

## Diagnosis of Diabetic Retinopathy using Machine Learning

Swati Gupta\* and Karandikar AM

Department of Computer Science and Engineering, Shri Ramdeobaba College Of Engineering and Management, Nagpur, India

### Abstract

Diabetic retinopathy is the most common diabetic eye disease and a leading cause of blindness. Regular screening for early disease detection has been a highly labor and resource intensive task. Hence automatic detection of these diseases through computational techniques would be a great remedy. There are many features present in retina like exudates and micro aneurysm feature. The presence of micro aneurysms (MAs) is usually an early sign of diabetic retinopathy and their automatic detection from color retinal images somewhat tough job, so for that we are using green Chanel images. The objective of project is to detect retinal micro-aneurysms and exudates for automatic screening of DR using classifier. To develop an automated DR screening system detection of dark lesions and bright lesions in digital funds photographs is needed. To detect retinal micro- aneurysms and exudates retinal funds images are taken from Messidor dataset. After pre-processing, morphological operations are performed to find the feature and then features are get extracted such as GLCM and Splat for classification. In this we are achieve the sensitivity and specificity of 87% and 100% respectively with accuracy of 86%.

**Keywords:** Exudates; Micro aneurysm; Green channel; Bright lesions; Red lesion

### Introduction

Diabetic retinopathy (DR) is a serious eye disease that occurs due to diabetes mellitus and it has grown as the most common cause of blindness in the present world. Based on latest reports by 2030 there is an epidemic rise of 4.4% in the global prevalence of diabetes [1]. Patient's sight can be affected by diabetes which causes cataracts, glaucoma, and most importantly, damage to blood vessels inside the eye, a condition known as "diabetic retinopathy". Effective treatments for DR are available though it requires early diagnosis and the continuous monitoring of diabetic patients. Diagnosis of DR is performed by the evaluation of retinal (fundus) images. Manual grading of these images to determine the severity of DR is rather slow and resource demanding [2]. It occurs when diabetes damages the tiny blood vessels inside the retina, the light sensitive tissue at the back of the eye. This tiny blood vessel will leak blood and fluid on the retina forms features such as micro-aneurysms, hemorrhages, hard exudates, cotton wool spots or venous loops [1].

Diabetic retinopathy can be broadly classified as non-proliferative diabetic retinopathy (NPDR) (Figure 1b) and proliferative diabetic retinopathy (PDR). Depending on the presence of features on the retina, which is said above, the stages of DR can be identified. A normal retina of the eye does not have any of the above said features (Table 1) [3-7]. In the NPDR stage, the disease can advance from mild, moderate to severe stage with various levels of features said above except less growth of new blood vessels. PDR is the advanced stage where the fluids sent by the retina for nourishment trigger the growth of new blood vessels. They grow along the retina and over the surface of the clear, vitreous gel

that fills the inside of the eye. If they leak blood, severe vision loss and even blindness can result [1].

Therefore regular screening of diabetic patients' retina is very essential and automated or computer-assisted analysis of diabetic patients' retina can help eye care specialist to screen larger populations of patients. Since 1982, the quantification of diabetic retinopathy and detection of features such as exudates and blood vessels (Figure 2) on fundus images were studied [8-12]. A computer based algorithms to detect the individual features of the fundus image were developed. In all these work, image processing techniques such as image preprocessing then segmentation after that 2D matched filters and image thresholding techniques were widely used for more clarify the image. An automatic system for the detection of symptoms pertaining to the abnormality is proposed [13-16]. Global and local thresholding values were used to segment exudates bright lesions from the red-free images. All these methods cannot be used to develop the features for automatic detection system as it is difficult to set a constant threshold value. The system proposed can detect and quantify only micro-aneurysms which can be useful in detecting only mild stages. Methods proposed in this paper use morphological image processing techniques. This method also cannot work for a constant threshold value when converted binary image. A prototype on automated diagnosis and understanding of retinal images was presented. The methods proposed in all the researches discussed

1	MA<5	Normal
2	5<MA<15	Mild
3	MA>15	Severe

**Table 1:** Show the diabetic retinopathy severity or grade as normal, mild and severe.



**Figure 1:** The fundus images of a) normal b) NPDR c) PDR.

\*Corresponding author: Swati Gupta, Department of Computer Science and Engineering, Shri Ramdeobaba College Of Engineering and Management, Nagpur, India, Tel: +91-8888602003; E-mail: [sswatigupta19@gmail.com](mailto:sswatigupta19@gmail.com)

Received June 07, 2015; Accepted July 24, 2015; Published August 31, 2015

**Citation:** Gupta S, Karandikar AM (2015) Diagnosis of Diabetic Retinopathy using Machine Learning. J Res Development 3: 127. doi:10.4172/2311-3278.1000127

**Copyright:** © 2015 Gupta S, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

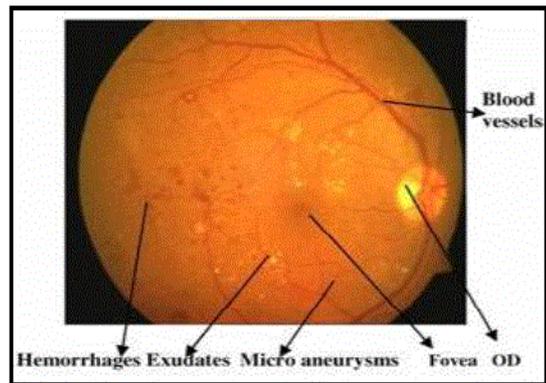


Figure 2: Diabetes Features In Defected Human Eye.

blood vessels that nourish the retina are blocked and this stage is called moderate NPDR. The next stage is severe NPDR during which many more blood vessels are blocked.

## Materials and Methods

In this paper, the evaluation of the automated diagnosis system of diabetic retinopathy has been performed by using a set of 94 images which were captured by retinal fundus camera and the captured images are stored in a JPG image format files with the size of 1500 × 1152 pixels at 24 bits pixel length.

In given proposed methodology in Figure 2 shows that image is get input from retinal fundus data set then pre-processing techniques are applied on retinal image. After that morphological operations are performed to identify exudates and micro-aneurysms features. Finally using multiclass SVM and KNN classifier for giving severity or grade of abnormality. For given methodology the input images are taken from MESSIDOR, Diabeticret DB1.

## Preprocessing

The three-stage algorithm to automatically detect and grade the severity of DR using ophthalmic fundus images is shown in Figure 2. In preprocessing stage, the image (Figure 4) is get rectified some problems such as blurring, non-clarity and size of image. In this stage resizing of the image is done and then color space conversion problem, image restoration and finally enhance the image. In color space conversion the input color fundus image (Figure 4) is converted into HSI (Figure 5) i.e. hue saturation and intensity model. In HSI the intensity component, color model space is decoupled from color images. The preprocessing module proceeds by histogram equalization and contrast enhancement and followed by scaling all pixel intensities in the range (0, 1) (Figures 6 and 7).

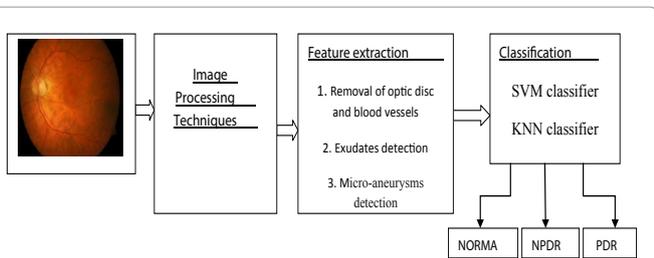


Figure 3: Proposed systems for detection and classification of different stages of Diabetic Retinopathy.

above are not reliable and robust as it does not provide any objective measurement on the features. As all the above discussed methods are mainly useful in analysis of the specific features on the retina.

We are proposing a system for automated classification of three types first one is normal then second one is NPDR (non proliferative diabetic retinopathy) and third is PDR (proliferative diabetic retinopathy) retinal images by automatically detecting the blood vessels hard exudates and GLCM features. The proposed system is shown in block diagram which is shown in Figure 3. The target of it, to measurement such as blood vessels and remove it, calculates exudates area and contrast is computed from the processed retinal images. These objective measurements are finally fed to an SVM classifier. Materials and methods section explains the methods in detail and results are tabulated and explained in Results section. Discussion and conclusions are presented in Discussion and Conclusions sections respectively.

## Diabetic retinopathy

According the presence and extent of the features such as hard exudates, cotton wools spots micro aneurysms or hemorrhages due to leakage of fluid and blood from the blood vessels which are shown in Figure 1. DR can be classified into normal, NPDR and PDR. Diabetic retinopathy is of two type's namely non proliferative i.e., NRDR and proliferative type i.e., PDR. Non proliferative is the early stage of the disease characterized by the presence of micro aneurysms. As the disease progresses the retina is deprived of oxygen and new blood vessels are formed leads to clouding vision. There are also three categories that are mild NPDR, moderate NPDR and severe NPDR.

In mild NPDR, micro aneurysms are small areas of balloon-like swellings in the retina's tiny blood vessels. As the disease growths, some



Figure 4: Original image.

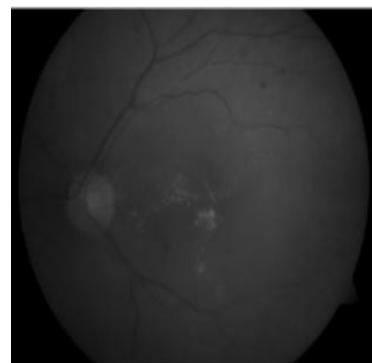


Figure 5: HIS model image.

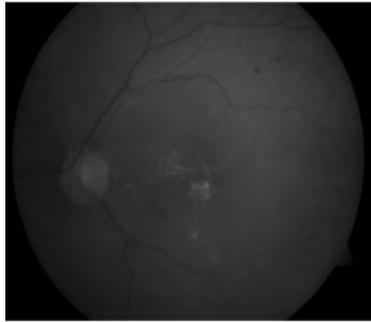


Figure 6: Pre-processed Image.

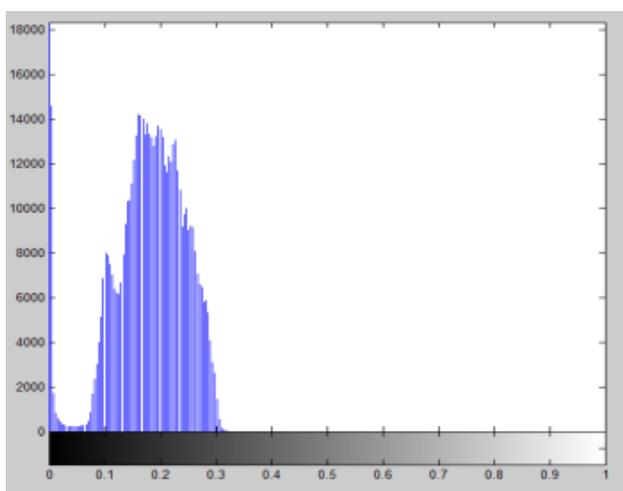


Figure 7: Histogram of pre-processed Image.

Next, is median filtered to obtain the enhanced images. These three detection stages are applied for DR severity detection per fundus image. In this the transformation equations used in the conversion of RGB to HSI are

$$H' = \begin{cases} \text{undefine} & \text{if } c = 0 \\ \frac{g-b}{c} \text{ mod } 6 & \text{if } M = r \\ \frac{b-r}{c} + 2 & \text{if } M = g \\ \frac{r-g}{c} + 4 & \text{if } M = b \end{cases} \quad (1)$$

$$h = 60^\circ * H' \quad (2)$$

Where r=red, g=green, b=blue and the saturation of component (S) is given by formula

$$S = 1 - 3 / [(r + g + b) * \min(r, g, b)] \quad (3)$$

Finally intensity (I) component is given by:

$$I = 1 / [3 * (r + g + b)] \quad (4)$$

The converted images are then filtered using hybrid median filter. This filter is used to remove noise like pepper and salt occur during image acquisition. Specialty of hybrid median filter is to provide better edge corner prevention by smoothing the quality of image and reduced

the noise appears due to thickness and thinness of boundaries of features. After filtering we are using CLACHE technique, which means the contrast limited adaptive histogram equalization for contrast enhancement and it improves the quality of images. The Figures 5 and 6 show the histogram equalization and image after preprocessing.

### Features extraction

In candidate extraction like Morphological operations such as erosion, dilation, closing, and opening are performed for finding micro-aneurysms and exudates features. Then invert image method is applied to for inverting the image. Then holes are filled in the image.

**Optical disc elimination:** The optical disc is brightest part of the normal eye in fundus images and approximately it is oval or elliptical in shape. In colored fundus images the OD appears as a bright yellowish or white region in fundus images. Exudates have high and similar intensity values of optic disc. So it is necessary to eliminate the optic disc from the retinal image. This brighter optic disc should be masked and removed using region properties and area finding. After preprocessing edge detection algorithm is applied for detection of optical disc and blood vessels. Canny edge detection is used for counter detection [3]. Canny edge detection algorithm enhances blurred edges by preserving all local maxima the gradient, through this it detects optimally the boundaries of features.

The logical black and white function is used to create and then invert the image to create mask image as shown in Figure 6. In this the mask image is created and then mask image is subtracted from the edge detected image.

**Blood vessels extraction and removal:** For the detection of micro-aneurysms and exudates remove blood vessel and optical disc is must from retinal image because it have similar concentration level like micro-aneurysms and exudates respectively. Dilation operation on the intensity image helps to remove high level of contrasts vessels of blood. After that we are using the structuring element to fill the small holes in images by using dilation operation. Structure elements (SE) have numbers of shapes like diamond, disc, round etc., but in our project we are using flat disc shaped structure which is used to remove the optical disc and blood vessels. If SE start from bright pixel then there will be no change and goes to the next pixel. On otherwise SE start from black pixels then SE covered whole image through dilation operation which adds pixels to pixels and create boundaries of objects in fundus image. Whereas erosion operation is used to remove pixels on object boundaries. Then the dilated image obeys the following equation.

$$A \oplus B = \{Z | (B)Z \cap A \neq \emptyset\} \quad (5)$$

Disk Shaped SE (B) on A, Erosion operator is used to remove completely blood vessels from images (Figure 8) without affecting other portions which is calculated by following formula:

$$A \odot B = \{Z | (B)_Z \leq A\} \quad (6)$$

**Detection of exudates and micro-aneurysms:** After removing blood vessels and optical disc from image, detect the exudates features. Exudates are a bright lesion of retina image (Figure 9). This feature is detected by using morphological closing operation and this closing operation applied on eroded fundus image. For the detection of exudates features, closing dilation is followed by erosion operator are used.

$$P \cdot Q = (P \oplus Q) \odot Q \quad (7)$$

Micro-aneurysms are other features of DR. For micro-aneurysms

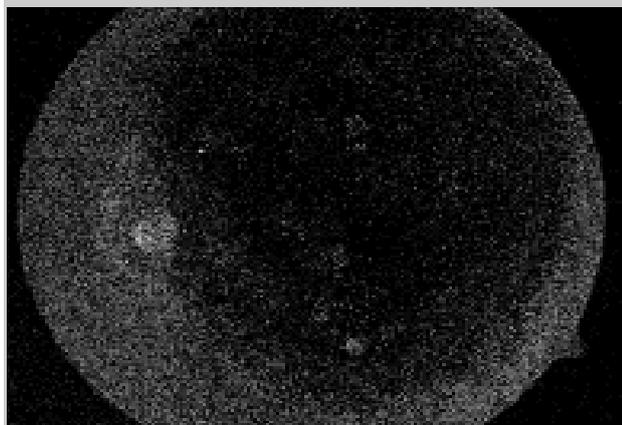


Figure 8: Blood Vessels Removed Image.

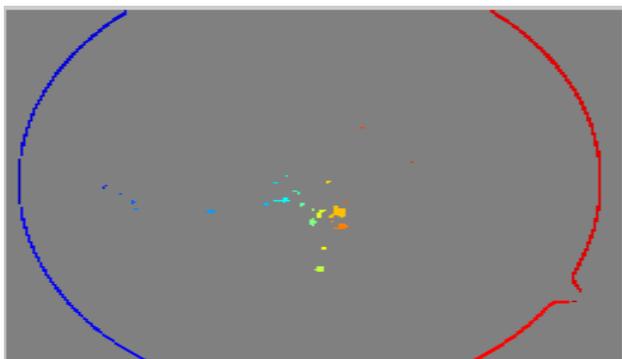


Figure 9: Exudates Features.



Figure 10: Binary image.

detection opening morphological operation is performed in which erosion is followed by dilation whereas the micro-aneurysms are appearing as red spot which get swell in the retina. Invert image method is used and then morphological opening operation is performed for detection of micro-aneurysms. So, we can easily count the micro-aneurysms values (Figure 10).

**Classification:** After detecting exudates and micro-aneurysms

(Figure 11) in color image the features get extracted from fundus image. All these features are calculated and applied to SVM and KNN classifier. SVM classifier gives better results than KNN classifier. In this paper we analyze a set of 17 features that consists of most of the features from GLCM (gray level co-occurrence matrix) (Figure 12) features and along with additional structural features that are obtained using the "regionprops" command in MATLAB. The features that we are using like energy, contrast, entropy, homogeneity, area, major axis, major axis length, convex image, etc. In this study we selected the 11 features out of 17 features that were calculating the probably of occurrence generated by ANOVA (analysis of variance between group) test. The ANOVA test uses variances to decide whether the mean are different.

These features are selected for reducing noise and enhancing the result of classifier accuracy. The result is shown in Figure 12. If the MA >15 then it treated as severe and if total  $5 < MA \leq 15$  so it treated as moderate. If the value is less than 5 then it is normal image (Figure 13).

**SVM classifier:** The SVM approach basically used for binary classification problems. In this we are using multi-class pattern recognition problem. Basically there are two types to solve multi-class problems by using binary SVM classifiers: The first one is "one-against-one" in this method for each possible pair of classification a binary classifier is calculated. Other approach is the "one-against-rest" in this method N different classifiers are constructed, one for each class. In this paper we built a two-class classifier over a feature vector  $\phi(x, y')$  derived from the pair consisting of the input features and the class of the data. At test time the classifier chooses the class (Figure 14).

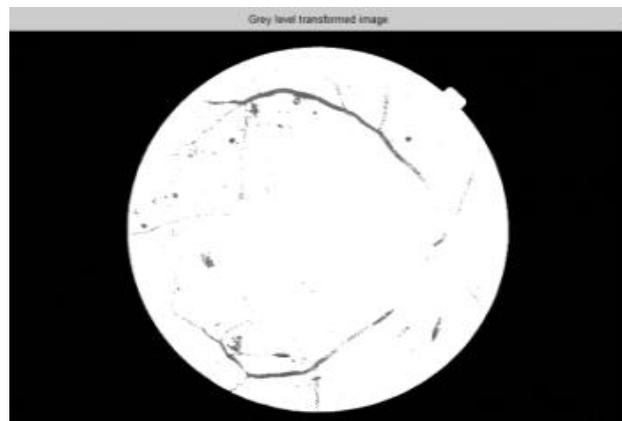


Figure 11: Gray level transformed image.

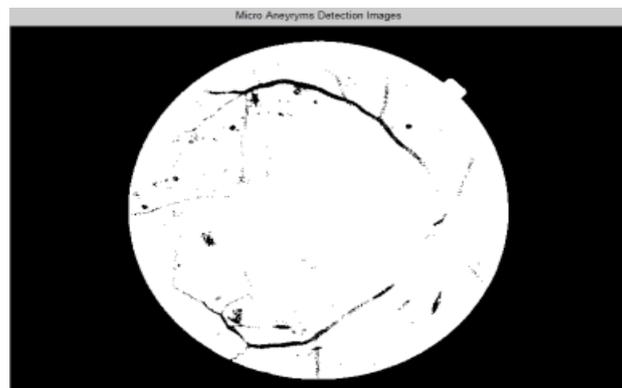


Figure 12: Micro-aneurysms Detection Image.

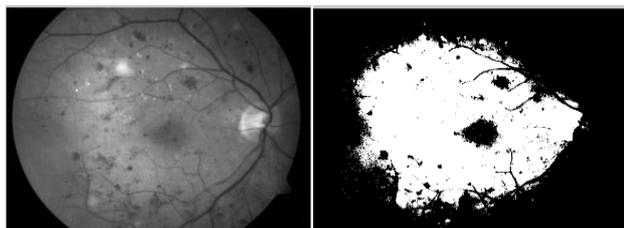


Figure 13: (a) original image (b) Severe NPDR.

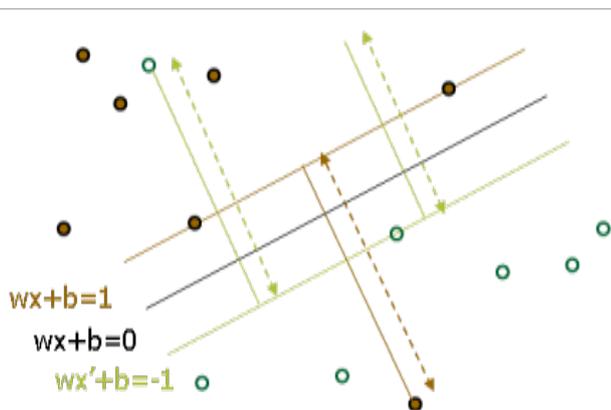


Figure 14: Hyper plane representation for multi-classification.



Figure 15: SVM Classifier: 70% Training ; 30% Testing.

$$y = \arg \max_y \bar{W}^T \varphi(\bar{X}, Y') \quad (8)$$

The margin during training is the gap between this value for the correct class and for the nearest other class. In the SVM world, such work comes under the label of structural SVMs (Figure 15).

**KNN classifier:** KNN classifier- The k-nearest-neighbor classifier is normally based on the Euclidean distance between a test sample and the specified training samples (Figure 16). The Euclidean distance between sample xi and xl(l=1,2,...,n) is defined as

$$d(x_i, x_l) = \sqrt{(x_{i1} - x_{l1})^2 + (x_{i2} - x_{l2})^2 + \dots + (x_{ip} - x_{lp})^2}$$

$$A R_i = \{x \in R_p : d(x, x_i) \leq d(x, x_m), \forall i \neq m\}$$

Voronoi cell encapsulates all neighboring points that are nearest to each sample and is defined as above. Where Ri is the Voronoi cell for sample xi and x represents all possible points within Voronoi cell Ri.

## Results

The final goal of our DR detection system is to classify the fundus images that are free from retinopathy lesions as normal, and to classify the abnormal images according to its severity as non PDR and PDR. Using the data obtained from the images their data are divided into two portions that is training and testing. In this we are using 70-30 ratio during the training and testing of classifiers. In our project we are doing comparative study with classifier SVM and KNN. SVM classifier gives better result than KNN which is shown in Tables 2 and 3.

Table 4 gives the information about training testing and accurately classified data and its percentage of accuracy is given in SVM classifier achieves 85.60% accuracy and whereas KNN classifier achieves 55.17% accuracy. Table 5 as shown is the breakdown of the images according to

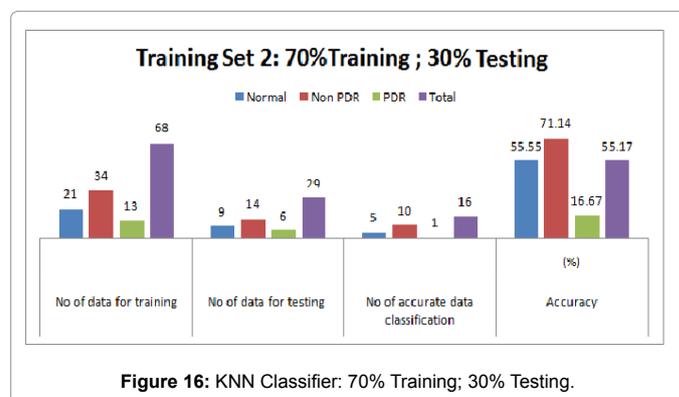


Figure 16: KNN Classifier: 70% Training ; 30% Testing.

Classes number	No. of data for training	Data for testing	Number of accurate data classification	Percentage accuracy
Normal	19	11	9	81.8181
Non PDR	32	8	7	87.5
PDR	22	8	7	87.5
Total	73	27	23	85.6060

Table 2: Results of support vector machine classifier.

Classes number	No. of data for training	Data for testing	Number of accurate data classification	Percentage accuracy
Normal	21	9	5	55.55
Non PDR	34	14	10	71.14
PDR	13	6	1	16.67
Total	68	29	16	55.17

Table 3: Results of K-nearest neighborhood.

True Predicated	Normal	Non PDR	PDR
Normal	9	0	0
Non PDR	0	1	4
PDR	0	5	1

Table 4: Breakdown results of support vector machine classifier.

True Predicated	Normal	Non PDR	PDR
Normal	9	0	0
Non PDR	0	1	4
PDR	0	5	1

Table 5: Breakdown results of K nearest neighbor.

the actual class and the classes predicted by SVM and KNN for training set.

Here from the calculation of Sensitivity and Specificity for the SVM and K-NN classifier are good for SVM sensitivity of 91 percent and specificity of 100 percent and a classification accuracy of 86.67 percent.

For KNN Classifier sensitivity of only 32.80 percent and specificity of 100 percent and a classification accuracy of 55 percent.

## Conclusion

The Graphic User Interface has been successfully developed. It is capable to input a test image to be processed and extract texture features of the image and display these values and the results of the classification of both SVM and KNN classifiers. This project, a reliable fundus image analysis system is developed to give an ophthalmologist the most comprehensive view of the retina state to diagnose diabetic retinopathy. This system extracts the fundal features such as retinal exudates and micro-aneurysms even for a low quality color fundus image. the fundus image analysis system detects and grades the severity of diabetic retinopathy. Hence this system can also function as an automatic tool for the mass screening of diabetic retinopathy. A morphological operation is done to detect dark lesions in digital fundus images. This method combines morphological based dark lesion detection with inverting and rage of pixels. To evaluate the performance of the proposed the fundus image analysis system detected diabetic retinopathy with a high sensitivity and specificity of 87% and 100% respectively with accuracy of 86% using SVM classifier.

## Acknowledgement

Authors would like to thank them for providing the data to do research on the area of detection of exudates, micro-aneurysms and for Diabetic Retinopathy.

## References

1. Gandhi M, Dhanasekaran R (2013) Diagnosis of Diabetic Retinopathy Using Morphological Process and SVM Classifier, IEEE International conference on Communication and Signal Processing, India pp: 873-877.
2. Li T, Meindert N, Reinhardt JM, Garvin MK, Abramoff MD (2013) Splat Feature Classification with Application to Retinal Hemorrhage Detection in Fundus Images, IEEE Transactions on Medical Imaging 32: 364-375.
3. Ramani RG, Geetha R, Lakshmi B, Jacob SG (2012) Data Mining Method of Evaluating Classifier Prediction Accuracy in Retinal Data. IEEE International Conference on Computational Intelligence and Computing Research pp: 1-4.
4. Rocha A, Carvalho T, Jelinek HF, Goldenstein S, Wainer J (2012) Points of Interest and Visual Dictionaries for Automatic Retinal Lesion Detection. IEEE Transactions on Biomedical Engineering 59: 2244-2253.
5. Giancardo L, Meriaudeau F, Karnowski T, Tobin K, Grisan E, et al. (2011) Textureless macula swelling detection with multiple retinal fundus images. IEEE Trans Biomed Eng 58: 795-799.
6. Giancardo L, Meriaudeau F, Karnowski TP, Li Y, Garg S, et al. (2012) Exudate-based diabetic macular edema detection in fundus images using publicly available dataset. Medical Image Analysis 16: 216-226.
7. Agurto C, Murray V, Barriga E, Murillo S, Pattichis M, et al. (2010) Multiscale Am-Fm methods for diabetic retinopathy lesion detection. Medical Imaging, IEEE Transactions 29: 502-512.
8. Kumar A, Gaur AK, Srivastava M (2012) A Segment based Technique for detecting Exudate from Retinal Fundus image. Procedia Technology 6: 1-9.
9. Sopharak A, Dailey MN, Uyyanonvara B, Barman S, Williamson T, et al. (2008) Machine learning approach to automatic Exudates detection in retinal images from diabetics patients. Journal of Modern Optics 57: 124-135.
10. Yazid H, Arof H, Isa HM (2012) Exudates segmentation using inverse surface adaptive thresholding. Measurement 45: 1599-1608.
11. Osare A, Shadgar B, Markham R (2009) A Computational Intelligence-Based Approach for Detection of Exudates in Diabetic Retinopathy Images. IEEE Transactions on Information Technology in Biomedicine 13: 535-545.
12. Kumari CJ, Maruthi R (2012) Detection of Hard Exudates in Color Fundus Images of the Human Retina. Procedia Engineering 30: 297-302.
13. Saleh MD, Eswaran C (2012) An automated decision-support system for non-proliferative diabetic retinopathy disease based on MAs and Has Detection. Computer Methods and Programs in Biomedicine 108: 186-196.
14. Lazar I, Hajdu A (2013) Retinal Microaneurysm Detection Through Local Rotating Cross-Section Profile Analysis. IEEE Transactions On Medical Imaging 32: 400-407.
15. Antal B, Hajdu A (2012) An Ensemble Based System for Microaneurysm Detection and Diabetic Retinopathy Grading. IEEE Transactions On Biomedical Engineering 59: 1720-1726.
16. Deepak KS, Sivaswamy J (2012) Automatic Assessment of Macular Edema from Colour Retinal Images, Medical Imaging. IEEE Transactions 31: 766-776.