A biologically inspired saliency model for color fundus images

Samrudhdhi B. Rangrej*
CVIT, IIIT-Hyderabad
Hyderabad, India.
rangrej.bharat@research.iiit.ac.in

Jayanthi Sivaswamy
CVIT, IIIT-Hyderabad
Hyderabad, India.
jsivaswamy@iiit.ac.in

ABSTRACT

Saliency computation is widely studied in computer vision but not in medical imaging. Existing computational saliency models have been developed for general (natural) images and hence may not be suitable for medical images. This is due to the variety of imaging modalities and the requirement of the models to capture not only normal but also deviations from normal anatomy. We present a biologically inspired model for colour fundus images and illustrate it for the case of diabetic retinopathy. The proposed model uses *spatially-varying* morphological operations to enhance lesions locally and combines an ensemble of results, of such operations, to generate the saliency map. The model is validated against an average Human Gaze map of 15 experts and found to have 10% higher recall (at 100% precision) than four leading saliency models proposed for natural images. The F-score for match with manual lesion markings by 5 experts was 0.4 (as opposed to 0.532 for gaze map) for our model and very poor for existing models. The model's utility is shown via a novel enhancement method which employs saliency to selectively enhance the abnormal regions and this was found to boost their contrast to noise ratio by $\sim 30\%$.

CCS Concepts

•Computing methodologies → Interest point and salient region detections; Modeling methodologies; •Applied computing → *Life and medical sciences;*

Keywords

Colour fundus image, diabetic retinopathy (DR), saliency, spatially-varying morphology, gaze map, selective-enhancement.

1. INTRODUCTION

Human attention is attracted by most prominent or visually salient objects in a scene. Saliency of an object is modulated by the task at hand [25] and we selectively attend to most informative regions of visual field while ignoring some unimportant regions. Much effort has been made to understand task-specific visual attention and perception, resulting in cognitive modeling of visual saliency [24]. Computational modeling of the same has been of interest to computer vision community for years [4] leading to their use in segmentation [1], object recognition [9], retrieval [23], image/video compression [10] and context aware image editing [11]. Many such applications are of interest in medical domain as well.

Medical experts look for specific type of abnormalities often at specific locations, ignoring irrelevant areas and artifacts, while reviewing images for diagnosis. Interest in medical image perception is on the rise partly to identify best practices for accurate diagnosis. Cognitive and psychophysical studies have attempted to understand how medical experts inspect images via eye tracking [18, 7]. Deriving saliency models for medical images is a difficult task due to the variability in modalities, anatomy and type of disease. Very little work has been done to devise computational models of saliency for medical images [15, 3, 27]. Saliency has been used in x-ray image classification[6], segmentation and registration of MRI [20, 21] and anatomical plane classification from fetal ultrasound [19].

Computational models of visual attention are broadly classified into 2 groups: bottom-up and top-down [4]. Bottom-up models are stimuli driven, whereas top-down models are intention, task or goal driven and based on prior knowledge. Depending on the medical modality, general bottom-up models may or may not be successful in explaining how experts review images. Existing top-down models for general images are inappropriate as the tasks are different. Saliency models need to be designed for specific modalities and type of lesions.

We propose a saliency model for colour fundus images of eyes and showcase it for diabetic retinopathy (DR). Our model handles normal and abnormal images with bright lesions (hard exudates, cotton wool spots) and dark lesions (hemorrhages, microaneurysms). Potential applications of this work are: (a) Selective enhancement of salient regions (discussed in section 4), (b) Tools which assist/train readers (in image reading centres) and clinical residents. Here, saliency value of an image region can be mapped to a confidence measure about the presence of abnormality. (c) Fully automatic solution development for lesion detection, with saliency aiding candidate region detection.

Existing computational saliency models include ones which are biologically plausible [14] or based on spectral-analysis

*corresponding author.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

ICVGIP, December 18 - 22, 2016, Guwahati, India

© 2016 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ISBN 978-1-4503-4753-2/16/12...$15.00

DOI: http://dx.doi.org/10.1145/3009977.3010041

[13], information- and decision-theory [5], pattern classification [16], graphs [12], etc. The proposed model is biologically motivated and is based on morphological processing. The model has been evaluated against *gaze maps* of retina specialists as well as manual lesion markings and compared with four existing bottom up saliency models developed for natural images. The model is showcased in a novel selective enhancement application.

2. METHOD

Human fixations follow a Gaussian distribution [26]. Visual information near a gaze-point is attended to more than those that are away. In terms of visual processing, this implies that information in a region proximal to a gaze-point is given higher importance than that are not. The inspection and understanding of the entire visual scene is done gradually as we move our eyes. We use this idea in developing the proposed model.

We propose an approach to saliency computation for a specific task, namely to detect abnormalities (bright and dark lesions) in a given image, by boosting their prominence at a local level. Let us consider a gaze-point p in the image. Our strategy is to boost the prominence of lesions or abnormalities in the image based on their proximity to p. Specifically, the boosting is higher for proximal as opposed to distant lesions. The boosting is achieved using *spatially varying* morphological processing. The final saliency map is derived by using a set of gaze-points p_i and combining the boosted results at these points.

The processing pipeline has 3 stages: (1) Preprocessing (2) Generation of an ensemble of pre-saliency maps (3) ensemble integration to produce the final saliency map. The details are presented next.

2.1 Preprocessing

Given a colour fundus image, all processing was restricted to the green channel. The vessel network and the optic-disk were detected and inpainted. The fundus was extended to cover the mask region.

2.2 Pre-saliency map generation

Given an image I and a gaze-point $p = (a, b)$, the origin is shifted to p. The resulting image $I_s\,(x, y) = I(x - a, y - b)$ which in polar coordinates, is denoted as $I_s(r, \theta)$. For any given $\theta \in (-\pi, \pi]$, we denote the 1-D image $I_\theta(r) = I_s(r, \theta)$. Consider a structuring element,

$$b(\rho) = \begin{cases} 0, & \rho \in D_b \\ -\infty, & \text{otherwise} \end{cases} \quad (1)$$

where, D_b is domain of structuring element b. Grayscale dilation and erosion of I_θ with b is expressed respectively as,

$$(I_\theta \oplus b)(r) = max\{I_\theta(r - \rho) \mid \rho \in D_b\} \quad (2)$$
$$(I_\theta \ominus b)(r) = min\{I_\theta(r + \rho) \mid \rho \in D_b\} \quad (3)$$

In order to introduce spatially-varying processing, the domain D_b is made to vary with r and hence the dilation/erosion is defined as follows,

$$D_b(r_n) \equiv [-f(r_n), f(r_n)] \quad (4)$$
$$(I_\theta \oplus b)(r_n) = max\{I_\theta(r_n - \rho) \mid \rho \in D_b(r_n)\} \quad (5)$$
$$(I_\theta \ominus b)(r_n) = min\{I_\theta(r_n + \rho) \mid \rho \in D_b(r_n)\} \quad (6)$$

We have chosen $f(r) = \lambda G_\sigma(r)$, where $G_\sigma(r)$ is a Gaussian distribution with zero mean and variance σ^2. λ is a free parameter which controls the domain length. Eq. 5 and 6 can also be interpreted as rank filtering with a filter/window whose length is a function of r_n. At any point r_n, a window centred at r_n and of length $2\lambda G_\sigma(r_n)$ is used to do the ranking operation. The above $1D$ operation (dilation/erosion) is performed on $I_\theta(r)$, $\forall\theta \in (-\pi, \pi]$. i.e. in each direction. Fig.1 shows the windows of varying length along different directions about a gaze point $P : (r, \theta)$. Windows for $I_{\theta=0}(r)$ are shown to be derived from Gaussian distribution.

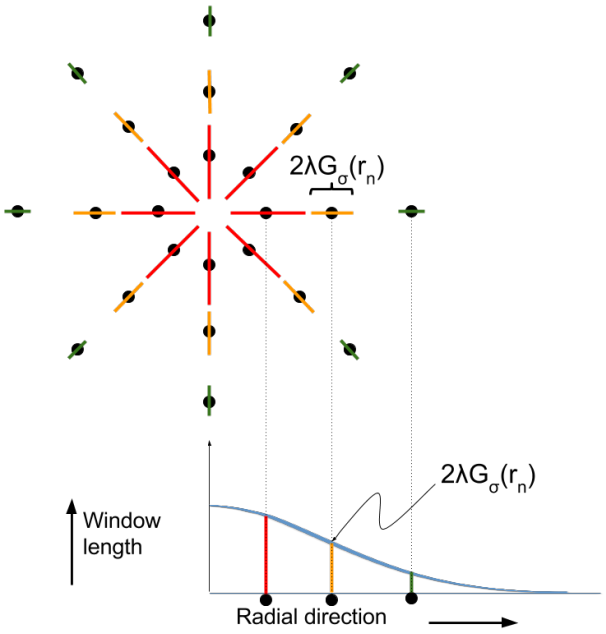


Figure 1: Window length at different points in the image.

Processed images are shifted back to get the final pre-saliency (PS) maps: $\text{PS}_q^i(x, y) = I_q(x + a_i, y + b_i)$ where q represents dilation or erosion operation. Dilation (erosion) will boost the saliency of bright (dark) lesions.

We illustrate this idea with a phantom in Fig.2. Here, an image patch is modeled as gray-scale texture with idealized lesions appearing as dots of appropriate colour: white dot (hard exudate or HE), white blurred spot (cotton wool spot or CWS) and dark dot (hemorrhage or HM) (Fig.2a). Four gaze-points p_i are selected on the three lesions and background (Fig.2b). The PS maps obtained with different p_i are dramatically different as seen in Fig.2c-j. Overall, it can be seen that a lesion is spatially extended while the background remains unchanged, in the PS map. It is noteworthy that when p is on the background (shown in yellow), both dilation and erosion has no effect since the lesions are not proximal enough to p in any direction. Hence, the original image and PS are almost identical in appearance. Similar behaviour is shown in real images next.

A sample fundus image patch with a HE and HM is shown in Fig.3 along with the PS maps for dilation/erosion derived with three p_j (shown in green, blue and black). In Fig.2 and 3, the p_j were selected on true lesions deliber-

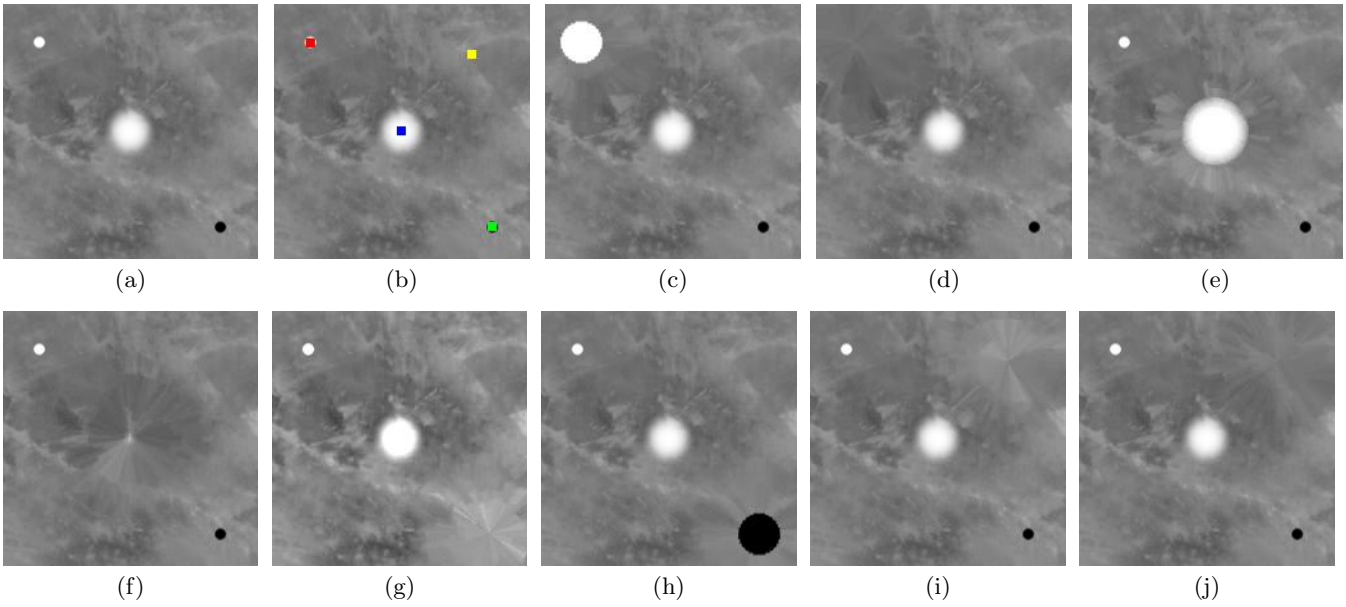


Figure 2: Pre-saliency maps for a phantom image (a) at four gaze-points considered (b). The maps with dilation - erosion for the red, blue, green and yellow gaze-points are shown in (c)-(d),(e)-(f),(g)-(h) and (i)-(j) respectively.

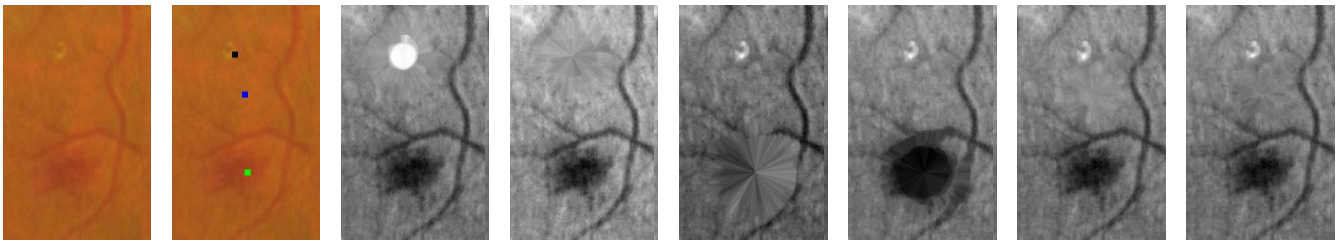


Figure 3: Pre-saliency maps for a fundus image patch (a) at three gaze-points (b). The maps with dilation - erosion for the black, green and blue gaze-points are shown in (c)-(d);(e)-(f) and (g)-(h) respectively.

ately to show the effect of the proposed spatially-varying morphological processing. In reality, the lesion locations are unknown. Hence, a set of $\{p_j \mid j = 1, 2, 3, ..., J\}$ at randomly selected locations are used to generate an ensemble of PS maps. Since both bright and dark lesions are of interest, both erosion and dilation are applied separately at each p_j, to obtain $2J$ PS maps. A judicious combination of these maps can help derive the desired saliency map, which is explained next.

2.3 Integration

In order to combine the J PS maps, we follow the strategy used in [22]. The J maps are summed to create a combined map C_q. The variance at every pixel location is also computed to derive a variance map V_q.

$$\mathrm{C}_q = \sum_{i=1}^{J} \mathrm{PS}_q^i \tag{7}$$

$$\mathrm{V}_q = \mathrm{Variance}(\mathrm{PS}_q^i); i \in [1, 2, 3, ..., J] \tag{8}$$

These two maps provide evidence for a location to be salient. An explicit Evidence map (E_q) is computed by exponential weighting of C_q by V_q as,

$$\mathrm{E}_q = \mathrm{C}_q \times e^{(\tau \times \mathrm{V}_q)}, \tau \in \mathbb{R} \tag{9}$$

Here, τ helps control the contribution of the variance at a pixel to the final evidence. A negative (positive) value of τ is chosen for bright (dark) lesion. Higher absolute value of τ boosts the saliency of even less prominent lesions. Separate E_qs, one for dilation and other for erosion, are extracted.

2.3.1 *Saliency map computation*

A Center surround (CS) filter is applied to handle the variable conspicuity of lesions. This is applied to both E_q separately and they are combined to get the Saliency map (S) as,

$$S = max\{\mathrm{E}_{dilation}^{cs}, \mathrm{E}_{erosion}^{cs}\} \tag{10}$$

where, E_q^{cs} is CS filtered E_q and $max\{.\}$ is a pixel-wise max operation. S is finally smoothed with a Gaussian filter. The proposed pipeline is shown in Fig.4 with all intermediate results.

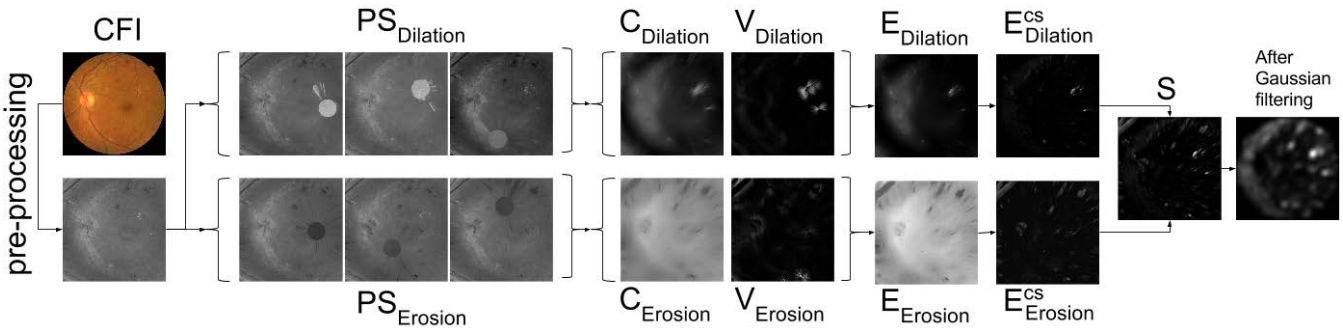


Figure 4: Proposed model for computing saliency with intermediate results.

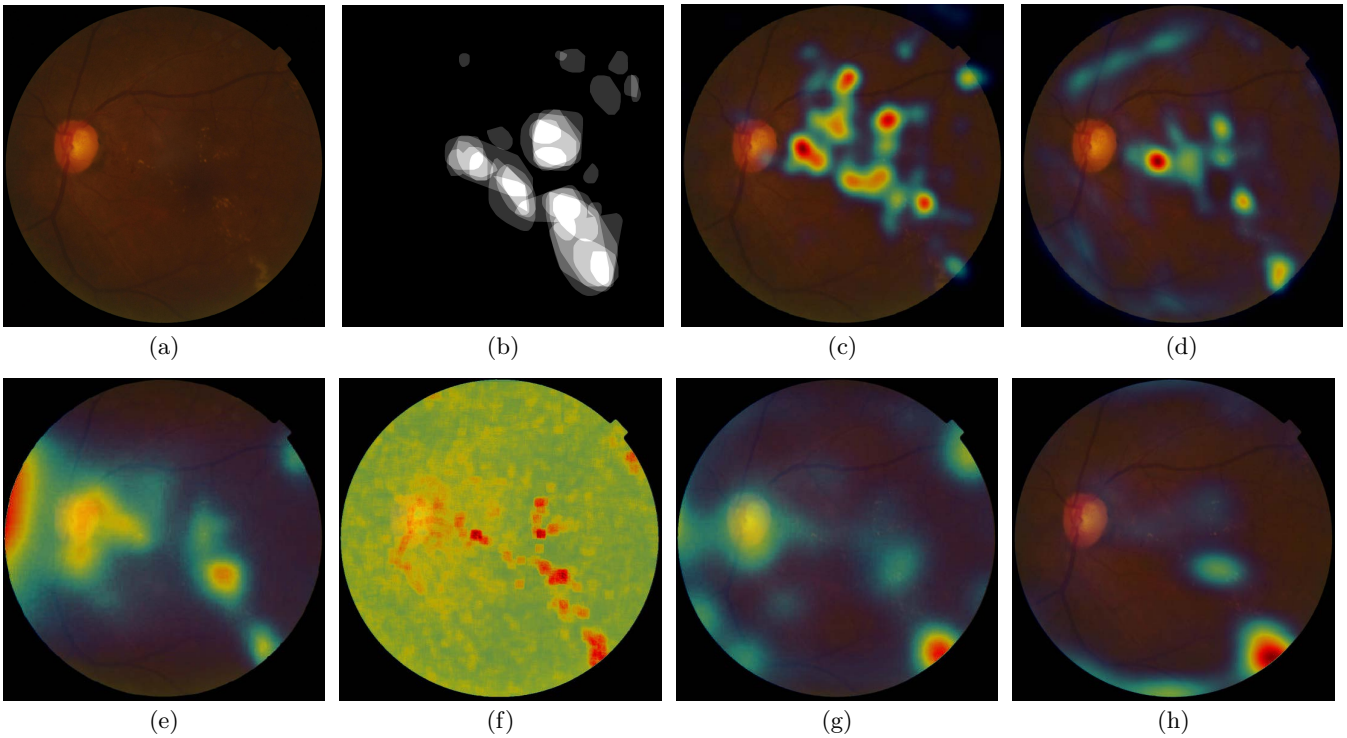


Figure 5: Comparison of saliency models (a) Sample fundus image (b) Markings from 5 experts (c) GM map(d) SED (Our model) (e) GBVS (f) AIM (g) IK (h) SR

3. RESULTS

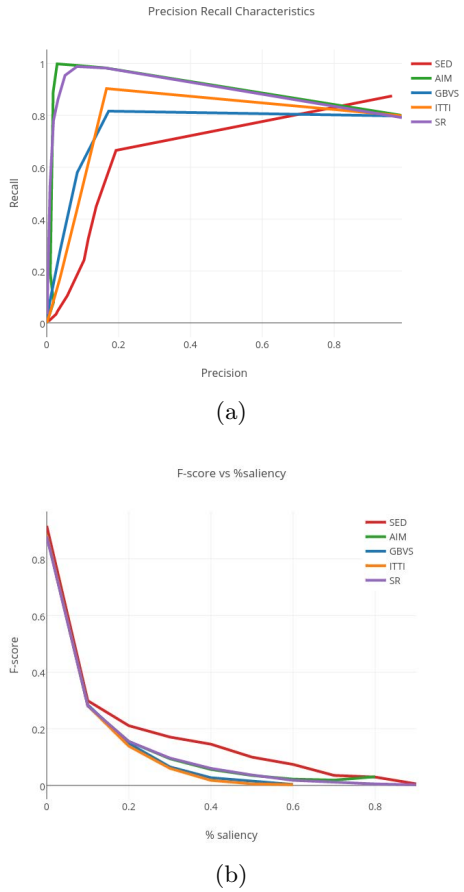
3.1 Evaluation for abnormal cases

The proposed Spatially-varying Erosion and Dilation (or SED) model for saliency was tested on images collected from a local eye hospital. Along with the images, manual marking (GT) of lesion regions was collected from 5 retina experts. The model was evaluated in two ways: i) against gaze maps derived from eye tracking and ii) GT. The proposed model was also benchmarked against 4 different bottom up models of saliency. The preprocessed images (section 2.1) were taken as input for all the saliency models for a fair comparison. The models taken for comparison were based on different approaches: SR [13], Itti-Koch (IK) [14], GBVS [12] and AIM [5]. Of these models, only the AIM model is based on learning (from patches of a large number of *natural* images).

3.1.1 Evaluation against Gaze-map

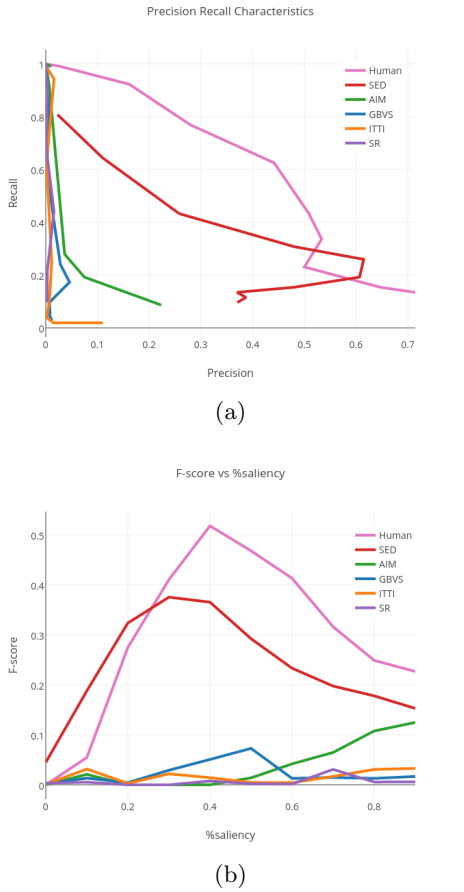
We begin with the evaluation against average human gaze map (GM). An eye tracking experiment was performed on 15 retina experts while they were reviewing the images exhaustively. To avoid human fatigue and hence low accuracy of eye tracking data, dataset size was limited to 10 images, all containing DR lesions. GMs of all experts were denoised and averaged to derive an average GM for each image.

A sample image and its softmap GT are shown in Fig.5a-b. The GM and the computed saliency maps are presented in Fig.5c-h. SED and AIM appear to be similar to both GT as well as GM, though there are some false positive saliency regions as well. The saliency maps from GBVS, IK and SR



(a) (b)

Figure 6: Comparative performance of various saliency models against gaze maps (a) Precision-Recall characteristics (b) F-score for saliency models.



(a) (b)

Figure 7: Comparative performance of various saliency models against ground truth (a) Precision-Recall characteristics (b) F-score for saliency models.

are sparse and do not have as much overlap with the GT and GM. A quantitative analysis was also done to compare the performances. The results are presented next.

First, we compared the saliency maps (S) with GMs. Precision (P) and recall (R) were computed by thresholding S in the 0-90% range in steps of 10%. The resultant plots are presented in Fig.6a. It can be observed that for $P < 0.2$ the R values of existing models show an increasing trend whereas for $P > 0.2$ the trend is a decreasing one. In contrast, SED shows an increasing trend for all $R \in [0, 1]$. At $P = 1$, SED outperforms the existing methods by 10%. This relative improvement is also seen in F-score which is presented in Fig.6b.

3.1.2 Evaluation against Ground Truth

In the medical domain, unlike general vision, high level (or top down) knowledge is used during image scrutiny as the end goal is diagnosis. It is quite possible that overt attention is not used in determining if a region has abnormalities or not. Hence, it is of interest to determine the degree of overlap between the GM and GT maps. Further, since a saliency map can have multiple applications (see section 1), we also evaluated all the models against GT. The GT softmap was thresholded at 50% agreement to generate a binary map. Again, all saliency maps and GMs were thresholded from 0-90% of saliency in steps of 10% to generate the PR plot and F-score vs %saliency plot presented in Fig.7. The plots reveal that SED and GM outperform all saliency models. This underscores the fact that a saliency model designed for natural images is not appropriate for medical images especially in helping to detect abnormalities.

3.2 Evaluation for normal cases

The saliency maps were also computed for images without any abnormalities i.e. *normal* images. These are shown for a sample image for all models in Fig. 8. The GM was collected separately for normal images and the one for Fig. 8a is shown in Fig. 8b. This reveals that experts scrutinize even normal cases with a set of fixations before making a diagnosis. The GM also indicates the macula to be the only region with significant foveation. This is to be expected as macula is responsible for sharp colour vision and is hence a danger zone; any presence of abnormality here calls for swift intervention. It can be seen that only SED is able to reject normal regions and generate a very sparse saliency map as desirable. The scrutiny of normal images is guided by

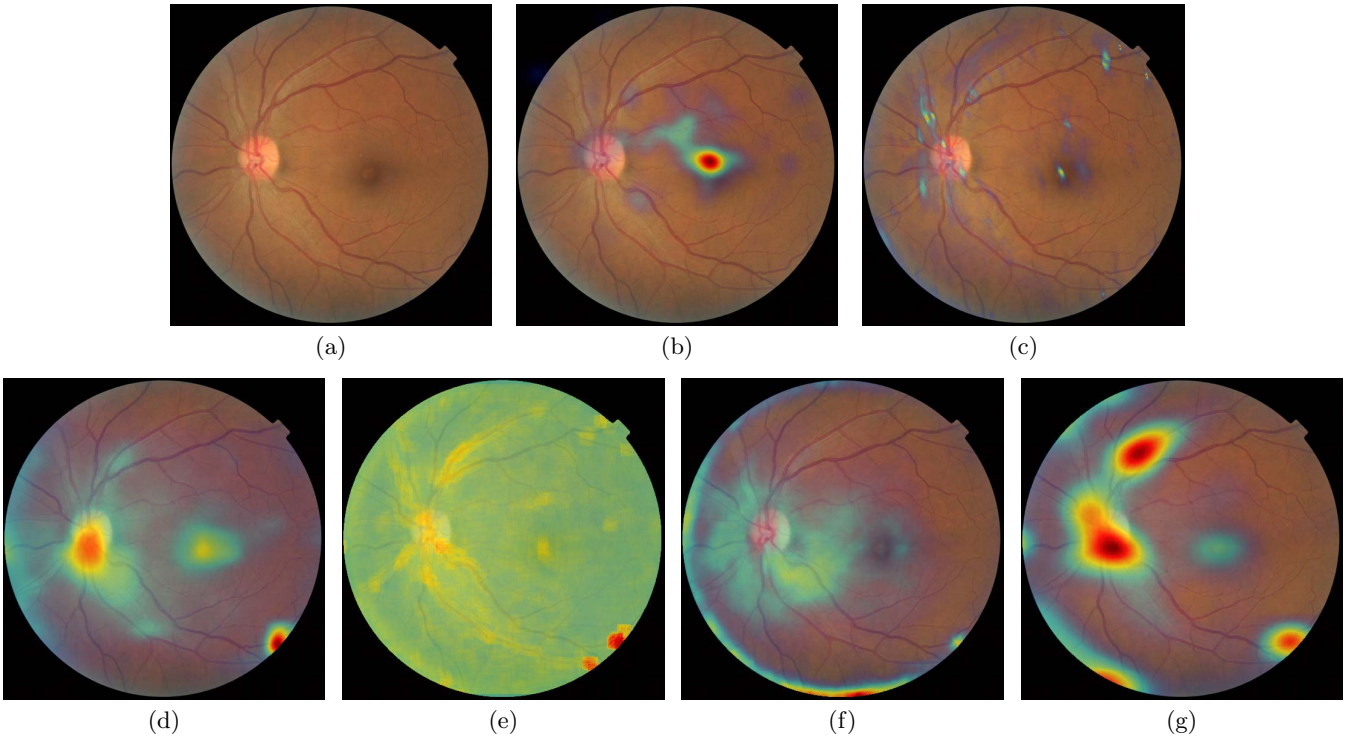


Figure 8: Comparison of saliency models for normal fundus image. (a) Original fundus image (b) Average gaze map (c) SED (Our model) (d) GBVS (e) AIM (f) IK (g) SR

more complex knowledge about the normal anatomy, danger zones and an expert's scrutiny style. Hence, a quantitative analysis is not appropriate for such cases.

4. APPLICATION: SELECTIVE ENHANCEMENT

Saliency can be employed in many applications as mentioned in section 1. We next consider the problem of selective enhancement of color fundus images in greater depth. Existing enhancement techniques include single channel(mostly green channel) enhancement or color enhancement which uses global information. Such methods alter perceptual quality of an image [8]. Such enhanced images may be useful for computational purposes but unsuitable for visual presentation in real word scenario where a human image reader is diagnosing images. Fully-automatic enhancement also does not allow a reader to apply minor adjustment as is often desirable. For example, radiologists routinely vary the window-center and window-length to adjust brightness and contrast of CT images for better visibility [2].

We propose a semi-automatic solution with Interactive Selective Enhancement (ISE). ISE enhances salient regions locally, with minimum alteration to perceptual quality. ISE provides interactive control over degree of enhancement i.e. reader can vary parameters to improve results generated by default parameter settings and can observe changes in real-time. ISE however, does not aim to correct non-uniform illumination or blur. Details are provided next.

Given an image I and evidence maps $E^{cs}_{dilation}$ and $E^{cs}_{erosion}$, enhanced image X is computed over multiple scales and fused as follows,

$$X_k = \sum_i \Big\{ [(1-\alpha)I_k - \alpha((a_k\, E^{cs}_{erosion} \times I_k) * G_i)] + [(1-\beta)I_k + \beta((b_k\, E^{cs}_{dilation} \times I_k) * G_i)] \Big\} \quad (11)$$

Here, the first and second terms help enhance dark and bright lesions respectively. k is index of color channel and i is scale. a and b control color shade in the enhanced image. α and β are mixing parameters which control the degree of enhancement. α, β, a, b are control parameters which can be varied by the reader. $E^{cs}_{erosion}$(or $E^{cs}_{dilation}$) masks out all the background and picks only dark (or bright) lesions from I, which is then mixed with I over multiple scales. Negative sign in the first term effects a darkening of the salient regions. Similarly, positive sign in the second term has the effect of brightening the salient regions.

The publicly available DiaretDB1[17] dataset is chosen for validation as it contains both bright and dark lesions and also provides lesion-level GT. We use 47 abnormal images out of given 89 images for the evaluation. A set of experiments are reported here for different settings of mixing parameters. a and b are fixed for all the experiments. First, mixing parameter α and β are varied one at a time while the fixed parameter is set to zero. Varying values of α with $\beta = 0$ results in enhancement of only dark lesions. Similarly, varying values of β with $\alpha = 0$ results in enhancement of only bright lesions. Figures 9 and 10 show qualitative results of IES for dark and bright lesions respectively. An excessive enhancement of dark (bright) lesions is seen to relatively brighten (darken) the background. If, both α and

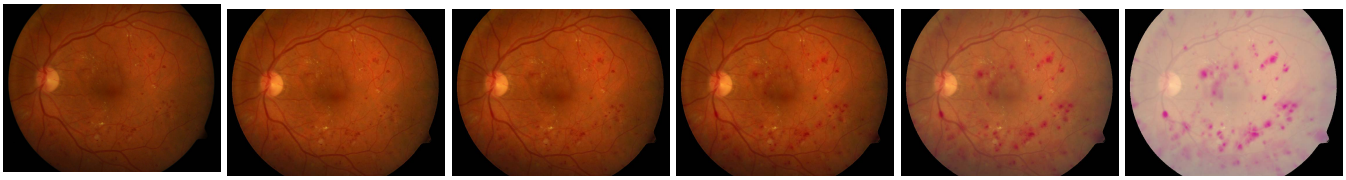


Figure 9: Selective enhancement of dark lesions for different α ($\beta = 0$).

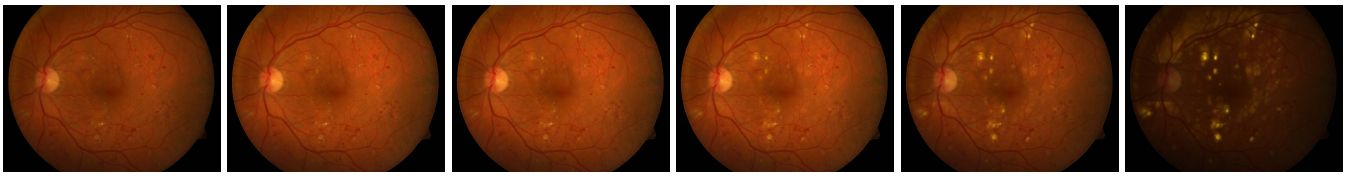


Figure 10: Selective enhancement of bright lesions for different β ($\alpha = 0$).

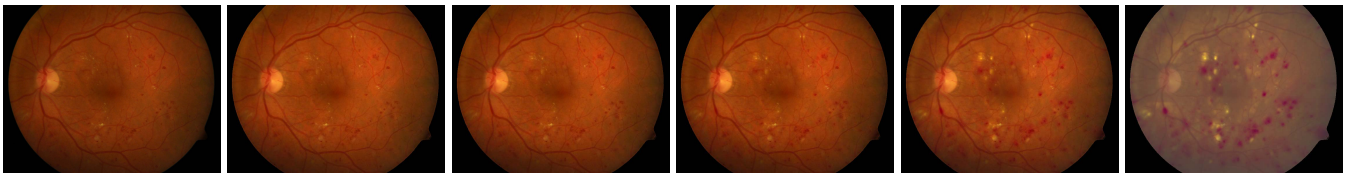


Figure 11: Balanced enhancement of both bright and dark lesions for varying values of α and β.

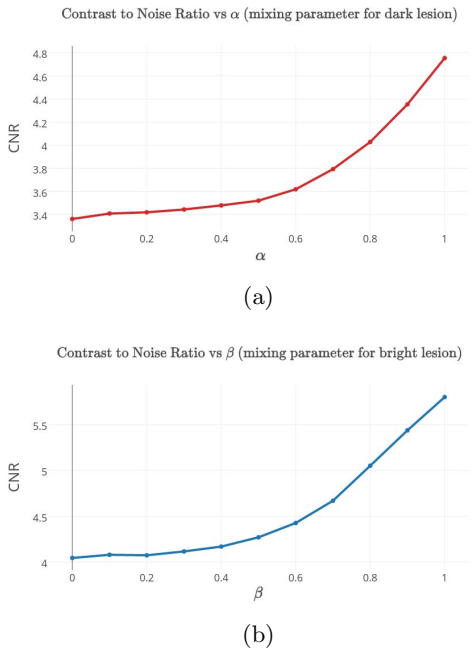


Figure 12: Contrast to noise ratio as a function of mixing parameter: (a) dark lesions (b) bright lesions.

β are varied simultaneously, this results in a balanced enhancement of both dark and bright lesions. Fig. 11 shows this on a sample image.

A quantitative evaluation of IES was performed by computing the contrast to noise ratio (CNR) for varying values of mixing parameters. CNR is defined as

$$CNR = \frac{|m_f - m_b|}{\sigma_b} \tag{12}$$

where, m_f and m_b are mean intensity of foreground (lesions in our case) and background respectively. σ_b is the standard deviation of background intensity. Fig.12 shows the relation between CNR and mixing parameter which is strictly increasing. Varying mixing parameter from 0 to 1 increases CNR by $\sim 30\%$ for both dark and bright lesions (from 4.04 to 5.8 for bright lesions and from 3.36 to 4.75 for dark lesions).

Figures 9, 10 and 11 suggest that certain range of α and β give the best perceptual quality. It is possible to dynamically compute the optimum parameters of ISE for a given image and use it as default settings subject to minor adjustments by reader. Optimized ISE can also be used for computational purposes and is beyond the scope of this paper.

5. CONCLUSION

Saliency computation has not received much attention in medical imaging. We presented a saliency model which is biologically inspired and is based on combining results of

spatially-varying morphological processing centred at randomly chosen gaze points. The poor performance of saliency models such AIM, GBVS etc., in contrast to the good performance of our model in predicting gaze points and identifying abnormal regions, underscore the point that generic saliency models are unsuitable for the medical domain. As opposed to the strategy we used where the gaze points were chosen randomly, a *guided* spatial arrangement of gaze-points would be an interesting variant worth exploring. Saliency maps can be used for many applications. We presented one which aids interactive and selective enhancement of regions of abnormalities. This offers an efficient and highly interactive environment for image readers in screening centres and also aid candidate selection in automatic processing.

6. ACKNOWLEDGEMENT

This work was supported by the Dept. of Electronics and Information Technology, Govt. of India under Grant: $DeitY/R\&D/TDC/13(8)/2013$.

7. REFERENCES

[1] R. Achanta, F. Estrada, P. Wils, and S. Süsstrunk. Salient region detection and segmentation. In *International conference on computer vision systems*, pages 66–75. Springer, 2008.

[2] N. E. M. Association et al. Digital imaging and communications in medicine (dicom), part 14: Gray-scale standard display function. 2001.

[3] S. Banerjee, S. Mitra, B. U. Shankar, and Y. Hayashi. A novel gbm saliency detection model using multi-channel mri. *PloS one*, 11(1):e0146388, 2016.

[4] A. Borji and L. Itti. State-of-the-art in visual attention modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(1):185–207, 2013.

[5] N. D. Bruce and J. K. Tsotsos. Saliency, attention, and visual search: An information theoretic approach. *Journal of vision*, 9(3):5, 2009.

[6] Z. Camlica, H. Tizhoosh, and F. Khalvati. Medical image classification via svm using lbp features from saliency-based folded data. *arXiv preprint arXiv:1509.04619*, 2015.

[7] L. Cooper, A. Gale, I. Darker, A. Toms, and J. Saada. Radiology image perception and observer performance: How does expertise and clinical information alter interpretation? stroke detection explored through eye-tracking. In *SPIE Medical Imaging*, pages 72630K–72630K. International Society for Optics and Photonics, 2009.

[8] M. J. Cree, E. Gamble, and D. Cornforth. Colour normalisation to reduce inter-patient and intra-patient variability in microaneurysm detection in colour retinal images. 2005.

[9] D. Gao and N. Vasconcelos. Discriminant saliency for visual recognition from cluttered scenes. In *Advances in neural information processing systems*, pages 481–488, 2004.

[10] C. Guo and L. Zhang. A novel multiresolution spatiotemporal saliency detection model and its applications in image and video compression. *IEEE TIP*, 19(1):185–198, 2010.

[11] A. Hagiwara, A. Sugimoto, and K. Kawamoto. Saliency-based image editing for guiding visual attention. In *Proceedings of the international workshop on pervasive eye tracking & mobile eye-based interaction*, pages 43–48. ACM, 2011.

[12] J. Harel, C. Koch, and P. Perona. Graph-based visual saliency. In *Advances in neural information processing systems*, pages 545–552, 2006.

[13] X. Hou and L. Zhang. Saliency detection: A spectral residual approach. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8. IEEE, 2007.

[14] L. Itti and C. Koch. Computational modelling of visual attention. *Nature reviews neuroscience*, 2(3):194–203, 2001.

[15] V. Jampani, J. Sivaswamy, V. Vaidya, et al. Assessment of computational visual attention models on medical images. In *ICVGIP*, page 80. ACM, 2012.

[16] T. Judd, K. Ehinger, F. Durand, and A. Torralba. Learning to predict where humans look. In *IEEE International Conference on Computer Vision*, pages 2106–2113. IEEE, 2009.

[17] T. Kauppi, V. Kalesnykiene, J.-K. Kamarainen, L. Lensu, I. Sorri, A. Raninen, R. Voutilainen, H. Uusitalo, H. Kälviäinen, and J. Pietilä. The diaretdb1 diabetic retinopathy database and evaluation protocol. In *BMVC*, pages 1–10, 2007.

[18] E. A. Krupinski. Current perspectives in medical image perception. *Attention, Perception, & Psychophysics*, 72(5):1205–1217, 2010.

[19] A. Kumar, P. Sridar, A. Quinton, R. K. Kumar, D. Feng, R. Nanan, and J. Kim. Plane identification in fetal ultrasound images using saliency maps and convolutional neural networks. In *ISBI*, pages 791–794. IEEE, 2016.

[20] D. Mahapatra and J. M. Buhmann. Visual saliency-based active learning for prostate magnetic resonance imaging segmentation. *Journal of Medical Imaging*, 3(1):014003–014003, 2016.

[21] D. Mahapatra and Y. Sun. Registration of dynamic renal mr images using neurobiological model of saliency. In *ISBI*, pages 1119–1122. IEEE, 2008.

[22] A. Ujjwal, J. Sivaswamy, et al. An assistive annotation system for retinal images. In *IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 1506–1509. IEEE, 2015.

[23] W. Wang, Y. Song, and A. Zhang. Semantics-based image retrieval by region saliency. In *International Conference on Image and Video Retrieval*, pages 29–37. Springer, 2002.

[24] J. M. Wolfe, K. R. Cave, and S. L. Franzel. Guided search: an alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human perception and performance*, 15(3):419, 1989.

[25] A. L. Yarbus. *Eye movements during perception of complex objects*. Springer, 1967.

[26] Q. Zhao and C. Koch. Learning a saliency map using fixated locations in natural scenes. *Journal of vision*, 11(3):9, 2011.

[27] X. Zou, X. Zhao, Y. Yang, and N. Li. Learning-based visual saliency model for detecting diabetic macular edema in retinal image. *Computational intelligence and neuroscience*, 2016:1, 2016.