Classification of SD-OCT Volumes with LBP: Application to DME Detection

Guillaume Lemaître⋆[1,2], Mojdeh Rastgoo[1,2], Joan Massich[2], Fabrice Mériaudeau[2], and Désiré Sidibé[2]

[1] ViCOROB, Universitat de Girona, Campus Montilivi, Edifici P4, 17071 Girona, Spain,
[2] LE2I UMR6306, CNRS, Arts et Métiers, Univ. Bourgogne Franche-Comté, 12 rue de la Fonderie, 71200 Le Creusot, France

Abstract. This paper addresses the problem of automatic classification of Spectral Domain OCT (SD-OCT) data for automatic identification of patients with Diabetic Macular Edema (DME) versus normal subjects. Our method is based on Local Binary Patterns (LBP) features to describe the texture of Optical Coherence Tomography (OCT) images and we compare different LBP features extraction approaches to compute a single signature for the whole OCT volume. Experimental results with two datasets of respectively 32 and 30 OCT volumes show that regardless of using low or high level representations, features derived from LBP texture have highly discriminative power.
Moreover, the experiments show that the proposed method achieves better classification performances than other recent published works.

Keywords: Diabetic Macular Edema, Optical Coherence Tomography, DME, OCT, LBP.

1 Introduction

Eye diseases such as Diabetic Retinopathy (DR) and Diabetic Macular Edema (DME) are the most common causes of irreversible vision loss in individuals with diabetes. Just in United States alone, health care and associated costs related to eye diseases are estimated at almost \$500 M [18]. Moreover, the prevalent cases of DR are expected to grow exponentially affecting over 300 M people worldwide by 2025 [24]. Early detection and treatment of DR and DME play a major role to prevent adverse effects such as blindness. Indeed, the detection and diagnosis of retinal diseases are based on the detection of vascular abnormalities or lesions in the retina.

In past decades, Computer Aided Diagnosis (CAD) systems devoted to ophthalmology, have been developed focusing on the automatic analysis of fundus images [1, 21]. However, the use of fundus photography is limited to the detection of signs which are correlated with retinal thickening such as hard and soft

⋆ Corresponding authors: `g.lemaitre58@gmail.com`, `mojdeh.rastgoo@gmail.com`, `mailsik@gmail.com` - Source available in GitHub [12]

exudates, hemorrhages or micro-aneurysms. However, DME is characterized as an increase in retinal thickness within 1 disk diameter of the fovea center with or without hard exudates and sometimes associated with cysts [10]. Therefore, fundus photography cannot always identify the clinical signs of DME; for example cysts, which are not visible in the retinal surface. In addition, it does not provide any quantitative measurements of retina thickness or information about cross-sectional retinal morphology.

Recently, Optical Coherence Tomography (OCT) has been widely used as a valuable diagnosis tool for DME detection. OCT is based on optical reflectivity and produces cross-sectional and three-dimensional images of the central retina, thus allowing quantitative retinal thickness and structure measurements. The new generation of OCT imaging, namely Spectral Domain OCT (SD-OCT) offers higher resolution and faster image acquisition over conventional time domain OCT. SD-OCT can produce $27,000$ to $40,000$ A-scans/seconds with an axial resolution ranging from $3.5\,\mu m$ to $6\,\mu m$ [6]. Many of the previous works on OCT image analysis have focused on the problem of retinal layers segmentation, which is a necessary step for retinal thickness measurements [7, 11], and few have addressed the specific problem of DME and its associated features detection from OCT images.

In this research we focus on the latter problem and propose an automatic framework for identification of DME patients versus normal subjects using OCT volumes. The proposed method is based on Local Binary Patterns (LBP) features to describe the texture of OCT images and dictionary learning using the Bag-of-Words (BoW) models [19]. We propose to extract 2D and 3D LBP features from OCT images and volumes, respectively. The LBP descriptors are further extracted from the entire sample or local patches within individual samples. In this research beside the comparison of 2D and 3D features, we also compare the effects of common pre-processing steps for OCT data, and different classifiers.

In the following of this paper, fist in Sect. 2 a summary of the related studies is presented.

This paper is organized as follows, Section 2 presents a summary of the related studies. The proposed framework is explained in Sect. 3, while the experiments and results are discussed in Sect. 4. Finally, the conclusion and avenue for future directions are drawn in Sect. [**?**].

2 Background

review, add previous version of the paper

Quellec *et al.* proposed a method for the identification of fluid-filled regions in SD-OCT images of the macula based on texture features extracted in the pre-segmented retinal layers [16].

The authors in [20] proposed a classification method to distinguish DME, Age-related Macular Degeneration (AMD) and normal SD-OCT volumes. The OCT images are pre-processed by reducing the speckle noise by enhancing the

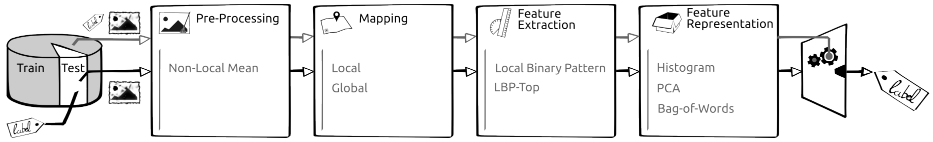


Fig. 1. Machine learning classification basic scheme

sparsity in a transform-domain and flattening the retinal curvature to reduce the inter-patient variations. Then, Histogram of Oriented Gradients (HOG) are extracted for each slice of a volume and a linear Support Vector Machines (SVM) is used for classification. On a dataset of 45 patients equally subdivided into the three aforementioned classes, this method leads to a correct classification rate of 100%, 100% and 86.67% for normal, DME and AMD patients, respectively.

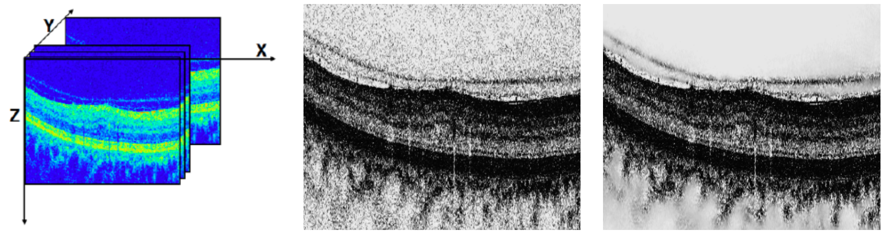
Venhuizen *et al.* also proposed a method for OCT images classification using the BoW models [23]. The method starts with the detection and selection of keypoints in each individual B-scan by keeping the most salient points corresponding to the top 3% of the vertical gradient values. Then, a texton of size 9×9 pixels is extracted around each keypoint, and Principal Component Analysis (PCA) is applied to reduce the dimension of every texton to get a feature vector of size 9. All extracted feature vectors are used to create a codebook using k-means clustering, and the obtained codebook from the training is used to represent each OCT volume as a feature vector occurrence histogram. Finally, this histogram is used as feature vector to train a Random Forest (RF) with a maximum of 100 trees. The method was used to classify OCT volumes between AMD and normal cases and achieved an Area Under the Curve (AUC) of 0.984 with a dataset of 384 OCT volumes.

The most similar work to ours is the work of Liu *et al.* who proposed a method for macular pathology detection in OCT images using LBP and gradient information as attributes [13]. The method starts by aligning and flattening the images, then a 3-level multi-scale spatial pyramid is created and edge and LBP histograms are extracted in each block at every level of the pyramid. All obtained histograms are concatenated into a global descriptor whose dimensions are reduced using PCA. Finally a SVM is used as classifier. The method achieved good results in detection OCT scan containing different pathology such as DME or AMD, with an AUC of 0.93 using a dataset of 326 OCT scans.

3 Materials and Methods

In my opinion this section should contain a description of the things we use in a generic manner. Describe the data, all the blocks in the figure but not their relations

The proposed method, as well as, its experimental set-up for OCT volume classification are outlined in Fig. 1. The methodology is formulated as a standard classification procedure, through 5 steps which are explained in the following.



(a) (b) (c)

Fig. 2. OCT: (a) Organization of the OCT data - (b) Original image - (c) NL-means filtering.

First, the OCT volumes are pre-processed as presented in details in Sect. 3.1. The mapping stage is used to determine a discrete set of elements (or structures) which is used for representing the OCT volume. Thereafter, two mapping strategies are defined: (i) *global* and (ii) *local* mapping. In the global mapping approach, a single structure is computed for the image/volume while in the local mapping, a set of structures is defined by sliding a window through the image/volume. Then, a descriptor is computed for each structure. The feature extraction and representation are presented in depth in Sect. 3.2 and Sect. 3.3, respectively. Finally the classification step is presented in Sect. 3.4.

3.1 Image pre-processing

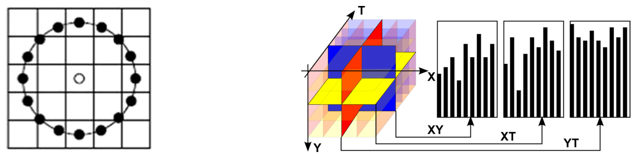
describe the new preprocessing strategies. They should all be explained as independent stuff, and in experimentation we'll describe which are we using for each experiment scenario.

OCT images are known to be affected by a speckle noise [17]. Subsequently, Non-Local Means (NL-means) [4] filtering has been successfully used in Ultra-Sound (US) images to filter similar noise [8] and is used in our framework to denoise each B-scan (i.e. each $x - z$ slice) of the OCT volumes (see in Fig. 2(a)). NL-means filtering offers the advantage to use all the possible self-predictions that the image can provide rather than local or frequency filters such as Gaussian, anisotropic or Wiener filters [4]. An example of filtering using NL-means filter on OCT image is depicted in Fig 2(b) and Fig. 2(c).

3.2 Features extraction

LBP is a texture descriptor based on the signs of the differences of a central pixel with respect to its neighboring pixels [15]. These differences are encoded in terms of binary patterns as in Eq. (1):

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p \ , \qquad s(\cdot) = \begin{cases} 1 & \text{if } (g_p - g_c) \geq 0 \\ 0 & \text{otherwise} \end{cases} \ , \tag{1}$$



(a) (b)

Fig. 3. The different LBP descriptors: (a) LBP with $(R = 2, P = 16)$ - (b) LBP-TOP [25].

where g_c, g_p are the intensities of the central pixel and a given neighbor pixel, respectively. P is the number of sampling points in the circle of radius R. Figure 3(a) illustrates the meaning of P and R.

Ojala *et al.* further extend the original LBP formulation to achieve rotation invariance at the expense of limiting the texture description to the notion of circular "uniformity" [15]. Volume encoding is later proposed by Zhao *et al.* by computing LBP descriptors in each orthogonal planes, so called LBP from Three Orthogonal Planes (LBP-TOP) [25].

3.3 Feature representation

Each OCT volume can be described by its texture and we employed two strategies.

Low-level representation The texture descriptor of an OCT volume is defined as the concatenation of the LBP histograms. Regarding the LBP-TOP, the feature descriptor is computed through the concatenation of the LBP histograms of the three orthogonal planes. Furthermore, the size of this entire feature vector is defined according to the mapping strategy chosen (see Fig. 1).—

High-level representation According to the chosen mapping strategy, the low-level representation can lead to a high dimensional feature space. High-level representation simplifies this high dimensional feature space into a more discriminant lower space. BoW approach is used for this purpose [19]. This model represents the features by creating a visual dictionary or "codebook", from the set of low-level features. The set of low-level features are clustered using k-means to create the codebook with k defining the number of visual words. After creating the codebook, each of the training example is represented as a histogram of size k obtained by calculating the frequency of occurrences of each of the k words in the extracted features from the training example.

3.4 Classification

Classification is a supervised learning method which intends to find a mapping $f(.)$ which relates a set of input x to a set of categorical outputs y. The learning is comprehended using a training set, which contains a set of N samples with their associated labels [14].

In this study we compare the performance of different classifiers including: RF, Gradient Boosting (GB), SVM, Logistic Regression (LR), and k-Nearest Neighbor (NN).

Random Forest is an ensemble of decision trees and was introduced by [3]. The ensemble uses each tree to predict an output and finalizes the ultimate prediction by aggregating the outputs of all tress. This classifier learns the data by training multiple decision trees on bootstrap samples of the original data. Each bootstrap of D dimension is used for training one decision tree and at each node, the best split among randomly ($d << D$) selected subset of descriptors is chosen. Each tree is grown to its maximum length without any pruning. In the prediction stage a sample is voted by each tree and it is labeled by considering the majority of the votes.

Gradient Boosting is a generalization form of AdaBoost (AdB), which is able to use real-value weak learners and minimizes different loss functions [26]. GB builds the ensemble in a greedy manner. It iteratively selects the best pair of real-valued weak learners and adjust their weights so that they minimize a given differentiable loss function. Common choice for the weak learner is decision stumps or regression trees while the loss function is generally an exponential loss or a logarithmic loss [2]. This minimization is carried out via gradient descent or quadratic approximation.

Support Vector Machines [22] is a sparse kernel method which aims to separate two classes by finding the best hyperplane which maximizes the margin between the two classes. Maximizing the margin is equivalent to minimizing the norm of the normal vector of the hyperplane:

$$\min_{\mathbf{w},\omega_0} \frac{1}{2}\|\mathbf{w}\|^2 \qquad \text{s.t.} \quad y_i(\mathbf{w}^T\mathbf{x_i} + \omega_0) \geq 1, i = 1 : N \tag{2}$$

This constraint intends to force all the point to be in the correct side of the decision boundary (hyperplane) with a minimum distance of 1. This assumption is only valid if the data is linearly separable. Thus for general cases a slack variable $\xi_i \geq 0$ is introduced, which is $\xi_i = 0$, if the points are on/or inside the correct margin boundary, is $0 < \xi_i \leq 1$ if the points are inside the margin but in the correct side of the decision boundary and otherwise if they lie in the wrong side of decision boundary is $\xi_i > 1$. This assumption introduce the *soft margin constraints.*

$$\begin{aligned} &\min_{\mathbf{w},\omega_0,\xi} \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{N}\xi_i \\ \text{s.t.} \quad &\xi_i \geq 0, \quad y_i(x_i^T\mathbf{w} + \omega_0) \geq 1 - \xi_i, i = 1 : N \end{aligned} \tag{3}$$

In this regard, the $\sum_i \xi_i$ term in Eq. 3, describes the upper bound on the number of misclassified points and C is the regularization parameter that controls the tolerance of the classifiers on the number of errors [14].

Logistic Regression [9] is another supervised learning which can provide associated probability of each prediction. As the name suggests, this classifier uses a logistic function, to estimate the probabilities. By defining the posterior probability for one of the classes, using logistic function (see Eq. 4), the probability of other class is defined as $p(c_2|x_i) = 1 - p(c_1|x_i)$.

$$p(c_1|x_i) = \frac{1}{1 + \exp(-w^T x_i)} \tag{4}$$

Here w is a vector of regression parameters, which allows to obtain a linear combination of the input feature vector x_i. Using this model the unlabeled sample is assigned to the class which maximizes the posterior probability.

$$C(x_i) = \arg\max_k p(C = k|x_i) \tag{5}$$

In this method, finding an optimal set of parameters for w is essential. The vectors of parameters w can be inferred by finding the maximum likelihood estimates via optimization methods such as quasi-Newton method [5].

k-Nearest Neighbor is one the simplest supervised machine learning classification methods. In this method a new unlabeled vector is assigned to the most represented class from k nearest-neighbors in the features space. In order to avoid a tie case, the parameter k is usually an odd number.

4 Experiments and Validation

4.1 Datasets

In this work, we validated our classification framework using two different datasets.

SERI - datasets were acquired by Singapore Eye Research Institute (SERI), using CIRRUS TM (Carl Zeiss Meditec, Inc., Dublin, CA) SD-OCT device. The datasets consist of 32 OCT volumes (16 DME and 16 normal cases). Each volume contains 128 B-sane with dimension of 512 × 1024 pixels. All SD-OCT images are read and assessed by trained graders and identifies as normal or DME cases based on evaluation of retinal thickening, hard exudates, intraretinal cystoid space formation and subretinal fluid.

Duke - datasets published by Srinivasan et al. [20] were acquired in Institutional Review Board-approved protocols using Spectralis SD-OCT (Heidelberg Engineering Inc., Heidelberg, Germany) imaging at Duke University, Harvard University and the University of Michigan. This datasets consist of 45 OCT volumes (15 AMD, 15 DME and 15 normal). In this study we only consider a subset of the original data containing 15 DME and 15 normal OCT volumes.

4.2 Experiments & Results

Both datasets are filtered to attenuate the effect of speckle noise. SIRE dataset is processed using NL-means as stated in Sect. 3.1. The different parameters were empirically tested and fixed such that the patch size, the search window and the filtering parameter were set to (15×15), (35×35) and 0.4, respectively. However, Duke dataset is already filtered using BM3D method [20]. For both datasets, LBP and LBP-TOP features are extracted for different sampling points of 8, 16 and 24 for radius of 1, 2 and 3, respectively. Two different mapping strategies are used: (i) *global* mapping corresponding to the 2D B-scan for LBP or the 3D volume for LBP-TOP and (ii) *local* mapping considering to a set of 2D P of size (7×7) for LBP or the 3D sub-volume for LBP-TOP of size $(7 \times 7 \times 7)$. For the high-level representation, when PCA is applied, the eigenvectors associated with the largest 99% cumulative eigenvalues are selected to reduce the number of dimensions. In BoW approach, an empirical search was performed to find the optimal number of visual words which is finally fixed to 32. The number of trees for each RF classifier was fixed to 100.

For evaluation purposes, all the results are expressed in terms of Sensitivity (SE) and Specificity (SP) using a Leave-One-Patient Out Cross-Validation (LOPO-CV) strategy. Thus, at each round a pair DME-normal volume is selected for testing while the rest are used for training. The use of LOPO-CV implies that no variance in SE and SP can be reported. However, and despite this limitation, LOPO-CV has been employed due to the small size of the datasets.

Experiment #1 is carried out on SERI dataset. Both low and high level feature representation are extracted and tested. The results are reported in Table 1.

Experiment #2 is carried out on the Duke dataset [20]. The OCT volumes provided by this dataset are cropped, with different sizes. Subsequently, the experiments involving the mapping using 2D B-scan do not comply with these requirements and thus are not carried out. The obtained results for this experiment are shown in Table 2.

Experiment #3 presents a comparison of our best approaches with the method reported in [23] in-house implemented and are expressed in Table 3.

Table 1. Obtained results using SERI datasets.

Features	8^{riu2}		16^{riu2}		24^{riu2}	
	SE	SP	SE	SP	SE	SP
LBP	43.75	43.75	37.50	50.00	50.00	62.50
LBP-TOP	56.25	62.50	**87.50**	**75.00**	68.75	68.75
LBP+PCA	50.00	62.50	56.25	37.50	68.75	68.75
LBP+BoW	50.00	81.25	57.50	68.75	50.00	50.00
LBP+BoW+P	75.00	87.50	**81.25**	**75.00**	68.75	62.5
LBP-TOP+BoW+P	62.50	68.75	56.25	37.50	37.50	43.75

Table 2. Obtained results using Duke datasets.

Features	8^{riu2}		16^{riu2}		24^{riu2}	
	SE	SP	SE	SP	SE	SP
LBP-TOP	80.00	93.33	73.33	86.67	73.33	86.67
LBP+BoW+P	80.00	86.67	86.67	100	93.33	86.67
LBP-TOP+BoW+P	80.00	86.67	86.67	86.67	60.00	80.00

Table 3. Comparing the proposed method by [23] on SERI and Duke datasets.

Data sets	SERI		Duke	
	SE	SP	SE	SP
Venhuizen *et al.* [23]	61.53	58.82	71.42	68.75
{LBP+BoW+P},16^{riu2}	81.25	75.00	86.67	100.00
{LBP-TOP},16^{riu2}	87.50	75.00	73.33	86.76

5 Conclusions

The work presented here addresses the automatic classification of SD-OCT data to identify subjects with DME versus normal. Based on the reported results, the low level volume 3D features and high level 2D features using patches achieve the most desirable results in the experimental setup presented here. The comparison against different datasets and methodologies, highlights that: regardless of using low or high level representations, volume signatures derived from LBP texture show high discriminative power for distinguishing DME vs normal volumes.

References

1. Abramoff, M.D., Garvin, M.K., Sonka, M.: Retinal image analysis: a review. IEEE Review Biomed. Eng. 3, 169–208 (2010)
2. Becker, C., Rigamonti, R., Lepetit, V., Fua, P.: Supervised feature learning for curvilinear structure segmentation. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013, pp. 526–533. Springer (2013)
3. Breiman, L.: Random forests. Machine learning 45(1), 5–32 (2001)
4. Buades, A., Coll, B., Morel, J.M.: A non-local algorithm for image denoising. In: Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on. vol. 2, pp. 60–65. IEEE (2005)
5. Byrd, R.H., Nocedal, J., Schnabel, R.B.: Representations of quasi-newton matrices and their use in limited memory methods. Mathematical Programming 63(1-3), 129–156 (1994)
6. Chen, T.C., Cense, B., Pierce, M.C., Nassif, N., Park, B.H., Yun, S.H., White, B.R., Bouma, B.E., Tearney, G.J., de Boer, J.F.: Spectral domain optical coherence tomography: ultra-high speed, ultra-high resolution ophtalmic imaging. Arch. Ophtalmol. 123(12), 1715–1720 (2005)

7. Chiu, S.J., Li, X.T., Nicholas, P., Toth, C.A., Izatt, J.A., Farsiu, S.: Automatic segmentation of seven retinal layers in sd-oct images congruent with expert manual segmentation. Optic Express 18(18), 19413–19428 (2010)
8. Coupe, P., Hellier, P., Kervrann, C., Barillot, C.: Nonlocal means-based speckle filtering for ultrasound images. IEEE TIP pp. 2221–2229 (Oct 2009)
9. Cox, D.R.: The regression analysis of binary sequences. Journal of the Royal Statistical Society. Series B (Methodological) pp. 215–242 (1958)
10. Early Treatment Diabetic Retinopathy Study Group: Photocoagulation for diabetic macular edema: early treatment diabetic retinopathy study report no 1. Arch. Ophtalmol. 103(12), 1796–1806 (1985)
11. Kafieh, R., Rabbani, H., Abramoff, M.D., Sonka, M.: Intra-retinal layer segmentation of 3d optical coherence tomography using coarse grained diffusion map. Medical Image Analysis 17, 907–928 (2013)
12. Lemaître, G., Rastgoo, M., Massich, J.: retinopathy: Miccai-omia-2015 (Jul 2015), http://dx.doi.org/10.5281/zenodo.22195
13. Liu, Y.Y., Chen, M., Ishikawa, H., Wollstein, G., Schuman, J.S., M., R.J.: Automated macular pathology diagnosis in retinal oct images using multi-scale spatial pyramid and local binary patterns in texture and shape encoding. Medical Image Analysis 15, 748–759 (2011)
14. Murphy, K.P.: Machine learning: a probabilistic perspective. MIT press (2012)
15. Ojala, T., Pietikäinen, M., Mäenpää, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. Pattern Analysis and Machine Intelligence, IEEE Transactions on 24(7), 971–987 (2002)
16. Quellec, G., Lee, K., Dolejsi, M., Garvin, M.K., Abramoff, M.D., Sonka, M.: Three-dimensional analysis of retinal layer texture: identification of fluid-filled regions in sd-oct of the macula. IEEE Trans. on Medical Imaging 29, 1321–1330 (2010)
17. Schmitt, J.M., Xiang, S., Yung, K.M.: Speckle in optical coherence tomography. Journal of biomedical optics 4(1), 95–105 (1999)
18. Sharma, S., Oliver-Hernandez, A., Liu, W., Walt, J.: The impact of diabetic retinopathy on health-related quality of life. Curr. Op. Ophtal. 16, 155–159 (2005)
19. Sivic, J., Zisserman, A.: Video google: a text retrieva approach to object matching in videos. In: IEEE ICCV. pp. 1470–1477 (2003)
20. Srinivasan, P.P., Kim, L.A., Metttu, P.S., Cousins, S.W., Comer, G.M., Izatt, J.A., Farsiu, S.: Fully automated detection of diabetic macular edema and dry age-related macular degeneration from optical coherence tomography images. Biomedical Optical Express 5(10), 3568–3577 (2014)
21. Trucco, E., Ruggeri, A., Karnowski, T., Giancardo, L., Chaum, E., Hubschman, J., al Diri, B., Cheung, C., Wong, D., Abramoff, M., Lim, G., Kumar, D., Burlina, P., Bressler, N.M., Jelinek, H.F., Meriaudeau, F., Quellec, G., MacGillivray, T., Dhillon, B.: Validation retinal fundus image analysis algorithms: issues and proposal. Investigative Ophthalmology & Visual Science 54(5), 3546–3569 (2013)
22. Vapnik, V., Lerner, A.: Generalized portrait method for pattern recognition. Automation and Remote Control 24(6), 774–780 (1963)
23. Venhuizen, F.G., van Ginneken, B., Bloemen, B., van Grisven, M.J.P.P., Philipsen, R., C., H., Theelen, T., Sanchez, C.I.: Automated age-related macular degeneration classification in oct using unsupervised feature learning. In: SPIE Medical Imaging. vol. 9414, p. 94141l (2015)
24. Wild, S., Roglic, G., Green, A., Sicree, R., King, H.: Global prevalence of diabetes estimates for the year 2000 and projections for 2030. Diabetes Care 27(5), 1047–1053 (2004)

25. Zhao, G., Ahonen, T., Matas, J., Pietikäinen, M.: Rotation-invariant image and video description with local binary pattern features. Image Processing, IEEE Transactions on 21(4), 1465–1477 (2012)
26. Zheng, Z., Zha, H., Zhang, T., Chapelle, O., Chen, K., Sun, G.: A general boosting method and its application to learning ranking functions for web search. In: Advances in neural information processing systems. pp. 1697–1704 (2007)