Fundus Lesion Detection Based on Visual Attention Model

Baisheng Dai[1], Wei Bu[2], Kuanquan Wang[1], and Xiangqian Wu[1(✉)]

[1] School of Computer Science and Technology,
Harbin Institute of Technology, Harbin 150001, China
{bsdai,wangkq,xqwu}@hit.edu.cn
[2] Department of New Media Technologies and Arts,
Harbin Institute of Technology, Harbin 150001, China
buwei@hit.edu.cn

Abstract. Reliable detection of fundus lesion is important for automated screening of diabetic retinopathy. This paper presents a novel method to detect the fundus lesion in retinal fundus image based on a visual attention model. The proposed method intends to model the visual attention mechanism of ophthalmologists during observing fundus images. That is, the abnormal structures, such as the dark and bright lesions in the image, usually attract the most attention of experts, however, the normal structures, such as optic disc and vessels, have been usually selectively ignored. To measure the visual attention for abnormal and normal areas, the incremental coding length is computed in local and global manner respectively. The final saliency map of fundus lesion is a fusion of attention maps computed for the abnormal and normal areas. Experimental results conducted on the publicly DiaRetDB1 dataset show that the proposed method achieved a sensitivity of 0.71 at a specificity of 0.82 and an AUC of 0.76 for fundus lesion detection, and achieved an accuracy of 100 % for normal area (optic disc) detection. The proposed method can assist the ophthalmologists in the inspection of fundus lesion.

Keywords: Diabetic retinopathy · Fundus lesion detection · Visual attention · Incremental coding length

1 Introduction

Diabetic Retinopathy (DR) is one of the main causes of blindness, and it is usually asymptomatic until the disease is at a late stage. The early detection of DR is thus important for patients to prevent visual loss. In fundus image, the most common signs of DR are the dark lesions, such as microaneurysms and hemorrhages, and the bright lesions, such as exudates and cotton-wool spots. The existence of these lesions can reflect the severity of DR. Currently, the inspection of these lesions is usually performed with the naked eye of ophthalmologists,

W. Che et al. (Eds.): ICYCSEE 2016, Part I, CCIS 623, pp. 384–394, 2016.
DOI: 10.1007/978-981-10-2053-7_34

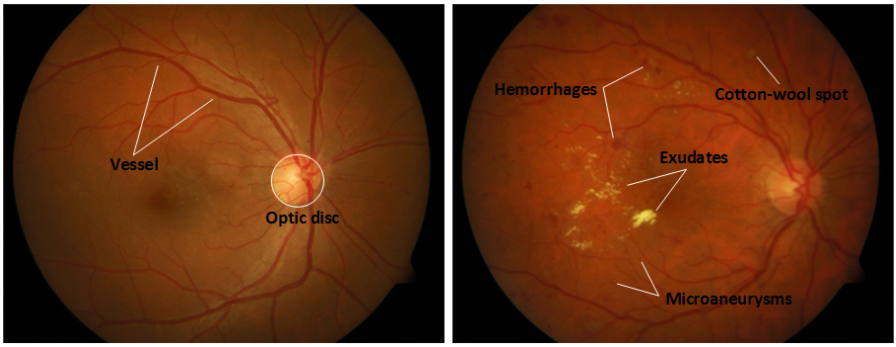


Fig. 1. Examples of retinal fundus images. (a) A normal fundus image. (b) A fundus image contains lesions.

such manual screening of DR, however, is time-consuming, subjective and error-prone. To diagnosis DR timely and reliably, the automated screening of DR is extensively investigated in past decades.

The first step of automated DR screening is the detection of fundus lesion, and because of the variability in appearance of those lesions, different algorithms have been designed to detect each type of those lesions separately. For microaneurysms detection, such as multi-scale correlation filtering [1] and diameter closing [2] have been presented. For hemorrhages detection, a splat feature classification approach has been reported in [3]. For the exudates detection, such as dynamic thresholding [4] and clustering based approach [5] have been used to detect the lesion. And with respect to the detection of cotton-wool spots, an improved Fuzzy C-Means approach has been presented in [6].

In fact, during clinically examining of the fundus image, all of those lesions will attract the most visual attention of experts, since they have irregular patterns compared with other areas. On the other hand, the normal structure such as optic disc will be selectively ignored since they are the locations that have been recently attended by the visual attention of many normal retinal images, which is so-called inhibition of return (IoR) [7]. Figure 1 shows two examples of retinal fundus images, for the task of DR screening, the lesions in Fig. 1(b) will attract the most attention of the experts.

Inspired by the biological vision mechanism, in this paper, a novel method for fundus lesion detection is presented based on a visual attention model, which intends to detect both of the dark and bright lesions in one unified framework.

The proposed method measures the "irregularity" of abnormal areas by incremental coding length (ICL) [8,9] in a local manner, and models the inhibition of return of normal areas also with ICL but in a global manner.

The rest of this paper is structured as follows. Section 2 describes our method to detect fundus lesion based on visual attention model. Experimental results are presented in Sect. 3. Section 4 concludes the paper.

2 Methodology

2.1 Abnormality Attention Computation

Here we assume that visual attention is driven by the predictive coding principle, i.e., the optimization of metabolic energy consumption in the brain [10]. The area attracted more attention of the visual system, i.e., the abnormal structure, is regarded as the location that need a more expensive neural code to be represented by its surrounding areas.

In this work, we use the incremental coding length (ICL) [8,9] to measure which area is more abnormal. Given an input image I, and p is a patch of image I. Let $\mathcal{N}_p$ be the spatial surrounding patches of p, which are captured in an overlapping manner. If the patches of image I are encoded with a lossy coding method with fixed distortion ε, $L_\varepsilon(\mathcal{N}_p)$ and $L_\varepsilon(\mathcal{N}_p \cup p)$ denotes the coding length of patches $\mathcal{N}_p$ and $\mathcal{N}_p \cup p$ respectively. The incremental coding length of patch p is then defined as [8,9]:

$$\delta L_\varepsilon(p) = L_\varepsilon(\mathcal{N}_p \cup p) - L_\varepsilon(\mathcal{N}_p) \tag{1}$$

where $\delta L_\varepsilon(p)$ is larger, the patch p is more abnormal than its surrounding areas.

To compute $\delta L_\varepsilon(p)$, we use the same scheme as suggested in [9], i.e.,

$$\delta L_\varepsilon(p) \approx L_\varepsilon(p|\mathcal{N}_p) \tag{2}$$

where $p|\mathcal{N}_p$ is the sparse representation of patch p over its surrounding patches $\mathcal{N}_p$.

Let $\mathbf{x} \in \mathbb{R}^n$ denotes the feature vector of the patch p, and $\mathbf{D} \in \mathbb{R}^{n \times m}$ is the dictionary matrix whose columns are the feature vectors of surrounding patches $\mathcal{N}_p$, as shown in Fig. 2. Finding the spare representation of p is to seek the optimal sparse code vector $\alpha \in \mathbb{R}^m$, which solves the following optimization problem:

$$\min_{\alpha} \frac{1}{2} \|\mathbf{x} - \mathbf{D}\alpha\|_2^2 + \lambda \|\alpha\|_1 \tag{3}$$

where the first term is the square reconstruction error, the second term is a ℓ_1 sparsity penalty on the code and λ is a coefficient controlling the sparsity penalty.

The incremental coding length $\delta L_\varepsilon(p)$ now is proportional to the number of non-zero entries of α, and in this work, it is computed as $\|\alpha\|_1$. The larger that is, the more abnormal the patch p will be. The final map of abnormality attention is then generated by accumulating $\delta L_\varepsilon(p)$ per pixel. Figure 4 illustrates examples of the map of abnormality attention, as shown in this figure, both dark and bright lesions have the higher response of the abnormality attention. However, the normal area, especially like the optic disc, also has higher response since it appears as irregular pattern in a single image. To suppress the unexpected high response in those normal areas, we model the IoR in following section.

2.2 Inhibition of Return Modelling

In the clinical task of DR screening, the normal areas, such as optic disc (OD) and vessels, will be selectively ignored by the experienced experts. At this point, we intend to model the IoR of normal areas. It is worth noting that, the normal vessel structures are networks and appearing as piece-wise linear structures, the patch of vessel generally can be sparely represented by its surroundings, the coding length $\delta L_\varepsilon(p)$ of majority of vessels is usually small. We hence mainly focus on modeling the inhibition of return for OD.

In light of the work in [10], we extend the computation of above-mentioned ICL in a global manner, as shown in Fig. 3. Instead of considering the surrounding patches of OD, we here introduce a set of reference set of OD $\mathcal{N}_{od}$, from which a dictionary $\mathbf{D}_{od} \in \mathbb{R}^{j \times k}$ is learned to yield a sparse representation of OD. Given an input patch p_{od}, its feature vector is denoted as $\mathbf{y}$, the spare code vector β is the optimal solution of the following problem:

$$\min_{\beta} \frac{1}{2} \|\mathbf{y} - \mathbf{D}_{od}\beta\|_2^2 + \gamma \|\beta\|_1 \tag{4}$$

where γ is a coefficient controlling the sparsity penalty.

Since ODs have a common appearance in different fundus images, the β of the patch located at the OD will be sparser than the one of the patch at other positions. Additionally, under a same level of sparsity, the square reconstruction error of the patch located at the OD will be smaller than the one of the patch at other positions. A weighted ICL of p_{od} is then computed by

$$\delta L_\varepsilon(p_{od}) \approx L_\varepsilon(p_{od}|\mathcal{N}_{od}) = \|\mathbf{y} - \mathbf{D}_{od}\beta\|_2^2 \cdot \|\beta\|_1 \tag{5}$$

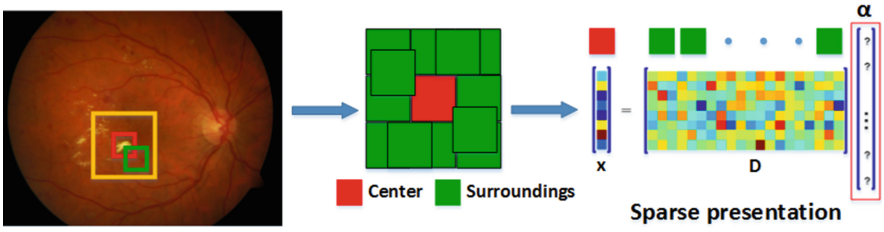


Fig. 2. Sparse representation of abnormal structures. (Color figure online)

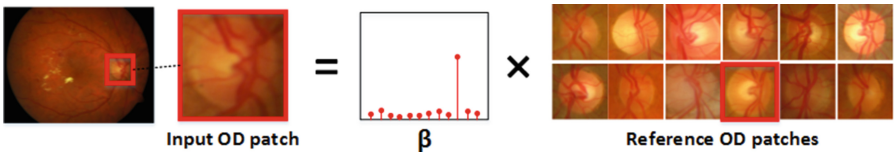


Fig. 3. Sparse representation of optic disc.

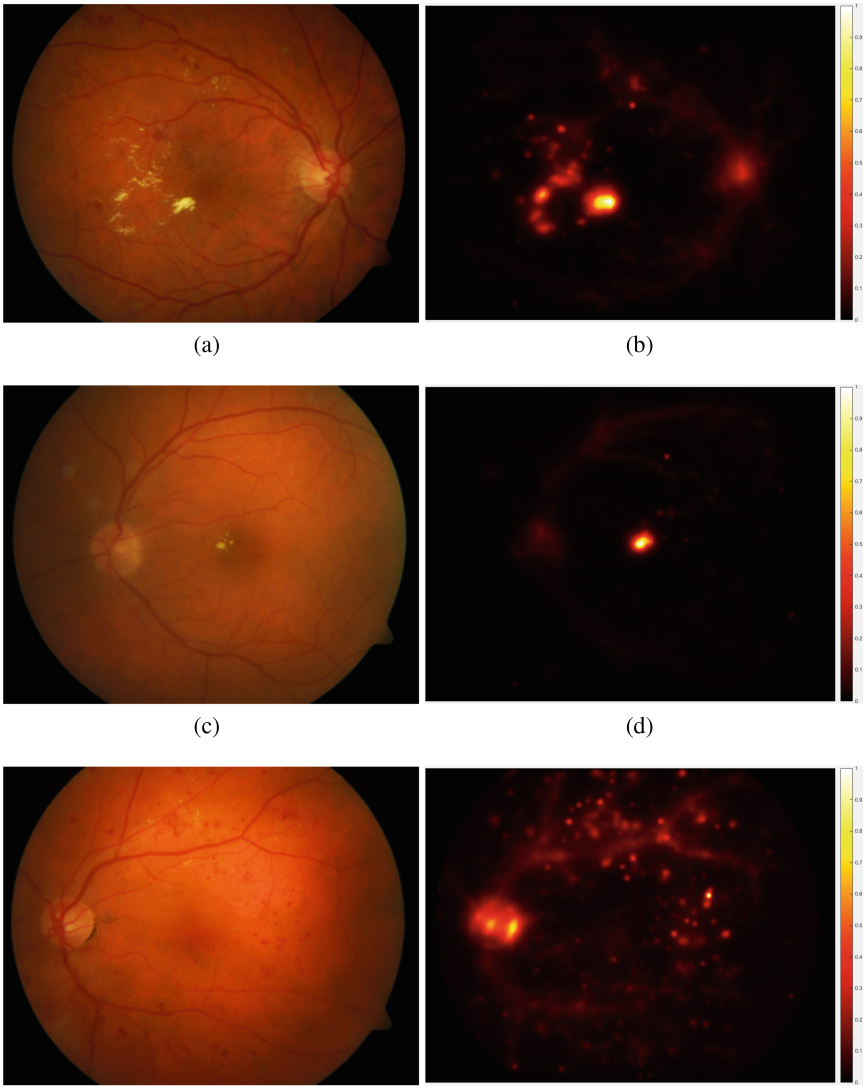


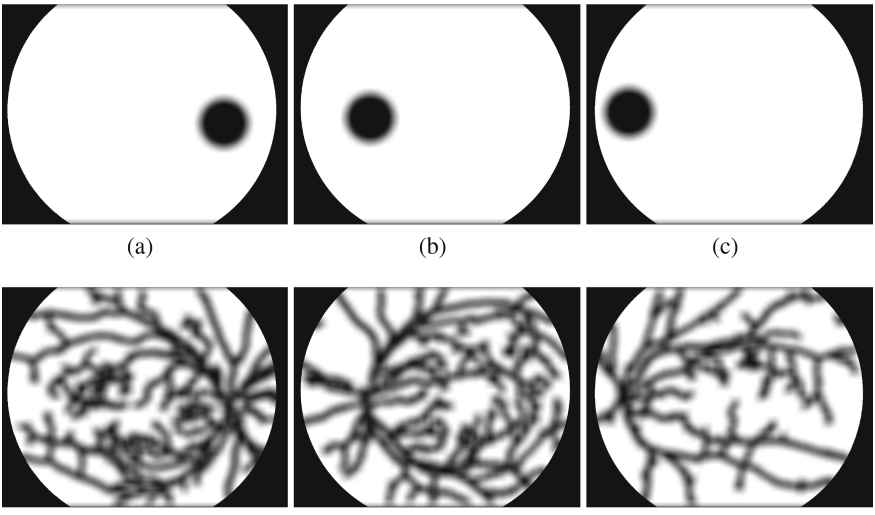
Fig. 4. Examples of abnormality attention map. (a), (c) and (e) Fundus images contains lesions. (b), (d) and (f) The map of abnormality attention.

For the patch located at the center of OD, its coding length $\delta L_\varepsilon(p_{od})$ is thus short, which can be used to reflect the level of inhibition of return. To avoid the overflow of the coding length, we introduce a sigmoid function and let

$$\delta L_\varepsilon(p_{od}) = \frac{1}{1+\exp^{-\Lambda}} - 1,\ \Lambda = \|\mathbf{y} - \mathbf{D}_{od}\beta\|_2^2 \cdot \|\beta\|_1 \tag{6}$$

(a) (b) (c)

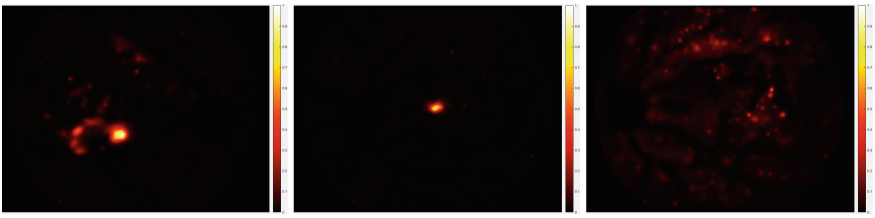
Fig. 5. The weighted ICL map of optic disc of Fig. 4(a), (c) and (e).



(a) (b) (c)

(d) (e) (f)

Fig. 6. IoR maps of normal areas. (a)–(b) IoR maps of the OD. (e)–(f) IoR maps of the vessels.



(a) (b) (c)

Fig. 7. Final lesion saliency map of Fig. 4(a), (c) and (e).

Figure 5 shows the weighted ICL map of the OD, where the OD region has the lowest response compared to others. Notice that there is only one OD in one fundus image, after computing the weighted ICL for each patch of the image, we select the point with the minimum weighted ICL as the center of the OD, and adopt an OD segmentation method [11] to segment the region of OD. The OD boundary is then approximated with a circle. The map of IoR of the OD is generated by setting the pixels inside the region of the segmented OD region to zero, while setting others to one. Further, to eliminate some interference of vessels, we here simply segment the vessel structures from the input fundus image with an adaptive thresholding method [1], and also construct an IoR map of the vessel structures by setting the pixels inside the segmented vessels to zero, while setting others to one. These maps are both smoothed with a Gaussian filter, and the final IoR maps are shown in Fig. 6.

2.3 Lesion Saliency Map Construction

After computing the abnormality attention and the IoR of input image, we next construct the final lesion saliency map. In this work, we multiply the map of abnormality attention and the two maps of IoR pixel-wisely, and let the result image as the final lesion saliency map, as shown in Fig. 7.

3 Experiments

3.1 Dataset

We tested our method on the public DiaRetDB1 V2.1 dataset [12], which contains 89 color fundus images with the fixed 1500×1152 resolution and are captured by the 50° FOV digital fundus camera. The images in this database involve four kinds of lesion, such as microaneurysms and hemorrhages, exudates and cotton-wool spots. For each image, four ground truth annotated by four different medical experts are provided. Since there are disagreements among four experts' annotations, we take a consensus of 75 % agreement as the fusion ground truth, as suggested in [12]. According to this ground truth, there are 51 images which contain lesions and 38 images without any lesions.

3.2 Implementation Details

In our experiments, the input image is contrast enhanced with the method in [13]. And for computing abnormality attention, the feature of p and $\mathcal{N}_p$ were extracted by directly vectorizing the red and green channel of the preprocessed RGB image patch with the size of $8 \times 8 \times 2$. The blue channel of RGB patch was abandoned since it is rather dark and short of useful information. In addition, because both dark and bright lesions in the fundus image have a variety of size, the computation of coding length was performed on image pyramid with 10 levels.

For computing the IoR of the OD region, the feature vector included three local features, i.e., the Dense SIFT [14], LIOP [15] and HoG [16], which were extracted from the green channel of RGB image patch, and the dictionary $\mathbf{D}_{od}$ is learned with the K-SVD algorithm [17].

3.3 Results

The performance of the proposed lesion detection was evaluated at image level as most of previous works in bright lesion detection. In this work, a detection of the image is considered as a true positive (TP) if finding out at least one kind of lesions in the image with lesions; a false negative (FN) if there is no founding of any lesions or only finding out the normal objects in the image with lesions; a false positive (FP) if finding out some objects in the image without any lesions; a true negative if there is no founding of any lesions in the image without any lesions. The sensitivity (SE) of the proposed method was then computed as $\mathrm{TP}/(\mathrm{TP}+\mathrm{FN})$, the specificity (SP) was computed as $\mathrm{TN}/(\mathrm{TN}+\mathrm{FP})$.

The proposed method achieves a SE of 0.71 at a SP of 0.82. To the best of our knowledge, no corresponding quantitative results have been reported for

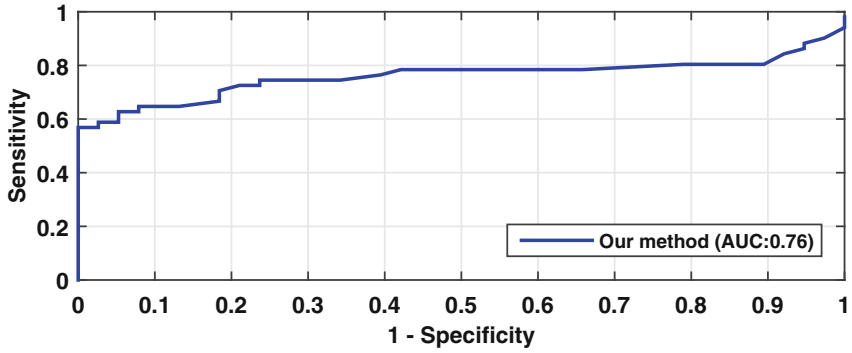


Fig. 8. ROC curve of fundus lesion detection.

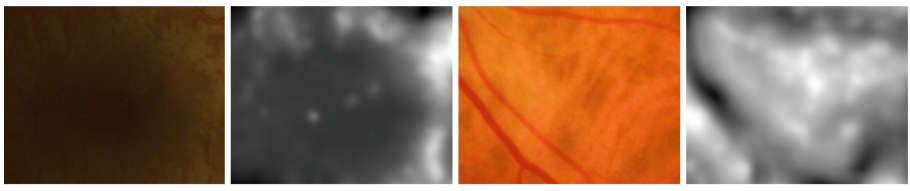


Fig. 9. The interference of lesion detection on normal images. (a) Image patch with normal macular reflection. (b) The corresponding saliency map patch of (a). (c) Image patch with the irregular background of tessellated retina. (b) The corresponding saliency map patch of (d).

Table 1. The performance of OD detection.

Methods	#Images	Success	ACC
Soares *et al.* [21]	89	88	98.88 %
Mahfouz and Fahmy [22]	89	87	97.75 %
Ramakanth and Babu [23]	89	88	98.88 %
The proposed method	**89**	**89**	**100 %**

detecting the dark and bright lesion at the same time. For comparison, here we simply exhibit some results of the methods only detect the dark or bright lesions. For bright lesion (exudates) detection, [18] reported a SE of 0.70 at a SP of 0.85 conducted on the same dataset, while for dark lesion (hemorrhages) detection, [19] reported a SE of 0.85 at a SP of 0.21 conducted on a private dataset. Noted that our method detects the dark and bright lesions at the same time, and can achieve a competitive performance with these methods, which proved that the proposed method is competent for the fundus lesion detection. The receiver operating characteristic (ROC) curve of the proposed lesion detection is also shown in Fig. 8, and the area under the ROC curve (AUC) is 0.76.

The proposed method is sensitive to the abnormal pattern of the fundus image, and many false positives occur on the normal macular reflection of the fundus image or the background region of the tessellated retina, as shown in Fig. 9. A further study on the lesion classification with supervised knowledge will improve the performance. We will attempt to resolve this problem as part of the future work.

Considering the computation of IoR for the OD can also be regarded as a method of OD localization, which can be used to analysis the cup-to-disc ratio for glaucoma diagnosis [20]. We also conducted an evaluation of our method to locate the OD on the database. Table 1 lists the accuracies of the proposed OD localization method and other approaches conducted on DiaRetDB1 V2.1 database, as can be seen, the proposed method can locate the OD region robustly, and outperforms other OD localization approaches.

4 Conclusions

In this paper, we proposed a novel method based on visual attention model to detect fundus lesions from the color retinal image. The abnormality attention of fundus lesion was measured with the incremental coding length computed in a local manner, while the IoR of the OD area was measured with the weighted incremental coding length computed in a global manner. The proposed method was evaluated on the DiaRetDB1 V2.1 database. The results revealed that the proposed method was an efficient scheme for automated lesion detection, and it was also an alternate method to locate the OD for the purpose of automated glaucoma diagnosis.

Acknowledgments. The authors would like to thank those who provided materials that were used in this study. This work was supported in part by the Natural Science Foundation of China under Grant 61472102, in part by the Fundamental Research Funds for the Central Universities under Grant HIT.NSRIF.2013091, and in part by the Humanity and Social Science Youth foundation of Ministry of Education of China under Grant 14YJC760001.

References

1. Zhang, B., Wu, X., You, J., Li, Q., Karray, F.: Detection of microaneurysms using multi-scale correlation coefficients. Pattern Recognit. **43**(6), 2237–2248 (2010)
2. Walter, T., Massin, P., Erginay, A., Ordonez, R., Jeulin, C., Klein, J.C.: Automatic detection of microaneurysms in color fundus images. Med. Image Anal. **11**(6), 555–566 (2007)
3. Tang, L., Niemeijer, M., Reinhardt, J.M., Garvin, M.K., Abràmoff, M.D.: Splat feature classification with application to retinal hemorrhage detection in fundus images. IEEE Trans. Med. Imaging **32**(2), 364–375 (2013)
4. Sánchez, C.I., García, M., Mayo, A., López, M.I., Hornero, R.: Retinal image analysis based on mixture models to detect hard exudates. Med. Image Anal. **13**(4), 650–658 (2009)
5. Hsu, W., Pallawala, P., Lee, M.L., Eong, K.G.A.: The role of domain knowledge in the detection of retinal hard exudates. In: Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001, CVPR 2001, vol. 2, pp. II-246–II-251. IEEE (2001)
6. Xiaohui, Z., Chutatape, O.: Detection and classification of bright lesions in color fundus images. In: 2004 International Conference on Image Processing, 2004, ICIP 2004, vol. 1, pp. 139–142. IEEE (2004)
7. Pratt, J., Abrams, R.A.: Inhibition of return to successively cued spatial locations. J. Exp. Psychol. Hum. Percept. Perform. **21**(6), 1343 (1995)
8. Wright, J., Ma, Y., Tao, Y., Lin, Z., Shum, H.Y.: Classification via minimum incremental coding length. SIAM J. Imaging Sci. **2**(2), 367–395 (2009)
9. Li, Y., Zhou, Y., Xu, L., Yang, X., Yang, J.: Incremental sparse saliency detection. In: 2009 16th IEEE International Conference on Image Processing (ICIP), pp. 3093–3096. IEEE (2009)
10. Hou, X., Zhang, L.: Dynamic visual attention: searching for coding length increments. In: Advances in Neural Information Processing Systems, pp. 681–688 (2009)
11. Wong, D., Liu, J., Lim, J., Jia, X., Yin, F., Li, H., Wong, T.: Level-set based automatic cup-to-disc ratio determination using retinal fundus images in argali. In: 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2008, EMBS 2008, pp. 2266–2269 (2008)
12. Kauppi, T., et al.: The DiaRetDB1 diabetic retinopathy database and evaluation protocol. In: Proceedings of BMVC, pp. 1–10 (2007)
13. Farbman, Z., Fattal, R., Lischinski, D., Szeliski, R.: Edge-preserving decompositions for multi-scale tone and detail manipulation. ACM Trans. Graph. **27**(3), 67:1–67:10 (2008)
14. Yang, J., Yu, K., Gong, Y., Huang, T.: Linear spatial pyramid matching using sparse coding for image classification. In: IEEE Conference on Computer Vision and Pattern Recognition, 2009, CVPR 2009, pp. 1794–1801. IEEE (2009)
15. Wang, Z., Fan, B., Wu, F.: Local intensity order pattern for feature description. In: Proceedings of ICCV, pp. 603–610. IEEE (2011)

16. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: Proceedings of CVPR, vol. 1, pp. 886–893. IEEE (2005)
17. Aharon, M., Elad, M., Bruckstein, A.: K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. IEEE Trans. Signal Process. **54**(11), 4311–4322 (2006)
18. Rocha, A., Carvalho, T., Jelinek, H.F., Goldenstein, S., Wainer, J.: Points of interest and visual dictionaries for automatic retinal lesion detection. IEEE Trans. Biomed. Eng. **59**(8), 2244–2253 (2012)
19. Hatanaka, Y., Nakagawa, T., Hayashi, Y., Mizukusa, Y., Fujita, A., Kakogawa, M., Kawase, K., Hara, T., Fujita, H.: CAD scheme for detection of hemorrhages and exudates in ocular fundus images. In: Medical Imaging, pp. 65142M–65142M. International Society for Optics and Photonics (2007)
20. Lowell, J., Hunter, A., Steel, D., Basu, A., Ryder, R., Fletcher, E., Kennedy, L.: Optic nerve head segmentation. IEEE Trans. Med. Imaging **23**(2), 256–264 (2004)
21. Soares, I., Castelo-Branco, M., Pinheiro, A.: Optic disk localization in retinal images based on cumulative sum fields (2015)
22. Mahfouz, A.E., Fahmy, A.S.: Fast localization of the optic disc using projection of image features. IEEE Trans. Image Process. **19**(12), 3285–3289 (2010)
23. Ramakanth, S.A., Babu, R.V.: Approximate nearest neighbour field based optic disk detection. Comput. Med. Imaging Graph. **38**(1), 49–56 (2014)