LBP and Machine Learning for Diabetic Retinopathy Detection

Jorge de la Calleja, Lourdes Tecuapetla, Ma. Auxilio Medina, Everardo Bárcenas, and Argelia B. Urbina Nájera

Universidad Politécnica de Puebla, Puebla, México, C.P. 72640
{jorge.delacalleja,lourdes.tecuapetla,
maria.medina,ismael.barcenas,argelia.urbina}@uppuebla.edu.mx

Abstract. Diabetic retinopathy is a chronic progressive eye disease associated to a group of eye problems as a complication of diabetes. This disease may cause severe vision loss or even blindness. Specialists analyze fundus images in order to diagnostic it and to give specific treatments. Fundus images are photographs taken of the retina using a retinal camera, this is a noninvasive medical procedure that provides a way to analyze the retina in patients with diabetes. The correct classification of these images depends on the ability and experience of specialists, and also the quality of the images. In this paper we present a method for diabetic retinopathy detection. This method is divided into two stages: in the first one, we have used local binary patterns (LBP) to extract local features, while in the second stage, we have applied artificial neural networks, random forest and support vector machines for the detection task. Preliminary results show that random forest was the best classifier with 97.46% of accuracy, using a data set of 71 images.

Keywords: machine learning, local binary patterns, medical image analysis.

1 Introduction

Currently there are approximately 382 million people worldwide with diabetes mellitus or simply diabetes. It is expected within 11 to 25 years, that this disease will increase to 592 million people according to the World Health Organization (WHO) and the International Diabetes Federation (IDF). Also, diabetes is the leading cause of blindness in people of working age between 40 and 60 years old, generating a health care spending of about $548 billion [7], [15].

The main complications of a person diagnosed with diabetes is that glucose levels are high, affecting several organs including kidney, nerves, brain and eyes. Particularly, diabetes can cause damage to the retina, known as diabetic retinopathy. At least the 80% of diabetic persons with 10-20 years of being diagnosed present any symptoms related to diabetic retinopathy [15].

Several works using machine learning and image processing methods have been proposed in order to classify some types of eye diseases, including diabetic

E. Corchado et al. (Eds.): IDEAL 2014, LNCS 8669, pp. 110–117, 2014.

retinopathy. In 2000 Ege B.M. et al [4] introduced a tool to provide automatic analysis of digital images of diabetic retinopathy. They tested several statistical classifiers, such as Bayesian, Mahalanobis and k-nn. Their best results were obtained with the Mahalanobis classifier. Osareh A. et al [13] proposed an automatic method to detect exudate regions. They introduced comprising image color normalization, enhancing the contrast between the objects and background, segmenting the color retinal images into homogenous regions using fuzzy c-means clustering, and classifying the regions into exudates and non exudates patches using artificial neural networks. They reported 92% of sensitivity and 82% of specificity. Niemeijer M. et al [10] proposed an automated system able to detect exudates and cotton-wool spots in color images. In their work they used k-nearest neighbors, reporting a operating characteristic curve (ROC) of 0.95 and sensitivity/specificity pairs of 0.95/0.88 for the detection of bright lesions of any type; 0.95/0.86, 0.70/0.93 and 0.77/0.88 for the detection of exudates, cotton-wool spots and drusen, respectively. In 2010, Silberman N. et al [14] proposed an automated system to detect diabetic retinopathy from retinal images. This approach used support vector machines to recognize exudates obtaining 98.4% of accuracy. In 2011, Gowda A. et al [5] attempted to detect exudates using backpropagation neural networks. The significant features were identified from preprocessed images by using two methods: decision trees and genetic algorithms. Their approach showed a classification accuracy of 98.45%. Kavitha S. and Duraiswamy K. [8] focused on automatic detection of diabetic retinopathy exudates in color fundus retinal images. Experiments on classification of hard and soft exudates were performed using image processing techniques. Exudates were detected with the aid of thresholding color histogram. The overall sensitivity, specificity and accuracy were 89.78%, 99.12% and 99.07%, respectively.

In contrast to previous approaches, we present a method to detect only diabetic retinopathy based on fundus images. Also, with the purpose of knowing how well local binary patterns performs on this domain, we have applied it to obtain local features from the images. The well-known machine learning methods of artificial neural networks, random forest and support vector machines were applied for the detection task. The paper is organized as follows: Section 2 describes a brief introduction of diabetic retinopathy. In Section 3 we describe the methods. In Section 4 we show preliminary experimental results and finally in Section 5 conclusions and future work are presented.

2 Diabetic Retinopathy

Diabetes mellitus is a metabolic disorder which results from a defect in insulin synthesis and secretion or from a resistance of the receptors on target tissues for this hormone. The most significant complication of diabetes mellitus involving the eye and which develops in 85% of all diabetics eventually is the retinopathy [6].

Diabetic retinopathy is an abnormality involving the small blood vessels that targets the central region, for example the macula. In fact, the diabetic retinopathy is a progressive disease and this is the main factor that causes blindness.

Classification of diabetic retinopathy according to the Early Treatment Diabetic Retinopathy Study (ETDRS) [1] is divided as: Diabetic Retinopathy Non Proliferative (DRNP), which is subdivided into mild, moderate, severe and very severe; and Diabetic Retinopathy Proliferative (DRP), which is subdivided into early, high-risk and advanced.

3 The Method

In order to classify fundus images as healthy or sick, we propose an automated method that is divided into two stages: 1) Extraction of features and 2) Classification of the images (Figure 1). In the first stage the characterization of the images is performed using local binary patterns considering uniform patterns, rotation invariant and rotation invariant uniform patterns. Then, a set of features is provided to machine learning algorithms to perform the classification task. In next subsections we explain the algorithms.

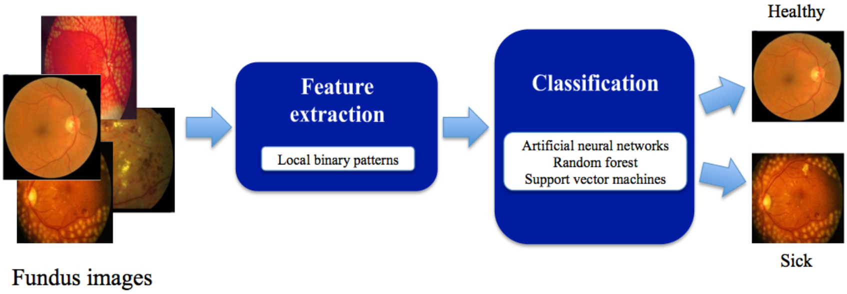

Fig. 1. Stages of the proposed method

3.1 Feature Extraction

The general idea for image classification is to extract relevant information (features) in a image, encode it as efficiently as possible, and compare one image encoding with a data base of similarly encoded images. Traditional approaches such as principal component analysis or linear discriminant analysis, look at the whole image as a high-dimensional vector in order to extract features. High dimensionality is not good, thus a lower-dimensional subspace is identified where useful information is preserved. Other alternative is extracting local features from images, i.e. instead of process the whole image as a high-dimensional vector, the idea is to describe only local features of an object. A method for performing this local feature extraction is local binary patterns explained below.

Local Binary Patterns. LBP was introduced as a texture description operator by Ojala et al. in 1996 [11]. This technique labels the pixels of an image by thresholding a 3×3 neighborhood of each pixel with the center value to yield a binary number (Figure 2). These binary numbers, called local binary patters

(LBP), codify local primitives including several types of curved edges and flat areas. The basic LBP operator only consider a small neighborhood, therefore it can not capture dominant features with large scale structures. Hence this operator was extended by using circular neighborhoods and interpolating the pixel values in order to allow any radius and number of pixels in the neighborhood [12].

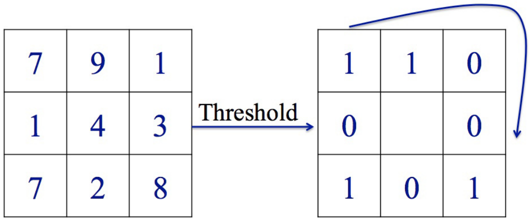

Fig. 2. The basic LBP Operator. The binary pattern 11001010 is equivalent to decimal 202.

The LBP operator produces 2^P different binary patterns for the P neighbors pixels, some of them contain more information than others, thus it is possible to use only a subset of fundamental patterns called uniform patterns to describe the texture of images. A LBP is uniform if it contains at most two bitwise transitions from 0 to 1 or vice versa when the binary string is considered circular. For example, 00000000, 01111110 and 11100111 are uniform patterns [12]. Accumulating the patterns which have more than 2 transitions into a single bin yields an LBP operator, with less than 2^P bins. For example, the number of labels for a neighborhood of 8 pixels is 256 for the standard LBP. After labeling a image with the LBP operator, a histogram of the labeled image $f_l(x,y)$ can be defined as

$$H_i = \sum_{x,y} I(f_l(x,y) = i), i = 0, ..., n-1. \tag{1}$$

where n is the number of different labels produced by the LBP operator and

$$I(F) = \begin{cases} 1 & F \text{ is true} \\ 0 & F \text{ is false} \end{cases} \tag{2}$$

This LBP histogram contains information about the distribution of the local micro-patterns, thus they can be used to describe image characteristics.

Fundus images can be seen as a composition of micro-patterns which can be described by the LBP histograms. Therefore, we have used LBP features to represent fundus images, which were equally divided into small regions $R_0, R_1, ...R_m$ to extract LBP histograms. The LBP features extracted from each sub-region are concatenated into a single histogram (Figure 3),

$$H_{i,j} = \sum_{x,y} I(f_l(x,y) = i) I(x,y) \in R_j, i = 0, ..., n-1; j = 0, ..., m-1. \tag{3}$$

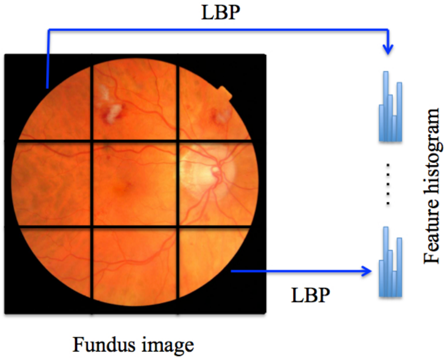

Fig. 3. A fundus image is divided into small regions. The LBP histograms are extracted and concatenated into a single one.

3.2 Machine Learning Algorithms

For the classification task we have used supervised learning, i.e. we have used a set of labeled images classified as healthy or sick. In next subsections we briefly describe artificial neural networks, random forest and support vector machines; we recommend the reader to review the references for details of these algorithms.

Artificial Neural Networks (ANN). This method has been one of the most used for learning tasks. ANN are based on the observation of biological neural systems which are formed by sets of units called neurons or nodes that are interconnected. Generally the architecture of an artificial neural network is divided into three parts called layers. These layers are the input layer, hidden layers, and the output layer. The way in which an artificial neural network works is as follows: each node (except for the input nodes) receives the output of all nodes in the previous layer and produces its own output, which then feeds the nodes in the next layer [9].

Random Forest (RF). This algorithm is an ensemble of unpruned classification trees, induced from bootstrap samples of the training data, using random feature selection in the tree induction process [2]. Prediction is made by aggregating the predictions of the ensemble. Random forest generally yields better performance than single tree classifiers such as C4.5.

Support Vector Machines (SVM). This method is based on the structural risk minimization principle from computational learning theory [3]. This principle provides a formal mechanism to select a hypothesis from a hypothesis space for learning from finite training data sets. The aim of SVMs is to compute the hyperplane that best separates a set of training examples. Two cases are analyzed: the linear separable case and the non-linear separable case. In the first case we are looking for the optimal hyperplane in the set of hyper-planes separating the given training examples. The optimal hyperplane maximizes the sum of the distances to the closest positive and negative training examples (considering only two classes). The second case is solved by mapping training examples to a high-dimensional feature space using kernel functions. In this space the decision boundary is linear and we can apply the first case. There are several kernels such as polynomial, radial basis functions, neural networks, Fourier series, and splines, among others; that are chosen depending on the application.

4 Experimental Results

We tested our method using images from the Messidor database[1]. These images were taken using a color video 3CCD camera on a Topcon TRC NW6 non-mydriatic retinograph with a 45 degree field of view. From the original data set of 1200 images, we selected a subset of 100 images at random. However, as we commented previously, specialists classify the images based on their ability and experience, but also considering the quality of them. Thus, we had to se-lected only 71 images due to their quality and visual information. Therefore we experimented with two data sets: 100 and 71 images.

In order to experiment with different binary patterns we used uniform patterns (59 bins), rotation invariant patterns (36 bins) and rotation invariant uniform patterns (9 bins) [12] considering 8 pixels in a 3×3 neighborhood.

We used artificial neural networks, random forest and support vector machines that are implemented in Weka[2] using default parameters and uniquely modifying the seeds in each iteration.

The metrics used to evaluate the performance of machine learning algorithms were accuracy and f-measure defined as $accuracy = (TP+TN)/(TP+FP+TN+FN)$, $f-measure = 2 \times (recall \times precision)/(recall + precision)$, $precision = TP/(TP+FP)$, $recall = TP/(FN+TP)$; where TP (True Positive) is the number of correct predictions of a positive example, FP (False Positive) is the number of incorrect predictions of a positive example, TN (True Negative) is the number of correct predictions of a negative instance and FN (False Negative) is the number of correct predictions of a positive instance.

Tables 1 and 2 show the results for each machine learning algorithm and local binary approach. These results were obtained by averaging ten runs of 10-fold cross-validation for each algorithm; that is, we randomly divided the original data set into ten equally sized subsets and performed 10 experiments, using in each experiment one of the subsets for testing and the other nine for training. The columns UP, RIP and RIUP denote uniform patterns, rotation invariant patterns and rotation invariant uniform patterns, respectively.

[1] MESSIDOR is a project within the scope of diabetic retinopathy. `http://messidor.crihan.fr/download-en.php`

[2] Weka is a collection of machine learning algorithms for data mining tasks. `http://www.cs.waikato.ac.nz/ml/weka/`

Table 1. Accuracy and F-measure for the algorithms using a data set of 100 images

	UP		RIP		RIUP	
	Acc.(%)	$F-m$	*Acc.*(%)	$F-m$	*Acc.*(%)	$F-m$
ANN	74.2	0.7408	77.6	0.7806	76.6	0.7668
RF	73.2	0.7274	74.2	0.7374	73.2	0.7306
SVM	77.2	0.7786	77.6	0.7822	78.8	0.7917

As we can see in Table 1, when using a data set of 100 images, the best classifier was support vector machines with 78.8% of accuracy and 0.7917 for $f-measure$ using rotation invariant uniform patterns. However, analyzing the results of the Table 2, we can observe that random forest obtained the best results with 97.4% of accuracy using uniform patterns and rotation invariant uniform patterns; and 0.974 for $f-measure$. We want to highlight that for this data set, all algorithms significantly improve their performance with over 93% for accuracy and 0.93 for $f-measure$. We can also comment that using a set of images with better quality for training, helps to obtain better results, i.e. the improvement of correct classification increases almost 18%.

Table 2. Accuracy and F-measure for the algorithms using a data set of 71 images

	UP		RIP		RIUP	
	Acc.(%)	$F-m$	*Acc.*(%)	$F-m$	*Acc.*(%)	$F-m$
ANN	93.2	0.9328	94.6	0.9468	94.9	0.9496
RF	97.4	0.9740	96.6	0.9660	97.4	0.9740
SVM	95.7	0.9580	95.7	0.9580	95.7	0.9580

5 Conclusions

We have presented a method for diabetic retinopathy detection using fundus images. Our preliminary experimental results show that the best machine learning algorithm was random forest with above 97% of accuracy. In addition, local binary patterns was useful in reducing data, in fact 9 bins using rotation invariant uniform pattern produced the best classification results. Future work includes: repeating the experiments with a larger data set, classifying several types of diabetic retinopathy, comparing other machine learning algorithms and feature extraction techniques, and finally to develop a software tool for specialists.

References

1. Bonafarte, S., et al.: Retinopatía diabética. Elsevier (1997)
2. Breiman, L.: Random Forests. Machine Learning 45(1), 5–32 (2001)

3. Burges, C.: A tutorial on support vector machines for pattern recognition. Proceedings of Data Mining and Knowledge Discovery 2, 121–167 (1998)
4. Ege, B.M., Hejlesen, O.K., Larsen, O.V., Moller, K., Jennings, B., Kerr, D., Cavan, D.A.: Screening for diabetic retinopathy using computer based image analysis and statistical classification. Computer Methods and Programs in Biomedicine 62(3), 165–175 (2000)
5. Gowda, A., et al.: Exudates detection in retinal images using back propagation neural networks. International Journal of Computer Applications 25 (2011)
6. Hem, K., et al.: Fluorescein angiosgraphy: A user's manual (2008)
7. International Diabetes Federation. IDF Diabetes Atlas153, 7–68 (2013)
8. Kavitha, D., Duraiswamy, K.: Automatic detection of hard and soft exudates in fundus images using color histogram thresholding. European Journal of Science Research 48, 493–504 (2011)
9. Mitchell, T.: Machine learning. McGrawHill (1997)
10. Niemeijer, M., et al.: Automatic detection and differentiation of drusen, exudates and cotton-wool spots in digital color fundus photographs for diabetic retinopathy diagnosis. Investigative Opthalmology and Visual Science 48 (2007)
11. Ojala, T., Pietikainen, M., Harwood, D.: A comparative study of texture measures with classification based on featured distribution. Pattern Recognition 29(1), 51–59 (1996)
12. Ojala, T., Pietikainen, M., Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence 24(7), 971–987 (2002)
13. Osareh, A., et al.: Automatic recognition of exudative maculopathy using fuzzy c-means clustering and neural networks. Medical Image Understanding and Analysis 3, 49–52 (2001)
14. Silberman, N., et al. Case for automated detection of diabetic retinopathy. In AAAI Spring Symposium on AI for Development, (2010)
15. World Health Organization. Diabetes mellitus. Media Centre (2014)