IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.8, Issue.6, June 2020 E-ISSN: 2347-2693

# Proliferative Diabetic Retinopathy Detection Using Machine Learning

**Neha Tamboli[1*], G.S. Malande[2]**

[1,2]Dept. of Computer Science and Engineering, Shri Guru Gobind Singhji Institute of Engineering and Technology, Vishnupuri, Nanded, India

**Corresponding Author:2018mns010@sggs.ac.in*

DOI: https://doi.org/10.26438/ijcse/v8i6.106111 | Available online at: www.ijcseonline.org

Received: 10/June/2020, Accepted: 20/June/2020, Published: 30/June/2020

***Abstract***— In this paper, the method for detection of neovascularization from fundus retinal image is presented. Neovascularization is the type of proliferative diabetic retinopathy and it is characterized by new, fragile retina vessels. It poses high risk for sudden vision loss. To avoid this risky situation, an early detection, proper treatment and diagnosis is essential. Therefore, we cannot underestimate the significance of accurate and timely detection of NV. We propose a method to detect NV which is based on automatic image processing that involves vessel segmentation using K-means, Vessel morphology, texture based features extraction and classification of images with support vector machine(SVM) and we achieved an average accuracy of 99 % on the selected test set.

***Keywords***— Feature Extraction, K-means clustering, morphological image processing, Neovascularization, Support vector machine**.**

## I. INTRODUCTION

Diabetes is a disease and organ condition that develops when the pancreas is no longer able to secrete enough insulin or pancreas cannot work properly. Over time, the entire cyclic system comprising the retina is affected by diabetes. Diabetic Retinopathy (DR) is a medical condition that causes retina damage. Diabetic Retinopathy (DR) is an eye focused abnormality caused by long term diabetes and is the most common reason for blindness before age of 50 years. It is the most significant cause of blindness and the more severe diabetic eye condition. Close to 415 million patients with diabetes are prone to blindness due to diabetes. The problem of diabetic retinopathy that usually occurs due to high blood sugar levels in diabetic patients. It occurred due to damage of light touchy tissue behind the eye, which is also called retina. The blood vessels may leak, or new blood vessels sometimes tend to grow on the retina.

The major parts of the eye are the iris, cornea, retina, sclera, nerve fibres, optic nerve, etc. as shown in the Figure 1.

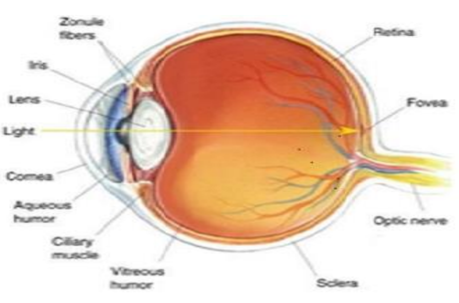


Figure 1. Human Eye Structure

DR's effect is set to triple from 2005- 2050 by the centre of disease control (CDC), the US agency. Different factors leading to DR include blood glucose, diabetes, blood pressure, etc.

Diabetic Retinopathy is based on the Greek word (die-uh-BET-ikert-ih-nOP-uh-thee). because, it is sensitive to light it directly affects the human eyes and can cause damage to blood venous. DR is mainly divides into two kinds of diabetic retinopathy, which is non-proliferative diabetic retinopathy(NPDR), it is early DR stage, with mild symptoms, or sometimes non-existent symptoms. Microaneurysm, which leaks fluid and blood into the retina, are the example of NPDR. It is the earliest clinical sign of DR. Another type is the more advanced stage of the disease is Proliferative Diabetic Retinopathy(PDR). Which is also known as neovascularization. The formation of new, fragile, abnormal and thin blood vessels due to excessive oxygen lacks in the retina. It sometimes leaks and bleeds blood into the retina region because new vessels are weak and fragile, leading to vision loss and possibly blindness.

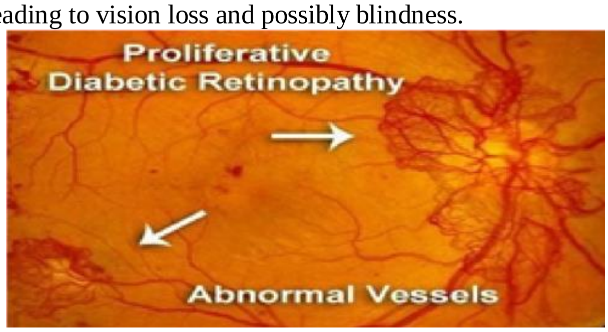


Figure 2. A human retina which is affected with diabetes-2

Figure 2 shows a neovascularization which is affected due to diabetes- 2. It shows a network of fragile, thin, abnormal new blood vessels that usually grow on the surface of the optic nerve or on the optic disc and in other areas of the retina. Proliferative diabetic Retinopathy may also lead to neo-vascular glaucoma, where the iris (the coloured part of the eye) is obstructed due to new abnormal blood vessels overloading and increasing eye pressure of the drainage channels. This high pressure can damage the optic nerve quickly leading to permanent blindness. Because of its severity, it is important to detect neovascularization at early stage.

## II. LITERATURE REVIEW

The retinal image analysis, based on computing, was first implemented in 1974. Today, retinal disease such as diabetic retinal disease (DR) are becoming a mainstream technique. In the fundus images there are several key features, like the retinal blood vessels, the optical disk and fovea. There are different kinds of eye diseases. DR is one of the deadliest disease of the eyes. By the WHO (World Health Organization) –united nations survey this disease is the second largest disease among humans[1].

Automated analysis system is suitable for processing digital images. Image analysis for fundus is a complex task because of the variability of colour or gray background images, the morphology and presence of certain features in various patients which can lead to a misinterpretation of the anatomical structures of the retina. For this, image processing is now become practical tool to process fundus image. Many examples of the use of digital imaging techniques used to identify DR can be found in this literature.

Neovascularization may be observed in ocular, corneal and retinal regions of the eye. Of these, the most serious PDR-related abnormality is retinal neovascularization. In spite of that, research in the field of exudates and microaneurysms and hemorrhages is not as extensive.

An evaluation method for the diagnosis of automatic retinopathy are the key input of this work. The basic truth and evaluation protocol is proposed as part of a prototype database. There are reporting experimental for a literature method. This study offer the means for the reliable assessment of automatic retinopathy. The idea for this approach comes from the strict rules for evaluating the biometric authentication method, for example the FERET protocol for methods of face recognition.

Goatman et al proposed first work in this area, using vessel based features such as shape, position, orientation and line density etc. After that normal and abnormal segments of the vessels are differentiated by the support vector machine [2]. This method has been tested on 38NVD images and 71 non NVD images and reaches a value of 0.911 area under curve (AUC).

In order to derive two distinct feature sets from binary vessel maps. Welikala et al suggested a dual classification system from the regular line operator and modified line operator [3],[4].The final result is entirely based on two SVM classification because vessel characteristics have been extracted from each two vessel map.

Hassan et al. also observed NV as features in the vessel and vessel area number scan window[5]. Average 89.4 percent specificity and a sensitivity of 63.9 were achieved on 11 images.

Agurto et al. developed an NVD detection system using features that have been extracted by AMFM, fractal analysis, granulometry on scale and amplitude scale [6]. The AUC value of 0.93% was achieved with the proposed feature set based on a data set that includes 100 NVD and 200 Standard Images.

The statistical structural analysis of the neovascularization, spectrum analysis and fractal analysis have been extracted by Lee et al. It reached 98.5 percent precision. Tested on 27 NV images[7].

The short survey of the framework for neovascularization detection reported by eswari and rajeswari was carried out by saranya et al. [8],[9], kasurde et al and RamaSubramanian et al. [10],[11].

Gupta et al. also proposed that a random forest classification is used to train labelled patches consuming to create a marked earth truth for retina vessels with texture features and vessel characteristics taken from the local patches [12].

Shuang Yu, proposed a method for automatic neovascularization detection on optic disc region using multi-level gabor filtering and achieves an accuracy of 95.23 % tested with support vector machine[13].

## III. MATERIALS AND METHODOLOGY

Figure 3 shows an overall framework of proposed methodology.

### A. Image Acquisition

Fundus image plays an important role in diabetes monitoring as retinal disorders are frequent and severe. Yet, because the eye fundus tends to be prone to vascular disorders, the fundus images is

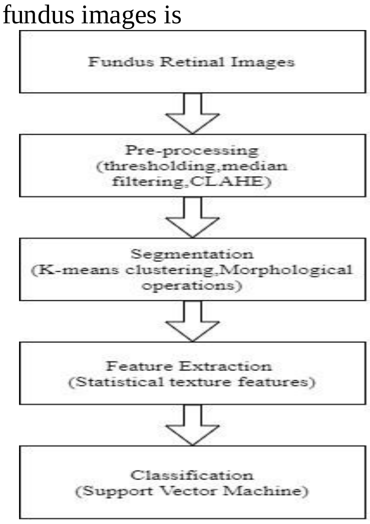


Figure 3. Framework for NV Detection Procedure

 

Considered an aggressive diabetes screening candidate. The success rate of screening is depending on precise fundus image capturing, especially on precise and reliable image processing algorithms for abnormalities detection.
Several abnormalities of fundus images were collected from two online databases. Below is image dataset from each source.

- IMAGERET-diaretdb0(2006) and diretdb1(2007): Diretdb0 consisted of 130 colour fundus images of which 20 were normal and 110 showed DR signs in the form of hard exudates, soft exudates , microaneurysms and haemorrhoids, and 20 were found to exhibit neovascularization[15].
- The diretdb1 database consisted of 89 colour fundus images, of which 84 shows indications of mild NPDR i.e. microaneurysms, 5 neovascularization images and 5 normal samples were found. The images were captured at 1500 * 1152 pixels in PNG format using 8 bits per color plane.

**B. Pre-processing**

The blood vessels of neovascularization are in twisted and thin nature. The intense pattern of development and irregular lighting leads to extremely difficult for abnormal blood vessels separation from background of the fundus. The proposed technique is used for detecting these blood vessels.

Method of illumination correction is used to remove uneven illumination (it is the amount of light source incident on the scene). The illumination correction procedure is used to remove uneven brightness. Because of the non-uniformities in the scene illumination, fundamental image processing operations such as thresholding are often hard to use. In Image processing, thresholding is a process where the images are divided into the foreground and background. It is one of the technique used for image segmentation, which is simple and effective. However, these illumination variations can be reduced or eliminated entirely if we use background subtraction and filtering prior to level-sensitive processing. It is based on background subtraction where scene is composed of homogeneous background and contains small objects darker than the background. Figure 4 shows an original fundus image sample with an uneven intensity values. Image filtering is used to get general intensity background image. Image filtering is applicable in many applications, which include edge detection, noise removing, sharpening and smoothing, image enhancement. Filters are building block of image processing methods.

The filter is represented by a kernel, a small array applied to and around each pixel of image. kernel is a small convolution matrix or a small matrix for blur, sharpen and edge detection in the image processing [14]. The application of filters to an image is like a convolution procedure, which is a linear filter type, represented by the multiplication of the matrix. While thresholding and image equalization are type of a non-linear filter as median filter. In the propose system median filter or morphological filter is used to obtain general intensity background image.

Median filter is a nonlinear digital filter, normally used to remove noise from the image or signal at the edges. This noise removal was use to improve the pre-processing results. Usually, it is used to remove “salt and pepper noise”. The noise of pepper and salt is also referred to as impulsive noise. There are sharp and sudden image signal disturbances which occur in the image as white and black pixels sparingly occurred. The value of salt noise in terms of frequencies is high ( 255......200) and value of pepper noise is low ( 5.....0). it is occurred due to transferring of data.

Median filter works, by pixel by pixel, by replacing each value with the average value of the next pixel. The pattern of the neighbour is called a window that slides the pixel in the image. This determines the centre by grouping all the pixels in numerical order from the window and then replaces the broadcast reference value for the pixel. The light is corrected by extracting the background image from the original image. Image subtraction is a process in which the numerical value of one pixel or entire pixel image is subtracted, which produce third image as output.

Contrast limited adaptive histogram equalization(CLAHE) is applied on the filtered component of the image [17]. Essentially, it is used to enhance image contrast. At this point, it optimizes the contrast enhancement by dividing and conquering local image data and thus effectively addresses global noise. In other words, the basic idea behind the algorithm is to divide the image into a number of small, non-overlapping contextual regions, called “Tiles”. Figure 5 shows pre-processed output after applying pre-processing techniques.

**C. Vessel segmentation**

In the medical field of disorder detection segmentation plays an important role. Segmentation is the process of automatic or semi-automatic detection of boundaries within an image. It divides an image into areas on a specified description, such as segmenting body organ/tissues for edge detection, disease detection, in the medical application. As the presence of abnormal thin vessels is characteristic of neovascularization, accurate retinal vessel segmentation is also very important for neovascularisation detections. New veins are less visible and smaller in size than normal blood veins. In this paper we have utilized morphological and k-means clustering based image segmentation techniques.

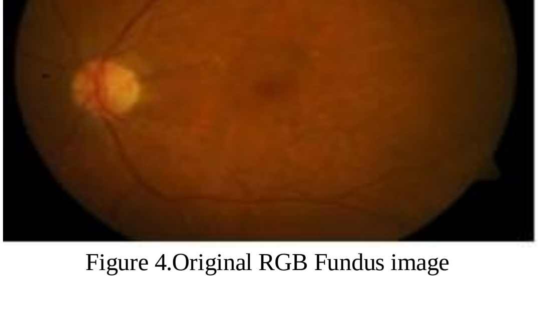

Figure 4.Original RGB Fundus image

 

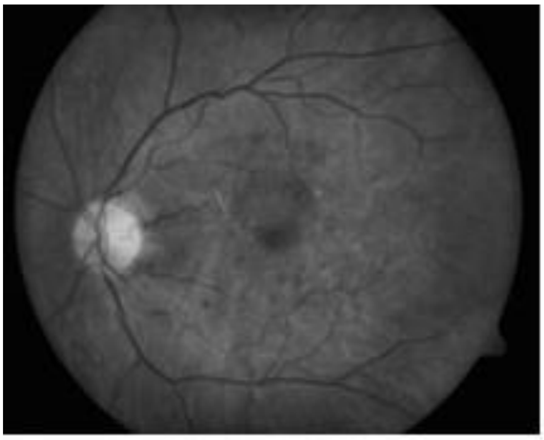

Figure 5. Pre-processed image

K -means clustering algorithm is an unsupervised algorithm and it is used to segment the area of interest from the background [19]. Thus, subtractive cluster is used to generate the initial centers and these centers are used in k-means algorithm for the segmentation of image. for segmentation purpose the algorithm extract texture features of an image and regions with similar texture features are grouped together to obtain the shape of the object. Figure 6 shows a candidate image.

To extract blood vessels from the images, the morphological operations are used. A series of nonlinear processes related to the shape or morphology of values is a morphological element in the processing of images rather than their numerical values. Morphological techniques are used to test a small form or template image called as structuring element. In comparison with the pixel neighbourhood, the structuring element is placed in every position in the image. Some operations check whether the elements fit into the neighbourhood, and others check whether the element intersects or hits the neighbourhood.

A binary image morphology produces a new binary image with a non-zero pixel value. A small binary image, a small matrix of pixels with each zero or one value is the structuring element. The dimensions of the matrix determine the dimension of the structured element, the patterns of the structuring element, and the form of nulls. It is common practice that the structural matrix is defined in strange dimensions and its origin is the centre of the matrix. Throughout the linear image filtering process, structural elements play the same function as convolution kernels in morphological image processing [18]. Each of its pixels is associated with the corresponding neighbourhood pixel under the structural elements. The structuring element fits the image when the corresponding pixel for each of its pixels is 1 as well. The structuring element is said to hit or intersect an image if one of its pixels corresponding to image pixel is 1. The structuring element is ignored with zero-rated pixels; this means, indicating points where the image value for this is irrelevant. The morphological operation used for vessel extraction is explained below. Figure 7 shows an extracted blood vessels of retinal fundus image.

**1) Opening**: In mathematical morphology, the expansion of erosion of set A through structural element B is the opening. Opening takes little objects from the front of image (usually taken as bright pixels) puts them in the background and takes away small objects in the foreground while closing , turning the small background islands into the foreground. This technique can also be used in an image to find certain shapes. Opening can be used for determining what a particular structuring element can fit into (edges, corners, etc.).

The opening of an image F by structuring element S is an erosion followed by a dilation. It is denoted by F ○ S.

$$F \bigcirc S = (F \ominus S) \oplus S \quad (1)$$

Opening is so called because, an object connected through a thin pixel bridge can open a gap. Any region surviving the erosion is restored by using dilation to its original size. Opening is an idempotent operation. Once an image is opened, there will be no further effect on that image for further openings with the same structuring element formulated as,

$$(F \bigcirc S) \bigcirc S = F \bigcirc S) \quad (2)$$

**2) Closing**: The closing procedure expands an image and erodes the extended image with the same structuring element for the two operations. Morphological closing helps to fill up small image hole while maintaining the shape and dimension of the objects in the image. The closure of a set (binary image) A by an element B of structure is an erosion of the dilation of that set in mathematical morphology. Opening removes objects and closing removes holes for small objects. A dilation followed by erosion is the closing of an image F with a structural elements S (referred by F • S ).

$$F \bullet S = (F \oplus Sort) \ominus Sort \quad (3)$$

It is named as closing because, gaps in the regions may be filled while preserving the original area size. Closing is also idempotent like opening (F • S) • S = F • S, and it is dual operation of opening formulated as,

$$F \bullet S = (F^c \bigcirc S)^c \quad (4)$$

$$F \bigcirc S = (F^c \bigcirc S)^c \quad (5)$$

A binary image is interpreted morphologically, by applying composite operations such as the opening and closing operations. They will act as in the forms of filters. The size of the structuring item is most critical for the removal of noisy data, but not for damage to items of interest. The kernel used for segmenting the blood vessels is in elliptical shaped kernel with size of 5 X 5 , 11 X11, and 23 X 23.

 

Figure 6. Extracted Blood Vessels

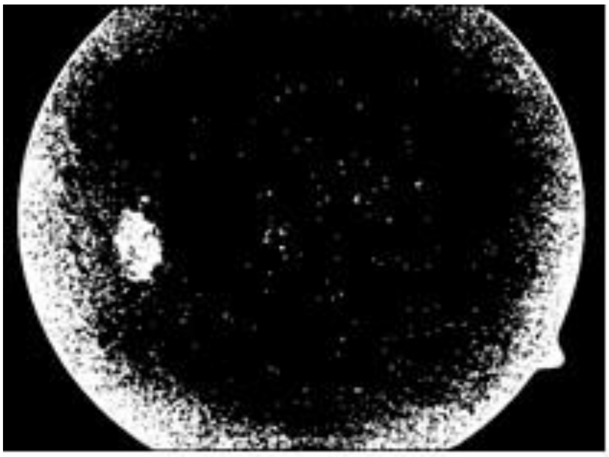

Figure 7. Candidate Image

**D. Feature Extraction**

Feature extraction is highly necessary for the classification of images. The Medical image dataset typically suffers from limited data space, caused by the high cost of equipment and personnel for image processing and for labelling.

As its consequences, the functional aspect usually constrained on implementation. Alternatively, the learning method will have a high variance and cannot be optimally trained to detect NVD, so here the extraction of features is essential. Feature extraction is also a way of simplifying the issue of a segmentation.

There are different properties of new and normal vessels of retinal image. The new vessels are thin, tortuous, twisted and reflect the statistical characteristics of the vessels. By defining the associated properties of the image, you can extract and condense the input type characteristics of the classifier into a specific dimension (n). The gray-scale image pixels are transformed into an X=[x1, x2 .. xn] where each Xi is unique and n is the dimensions of the vector where i=0,1,2,......,. As to detect vessel abnormalities texture properties is very important. The texture of an image region is defined by the distribution of gray-scale over the region of pixels. The following are the first order statistical features are extracted based on texture [21].

1) **Gradient Variation:** It is feature based on an abnormal vessel's greater contrast variability than normal vessels. It is measured using sobel gradient.
2) **Mean:** The average value of all matrix components, i.e. image pixels.

$$Mean = \mu = \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n} p(i,j) \quad (6)$$

Here N is the number of image pixels and i, j is row and column of image. m and n are final row and column values of image, p(i, j) is the image matrix.

3) **Variance:** It is a measurement of gray level deviation from mean of image.

$$\text{Variance}= \sum_{i=0}^{m}\sum_{j=0}^{n}(i-\mu)^2 \, p(i,j) \quad (7)$$

4) **Standard Deviation:** It is the square root of the variance of components in the matrix.

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{i=1}^{m}\sum_{j=1}^{n}[p(i,j)-\mu]^2} \quad (8)$$

5) **Entropy:** A retinal image entropy can be defined as a measure of the uncertainty associated with a random variable. There is high entropy in homogeneous areas.

$$Entropy = -\sum_{i=0}^{m}\sum_{j=0}^{n}[p(i,j]\log_2[p(i,j)] \quad (9)$$

**E. Classification**

Upon obtaining the feature vector, a support vector machine (SVM) is used to classify NV images and normal images optimally. SVM is supervised learning algorithm that aims to reach a decision limit or a hyperplane separation with the greatest margin of training [14],[18]. Here SVM is used as binary classifier. It fits the hyperplane between the classes and it is providing maximum separation margin.

## IV. EXPERIMENTAL ANALYSIS

An accuracy of classifier can be formulated as,

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (10)$$

Apart from Accuracy parameter there are also some statistical techniques used for result analysis.

$$Sensitivity = \frac{TP}{TP+FP} \quad (11)$$

$$Specificity = \frac{TN}{TN+FN} \quad (12)$$

The proposed method achieved a specificity of 99% and sensitivity of 0% on selected dataset. The performance measure in terms of accuracy shows in TABLE I.

Table 1 Comparative Result of Methods

| Author | Methods | Accuracy |
|---|---|---|
| Goatman et al | Vessel Based Features | 91.10% |
| Hassan et al | Fixed Size Scanning window | 89.40% |
| Agurto et al | AM-FM, fractal analysis,granulometry | 93% |
| Lee et al | Statistical texture spectrum and fractal analysis | 98.50% |
| Gupta et al | Texture and Vesselness | |

 

| | | |
|---|---|---|
| | features | 95.75% |
| Shuang Yu | Multi-level Gabor Filter | 95.23% |
| **Proposed Method** | K-means clustering,Morphological operations, Statistical texture features | 99% |

## V. CONCLUSION

In the proposed algorithm we have used image pre-processing techniques, k-means based image segmentation techniques and morphological operations to extract blood vessels for detecting the neovascularization using support vector machine. Timely and accurate detection of NV is important as NV is severe type of diabetic retinopathy. For performing these techniques, we have used publicly available database diretdb0, diaretdb1. The algorithm achieves an average accuracy of 99 % on randomly selected dataset.

The future work of this procedure will expand to extract additional features with the help of GLCM(Gray-Level Co-Occurrence Matrix). Which will be useful to extract second order statistical texture features. So that, it will be helpful for analysis and detection of DR.

## REFERENCES

[1] M. D. Abr'amoff, M. K. Garvin, and M. Sonka, "Retinal imaging and picture analysis," IEEE Reviews in Medicine Engineering, vol. 3, pp. 169–208, 2010.

[2] K. A. Goatman, A. D. Fleming, S. Philip, G. J. Williams, J. A. Olson, and P. F. Clear, 'New color fundus picture detection vessels'" IEEE transactions on medical imaging, vol. 30, no. 4, pp. 972–9, 2011.

[3] R. Welikala, M. Fraz, J. Dehmeshki, A. Hoppe, V. Tah, S. Mann, T. Williamson, and S. Barman, "Genetic algorithmic program primarily based on various features combined with two classification for the machine driven detection of proliferative diabetic retinopathy,"processed Medical Imaging and Graphics, vol. 43, pp.64-77, 2015.

[4] R. Welikala, J. Dehmeshki, A. Hoppe, V. Tah, S. Mann, T. Williamson,and S. Barman, "Automated detection and dual classification of the proliferative diabetic retinopathy" Computer Methods and programs in biomedicine, vol. 114, no. 3, pp. 247-261,2014

[5] S. S. A. Hassan, D. B. L. Bong, and M. Premsenthil, "Detection of Neovascularization in Diabetic Retinopathy," Journal of Digital maging, vol. 25, no. 3, pp. 437–444, 2012.

[6] C. Agurto, H. Yu, V. Murray, M. S. Pattichis, S. Barriga, W. Bauman, and P. Soliz, "Detection of neovascularization at intervals the optic disc follow associate AM-FM illustration, granulometry, and vessel segmentation", Annual International Conference of the IEEE Engineering in and Biology Society, vol.2012, pp. 4946-9,2012.

[7] Q. Li, "Neovascularity detection based upon fractal analysis and texture analysis with diabetic retinopathy interaction effects," PLoS ONE, vol. 8, no. 12, 2013.

[8] S. Eswari and S. Rajeswari, "Recent vessels detection method in retinal fundus images A Survey," in In Computing Technologies and Intelligent Data Engineering (ICCTIDE), International Conference on, pp. 1–6, 2016.

[9] K. Saranya, B. Ramasubramanian, and S. Kaja Mohideen, "A novel approach for the detection of new vessels in the retinal images for screening Diabetic Retinopathy," 2012 International Conference on Communication and Signal Processing, pp. 57–61, 2012.

[10] S. D. Kasurde and S.N.Randive, "An Automatic Detection of Proliferative Diabetic Retinopathy," in Energy Systems and Applications, 2015 International Conference on, pp. 86–90, 2015.

[11] B. Ramasubramanian and G. Anitha, " An Effective Approach to Detect New Vessels in Diabetic Retinopathy images," International Journal of Engineering and Innovative Technology(IJEIT), vol. 2, no. 3, pp. 240-244, 2012.

[12] G. Gupta, S. Kulasekaran, K. Ram, M. Sivaprakasam, M. N. Joshi and R. Gandhi, "Local Neovascularization Characterisation and Identification of retinal retinopathy in proliferative diabetic fundus Images, "Medical Imaging and Graphics Computerised, vol. 55, 55 Pp. 132– 124, 2017.

[13] "Automatic Neovascularization Identification of Optic Disc Region based on Machine Learning" IEEE Transaction.

[14] N. Cristianini and J. Shawe-Taylor, An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. Cambridge University Press, 2000.

[15] T. Kauppi, V. Kalesnykiene, J.-k. Kamarainen, L. Lensu, and I. Sorri, "DIARETDB0 : Evaluation Database and Methodology for Diabetic Retinopathy Algorithms," Machine Vision and Pattern Recognition Research Group, Lappeenranta University of Technology, Finland.,pp. 1–17, 2006.

[16] David A. Clausi, "An analysis of the co-occurrence texture statistics based on gray levels", Can. J. Remote Sensing, vol. 28, no. 1, pp. 45-62, 2002.

[17] M.S.Miri, and A. Mahloojifar "A Review of Contrast to Determine Retinal Image Enhancement Techniques "ICSIPA (2009) 90 – 94.

[18] M.Rajesh Babu, BVNR Siva Kumar. Morphological process and SVM classification of diabetic retinopathy diagnosis. September 2016 , IJIRT ,Volume 3 Issue 4 ,ISSN: 2349- 6002.

[19] J.A. Hartigan and M.A. Wong, "Algorithm AS 136: A algorithm of k-means clustering ," Applied Statistics, vol. 28, pp. 100-108, 1979.

[20] Kaggle,"Diabetic Retinopathy Detection Challenge."https://www.kaggle.com/c/diabetic-retinopathydetection. Accessed: 2016-07-05.

[21] Vijay Kumar, Priyanka Gupta, "Importance of Statistical Measures in Digital Image Processing," International Journal of Emerging Technology and Advanced Engineering, ISBN: 22502459, vol. 2(8), August 2012.

### Authors Profile

***Ms. Neha Tamboli*** has completed her Graduation in Computer Science and Engineering in 2018. Currently she is doing M.Tech, in Computer Network and Information Security at SGGSIE&T, Nanded. Email:2018mns010@sggs.ac.in

***Prof. G.S. Malande*** is Assistant professor in the department of Computer Science and Engineering at SGGSIE&T, Nanded. Email: gsmalande@sggs.ac.in