35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Comparison of Logistic Regression and Neural Network Classifiers in the Detection of Hard Exudates in Retinal Images

María García, *Member, IEEE*, Carmen Valverde, María I. López, Jesús Poza, *Member, IEEE* and Roberto Hornero, *Senior Member, IEEE*

Abstract— **Diabetic Retinopathy (DR) is a common cause of visual impairment in industrialized countries. Automatic recognition of DR lesions in retinal images can contribute to the diagnosis and screening of this disease. The aim of this study is to automatically detect one of these lesions: hard exudates (EXs). Based on their properties, we extracted a set of features from image regions and selected the subset that best discriminated between EXs and the retinal background using logistic regression (LR). The LR model obtained, a multilayer perceptron (MLP) classifier and a radial basis function (RBF) classifier were subsequently used to obtain the final segmentation of EXs. Our database contained 130 images with variable color, brightness, and quality. Fifty of them were used to obtain the training examples. The remaining 80 images were used to test the performance of the method. The highest statistics were achieved for MLP or RBF. Using a lesion based criterion, our results reached a mean sensitivity of 95.9% (MLP) and a mean positive predictive value of 85.7% (RBF). With an image-based criterion, we achieved a 100% mean sensitivity, 87.5% mean specificity and 93.8% mean accuracy (MLP and RBF).**

I. Introduction

Diabetic retinopathy (DR) is an important cause of visual impairment among people of working age in industrialized countries [1]. Treatment with laser photocoagulation can reduce the risk of blindness and moderate vision loss by more than 90%, but the treatment does not restore lost vision [2]. Therefore, early DR detection is essential to prevent serious sight damage. However, this is not an easy task because the patient perceives no symptoms of the disease until the advanced stages. To ensure that treatment is received on time, diabetic patients should undergo periodic eye examinations [2]. The growing incidence of diabetes, the high cost of examinations and the lack of specialists prevent many patients from receiving effective treatment. Computer aided diagnosis of DR could help ophthalmologists in the diagnosis and follow-up of the disease, with the subsequent cost and time savings. Efforts have been made to automatically detect early clinical signs of DR, such as hard exudates (EXs) [3]. EXs are lipid and lipoprotein deposits, white, yellowish or waxy, which appear as compact patches with well-defined boundaries in retinal images. Many different methods to detect these lesions can be found in the literature. Edge detectors were applied to extract EXs borders [4], [5]. Other authors [6] distinguished among bright lesions using *k*-NN and linear discriminant analysis classifiers. Neural networks (NNs) have been also applied for EXs segmentation [7-10].

The present research introduces a comparison among three different classifiers for EXs detection. These are a logistic regression (LR) model and two neural network classifiers: multilayer perceptron (MLP) and radial basis function (RBF). Hence, the aims of the present study were: (i) to develop an automatic method for the detection of EXs in retinal images, (ii) to compare the performance of three classifiers in this context, (iii) to assess the diagnostic potential of the proposed method in DR detection.

II. Material and Methods

A. Image database

In this study, a total of 130 images were used. They were provided by the "Instituto de Oftalmobiología Aplicada" of the University of Valladolid, Spain. All of them belonged to healthy retinas or to patients with mild to moderate DR, according to a senior ophthalmologist. Thirty of these images were captured with a TopCon TRC-NW6S non-mydriatic retinal camera at a field-of-view (FOV) of 45º. The remaining 100 images were captured with a TopCon TRC-50IX mydriatic retinal camera at 50º FOV. No bad quality images were discarded form the study. Image resolution was 576x768 pixels in 24 bit JPEG format. A senior ophthalmologist manually marked the EXs in the images. The results of our algorithm were compared with these hand labeled images. The images were randomly divided into a training set and a test set. The training set contained EX and non-EX examples from 50 images in our database. The remaining 80 images formed the test set.

B. Luminosity and contrast normalization

The physical features of a patient and the image acquisition process produce great variability within and between retinal images. Thus, it is harder to distinguish retinal features and lesions in some areas. Preprocessing is necessary to normalize the images and increase the contrast between EXs and the background. We followed the method

* This research was partially supported by: the *Ministerio de Ciencia e Innovación* and *Instituto de Salud Carlos III* under project PI10/02664, the *Ministerio de Economía y Competitividad* and FEDER under project TEC2011-22987; the 'Proyecto Cero 2011 on Ageing' from *Fundación General CSIC*, *Obra Social La Caixa* and CSIC; and project VA111A11-2 from *Consejería de Educación* (*Junta de Castilla y León*)..

M. García, J. Poza, and R. Hornero are with the Biomedical Engineering Group, University of Valladolid, Paseo de Belén 15, 47011 Valladolid, Spain, (e-mail: margar@tel.uva.es, jespoz@tel.uva.es, robhor@tel.uva.es).

C. Valverde is with the Hospital of Medina del Campo, Valladolid, Spain (e-mail: carmenvalverde12@hotmail.com)

M. I. López is with the Instituto de Oftalmobiología Aplicada, University of Valladolid, and Hospital Clínico Universitario, Valladolid, Spain (e-mail: maribel@ioba.med.uva.es).

proposed by Foracchia et al. [11], which computes the original undistorted image from estimates of the luminosity and contrast drifts of the observed image. Normalization yielded a color image, with color components $\hat{R}^o\hat{G}^o\hat{B}^o$. The effect of normalization is shown in Fig. 1(a-b).

C. Segmentation

The aim of the segmentation stage was to separate all the possible EX regions in the image. Segmentation was accomplished using the properties of the global and local histograms in $\hat{G}^o$. These histograms were usually bell-shaped. The maximum corresponded to the background and the right tail corresponded to the bright regions we wanted to segment. We set a threshold at the gray level of the right tail for which the histograms (global and local) decreased to the 10% of the maximum. To calculate local histograms, the image was partitioned into square blocks of side 200 pixels (zero-padding when necessary). The resulting images were combined using the binary AND operation to obtain the candidate EX regions [9], [10]. The papillary region was also masked to reduce the computational cost of the classifiers. We used a method that combines mathematical morphology and Hough transform [12]. Fig. 1(c) is the result of segmentation.

D. Feature extraction

In order to classify the candidate regions into EX or non-EX classes, a set of significant features was extracted from each region. We focused on those characteristics that help ophthalmologists to visually distinguish EXs. We selected 24 features previously used in this context [6-10], [13]:

- Mean $\hat{R}^o\hat{G}^o\hat{B}^o$ values inside the region (1-3).
- Standard deviation of the $\hat{R}^o\hat{G}^o\hat{B}^o$ values inside the region (4-6).
- Mean $\hat{R}^o\hat{G}^o\hat{B}^o$ value around the region (7-9).
- Standard deviation of the $\hat{R}^o\hat{G}^o\hat{B}^o$ value around the region (10-12).
- $\hat{R}^o\hat{G}^o\hat{B}^o$ values of the region centroid (13-15).
- Region size (16).
- Region compactness (17).
- Region edge strength (18).
- Homogeneity of the region, measured as the Shannon's entropy of the $\hat{R}^o\hat{G}^o\hat{B}^o$ values inside the region (19-21).
- Color difference of the $\hat{R}^o\hat{G}^o\hat{B}^o$ values (22-24).

E. Feature selection

Feature selection is aimed at choosing the most relevant subset of the extracted features for a specific problem. This is a useful step because misclassification probability tends to increase with the dimensionality of the input space and the structure of the classifier is more difficult to interpret [14]. LR is a statistical method commonly used for this task and is adequate for data sets that are not multivariate normal and homocedastic [15]. It analyzes the relationship between a dichotomous dependent variable (the class of an object) and several independent variables (the extracted features) [16]. If the possible values of the dependent variable are 0 and 1, and if $\mathbf{X}$ is the vector of independent variables, the *logit* transformation can be modeled under the LR model as [16]:

$$\text{logit}(Y=1) = \ln\left[\frac{P(Y=1)}{P(Y=0)}\right] = \alpha + \boldsymbol{\beta}\mathbf{X}, \quad (1)$$

where Y is the dependent variable, 1 is the desired outcome, $\mathbf{X}$ is the independent variable vector and α and $\boldsymbol{\beta}$ are the parameters of the model to be identified by the maximum likelihood method [16]. For model selection, we used a stepwise forward selection method, where the inclusion of variables was based on the score test and the elimination of variables was based on the likelihood ratio test [16].

F. Classification

LR can be also applied to separate patterns in two classes. First, α and $\boldsymbol{\beta}$ are set using the training data as described before. Then, the probability $p(Y=1)$ for a test pattern can be obtained as [16]:

$$\text{p}(Y=1) = \frac{1}{1+e^{-(\alpha+\boldsymbol{\beta}\mathbf{X})}} \quad (2)$$

The class of a test example can be determined by comparing $p(Y=1)$ with a threshold. In this study, if $p(Y=1) > 0.5$, the object was labeled in class 1 (class EX); otherwise, it was labeled as class 0 (class non-EX). The selection of this threshold was motivated by the interpretation of the outputs as posterior probabilities.

NNs have been also successfully used in previous studies after a feature selection stage based on LR [10]. Thus, we included also two additional classifiers based on the most representative paradigms of NNs. MLP is an important class of NN classifiers [17], [18]. A single-layered MLP can approximate any function provided it contains enough hidden units and their activation function satisfies certain constraints [19]. In the hidden layer, we used a hyperbolic tangent sigmoid activation function, which improves the learning speed of MLP [17]. In the output layer, we used the logistic sigmoid (*log-sigmoid*) activation function. The outputs of log-sigmoid lie in the range (0, 1), and can be interpreted as posterior probabilities [18]. MLP training can be viewed as the minimization of an error function. We chose a cross-entropy error function because it simplifies the optimization process when *log-sigmoid* is used in the output layer. As minimization algorithm, we selected the scaled conjugate gradients method because of its fast convergence. To avoid overfitting, we used a weight decay regularizer [18]. The regularization parameter, υ, and the number of hidden neurons were experimentally determined.

RBF NNs are also universal approximators and contain three layers of neurons [17], [18]. The input layer receives the feature vectors. The hidden layer in RBF performs a nonlinear transform of the input space into a high

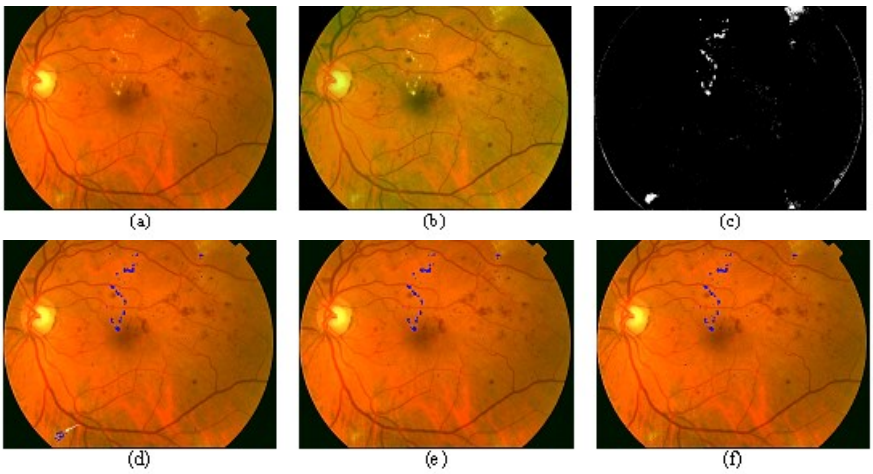

Figure 1. (a) Original retinal image. (b) Normalized image. (c) Candidate EX regions obtained after segmentation, in white. (d) Final result obtained with LR, in blue, superimposed over the original image. A cotton wool spot, indicated with a white arrow, has been incorrectly detected as EX. (e) Final result obtained with MLP, in blue, superimposed over the original image. (f) Final result obtained with RBF, in blue, superimposed over the original image.

dimensional space, where the patterns are more likely to be linearly separable. The output layer is linear and provides the NN response [17], [18]. As hidden layer activation functions we chose multidimensional Gaussian functions, since they are translationally and rotationally invariant [17]. They are characterized by their centers and *spread*, which represents their width. For RBF training, we used the orthogonal least squares algorithm [20]. The number of radial basis functions and their *spread* were experimentally set. Classification examples are shown in Fig. 1(d-f).

III. Results

A. Selected features

We computed the 24 features described in Section II-C for 940 EX regions and 940 non-EX regions extracted after the segmentation of our training set. We used the statistical package SPSS (version 20; IBM Corp., Armonk, NY, USA) to perform LR on these data. The method selected 13 features: (1) mean of the $\hat{R}^o$ value inside the region, (2) standard deviation of the $\hat{R}^o$ value inside the region, (3-5) mean values of the three color channels around the region, (6-8) standard deviation of the three color channels around the region, (9) $\hat{G}^o$ value of the region centroid, (10) size, (11) edge strength, (12) homogeneity of the region in $\hat{G}^o$ and (13) color difference of the region in $\hat{R}^o$.

B. Optimization of the parameters of MLP and RBF

Different configurations of the NN classifiers were tested to find the architecture that led to the minimum prediction risk [19]. A 13-dimensional feature vector, which represented the selected features, was computed for each of the aforementioned 940 EX regions and 940 non-EX regions. To improve the behavior of NNs, the inputs were normalized (mean=0, standard deviation=1) and the data were randomly presented to the input of the network [17]. We used 10-fold cross-validation to asses the generalization ability of the NNs, as it is an estimator of the prediction risk [21]. For the different architectures, we measured the mean sensitivity (*SE*), specificity (*SP*) and accuracy (*AC*) obtained for the validation set.

The optimum operating point for MLP was chosen with 35 hidden neurons and $\upsilon = 2$ (*SE*=93.0%, *SP*=92.8%, *AC*=93.0%). Regarding RBF, the optimum operating point was obtained with 90 hidden neurons and *spread* =3.5 (*SE* =92.7%, *SP* =91.8%, *AC* =92.2%).

B. Performance of the algorithm

The performance of the complete algorithm was tested on 80 unseen images (test set). Forty of them contained EXs and in the remaining 40 images the ophthalmologists did not mark any EXs. Our results were obtained in terms of a lesion based criterion (pixel resolution) and an image-based criterion [8-10]. With the lesion based criterion, we measured the mean sensitivity (SE_l) and positive predictive value (PPV_l). With an image-based criterion, we measured the mean sensitivity (SE_i), specificity (SP_i) and accuracy (AC_i). In order to improve the performance of the system, we identified those images where less than 30 pixels (0.0068% of the total number of pixels) had been detected as EXs and considered them as belonging to healthy retinas [8]. Our results are summarized in Table I. They were obtained averaging the results for each image in the test set and considering only EXs as a marker for DR.

IV. Discussion and Conclusion

In this study, we propose an automatic method to extract EXs in fundus images. The algorithm was tested on 40 images with EXs and 40 images without EXs. A set of 24 features was statistically analyzed using LR to obtain a subset of 13 features with the maximum discriminatory power. These features were used to configure and train three classifiers: LR, MLP and RBF. To the best of our knowledge, the comparison of these classifiers has not been previously investigated in this context.

The statistics in Table I show that we obtained better results using MLP or RBF than using LR with both criteria, even if feature selection relied on the results for the LR

TABLE I. PERFORMANCE OF LR AND MLP FOR EX DETECTION

Neural Network	Lesion-based criterion		Image-based criterion		
	SE_l (%)	PPV_l (%)	SE_i (%)	SP_i (%)	AC_i (%)
LR	95.1	84.0	100	77.5	88.8
MLP	95.9	84.5	100	87.5	93.8
RBF	86.9	85.7	100	87.5	93.8

SE_l = lesion-based sensitivity, PPV_l = lesion-based positive predictive value, SE_i = image-based sensitivity, SP_i = image-based specificity, AC_i = image-based accuracy.

classifier. This fact highlights the utility of NNs in retinal image analysis. The disadvantage of MLP and RBF over LR is that they require longer training times, especially when cross-validation is applied. Besides, the values of SE_i in Table I show that we detected all images with EXs using all classifiers. These results can be considered satisfactory according to [22], where the authors stated that SE_i equal or greater than 60% maximizes cost-effectiveness in screening for DR. We also detected some false positives, as SP_i values in Table I indicate. However, for computer-aided DR screening, it is more important to correctly classify all patients with sight-threatening DR (high sensitivity), even if it involves misclassifying some healthy subjects.

The detection of EXs in retinal images has also been analyzed in previous studies. Walter et al. [4] report lesion-based and image-based statistics. They obtained SE_l=92.8% and PPV_l=92.4%, SE_i=100% and SP_i=86.7%. However, the distinction between EXs and other bright lesions was not addressed. Li and Chutatape [5] obtained SE_i=100% and AC_i=74% using edge detectors. The differentiation among different bright lesions is addressed in [6]. They obtained SE_i=95.0% and SP_i=86.0%. A set of 86 features was studied in the final classification stage. Other authors [7] used a Support Vector Machine (SVM) classifier. They obtained SE_l=88.0% and PPV_l=84.0%. The results given for the work by Osareh [8] reach SE_l =90.0%, PPV_l =89.3%, SE_i=95.0% and SP_i=88.9%. They studied MLP and SVM classifiers. The results in some of these works are above the statistics in Table I. However, it is impossible to objectively contrast our results with those reported in literature because the databases and performance measures vary among studies.

The proposed method presents some limitations that merit consideration. Despite the normalization step, the color and size of EXs can differ even within the same eye. This makes it difficult to detect all of them, specially the most subtle lesions. Besides, it would be desirable to test whether the proposed method is appropriate to analyze images from different cameras and showing a wide variety of lesions. In this sense, future works would be aimed at testing the proposed methodology on publicly available databases like DIARETDB0 and DIARETDB1. We would also try to study additional features and detect other types of lesions to grade the evolution of DR.

In conclusion, the use of automatic methods based on NNs can be adequate for lesion detection in retinal images and could serve as a diagnostic aid for ophthalmologists in the screening for DR.

REFERENCES

[1] G. L. Ong, L. G. Ripley, R. S. Newsom, M. Cooper, and A. G. Casswell, "Screening for sight-threatening diabetic retinopathy: comparison of fundus photography with automated color contrast threshold test," *Am. J. Ophthalmol.*, vol. 137, pp. 445–452, 2004.

[2] D. S. Fong, L. Aiello, T. W. Gardner, G. L. King, G. Blankenship, J. D. Cavallerano, F. L. Ferris, and R. Klein, "Diabetic retinopathy," *Diabetes Care*, vol.26, pp.226–229, 2003.

[3] R. Klein, B. E. Klein, S. E. Moss, M. D. Davis, and D. L. DeMets, "The Wisconsin epidemiologic study of diabetic retinopathy VII. Diabetic nonproliferative retinal lesions," *Ophthalmol*, vol. 94, pp. 1389–1400, 1987.

[4] T. Walter, J. C. Klein, P. Massin, and A. Erginay, "A contribution of image processing to the diagnosis of diabetic retinopathy – Detection of exudates in color fundus images of the human retina," *IEEE Trans. Med. Imaging*. vol. 21, pp. 1236–1243, 2002.

[5] H. Li, and O. Chutatape, "Automated feature extraction in color retinal images by a model based approach," *IEEE Trans. Biomed. Eng*. vol.51, pp.246–254, 2004.

[6] M. Niemeijer, B. van Ginneken, S. R. Russell, M. S. A. Suttorp-Schulten, and M. Abràmoff, "Automated detection and differentiation of drusen, exudates, and cotton-wool spots in digital color fundus photographs for diabetic retinopathy diagnosis," *Invest. Ophthalmol. Vis. Sci*., vol.48, pp. 2260–2267, 2007.

[7] X. Zhang and O. Chutatape, "Top-down and bottom-up strategies in lesion detection of background diabetic retinopathy," in *Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition*, 2005, vol. 2, pp. 422–428.

[8] A. Osareh, "Automated identification of diabetic retinal exudates and the optic disc," Ph.D. dissertation, Dept. Comput. Scienc., University of Bristol, Bristol, U.K., 2004

[9] M. García, C. I. Sánchez, M. I López, D. Abásolo, and R. Hornero, "Neural network based detection of hard exudates in retinal images," *Comput. Methods Programs Biomed.* vol. 93, no. 1, pp.9–19, 2009.

[10] M. García, C. I. Sánchez, J. Poza, M. I López, and R. Hornero, "Detection of hard exudates in retinal images using a radial basis function classifier," *Ann. Biomed. Eng.*, vol. 37, pp.1448–1463, 2009.

[11] M. Foracchia, E. Grisan, and A. Ruggeri, "Luminosity and contrast normalization in retinal images," *Med. Imag. Anal.*, vol. 9, pp. 179–190, 2005.

[12] C. I. Sánchez, R. Hornero, M. I. López, M. Aboy, J. Poza, and D. Abásolo, "A novel automatic image processing algorithm for detection of hard exudates based on retinal image analysis," *Med. Eng. Phys.*, vol. 30, pp. 350–357, 2007.

[13] J. W. Sleigh, D. A. Steyn-Ross., M. L. Steyn-Ross., C. Grant, and G. Ludbrook, "Cortical entropy changes with general anaesthesia: theory and experiment," *Physiol. Meas.*, vol. 25, pp. 921–934, 2004.

[14] M. Nadler and E. Smith, Pattern Recognition Engineering, New York: John Wiley, 1993.

[15] J. D. Jobson, *Applied multivariate data analysis. Volume I: regression and experimental design*. New York: Springer, 1991.

[16] D. W. Hosmer and S. Lemeshow, *Applied logistic regression*. New York, NY: John Wiley & Sons, 1989.

[17] S. Haykin, *Neural Networks: A comprehensive foundation*. Upper Saddle River, NJ: Prentice-Hall International, 1999.

[18] C. M. Bishop, *Neural networks for pattern recognition*. Oxford: Oxford University Press, 1995.

[19] G.-B. Huang, Y.-Q. Chen, and H. A. Babri, "Classification ability of single hidden layer feedforward neural networks," *IEEE Trans. Neural Netw*., vol. 11, pp. 799–801, 2000.

[20] S. Chen, C. F. N. Cowan, P. M. Grant, "Orthogonal least squares learning algorithm for radial basis function networks," *IEEE Trans. Neural Netw.*, vol. 2, pp. 302–309, 1991.

[21] J. Moody, "Prediction risk and architecture selection for neural networks," in: *From Statistics to Neural Networks: Theory and Pattern Recognition Application*, V. Cherkassky, J. H. Friedman and H. Wechsler Eds. Berlin: Springer-Verlag, 1994.

[22] J. C. Javitt, J. K. Canner, R. G. Frank, D. M. Steinwachs, and A. Sommer, "Detecting and treating retinopathy in patients with type I diabetes mellitus. A health policy model," *Ophthalmol.*, vol. 97, pp. 483–494, 1990.