T. Spencer, P. Ross, P. Sharp, and J. Forrester, "Quantification of diabetic maculopathy by digital imaging of the fundus," *Eye*, vol. 5, pp. 130–137, 1991.

[15] J. Saito, D. Roxburgh, D. Sutton, and A. Ellingford, "A new method of image analysis of fluorescein angiography applied to age-related macular degeneration," *Eye*, vol. 9, pp. 70–76, 1995.

[16] E. Trucco, C. R. Buchanan, T. Aslam, and B. Dhillon, "Contextual detection of ischemic regions in ultra-wide-field-of-view retinal fluorescein angiograms," in *Proc. IEEE EMBC*, 2007, pp. 6739–6742.

[17] C. L. Tsai, C. L. Yang, S. J. Chen, C. H. Lin, K. S. Chan, and W. Y. Lin, "Automatic characterization of classic choroidal neovascularization by using adaboost for supervised learning," *Invest. Ophthalmol. Vis. Sci.*, vol. 52, pp. 2767–2774, 2011.

[18] Y. Yuan, J. Wang, B. Li, and M. Meng, "Saliency based ulcer detection for wireless capsule endoscopy diagnosis," *IEEE Trans. Med. Imag.*, Apr. 2015.

[19] D. Mahapatra and Y. Sun, "Mrf-based intensity invariant elastic registration of cardiac perfusion images using saliency information," *IEEE Trans. Biomed Eng.*, vol. 58, no. 4, pp. 991–1000, Apr. 2011.

[20] U. S. Deepak, A. Chakravarty, and J. Sivaswamy, "Visual saliency based bright lesion detection and discrimination in retinal images," in *Proc. IEEE ISBI*, 2013, pp. 1436–1439.

[21] X. Hou and L. Zhang, "Saliency detection: A spectral residual approach." in *Proc. IEEE CVPR*, 2007, pp. 1–8.

[22] V. Jampani, Ujjwal, J. Sivaswamy, and V. Vaidya, "Assessment of

computational visual attention models on medical images," in *Proc. Indian Conf. Comput. Vis. Graph. Imag. Process.*, 2012, pp. 1–8.
[23] D. Gao, V. Mahadevan, and N. Vasconcelos, "On the plausibility of the discriminant center-surround hypothesis for visual saliency," *J. Vis.*, vol. 8, no. 7, pp. 1–18, 2008.
[24] N. Bruce and J. Tsotsos, "Saliency based on information maximization." in *Proc. NIPS*, vol. 18, 2006, pp. 155–162.
[25] J. Harel, C. Koch, and P. Perona., "Graph-based visual saliency," in *Proc. NIPS*, 2006, pp. 545–552.
[26] M. Aziz and B. Mertsching, "Fast and robust generation of feature maps for region-based visual attention," *IEEE Trans. Image Process.*, vol. 17, pp. 633–644, 2008.
[27] J. Shen, Y. Du, W. Wang, and X. Li, "Lazy random walks for superpixel segmentation," *IEEE Trans. Image Process.*, vol. 23, no. 4, pp. 1451–1462, 2014.
[28] P. Martins, P. Carvalho, and C. Gatta, "Context-aware features and robust image representations," *J. Vis. Comun. Image Represent.*, vol. 25, pp. 339–348, 2014.
[29] Y. Zhao and Y. Liu, "Using region-based saliency for 3d interest points detection," in *Proc. CAIP*, 2013, pp. 108–115.
[30] C. Lang, T. Nguyen, H. Katti, K. Yadati, M. Kankanhalli, and S. Yan, "Depth matters: Influence of depth cues on spatial saliency," in *IEEE Trans. Pattern Anal. Mach.Intell.*, 2012, pp. 101–115.
[31] C. Li, C. Xu, C. Gui, and M. Fox, "Distance regularized level set evolution and its application to image segmentation," *IEEE Trans. Image Process.*, vol. 19, pp. 3243–3254, 2010.
[32] B. Ko and J. Nam, "Object-of-interest image segmentation based on human attention and semantic region clustering," *J. Opt. Soc. Am. A*, vol. 23, pp. 2462–2470, 2006.
[33] O. Daněk, P. Matula, M. Maška, and M. Kozubek, "Smooth Chan-Vese segmentation via graph cuts," *Pattern Recogn. Lett.*, vol. 33, pp. 1405–1410, 2012.
[34] Y. Zhao, Y. Liu, X. Wu, S. Harding, and Y. Zheng, "Retinal vessel segmentation: An efficient graph cut approach with retinex and local phase," *PLoS ONE*, vol. 10, p. e0122332., 2015.
[35] J. Shen, Y. Du, and X. Li, "Interactive segmentation using constrained laplacian optimization," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 7, pp. 1088–1100, 2014.
[36] X. Dong, J. Shen, and L. Shao, "Submarkov random walk for image segmentation," *IEEE Trans. on Image Process.*, vol. 25, no. 2, pp. 516–527, 2016.
[37] Y. Boykov and V. Kolmogorov, "Computing geodesics and minimal surfaces via graph cuts," in *Proc. 9th IEEE ICCV*, 2003, pp. 26–33.
[38] Y. Zhao, L. Rada, K. Chen, S. P. Harding, and Y. Zheng, "Automated vessel segmentation using infinite perimeter active contour model with hybrid region information with application to retinal images," *IEEE Trans. Med. Imag.*, vol. 34, pp. 1797–1807, 2015.
[39] Y. Zhao, I. J. MacCormick, D. Parry, S. Leach, N. Beare, S. P. Harding, and Y. Zheng, "Automated detection of vessel abnormalities on fluorescein angiogram in malarial retinopathy," *Sci. Rep.*, vol. 5, p. e11154, 2015.
[40] X. Hong, S. Chen, and C. Harris, "A kernel-based two-class classifier for imbalanced data sets," *IEEE Trans. Neural Netw.*, vol. 18, pp. 28–41, 2007.
[41] S. Goferman, L. Zelnik-Manor, and A. Tal, "Context-aware saliency detection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 34, no. 10, pp. 1915–1926, Oct. 2012.
[42] A. Borji, D. Sihite, and L. Itti, "Salient object detection: A benchmark," in *Proc. ECCV*, 2012, pp. 414–429.