

## Glaucoma Detection Using Enhanced K-Strange Points Clustering Algorithm and Classification

Vaishnavi Kamat<sup>1</sup>, Shruti Chatti<sup>2</sup>, Alvira Rodrigues<sup>2</sup>  
Chinmayee Shetty<sup>2</sup>, Anusaya Vadji<sup>2</sup>

<sup>1</sup>(Head, Department of Computer Engineering, Agnel Institute of Technology and Design, Goa University, India)

<sup>2</sup>(B.E Students, Department of Computer Engineering, Agnel Institute of Technology & Design, Goa, India)

---

**Abstract :** Glaucoma is an eye disorder that majorly affects the optic nerve head in the retina. The damage caused to optic disc leads to gradual loss of peripheral vision which may further result in complete blindness. Glaucoma cannot be cured, hence early and accurate detection is necessary. This paper proposes a method to detect Glaucoma using fundus images. Enhanced K-Strange Points Clustering (EKSTRAP) algorithm is applied to obtain cup, disc and the blood vessels from the Neuro-Retinal Rim (NRR). Further elliptical fitting method is used to compute cup to disc (CDR) ratio. The Inferior-Superior-Nasal-Temporal (ISNT) ratio is obtained using masking. CDR and ISNT are used as inputs to the Naïve Bayes classifier.

**Keywords** - Cup to Disc Ratio (CDR), Enhanced K-Strange Points Clustering (EKSTRAP), Glaucoma, Inferior, Superior, Nasal, Temporal quadrants (ISNT), Neuro Retinal Rim (NRR)

---

### I. INTRODUCTION

The eye is the most complex and delicate sense organ in the human body. The most leading causes of blindness are cataract, glaucoma and diabetes retinopathy. The second leading cause of blindness is glaucoma. Glaucoma is caused due to the increase in the Intra-Ocular Pressure (IOP) which damages the optic nerve. The IOP increases due to the build-up of excess fluid called the aqueous humor. Glaucoma mainly affects the optic disc by increasing the cup size thereby decreasing the NRR region.

Narasimhan et al. proposed a method for detection of glaucoma based on CDR and ISNT ratio wherein K-Means Clustering Algorithm is applied to obtain the optic disc and cup region and an elliptical fitting technique is applied to compute the CDR values. [1] Suryawanshi adopted a technique in which the CDR was obtained and boundary smoothening algorithm were applied to obtain perfect shape of the cup and disc. If the cup to disc ratio exceeds 0.3 then it indicates the abnormal condition that is the presence of glaucoma. [2] Pachiyappan et al. proposed a method for diagnosing Glaucoma using fundus images of the eye and the optical coherence tomography (OCT). The Retinal Nerve Fiber Layer (RNFL) can be broadly classified into top layer of RNFL, bottom layer of RNFL and the distance between the two boundaries. Based on the thickness of the nerve fiber layer which is nearly 105µm it was classified as Glaucomatous and Non-Glaucomatous. [3]

This paper proposes a new method for the diagnosis of glaucoma using fundus images. The ratio of the optic cup to disc (CDR) in fundus images is one of the primary parameters for the diagnosis of glaucoma. The EKSTRAP algorithm is recursively applied to extract the optic disc and optic cup region. The diameters of the optic cup and optic disc that are required to compute the CDR ratio is obtained by applying elliptical fitting technique. The blood vessels in the optic disc region are detected by using edge detection technique. The ratio of area of blood vessels in the inferior- superior side to area of blood vessels in the nasal-temporal side is known as the ISNT ratio. Masking technique is used to obtain the area of blood vessels in each of the regions. The CDR and ISNT values are the inputs to the Naive Bayes Classifier which labels the input fundus image as Glaucomatous, Suspected or Normal.

### II. CONCEPTS

#### 2.1 Pre-processing

Image pre-processing are techniques that are initially applied on the input image to eliminate noise and unwanted features in order to obtain the appropriate result. The fundus image is a RGB image. The red channel is the brightest channel and tends to be saturated. The green channel has the highest contrast and the blue channel appears to be empty. In the proposed work, green channel of the optic disc is extracted as it has the highest contrast.

**2.2 Processing**

**Morphological Operation**

Morphological operations are non-linear operations that are related to the shape or morphology of an image and its features. Performing morphological operation on a binary image results in creating a new binary image and if the test is successful the pixel will have a non-zero value at the location in the input image.

**Dilation**

Image A dilated by structuring element B (denoted by  $A \oplus B$ ) produces a binary image,  $F = A \oplus B$  with 1 in all locations (x, y) of a structuring element's origin at which that structuring element hits the input image A and 0 otherwise. i.e for every point  $x \in B$  translate A by these coordinates, and then, take the union of all these translations. [5]

$$A \oplus B = \bigcup_{b \in B} (A)_b$$

Where, b is any point that belongs to image B.

**Erosion**

A eroded by structuring element B (denoted by  $A \ominus B$ ) produces a binary image  $F = A \ominus B$  with 1 in all locations (x, y) of a structuring element's origin at which the structuring element B fits the input image A and 0 otherwise.

$$A \ominus B = \bigcap A_B$$

The original image and a structuring element are taken as inputs. Eroding the original image by a structuring element results in a thinned output. [5]

**Elliptical fitting method**

Elliptical fitting technique is applied on the extracted cup and the disc to get the diameter of the cup and disc.

The area of the ellipse is calculated by using the following formulae. [7]

$$X = a * (\cos\alpha \cos\beta) - b * (\sin\alpha \sin\beta) \quad (1)$$

$$Y = a * (\cos\alpha \sin\beta) + b * (\sin\alpha \cos\beta) \quad (2)$$

Here,  $\beta = 0^\circ$  and  $\alpha$  varies from  $0^\circ$  to  $360^\circ$

$$\text{Area} = \pi ab \quad (3)$$

Where,

a: the major axis length (half of the rectangle width)

b: the minor axis length (half of the rectangle height).

**ISNT**

Increase in the cup size leads to decrease of the NRR. The rim width is calculated using ISNT rule. Rim area is calculated using masking technique where ISNT forms the 4 quadrants of the rim. For a Glaucomatous eye the area covered by blood vessels in the Inferior-Superior region should be greater than the temporal-nasal region. The mask is rotated by  $90^\circ$  each time and is used to obtain the area covered by blood vessels in each quadrant. The mask fits the image perfectly as both the cropped image and the mask have the same dimensions. [6]

**III. CLUSTERING TECHNIQUES**

**3.1 K-means Clustering**

K-means algorithm begins by randomly selecting the seeds of M clusters and groups the input samples around the clusters to define m centroids, one for each cluster. Using the m centroids, the Euclidean distance is calculated of the input sample. Next, m new centroids are calculated as the mean of the clusters derived from the previous step. This process is repeated until the means of the consecutive iterations and the cluster membership remains the same. K-means clustering plays a vital role in computing the CDR. It is an unsupervised learning algorithm that solves the well-known clustering problem. The time complexity is O (n).

The distance measure used is:

$$\text{Euclidean distance} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + \dots + (x_n - y_n)^2}$$

Where,  $x_1, y_1, \dots, x_n, y_n$  are attributes.

**K-Means Algorithm:**

1. Choose the number of clusters 'M'.
2. Randomly pick 'm' centroids for each of the 'M' clusters.
3. Calculate the Euclidean distance for each of the object in the dataset from each of the centroids.
4. Based on the distance computed assign each object to the cluster it is nearest to.
5. Calculate the mean of the attribute values of the objects in each cluster. Set these means as the new centroids.

6. Check for the stopping criteria i.e. if the cluster membership is unchanged or the consecutive means are the same.
7. If yes, go to Step 9.
8. Else, go to Step 3.
9. Stop

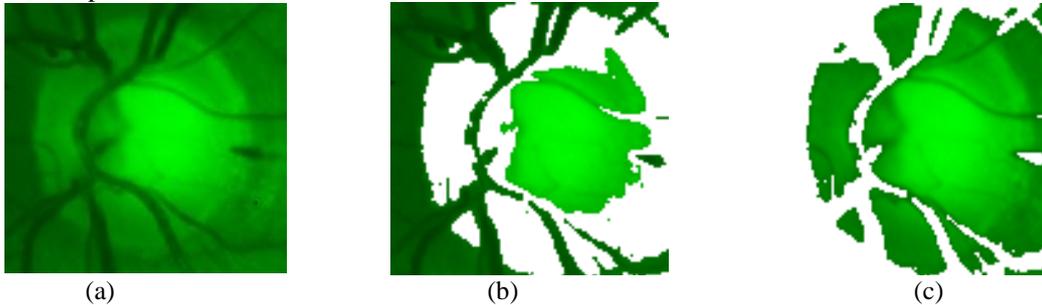


Fig 1: (a) Input image (b) Extracted cup (c) Extracted disc

### 3.2 Enhanced K-Strange Points Clustering

The EKSTRAP Algorithm is about finding strange points which are maximally apart from each other. This algorithm begins by finding the minimum  $X_{min}$  from the dataset. Later the maximum  $X_{max}$  point is calculated by computing the distance between all the points and the  $X_{min}$  using the Euclidean distance formula. After obtaining the  $X_{min}$  and  $X_{max}$  point next step is to compute strange points  $X_{str}$  which may be present anywhere. In order to locate its exact position, shifting procedure is applied as shown in algorithm. This method avoids the need of selecting the two random points by calculating the Euclidean distances between all the points from the dataset, thereby improving the performance of the algorithm. It requires  $O(n)$  time for finding the  $X_{min}$  and  $X_{max}$  point. [4]

#### Enhanced K-Strange Points Clustering Algorithm:

1. Find the minimum of the dataset,  $X_{min}$ .
2. Find a point  $X_{max}$ , which is at a maximum distance from  $X_{min}$ .
3. Locate a third point D which is farthest from  $X_{min}$  and  $X_{max}$
4. If  $(d(X_{min}, D) = d(X_{max}, D))$

$$X_{str} = D$$

else if  $(d(X_{min}, D) < d(X_{max}, D))$

$$X_{str} = X_{strprv} + X_m \left[ \frac{|X_{max} - X_{strprv}|}{X-1} \right]$$

else if  $(d(X_{min}, D) > d(X_{max}, D))$

$$X_{str} = X_{min} + X_m \left[ \frac{|X_{strprv} - X_{min}|}{X-1} \right]$$

5. Repeat the above steps until X strange points are found.
6. Assign the remaining points in the dataset into clusters formed by these non collinear X-strange points.
7. Output Xclusters

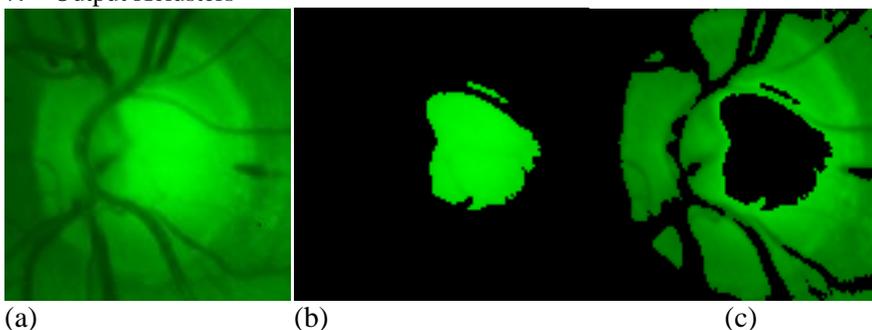


Fig 2: (a) original image (b) Extracted cup (c) Extracted disc

#### IV. CLASSIFICATION

Naive Bayes classifier is a probabilistic classifier based on Bayes Theorem and assumes independence among features. [8]

Bayes theorem calculates posterior probability  $P(c|x)$  from  $P(c)$ ,  $P(x)$  and  $P(x|c)$ .

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)}$$

Likelihood
Class Prior Probability  
Posterior Probability
Predictor Prior Probability

$$P(c|X) = P(x_1|c) \times P(x_2|c) \times \dots \times P(x_n|c) \times P(c)$$

Where,

$P(c|x)$ : posterior probability of class,  $c$  given predictor,  $x$ .

$P(c)$ : class prior probability

$P(x|c)$ : likelihood probability of predictor,  $x$  given class,  $c$ .

$P(x)$ : prior probability of predictor.

**Algorithm:**

Step 1: Let  $D$  be the training set of tuples along with their class labels.

Step 2: Given  $n$  classes, the classifier will predict that the unknown tuple  $X$  belongs to the class  $C_i$  only if,

$$P(C_i|X) > P(C_j|X) \text{ for } 1 \leq j \leq n, j \neq i$$

Maximize  $P(C_i|X)$ . The class  $C_i$  for which  $P(C_i|X)$  is maximized is called the maximum posterior hypothesis.

By Bayes theorem,

$$P(C_i|X) = \frac{P(X|C_i)P(C_i)}{P(X)}$$

Step 3: As  $P(X)$  is constant for all classes, only  $P(X|C_i)P(C_i)$  needs to be maximized. The class prior probabilities are obtained by,

$$P(C_i) = \frac{|C_i, D|}{|D|}$$

Where,  $|C_i, D|$  is the number of training tuples of class  $C_i$  in  $D$ .

Step 4: Naïve assumption of class conditional independence is made i.e. the attribute values are conditionally independent of one another, given the class label of the tuple.

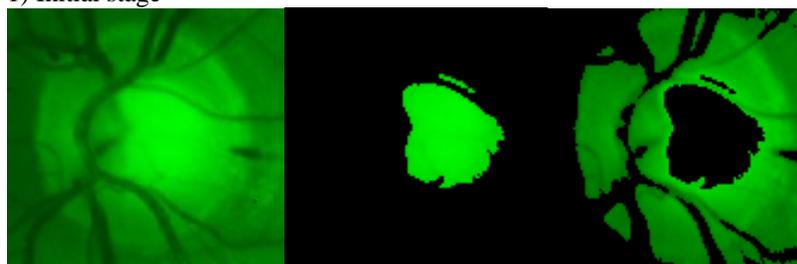
$$P(X|C_i) = \prod_{k=1}^n P(x_k|C_i)$$

#### V. PROPOSED METHODOLOGY

The pre-processed image is segmented to obtain clusters of optic disc, cup and the blood vessels using EKSTRAP algorithm. Image is converted into a binary image for further processing. Morphological operation as mentioned in section 2.2 is applied on the segmented output to smoothen the boundaries.

Elliptical fitting method which is a boundary smoothing technique is applied to obtain the diameters needed to compute the CDR and the ISNT ratio is computed using the masking technique as described earlier. The CDR and the ISNT ratio is given as input to Naive Bayes classifier which classifies whether the image is glaucomatous or suspected or normal.

1) Initial stage



(a)

(b)

(c)

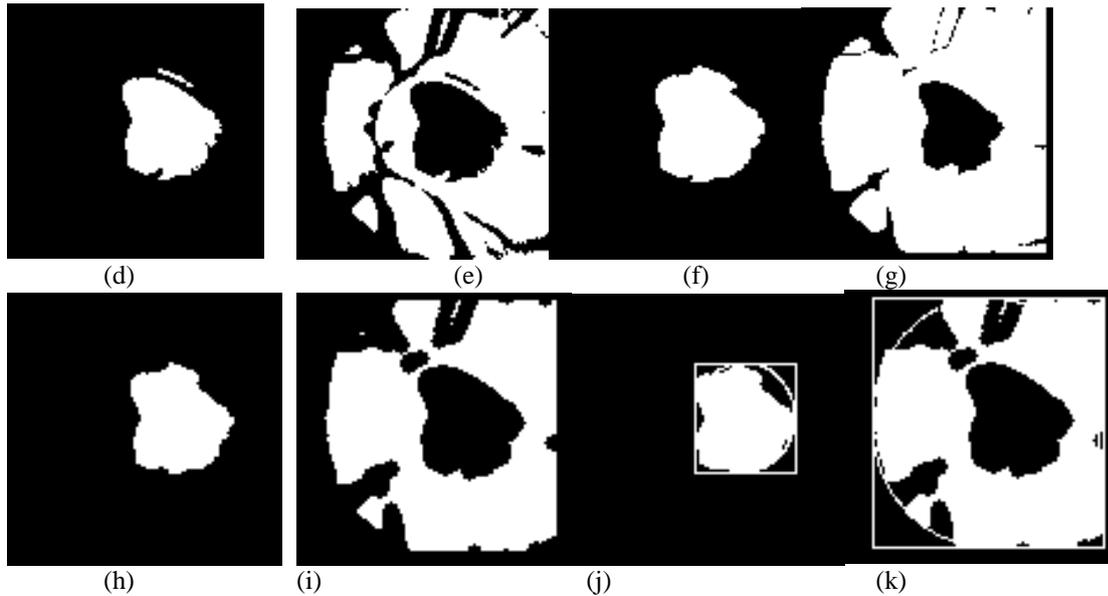


Fig 3:(a)Original image (b) Extracted cup (c) Extracted disc (d) Binary cup (e) Binary disc (f) Dilated cup (g) Dilated disc (h) Eroded cup (i) Eroded disc (j) Elliptical fitting on cup (k) Elliptical fitting on disc

To calculate ISNT ratio, NRR has to be extracted. To extract neuro retinal rim, the cup is subtracted from the disc and to detect the blood vessels in each of the quadrants edge detection technique is used. This is then converted to binary and masking is applied.

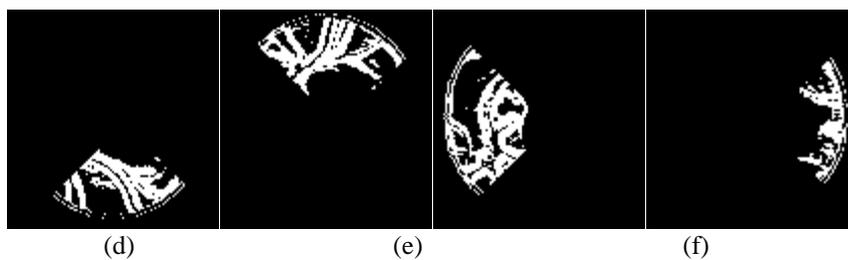
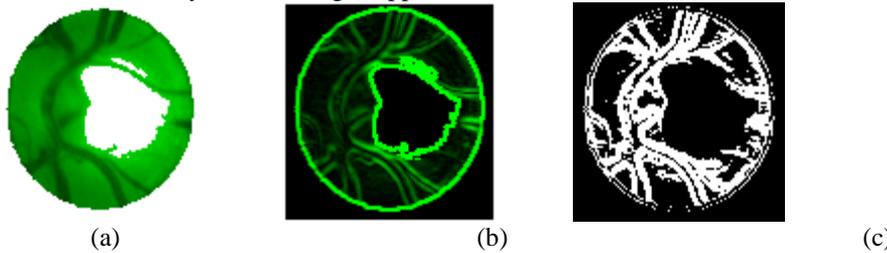


Fig 4: (a)Extracted NRR (b) Applying edge detection technique (c)Conversion to Binary (d) Blood vessels in I quadrant (e) Blood vessels in S quadrant (f) Blood vessels in N quadrant (g) Blood vessels in T quadrant

## VI. RESULT

K-MEANS	EKSTRAP
<p style="text-align: center;"><b>CUP</b></p> <p style="text-align: center;"><b>DISC</b></p>	<p style="text-align: center;"><b>CUP</b></p> <p style="text-align: center;"><b>DISC</b></p>

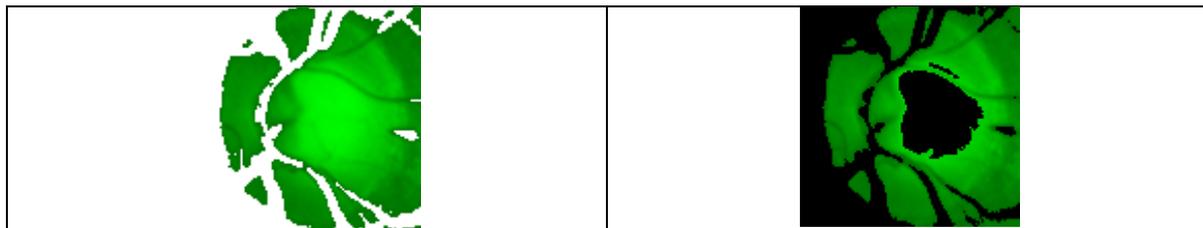


Table 1: Comparison between K-means and Enhanced K-Strange Points Clustering Algorithm

## VII. CONCLUSION

K-means algorithm does not always converge as it begins by initially selecting random seeds. If the seeds are not selected properly then the method has to be repeated entirely. Where as in case of EKSTRAP this is not the case. So, EKSTRAP gives better cluster outputs.

The CDR and ISNT ratio calculated from the clustering results were given to Naïve Bayes Classifier for classification.

Better segmentation techniques and classifier can be used. More inputs can be given to classifier to obtain more accurate results. Further, different stages of glaucoma can be classified.

## REFERENCES

- [1] K.Narasimhan, Dr.K.Vijayarekha, "An efficient automated system for glaucoma detection using fundus image", *Journal of Theoretical and Applied Information Technology* Vol. 33 No.1, ISSN: 1992-8645, November 2011
- [2] Preeti Kailas Suryawanshi: "An approach to glaucoma using image segmentation techniques" *International Journal Of Engineering Sciences & Research Technology* 2(9):September, 2013.
- [3] ArulmozhivarmanPachiyappan, Undurti N Das, Tatavarti VSP Murthy and RaoTatavarti, "Automated diagnosis of diabetic retinopathy and glaucoma using fundus and OCT images" *Lipids in Health and Disease* 2012, 11:73.
- [4] Terence Johnson, Dr.Santosh Kumar Singh, "Enhanced K Strange Points Clustering Algorithm", *Proceedings of the 2<sup>nd</sup> International Research Conference on Emerging Information Technology and Engineering Solutions. EITES 2015, 978-1-4799-1838-6/15, IEEE Computer Society Washington, DC USA© 2015 IEEE, DOI 10.1109/EITES.2015.14, indexed in ACM Digital Library, pp 32-37*
- [5] Gonzalez, R.C. and R.E. Woods, *Digital Image Processing*. (Reading, Massachusetts: Addison-Wesley, 1992).
- [6] KurnikaChoudhary, ShamikTiwari,"ANN Glaucoma Detection using Cup-to-Disk Ratio and Neuroretinal Rim ", *International Journal of Computer Applications* (0975 – 8887) Volume 111 – No 11, February 2015
- [7] T.R.GANESHBABU," Automated Detection of Optic Disc Using Fundus Image ", *International Journal of Advances in Engineering*, 2015, 1(1), 13-17 ISSN: 2394-9260 (printed version); ISSN: 2394-9279 (online version)[url:http://www.venuspublications.com/ijae.html](http://www.venuspublications.com/ijae.html).
- [8] Pang-Ning Tan, Michael Steinbach, Vipin Kumar, *Introduction to Data Mining* (Pearson Education, ISBN:81-317-1472-1)