

# IMPROVED AUTOMATIC DETECTION OF GLAUCOMA USING CUP-TO-DISK RATIO AND HYBRID CLASSIFIERS

Deepthi K Prasad<sup>1</sup>, L. Vibha<sup>2</sup> and K.R. Venugopal<sup>3</sup>

<sup>1,2</sup>Department of Computer Science and Engineering, BNM Institute of Technology, India

<sup>3</sup>University Visvesvaraya College of Engineering, Bangalore University, India

## Abstract

*Glaucoma is one of the most complicated disorder in human eye that causes permanent vision loss gradually if not detect in early stage. It can damage the optic nerve without any symptoms and warnings. Different automated glaucoma detection systems were developed for analyzing glaucoma at early stage but lacked good accuracy of detection. This paper proposes a novel automated glaucoma detection system which effectively process with digital colour fundus images using hybrid classifiers. The proposed system concentrates on both Cup-to Disk Ratio (CDR) and different features to improve the accuracy of glaucoma. Morphological Hough Transform Algorithm (MHTA) is designed for optic disc segmentation. Intensity based elliptic curve method is used for separation of optic cup effectively. Further feature extraction and CDR value can be estimated. Finally, classification is performed with combination of Naive Bayes Classifier and K Nearest Neighbour (KNN). The proposed system is evaluated by using High Resolution Fundus (HRF) database which outperforms the earlier methods in literature in various performance metrics.*

## Keywords:

*Glaucoma, Optic Nerve, Cup-to-Disk Ratio, HRF Database, Hybrid Classifier*

## 1. INTRODUCTION

Glaucoma is a group of disease associated with human eye that causes permanent blindness without any symptoms and warnings. Initially, glaucoma increases Intraocular Pressure (IOP) [2] [9] [10] and if not diagnosed at this stage, it automatically destroys the optic nerve and eventually leads to blindness. Optic nerve is located at the back of eye which carries visual information from eye to brain. The Fig.1(a) represents as image of the back of healthy eye, where optic nerve is located. The Fig.1(b) shows the damaged optic nerve which affects by glaucoma. Glaucoma can affect people at any age but there is a high risk for people aged over 60, people who use steroids, people who have eye pressure and eye injury. Glaucoma affects peripheral visual field at an early stage [3]. In advance, it totally affects the vision and results in loss of visual acuity, which causes severe sight destruction and complete vision loss. There are two main types of glaucoma namely, Open-angle glaucoma and Angle-Closure glaucoma. Open-angle glaucoma develops slowly and Angle-Closure glaucoma develops quickly and these types are caused by IOP. In both the types, glaucoma is asymptomatic in the beginning.

The possible symptoms of Glaucoma are:

- Eye and head pain
- Hazy vision
- Vomiting or Nausea
- Sudden sight loss

Most glaucoma affected people remain asymptomatic until advanced stage, so early detection is essential to prevent visual damage. Commonly glaucoma [13] [14] diagnosis is based on the visual field loss tests, IOP, optic nerve and manual assessment via ophthalmoscope and fundus images. In addition to objective, glaucoma stages and its structure of Optic Nerve Heads (ONH) changes for identifying the glaucoma disease. Structural changes occurred in internal eye are one of the vital sources to detect glaucoma. ONH structures are Optic Disc Diameter, cup diameter, CDR, structural and unstructured features, etc.

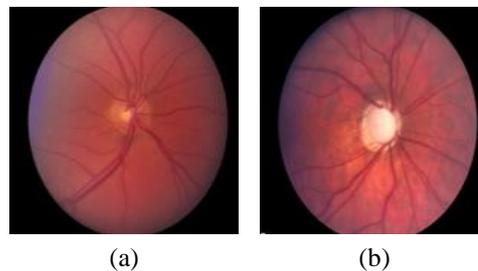


Fig.1. (a) Normal optic nerve, (b) Glaucoma detected optic nerve

Optic Disc (OD) [11] location is main region of fundus which is used to determine the severity of glaucoma. It is usually the bright component in retinal fundus images and optic disc centre is the origin of blood vessels. OD can be examined by colour, size, sharpness of margins, size and shape of cup and pattern of disc vessels. Based on the above structure of OD, segmentation is performed with template matching [16] [19], Snake based Contour Refinement [11], shape regression method [6], firefly intelligent algorithm [22] and Particle Swarm Optimization (PSO) [15].

OD can be separated into two distinct zones such as central bright zone called Optic Cup (OC) and a peripheral region as neuroretinal rim. OC is originated into the center of neuroretinal rim that intensity is high in OD. The enlarge intensity of OC is an important indicator of glaucoma succession. Cup separation can be processed by r-bends information [17], convex hull [2], shape regression method [6]; threshold level set approach [7] and boundary segmentation method [9].

Automated segmentation of cup and disc diameter is most important part towards automated disease classification. The CDR [14] indicates the diameter of the cup expressed as a fraction of the diameter of the disc. This ratio supports to determine the damage occurrence on nerve fibers. For automatic calculation of CDR [1] [4] [5] [24], OD and OC are needed to be segmented. Standard threshold CDR value of glaucoma is 0.3, while for normal it is less or equal to 0.3. The Fig.2 shows the area of optic disc and cup in retina fundus image.

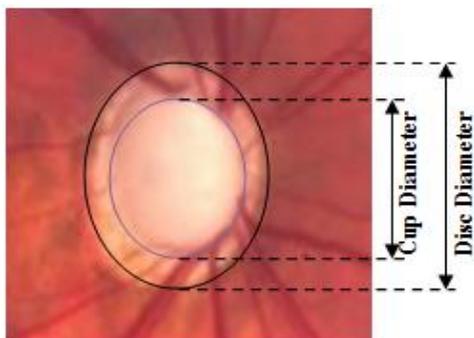


Fig.2. Optic Disc-Cup Ratio

To improve accuracy of classification feature selection is also processed along with CDR. Feature extraction is based on structured and unstructured features. Then, a feature selection algorithm is developed that can eliminate those features that affects classification accuracy. Feature extraction can be processed by using Principal Component Analysis (PCA) [18].

CDR and selected features are important factors for identifying glaucoma. These factors help to improve the classification process with different levels of stages. Classification process are using different machine learning algorithms such as K-Nearest Neighbour (KNN) [19], support vector machine (SVM) [1], Multilayer Perception and Radial Basis Function (RBF) [8] and probabilistic based method [13].

The contribution of the proposed framework is summarized as follows:

- Introducing a better pre-processing method by using RGB to Gray scale conversion, Normalization, Noise removal and contrast enhancement which all enhancing the quality of image.
- Effectively segment the optic disc by MHT algorithm and intensity based elliptic curve is used for separating cup for measuring the CDR of the retinal fundus image.
- To improve the classification, feature selection is processed by best first search, genetic search and rank search
- Finally, classification is processed based on severity and damage to the optic fiber with the help of KNN and Naïve Bayes classifiers.

The organization of this paper is as follows. In section 2, literature review in the area of glaucoma diagnosis is presented. Section 3 describes the architecture and modelling of the proposed glaucoma diagnosis system. Section 4 comprises of the graphical improvements with previous work. Section 5 presents the conclusion and future enhancements of this work.

## 2. RELATED WORK

Retinal fundus images are commonly used to detect glaucoma in earlier stage. Many researchers have proposed algorithms and automatic techniques to classify the early stage of glaucoma. Initially, pre-processing is necessity to enhance the image quality which suppresses unwanted distortions for further processing. In [1] [5][15], pre-processing technique is applied with low-pass filter and histogram equalization, then Anisotropic Diffusion Filtering [7] is used for reducing the speckle of the image and

enhance the contrast of the edges. However these techniques blur some important features in images, which degrade the performance in diagnosis.

After completion of pre-processing, Optic disc segmentation is processed with snake based contour refinement [11]. It is based on snake evolution for considering distance between OD center and OD diameter by using template matching. This method does not consider the OD size and its shape, so accuracy is low. To overcome this problem template matching [16] was used to approximately locate the optic disc center with the help of intensity information in HIS colour space. Then, shape prior term, distance regulation term and edge-based term is designed to identify the boundary of OD. However, this method suffers due to the vessel edges present in and around the OD region. To handle this, Isotropic Undecimated Wavelet Transform (IUWT) [1] and edge location refinement are proposed for segmenting the OD with many features. It is masking the blood vessels for facilitate the segmentation of different images. However, this process was effectively processed with boundary regions. Shape regression method [6] is processed with OD shapes by using gradient based regression trees, which is used to learn each cascade in temporal and nasal area over OD. Then, PCA is performed on the covariance matrix of the normalized training shapes. However, covariance matrix is difficult to solve accurately and it is also time consuming.

Separation of OC from OD is important for analyzing CDR accurately without any proper refinement. Authors of paper [2] concentrated on convex hull, which extract the cup by using green plane with structuring element of cup size. It also focused on brightest intensity which appears in center point of pixels in OD. It is only processed based on intensity and not concentrated on the cup boundary. To overcome this problem, neural retina cup detection [23] is used to improve the accuracy of boundary estimation. Ellipse fitting was performed for obtained disc boundary and convex hull was used for selecting feature points around neural-retina cup region.

Methods of extracting features in fundus images have been explained in [20], [21], [8] and [2]. K-Means clustering [21] was applied for feature extraction based on k-centroid. Distances are measured from their respective centers. Then, find out the relevant point that belongs to associate the nearest centroid, which aim to minimize squared error function. This process is difficult to predict k-value because it does not work well with clusters of different size and density. To address this problem, texture features are extracted [8] such as contrast, homogeneity, correlation, entropy and energy. Gray-level-Co-occurrence matrix is used to extract these texture features in different colour models. To enhance the feature extraction both texture and intensity based features [2] are selected, which includes Local Binary patterns (LBP), wavelet based features, colour moments and auto correlograms. Feature selection was important to select the subset of original features. Minimum Redundancy Maximum Relevance (mRMR) [24] feature selection process was proposed for selecting the relevant features. It was processed with the help of mutual information values depends on each feature and its target feature set. Then, ranking process is conducted based on different criteria.

Several classification methods have been proposed for classifying the glaucoma diseases. Most of the research papers

used machine learning algorithms to solve classification problems. Support Vector Machine (SVM) [1] [2] is used, which is a linear classification method that determines maximum-margin and soft hyper plane for separating classes. Linear kernel classifier provides better result for glaucoma diagnosis. This classifier provides poor performance when numbers of features are increased. Artificial Neural Networks (ANN) [25] is involved in classification process with forward phase and backward phase. Input layer can be conveyed to the output layer through forward layer whereas output layer moving back towards input layer for calculating error and gradient value. Then, update the weighted values to classify the image as normal or glaucoma. Thus aforementioned techniques and methods result in several problems which lack in several areas like accuracy of detection mechanism, absence of appropriate selection method and not imposing any constraints classification mechanisms.

### 3. PROPOSED SYSTEM

#### 3.1 OVERVIEW OF THE PROPOSED SYSTEM

The proposed system produces a new solution for glaucoma detection based on severity of stages. High Resolution Fundus (HRF) database image is used as input image for detection. First, pre-processing is an important process in automatic detection that consists of the following steps such as RGB to Gray scale conversion, Normalization, Noise removal and Contrast enhancement. OD Segmentation is used to isolate the relevant part of the retina image and to calculate the CDR. MHT algorithm is used to segment the optic disc, then separation of OC is performed with intensity based elliptic curve method. CDR value is calculated based on the optic disc and optic cup and also normal value is 0.3. A larger ratio may result for glaucoma indication. Further, feature extraction is processed with following features such as texture, spectral and intensity features. Then, feature selection is performed with best first search, genetic search and rank search. The proposed feature selection process provides the best set of features, which effectively used for classification with low error rate. For classification process, the hybrid classifiers such as Naïve Bayes classifier and K-Nearest Neighbor are used. These combined classification techniques improve the classification process based on severity of disease and damage to optic fiber. The Fig.3 represents the architecture of the proposed system.

#### 3.2 PRE-PROCESSING

Pre-processing is an important step for enhancing the quality of retinal image. In the proposed method, image pre-processing is used for segmentation of OD and OC for accurate estimation of Cup to Disk Ratio (CDR). Pre-processing step has four stages such as

- RGB to Gray Scale
- Normalization
- Noise Removal
- Contrast Enhancement

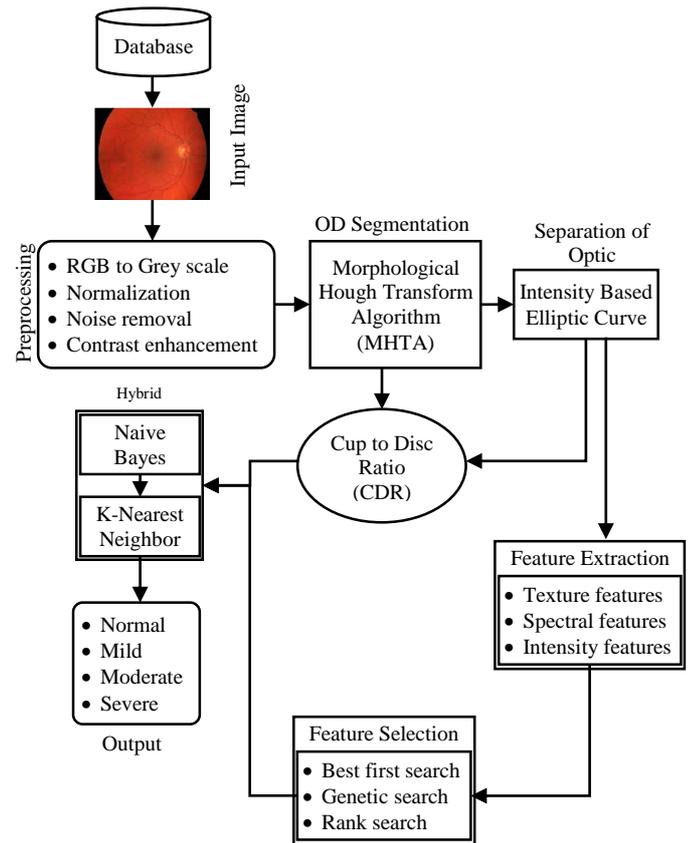


Fig.3. Architecture and Modelling

##### 3.2.1 RGB to Gray Scale:

To enhance the quality of image, RGB image is converted to Gray scale to reduce computational requirements. RGB to gray scale conversion is based on following equation,

$$I_Y = 0.333 * F_R + 0.5 * F_G + 0.1666 * F_B$$

Here,  $F_R$ ,  $F_G$ ,  $F_B$  are the intensity of  $R$ ,  $G$ ,  $B$  components respectively and  $I_Y$  is the intensity of equivalent gray level image of RGB.

##### 3.2.2 Normalization:

Normalization is used to eliminate the irrelevant principle components in images and increases the pixel intensity value. This greatly improves the image analysis performance. The main aim of normalization is to minimize the redundancies for producing well-structured relations of database images.

##### 3.2.3 Noise Removal:

Noise in images leads to performance degradation. In this work, adaptive median filter is used to remove the noise from image. Adaptive median filtering [26] has been widely used as an advanced method for standard median filtering. The main objectives of this filtering are to remove impulse noise, reduce distortion (excessive thinning or object boundary thinning) and smoothing of other types of noise. Here the noise detection accuracy is analyzed with the help of pollution level,

$$\xi = |N|/|I| \quad (1)$$

where,

$$|N| = \sum_{Y_{ij}=1} Y_{ij} \text{ that specifies the number of noise set } N \text{ and } |I| \text{ is the number of image set } I.$$

Then the median filter window sizes are initialized and set as noise pollution levels. Here there is need to calculate the minimum distance between pollution level and a set of pollution levels,

$$md = \min_{S_i \in S} \|\xi - NP_i\| \quad (2)$$

Here,  $NP$  is the set of noise pollution levels where the levels are 0.3, 0.5, 0.7, 0.9, 1 and  $S$  is the window sizes. The size of filtering window is identified as,

$$ms = S_i \|\xi - NP_i\| = md \quad (3)$$

The window size of adaptive median filter is  $m \times m$ . finally the noise from the image is removed.

### 3.2.4 Contrast Enhancement:

This method is very important for acquiring better results on automatic detection of glaucoma. In the proposed technique CLAHE (Contrast Limited Adaptive Histogram Equalization) technique is used for enhancing the image contrast as it helps in identifying the accurate optic cups. Here the contrast amplification near a specific pixel value is given by the slope of transformation function. This is directly proportional to slope of neighbourhood cumulative distribution function (CDF). CLAHE limits amplification based on clipping the histogram at predefined value while computing the CDF. Thus, this limits the slope of CDF followed by transformation function. The clip limit depends on normalization of histogram and on the size of neighbourhood region.

## 3.3 OPTIC DISC SEGMENTATION

Segmentation is defined as the process of partitioning of a digital image into several segments (number of pixels). The main goal of segmentation is to simplify the representation of an image that helps for analyzing the image easier. In the proposed work, Optic Disc Segmentation is used for isolating the relevant part of retina image and helps for calculating the CDR. Morphological Hough Transformation (MHT) Algorithm is used to segment the optic disc.

### 3.3.1 Contrast Enhancement:

Morphological operations such as erosion and dilation are focused in proposed method for eliminating blood vessels. Morphological operations can be mathematically defined by,

Erosion of an image  $I$  by the structural element  $A$  is specified by the set of operation such as,

$$I \ominus A = \{p \in Z^2 | (p+q) \in I, \text{ for every } q \in A\} \quad (4)$$

Dilation of the image by the structural element  $A$  is specified by the set of operation such as,

$$I (+) A = \{(p+q) | p \in I, q \in A\} \quad (5)$$

where,  $A$  is the original image,  $p$  and  $q$  are the locations of a pixels on the image  $I$ . Based on these operations, blood vessels are segmented from the original image. The Fig.4 describes the removal of blood vessel.

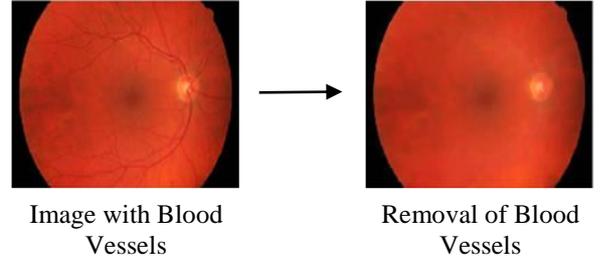


Fig.4. Removal of blood Vessels

The removal of blood vessels is more helpful in segmenting the optic discs very clearly. The portion of image that includes only the optic disc is identified. Circular Hough transformation is used here and can be expressed as,

$$y = k \cdot x + m \quad (6)$$

where,  $k$  is the gradient and  $m$  is the intersection with  $y$ -axis. The representation of a circle is alternatively by,

$$r^2 = (x-a)^2 + (y-b)^2 \quad (7)$$

Here  $r$  is the radius from the origin whereas  $a, b$  are the center of circle direction of  $x$  and  $y$ . The parameter representation of a circle is,

$$x = a + r \cos \theta \quad (8)$$

$$y = b + r \sin \theta \quad (9)$$

Finally, the accumulated array is formed for identifying the shape of the disc. Based on Hough transform optic disc is segmented from the fundus image. Then a matrix is created where each column corresponds to angles and rows correspond to bins or intervals of resulting distance  $r$ . For every point of image  $r, \theta$  values are computed and increase the values of corresponding elements of accumulator matrix. Here canny edge detector is used for collecting the edge points. The algorithm for MHT is given in Algorithm 1 as follows:

#### Algorithm 1. MHT Algorithm

**Step 1:** Select image

**Step 2:** morphological erosion process (Eq.(4))

**Step 3:** morphological dilation process (Eq.(5))

**Step 4:** Form ARR and initialize with 0

Use canny edge detector for collecting edge points

**Step 5:** for every edge point,

Compute  $\Omega(x)$

$x_c = \text{Eq.}(8)$

$y_c = \text{Eq.}(9)$

$\text{ARR}(x_c, y_c) = \text{ARR}(x_c, y_c) + 1$

End up to possible edge point

**Step 6:** Collect candidate circles

**Step 7:** Draw a circle

### 3.3.2 Separation of Cup and Disc:

After identifying the shape of optic disc, optic disc and cup have to be separated. The intensity of the disc image is calculated. Inside the disc, intensity of cup area is higher, which helps to identify the exact area of cup. The intensity is calculated by,

$$p(i, j) = \frac{1}{x-y} \sum_{x=0}^{x-j} \sum_{y=0}^{y-j} q_i(i-x, j-y) \quad (10)$$

Here  $p$  and  $q$  are the regions for every  $(i, j)$  point of the image. The optic cup is a center area of optic disc which is very important for identifying the CDR. Elliptic curve method is used for identifying the boundary of the cup. The mathematical formula for elliptical curve is,

$$y^2 = x^2 + ax + b \quad (11)$$

The optic discs and cup are segmented. The separation of disc and cup boundary is a significant part of automated disease detection. Here the CDR ratio [4] is evaluated based on the optic disc and optic cup ratio as,

$$\hat{r} = \frac{1}{1_w^T} r^T w \quad (12)$$

where,  $w$  is the linear reconstruction coefficient,  $T$  is the transpose;  $1_w$  is a vector of 1's with equal lengths to the  $w$ . Here the threshold value is specified for automatic detection. Normal CDR Value is 0.3 which specifies the safer side of disease whereas the greater ratio specifies the glaucoma disease.

### 3.4 FEATURE EXTRACTION

The feature extraction is an important process for classifying the images into two or more categories. The proposed method involves extraction of three types of features such as

- Texture Features
- Shape Features
- Intensity Features

#### 3.4.1 Texture Features:

Texture is a most needed for images which is regional descriptor that plays a major role in retrieval of features. There are several texture features which are used for the proposed system. The texture features are extracted based on co-occurrence metric  $C(i, j)$  which counts the co-occurrence of pixels with grey values ' $i$  and  $j$ ', distance  $d$ , orientation " $\theta$ " that have degree values such as 0, 45, 95 and 135. Some of the features are listed as follows:

**Contrast:** This feature returns a measure between intensity of pixel and its neighbour,

$$Contrast = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} (i-j)^2 p(i, j, \delta, \theta) \quad (13)$$

**Homogeneity:** This feature provides a value which represents proximity of distribution of elements. This has the relation of co-occurrence matrix diagonal,

$$Homogeneity = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \frac{1}{1+(i-j)^2} p(i, j, d, \theta) \quad (14)$$

**Correlation:** This measures the correlation of a pixel with its neighbour.

$$Correlation = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \frac{ij p(i, j, d, \theta) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (15)$$

**Entropy:** This specifies the measure of information contained in the degree of gray levels dispersion.

$$Entropy = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} p(i, j, \delta, \theta) \log_2 p(i, j, \delta, \theta) \quad (16)$$

**Energy:** The energy provides the total of elements with Elevated Square on the matrix of co-occurrence of gray levels.

$$Energy = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} [p(i, j, \delta, \theta)^2] \quad (17)$$

#### 3.4.2 Shape Features:

For identifying the shape of optic cups, some common shape features like circularity, aspect ratio, Elongation shape factor and compactness shape factor are extracted.

**Circularity:** This is the common shape factor with a function of an area ( $A$ ) and perimeter ( $P$ ),

$$C_{circle} = 4\pi A/P^2 \quad (18)$$

**Aspect ratio:** The aspect ratio specifies the largest diameter and smallest diameter.

$$AR = d_{min}/d_{max} \quad (19)$$

**Elongation Shape factor:** This is defined as the square root of the two moments ( $I$ ) of the pixel with the principal axes.

$$Elong = \sqrt{\frac{i_2}{i_1}} \quad (20)$$

**Compactness Shape factor:** It is a function of polar moment ( $i$ ) of a particle and a circle of equal area  $A$ .

$$Comp = \frac{A^2}{4\pi \sqrt{i_1^2 + i_2^2}} \quad (21)$$

#### 3.4.3 Intensity Features:

The intensity features are defined as the statistical level features with first order [27]. Consider  $I$  represent the gray levels of image region. The histogram  $P(I)$  is calculated by,

$$P(I) = \frac{(\text{Number of pixels with gray level } I)}{(\text{Total no. of pixels in region})} \quad (22)$$

Using  $P(I)$ , the mean  $M$  and central moments of  $I$  are defined by,

$$M = E[I^1] = \sum I^1 P(I) \quad (23)$$

$$\mu_k = E[(I - E[I])^k] = \sum_{I=0}^{N_g-1} (I - M)^k P(I) \quad (24)$$

Here  $k = 2, 3, 4$  and  $N_g$  is the number of gray levels. Based on above mentioned equations, the features of the optic cup are extracted. Then according to proposed work feature selection is performed.

### 3.5 FEATURE SELECTION

Feature selection helps to reduce the complication of classification process. With the help of above extracted features like texture, shape and intensity it is possible to automatically detect glaucoma disease. Generally feature selection process selects the relevant features which are most informative by reducing other features. Thus the reduction of features helps to reduce computation time and storage time. Features are selected

using the best first search [30], genetic search [28] and rank search [29]. All the three search techniques are implemented for feature selection. The result of search techniques is collected and the higher priority features are selected as the optimal features. Here the Local Genetic Algorithm search [28] involves reducing the search space for the classification process based on using normal mutation process or a best first search during the calculation of fitness function. The fitness function is calculated by,

$$FF = \frac{\text{Concate}(\text{feature 1, feature 2, feature 3, ...})}{\text{Total Feature}} \quad (25)$$

Best first search [30] is a general tree-search or graph search algorithm that is a feature is selected based on heuristic function. The best first search is proposed for exploring the correct feature which used for selecting the optimum feature. Rank search algorithm [29] based on improved page rank algorithm for selecting optimum feature. Based on these algorithms best results are combined on these search algorithms to select the best features. The algorithm for feature selection is given in Algorithm 2.

### 3.6 CLASSIFICATION

Classification is a decision making process that helps to identify the images given based on features. It is one of the multivariate techniques. The proposed method uses a classifier called as Hybrid Classifiers in order to acquire higher accuracy. This hybrid classifiers works by integrating Naive Bayes classifier and K nearest neighbour.

#### 3.6.1 Naïve Bayes Classifier:

Naive Bayes Classifier is a probabilistic model which is based on Bayes rule along with strong independence assumptions. This involves a simplifying conditional independence assumption which does not affect the accuracy in classifications and it works in a faster manner. The probability model for a classifier is,

$$P(Z | x_1, \dots, x_n)$$

On a dependent class variable  $Z$  with a small number of classes, conditional on several variables  $x_1$  through  $x_n$ . Based on Bayes rule,

$$P(Z | x_1, \dots, x_n) = \frac{P(Z)(x_1, \dots, x_n | Z)}{P(x_1, \dots, x_n)}$$

Here the numerator value is equal to joint probability model  $P(Z | x_1, \dots, x_n)$  that can specify with definition of conditional probability.

$$P(Z, x_1, \dots, x_n) = P(Z)(x_1, \dots, x_n | Z)$$

$$= P(Z)P(x_1 | Z)(x_2, \dots, x_n | Z, x_1)$$

$$= P(Z)P(x_1 | Z)P(x_2 | Z, x_1)(x_3, \dots, x_n | Z, x_1, x_2)$$

$$= P(Z)P(x_1 | Z)P(x_2 | Z, x_1)P(x_3 | Z, x_1, x_2)(x_4, \dots, x_n | Z, x_1, x_2, x_3)$$

Now assume that each attribute  $x_i$  is conditionally independent of every other attribute  $x_j$  for  $j \neq i$ .

$$P(Z, x_1, \dots, x_n) = P(Z)P(x_1 | Z)P(x_2 | Z)P(x_3 | Z)$$

$$= P(Z) \prod_{i=1}^n P(x_i | Z) \quad (26)$$

This specifies above assumptions, were the conditional distributions on a class variable  $z$  can be expressed by,

$$P(Z | x_1, \dots, x_n) = \frac{1}{S} P(C) \prod_{i=1}^n (x_i | Z) \quad (27)$$

Here  $S$  is a scaling factor that which depends on  $x_1, \dots, x_n$  i.e., a constant whether the values of feature variables are known.

#### 3.6.2 K- Nearest Neighbour:

This algorithm is a non-parametric lazy learning algorithm because it does not use any assumptions on underlying data distributions. Based on the proposed method, Naive Bayesian classifier values are included. Here the probability of features is taken as input for searching the nearest neighbor.

#### Algorithm 2. Feature Selection Algorithm

**Step 1:** Extract features from above section 3.4

**Step 2:** Assign them in order  $F_1, F_2, F_3 \dots$

**Step 3:** Genetic algorithm [28]:

Start

Evaluate fitness function using Eq.(25)

While (stop  $\rightarrow$  not full filled)

Select  $F_i$

Crossover  $F_i$

Mutate  $F_i$

$F_i \rightarrow BF(G)$

End

**Step 4:** Best First Search [30]:

Start

$BF \rightarrow \text{Null}$

Select  $F$  random

For each  $F_i \rightarrow$  selected

If  $F_i \rightarrow$  Adjacent  $\forall F_i$

Move to queue

$F_i \rightarrow BF(BFS)$

Else select next  $F_i$

End

**Step 5:** Page rank algorithm [29]:

Start

Rank all  $F_i \rightarrow 1$

Calculate ranks for features,

$$FR = 0.15 + 0.85 \left( FR \frac{F_1}{C(F_1)} + FR \frac{F_2}{C(F_2)} + \dots + FR \frac{F_n}{C(F_n)} \right)$$

Calculate mean value  $\forall F_i$

Mean ( $F_i$ ) = ( $\sum FR$ )/(Total  $F_i$ )

Normalize  $FR$  by,

Norm  $FR(F_i) = FR(F_i)/\text{Mean}(F_i)$

Assign  $FR(F_i) = \text{Norm } FR(F_i)$

If  $FR(F_i)$  are same for two iterations

Then  $F_i \rightarrow BF(RS)$

Repeat  $FR$

End If

End

**Step 6:** Combine BF  $\Rightarrow$   
 $[(BF(RS)) \cap BF(GA)] \cup [(BF(RS)) \cap (BF(BFS))] \cup [(BF(GA)) \cap (BF(BFS))]$

Best BF

End

Here  $F_i$  is the features,  $BF$  is the best feature, and  $FR$  is the feature of rank,  $F_1$  and  $F_2$  are the rank for features  $F_i = \{F_1, F_2 \dots\}$  and  $C(F_i)$  is the total number of features on the images.

Here this involves the training set  $D$  and a test object  $y = (a', b')$ . Then KNN algorithm computes the similarity between  $C$  and all training objects  $(a, b) \in D$  for determining the nearest neighbour list ( $D_z$ ). Here,  $a'$  is the data of test object where  $b'$  is its class,  $a$  is the data of training object where  $b$  is its class. So similarity can be done between  $C$  and every object  $(a, b) \in D$  that are,

$$D_z(a, b) = \sqrt{\sum_{i=0}^n (a' - a)^2} \quad (28)$$

The features are classified as Mild, Moderate, Normal and Severe. This involves the classification with both CDR ratio and feature selection result. According to classification, the result is computed as follows:

- Normal value  $\leq 0.3$
- Mild value  $\leq 0.5$
- Moderate value  $\leq 0.7$
- Severe value  $\leq 0.8$

The proposed model automatically detects the glaucoma disease. The Fig.5 describes the proposed architecture of hybrid classification.

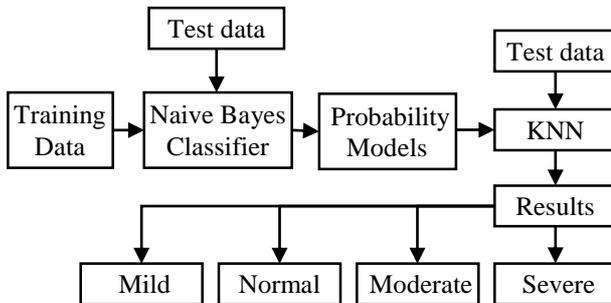


Fig.5. Proposed Architecture of Hybrid Classifier

## 4. EXPERIMENTAL EVALUATION

The proposed process involves improved automatic detection of glaucoma disease with the help of CDR ratio that is obtained by optic disc segmentation and classification procedures. The proposed system is implemented in MATLAB (MATrix LABoratory).

### 4.1 DATABASE

HRF (High Resolution Fundus) database was established by collaborative research group and the main goal of this HRF dataset is supporting comparative studies on automatic

segmentation algorithms on the retinal images that are worked with several segmentation and classification algorithms on retinal fundus images. The proposed method used HRF images for automatic detection process. This public dataset contains the 15 images of healthy patients, 15 images of diabetic retinopathy patients and 15 images of glaucomatous patients. In this dataset, every image has binary gold standard segmentation. These images are obtained with CANON CF-60 UVi cameras that are equipped with Canon EOS-20D with degree (60) field of view (FOV). The size of HRF images are  $3506 \times 2336$  pixels, 24-bits per pixel and can be stored in JPEG format. For every image, a binary mask determining the FOV is provided.

### 4.2 PERFORMANCE METRICS

The performance of the segmentation and classifier can be measured in form of Sensitivity (or) Recall, Specificity and accuracy. Here the outcomes are specified by True Positive, True Negative, False Positive and False Negative. Usually, True Positive (TP) is defined as the both training algorithm and testing algorithm results are positive, True Negative (TN) is defined as the both training algorithm and testing algorithm results are negative, False Positive (FP) is that training algorithm result is positive and testing algorithm are negative and finally False Negative (FN) is the training algorithm result is negative and testing algorithm results are positive. Performance metrics are described as follows:

- *Sensitivity*: It is defined as the measure of the proportion of positives whereas the disease affected peoples are identified correctly, which can be expressed in formulation as follows:

$$Sensitivity(\%) = \frac{TP}{TP + FN} \times 100\% \quad (29)$$

- *Specificity*: This can be measured by proportion of negatives where the humans are not affected which are identified correctly. This can be formulated by,

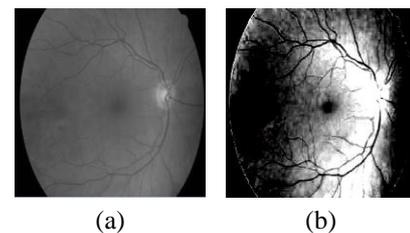
$$Specificity(\%) = \frac{TN}{TN + FP} \times 100\% \quad (30)$$

- *Accuracy*: This can be measured by the overall performance of the proposed process that includes classification and segmentation. This can be computed by,

$$Accuracy(\%) = \frac{TN + TP}{TN + TP + FP + FN} \times 100\% \quad (31)$$

### 4.3 RESULT AND DISCUSSIONS

Segmentation algorithms (such as MHT) and several classification algorithms (hybrid classifiers) are compared in this work. The Fig.6 describes the pre-processing results such as RGB to grey scale conversion, normalization, noise removal and contrast enhancement.



(a)

(b)

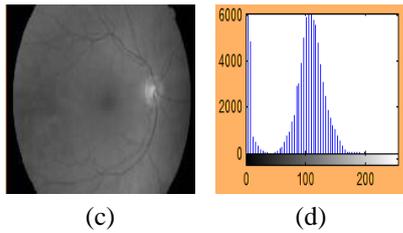


Fig.6. Pre-processing Results: (a) RGB to Gray (b) Normalization (c) Noise Removal and (d) Contrast Enhancement

In optic disc segmentation, the MHT algorithm is used to remove blood vessels and separation of optic disc and cup is done by using the intensity based elliptic curve method. This method helps in improving the accuracy. Results of optic disc segmentation are depicted in Fig.7.

The feature extraction process extracts several kinds of features such as textural features, Shape features and intensity features. After extracting the features, selection of optimal features is done using rank search, genetic search and best first search techniques Hybrid classification technique combines Naive Bayes and K-Nearest Neighbour for the automatic detection of results like mild, normal, moderate and severe.

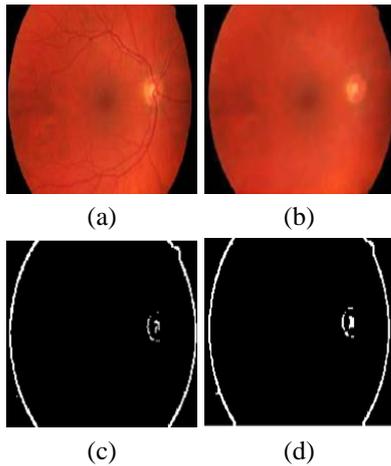


Fig.7. Optic Disc Segmentation: (a) Contrast Enhanced Image (b) Removal of Blood Vessel (c) Optic Disc and (d) Optic Disc Edges

The Fig.8 denotes the hybrid classification results according to regions.

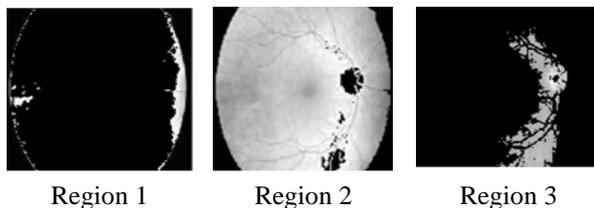


Fig.8. Hybrid classifier results according to regions

The Table.1 describes the comparison result obtained from the proposed experiment. Here proposed hybrid classification procedure is compared with existing single classifier [31] called Naive Bayes Classifier. The obtained results from both classifiers are compared. When a single classifier is used, it results in lower

accuracy, sensitivity and specificity for 45 images including the healthy patients and glaucoma patients. In hybrid classification approach best results are obtained using HRF database than compared to a single classifier.

The Fig.9 describes the graphical result of classification approaches and Fig.10 describes the samples results for classification techniques.

Table.1. Comparison of different classifiers

Classifiers	Sensitivity	Specificity	Accuracy
Naïve Bayes Classifier [31]	84.16	40	75.4
Hybrid Classifier (Naïve Bayes + KNN)	86.41	48	78.91

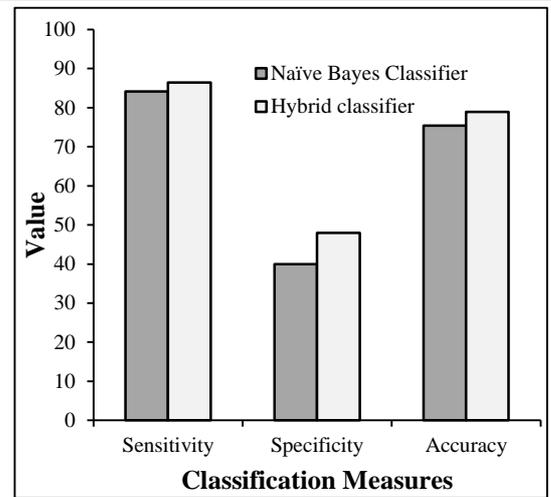


Fig.9. Comparison of Classification Metrics

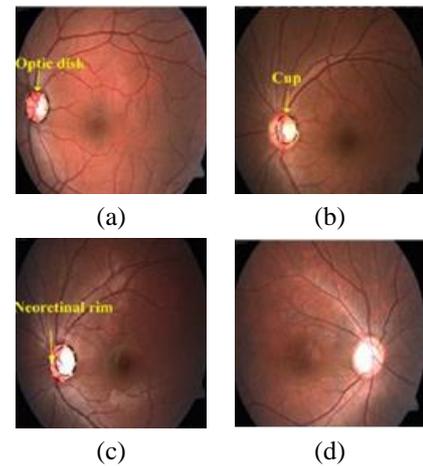


Fig.10. Sample results for classification techniques (a) Normal (b) Mild (c) Moderate (d) Severe

### 5. CONCLUSION

Glaucoma is kind of eye disease which leads to damage of optic nerve and vision loss. It is asymptotic in the beginning but eventually leads to loss of vision, if untreated. To identify the

presence of glaucoma at early stage, this work proposes a new approach for automatic detection of glaucoma disease. Pre-processing stage includes steps to reduce noise in image, contrast enhancement and normalization.

After pre-processing, optic disc is segmented based on MHT algorithm and cup is segmented from optic disc using intensity based elliptic curves. Various features like texture, shape and intensity features are extracted from the images. The best features are selected using three searching techniques namely, best first search, genetic search and rank search. For classification, hybrid classification method is used which results in higher accuracy compared to a single classifier. Thus the proposed process automatically detects the glaucoma disease of the human being which helps for medical treatment according to the classification result.

The future work aims to extract the additional features such as LBP and colour moments, multi-wavelets, haralick features, wavelets for detecting the glaucoma disease. These features can be extracted in separate manner or combination of two or three features, for example wavelets and multi-wavelets. Thus it helps to identify the glaucoma automatically with an increased accuracy.

## REFERENCES

- [1] A. Martinez-del-Rincon, J. Miller, and A. Azuara-Blanco, "Automatic Analysis of Digital Images for Glaucoma Detection", *Proceedings of International Conference on Irish Machine Vision and Image*, pp. 223-229, 2014.
- [2] Anum Abdul Salam, M. Usman Akram, Kamran Wazir, Syed Muhammad Anwar and Muhammad Majid, "Autonomous Glaucoma Detection from Fundus Image using Cup to Disc Ratio and Hybrid Features", *Proceedings of IEEE International Symposium on Signal Processing and Information Technology*, pp. 370-374, 2015.
- [3] Debjit Bhowmik, K.P. Sampath Kumar, Lokesh Deb, Shraavan Paswan and A.S. Dutta, "Glaucoma-A Eye Disorder Its Causes, Risk Factor, Prevention and Medication", *The Pharma Innovation*, Vol. 1, No. 1, pp. 66-81, 2012.
- [4] Jun Cheng, Fengshou Yin and Damon Wing Kee Wong, "Sparse Dissimilarity-Constrained Coding for Glaucoma Screening", *IEEE Transactions on Biomedical Engineering*, Vol. 62, No. 5, pp. 1359-1403, 2015.
- [5] Jiwanpreet Kaur Virk, Mooninder Singh and Mandeep Singh, "Cup-to-Disk Ratio (CDR) Determination for Glaucoma Screening", *Proceedings of 1<sup>st</sup> International Conference on Next Generation Computing Technologies*, pp. 504-507, 2015.
- [6] Suman Sedai, Pallab K. Roy, Dwarikanath Mahapatra and Rahil Garnavi, "Segmentation of Optic Disc and Optic Cup in Retinal Fundus Images Using Shape Regression", *Proceedings of 38<sup>th</sup> Annual International Conference of the IEEE Engineering in Medicine and Biology*, pp. 3260-3264, 2016.
- [7] Sobia Naz and Sheela N Rao, "Glaucoma Detection in Color Fundus Images Using cup to Disc Ratio", *International Journal of Engineering and Science*, Vol. 3, No. 6, pp. 51-58, 2014.
- [8] Maila Claro, Leonardo Santos, Wallinson Silva, Flavio Araujo and Nayara Moura "Automatic Glaucoma Detection based on Optic Disc Segmentation and Texture Feature Extraction", *CLEI Electronic Journal*, Vol. 19, No. 2, pp. 1-10, 2016.
- [9] P.Y. Shruti and A. Sharangouda, "Automatic Retina Feature Analysis for Glaucoma Detection using Cup to Disk Ratio Based on Morphology and Hough Circle Based Techniques", *International Journal of Advanced Research in Computer and Communication Engineering*, Vol. 4, No. 6, pp. 163-166, 2015.
- [10] Imran Qureshi, "Glaucoma Detection in Retinal Images Using Image Processing Techniques: A Survey", *International Journal Advanced Networking and Applications*, Vol. 7, No. 2, pp. 2705-2718, 2015.
- [11] Andrea Giachetti, Lucia Ballerini and Emanuele Trucco, "Accurate and Reliable Segmentation of the Optic Disc in Digital Fundus Images", *Journal of Medical Imaging*, Vol. 1, No. 2, pp. 1-11, 2014.
- [12] Georg Michelson, Joachim Hornegger, Simone Wärtges and Berthold Lausen, "The Papilla as Screening Parameter for Early Diagnosis of Glaucoma", *Deutsches Arzteblatt International*, Vol. 105, pp. 583-590, 2008.
- [13] Rudiger Bock, Jorg Meier, Laszlo G. Nyul, Joachim Hornegger and Georg Michelson, "Glaucoma Risk index:Automated Glaucoma Detection from Colour Fundus Images", *Medical Image Analysis*, Vol. 14, No. 3, pp. 471-481, 2010.
- [14] Javeria Ayub, Jamil Ahmad, Jan Muhammad Lubna Aziz, Sara Ayub, Usman Akram and Imran Basit, "Glaucoma Detection through Optic Disc and Cup Segmentation using K-mean Clustering", *Proceedings of International Conference on Computing, Electronic and Electrical Engineering*, pp. 1-5, 2016.
- [15] Thresiamma Devasia, Poulouse Jacob and Tessamma Thomas, "Automatic Optic Disc Localization and Segmentation using Swarm Intelligence", *World of Computer Science and Information Technology Journal*, Vol. 5, No. 6, pp. 92-97, 2015.
- [16] Chuang Wang, Djibril Kaba and Yongmin Li, "Level Set Segmentation of Optic Discs from Retinal Images", *Journal of Medical and Bioengineering*, Vol. 4, No. 3, pp. 1-8, 2015.
- [17] Gopal Datt Joshi, Jayanthi Sivaswamy, and S.R. Krishnadas, "Optic Disk and Cup Segmentation from Monocular Colour Retinal Images for Glaucoma Assessment", *IEEE Transactions on Medical Imaging*, Vol. 30, No. 6, pp. 1-14, 2015.
- [18] Anum A. Salam, Tehmina Khalil, M. Usman Akram, Amina Jameel and Imran Basit, "Automated Detection of Glaucoma using Structural and Non-Structural Features", *Springerplus*, Vol. 5, No. 1, pp. 1-21, 2016.
- [19] Arturo Aquino, Manuel Emilio Gegundez-Arias and Diego Marín, "Detecting the Optic Disc Boundary in Digital Fundus Images using Morphological, Edge Detection, and Feature Extraction Techniques", *IEEE Transactions on Medical Imaging*, Vol. 29, No. 11, pp. 1-10, 2010.
- [20] Karthikeyan Sakthivel and Rengarajan Narayanan, "An Automated Detection of Glaucoma using Histogram Features", *International Journal of Ophthalmology*, Vol. 8, No. 1, pp. 194-200, 2015.

- [21] K. Narasimhan and K. Vijayarekha, "An Efficient Automated System For Glaucoma Detection using Fundus Image", *Journal of Theoretical and Applied Information Technology*, Vol. 33, No. 1, pp. 104-110, 2011.
- [22] Javad Rahebi and Firat Hardalaç, "A New Approach to Optic Disc Detection in Human Retinal Images using the Firefly Algorithm", *Medical and Biological Engineering and Computing*, Vol. 54, No. 2-3, pp. 1-9, 2015.
- [23] Zhuo Zhang et al., "Neuro-Retinal Optic Cup Detection in Glaucoma Diagnosis", *Proceedings of 2<sup>nd</sup> International Conference on Biomedical Engineering and Informatics*, pp. 1-4, 2009.
- [24] Yuji Hatanaka, Atsushi Noudo, Chisako Muramatsu, Akira Sawada, Takeshi Hara, Tetsuya Yamamoto and Hiroshi Fujita, "Automatic Measurement of Cup to Disc Ratio Based on Line Profile Analysis in Retinal Images", *Proceedings of 33<sup>rd</sup> Annual IEEE International Conference on Medicine and Biology Society*, pp. 3387- 3390, 2011.
- [25] R. Gayathri, P.Y. Rao and S. Anma, "Automated Glaucoma Detection System based on Wavelet Energy features and ANN", *Proceedings of IEEE International Conference on Advances in Computing, Communications and Informatics*, pp. 2808- 2812, 2014.
- [26] Jiafu Jiang and Jing Shen, "An Effective Adaptive Median Filter Algorithm for Removing Salt and Pepper Noise in Images", *Proceedings of International Symposium on Photonics and Optoelectronics*, pp. 1-4, 2010.
- [27] Namita Aggarwal and R.K. Agrawal, "First and Second Order Statistics Features for Classification of Magnetic Resonance Brain Images", *Journal of Signal and Information Processing*, Vol. 3, pp. 146-153, 2012.
- [28] Carlos Garcia-Martinez and Manuel Lozano, "Local Search Based on Genetic Algorithms", Available at: [https://sci2s.ugr.es/sites/default/files/ficherosPublicaciones/0609\\_garciaSpringerBook08.pdf](https://sci2s.ugr.es/sites/default/files/ficherosPublicaciones/0609_garciaSpringerBook08.pdf).
- [29] D. Hema and B.N. Roy, "An Improved Page Rank Algorithm based on Optimized Normalization Technique", *International Journal of Computer Science and Information Technologies*, Vol. 2, No. 5, pp. 2183-2188, 2011.
- [30] Van Dang and W. Bruce Croft, "Feature Selection for Document Ranking using Best First Search and Coordinate Ascent", *Proceedings of International Woodstock Conference*, pp. 1-5, 2008.
- [31] Apeksha Avinash, K. Magesh and C. Vinoth Kumar, "Sift Feature Based Detection of Glaucoma", *Proceedings of IRF International Conference*, pp. 11-14, 2016.