AUTOMATIC SYSTEM FOR DIABETIC RETINOPATHY SCREENING BASED ON AM-FM, PARTIAL LEAST SQUARES, AND SUPPORT VECTOR MACHINES

[1,2]E. Simon Barriga, [1,2]Victor Murray, [1,2]Carla Agurto, [2]Marios Pattichis, [3]Wendall Bauman, [1]Gilberto Zamora, and [1,4]Peter Soliz

[1]VisionQuest Biomedical, Albuquerque, NM, USA, [2]University of New Mexico, Electrical and Computer Engineering Department, NM, USA, [3]Retina Institute of South Texas, San Antonio, TX, USA, [4]University of Iowa Department of Ophthalmology and Visual Sciences, Iowa City, IA

ABSTRACT

Diabetic retinopathy (DR) is a disease that affects over 170 million people worldwide. In the United States, it is estimated that over 10 million diabetics do not receive the recommended annual eye examinations, significantly increasing their risk of vision loss. In this paper we present an automatic system to detect the presence of DR by analyzing a photograph of the central field of the retina. The system applies Amplitude Modulation-Frequency Modulation (AM-FM) for feature extraction, and partial least squares (PLS) and support vector machines (SVM) for classification. We tested the system on a total of 400 images, and obtained an area under the ROC curve (AUC) of 0.86, and corresponding sensitivity/specificity of 98%/68%. We also tested the accuracy of the system for patients needing immediate referral to a specialist, obtaining an AUC of 0.98.

Index Terms— Diabetic retinopathy, amplitude modulation-frequency modulation, AM-FM, partial least squares, support vector machines

1. INTRODUCTION

Most other approaches to computer-based DR screening try to replicate the manual process of detecting individual lesions, referred to as bottom up. These algorithms attempt to detect and count lesions such as microaneurysms, exudates, hemorrhages, neovascularization, etc. The focus on specific lesions that are associated with DR is at the expense of comprehensively evaluating the patient's retinal images for any abnormality, including drusen (one sign of age-related macular degeneration) and other retinal abnormalities. Ours is a top down approach that faithfully adheres to the goal of screening patients for any retinal disease and leaving the diagnosis to the expert. Our algorithm screens normals and deduces that the remaining images have some form of pathology. These images are then presented to a human reader or clinician to determine the specific pathology. This hybrid strategy has several other advantages: (1) is better accepted by eye care providers; (2) is more comprehensive in terms of screening for other eye disease; and (3) has significant economic implications.

Several investigators have studied the effectiveness of computer-based eye disease screening, including one of the first commercial attempts (Retinalyze by Larsen et al. [1]). Retinalyze, limited to DR screening, was reported to have a sensitivity of 88% and a specificity of 52%. Though having used substantially more cases to test his DR screening algorithm, Abràmoff et al. [2] give results that are consistent with the work of Fleming [3], Lee [4], Sanchez [5], and numerous others. In the U.S., Tobin, Chaum, and Karnowski [6] have reported on a system for DR screening that is based on the concept of content-based image retrieval (CBIR). Abràmoff et al. concluded in their study that "published algorithms cannot yet be recommended for clinical practice." The basis of this statement was that current algorithms fail to screen comprehensively for all important lesions, including neovascularization and macular edema.

In this paper we present a diabetic retinopathy screening system based on Amplitude modulation - frequency modulation (AM-FM) and two statistical classifiers: partial least squares (PLS) and support vector machines (SVM). This system is evaluated in a set of 400 images, where we demonstrate high accuracy in the detection of patients with different levels of DR.

2. DATASET

For this study we have used 400 high-quality retinal images from the Messidor dataset [7]. All images are color 8-bit per channel on TIFF uncompressed format. 400 images are taken with a mydriatic camera (with pupil dilation), centered on the macula, and have size 2240-by-1488 pixels.

3. MULTISCALE AM-FM OF RETINAL IMAGES

3.1. AM-FM decomposition

A digital image can be represented by its amplitude and frequency components as:

$$I(x,y) \approx \sum_{n=1}^{M} a_n(x,y)\cos\varphi_n(x,y) \qquad (1)$$

where M is the number of AM-FM components, $a_n(x,y)$ represent the instantaneous amplitude functions (IA), and $\varphi_n(x,y)$ represent the instantaneous phase functions [8].

For each AM-FM component, the instantaneous frequency (IF) is defined in terms of the gradient of the phase φ_n:

$$\nabla\varphi_n(x,y) = \left(\frac{\partial\varphi_n(x,y)}{\partial x}, \frac{\partial\varphi_n(x,y)}{\partial y} \right). \quad (2)$$

3.2. Frequency scales

AM-FM components are extracted from different frequency bands. We consider the use of 31 bandpass channel filters associated with six frequency scales (see Table 1). These six filters (or "scales") can be used in conjunction to create thirteen Combinations of Scales (CoS) which span the image frequency spectrum. We estimate a single AM-FM component over each combination of scales using Dominant Component Analysis [9].

Figure 1a shows an image from the Messidor database containing abnormal features typical of diabetic retinopathy, such as exudates, hemorrhages, and micro aneurysms. Figure 1b shows the instantaneous amplitude of the AM-FM representation for the very low (VL) frequencies filter, and Figure 1c shows the IF magnitude for the LPF filter. As seen in these images, most of the features of the image, such as vessels, optic disc, and abnormal features, are well represented by the AM-FM features.

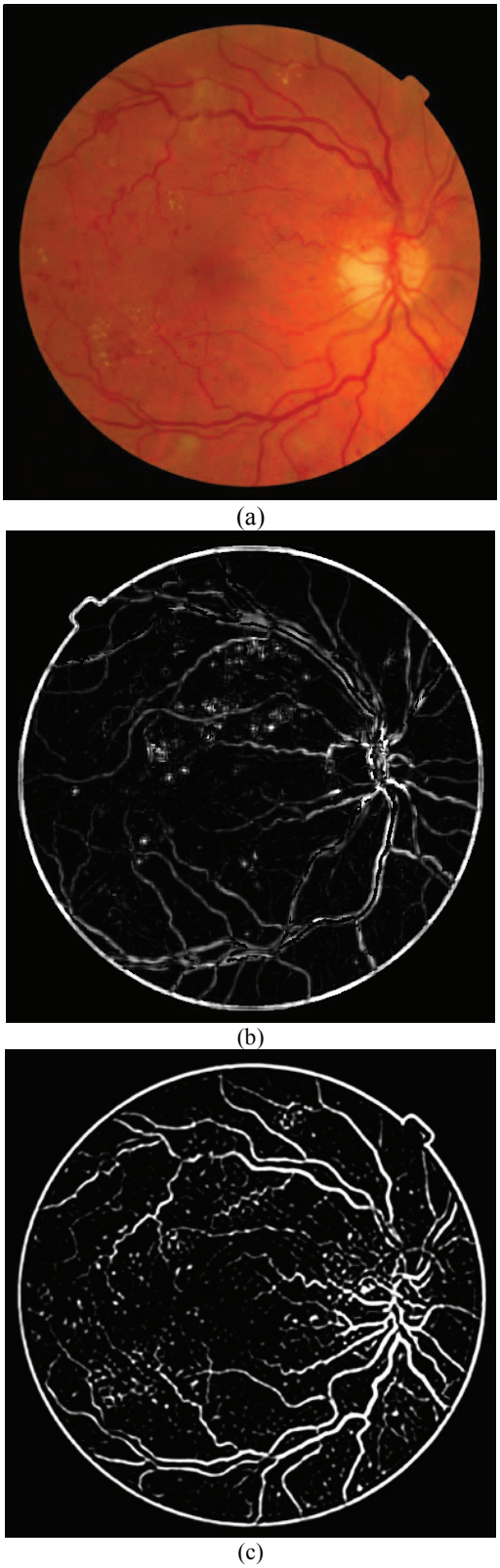
TABLE 1
BAND PASS FILTERS ASSOCIATED WITH MULTIPLE IMAGE SCALES.

Frequency Scale Band	Filters	Range in pixels
Low Pass Filter (LPF)	1	45.3 to ∞
Ultra Low Frequencies (UL)	26-31	22.6 to 64
Very Low Frequencies (VL)	20-25	11.3 to 32
Low Frequencies (L)	14-19	5.7 to 16
Medium Frequencies (M)	8-13	2.8 to 8
High Frequencies (H)	2-7	1.4 to 4

4. CLASSIFICATION SYSTEM

The objective of the DR screening system is to assign an image to one of two classes: "normal", if the patient does not present any signs of DR, and "abnormal" if the patient presents any characteristics of DR, regardless of the severity of the disease. The main objective of the screening system is not to diagnose the disease, but to refer the patient to a specialist only when a disease is present, thus improving the throughput of DR cases on the clinic.

Figure 2 shows a block diagram of the screening system. Once the AM-FM features are calculated, each image is separated in a number of regions of interest (ROI),



(a) (b) (c)

Figure 1. (a) Image from Messidor database. (b) Instantaneous Amplitude image for VL filter. (c) Instantaneous Frequency Magnitude for LPF filter.relevant

and histograms are calculated for each ROI. Those features are then clustered to reduce dimensionality. We used k-means clustering (an unsupervised classification method) to cluster the information in 30 groups so a feature vector for each image can be obtained. This vector represents the number of regions in each of the 30 clusters for each image. Finally we use a classifier to determine if the image belongs to the "normal" or "abnormal" category.

4.1. Classifier

We have used two different classifiers to determine which one produces the best results in determining if the images are normal or abnormal. Partial least squares (PLS) and support vector machines (SVM) are the two classifiers presented in this work. A brief explanation of both follows.

4.1.1. Partial least squares (PLS)

PLS is based on reducing the matrix of independent variables to a lower dimensional sub-space. To better explain this, the regression problem is given by

$$y = X\beta + \varepsilon \qquad (3)$$

where y is a $nx1$ vector of dependent variables, X is a nxp matrix of independent variables, β is a $px1$ vector of regression weights, and ε is a $nx1$ vector of residuals. PLS is based on a reduction of the independent variable X to a lower dimensional subspace. The first step on the method is to factor X as

$$X = T * L \qquad (4)$$

where T is an orthogonal nxp matrix of T-scores and L is a pxp matrix of factor loadings. Such a factorization is not unique and the definition of PLS is given by the criteria of the factorization.

PLS relies on a spectral decomposition which contains the correlation of the independent variables X with the dependent variables y. This way, PLS tries to explain the relationship of X with y as parsimoniously as possible [12].

4.1.2. Support Vector Machines (SVM)

SVM are a set of supervised learning methods for classification, based on the construction of hyperplanes for the separation of classes. Originally, hyperplanes could only be used for linear classification, but it is possible to solve non-linear classification problems using kernels [10].

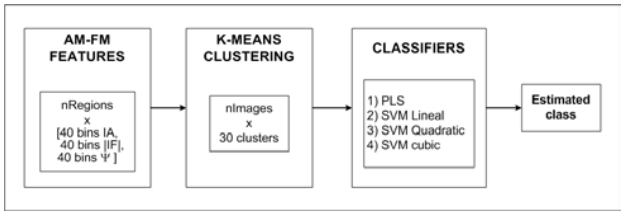


Figure 2. Block diagram of the DR screening system.

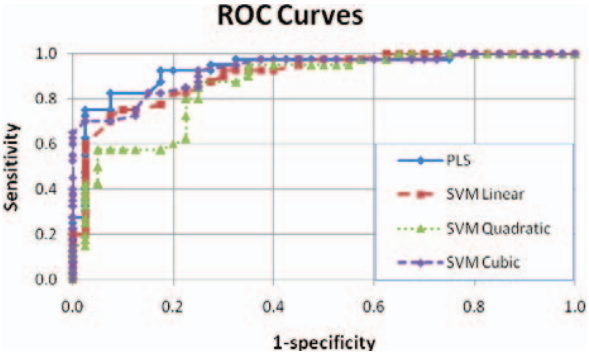


Figure 3. ROC curves for best results of the screening system. Classification using PLS achieves slightly better results than SVM methods.

For this paper we have used the linear, quadratic, and cubic polynomial kernels. All of those can be derived from the following equation:

$$k(x_i, x_j) = (x_i \bullet x_j)^n \qquad (5)$$

where k is the kernel function, x_i and x_j are data points, and n is the degree of the polynomial kernel function.

5. RESULTS AND DISCUSSION

We applied the screening algorithm on two sets of 400 images from the Messidor database. For each set, half were randomly selected for training and the rest were used for testing. We repeated this process 20 times to avoid any bias in the results.

Table 3: Messidor data classification results

	Average AUC	Max AUC	Min AUC	Best Sens/Spec
PLS	0.86	0.95	0.77	98%/67%
SVM Linear	0.82	0.93	0.73	95%/55%
SVM Quadratic	0.71	0.86	0.59	95%/65%
SVM Cubic	0.84	0.93	0.72	98%/62%

Table 3 presents the results of the classification for the first set of Messidor images. Here, we see that PLS achieves the best results, with an average AUC of 0.86, and maximum and minimum AUCs of 0.95 and 0.77, respectively. Its corresponding best sensitivity and specificity pairs are 98% and 67%, respectively. The best results using SVM were obtained with the cubic kernel function, which achieved an average AUC of 0.84, with maximum and minimum AUCs of 0.93 and 0.72, respectively. Best sensitivity and specificity of 98% and 62% were achieved in this case. Figure 3 shows the ROC curves for the best results from table 3.

Table 4 presents the results of classifying only the patients needing immediate referral to an eye clinic due to the suspected presence of immediate vision threatening conditions such as clinically-significant macular edema or

Table 4: Messidor data immediate referrals classificattion				
	Average AUC	Max AUC	Min AUC	Best Sens/Spec
PLS	0.98	1	0.83	100%/100%
SVM Linear	0.96	1	0.87	100%/100%
SVM Quadratic	0.82	0.90	0.68	100%/65%
SVM Cubic	0.95	0.99	0.85	100%/97%

proliferative diabetic retinopathy. In this case, only images from immediate referral patients and normal subjects were used to produce the classification results.

From table 4, we see that our system obtains high accuracy when determining the subjects with advanced stages of retinopathy. In the best cases, we have obtained a sensitivity of 100%. Comparing tables 3 and 4, we note that most of the cases where abnormal patients are missed by the system are consistent with images presenting a small number of lesions corresponding to patients who do not require a referral to the specialist and are not at risk for vision threatening progression. In this table we also see that PLS again produces the best results for classification, closely followed by SVM with linear and cubic kernels.

6. CONCLUSIONS

We have presented a system for diabetic retinopathy screening that achieves high sensitivity and specificity when classifying patients with abnormal retinal features. The average result over 20 runs of the algorithm on a randomized sample of our database has yielded an AUC = 0.86, which is comparable to the results of the state of the art system presented by Niemeijer, et al. [11]. Neimeijer's system uses a "bottom up" approach, which is based on marking every abnormal lesion present of the retina, as opposed to our "top down" approach which does not need marking and detection of every individual lesion.

The advantage of our approach is the ability to apply this methodology to any arbitrary set of data with different characteristics, such as pixel density, resolution, compression, etc. without the laborious task of creating a reference database of images which have had each lesion annotated by a trained analysis or ophthalmologist. Today, there are countless numbers of retinal imagers collecting data under various conditions which may require individualized or specifically trained algorithms. A methodology is needed that easily transports to each of these databases with the need only for a top-level grading of an image as normal or pathological. This task is commonly performed by trained technicians in a matter of seconds per image. Our approach has the potential for easily being transported to any imaging system.

Future work will be oriented on selecting the most relevant AM-FM features for the system, as some of them can be extracting redundant information that only adds to the complexity of the classification system. Also, we will explore different kernels for SVM that could potentially improve the results.

ACKNOWLEDGEMENTS

This project was supported in by the National Eye Institute under Grants EY018280 and EY020015.

We would like to thank Méthodes d'Evaluation de Systèmes de Segmentation et d'Indexation Dédiées à l'Ophtalmologie Rétinienne (MESSIDOR) for allowing us to use their database in this study.

7. REFERENCES

[1] Larsen N, Godt J, Grunkin M, Lund-Andersen H, Larsen M. "Automated detection of diabetic retinopathy in a fundus photographic screening population." Invest Ophthalmol Vis Sci. 2003;44(2):767-71.

[2] Abràmoff, M.D., et al., "Evaluation of a system for automatic detection of diabetic retinopathy from color fundus photographs in a large population of patients with diabetes." Diabetes Care, 2008. 31(2): p. 193-8.

[3] Fleming, A.D., et al., "Automatic detection of retinal anatomy to assist diabetic retinopathy screening." Phys Med Biol, 2007. 52(2): p. 331-45.

[4] Lee, SC, Lee ET, Wang Y, Klein R, Kingsley RM, Warn A, "Computer classification of non-proliferative diabetic retinopathy," Archives of Ophthalmology, June 123(5), 2005.

[5] Sánchez CI, Hornero R, López MI, Poza J, "Retinal image analysis to detect and quantify lesions associated with diabetic retinopathy," Conf Proc IEEE Eng Med Biol Soc. 2004;3:1624-7.

[6] Tobin KW, Abramoff MD, Chaum E, Giancardo L, Govindasamy V, Karnowski TP, Tennant MT, and Swainson S, "Using a patient image archive to diagnose retinopathy," Conference proceedings : . Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference 1:5441-4 2008.

[7] TECHNO-VISION Project, "MESSIDOR: methods to evaluate segmentation and indexing techniques in the field of retinal ophthalmology." Available: http://messidor.crihan.fr/.

[8] M. S. Pattichis and A. C. Bovik, "Analyzing image structure by multidimensional frecuency modulation", IEEE Trans. Pattern Anal. Mach. Intell., 2007; no. 5, pp. 753-766.

[9] V. Murray, P. Rodriguez and M.S. Pattichis, "Multi-scale AM-FM Demodulation and Reconstruction Methods with Improved Accuracy," accepted for publication, IEEE Transactions on Image Processing, 2009.

[10] Cristianini, N. and Shawe-Taylor, J. (2000). "An Introduction to Support Vector Machines and Other Kernel-based Learning Methods," First Edition, Cambridge: Cambridge University Press.

[11] M. Niemeijer, M.D. Abramoff, B. van Ginneken, "Information Fusion for Diabetic Retinopathy CAD in Digital Color Fundus Photographs", IEEE Transactions on Medical Imaging, 2009, vol. 28, pp. 775-785.

[12] Wold, H., "Personal memories of the early PLS development," Chemometrics and Intelligent Laboratory Systems, 58:83–84, 2001.