Computer Vision and Machine Learning techniques hard exudates detection eye-fundus images

Tiago Carvalho
Institute of Computing
UNICAMP
Campinas, Brazil
tjose@ic.unicamp.br

Anderson Rocha
Institute of Computing
UNICAMP
Campinas, Brazil
anderson.rocha@ic.unicamp.br

Jacques Wainer
Institute of Computing
UNICAMP
Campinas, Brazil
wainer@ic.unicamp.br

Siome K. Goldenstein
Institute of Computing
UNICAMP
Campinas, Brazil
siome@ic.unicamp.br

***Abstract*—*Diabetic retinopathy* (DR) is a diabetes development that affects the retina's blood flow. If untreated, DR eventually leads to blindness in persons in the age range of 20 to 74 years in developed countries. One of the most successful ways to fight DR is early diagnosing through the analysis of *ocular-fundus* images of the human retina. In this work, we present a new approach to detect Hard Exudates (HE) (a kind of DR lesion) from ocular-fundus images. Our work is geared towards the automatic triage scenario, where patients whose retina is considered not-normal by the system will see a specialist. This implies that automatic screening needs an evaluation criteria that rewards low false negative rates, i.e., we should avoid images incorrectly classified as normal as much as possible. Our solution constructs a visual dictionary based on normal and hard exudate images and classifies whether an *ocular-fundus* image is normal or a HE candidate. We evaluate the methodology on test different parameter configurations, and demonstrate the robustness and reliability of the approach performing five folds cross-validation[1].**

***Keywords*-Hard Exudate; Diabetic Retinopathy; Retinoscopy Images; Automatic Screening; Preventable Blindness; Visual Dictionaries.**

I. Introduction

Diabetes Mellitus (DM) is a systemic, chronic, and life-threatening disease that can cause micro and macro vascular changes. According to World Health Organization (WHO) [1], it affects 135 million people worldwide, and may affect as many as 366 million in the year 2030 [2].

Diabetic Retinopathy (DR) is mainly the result of microvascular retinal changes triggered by diabetes. If not treated in time, DR can lead to the complete loss of sight. Recent reports account that about 25 thousand people with diabetes go blind every year in the US due to diabetic retinopathy [3]. According to Abràmoff et al. [3], in the US and Europe, DR is the major cause of blindness for the economically active population[2] and, according to a 2002 report, it is estimated that DR is responsible for 5% of all the world's blindness cases [4].

Since early diagnostic and treatment of DR may prevent blindness [5], systematic screening (by specialists) of diabetic patients is a cost-effective procedure to reduce DR-caused blindness [6].

[1]Master's thesis

[2]In some developing countries (e.g., Brazil) this is not yet the case but it is rapidly going in this direction [4].

Hard exudates (HE) are one of many kind of lesions caused by DR, which appear at early disease stages and the detection is the focus of many works involving DR lesions recognition.

In this context, this work proposes an approach to detect HE based on points of interest and visual dictionaries. We validate our approach with a series of experiments on publicly available data, and compare the results against the state-of-the-art techniques.

Contributions: The proposed method shows how visual dictionaries can be used to detect HE in ocular fundus images. Our method have some important features, like, uses specialist knowledge in training stage and no uses any preprocessing in images. In contrast with literature approachs, that use specific methods to detect each kind of lesion, our method do not need many efforts to be applied in a different lesion, it just needs a new dictionary based on target lesion like showed in [7]. Other contribution is to shed light over that visual dictionaries can be used in problems with only two classes.

A. Related work

Most of the state-of-the-art approaches focus on single anomalies detection and, in most of them, ts extensions to other pathologies or a more general case are not straightforward or even possible.

There has been a few recent reviews in automatic retina analysis, with emphasis on DR detection, that describe some of the major research results in the area. We will refer the reader to these papers for a general description of the algorithms. Abràmoff et al. [8] is a more general review in retina image processing, including ocular-fundus image processing, while [9] and [10] are more specific to DR.

Most of the techniques used to detect lesions are based on specific segmentation or feature extraction techniques developed for *each* specific lesion. Thus, although these techniques have been achieving high accuracy rates, a technique developed for one kind of lesion cannot be used to detect another one.

There are many metrics to measure the success of a detection algorithm. Techniques interested in detecting each specific lesion, to be used, for example, as a decision support tool for specilists, usually report a *per lesion* metric, for example, sensitivity (number of true lesions detected over the number of lesions) and false positive rates per image (number of regions falsely marked as a lesion). Other approachs are interested in *per image* metrics, such as sensitivity (number of images tagged as having a lesion over the total number of images with lesion), and specificity (number of images tagged as normal over the total number of normal images).

Table I report some of the results of recent published results in hard exudate detection.

Table I
HARD EXUDATE DETECTION METHODS (SENSITIVITY *vs.* SPECIFICITY)
SEE TEXT FOR DETAILED EXPLANATION OF THE CASES.

Technique	**Sensitivity**	**Specificity**	**# of images in the data set**
Sopharak et al. [11]	80%	99.5%	60
Yun et al. [12]	80%	99.5%	124
Welfer et al. [13]	70.5%	98.8%	89
Wang et al. [14]	100%	70%	154
Garcia et al. [15]	88%	84%	117
Sopharak et al. [16]	87.3%	99.3%	60
Fleming et al. [16]	95%	86.6%	13,219
Sanchez et al. [17]	100%	90%	106
Sanchez et al. [18]	91.0%	70%	144

B. Technique overview

Figure 1 depicts the sequence of steps of the proposed approach (training and testing).

II. TECHNICAL BACKGROUND

In this section, we detail some essential concepts used in our approach.

A. Hard Exudate

Figure 2 depicts an ocular-fundus image with the retina's main regions highlighted.

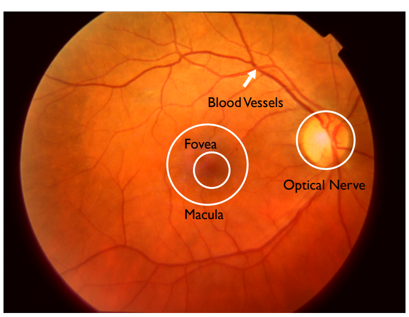

Figure 2. Ocular-fundus image and the retina's main regions.

In this work, we focus on the DR pathology (also called *lesions*), that is one of main indicative of an early stage of the disease. We are interested in lesions called *intra-retinal lipid exudates (hard exudates)*, which reflect the breakdown of the blood-retinal barrier. Failures in the blood-retinal barrier enable access of fluid rich in lipids and proteins to the parenchyma, causing retinal edema and exudation (Figure 3).

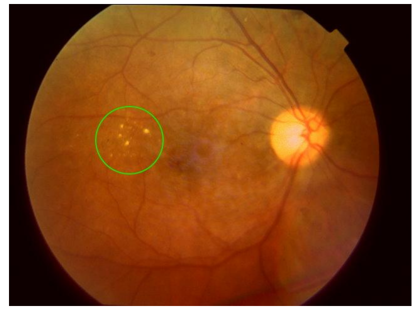

Figure 3. Intraretinal lipid exudates showed inside the green circle.

B. Local Features

In order to represent the visual content of a given image, we find a set of points of interest (PoI) in such images and characterize their surrounding regions. It is desired to choose scale-invariant interest points in order to achieve a representation robust to some possible image transformations. Among to perform such task, we can use several different approaches.

Some of this approaches are *Speeded-Up Robust Features* (SURF) [19] and *Scale-Invariant Features Transform* (SIFT) [20]. The SURF algorithm is based on the Hessian matrix, and your descriptor uses a distribution of Haar-wavelet responses within the interest point's neighborhood.

The SIFT algorithm use a PoI detector based on the DoG in multiple scales, and compose the descriptor using the gradient of the interest point's neighborhood in a local image.

C. Visual Dictionaries

As we have discussed in Section II-B, SURF (and SIFT) are good low-level representative feature detectors. However, this distinctiveness power comes with a price: as these solutions are often designed for exact matching, they do not translate directly into good results for image classification in broad or even constrained domains.

However, these approaches are not well suited for direct use when classifying exudate and normal images. To preserve the distinctiveness power of such descriptors while increasing their generalization, we use the concept of *visual dictionaries*.

Visual dictionaries are a strong images representation where each image is just a collection of regions and no matter the spacial region's form. The only considered information is the region's appearence [21].

The main point to build a visual dictionary is learn, using a training set, the generator model [22] that selects the r more representatives regions of the problem, so that, the number of regions must be enough to discern relevant changes on images without discern irrelevant changes, like noise [23], [24].

D. Support Vector Machine

Support Vector Machine (SVM) [25] is a useful technique for data classification. Given a training set of instance-label pairs $(\mathbf{x}_i, \mathbf{y}_i)$, $i = 1, ..., l$ where $\mathbf{x}_i \in \mathbb{R}^n$ and $\mathbf{y} \in \{1, -1\}^l$, the SVM solves the following optimization problem:

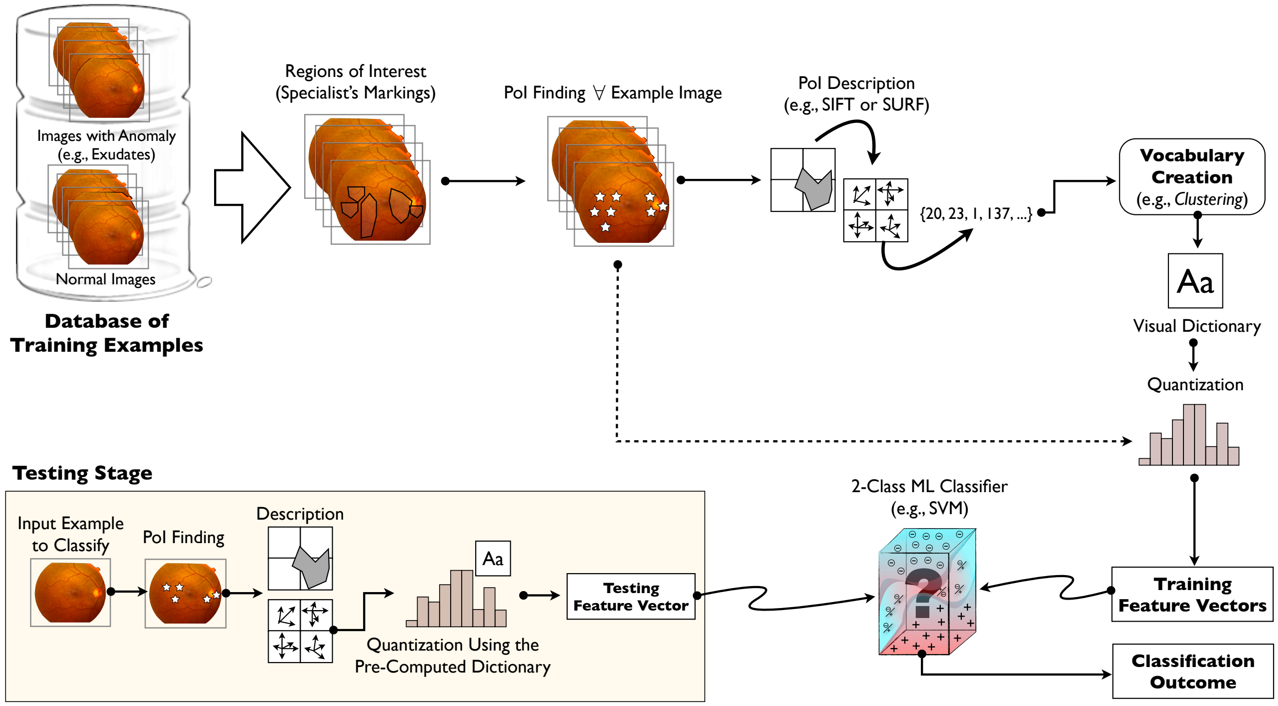

Figure 1. Sequence of steps for classifying HE lesions from retinoscopy ocular-fundus images.

$$\min_{\mathbf{w},b,\xi} = \left\{ \frac{1}{2}\mathbf{w}^T\mathbf{w} + \mathbf{C}\sum_{i=1}^{l}\xi_i \right\} \quad (1)$$

such that

$$\mathbf{y}_i\left(\mathbf{w}^T\phi\left(\mathbf{x}_i\right) + b \geq 1 - \xi_i\right)$$

and

$$\xi \geq 0.$$

The training vectors $\mathbf{x}_i$ are mapped into a higher dimensional space by function ϕ. Then, SVM finds a linear separating hyperplane with maximal margin in the referred dimensional space. In Eq. 1, $\mathbf{C} > 0$ is the penalty parameter of the error term. Furthermore, $\mathbf{K}\left(\mathbf{x}_i, \mathbf{x}_j\right) = \phi\left(\mathbf{x}_i\right)^T\phi\left(\mathbf{x}_j\right)$ is called the kernel function.

III. Methodology

To solve the problem of detecting exudates in ocular-fundus images, we select and create a database of training examples comprising training positive images with exudates and negative images considered normal by specialists. Furthermore, the specialists mark in normal images, regions that represents normal regions, and in images with exudates, they mark the best regions that represents exudates. All these regions (normal and exudate) are called “regions of interest” (RoI). The training stage than locate the points of interest within these RoI, using a local features algorithm.

After finding the PoIs, we need to create the dictionary or codebook representing distinctive features of images with exudates as well as images tagged as normal by specialists. For that, we need to choose the size (number of words) k of the dictionary. Thus all the PoI in normal RoI are clustered into $k/2$ groups, as are the PoI in non-normal RoI.

After creating the dictionary, we perform the assignment of each of the training images' PoIs to the closest visual word of the dictionary. This step is known as *quantization*. In the end of the quantization process, one left with a set of feature vectors representing the histogram of the selected visual words for each image.

In order to perform the final classification procedure, we use SVM with two-class. For training the classifier, we feed it with feature vectors calculated using the training images containing positive (e.g., images containing exudates) and negative (normal images) examples. Once the dictionary is created, and the classifier trained, the training step is completed. To classify a new image, one proceed as follows:

- compute the PoIs of the image using SURF;
- map each PoI to its closest word in the dictionary.

IV. Experiments and Results

We validate our technique through three rounds of experiments, where each experiment show us the best parameters configuration to choice at final approach configuration. All rounds have been conducted about DR1 dataset (all the reported results are the average of a 5-fold cross-validation procedure) and using ROC curves to display the results.

A. DR1 dataset

This is a dataset from the Ophthalmology Department of Federal University of São Paulo (Unifesp), collected across several months. The dataset comprises 932 images from which 687 are normal retinas and 245 images contain HE images. All the images in this dataset were manually annotated by three medical specialists. The average resolution of the images is 640×480 pixels. Aiming to allow researchers to perform fair comparisons with respect to the approach discussed in this paper, we make this dataset publicly available[3].

B. SVM parameters

To adjust the SVM parameters, a grid search using RBF kernel was perfomed on training set in each round of experiments. The only parameter that not changed in all rounds was the class weight. This weight was calculated using the number of images in each class on training set.

C. ROC curves

In our ROC curves, the *sensitivity* depicts the true positive rates (true positive are HE images correctly classified) and *1 - specificity* depicts the false negative rates (false negative are normal images incorrectly classified). The ROC curve is generated according to a standard practice in which we perform a classification over the test examples using the chosen classifier (SVM in this case) and then move the decision boundary margin in order to produce the characteristic curve.

D. Rounds of Experiments

Round 1: In this round, we discuss which is the best approach to select local features in our methodology. As such, SIFT and SURF algorithms have been tested using 500 visual words random selected in RoI. The results are depicted in Figure 4.

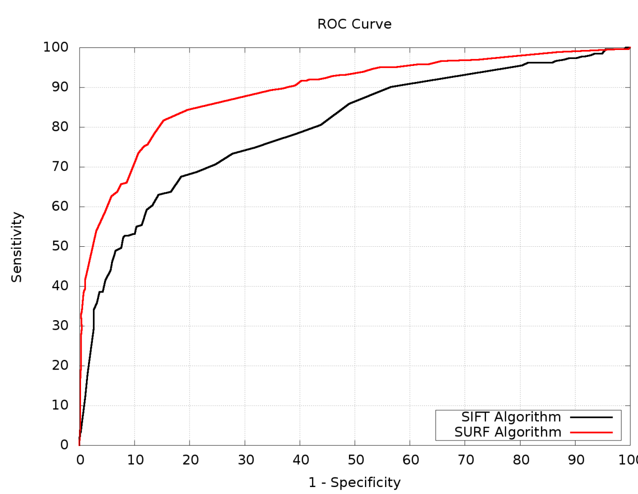

Figure 4. ROC curves using SIFT and SURF algorithms to select local features.

[3]Please visit http://www.recod.ic.unicamp.br/site/asdr/

Round 2: Since the SURF was selected to be our local features descriptor, in round two, we investigate which is the best way to select the visual dictionary words. In our first try, we select 500 (250 in normal images and 250 in HE images) randomly visual words in RoIs of the training set images. Then, in our second try, we used clustering (k-means) to select 250 more representatives words for normal images and 250 more representatives words for lesion's words at the same RoIs of the first try. The Figure 5 shows this round results.

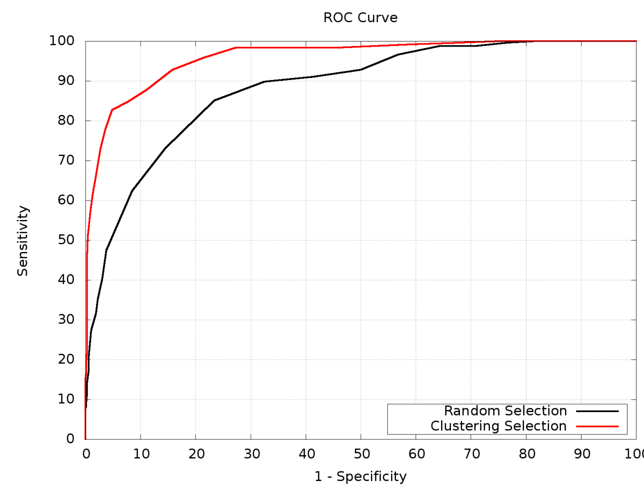

Figure 5. ROC curves using clustering and randomized selection to select the visual dictionary.

Round 3: Using the best choices of rounds 1 and 2 (SURF and clustering), in round 3 we have checked how the number of visual words affects our methodology. The results are depicted in Figure 6.

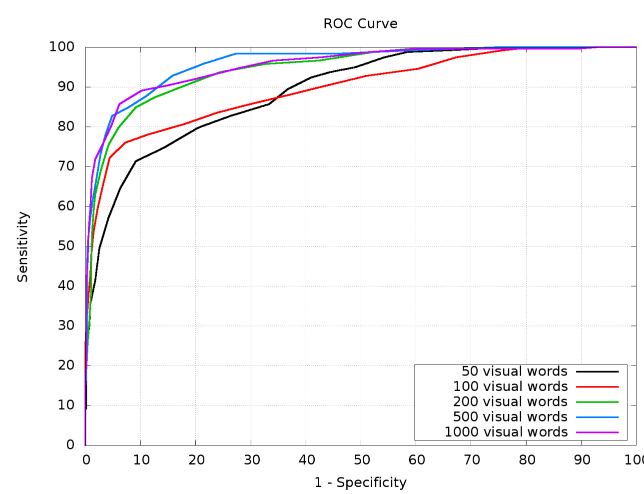

Figure 6. ROC curves using visual dictionaries with different sizes.

Based in all previously described experiments, the best-performing is achieved using the SURF algorithm like local features, clustering to select visual words that will be compose the visual dictionary, and using a dictionary with size consists of 500 words (250 HE lesions-based words plus 250 normal-based words). In particular, at the 90% sensitivity, the methodology achieves 87% specificity. The equal error rate (EER) for the bright lesions classifier considering a dictionary with 500 words is 88% and the corresponding area under the curve is AUC = 95.3%.

If we compare our results with the results presented in Table I at Section I, is clearly shown that our results are equivalents with the most of results showed. However, the big advantage of our approach is that it can be easyly applied to detect other kinds of anomalies, like showed in [7].

V. Conclusion

One of the most successful means for fighting DR in developed and developing countries continues to be its early diagnosis through the analysis of ocular-fundus images of the human retina.

In this direction, in this work we have used a visual dictionary approach for detecting hard exudate lesions in retinoscopy ocular-fundus images. Visual dictionary is an elegant method to learn and represent important features of hard exudate images, and allows us to classify whether an ocular-fundus image is normal or a HE candidate.

We have performed experiments on a data set with more than 900 images, using five fold cross-validation, and we showed the effectiviness of our contribution. Using just a few have 500 visual words, we are already able to achieve good classification results. In order to create the visual dictionary representing an ocular-fundus image with HE anomaly and its normal counterparts, we evaluate two possibilities: random selection and clustering in RoIs. To find these RoIs, a specialist just needs to mark a few regions on images to provide the system with the basis for classification. This procedure is performed just once during the training stage.

Our best results were achieved using SURF for the detection and description of the points of interest and a visual dictionary of 500 words, selected using clustering in RoIs.

An important characteristic of our approach is that it is not specific to hard exudates, and we shown in other works [26], [27] that the technique can be readily applied to detect microaneurysms and deep hemorrhages with very competitive results.

Our future work includes the combination of simple detectors specialized in detection of simple anomalies such as: exudate, microaneurysms, deep hemorrhages and other detectors to create a final high-level detector for normal *vs.* non-normal ocular-fundus images. We believe the combination of such simple detectors will provide high classification results representing a step forward in computer aided diagnosing using ocular-fundus retinoscopy images.

Acknowledgment

We thank the Microsoft Research and Fapesp financial support under the grants MSR-Fapesp 2008/54443-2 and Fapesp 2010/05647-4. We also thank the Federal University of São Paulo medical team for helping us to collect and tag the ocular-fundus images. In particular, we express our gratitude to Dr. Eduardo Dib, from Federal Univ. of São Paulo, for taking the lead in grading the training images.

References

[1] World Health Organization, "Diabetes programme," Online, August 2010, `http://www.who.int/diabetes/en`.

[2] S. Wild, G. Roglic, A. Green, R. Sicree, and H. King, "Global prevalence of diabetes estimates for the year 2000 and projections for 2030," *Diabetes Care*, vol. 27, no. 5, pp. 1047–1053, 2005.

[3] M. D. Abràmoff, M. Niemeijer, M. S. Suttorp-Schulten, M. A. Viergever, S. R. Russell, and B. van Ginneken, "Evaluation of a system for automatic detection of diabetic retinopathy from color fundus photographs in a large population of patients with diabetes," *Diabetes Care*, vol. 31, no. 2, pp. 193–198, 2008.

[4] S. R. Salomão, M. R. K. H. Mitsuhiro, and R. Belfort Jr., "Visual impairment and blindness: an overview of prevalence and causes in Brazil," *Anais da Academia Brasileira de CiÃªncias*, vol. 81, pp. 539–549, 2009.

[5] Q. Mohamed, M. C. Gillies, and T. Y. Wong, "Management of diabetic retinopathy: A systematic review," *JAMA*, vol. 298, no. 8, pp. 902–916, 2007.

[6] M. James, D. A. Turner, D. M. Broadbent, J. Vora, and S. P. Harding, "Cost effectiveness analysis of screening for sight threatening diabetic eye disease," *BMJ*, vol. 320, p. 1627, 2000.

[7] T. J. de Carvalho, "Aplicação de técnicas de visão computacional e aprendizado de máquina para a deteção de exsudatos duros em imagens de fundo de olho," Master Thesis, Universidade Estadual de Campinas, 2010.

[8] M. D. Abràmoff, M. K. Garvin, and M. Sonka, "Retinal imaging and image analysis," *IEEE Reviews in Biomedical Engineering*, vol. 3, pp. 169–208, 2010.

[9] R. Winder, P. Morrow, I. McRitchie, J. Bailie, and P. Hart, "Algorithms for digital image processing in diabetic retinopathy," *Computerized Medical Imaging and Graphics*, vol. 33, pp. 608 – 622, 2009.

[10] O. Faust, R. Acharya, E. Y. K. Ng, K.-H. Ng, and J. S. Suri, "Algorithms for the automated detection of diabetic retinopathy using digital fundus images: A review," *Journal Medical Systems*, 2011, dOI: 10.1007/s10916-010-9454-7.

[11] A. Sopharak, B. Uyyanonvara, S. Barman, and T. H. Williamson, "Automatic detection of diabetic retinopathy exudates from non-dilated retinal images using mathematical morphology methods," *Computerized Medical Imaging and Graphics*, vol. 32, p. 8, 2008.

[12] W. L. Yun, A. Rajendra, Y. V. Venkatesh, C. Chee, L. Min, and E. Y. K. Ng, "Identification of different stages of diabetic retinopathy using retinal optical images," *Intl. Journal on Information Sciences*, vol. 178, pp. 106–121, 2008.

[13] D. Welfer, J. Scharcanski, and D. R. Marinho, "A coarse-to-fine strategy for automatically detecting exudates in color eye fundus images," *Computerized Medical Imaging and Graphics*, vol. 34, no. 3, pp. 228–235, 2010.

[14] H. Wang, W. Hsu, K. G. Goh, and M. L. Lee, "An effective approach to detect lesions in color retinal images," in *IEEE Intl. Conference on Computer Vision and Pattern Recognition*, 2000, pp. 181–186.

[15] M. Garcia, C. I. Sanchez, M. I. Lopez, and R. H. Daniel Abasolo, "Neural network based detection of hard exudates in retinal images," *Computer Methods and Programs in Biomedicine*, vol. 93, pp. 9–19, 2009.

[16] A. Sopharak, B. Uyyanonvara, and S. Barman, "Automatic exudate detection from non-dilated diabetic retinopathy retinal images using fuzzy c-means clustering," *Sensors*, vol. 9, no. 3, pp. 2148–2161, 2009.

[17] C. I. Sanchez, M. Garcia, A. Mayo, M. I. Lopez, and R. Hornero, "Retinal image analysis based on mixture models to detect hard exudates," *Medical Image Analysis*, vol. 13, pp. 650–658, 2009.

[18] C. I. Sanchez, M. Niemeijer, M. S. A. S. Schulten, M. Abràmoff, and B. van Ginneken, "Improving hard exudate detection in retinal images through a combination of local and contextual information," in *IEEE Intl. Symposium on Biomedical Imaging*, 2010, pp. 5–8.

[19] H. Bay, T. Tuytelaars, and L. V. Gool, "SURF: Speeded up robust features," in *European Conference on Computer Vision*, 2006, pp. 1–14.

[20] D. Lowe, "Distinctive image features from scale-invariant keypoints," *Intl. Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[21] J. Sivic and A. Zisserman, "Video google: A text retrieval approach to object matching in videos," in *IEEE Intl. Conference on Computer Vision*, 2003, pp. 1470–1477.

[22] I. Ulusoy and C. M. Bishop, "Generative versus discriminative methods for object recognition," in *IEEE Intl. Conference on Computer Vision and Pattern Recognition*, vol. 2, 2005, pp. 258–265.

[23] F.-F. Li and P. Perona, "A bayesian hierarchical model for learning natural scene categories," in *IEEE Intl. Conference on Computer Vision and Pattern Recognition*, vol. 2, 2005, pp. 524–531.

[24] G. Csurka, C. R. Dance, L. Fan, J. Willamowski, and C. Bray, "Visual categorization with bags of keypoints," in *Workshop on Statistical Learning in Computer Vision*, 2004, pp. 1–8.

[25] C. M. Bishop, *Pattern Recognition and Machine Learning*, 1st ed. Springer, 2006.

[26] A. Rocha, T. Carvalho, S. Goldenstein, and J. Wainer, "Points of Interest and Visual Dictionary for Retina Pathology Detection," Institute of Computing, University of Campinas, Tech. Rep. IC-11-07, March 2011.

[27] H. F. Jelinek, A. Rocha, T. Carvalho, S. Goldenstein, and J. Wainer, "Machine learning and pattern classification in identification of indigenous retinal pathology," in *33rd IEEE Annual Intl. Conference on Engineering in Medicine and Biology (EMBC)*, 2011.