

**How to Cite:**

Swathi, P., & Jyothi, S. (2022). Deep feature extraction of retinal images for diabetic retinopathy enhanced diagnosis. *International Journal of Health Sciences*, 6(S9), 4194–4209.  
<https://doi.org/10.53730/ijhs.v6nS9.13666>

# Deep feature extraction of retinal images for diabetic retinopathy enhanced diagnosis

**P. Swathi**

Research Scholar, Dept. of Computer Science, Sri Padmavati Mahila University, Tirupati

Email: [swathivinubaby@gmail.com](mailto:swathivinubaby@gmail.com)

**S. Jyothi**

Professor, Dept. of Computer Science, Sri Padmavati Mahila University, Tirupati

Email: [Jyothi.spmv@gmail.com](mailto:Jyothi.spmv@gmail.com)

**Abstract**---Most people of working age groups globally around are affected by diabetic retinopathy, which is the most common cause of preventable vision impairment. Recent studies have improved our comprehension of the need for more effective and affordable methods of identifying, managing, diagnosing, and treating retinal disease in clinical eye care practice. The development of computer-aided diagnosis tools must consider the significance of diabetic retinopathy screening programmes and the challenge of obtaining a valid early diagnosis of diabetic retinopathy at an affordable price. Retinal image analysis using computer-aided disease diagnosis may make it easier to screen the population for diabetes mellitus on a mass scale and may also assist physicians to make better use of their time. The creation of image processing models using big data analysis may run into many technical challenges, including the visualization analysis needed by image processing technologies, the semantic expression, and storage of large image samples, the complexity of the algorithms needed for feature extraction, recognition, and prediction analysis using big data, as well as the time and memory requirements.

**Keywords**---Diabetic Retinopathy, Feature Extraction, Deep Feature Extraction, Image Processing.

## 1. Introduction

A common retinal consequence of diabetes is diabetic retinopathy (DR). It is a significant contributor to blindness in both middle-aged and older age groups. A total of 23.6 million people, or 7.8 percent of the US population, have diabetes, although only 17.9 million cases are officially diagnosed, according to the

National Diabetes Information (US) database. To prevent visual loss, it is crucial to find the condition early through routine screening. An automated DR diagnosis system can be incredibly helpful in this process because a huge population needs to be checked, and that screening needs to be repeated.

Ophthalmologists employ colour fundus images to research eye conditions including diabetic retinopathy. Figure 1 displays a typical retinal picture labelled with different DR characteristic components. Small red dots known as microaneurysms are saccular pouches that form because of localised capillary wall distension [14]. Haemorrhages, which are large blood clots, could potentially result for retinal infection. Lipid deposits that appear as bright yellow lesions are called hard exudates. The optic disc is the bright, spherical area from which the blood vessels project. The fovea, which is in retina's centre, is defined as a vision is sharpest. To assess the severity of diabetic retinopathy complications, consider the spatial distribution of exudates, microaneurysms, and hemorrhages, particularly in relation to the fovea.

The classification model and accuracy rate efficiency are influenced by the numerical characteristics of picture features, which are the data representations. Numerous feature extraction methods have been developed recently, and each method has its own advantages and disadvantages [15, 16, 17, 18]. Relevant features are provided by a advanced feature extraction technique. The diabetic Retinopathies (DR) picture dataset, which was compiled from a variety of patients, served as the basis for this paper's performance evaluation of the data models produced by various feature extraction methods.

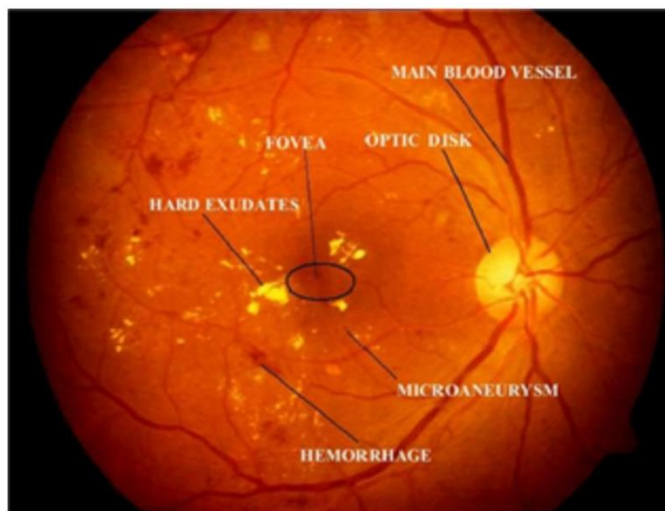


Figure 1: Typical retinal image with feature components of Diabetic Retinopathy

## 2. Literature review

The following details the related study that numerous researchers have conducted.

Blood vessels have been extracted using morphological techniques by Shilpa Joshi et al. [1]. For DR detection, Kittipol Wisaeng, et al. [2] have presented a method leveraging FCM. Priya R., et al computer-based technique [3] to distinguish between the various stages of DR by using several classifiers is illustrated. A technique developed by Madhura Jagannath, et al. [4] uses morphological procedures to identify blood vessels as well as textural cues to automatically detect DR. Using ANN (Artificial Neural Network) and SVM, Oliver Faust, et al. [5] examined techniques for classifying DR as normal, NPDR, and PDR. A method that uses a fuzzy inference system to identify exudates and categorise photos to determine the severity of DR has been proposed by M. Ponnibala, et al. [6]. (FIS).

A method for detecting exudates has been developed by Archana G, et al. [7] using segmentation based on brightness associative criteria. Textural feature estimation based on spatial parameter grey values has been presented by Robert M. Haralick, et al. [8]. These textural properties have been demonstrated to be useful in applications involving image categorization. There are a total of 28 texture features that may be extracted from gray-tone spatial-dependence matrices. Rafael C. Gonzalez, et al [9] detailed the various morphological processes and applications to the considered images. Different segmentation techniques have been presented in the chapter on image segmentation. It has been described how to extract texture features from an image.

A technique for noticing microaneurysms has been put out by C. Aravind, et al. [10] that involves identification feature followed by classification using SVM. The pre-processing stage includes the elimination of blood vessels and optic discs.

Using image processing, Manjiri B. Patwari, et al. [11] have described an approach for identifying and quantifying the microaneurysms. To detect microaneurysms, intensity modification, histogram equalisation, morphological operations and segmentation are followed. According to the region of orientation, solidity and convexity, C. Jayakumari and T. Santhanam [12] have developed a method for detection and categorisation of objects using modular neural networks. An approach to categorise the stages of DR using a Back Propagation Network (BPN) classifier has been proposed by P. Raghavi [13]. Before extracting the blood vascular network, the pre-processing stage entails image improvement and noise reduction using the Gabor filter. The locations of exudates, microaneurysms, and haemorrhages are the characteristics for classification.

### **3. Image Processing**

Image processing is the technology with more advancement. The fields of Agriculture, textile and transportation industries have effectively applied image modification, image coding, compression, segmentation, and other related technologies [22]. Traditional image processing techniques, on the other hand, can't handle the large number of image samples available today. This has caused a new technical field to be created with a focus on exploring image processing technology and creating a model for increasing the level and efficiency of image processing. Here Patients with diabetes and related retinal issues have their retinas imaged using MRI technology.

## 4. Big Data

Big Data is one of the efficient fields in the current data handling system for handling a heavy dataset [19]. The big data embraces videos, images, and any type of data. One may choose to apply Big Data analytics to maximise the utility of Big Data in the future. Analytics are frequently applied to both structured and unstructured data. Hardware, software, and algorithms must be used to analyse the original data for additional implementation. Today, it is more beneficial for businesspeople, healthcare providers, and military groups to analyse hidden patterns and discover unidentified correlations. MRI image data refers to the large image sizes being employed in Big Data and image data analytics.

## 5. Methodology

The method is a set of principles or instructions on how to go about obtaining and verifying knowledge about a certain subject. As a foundation for their study, various branches of science have produced quite diverse bodies of technique. Here Patients with diabetes and related retinal issues have their retinas imaged using MRI technology. The methodology for diabetic retinopathy image analysis is shown in figure 2.

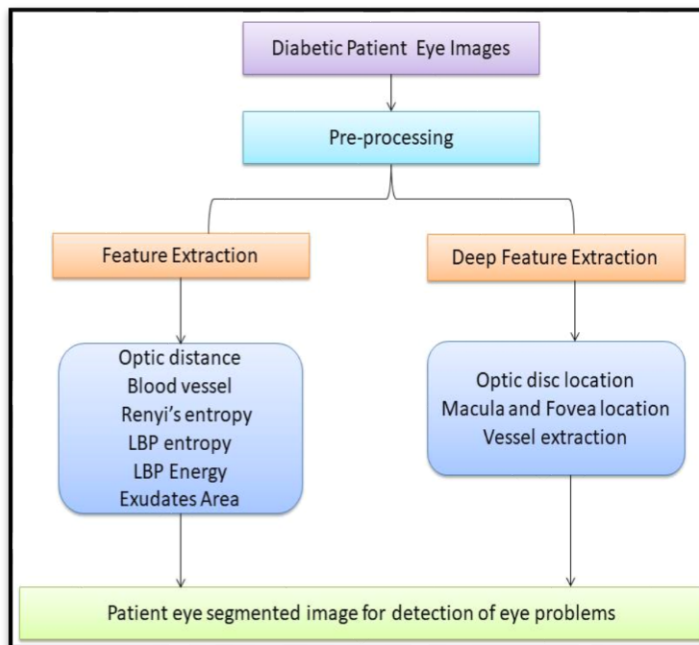


Figure 2: Methodology of Image Analysis

### 5.1. Feature extraction

To make the process of identifying the pattern simple, feature extraction describes the pertinent shape information that is included in a pattern [20]. Feature extraction is a particular kind of dimensionality reduction that is used in pattern recognition and image processing. Finding the most pertinent information in the

original data and representing it in a lower dimensional space are the fundamental objectives of feature extraction. The input data will be transformed into a reduced representation set of features when the input data to an algorithm is too vast to be processed and it is thought to be redundant (many data, but not much information) and it also named features vector. Feature extraction is the process of turning a set of features from the input data. If the extracted features are appropriately selected, it is anticipated that the features set will extract the necessary information from the input data to carry out the intended task utilising this condensed representation rather than the full size input. A new area of study in the science of image processing is pattern recognition. It has been used for a variety of tasks, including data entry, postal address reading, check sorting, tax reading, character recognition, document verification, reading bank deposit slips, obtaining information from checks, and applications for credit cards, health insurance, loans and tax forms.

In a character recognition system, feature extraction is carried out after the pre-processing stage. The main goal of pattern recognition is to correctly classify an input pattern as one of the potential output classes. The two main stages of this approach are feature selection and classification. Because a poorly chosen set of characteristics will prevent the classifier from being able to distinguish them, feature selection is essential to the entire process. According to Lippman's criteria for selecting features, they must "contain the information necessary to discriminate between classes, be insensitive to irrelevant variability in the input, and also be limited in number to permit efficient computation of discriminant functions and to minimize the amount of training data required." The goal of feature extraction, a crucial stage in the development of any pattern classification, is to obtain the pertinent data that defines each class. Relevant features are retrieved from objects and alphabets in this procedure to create feature vectors. Classifiers then employ these feature vectors to distinguish between the input unit and the desired output unit. By looking at these features, the classifier can more easily distinguish between various classes, making classification easier. The technique of extracting the most crucial information from raw data is known as feature extraction. Finding the group of parameters that perfectly and only describe a character's shape is known as feature extraction. Each character is represented by a feature vector during the feature extraction phase, which serves as its identification.

### **5.1.1. Feature Assortment Problem**

In classification problems, especially in handwriting identification, choosing the most meaningful features is essential because:

- a) It is time-consuming to find all possible feature subsets that can be formed from the initial set,
- b) every feature is significant for at least some discriminations, and
- c) Variation both within and between classes is not overly high. Beyond a certain point, adding further features results in worse performance rather than better performance.

### 5.1.2. Feature Extraction in Medical images

The threshold approach uses properties like grey features and picture pixels to segment data. This is how it is expressed:

$$G(x, y) = \begin{cases} 0, & f(x, y) < T, \\ 255, & f(x, y) > T. \end{cases}$$

The segmented picture from the formula is not suitable when complex image information is present because it only uses one threshold for segmentation. An adaptive segmentation technique based on the various picture attributes is the global threshold method. Here is how it works:

Step 1: Choose an ideal estimated value for the global threshold during initialization, and normally choose the average of the maximum and minimum values as stated in the following formula:

$$T = 0.5 (\min G (:) + \max G (:) ).$$

Step 2. Segment the image into two groups, G1 and G2, based on the threshold T. In G1, the grey value is more than T, whereas in G2, the grey value is less than T.

Step 3. Calculate the average grey value and the average value of the G1 and G2 image sets, which are  $m_1$  and  $m_2$ , respectively.

Step 4. As demonstrated in the following equation, update the T value:

$$T = \frac{1}{2} (m_1 + m_2) .$$

Until the T value is lower than the desired value, repeating Steps 1 through 4 is necessary.

### 5.2. Deep feature extraction

The three main categories of feature extraction techniques are as follows: A. Statistical Characteristics The statistical distribution of the points is used to generate these characteristics [21]. Great speed, low complexity, and style diversity management are the highlighted features. They could also be applied to shrink the size of the feature collection. The main statistical characteristics are as follows:

- 1) Zoning: The character's frame is divided into a few overlapping or non-overlapping zones, and the characteristics are formed by analysing the densities of the points and certain strokes in various regions. Character contour direction is measured by contour direction characteristics. Another illustration is bent points. Bending points are places in an image when a stroke has a strong curvature.

- 2) **Characteristic Loci:** The number of times the line segments these vectors intersect are employed as features, and for each white point in the character's background, vertical and horizontal vectors are generated.
- 3) **Crossing and Distances:** Crossing counts the instances in which pixels in the character picture change from the background to the foreground along both vertical and horizontal axes. Distances determine the separation between the first image pixel discovered and the image's top and bottom bounds along the horizontal lines.

### **5.2.1 Transformation and Expansion Features**

Global deformations like translation and rotation have little effect on these characteristics [21]. In general, a continuous signal has more information that must be represented for categorization. A signal can also be signified by a linear combination of simple and definite functions. Series expansion is a form of compact encoding made possible by the coefficients of the linear combination. These are typical transform and series expansion features:

- 1) **Fourier Transforms:** In general, features in an n-dimensional Euclidean space are chosen based on the magnitude spectrum of the measurement vector. The Fourier Transform's capacity to identify the position-shifted characters when it only looks at the magnitude spectrum and ignores the phase is one of its most alluring features. OCR has used Fourier Transforms in a variety of ways.
- 2) **Walsh Hadamard Transform:** Because all arithmetic computations require addition and subtraction, this characteristic is better suited for high-speed computing. The performance of the Walsh Hadamard transform strongly depends on the positioning of the characters, which is its main flaw.
- 3) **Rapid transform:** It is identical to the Hadamard Transform with the exception of the absolute value operation, which is thought to have solved the position shifting issue.
- 4) **Hough Transform:** is a method for document baseline detection. It is also used to describe character parameter curves.
- 5) **Gabor Transform:** A variant of the windowed Fourier Transform, the Gabor Transform. In this instance, the window is specified by a Gaussian function rather than being a discrete size.
- 6) **Wavelets:** A series expansion method called wavelet transformation enables us to represent the signal at various resolutions.
- 7) **Karhunen Loeve Expansion:** This Eigen vector analysis seeks to condense the size of the feature set by generating new features that are linear combinations of the original features.
- 8) **Moments:** Moment normalisation works to make identifying an object in an image independent of size translation and rotation.

### **5.2.2 Geometrical and Topological Features**

These characteristics have a high tolerance for distortions and stylistic changes and may represent both global and local character traits. The contour of the object may be encoded in these topological properties, or it may be necessary to know what kind of components make up the thing.

- 1) Strokes: Characters are made up of these basic building blocks. Arabic characters are made up of curves and splines, which are more sophisticated than the basic lines and arcs that make up Latin characters. A stroke is also described in on-screen character recognition as a line segment that runs from pen down to pen up.
- 2) Stroke Directions and Bays: When writing a character, the order in which the pen moves in different directions is employed to define its qualities.
- 3) Chain Codes are a feature that is created by mapping a character's strokes into a dimensional parameter space made up of codes.
- 4) End places where line segments and loops intersect.
- 5) Angle characteristics and strokes connections.

## 6. Execution

To diagnose the diabetic retinopathy various patient's eye MRI images sets are considered due to large amount of space is required for storing data big data MongoDB is used. At first, required packages are installed and then retrieved eye MRI images from MongoDB to python for diagnosis.

Initially, noise is removed from the eye MRI images. Feature extraction algorithm is subjected to the images to segment the regions and segregated depending on the initial pixels to diagnose the level of infection in eyes. Here different RGB colors are applied with feature extraction algorithm for easy analysis to differentiate optic disc, blood vessel, Renyi's entropy, LBP Entropy, LBP energy, Exudates Area to know the level of eye infection.

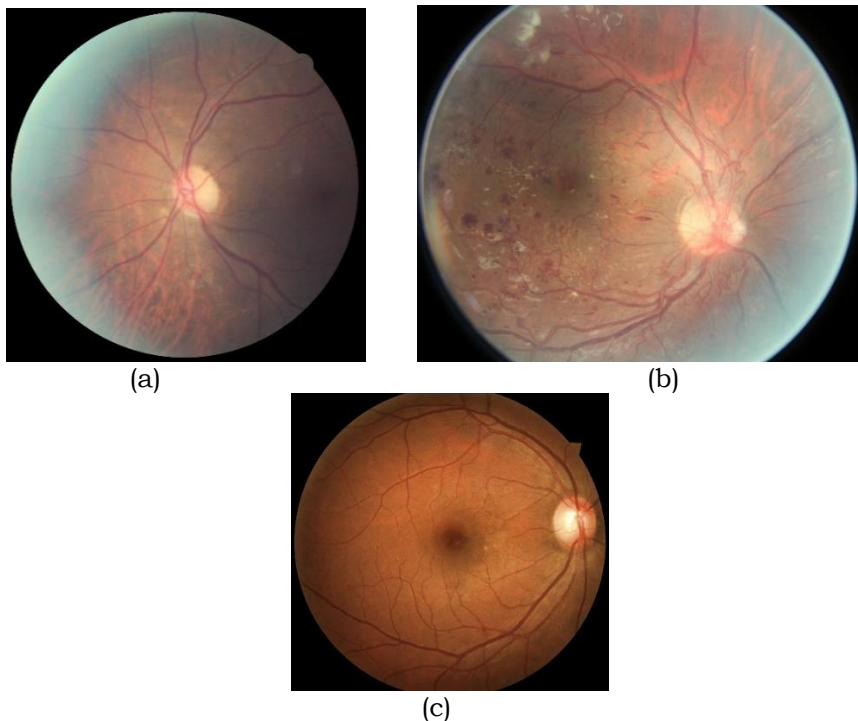


Figure 3: Diabetic Retinopathy Eye MRI image

Here figure 3(a), (b), (c) shows the input images of diabetic Retinopathy, the optic disc extracted image is shown in figure 4.

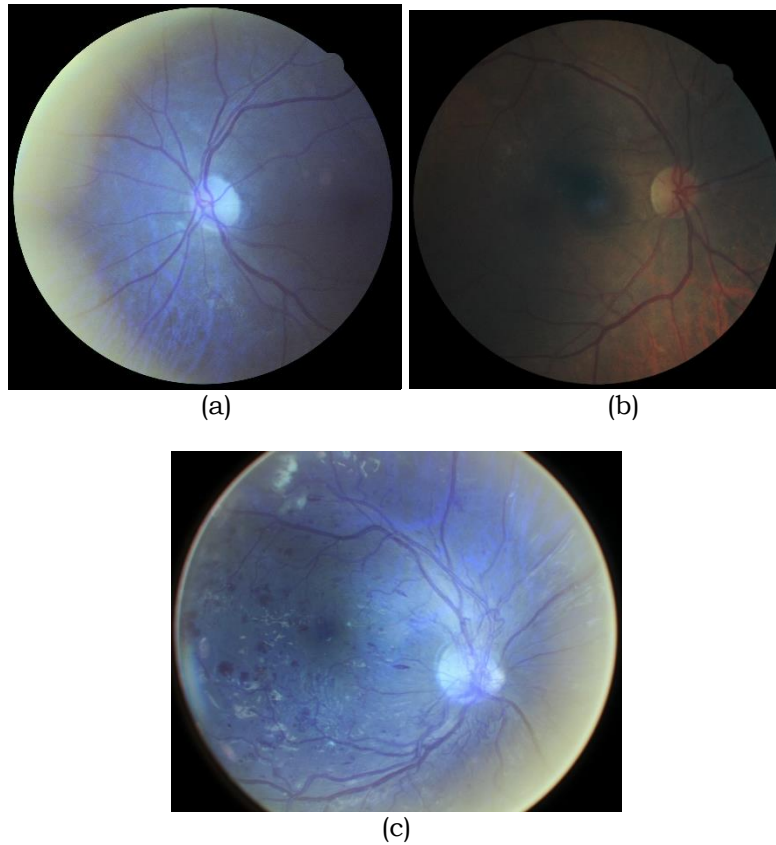


Figure 4: Feature Extraction to classify optical disc in DR Eye MRI Image

In figure 4 classifies the optical disc which is a nerve that point in the eye that connects to retina. Here, by observing optic distance in figure 4(a) is mildly infected, figure 4(b) is having moderately infected whereas figure 4(c) is proliferative diabetic retinopathy attacked

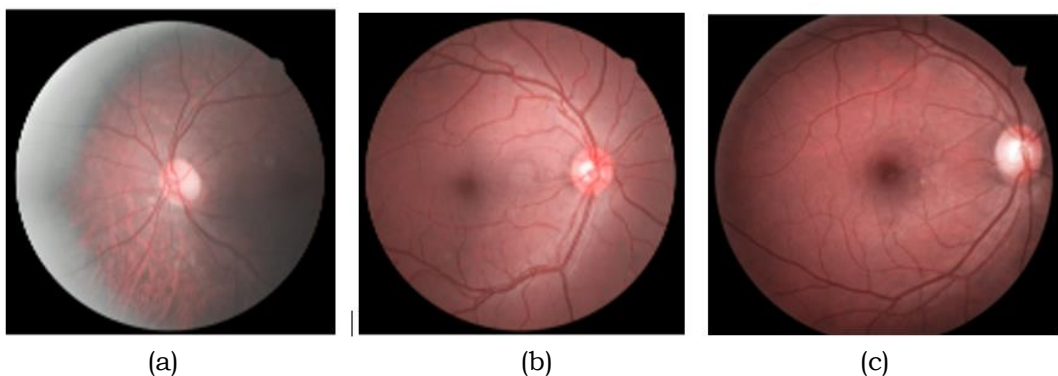


Figure 5: Feature Extraction to classify blood vessels in DR Eye MRI Image

The figure 3 classifies blood vessels in the images which is a beneath tissue that cover the white layer of the eye is analysed with redness which forms the subconjunctival hemorrhage infection. Here by observing figure 5(a) is affected with mild infection, figure 5(b) eye is affected moderately and figure 5(c) is severely affected with eye infection.

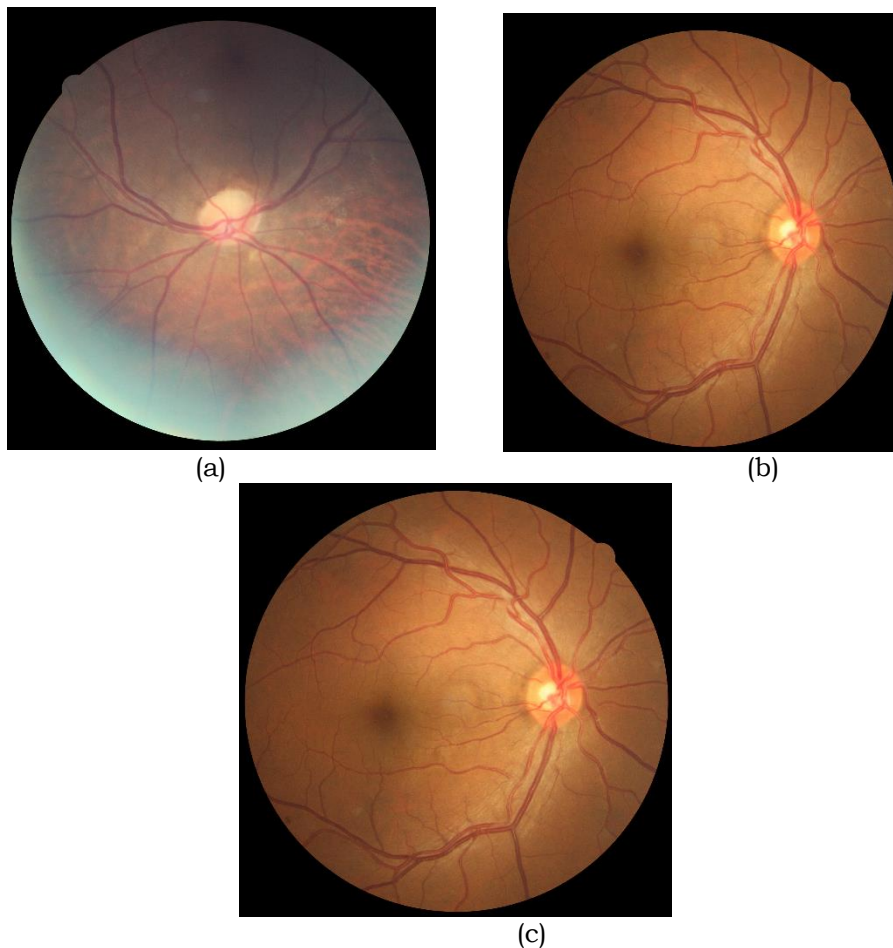


Figure 6: Feature Extraction to classify Renyis entropy in DR Eye MRI Image

The figure 6 represents Renyis entropy which detects artifacts in eye blinks that causes loss of eye. Here by observing figure 6(a) is affected with mild infection, figure 6(b) eye is affected moderately and figure 6(c) is severely affected with eye infection.

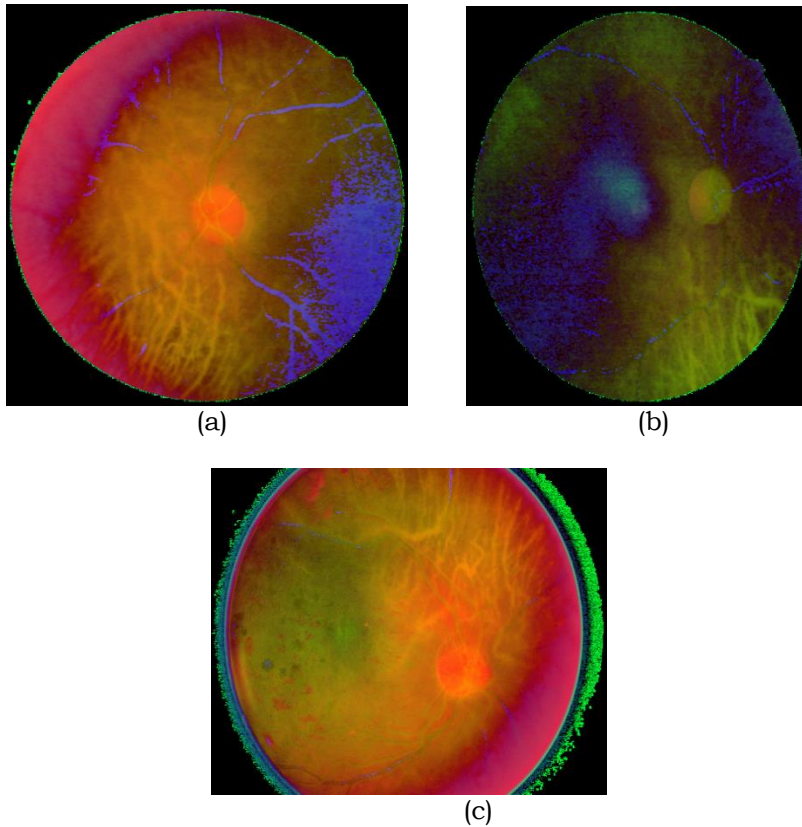
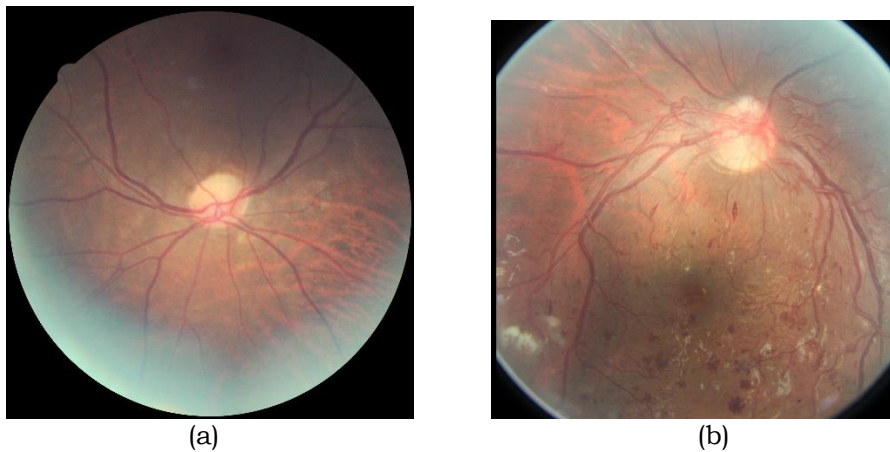
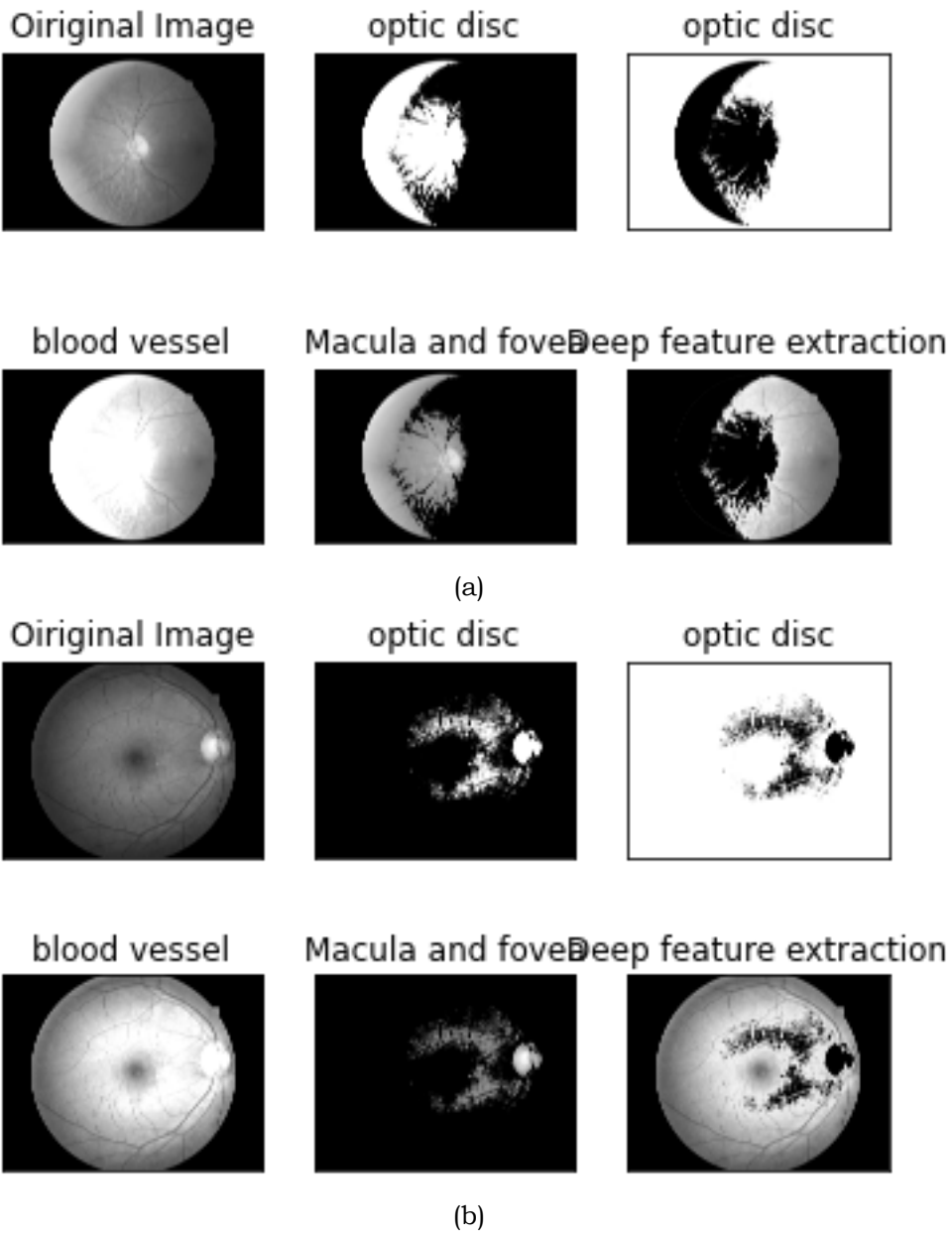


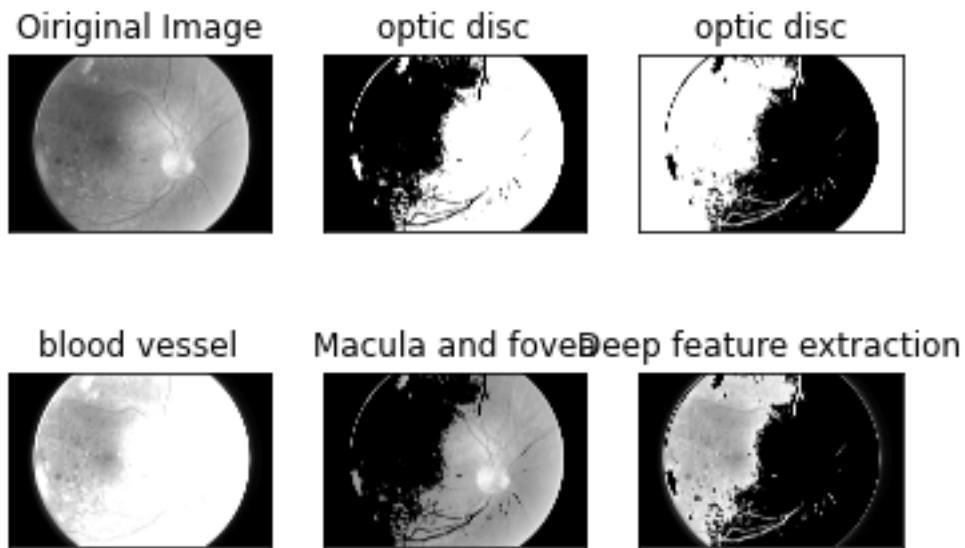
Figure 7: Feature Extraction to classify in DR Eye MRI Image

The figure 7 represents the Lateral Geniculate Bodies (LBP) entropy and LBP energy which tracks the visual of eye sight. In image analysis by using texture features it relives the stage of information with shape coarser texture variations that tells the glaucoma. Here by observing figure 7(a) is affected with mild infection, figure 7(b) eye is affected moderately and figure 7(c) is severely affected with eye infection.









(c)

Figure 10: Analysis of Optic disc, Macula & fovea location, Blood vessel of DR Eye

From Figure 10. Optic disc is classified to point nerve in the eye that connects to retina. Macula and fovea location represents cone cells are attached to ganglion cells in the retina, with finest sight called fovea vision. Blood vessel extraction image represents beneath tissue that cover the white layer of the eye is analysed with redness

In figure 10 (a) is Optic disc is affected mildly Macula and fovea location is moderately effected and blood vessels are moderately effected. Whereas, in figure 10 (b) is Optic disc is affected moderately Macula and fovea location is mildly effected and blood vessels are mildly effected, and in figure 10 (c) is Optic disc is affected moderately effected, Macula and fovea location is moderately effected and blood vessels severely effected. So figure 10(b) eye MRI image patient has immediately surgery has to be done.

## 7. Conclusion

The diabetic retinopathy MRI image analysis is a limited summary of diagnosis meant to be comprehensive, understands and assess potential diagnosis enables the investigator making a better prognosis and providing the best medication to the patients. In this study, feature extraction technique is applied to the input eye MRI images to extract the features of infections like optic disc, blood vessel, Renyi's entropy, LBP Entropy, LBP energy, Exudates area to know the level of eye infection. DFET is applied for thorough analysis of the rental MRI image. In depth optic disc, Macula and fovea location and blood vessels in the MRI image is analysed for knowing the stage of the rental image. By this analysis, specific patient has intended to tell us for need of medication or suitable medical advice, diagnosis or treatment.

## References

- [1] Shilpa J and Dr P.T. Karule (2012) "Retinal Blood Vessel Segmentation", IJEIT, Vol 1, Issue 3.
- [2] Kittipol Wisaeng, Nualsawat Hiransakolwong, and EkkaratPothiruk (2012) "Automatic Detection of Exudates in Diabetic Retinopathy Images", JCS, pg. 1304-1313, Science Publications.
- [3] Priya.R, Aruna.P (2013) "Diagnosis of Diabetic Retinopathy using Machine Learning Techniques", ICTACT, Vol 03.
- [4] Madhura Jagannath Paranjpe, Prof. M N Kakatkar (2013) "Automated Diabetic Retinopathy Severity Classification using Support Vector Machine", IJRSAT, Vol 3, Pg-086-091.
- [5] Oliver Faust, Rajendra Acharya U., E. Y. K. Ng, Kwan-Hoong Ng & Jasjit S. Suri (2010) "Algorithms for the Automated Detection of Diabetic Retinopathy Using Digital Fundus Images: A Review", Springer Science+Business Media, J Med Syst.
- [6] M. Ponni Bala, S. Vijayachitra (2012) "Computerised Retinal Image Analysis to Detect and Quantify Exudates Associated with Diabetic Retinopathy", IJCA, No. 2, Vol 54.
- [7] Archana .G, Avinaya V, Chegireddy Keerthi , Gayathri Shivaram, and S. Vasanthi, (2013) "Abnormality Detection and Its Severity Classification in Retinal Images", IJREAT, Vol 1, Issue 1.
- [8] Robert M. Haralick, K. Shanmugam, Its'hak Dinstein, (1973) "Textural Features for Image Classification", IEEE TSMC, Vol.3, pp.610- 621.
- [9] Rafael C. G, Richard C. W (2009) "Digital image processing", Edi-3, Chapter 9, Chapter 10, Chapter11, Pg.no.627-833., ISBN: 978-81- 317-2695-2, Prentice Hall.
- [10] C.Aravind, M.PonniBala, S. Vijayachitra (2013) "Automatic Detection of Microaneurysms and Classification of Diabetic Retinopathy Images using SVM Technique", IJCA, ICIIIOSP, pg. 0975 – 8887.
- [11] Manjiri B. Patwari, Ramesh R. Manza, Yogesh M. Rajput, Manoj Saswade, Neha Deshpande (2013) "Detection and Counting the Microaneurysms using Image Processing Techniques", IJAIS, Vol.6, FCS.
- [12] C. Jayakumari and T. Santhanam (2007) "Detection of Hard Exudates for Diabetic Retinopathy using Contextual Clustering and Fuzzy Art Neural Network", AJIT Medwell Journals, Pg. 842-846.
- [13] P. Raghavi (2013) "Classification and Grading of Diabetic Retinal Images for Implementation of Computer-Aided Diagnosis System", IJERT, Issue 8, Vol. 2.
- [14] R. J. Mullen, D. N. Monekosso and P. Remagnino (2013) "Ant algorithms for image feature extraction", Expert Systems with Applications, vol. 40, no. 11, pp. 4315-4332.
- [15] J. Qian, J. Yang and G. Gao (2013) "Discriminative histograms of local dominant orientation (D-HLDO) for biometric image feature extraction", Pattern Recognition, vol. 46, no. 10, pp. 2724-2739.
- [16] K. Xiao, A. L. Liang, H. B. Guan and A. E. Hassanien (2013) "Extraction and application of deformation-based feature in medical images", Neurocomputing, vol. 120, pp. 177-184.

- [17] C. Li, Q. Liu, J. Liu and H. Lu, (2013) "Ordinal regularized manifold feature extraction for image ranking", *Signal Processing*, vol. 93, no. 6, pp. 1651-1661, 2013.
- [18] P. L. Stanchev, D. Green Jr. and B. Dimitrov. (2003), "High level colour similarity retrieval", *International Journal of Information Theories and Applications*, vol. 10, no. 3, pp. 363-369.
- [19] Goutam, S. Gupta, S. Deepti, (2016) "Big Data Analytics: Image Enhancement Based Approach", *International Journal in Advance Research in Computer Science, Software Engineering* vol. 5, pp. 570 -573.
- [20] M. Saad (2008) "Low-level color and texture feature extraction for content-based image retrieval," Final Project Report, EE K 381, pp. 20-28.
- [21] T. Deselaers, D. Keysers and H. Ney (2008) "Features for image retrieval: an experimental comparison," *Information Retrieval* 11, vol. 2, pp. 77-107.
- [22] Celebi M. E., Kingravi H. A., Uddin B., et al. (2007) "A methodological approach to the classification of dermoscopy images". *Computerized Medical Imaging and Graphics*. 2007;31(6):362-373. Doi: 10.1016/j.compmedimag.2007.01.003.