Iterative Kernel PCA based Classification of Retinal Images for Diabetic Macular Edema

Vipin Krishnan C V
PG Scholar-M.E,CS
Hindusthan college of Engineering & Technology
Coimbatore,India

V.S.Jayanthi, PhD.
Professor-ECE
Hindusthan college of Engineering & Technology
Coimbatore,India

Jestin V Kunjummen
Assistant professor-ECE
Hindusthan college of Engineering & Technology
Coimbatore,India

ABSTRACT
Swelling in the macular region of retina which is also known as macular edema, is a complication of the eye often leading to reduced capacity of vision. Diabetic retinopathy is also a severe complication to vision. In this work, iterative kernel based PCA is proposed which is a novel method used for the classification purpose in diseased retinal images. Exudate detection is carried out via a supervised learning approach using the normal fundus images. Feature extraction is introduced to capture the global characteristics of the fundus images and discriminate the normal from diseased images. The performance of the proposed methodology with the conventional PCA is evaluated based on classification accuracy. Experimental results shows the superior nature of iterative kernel based PCA in terms of performance measures.

General Terms
Feature extraction, Classification

Keywords
Diabetic Macular Edema; Fundus images; Hard exudates; Iterative kernel based Principal Component Analysis; Retinal images.

1. INTRODUCTION

Vision is the most advanced human sense. So images play the most important role in human perception. The human eye is nearly in the shape of a sphere. Its average diameter is approximately 20 mm. The eye is made up of three coats, enclosing three transparent structures. The outer most layers are composed of the cornea and sclera. The middle layer consists of the choroid, ciliary body, and iris. The innermost is the retina, which gets its circulation from the vessels of the choroid as well as the retinal vessels, which can be seen in an ophthalmoscope. Retinal inner most membrane central portion is the fovea. Fovea is a circular indentation with diameter 1.5mm

Diabetic retinopathy, is retinopathy [1] (damage to the retina) caused by complications of diabetes, which can eventually lead to blindness. It is an ocular manifestation of systemic disease which affects up to 80% of all patients who have had diabetes for 10 years or more. Despite these intimidating statistics, research indicates that at least 90% of these new cases could be reduced if there was proper and vigilant treatment and monitoring of the eyes. The longer a person has diabetes, the higher his or her chances of developing diabetic retinopathy.

Diabetic macular edema [2] advanced symptom of diabetic retinopathy. Macular edema occurs when fluid and protein deposits collect on or under the macula of the eye (a yellow central area of the retina) and causes it to thicken and swell. The swelling may distort a person's central vision, as the macula is near the center of the retina at the back of the eyeball. This area holds tightly packed cones that provide sharp, clear central vision to enable a person to see detail, form, and color that is directly in the direction of gaze. Swelling in the macular region of retina which is also known as macular edema, is a complication of the eye often leading to reduced capacity of vision. DME caused due to diabetes is a high risk complication which can cause irreversible loss of vision. Early detection of even a minor sign of DME is essential as it may also appear without any external symptoms. Once detected during retinal examination, it demands immediate treatment ranging from glycemic and blood pressure control, to laser surgery. DME is generally detected directly or indirectly. Direct ways are using stereoscopy (for manual examination) or optical computed tomography images. Indirect method is by detecting the presence of hard exudates (HE) in the retina. HE are formed due to secretion of plasma from capillaries resulting from the complications of retinal vasculature and could lead to retinal swelling. In color fundus images they appear as yellow–white deposits. Detecting the presence of hard exudates (HE) in different areas of retina is now considered a standard method to assess DME from color fundus images.

The severity of the risk of edema is evaluated based on the proximity of HE to the macula, which is defined to be a circular region centered at fovea and with one optic disc (OD) diameter. The risk for DME increases when the HE locations approach the macula, with the risk being the highest when they are within the macula. This is an important factor in DME assessment for further referral of the patients to an expert.

2. PROPOSED METHODOLOGY

The framework for the proposed scheme for classification and severity detection is shown in Figure 1. The abnormal retinal images. are collected from ophthalmologists for disease identification system. The images are initially processed to enhance the contrast in order to accurately detect the anatomical structures. An extensive set of features are extracted from these anatomical structures.

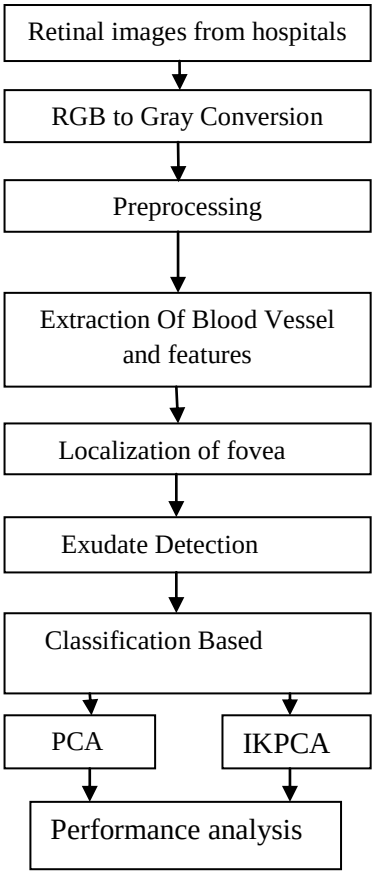


Figure 1. Flow diagram of the proposed work

The rest of this paper is organized as follows: Section III deals with the retinal image database and image pre-processing techniques, Section IV comprises the feature extraction methodologies and extraction of blood vessels Section V deals with the localization of fovea and Exudate detection and Section VI deals with PCA Classifier and Severity detection, Section VII shows Results and Discussions

3. RETINAL IMAGE DATABASE AND IMAGE PRE-PROCESSING

Most of the test images are collected from Eye Hospitals and MESSIDOR database. The image database that used in this work consists of digital retinal images obtained using the fundus florescent scan. The images are stored as colour JPEG images and are 2124×2056 pixels in size for all the objects. The intensity value of all the retinal images ranges from 0 to 255 (for each R, G & B planes). The real time images are collected for Diabetic Macular Edema having Cystoid macular edema

Pre-processing is an important and diverse set of image preparation program. Retinal images usually have pathological noise and various texture backgrounds, which may cause difficulties in extraction. So it should be removed. Due to correlation of color information in RGB space we first convert the color fundus image into Gray scale.. The raw retinal images usually have very low contrast which is signified by the grouping of large peaks in a small area on the histogram plot. The contrast of the retinal images is improved by histogram equalization which brings out details which are not clearly visible in the raw retinal images

4. EXTRACTION OF BLOOD VESSELS AND FEATURES

To classify images into normal or abnormal, it is represented by using relevant features.The purpose of feature extraction is to reduce the original data set by measuring certain properties, or features, that distinguish one input pattern from another pattern. The feature set should be selected such that the between-class discrimination is maximized while the within class discrimination is minimized.

Six features are used in this work among which five are based on texture of image (statistical features) and one is disease based features. The disease based features are extracted from fovea extracted images. Region based segmentation is done for the images. Even though seven statistical features are calculated, to the PCA classifier we are giving only the mean and median feature value

A. Statistical Features

Mean

It is the mean of pixels in image. The n^{th} moment of about mean is

$$\mu_{n=} \sum_{i=0}^{L-1}(z_i - m)^n \, p(z_i) \quad (1)$$

where m is the mean value of z(the gray level)

$$m=\sum_{i=0}^{L-1} z_i \, p(z_i) \quad (2)$$

Standard Deviation

$$SN = \frac{1}{N} \sum_{i=1}^{N}(x_i - x')^2 \quad (3)$$

Where x' is the mean value.

Variance

It is a measure of gray level contrast that can be used to establish descriptors of relative components.

$$\sigma(z)=\mu_2(z) \quad (4)$$

Entropy

$H(z) = -\int p(z) \ln p(z) dz$ is Shannon's entropy of the image window z, and p is the distribution of the gray levels in the considered window.

Energy

Returns the sum of squared elements in the Gray Level Co-occurrence matrix(GLCM)and range will be in [0 1].

$$\text{Energy} = \sum_{i,j} p(i,j)^2 \quad (5)$$

where *i* the number of rows, j is the number of columns and *p(i,j)* is the mean of the GLCM of the images

B. Disease based features

Area

It gives the area of the disease spread and objects in the eye.

$$\rho(i,j) = \frac{1}{M-N}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} b_i \, (i-m, j-n) \quad (6)$$

which supports M,X,N region for every (i,j) point of the image.

C. Blood vessel extraction

STEP1:To reduce correlated color information ,convert RGB to Gray
STEP2:Appaly morphological opening operation with disc shaped structuring element on gray scale image to reduce noise
STEP3:Use morphological closing operation to remove vessel
STEP4:Use Top-Hat transform to extract the vessel like structure
STEP5:Binarize the resultant image by thresholding
STEP6:By connected component analysis reduce the noise of arbitrary shape

5. LOCALIZATION OF FOVEA AND EXUDATE DETECTION

To localize the fovea region, we start with the image, containing only the blood vessels. Let, G be the approximated center of the OD. P be the point on the horizontal line passing through the center at a distance $2.5\times d$ in the direction of centroid. As indicated in the literature, P lies in the vicinity (may be above/below/within) of the macula region. In order to extract the fovea region, a strip of width k pixels through the point P (take P as middle of the strip) in a direction perpendicular to the line GP is considered

We rely on the fact that the fovea region is free from any vessel. A sliding window of size k×k is applied along the strip starting from point P in upward and down ward direction. A chain of number is obtained where the number denotes the count of black pixels lying in the window. Finally, the maximum run length of zeros in the chain enables us to localize the fovea region.

Let, S and E be the start and end position of the maximum run length of zero's respectively and D be the mid position of S and E. The circular region with radius DS and centered at D is the region of interest. To determine the said region of interest, a binary image *BW* of size same as that of the input image with single black pixel at position D and it is dilated by a flat disc of radius DS to obtain *BWd*. Pixels extracted from original gray-scale image, corresponding to the black region in *BWd* will form the region of interest

STEP 1 Locate a point P horizontally at a distance 2.5×d from G towards the centroid.

STEP 2 Consider a vertical strip of width k pixels around P perpendicular to GP.

STEP 3 Apply a k×k sliding window along the strip and form the chain of numbers denoting the black pixels in the window.

STEP4 Find the maximum run length of zeros, L in the number chain.

STEP 5 Let S and E are the start and end position corresponding to L and D is the mid position of S and E.

STEP 6 Consider a binary image *BW* of size same as the input image with only a black pixel at position D. Dilate *BW* by a disc of radius DS to obtain *BWd*.

STEP 7 Obtain R as the portion of the gray-scale image, *I*1 corresponding to the black region in *BWd*.

STEP 8 Binarize R to approximate macula region.

STEP 9 Refine binarized R by removing noise and fitting the circle to obtain final macula region.

6. PCA BASED CLASSIFIER

PCA of a multivariate Gaussian distribution centered at (1,3) with a standard deviation of 3 in roughly the (0.878, 0.478) direction and of 1 in the orthogonal direction. Principal component analysis (PCA) is a mathematical procedure that uses orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components.

7. PCA ALGORITHM FOR IMAGE CLASSIFICATION

The following steps is the algorithm for PCA

STEP 1 Calculate the mean of feature values extracted.
STEP 2 Subtract the mean from each image feature values.
STEP 3 Form the covariance matrix.
STEP 4 Compute the Eigen values.
STEP 5 Compute the Eigen vectors.
STEP 6 First principal component is taken based on maximum variance for classification purpose.

8. IKPCA BASED CLASSIFIER

Conventional PCA fails to represent the underlying nonlinear structure of data, for it takes only second order correlations into account As a natural nonlinear extension of PCA iterative kernel PCA computes the principal component in a possibly high dimensional feature space.

9. IKPCA ALGORITHM FOR IMAGE CLASSIFICATION

The following steps is the algorithm for IKPCA

STEP 1 Calculate the mean of feature values extracted.
STEP 2 Subtract the mean from each image feature values.
STEP 3 Kernel matrix is created by taking the mask over the region in the image.
STEP 4 Eigen values and Eigen vectors are calculated by iterative or adaptive kernel methods.
STEP 5 Principal components are obtained from kernel matrix.

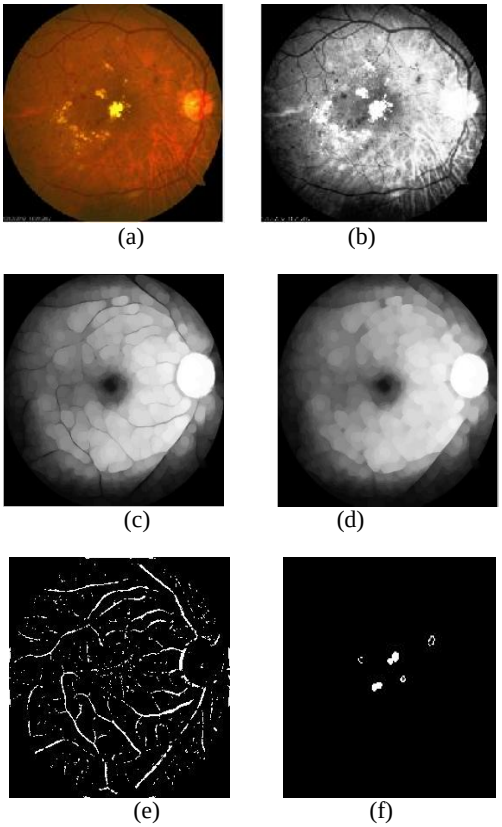
10. RESULTS AND DISCUSSIONS

Experiments are conducted on real time retinal images collected from hospital. After pre-processing, the features mentioned in section IV are extracted from segmented images.

Figure 2 shows PCA based exudates detection in which input image is enhanced, then morphological operation and

blood vessel extraction is carried out and it is given to PCA classifier.

Figure 3 shows IKPCA based exudates detection in which input image is enhanced, then morphological operation and blood vessel extraction is carried out and it is given to IKPCA classifier which qualitatively shows more area of exudates detection.

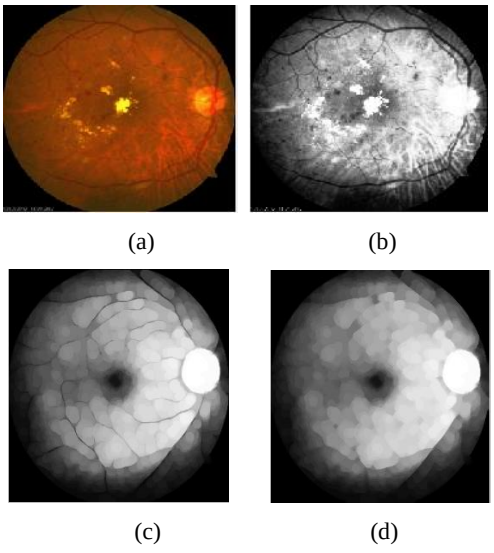


(a) (b)

(c) (d)

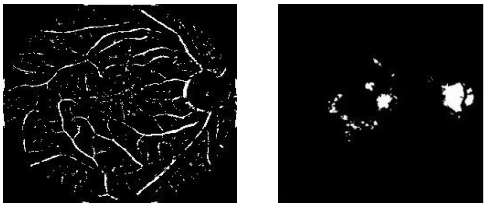
(e) (f)

Figure 2 (a) Input image (b) Enhanced image (c) Morphological open (d) Morphological close (e) Blood vessel extraction (f) PCA based exudate detection after fovea localization.



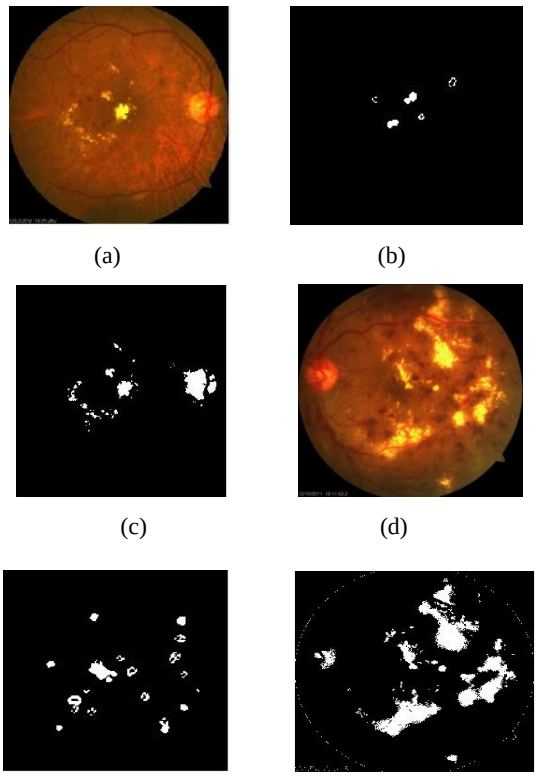
(a) (b)

(c) (d)



(e) (f)

Figure 3: (a) Input image (b) Enhanced image (c)Morphological open (d) Morphological close (e) Blood vessel extraction (f) IKPCA based exudates detection.



(a) (b)

(c) (d)

(e) (f)

Figure 4: Qualitative analysis of (a) Input image 1 (b) PCA based classification for input image 1 and (c) IKPCA based classification for input image 1 (d) Input image 2 (e) PCA based classification for input image 2 (f) IKPCA based classification for input image 2

Table 1 Performance analysis of PCA and IKPCA

	PCA	IKPCA
Classification Accuracy (%)	86.5	92.50

Classification accuracy means the number of correctly classified images to the total number of images.

From the quantitative and qualitative analysis of the images shown in figure 4, it is very clear that classification accuracy and more exudate detection is made possible with IKPCA compared to PCA and it gives the overall exudates present in an abnormal fundus image. Out of 200 images in the real time dataset, 173 images are classified correctly, and attained classification accuracy of 86.5% by PCA and 185 images are classified correctly which leads to classification accuracy of 92.5% by Iterative kernel based PCA.

11. CONCLUSION AND FUTURE WORK

In this work, iterative kernel based PCA is used for the classification of diseased retinal images from normal fundus images. IKPCA gives higher exudate detection by creating a mask or kernel over the diseased area thereby improves the classification accuracy. The performance of the proposed methodology was evaluated with the conventional PCA based on classification accuracy. Qualitative and quantitative analysis shows the superior nature of iterative kernel based PCA in terms of performance measures.

As a future work, feature optimization algorithms can be used for the retinal images and improve the performance parameters.

ACKNOWLEDGEMENT

The authors wish to thank Ahalia Hospital, Kerala, India for providing the real time test images to carry out the research work.

12. REFERENCES

[1] Nathan silberman ,Lakshminarayanan Subramanian " case for automated detection of diabetic retinopathy"advancement of Artificial Intelligence 2010

[2] K. Sai Deepak and Jayanthi Sivaswamy, *Member* IEEE "Automatic Assessment of Macular Edema From Color Retinal Images" , IEEE Transaction on medical imaging,VOL.31, NO. 3 march 2012

[3] C. P. Wilkinson, F. L. Ferris, R. E. Klein, P. P. Lee, C. D. Agardh, M.Davis, D. Dills, A. Kampik, R. Pararajasegaram, and J. T. Verdaguer,"Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales," Am. Acad. Ophthalmol., vol. 110,no. 9, pp. 1677–1682, Sep. 2003.

[4] A. Rocha, T. Carvalho, S. Goldenstein, and J. Wainer, Points of interest and visual dictionary for retina pathology detection Inst.Comput., Univ. Campinas, Tech. Rep. IC-11-07, Mar. 2011.

[5] K. Ram and J. Sivaswamy, "Multi-space clustering for segmentation of exudates in retinal color photographs," in Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., Sep. 2009, pp. 1437–1440

[6] K. S. Deepak, G. D. Joshi, and J. Sivaswamy, "Content-based retrieval of retinal images for maculopathy," in Proc. 1st ACM Int. Health Inf.Symp., 2010, pp. 135–143

[7] R. F. N. Silberman, K. Ahlrich, and L. Subramanian, "Case for automated detection of diabetic retinopathy," Proc. AAAI Artif. Intell. Development (AI-D'10*)*, pp. 85–90, Mar. 2010.

[8] P.M.Narendra, and K. Fukunaga, "A branch and bound algorithm for feature selection," *IEEE Transactions on Computers*, 26(9**)**, pp.917-922 1977.

[9] M. Verma, R. Raman, and R. E.Mohan, "Application of tele ophthalmologyin remote diagnosis and management of adnexal and orbital diseases," Indian J. Ophthalmol., vol. 57, no. 5, pp. 381–384, Jul. 2009

[10] C.Sanchez,A.Mayo,M.Lopez 'Automatic image processing algorithm to detect hard exudates based on mixturemodels'in.proc.28thannu.inf.conf.IEEE.Eng.Med. Biol.sep 2006

[11] J.Davidson,T.Ciulla 'How the diabetic eye loss the vision' Endocrine vol.32 Nov 2007

[12] C. Agurto, V. Murray, E. Barriga, S. Murillo, M. Pattichis, H. Davis,S. Russell, M. Abramoff, and P. Soliz, "Multiscale am-fm methods for diabetic retinopathy lesion detection," *IEEE Trans. Med. Imag.*, vol.29, no. 2, pp. 502–512, Feb. 2010.

[13] M. Kudo and J. Sklansky. "Comparison of Algorithms that select features for pattern classifiers," *Pattern Recognition*,33, pp.25-41, 2000.

[14] K. S. Deepak, G. D. Joshi, and J. Sivaswamy, "Content-based retrieval of retinal images for maculopathy," in Proc. 1st ACM Int. Health Inf.Symp., 2010, pp. 135–143

[15] Meindert Niemeijer, Micheal D. Abramoff, Member, IEEE,and Bram van Ginneken, Member, IEEE."Segmentation of the optic disc,macula and vascular arch in fundus photographs",IEEE Transactions on Medical imaging,Vol.26,No.1,pp.116-127,January 2007.