Bionature, 40(1) 2020 : 17-28 ISSN: 0970-9835 (P), 0974-4282 (O)

AUTOMATIC DETECTION OF RETINAL MICROANEURYSMS USING HYBRID KERNEL SVM CLASSIFIER

K. SINDHU[1*] AND S. A. PRAYLIN SELVA BLESSY[1]
[1]Department of Electronics and Communication Engineering, Bethlahem Institute of Engineering, Tamil Nadu, India.
Email: sindhukjd@gmail.com

Received: 15 March 2020
Accepted: 20 May 2020
Published: 24 May 2020

Original Research Article

ABSTRACT

Retinal microaneurysms(MA) are the earliest clinical sign of diabetic retinopathy (DR) disease. DR is diagnosed by inspecting fundus images. Early detection of DR protects patients from losing their vision. In this paper an automatic method for detection of MA is proposed. This system utilizes preprocessing of retinal image to correct for non-uniform illumination and enhance contrast in the area of interest. The detection of MAs using Local Convergence index Features (LCF) and Hybrid kernel Support Vector Machine (HKSVM) classifier tells the probability of being actual MAs. The intensity based features and shape based features are extracted and combined using ensemble classifier positives will be reduced during the classification phase. The MA candidates are extracted and classified using HKSVM Classifier. The proposed method has been evaluated by public databases: Retinopathy Online Challenge (ROC) and e-optha. The efficiency and effectiveness of the method processed has been demonstrated by experimental results and thus proving its potential as a diagnostic tool for DR.

Keywords: Diabetic retinopathy; microaneurysms; support vector machines; fundus images.

ABBREVIATIONS

AUC : Area under the Curve
CI : Convergence Index
DR : Diabetic Retinopathy
GW : Gradient Weighted
KNN : K-Nearest Neighbors
LCF : Local Convergence index Features
MA : Microaneurysms
RUS : Random under Sampling Technique
ROC : Retinopathy Online Challenge
HKSVM : Hybrid Kernel Support Vector Machine

INTRODUCTION

People with diabetes have an eye disease is called diabetic retinopathy (DR) which affects the retinal parts. It is the most common cause of vision impairment and blindness among the middle-aged population in the world. The uncontrolled and chronically high blood sugar can cause increased damage to tiny blood vessels that can lead to diabetic retinopathy. Lee et al. [1] Microaneurysms (MAs), is small swelling, red isolated dot near the tiny blood vessels sometimes leaking fluid and blood into the retina. MAs are rounded small shape structures which are less than 125 μm in diameter. The detection of MAs is considered as one of the most important clinical strategies for the early diagnosis of DR and blindness prevention in a cost-effective health care practice.

The ever increasing number of patients globally makes it imperative that a reliable

and early diagnostic method is identified at the earliest to assist in detection of diabetic retinopathy.

Automatic detection of diabetic retinopathy will lead to a large amount of savings of time and effort. T. Spencer et al. [2] proposed a mathematical technique to segment MA within fluorescein angiogram. J.H. Hipwell et al. [3] sed Gaussian matched filters to retain candidate MA for classification. Usher et al. [4] had classified MA based on a combination of Recursive Region Growing segment MA candidate regions and k-nearest neighbors (KNN) while Niemeijer et al. [5] suggested that a hybrid between mathematical morphology techniques and a pixel classification method be employed in candidate extraction and the true MA is identified using a KNN classifier with a set of shape and intensity features. The Seiffert et al. [6] discriminate MAs from non-MA candidates, using a hybrid sampling/boosting algorithm, called RUSBoost. Maher et al. [7] has already evaluated a decision support system for automatic screening of non-proliferative diabetic retinopathy and support vector machines were employed in the automated diagnosis of non-proliferative diabetic retinopathy [8]. Zhang et al. [9] proposed a hierarchical approach based on multiscale correlation filtering is used to detect all MAs from color retinal images. Giancardo et al. [10] detected MA candidates using a thresholding technique followed by a Radon transformation at various scanning angles and support vector machine and feature extracted from random space to identify MAs.

This paper is devoted to detecting MAs which are present from the earliest stages of DR and remain throughout the development of the disease. In the first stage of preprocessing we are improving the contrast and non-illumination of the retinal image for candidate extraction. Candidates are then extracted using a gradient weighting technique and an iterative thresholding approach. The next stage utilizes intensity and shape descriptors and in addition a new set of features based on local convergence index filters. The final stage involves the feeding of the collective set into hybrid sampling/booster classifier that discriminates between MAs from non MA candidates and method is evaluated with images of different resolutions and modalities.

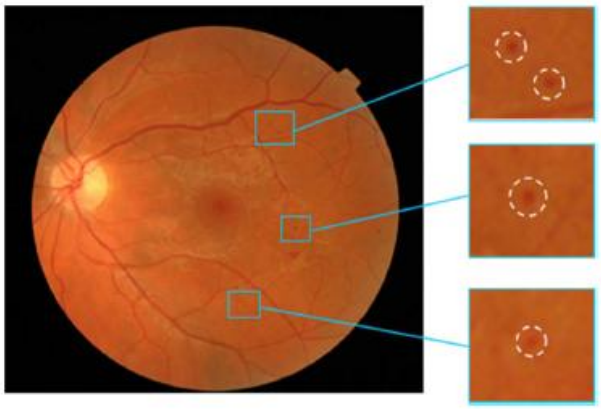

Fig. 1. An example of fundus image with several MAs

MATERIALS AND METHODS

Database

The proposed method has been tested on two public databases: e-ophtha and Retinopathy Online Challenge (ROC) database Niemeijer et al. [11]. The e-ophtha is collected by Decenciere et al. [12]. E-ophtha-MA comprises 148 fundus images with microaneurysms or small haemorrhages, which are manually commented by ophthalmology experts. In this paper, 74 fundus images are used as training set and the remaining 74 fundus images as test set. The ROC database is made up to 50 images for training and 50 images for testing and only the annotations of the training set are available to the public. Besides, Only 37 images of training set contain MAs, while the other 13 images do not contain any MA. Thus, in this paper, the hybrid kernel SVM applied to evaluate the ability of classifiers when we use this database. The proposed framework is implemented in MATLAB 2010 Rb software.

Method

DR is the leading cause of blindness in diabetic patients and chronic high blood sugar is associated with blood vessel damage which in the DR is characterized by increased leakage of retinal blood vessels in the early stage and late stage by formation of ne abnormal blood vessel that leads to scarring, cell loss in the retina. The diagnosis uses Digital retina fundus images, the color images of retinal fundus, tiny reddish isolated dots near blood vessels indicates MA. As the diagnosis depends on the image, the properties of the image as well as the imaging device such as resolution, modality, compression technique, illumination and contrast variation play a very crucial role. To discriminate MAs from non-MA candidates in fundus retinal images, we have proposed a system using Local Convergence index Features (LCF), and Hybrid Kernal Support Vector Machine (HKSVM) classifier. The method proposed is a reliable and cost effective which can aid ophthalmologists in early detection of DR. The flow chart of the proposed frame work is depicted in Fig. 2.

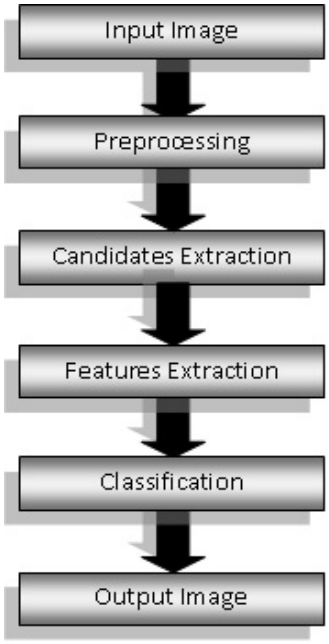

Fig. 2. Diagram of proposed framework

Image Preprocessing

The quality of a fund us image may vary due to factors such as eye movement, media opacity, small pupils, nonalignment of the camera; problems in-camera focusing and noise at the image acquisition stages. These factors actually affect image contrast, illumination, and color. Therefore, fundus image preprocessing is quite a mandatory step in computer-aided design systems. Steps of image preprocessing are explained below:

In the proposed method, the green channel of fundus images is used as the main input. The green channel provides the best MA-background contrast, while the red channel is saturated. The blue channel is the darkest color, channel and does not contain any information. Walter et al. [13] explained why the green channel contained most of the image information. The Green channel of images in RGB provides better contrast between microaneurysm and background, only use the green channel. In the acquisition process, the images are non-uniformly illuminated and exhibit local luminosity and contrast variability. In order to make the MA detection more robust, each image is preprocessed which normalizes luminosity and contrast. Luminosity and contrast variability in the background is estimated and then used for normalizing the whole image as follows:

$$I_N(x,y) = \frac{I_G(x,y) - L(x,y)}{C(x,y)} \quad (1)$$

Where C(x, y) and L(x, y) represent the deviation factor of contrast and lightness respectively, which can be estimated by the standard deviation and mean of the background pixels in the image. The visibilities of the lesions are improved by local reduction of luminosity and contrast variation of retinal images.

Candidate Extraction

The candidate extraction is one of the main steps in the automated detection of MAs. Candidate extraction is to reduce the computational burden by decreasing the number of objects for further analysis. Using this method unwanted regions other than the abnormalities are ignored. The candidate regions are identified by taking profile of the region along different directions with the predefined conditions. If the region is the candidate region, then it at least has one maximum due to the fact that MAs appear as the bright pixel. To overcome the false identification of noise region as candidate region, maximum values at different orientation is observed.

A Multi-scale multi-orientation gradient weighting and iterative thresholding technique is used to extract MA candidates. A multi-scale multi-orientation, weight for each pixel in a normalized image is computed on the gradient magnitude. Gradient magnitude values for different scales and different orientation are obtained by the convolution of the image. The weight of a pixel is inversely related to the gradient values of the pixel location, for pixels with small gradient magnitude (smooth regions), the weight is large, and for the pixels with large gradient magnitude (as those on the edges), the output weight is small. The gradient weighted image is thus obtained by

$$I_w(x,y,\sigma_G,\theta_G) = \frac{1 - I_M(x,y,\sigma_G,\theta_G)}{1 + I_M(x,y,\sigma_G,\theta_G)} \quad (2)$$

Samples of gradient weighted images of a small patch for different scales and orientations. The final multi-orientation GW image is defined as the sum of the individual gradient-weighted images in all orientations for each scale separately via

$$I_{WO}(x,y,\sigma_G) = \sum_{\theta_G=0}^{\frac{\pi}{2}} I_W(x,y,\sigma\ ,\theta_G) \quad (3)$$

The candidate is then extracted by multi-orientation gradient-weighted images are used in the successive thresholding process. To include all micro aneurysms sizes, several scales of multi-orientation gradient-weighted images are obtained and the final image is obtained as a summation of overall selected scales:

$$I_{WOS}(x,y) = \sum_{\sigma G=1}^{5} I_{WO}(x,y,\sigma_G) \qquad (4)$$

Green channel is mostly integrated among RGB dataset as it gives the best contrast between microaneurysms and background.

Iterative thresholding is applied on the multi-orientation GW images to obtain binary images corresponding to different threshold. Using the connected component analysis in image, the objects with a hole inside will satisfy the area, eccentricity and extent constraints are identified.

Feature Extraction

Extracting suitable features and descriptors for the candidate regions is an important step for the final classification stage. Since the MAs appear in different colors and sizes, several shape and intensity features are extracted. The feature set is completed by including the responses and the estimated radii of different LCF.

Intensity Based Features

These features are described indicating the darkness of MAs compared to their neighborhood background. Intensity features are extracted inside the center location of the whole candidate object. Also in a square neighborhood region, this is 3 times as large as the candidate area. The neighborhood region is centered on the center of candidate. For every candidate, the average, maximum and minimum of the green intensity values are obtained for the candidate and neighborhood regions, separately.

Shape Based Features

Shape-based features are extracted for each candidate region. From MAs are small and they appear as round structures with a diameter less than 125 µm, the following shape-based features are extracted for each candidate region:

- Area (S_{Area}): Area of the candidate region specified by the actual number of pixels.
- Convex area (S_{Cona}): Area of candidate convex region specified by the actual number of pixels.
- Solidity (S_{Sol}): Ratio of the area of candidate (S_{Area}) over the convex area (S_{Cona}).
- Extent (S_{Ext}): Ratio of the area of the pixels in the bounding box as shown by red color.
- Perimeter (S_{Per}): Distance around the boundary of the region by calculating the distance between each adjoining pair of pixels.
- Circularity (SCird): Diameter of a circle with the same area +as the region which is equal to $\sqrt{4S_{Area}} / \Pi$.
- Ellipticity ($S_{AxiA,}$ S_{AxiB}): Lengths of the major and minor axes of the ellipse that has the same normalized second central moments as the candidate region. The major and minor axes are depicted by red lines.
- Eccentricity (S_{Ecc}): Ratio of distance between the foci and the major axis length (SAxiA) of the ellipse with a same 2nd moment as the region.
- Euler number (S_{Eul}): Number of objects in the region minus the number of holes in those objects.

LCF Based Feature

The LCF filters work on the basis of gradient convergence and not intensity and hence can detect low contrast MAs. The convergence evaluation in a regional band allows the reduction of uncertainty caused

by noise. The convergence index (CI) filters are suitable for the detection of convex shapes and objects with a limited range of sizes. The CI filters detect and expresses the convergence, degree of gradient vectors within the local area. Given an input image I(x, y), for each pixel with spatial coordinates (x, y), the CI is defined by

$$CI(x,y) = \frac{1}{M}\sum_{(\theta i,m)es} \cos(\varphi(x,y,\theta_i,m)) \quad (5)$$

Where M is the number of points in the filter support region S, and ϕ (x, y, θ_i, m) is the orientation angle of the gradient vector of the polar coordinate (θ_i,m) with respect to the line, with the direction i, that connects (θ_i,m) to (x, y). The output of CI falls between −1 and +1. At the maximum value of +1 all the gradient vectors in S point toward the pixel of interest. This occurs if equi-contours of the intensity in the neighborhood of the pixel of interest are concentric. The angular difference ϕ is given by

$$\varphi(x,y,\theta_i,m) = \theta_i - \alpha(x,y,\theta_i,m), \alpha(x,y,\theta_i,m) \quad (6)$$

$$= \tan^{-1}\left(\frac{\frac{\partial}{\partial x} I(x+m\sin(\theta_i), y+m\cos(\theta_i))}{\frac{\partial}{\partial y} I(x+m\sin(\theta_i), y+m\cos(\theta_i))}\right) \quad (7)$$

Here α is image gradient orientation within the convergence filter support region. The support region polar coordinates are denoted by the radial coordinate m, the distance from point of interest (x, y) in pixel, and angular coordinate θ_i which is sampled with N equally spaced radial lines(θ_i = 2π / N (i − 1), i ∈ {1, ..., N}). The set of radial lines is emerging from the point where the filter is being applied to, and is equally distributed over a circular region centered at the point of interest (x, y). The result of applying on the input image I is the filter's response image. The radius of support region at that location for each candidate point can be obtained and shape estimation is done by finding radius of ring support region for each candidate in the given image.

Classification

To discriminate MAs from non-MA candidates, use a machine learning HKSVM classifier. The HKSVM networks are supervised learning models with associated learning algorithms that analyze data and recognize patterns, used for classification. HKSVM is to classify images into 2 classes. It is robust against outliers and over-fitting. There are two phases in the SVM namely, Learning/Training phase and Testing Phase. In the learning phase the classifier is made to learn the known set of images, here the feature vector of each known image is fed to the classifier and the output is labeled accordingly. The Testing Phase is wherein the unknown images feature vector is fed to the classifier and based on its mapping with the Learning Phase, the image is classified appropriately.

The HKSVM performs classification by constructing an N-dimensional hyper plane that optimally separates the data into two categories. A high dimensional feature space is created by mapping the input space and then a hyper plane is constructed that maximizes the margin of separation between the classes. The support vectors are pointing that lies close to the decision surface and have a direct effect on their location. When the classes are non-separable, the optimal hyper plane is the one that minimizes the probability of

classification error. So the goal of HKSVM modelling is to find the optimal hyper plane that separates clusters of vector in such a way those cases with different target varying lay on either side of the plane and the vectors near hyper plane are support vectors. In simple words, given a set of training examples, each marked as belonging to one of two categories, an HKSVM training algorithm builds a model that predicts whether a given new sample falls into one or the other category. To fit nonlinear curves to the data, HKSVM makes use of a kernel function to map the data into a different space where a hyperplane can be used to do the separation. In this work, we have used hybrid kernel which is given by,

$$K(x, x') = (x.x'+1)^{d} \quad (8)$$

Where x and x' are the training vectors, d is the kernel parameter.

RESULTS AND DISCUSSION

The proposed method is in the form of a pipeline of techniques where the performance of each step depends on the output of the previous step. Hence, the method is validated extensively in each one of three stages: candidate extraction, MA detection, and image classification. The Fig. 3 shows the experimental results of the proposed work.

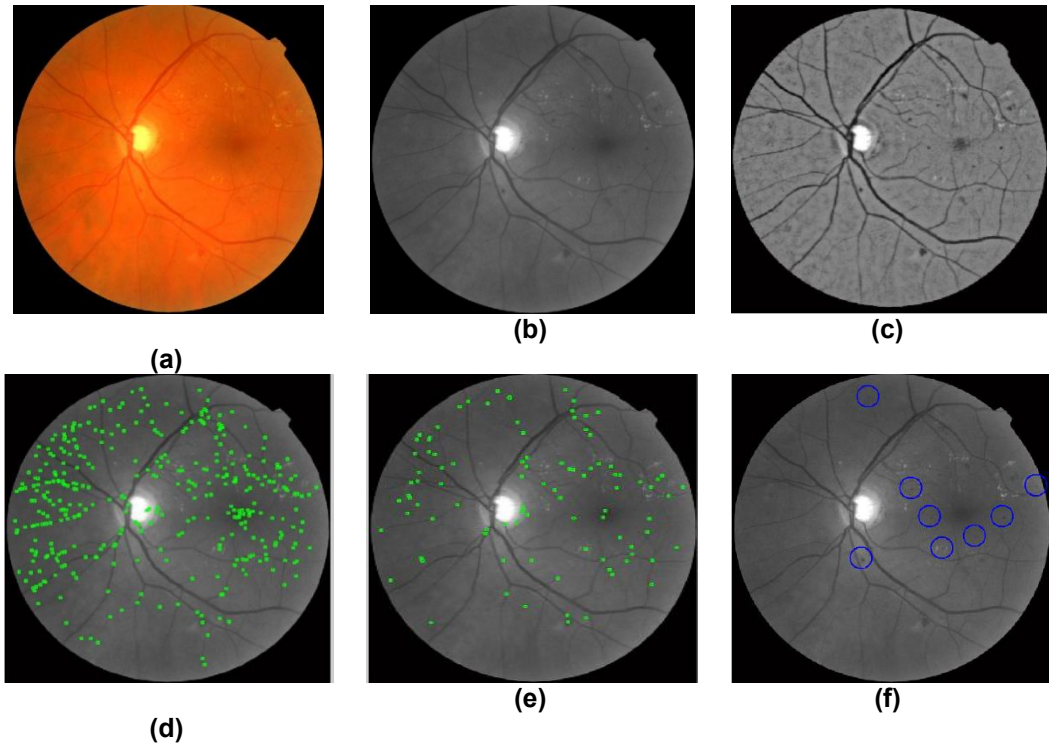

Fig. 3. Experimental results of (a) original retinal input image, (b) preprocessed green channel image, (c) normalized green channel image, (d) candidates extraction (e) MA candidate extraction (f) identification of MA or Non-MA

In order to make the MA detection more robust, each image is preprocessed which normalizes luminosity and contrast. The Fig. 3(a) shows digital color fundus images, MAs appear as tiny, reddish isolated dots near tiny blood vessels Fig. 3(b) shows the Green channel of images in RGB provides a better contrast between microaneurysm and back-ground, only use the green channel. As a result of the acquisition process, images are non-uniformly illuminated and exhibit local luminosity and contrast variability. The Fig. 3(c) shows the normalized green channel image. The Fig. 3(d-f) shows Candidate extracted output and identification of MA.

The sensitivity, specificity and accuracy measures were used to quantify the results achieved by the detection method. These measures were chosen based on the related work, where most publications use these measures. For a clear understanding of these measures, it is necessary to conceptualize some classifications used in laboratory tests for detection of diseases. When a test result is positive, the individual can manifest the disease, which is called "True Positive" (TP) or cannot, express it, which is known as "False Positive" (FP). On the other hand, when the result is negative, the individual cannot have the disease, which is known as "True Negative" (TN) or can have it, is called "False Negative" (FN).

The sensitivity is defined as the ability of a test to detect correctly people with a disease or condition. The sensitivity is calculated according to Eq.(9)

$$Sensitivit\ \ y = \frac{TP}{TP\ +\ FN} \qquad (9)$$

The specificity is defined as the ability of a test to exclude properly people without a disease or Condition. The specificity is calculated according to Eq.(10)

$$Specificit\ \ y = \frac{TN}{TN +\ FP} \qquad (10)$$

The accuracy of the given system is calculated according to Eq.(11)

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} \qquad (11)$$

It is important to have some balance between the values of sensitivity and specificity. Get high values for sensitivity and low values for specificity means that the method detected most MAs (true positives), but other structures were erroneously classified as MAs (false positives). On the other hand, get high values for specificity and low values for sensitivity means that the method deleted correctly regions without MAs (true negative) but was not able to identify all MAs (true positives).The table shows the evaluated results of existing methods with our proposed method.

Table 1 List some of the existing methodology for MA detection and our proposed method using public databases. Accurate localization of the MAs is essential for the diagnosis, it depends on the performance of the classifier which is evaluated using Receiver Operating Characteristics (ROC). Area under Curve (AUC) value for HKSVM is 93.83% and for RUSBoost it is 90.01% of the same training samples shown in Fig. 4. The proposed candidate extractor out performs the sensitivity value of 0.85 on the ROC dataset. The average of false positives per image is higher than the ones reported by other methods; however, it is less than 0.06% of the total number of pixels in one image. The true MAs as much as possible without considering the balance between sensitivity and specificity. The false positives are later discarded in the classification step.

Table 1. Results for different existing methods with proposed method for MAs detection

Authors	Dataset	Candidates extraction method	Classifier	Accuracy	Sensitivity	Specificity
Su et al. [14]	ROC, DiaretDB1	Dark object filtering technique	KNN classifiers	__	51.7%	__
Yuji et al. [15]	DiaretDB1	The double-ring filter, shape based on the Hessian matrix, and Gabor filter	DCNN classifier	__	84%	__
Kumar et al. [16]	DiaretDB1	PCA	SVM Classifier	_	96%	92%
Jiawei et al. [17]	Grampian	Microaneurysms turnover and pathological risk factors	SVM Classifier	__	89%	88%
Ling et al. [18]	DiaretDB1	Image-to-text mapping Model	CNN Classifier	96.1%	87.7%	__
Proposed Method	ROC, e-optha	Gradient weighting technique and an iterative thresholding	HKSVM Classifier	97.2%	98.5%	98%

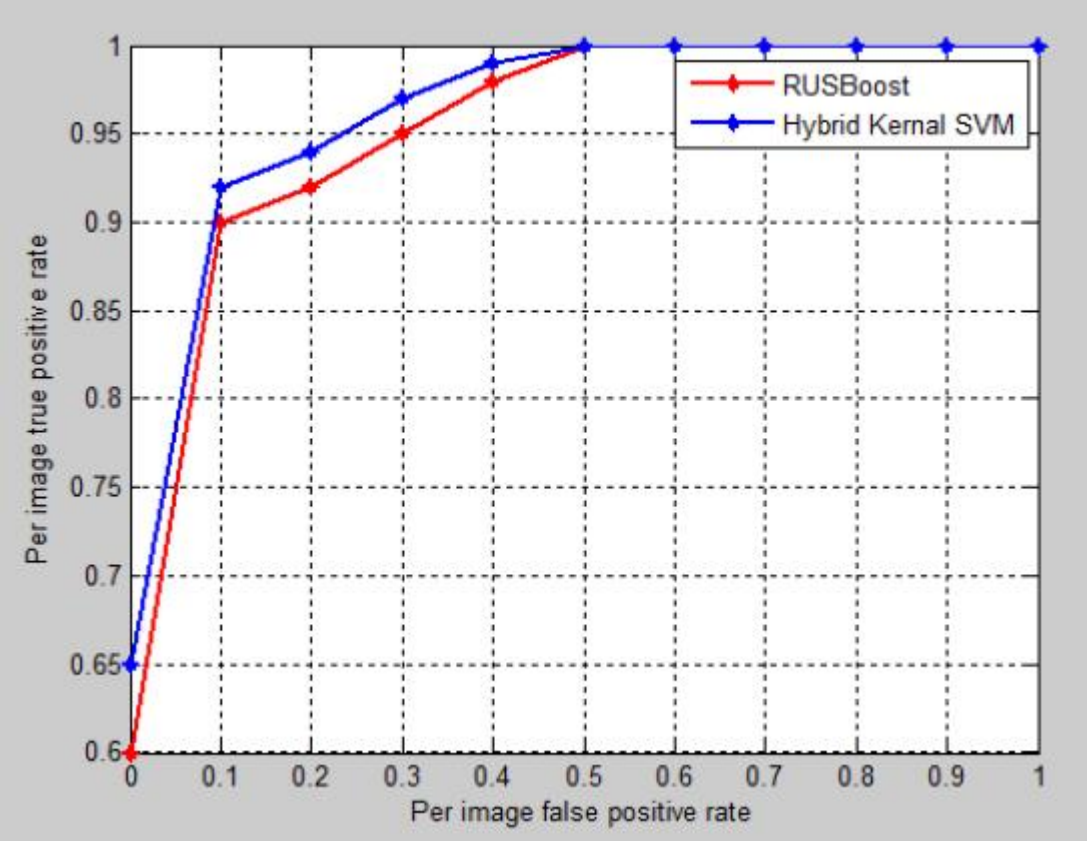

Fig. 4. ROC Plot for RUS boost and HKSVM

CONCLUSION

In this paper the input images from the database have poor quality. So pre-processing is done. For classification purposes features were extracted from pre-processed image. DR has been classified and grading is given using HKSVM. The research findings indicate that the performance of the proposed hybrid kernel SVM classification algorithm in terms of accuracy was improved and better when compared to the other existing data classification algorithms and this system give a successful diagnosis with accuracy 97.2%. The methodology can also be extended for the detection of other abnormal conditions due to DR like haemorrhages, and their augmentation with respect to time for grading the disease progression based on the measurement.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Lee R, Wong TY, Sabanayagam C. Epidemiology of diabetic retinopathy, diabetic macular edema and related

vision loss. Eye and Vision. 2015; 2(1):17.
2. Spencer T, Olson JA, Hardy MC, Sharp PF, Forrester JV. An image processing strategy for the segmentation and quantification of micro-aneurysms in fluorescein angiograms of the ocular fundus comp. Biomed Res. 1996;29:284-302.
3. Hipwell JH, Strachan F, Olson JA, MCHardy KC, Sharp PF, Forrester JV. Automated detection of micro aneurysms in digital red-free photographs: A diabetic retinopathy screening tool. Diabetic Medicine. 2000;17:588–594.
4. Usher D, Dumskyj M, Himaga M. Automated detection of diabetic retinopathy in digital retinal images: A tool for diabetic retinopathy screening. Diabetic Medicine. 2002;21:84–90.
5. Niemeijer M, Van Ginneken B, Staal J, Suttorp-Schulten MSA, Abr`amoff MD. Automatic detection of red lesions in digital color fundus photographs. IEEE Transactions on Medical Imaging. 2005;24(5):584–592
6. Seiffert C, Khoshgoftaar TM, Hulse JV, Napolitano A. RUSBoost: A hybrid approach to alleviating class imbalance. IEEE Trans. Syst., Man, Cybern., Syst. 2010;40(1):185–197.
7. Maher RS, Kayte SN, Meldhe ST, Dhopeshwarkar M. Automated diagnosis non-proliferative diabetic retinopathy in fundus images using support vector machine. International Journal of Computer Applications. 2015;125(15):7–10.
8. Maher R, Kayte S, Panchal D, Sathe P, Meldhe S. A decision support system for automatic screening of non-proliferative diabetic retinopathy. International Journal of Emerging Research in Management and Technology. 2015;4(10):18–24.
9. Zhang B, Karray F, Li Q, Zhang L. Sparse representation classifier for micro-aneurysm detection and retinal blood vessel extraction. Information Sciences. 2012;200:78–90.
10. Giancardo L, Meriaudeau F, Karnowski TP, Li Y, Tobin KW, Chaum E. Microaneurysm detection with radon transform-based classification on retina images. In International Conference of the IEEE Engineering in Medicine and Biology Society. 2011;5939–5942.
11. Niemeijer M, Van Ginneken B, Cree MJ, et al. Retinopathy online challenge: Automatic detection of micro aneurysms in digital color fundus photographs. IEEE Trans. Med. Imag. 2010;29(1):185–195.
12. Decenciere E, Cazuguel G, Zhang X, et al. Tele Ophta: Machine learning and image processing methods for teleophthalmology. IRBM. 2013;34(2): 196–203.
13. Walter T, Massin P, Erginay A, et al. Automatic detection of micro-aneurysms in color fundus images. Med. Image Anal. 2007;11(6):555–566.
14. Su Wang, Hongying Lilian Tang, Lutfiah Ismail Al turk, Yin Hu, Saeid Sanei, George Michael Saleh, Tunde Peto. Localizing microaneurysms in fundus images through singular spectrum analysis. IEEE Trans. Biomed. Eng. 2017;64(5):990–1002.
15. Yuji Hatanaka, Mitsuhiro Miyashita, Chisako Muramatsu. Automatic Micro-aneurysms Detection on Retinal Images Using Deep Convolution Neural Network. IEEE; 2018.
16. Kumar S, Kumar B. Diabetic retinopathy detection by extracting area and number of microaneurysm from colour fundus image. 5th International Conference on Signal

Processing and Integrated Networks (SPIN); 2018.
17. Jiawei Xu, Xiaoqin Zhang, Huiling Chen, Jing Li, Jin Zhang, Ling Shao, Gang Wang. Automatic analysis of micro-aneurysms turnover to diagnose the progression of diabetic retino-pathy, IEEE. 2018;6.
18. Ling Dai, Ruogu Fang, Huating Li, Xuhong Hou, Bin Sheng, Qiang Wu, Weiping Jia. Clinical report guided retinal microaneurysm detection with multi-sieving deep learning. IEEE Transactions on Medical Imaging. 2018;37(5).