

Document downloaded from:

<http://hdl.handle.net/10251/65172>

This paper must be cited as:

Morales, S.; Engan, K.; Naranjo Ornedo, V.; Colomer, A. (2015). Retinal Disease Screening through Local Binary Patterns. IEEE Journal of Biomedical and Health Informatics. (99):1-8. doi:10.1109/JBHI.2015.2490798.



The final publication is available at

<http://dx.doi.org/10.1109/JBHI.2015.2490798>

Copyright Institute of Electrical and Electronics Engineers (IEEE)

Additional Information

“© 2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.”

Retinal Disease Screening through Local Binary Patterns

Sandra Morales, Kjersti Engan, Valery Naranjo and Adrián Colomer

Abstract—This work investigates discrimination capabilities in the texture of fundus images to differentiate between pathological and healthy images. For this purpose, the performance of Local Binary Patterns (LBP) as a texture descriptor for retinal images has been explored and compared with other descriptors such as LBP filtering (LBP_F) and local phase quantization (LPQ). The goal is to distinguish between diabetic retinopathy (DR), age-related macular degeneration (AMD) and normal fundus images analysing the texture of the retina background and avoiding a previous lesion segmentation stage. Five experiments (separating DR from normal, AMD from normal, pathological from normal, DR from AMD and the three different classes) were designed and validated with the proposed procedure obtaining promising results. For each experiment, several classifiers were tested. An average sensitivity and specificity higher than 0.86 in all the cases and almost of 1 and 0.99, respectively, for AMD detection were achieved. These results suggest that the method presented in this paper is a robust algorithm for describing retina texture and can be useful in a diagnosis aid system for retinal disease screening.

Index Terms—Local Binary Patterns, Diabetic Retinopathy, Age-related Macular Degeneration, AMD, Diagnosis Aid System, Fundus Image, Retinal Image.

I. INTRODUCTION

The World Health Organization (WHO) estimates that in 2010 there were 285 million people visually impaired around the world [1]. In spite of the fact that the number of blindness cases has been significantly reduced in recent years, it is estimated that 80% of the cases of visual impairment are preventable or treatable [1].

Diabetic retinopathy (DR) and age-related macular degeneration (AMD) are nowadays two of the most frequent causes of blindness and vision loss [2]. In addition, these diseases will experience a high growth in the future due to diabetes incidence increase and ageing population in the current society. Their early diagnosis allows, through appropriate treatment, to reduce costs generated when they are in advanced states and may become chronic. This fact justifies screening campaigns.

Manuscript submitted May 4, 2015. This work was supported by NILS Science and Sustainability Programme (010-ABEL-IM-2013) and by the Ministerio de Economía y Competitividad of Spain, Project ACRIMA (TIN2013-46751-R).

S. Morales, V. Naranjo and A. Colomer are with the Instituto Interuniversitario de Investigación en Bioingeniería y Tecnología Orientada al Ser Humano, Universitat Politècnica de València, I3BH/LabHuman, Camino de Vera s/n, 46022 Valencia, Spain, and also with the Grupo Tecnologías de Informática Aplicadas a la Oftalmología, Unidad Conjunta UPV- FIS-ABIO, Spain (e-mail: smorales@labhuman.com; vnaranjo@labhuman.com; acolomer@labhuman.com). The work of A. Colomer has been supported by the Spanish Government under the FPI Grant BES-2014-067889.

K. Engan is with the Department of Electrical Engineering and Computer Science, University of Stavanger, N-4036 Stavanger, Norway (e-mail: kjersti.engan@uis.no).

However, a screening campaign requires a heavy workload for trained experts in the analysis of anomalous patterns of each disease which, added to the at-risk population increase, makes these campaigns economically infeasible. Therefore, the need for automatic screening systems is highlighted.

Based on these facts, a computer-aided diagnosis software capable of discriminating, through image processing, between a healthy fundus (without any pathology) and DR and AMD patients was developed. Thanks to high resolution of digital fundus images, they can be automatically processed providing invaluable help to clinicians in early diagnosis and disease prevention. Specifically, the final aim of the software proposed in this paper is to be used in an automatic screening of these diseases making the at-risk population assessment possible.

DR and AMD can be characterized by the presence of specific types of retinal lesions such as microaneurysms, exudates or drusen, among others. Figure 1 depicts some examples of these diseases in comparison with the fundus image from a healthy patient.

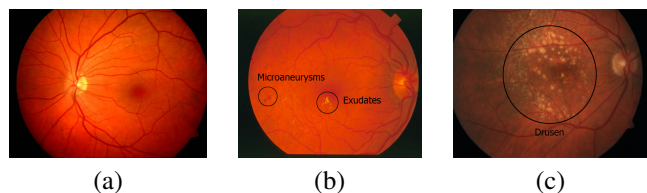


Fig. 1. Fundus images. (a) Healthy, (b) DR (with microaneurysms and exudates) and (c) AMD (with drusen).

This paper investigates discrimination capabilities in the texture of fundus to differentiate between pathological and healthy images. In particular, the main focus lies in exploring the performance of Local Binary Patterns (LBP) as a texture descriptor for retinal images. LBP technique has been given a lot of attention in recent years [3], [4]. It is based on looking at the local variations around each pixel, and assigning labels to different local patterns. Thereafter, the distribution of the labels is evaluated and used in the classification stage. There are many examples of the success of LBP used to describe and classify textures in general [5]–[7] and also in the case of medical imaging [8]–[10]. However, regarding fundus image processing, LBP have not been widely used. Most state-of-the-art works that use the LBP technique on fundus images focus on the segmentation of the retinal vessels [11], [12] rather than on a full diagnosis system, although some examples can be found in this direction [13]–[15]. In Mookiah *et al.* [13] abnormal signs were extracted from fundus images to detect normal fundus and two DR stages. Thirteen features, such

as area of hard exudates, area of blood vessels, bifurcation points, texture and entropies, fed three different classifiers (Probabilistic Neural Network (PNN), Decision Tree C4.5 and Support Vector Machine (SVM)). The texture is found by LBP and Laws energy. A previous segmentation of the exudates, optic disc and blood vessels is needed for feature extraction. The experiments are conducted on 156 subjects and the PNN is chosen as the best classifier with three-fold cross validation [13]. In more recent work of Mookiah *et al.* [16], a different methodology for AMD characterization is done through local configuration patterns (LCP) rather than by LBP. Linear configuration coefficients and pattern occurrence features are extracted and a linear SVM is used after feature selection. Krishnan and Laude combine LBP with entropies and invariant moments to generate an integrated index for diabetic retinopathy diagnosis. They demonstrated that there exist significant differences in the index for normal images and DR images and they emphasized that lesion segmentation was not required [14]. Garnier *et al.* deal with the AMD detection using LBP. The texture information on several scales is analysed through a wavelet decomposition and a LBP histogram is found from the wavelet coefficients. Linear Discriminant Analysis (LDA) is used for feature dimension reduction using the values of the entire LBP histogram as input features. Image classification on a set of 45 images is evaluated with a leave-one-out validation method [15].

The goal of this paper is to distinguish between DR, AMD and normal fundus images at the same time and avoiding any previous segmentation stage of retinal lesions. The texture of the retina background is directly analysed by means of LBP, and only this information is used to differentiate healthy patients and these two pathologies. A comprehensive study about what type of classifier obtains the best results is also undertaken. The performance of Logistic Regression, Neural Networks, SVM, Naive Bayes, J48, Rotation Forest, Random Forest and AdaBoost M1 is compared. This approach is different from previous works that use LBP. Mookiah *et al.* [13] require the segmentation of exudates in addition to segmentation of main structures (optic disc and vessels) for feature extraction and, although three different classes are identified, they only focus on DR detection. Krishnan and Laude and Garnier *et al.* [14], [15] do not need previous segmentations but only handle with a disease at time, in particular with DR and AMD diagnosis, respectively. Moreover, Krishnan and Laude did not provide values to determine the accuracy of the normal and DR discrimination.

Many operators for texture description have been defined in the literature. Some of them are modifications of the original LBP such as completed LBP (CLBP) [17], LBP filtering (LBPf) [18], dominant LBP (DLBP) [19], etc. Other state-of-the-art descriptors are completely different as Weber local descriptor (WLD) [20], local contrast patterns (LCPs) [21] or local phase quantization (LPQ) [22]. The LBP have been seen to be useful in many applications, and is simple and easy to compute. For these reasons we wanted to explore LBP for the present application. For comparison, experiments using LBPf and LPQ are presented as well.

The rest of the paper is organized as follows: in Section

II materials and methods are described and in Section III the proposed method is presented. Section IV shows how system validation was performed and as well as the obtained results. Finally, Section V provides discussion and Section VI conclusions and some future areas for work.

II. MATERIALS AND METHODS

A. Material

The material of this work are images previously diagnosed as normal (without known pathology), DR or AMD. The dataset used was composed of images from 4 different databases which included some of the categories under study: ARIA [23]–[25], STARE [26]–[28], E-OPHTHA [29], [30] and DIAGNOS [15]. ARIA database is formed by 143 colour fundus images (768×576 pixels), which are organised into three classes: age-related macular degeneration (AMD) subjects ($n=23$), healthy control-group subjects ($n=61$), and diabetic subjects ($n=59$). Trained image analysis experts have traced out the blood vessels, and also the optic disc and fovea where relevant. STARE database is a full set of 402 images (700×605 pixels) where thirteen different diagnoses were considered. From this dataset, three subsets were generated: age-related macular degeneration ($n=47$), normal ($n=37$), and diabetic retinopathy ($n=89$). E-OPHTHA is a database of fundus images especially designed for diabetic retinopathy screening. It contains 257 images with no lesion, 47 images with exudates and 148 with microaneurysms or small hemorrhages making a total of 174 images with diabetic retinopathy. Finally, DIAGNOS is a private database, property of DIAGNOS Inc., composed of 45 fundus images, 22 afflicted with AMD and 23 healthy. Both E-OPHTHA and DIAGNOS have a range of different image resolutions. The four databases experience a significant variability in color, illumination, resolution, quality, etc. both within and, even more, among the databases.

All images of the resulting dataset must comply with certain quality criteria. The following causes were considered reasons for exclusion:

- Images with severe artefacts, for example bright and circular spots produced by some dust in the camera lens.
- Images affected by a relative large amount of impulsive noise (salt and pepper noise).
- Vascular network is largely over-segmented by the method presented in [31].¹
- Images with a doubtful diagnosis.²
- Images with highlights around the vessels associated with young retinas.³
- Tessellated images due to the fact there are lesser amounts of pigment in the retinal pigment epithelium.³

Figure 2 depicts some of these cases. Most of these choices were done to determine if LBP were able to discriminate between healthy and pathological images in a normal situation or, in other words, without the presence of distracting elements.

¹Further explanation in Section III-B.

²Based on second opinion from a medical doctor.

³This changes the images dramatically, thus they are not included here. However they should be regarded separately in future studies since it is highly desirable to be able to include these types of patients.

If this hypothesis is confirmed, the method will be expanded in future work to include images of different appearance, for example the tessellated images.

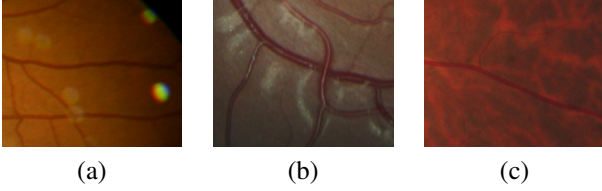


Fig. 2. Excluded images. (a) With artefacts, (b) With highlights and (c) Tessellated.

After exclusion, the resulting dataset used in this work is formed by a total of 251 images. This dataset was divided into two subsets, one for training and testing by cross validation (model set) and other purely for testing (validation set). The model set contains 80% of the images and the validation set the remaining 20%. A list of the images that compose the resulting dataset is made available at http://158.42.170.205/flexshare/acrima/Resulting_dataset to facilitate future fair comparisons. However, some images belong to a private database (DIAGNOS) and they are not publicly available. Table I details the number of images of each database after the exclusion criteria and also the content of each subset.

TABLE I
CONTENT OF EACH DATABASE: MODEL AND VALIDATION SET.

	ARIA	STARE	E-OPHTHA	DIAGNOS	Total
AMD	9	23	0	17	49
DR	8	35	37	0	80
NORMAL	30	13	79	0	122
<i>Total</i>	47	71	116	17	251
<i>Model set</i>					
	ARIA	STARE	E-OPHTHA	DIAGNOS	Total
AMD	7	18	0	14	39
DR	6	28	29	0	63
NORMAL	24	10	63	0	97
<i>Total</i>	37	56	92	14	199
<i>Validation set</i>					
	ARIA	STARE	E-OPHTHA	DIAGNOS	Total
AMD	2	5	0	3	10
DR	2	7	8	0	17
NORMAL	6	3	16	0	25
<i>Total</i>	10	15	24	3	52

B. Local binary patterns

Local binary patterns (LBP) are a powerful grey-scale texture operator used in many computer vision applications because of its computation simplicity [3], [4]. The first step in LBP is to produce a label for each pixel in the image where the label is found based on the local neighbourhood of the pixel which is defined by a radius, R , and a number of points, P . The neighbouring pixels are thresholded with respect to the grey value of the central pixel of the neighbourhood generating a binary string or, in other words, a binary pattern. The value of a LBP label is obtained for every pixel by summing the binary string weighted with powers of two as follows:

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

where g_p and g_c are the grey values of the neighbourhood and central pixel, respectively. P represents the number of samples on the symmetric circular neighbourhood of radius R . The g_p values are interpolated to fit with a given R and P . The values of the labels depend on the size of the neighbourhood (P).

2^P different binary patterns can be generated in each neighbourhood. However, the bits of these patterns must be rotated to the minimum value to achieve a rotation invariant pattern. In the case of $P = 8$, only 36 of the 2^P possible patterns are rotation invariant, i.e., $LBP_{8,R}$ can have 36 different values. Figure 3 shows how LBP are calculated for a circular neighbourhood of radius 1 ($R = 1$) and 8 samples ($P = 8$).

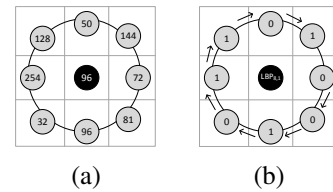


Fig. 3. LBP computation: (a) Grey values of a circular neighbourhood of radius 1 and 8 samples. (b) Thresholding between the grey value of the neighbourhood and the central pixel. The rotation invariant local binary pattern generated is 00101101 (the arrows indicate the order in which the string is formed). Specifically, the LBP label is obtained as follows $LBP_{8,1} = 0 \times 2^0 + 0 \times 2^1 + 1 \times 2^2 + 0 \times 2^3 + 1 \times 2^4 + 1 \times 2^5 + 0 \times 2^6 + 1 \times 2^7 = 180$.

When LBP are used for texture description, it is common to include a contrast measure by defining the rotational invariant local variance as follows:

$$VAR_{P,R} = \frac{1}{P} \sum_{p=0}^{P-1} (g_p - \mu)^2, \quad \mu = \frac{1}{P} \sum_{p=0}^{P-1} g_p. \quad (2)$$

The LBP and VAR measures are complementary and are combined to enhance the performance of the LBP operator. The implementation of both measures is publicly available online in http://www.ee.oulu.fi/~gyzhao/LBP_Book.htm.

III. PROPOSED METHOD

An algorithm for retina image classification without the need for prior segmentation of suspicious lesions was developed. Manual lesion segmentation is time consuming and automatic segmentation algorithms might not be accurate, thus removing the need for lesion segmentation can make the classification more robust. The algorithm is mainly based on the texture analysis of the retina background by means of LBP.

A. Pre-processing

Due to the fact that the images under study belong to different databases, the size of the images varies. As the LBP and VAR values depend on the radius of the neighbourhood, the images must be resized to a standardized size to obtain comparable texture descriptors. The images are resized using the length of the horizontal diameter of the fundus as reference [32]. Bicubic interpolation is used for resizing; the output pixel

value is a weighted average of pixels in the nearest 4-by-4 neighbourhood. Before feature extraction, a median filter for noise reduction is performed using a 3-by-3 neighbourhood.

Only the pixels of the retina background are considered significant for the texture analysis. Thus the main structures of the fundus (the vascular network and the optic disc), which are not related to the diseases under study, should not be taken into account when the fundus texture is analysed. Some preliminary tests showed that if these predominant structures were included in the texture analysis, the differences between healthy and pathological images were not appreciated due to the similar aspect of these structures. In Figure 4 the main structures present in a fundus image are identified. The optic disc and the vascular network are detected by our own methods. The method used for optic disc detection is mainly based on principal component analysis along with mathematical morphology operations such as stochastic and stratified watershed and geodesic transformations [33]. The algorithm for vessel segmentation combines the use of basic mathematical morphology operations with curvature evaluation [31]. The external mask is directly obtained by thresholding.

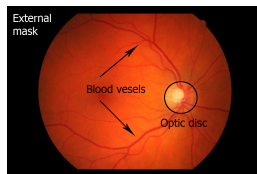


Fig. 4. Structures present in a fundus image.

Global masks including both external and structure masks for all images are publicly available in http://158.42.170.205/flexshare/acrima/Resulting_dataset. The external masks were eroded to ensure that the LBP neighbourhood was contained within the field of view of the retinal camera. The structure masks were dilated to avoid that the vessels and the optic disc were included in the LBP neighbourhood.

B. Feature extraction

The LBP and VAR operators described above are used to characterize the texture of the retina background. They are calculated for each pixel of the RGB images using $P = 8$ and different values of R ($R = \{1, 2, 3, 5\}$). The LBP and VAR values corresponding to pixel positions of the optic disc, vessels or outside the fundus are not considered.

The red, green and blue components of each image are independently analysed. One example of the aspect of the LBP and VAR images of an AMD fundus is depicted in Figure 5.

The resulting LBP and VAR images provide a description of the image texture. After masking the optic disc and vessel segments, the LBP and VAR values within the external mask of the fundus are collected into histograms, one for each color (RGB). Different statistical information is extracted from these histograms to use it as features in the classification stage. Concretely, the calculated statistical values are: mean, standard deviation, median, entropy, skewness and kurtosis. To sum up, 6 statistical values are calculated from each LBP and VAR

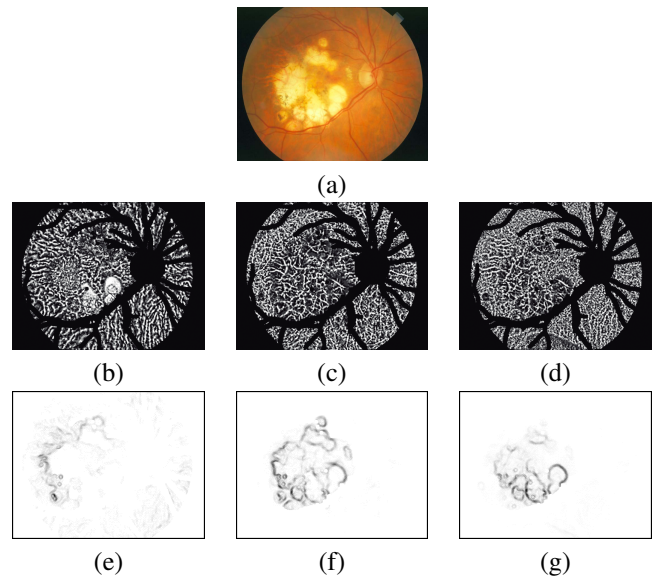


Fig. 5. Feature extraction using $P = 8$ and $R = 5$. (a) AMD fundus image, (b-d) LBP images calculated on R, G and B components, respectively. Optic disc and vessel segments are removed (black). (e-g) VAR images calculated on R, G and B components, respectively. Optic disc and vessel segments are removed (white).

histogram, giving place to 12 features for each radius used. Consequently, the total number of features is equal to 144 (12 features \times 4 radius \times 3 components). Figure 6 depicts the feature extraction flowchart.

C. Classification

Once the features are extracted, the data of the model set must be preprocessed before the classification stage. In the pre-processing, two tasks are carried out: data normalization and data resampling. The first one because the range of values of raw data varies widely and the second one because the dataset is clearly unbalanced and most machine learning algorithms would not work properly. In particular, the method used for the normalization is to standardize all numeric attributes in the given dataset to have zero mean and unit variance and, for the resampling, the Synthetic Minority Oversampling TEchnique (SMOTE) [34] is applied.

Afterwards, external cross validation (CV) [35], [36], also called nested CV, is performed on the model set so that the dimensionality of the data is reduced by feature selection before being passed on to a classifier. 10 folds are used in the external loop and 5 folds in the internal loop. The purpose of the internal loop is to select a feature subset and the used technique is a wrapper method [37] with forward (best first) selection. The same type of classifier is used in both the internal and external loops. The external loop divides the set into 10 non-overlapping pairs of training (90%) and test (10%) sets. For each fold of the external CV, the training set is further divided into 5 non-overlapping sets by the internal CV loop. The internal loop is done first to select the feature subset of this particular fold of the external loop. Thereafter, the external loop trains the classifier using this subset, and tests it on the remaining 10%. This is repeated for every fold. Notice that

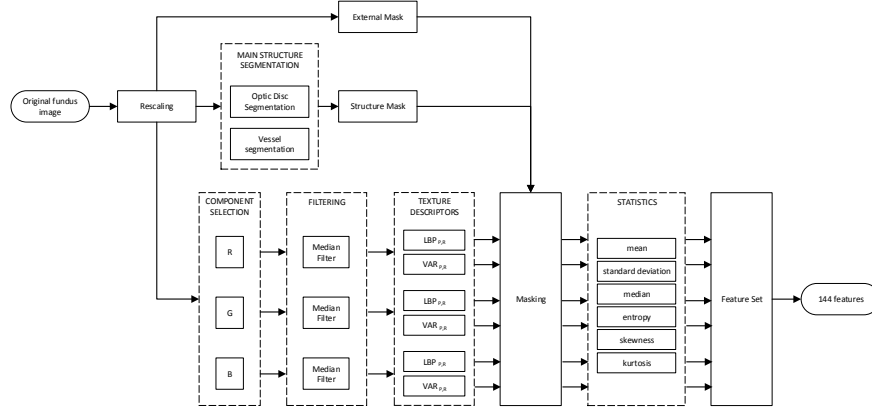


Fig. 6. Feature extraction flowchart. First, the original fundus image is rescaled based on its horizontal diameter. Secondly, LBP and VAR measures are performed on the three RGB components using $P = 8$ and $R = \{1, 2, 3, 5\}$. Then, the external mask of the fundus is determined and the optic disc and vessels are segmented generating a structure mask. Both, external and structure masks, are used for masking the result of the texture descriptors. The final feature set is formed by the statistical values of all texture descriptors after masking.

the feature set might vary with each external fold of the CV scheme. Thus doing an external or nested CV gives a measure of how well the method works for this dataset, where the method includes the feature subset selector and the choice of classifier. Figure 7 shows how the external CV is performed for the first fold.

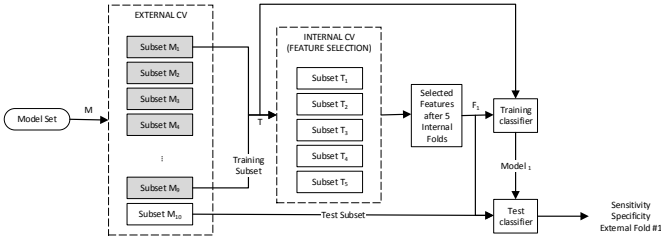


Fig. 7. External cross validation flowchart. The process is repeated iteratively for the 10 folds of the external loop.

Finally, a final classifier is made using the whole model set for feature subset selection and thereafter the whole model set is used for training the classifier. The validation set is tested on the final classifier. The process is summarized in Figure 8. The normalization parameters from the model set are saved as a part of the classifier, such that the validation set is normalized using these same parameters.

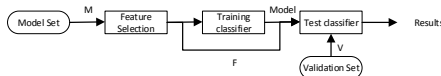


Fig. 8. Final validation flowchart.

Both, data preprocessing and classification, were carried out with Weka [38].

IV. RESULTS

Five experiments were conducted and validated with the proposed procedure: AMD - Normal, DR - Normal, Pathological (including AMD and DR) - Normal, AMD - DR and 3 class problem (AMD - DR - Normal).

For each experiment, different classifiers were tested: Logistic Regression [39], Neural Networks [40], SVM [41], Naive Bayes [42], C4.5 [43], Rotation Forest [44], Random Forest [45], and AdaBoost [46]. The performance of the algorithms was evaluated based on two concepts: sensitivity or true positive rate (TPR) and specificity or true negative rate (TNR). Sensitivity and specificity measure the proportion of positive and negative cases which are correctly identified as such, respectively. Table II details the sensitivity and specificity obtained on the model and validation sets in all the experiments with the different classifiers. In the case of the three-class problem, the average sensitivity and specificity are shown. The metrics computed after external cross validation are additive over the folds, i.e. the correctly or incorrectly classified cases are summed over the folds and divided by the total number of instances. The best results of each experiment on the model set are highlighted in bold.

This work makes use of the LBP operator but many others texture descriptors exist as mentioned in Section I. In particular, the performance of LBP was compared with two methods for texture classification: LBP filtering (LBPf) [18] and local phase quantization (LPQ) [22]. LBPf is a multi-resolution filtered version of the LBP that combines exponentially growing circular neighbourhoods with Gaussian low-pass filtering with the aim of avoiding aliasing effects caused by sparse sampling. The size of the Gaussian filters is increased according to R and P LBP values. LPQ is a blur insensitive method that is based on phase information of the discrete Fourier transform computed locally for every pixel. LPQ features were calculated using $n \times n$ windows ($n = \{3, 5, 7, 9\}$). Table III compares the results of the proposed method with those achieved by LBPf and LPQ. Only the texture descriptor of the feature extraction stage was modified in the procedure.

In machine learning, there exist two main approaches for dimensionality reduction. This is usually performed by selecting a subset of the original features or by constructing new features to replace the original ones. In Table IV two methods for dimensionality reduction were compared: feature selection

TABLE II
RESULTS OF THE DIFFERENT EXPERIMENTS ON THE MODEL SET (EXTERNAL CV) AND VALIDATION SET.

	<i>AMD-Normal</i>		<i>DR-Normal</i>		<i>Path.-Normal</i>		<i>AMD-DR</i>		<i>AMD-DR-Normal</i>		
	TPR	TNR	TPR	TNR	TPR	TNR	TPR	TNR	TPR	TNR	
MODEL SET	Logistic Regression	0.959	0.979	0.732	0.845	0.755	0.891	0.984	0.968	0.859	0.930
	Neural Networks	1.000	0.979	0.753	0.835	0.716	0.921	1.000	0.984	0.873	0.936
	SVM	1.000	0.990	0.784	0.928	0.775	0.901	0.937	0.984	0.876	0.938
	Naive Bayes	1.000	0.990	0.629	0.835	0.647	0.871	0.984	0.952	0.811	0.905
	C4.5	0.990	0.979	0.876	0.835	0.853	0.871	1.000	0.921	0.883	0.942
	Rotation Forest	0.990	0.969	0.825	0.856	0.873	0.931	0.984	0.952	0.890	0.945
	Random Forest	0.990	0.990	0.856	0.897	0.873	0.881	1.000	0.937	0.883	0.942
	AdaBoost	0.990	0.990	0.835	0.825	0.902	0.881	0.937	0.921	0.856	0.928
VALIDATION SET	Logistic Regression	1.000	1.000	0.706	0.960	0.741	0.800	1.000	0.882	0.865	0.902
	Neural Networks	1.000	0.960	0.765	0.920	0.815	0.960	1.000	0.941	0.750	0.816
	SVM	1.000	1.000	0.588	0.920	0.704	0.840	0.900	1.000	0.788	0.844
	Naive Bayes	1.000	0.960	0.471	0.880	0.704	0.880	1.000	1.000	0.673	0.814
	C4.5	1.000	1.000	0.588	0.840	0.741	0.920	1.000	1.000	0.769	0.837
	Rotation Forest	0.900	1.000	0.706	0.960	0.889	1.000	0.900	1.000	0.827	0.878
	Random Forest	1.000	1.000	0.588	0.920	0.815	0.880	1.000	1.000	0.769	0.842
	AdaBoost	1.000	1.000	0.588	0.960	0.815	0.880	0.900	0.941	0.769	0.842

TABLE III
COMPARISON WITH OTHER TEXTURE DESCRIPTORS.

	<i>AMD-Normal</i> (SVM)		<i>DR-Normal</i> (Random Forest)		<i>Path.-Normal</i> (AdaBoost)		<i>AMD-DR</i> (Neural Networks)		<i>AMD-DR-Normal</i> (Rotation Forest)	
	TPR	TNR	TPR	TNR	TPR	TNR	TPR	TNR	TPR	TNR
LPBF [18]	1.000	1.000	0.835	0.887	0.843	0.851	1.000	0.952	0.911	0.955
LPQ [22]	0.907	0.887	0.649	0.701	0.667	0.713	0.825	0.778	0.729	0.864
Proposed method	1.000	0.990	0.856	0.897	0.902	0.881	1.000	0.984	0.890	0.945

through wrapper strategy, as used in the other experiments of the paper, and principal component analysis (PCA) to transform the data from a high-dimensional space to a space of fewer dimensions. Only the method for dimensionality reduction was changed, the rest of the procedure was remained unaltered. The best classifier for each experiment was used. Such an approach will be repeated in the following tests.

The proposed method combines the features extracted from the three RGB components. This can be justified by comparing the proposed method of all three components with the corresponding method but used on only one color channel at the time. From the results depicted in Table V, it is seen that combining the color information in general performs slightly better than the best color band (green).

The proposed method was also compared with other state-of-the-art algorithms which are based on fundus image texture analysis for aided diagnosis. To the best of the author's knowledge, there is no other system that analyses the texture of the retina background and detects AMD and DR at the same time, therefore, it was only possible to compare the results from two class diagnosis, see Table VI. The results of the proposed method shown in this table are those achieved on the model set because the other compared works did not test their algorithms on an independent validation set. Note that the obtained results are not directly comparable because the dataset used in each case is different.

V. DISCUSSION

Many more AMD or DR detection techniques exist in the literature but most of them focus on lesion segmentation instead of a study of the retina background. This fact makes the

accuracy of the classification stage dependent on the accuracy of the lesion segmentation. Lesion segmentation involves a series of uncertainties and a non accurate segmentation may provoke important errors in the classification. The main advantage of the procedure proposed in this paper is that it gets a good performance without having to search different types of lesions. The only needed segmentation in the presented approach is to mask the significant structures (vessels and optic disc) but their accuracy has little influence on the final result.

With regard to the obtained results, the experiment with the best performance is "AMD-Normal" achieving a sensitivity and specificity greater than 0.99 and improving the results of other state-of-the-art methods. The rest of the experiments show that the lower values correspond to the detection of the DR. This is because the lesions of the DR are usually smaller than those of the AMD and, therefore, they are more difficult to detect only through texture analysis. Even so, the results are promising in the three-class diagnosis.

In addition to the results of the external CV, an independent set of 52 images was saved to test the proposed method. The validation results follow the same trend shown in the CV, i.e. the "AMD-Normal" experiment achieves the best performance.

To analyze the robustness of the LBP for describing fundus images, LBP were compared with other texture descriptors. From Table III it is seen that LBP works better than LPBF and LPQ descriptors in most of the experiments. Theoretically, the sparse sampling exploited by LBP operators with large neighbourhood radii may not result in an adequate representation of the image due to aliasing effects. For that reason, LPBF was expected to be a better solution. However, aliasing effects are not too relevant in the proposed procedure because

TABLE IV
COMPARISON OF DIFFERENT METHODS FOR DIMENSIONALITY REDUCTION.

	<i>AMD-Normal (SVM)</i>		<i>DR-Normal (Random Forest)</i>		<i>Path.-Normal (AdaBoost)</i>		<i>AMD-DR (Neural Networks)</i>		<i>AMD-DR-Normal (Rotation Forest)</i>	
	TPR	TNR	TPR	TNR	TPR	TNR	TPR	TNR	TPR	TNR
PCA	1.000	0.866	0.742	0.835	0.608	0.762	0.968	0.984	0.828	0.914
Proposed method	1.000	0.990	0.856	0.897	0.902	0.881	1.000	0.984	0.890	0.945

TABLE V
ANALYSIS OF THE RGB INFLUENCE.

	<i>AMD-Normal (SVM)</i>		<i>DR-Normal (Random Forest)</i>		<i>Path.-Normal (AdaBoost)</i>		<i>AMD-DR (Neural Networks)</i>		<i>AMD-DR-Normal (Rotation Forest)</i>	
	TPR	TNR	TPR	TNR	TPR	TNR	TPR	TNR	TPR	TNR
Red component	1.000	0.979	0.784	0.835	0.716	0.822	0.984	0.984	0.873	0.936
Green component	0.979	1.000	0.866	0.887	0.892	0.752	0.968	0.937	0.907	0.954
Blue component	0.990	0.959	0.825	0.876	0.716	0.861	0.762	0.889	0.808	0.904
Proposed method	1.000	0.990	0.856	0.897	0.902	0.881	1.000	0.984	0.890	0.945

TABLE VI
AMD-NORMAL AND DR-NORMAL COMPARISONS ON DIFFERENT DATABASES.

	TPR	TNR	Features	Classifier	Validation	#images	
AMD	Garnier <i>et al.</i> [15]	0.913	0.955	Wavelet transform and LBP	LDA	Leave-one-out	45
	Mookiah <i>et al.</i> [16]	0.980	0.975	LCP	SVM	10-fold CV	83
	Proposed method	1.000	0.990	LBP	SVM	10-fold external CV	136
DR	Mookiah <i>et al.</i> [13]	0.963	0.961	Blood vessel area, exudates area, bifurcation points, texture and entropy	PNN	3-fold CV	156
	Proposed method	0.856	0.897	LBP	Random Forest	10-fold external CV	160

vessels and OD are masked and, therefore, the presence of high frequencies in the image is significantly reduced. Moreover, it was demonstrated that the phase information utilized by LPQ is less representative than LBP for the characterization of AMD and DR fundus.

The influence of the method used for feature dimensionality reduction was also studied. The wrapper strategy used in the external CV is more time consuming than PCA but the sensitivity and specificity achieved in all the experiments is higher, as seen in Table IV. Therefore the wrapper-based strategy is considered a better way of doing feature selection.

It is well known in the literature that the green component of the fundus image provides a better visualization of the retinal structures compared to the other two color channels. However, it is shown in Table V that the use of the combined RGB features usually improves the results compared to using only the green component. The least significant component is sometimes the red and sometimes the blue one depending on the experiment.

VI. CONCLUSIONS

In this paper, a new approach for AMD and DR diagnosis was presented. It is based on analysing texture discrimination capabilities in fundus images to differentiate healthy patients from AMD and DR images. The performance of LBP along with different classifiers was tested and compared with other texture descriptors. The most important finding is that the proposed method is capable of discriminating the classes based on analysing the texture of the retina background, avoiding

previous segmentation of retinal lesions. Such lesion segmentation algorithms might be both time consuming and potential inaccurate, thus avoiding the segmentation is beneficial.

The obtained results demonstrate that using LBP as texture descriptor for fundus images provides useful features for retinal disease screening.

In future work, a larger test of the method with more images should be done. Moreover, some work should be carried out to develop strategies that enable the analysis of the type of images that were excluded from the initial database, such as tessellated fundus, images with highlights or typical artefacts. Other research line is to automatically determine the presence of biological image variation (tessellation, highlighting or other) prior to the classification step to train different classifiers and use different feature combinations for each specific case. We also wish to explore more texture descriptors. For example, the idea of LBP has been developed further into non-binary coding for texture description, and has provided good results recently [47]. In addition, recent literature describes new texture descriptors based on the co-occurrence method with promising results used on medical images [48].

ACKNOWLEDGEMENT

The authors would like to express their deep gratitude to Mickaël Garnier and DIAGNOS Inc. for facilitating access to the DIAGNOS dataset.

REFERENCES

- [1] World Health Organization (WHO), "Universal eye health: a global action plan 2014-2019," 2013.

- [2] World Health Organization (WHO), "Action plan for the prevention of avoidable blindness and visual impairment 2009-2013," 2010.
- [3] T. Ojala, M. Pietikinen, and T. Menp, "A generalized local binary pattern operator for multiresolution gray scale and rotation invariant texture classification," in *Advances in Pattern Recognition, 2nd International Conference on*, 2001, pp. 397–406.
- [4] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 7, pp. 971–987, 2002.
- [5] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 12, pp. 2037–2041, 2006.
- [6] M. Heikkil, M. Pietikinen, and C. Schmid, "Description of interest regions with local binary patterns," *Pattern Recognition*, vol. 42, no. 3, pp. 425 – 436, 2009.
- [7] Z. Yang and H. Ai, "Demographic classification with local binary patterns," in *Advances in Biometrics*, ser. Lecture Notes in Computer Science, S.-W. Lee and S. Li, Eds., 2007, vol. 4642, pp. 464–473.
- [8] L. Kotu, K. Engan, T. Eftestol, L. Woie, S. Orn, and A. Katsaggelos, "Local binary patterns used on cardiac MRI to classify high and low risk patient groups," in *Signal Processing Conference (EUSIPCO), Proceedings of the 20th European*, 2012, pp. 2586–2590.
- [9] K. Oppedal, K. Engan, D. Aarsland, M. Beyer, O. B. Tysnes, and T. Eftestol, "Using local binary pattern to classify dementia in MRI," in *Biomedical Imaging (ISBI), 9th IEEE International Symposium on*, May 2012, pp. 594–597.
- [10] L. Nanni, A. Lumini, and S. Brahmam, "Local binary patterns variants as texture descriptors for medical image analysis," *Artificial Intelligence in Medicine*, vol. 49, no. 2, pp. 117 – 125, 2010.
- [11] S. Zabih, M. Delgir, and H.-R. Pourreza, "Retinal vessel segmentation using color image morphology and local binary patterns," in *Machine Vision and Image Processing (MVIP), 6th Iranian*, 2010, pp. 1–5.
- [12] S. Dhanushkodi and M. Vasuki, "Diagnosis system for diabetic retinopathy to prevent vision loss," *Applied Medical Informatics*, vol. 33, no. 3, pp. 1–11, 2013.
- [13] M. Mookiah, U. R. Acharya, R. J. Martis, C. K. Chua, C. Lim, E. Ng, and A. Laude, "Evolutionary algorithm based classifier parameter tuning for automatic diabetic retinopathy grading: A hybrid feature extraction approach," *Knowledge-Based Systems*, vol. 39, no. 0, pp. 9 – 22, 2013.
- [14] M. M. R. Krishnan and A. Laude, "An integrated diabetic retinopathy index for the diagnosis of retinopathy using digital fundus image features," *Journal of Medical Imaging and Health Informatics*, vol. 3, no. 2, pp. 306–313, 2013.
- [15] M. Garnier, T. Hurtut, H. Ben Tahar, and F. Cheriet, "Automatic multiresolution age-related macular degeneration detection from fundus images," in *SPIE, Proceedings*, vol. 9035, 2014, pp. 903 532–903 532–7.
- [16] M. R. K. Mookiah, U. R. Acharya, H. Fujita, J. E. Koh, J. H. Tan, K. Noronha, S. V. Bhandary, C. K. Chua, C. M. Lim, A. Laude, and L. Tong, "Local configuration pattern features for age-related macular degeneration characterization and classification," *Computers in Biology and Medicine*, vol. 63, pp. 208 – 218, 2015.
- [17] Z. Guo, D. Zhang, and D. Zhang, "A completed modeling of local binary pattern operator for texture classification," *Image Processing, IEEE Transactions on*, vol. 19, no. 6, pp. 1657–1663, 2010.
- [18] T. Mäenpää and M. Pietikäinen, "Multi-scale binary patterns for texture analysis," in *Image Analysis*, ser. Lecture Notes in Computer Science, J. Bigun and T. Gustavsson, Eds., 2003, vol. 2749, pp. 885–892.
- [19] S. Liao, M. Law, and A. Chung, "Dominant local binary patterns for texture classification," *Image Processing, IEEE Transactions on*, vol. 18, no. 5, pp. 1107–1118, May 2009.
- [20] J. Chen, S. Shan, C. He, G. Zhao, M. Pietikainen, X. Chen, and W. Gao, "Wld: A robust local image descriptor," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, no. 9, pp. 1705–1720, 2010.
- [21] T. Song, H. Li, F. Meng, Q. Wu, B. Luo, B. Zeng, and M. Gabbouj, "Noise-robust texture description using local contrast patterns via global measures," *Signal Processing Letters, IEEE*, vol. 21, no. 1, pp. 93–96, 2014.
- [22] V. Ojansivu and J. Heikkil, "Blur insensitive texture classification using local phase quantization," in *Image and Signal Processing*, ser. Lecture Notes in Computer Science, A. Elmoataz, O. Lezoray, F. Nouboud, and D. Mammass, Eds., 2008, vol. 5099, pp. 236–243.
- [23] D. Farnell, F. Hatfield, P. Knox, M. Reakes, S. Spencer, D. Parry, and S. Harding, "Enhancement of blood vessels in digital fundus photographs via the application of multiscale line operators," *Franklin Institute, Journal of the*, vol. 345, no. 7, pp. 748 – 765, 2008.
- [24] Y. Zheng, M. H. A. Hijazi, and F. Coenen, "Automated disease / no disease grading of age-related macular degeneration by an image mining approach," *Investigative Ophthalmology & Visual Science*, vol. 53, no. 13, pp. 8310–8, 2012.
- [25] ARIA online, "Retinal image archive," 2006, http://www.eyecharity.com/aria_online.html. Last accessed on 18th September 2014.
- [26] A. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *Medical Imaging, IEEE Transactions on*, vol. 19, no. 3, pp. 203–210, 2000.
- [27] A. Hoover and M. Goldbaum, "Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels," *Medical Imaging, IEEE Transactions on*, vol. 22, no. 8, pp. 951–958, 2003.
- [28] STARE database, "STructure Analysis of the Retina," 2004, <http://www.ces.clemson.edu/~ahoover/stare/>. Last accessed on 18th September 2014.
- [29] E. Decencièrre, G. Cazuguel, X. Zhang, G. Thibault, J.-C. Klein, F. Meyer, B. Marcotegui, G. Quèllec, M. Lamard, R. Danno, D. Elie, P. Massin, Z. Viktor, A. Erginay, B. Lay, and A. Chabouis, "TeleOphta: Machine learning and image processing methods for teleophthalmology," *IRBM (Special issue: ANR TECSAN: Technologies for Health and Autonomy)*, vol. 34, no. 2, pp. 196–203, 2013.
- [30] E-Ophtha, "A color fundus image database," 2006, <http://www.adcis.net/en/Download-Third-Party/E-Ophtha.html>. Last accessed on 18th September 2014.
- [31] S. Morales, V. Naranjo, J. Angulo, J. J. Fuertes, and M. Alcañiz, "Segmentation and analysis of retinal vascular tree from fundus images processing," in *International Conference on Bio-inspired Systems and Signal Processing (BIOSIGNALS 2012)*, 2012, pp. 321 – 324.
- [32] X. Zhang, G. Thibault, E. Decencièrre, G. Quèllec, G. Cazuguel, A. Erginay, P. Massin, and A. Chabouis, "Spatial normalization of eye fundus images," in *ISBI 2012 : 9th IEEE International Symposium on Biomedical Imaging*, 2012.
- [33] S. Morales, V. Naranjo, J. Angulo, and M. Alcañiz, "Automatic detection of optic disc based on pca and mathematical morphology," *Medical Imaging, IEEE Transactions on*, vol. 32, no. 4, pp. 786–796, April 2013.
- [34] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *Artificial Intelligence Research, Journal of*, vol. 16, pp. 321–357, 2002.
- [35] T. Scheffer, "Error estimation and model selection," Ph.D. dissertation, Technischen Universitt Berlin, School of Computer Science, 1999.
- [36] S. Dudoit and M. J. van der Laan, "Asymptotics of cross-validated risk estimation in estimator selection and performance assessment," *Statistical Methodology*, vol. 2, no. 2, pp. 131 – 154, 2005.
- [37] R. Kohavi and G. H. John, "Wrappers for feature subset selection," *Artificial Intelligence*, vol. 97, no. 12, pp. 273 – 324, 1997.
- [38] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The weka data mining software: An update," *SIGKDD Explor. Newsl.*, vol. 11, no. 1, pp. 10–18, Nov. 2009.
- [39] L. S. Cessie and J. C. van Houwelingen, "Ridge Estimators in Logistic Regression," *Applied Statistics*, vol. 41(1), pp. 191–201, 1992.
- [40] R. Hecht-Nielsen, *Neurocomputing*. Addison-Wesley Longman Publishing Co., Inc., 1990.
- [41] C. C. Chang and C. J. Lin, "LIBSVM: A library for support vector machines," *Intelligent Systems and Technology, ACM Transactions on*, vol. 2, no. 3, pp. 27:1–27:27, 2011.
- [42] G. H. John and P. Langley, "Estimating continuous distributions in bayesian classifiers," in *Uncertainty in Artificial Intelligence, Proceedings of the Eleventh Conference on*, ser. UAI'95, 1995, pp. 338–345.
- [43] J. R. Quinlan, *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers Inc., 1993.
- [44] J. Rodriguez, L. Kuncheva, and C. Alonso, "Rotation forest: A new classifier ensemble method," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 10, pp. 1619–1630, 2006.
- [45] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [46] Y. Freund and R. E. Schapire, "Experiments with a new boosting algorithm," in *Machine Learning, Proceedings of the 13th International Conference on (ICML 1996)*, L. Saitta, Ed., 1996, pp. 148–156.
- [47] M. Paci, L. Nanni, A. Lahti, K. Aalto-Setälä, J. Hyttinen, and S. Severi, "Non-binary coding for texture descriptors in sub-cellular and stem cell image classification," *Current Bioinformatics*, vol. 8, no. 2, pp. 208–219, 2013.
- [48] L. Nanni, S. Brahmam, S. Ghidoni, E. Menegatti, and T. Barrier, "Different approaches for extracting information from the co-occurrence matrix," *PLoS ONE*, vol. 8, no. 12, p. e83554, 2013.